Ontological Approach towards Discovering and Recommending COVID-19 Therapeutics, Risk Factors, and Drug Interactions

Debanjali Bain[1,3][0000−0002−2811−6692] and Biswanath Dutta[2][0000−0003−3059−8202]

[1] DRTC, Indian Statistical Institute, Bangalore, India. debanjali@drtc.isibang.ac.in
[2] DRTC, Indian Statistical Institute, Bangalore, India. bisu@drtc.isibang.ac.in
[3] Department of Library and Information Science, Calcutta University, Kolkata, India

Bain, D., Dutta, B. Ontological approach towards discovering and recommending COVID-19 therapeutics, risk factors, and drug interactions. J Supercomput 81, 841 (2025). https://doi.org/10.1007/s11227-025-07302-x

Abstract.
In today's data-driven world, integrating diverse healthcare data sources into a unified framework is essential. The COVID-19 pandemic has underscored the critical need for extracting meaningful insights from fragmented clinical data, particularly in areas such as treatment efficacy, risk factor identification, and drug interactions. To address these challenges, we propose the COVID-19 Drug and Risk Ontology (COViDRO)—a formally developed OWL-DL ontology designed to model and integrate COVID-19 treatment options aligned with the "PRADiCT" framework (Patient Risk factors, Adverse effects, Drug interaction, Clinical findings, and Treatment procedure). We hypothesize that this ontology will assist healthcare professionals in discovering and recommending COVID-19 therapeutics tailored to individual patients by considering risk factors, underlying health conditions, ongoing medications, potential drug interactions, and adverse effects. To validate its reliability and effectiveness, COViDRO underwent a multi-tier evaluation process: (1) Quality-based assessment using the Ontology Pitfall Scanner (OOPS!) to detect and resolve modeling errors, benchmarking COViDRO against related ontologies based on structural, functional, and usability dimensions; (2) Structural and logical validation using OntoDebug for structural integrity checks and the Pellet reasoner for logical consistency verification; (3) Quantitative evaluation using the OntoMetrics framework to assess ontology metrics such as attribute richness, relation richness, and knowledge base complexity, comparing with related ontologies; and (4) Query-based evaluation using SPARQL to assess the ontology's reasoning and retrieval capacity. The evaluation results confirm that COViDRO effectively organizes 135 classes, 32 object properties, and 15 data properties, enabling structured clinical reasoning. SPARQL queries successfully demonstrate its ability to retrieve patient-specific therapeutic recommendations, assess risk factors, and generate drug interaction alerts, validating its practical utility in healthcare settings. As a formal, DL-enabled ontology, COViDRO can contribute to automated inference of treatment options, thereby enhancing decision support systems and knowledge-based applications. Its structured and extensible approach makes it a valuable resource not only for COVID-19 but also for future pandemics and infectious disease management, reinforcing its significance in healthcare informatics.

Keywords: COVID-19 Data Modeling · Knowledge Graph · Description Logic · Ontology · Automated reasoning · Drug Interactions · Risk Factors · Personalized Care

1 Introduction

The COVID-19 pandemic, an exceptional global crisis, has created a significant demand for accurate and comprehensive information related to treatments, risk factors, and drug interactions associated with the disease [1,2]. This urgent requirement poses a substantial challenge to healthcare professionals and researchers worldwide in understanding the effectiveness of various treatment options, identifying potential risk factors, and managing drug interactions in COVID-19 patients [3]. The complexity in treating COVID-19 arises from the wide range of available medications, their potential side effects, and the intricate interactions that can occur when multiple drugs are used together [4]. In this context, a deep understanding of drug interactions, risk factors, and treatment choices is crucial for providing effective patient care and reducing the risk of complications [5].

Importantly, the selection of medication should be tailored to the specific medical conditions of each patient, taking into account factors like hospitalization, pregnancy, obesity, and underlying health issues [7]. For example, consider Remdesivir, a commonly used drug for COVID-19 treatment. It is approved for both hospitalized and non-hospitalized individuals, including adults and pediatric patients over 28 days old and weighing at least 3 kilograms. It is particularly recommended for patients with mild-to-moderate COVID-19 symptoms who are at a higher risk of progressing to severe illness or needing hospitalization [8].

However, caution is necessary when administering Remdesivir to patients taking medications that may interact with it, such as certain steroids like dexamethasone and betamethasone, as it can affect the treatment's effectiveness and safety [9]. Adding to the complexity is the dispersal of information about COVID-19 treatment guidelines, drug interactions, and side effects across various websites and databases (see [a] to [f] of Appendix, section 9 for details). Manually processing all this information presents significant challenges in extracting valuable insights [6].

An ontology, "a formal, explicit specification of a shared conceptualization" [10,11], provides a structured framework for representing domain-specific knowledge by defining concepts, relationships, and properties, enabling semantic interoperability and structured data representation. A Knowledge Graph (KG) integrates ontologies with interconnected data instances [12,13], facilitating seamless data integration, retrieval, and structured analysis across domains [14]. In healthcare, ontology-based KGs interconnect treatment guidelines, drug interactions, and patient data, enhancing clinical decision-making [15,16,17]. Description Logic (DL), a formalism with logic-based semantics, underpins expressive ontology languages like OWL, enabling automated reasoning and knowledge inference [18,19,91]. Given that medical data is often unstructured, DL-based ontologies help organize and relate key information, improving decision-making through semantic reasoning [90]. In COVID-19 research, such ontologies model treatment variables, including patient conditions, drug interactions, and adverse effects, ensuring structured knowledge representation and retrieval [20,24].

Several notable efforts have contributed to the development of COVID-19 ontology, including the COVID-19 Surveillance Ontology [26], CIDO-COVID-19 [27], COVIDCRFRAPID [28], DRUGS4COVID19 [29], ROC [30], COVID-19 [31], CODO [32], and others (See, 2). However, none of these studies have comprehensively covered aspects like recommended treatments, drug interactions, adverse effects, patient risk levels, risk factors, underlying health conditions, and diagnoses [43]. To meet these crucial requirements, we present the *COVID-19 Drug and Risk Ontology (COViDRO)*, a formal ontology built on OWL-DL. COViDRO addresses the multifaceted challenges associated with COVID-19 treatment, risk factors, and drug interactions.

The knowledge for the COViDRO model has been extracted from diverse medical literature and treatment guidelines from reputable organizations (see Appendix, section 9 for details). These organizations include the World Health Organization (WHO), the National Institutes of Health (NIH), the Food and Drug Administration (FDA), and the Centers for Disease Control and Prevention (CDC). The information about therapeutics, adverse effects, and drug interactions is sourced from a wide range of authoritative and reliable medical literature. The process of knowledge extraction is further detailed in the section 3 of the model. Through the development of COViDRO, we envisioned that it would enhance patient care [45,47,48,50], support research initiatives [44,52,53,54], and contribute to the formulation of well-informed public health strategies [46,52,50].

COViDRO can be seamlessly integrated into KG-based information systems or recommender systems, aiding healthcare professionals in suggesting appropriate treatments based on a comprehensive assessment of factors, abbreviated as "*PRADiCT*" (Patient Risk factors, Adverse effects, Drug interaction, Clinical findings, and Treatment procedure). These factors encompass patient risk level; patient risk factors, (i.e., underlying health conditions, age, immunocompromised state, and occupation); drug interactions; drug adverse effects; clinical findings (i.e., diagnosis, signs, symptoms, and status); and treatment procedures [49]. COViDRO offers a standardized framework for organizing and integrating data from diverse sources, including clinical trials, medical literature, and real-world patient data. By representing *PRADiCT* in a structured and controlled manner, COViDRO can enhance informed decision-making [44,51], thereby elevating the quality of patient care [45,47].

While inspired by the CODO model, which concentrates on epidemiological aspects and COVID-19 transmission, COViDRO expands upon CODO's capabilities by seamlessly incorporating various aspects of *PRADiCT*. This inclusiveness proves effective in recommending suitable treatment options for COVID-19. The comprehensive approach positions COViDRO as a patient-centric solution for facilitating COVID-19 treatment options and personalized care based on individual patient characteristics. The *objectives* of COViDRO are as follows:

1. *Knowledge Organization*: The primary objective of COViDRO is to provide a comprehensive representation of *PRADiCT*. COViDRO organizes information about specific drugs, their classifications, interactions, dosages, and other relevant details. It also categorizes risk factors associated with COVID-19, such as age, comorbidities, and immunocompromised conditions.
2. *Data Integration*: With its standardized classes and properties, COViDRO can assist the user in the integration of data from various sources, including clinical trials, medical literature, and real-world patient data.
3. *Decision Support Systems*: COViDRO can serve as the foundation for developing decision support systems or clinical decision support tools. These tools assist healthcare providers in selecting appropriate treatment options based on *PRADiCT*. COViDRO's scope, technique, and ontological representational approach distinguish it from other initiatives in the field (detailed in Section 3).

The main *contributions* of this study are as follows: (i) *Design and Development of a Comprehensive Data Modeling Framework*: One of the primary contributions of this research is the design and development of a data modeling framework tailored to COVID-19. This framework encompasses treatment options, risk factors, and drug interactions, addressing the critical need for structured and organized information in the face of the pandemic's complexities.

(ii) *Creation of a Formal DL-Enabled Ontology*: This study introduces a formal DL-enabled ontology based on the OWL-W3C that specifically caters to *PRADiCT*. This ontology employs description logic, enhancing its ability to represent complex relationships and facilitate advanced inference.

(iii) *Enhancing Evidence-Based Decision-Making*: This involves defining the classes within the ontology that automate the inference of suitable treatment options for individual patients based on *PRADiCT*, offering healthcare professionals actionable insights. Additionally, we can incorporate and extend additional definitions as new knowledge evolves and requirements arise.

(iv) *Practical Applications of COViDRO*: The research showcases practical applications of the formal DL-enabled COViDRO ontology, illustrated through SPARQL queries. These examples vividly demonstrate how COViDRO's defined classes and automatic inference capabilities can inform essential decision-making processes when selecting and recommending specific drugs for COVID-19 patients.

(v) *Ontology Integration*: Another significant contribution of this study involves the seamless integration of the COViDRO model with the foundational CODO ontology through the designated URI (Uniform Resource Identifier) `https://w3id.org/codo`. This integration serves as a fundamental pillar for the development of COViDRO, ensuring a cohesive connection and extension of concepts. The clear identification of each concept within COViDRO using URIs, following the format `https://w3id.org/codo#COVIDRO_ConceptName`, establishes unambiguous referencing and structured representation.

For instance, if there is a concept class in COViDRO named "`Adverse effect`", it is distinctly represented as: `https://w3id.org/codo#COVIDRO_AdverseEffect`. Similarly, a concept object property named "`has adverse effect`" would be expressed as: `https://w3id.org/codo#COVIDRO_hasAdverseEffect`, and a concept data property named "`drug code`" as: `https://w3id.org/codo#COVIDRO_drugCode`, and so forth. This naming convention not only fosters clarity but also supports interoperability, allowing COViDRO to be seamlessly referenced and incorporated by other systems, applications, and ontologies.

The subsequent sections of this article are organized as follows: Section 2 discusses related work, highlighting COViDRO's distinctiveness in terms of its scope, technique, and ontological representational approach. Section 3 presents the methodology employed in constructing the COViDRO ontology, providing detailed insights into the development process. In Section 4, the COViDRO ontology is described, focusing on its ontological representation and highlighting key features such as classes, properties, and relationships. The role of defined classes in COViDRO has been discussed in Section 5. The efficacy of the COViDRO ontology is evaluated using various ontology evaluation methods in Section 6. Section 7 presents the discussions and limitations of COViDRO. Section 8 concludes the article by summarizing the main findings and outlining future directions for the ongoing development of the COViDRO ontology and KGs. This article also includes an Appendix section, Section 9 to detail the source information.

2 Related Work

In the field of COVID-19 ontology development, several works have contributed to the representation of COVID-19 knowledge and data. Notable ontologies in this domain include the COVID-19 Surveillance Ontology [26], CIDO-COVID-19 [27], COVIDCRFRAPID [28], DRUGS4COVID19 [29], ROC [30], COVID-19 [31], CODO [32] and others.

The COVID-19 Surveillance Ontology is an application ontology focusing on COVID-19 cases and related respiratory illnesses, utilizing data from electronic medical record systems. CIDO-COVID-19 adheres to recommended ontology development criteria and provides information on various aspects such as disease, diagnosis, transmission, symptoms, therapy, and prevention. The COVIDCRFRAPID ontology by the World Health Organization (WHO) serves as a semantic data model for the WHO's COVID-19 RAPID case record form, offering inferences for questions and answers. DRUGS4COVID19 describes relationships between medications and the COVID-19 virus, encompassing major classes such as drugs, effects, diseases, symptoms, and chemical substances. The ROC ontology has been developed to analyze the effectiveness and negative effects of government responses to COVID-19 in different countries. The COVID-19 Ontology supports text mining, data harmonization, and research on SARS-CoV-2, spanning virology, epidemiology, molecular interactions, and clinical aspects. It is particularly useful for drug repurposing, integrating extensive chemical data relevant to COVID-19 therapeutics.

Some other significant contributions beyond the existing ontologies include Drug Ontology (DrOn) [33], COVID-19 Infection Risk Ontology (CIRO) [34], Drug-Drug Interaction Ontology (DINTO) [35], Ontology of Drug Adverse Events (ODAE) [36], and Knowledge4COVID-19 [37].

DrOn is a modular and extensible ontology of drug products, their ingredients, and their biological activity, designed to enable comparative effectiveness and health services researchers to query National Drug Codes (NDCs) based on various criteria such as ingredient, molecular disposition, therapeutic disposition, and physiological effect.

CIRO aims to automate the assessment of COVID-19 infection risks for contact tracing and screening of close contacts, utilizing RDF and SPARQL queries for risk evaluation.

DINTO provides a machine-readable ontology for drug-drug interactions and their underlying mechanisms, addressing representation and inconsistency gaps in existing databases.

ODAE is an ontology-based systematic representation and analysis of drug adverse events, specifically focusing on adverse drug events (ADEs). Unlike COViDRO, ODAE is designed to capture and study the effects of various factors, such as patient age and the disease being treated, on the outcomes of ADEs. ODAE provides a general and extensible representation of ADEs associated with over 200 neuropathy-inducing drugs under different age and disease conditions.

Knowledge4COVID-19 is a semantic-based framework designed to construct KG that integrates diverse data sources, with a strong emphasis on drug-drug interactions, drug side effects, and treatment toxicities. The ontology includes detailed classes and properties for drugs (e.g., CovidDrug, DrugSideEffect), interactions (e.g., DrugDrugInteraction, PharmacodynamicDrugDrugInteraction), and adverse events (e.g., AdverseEvent), as well as annotations for publications and semantic relations. While it excels in integrating heterogeneous data and analyzing toxicity and drug interactions, it does not support personalized therapeutic recommendations or automated reasoning tailored to individual patients. In contrast, COVIDRO extends its scope beyond toxicity analysis to encompass a wider range of factors, including patient-specific conditions, treatment guidelines, and comprehensive risk assessments, all encapsulated within its PRADiCT model.

Unlike these previous studies, COViDRO distinguishes itself with its comprehensive representation of COVID-19-related treatment options, risk factors, and drug interactions. Table 1 provides a comparative overview that demonstrates how COViDRO fully integrates all PRADiCT factors, positioning it as a tool for generating personalized treatment recommendations. Unlike other ontologies that focus on isolated aspects such as disease tracking, drug interactions, or epidemiology, COViDRO offers a holistic solution by incorporating patient risk factors, drug adverse effects, treatment procedures, and real-world clinical data. It also supports automated reasoning and decision-making, which can empower healthcare providers to offer patient-specific treatment recommendations. Throughout its development, it follows best practices by reusing established vocabulary and ontologies i.e., SCHEMA [38], FOAF [39], ORG [40], SNOMED-CT [41], CODO [32], OBO ontologies [42], and many more, ensuring data consistency and compatibility. This reuse not only enhances data coherence but also transforms COViDRO into a linked data vocabulary, improving interoperability (further details are in section 3).

COViDRO builds upon the foundation laid by the COviD-19 Ontology for cases and patient information (CODO). CODO has been implemented in a variety of different works, including ROC, COPOMBOCY [86], COVIDO [87], COKPME [88], InBan CIDO [89], COVIDonto [18], and others. CODO primarily focuses on epidemiological aspects such as virus transmission, patient records, and test results, adhering to FAIR principles (Findable, Accessible, Interoperable, and Reusable) [15,16]. It supports advanced analytics, contact tracing, and semantic data organization. COViDRO extends CODO's capabilities by incorporating patient-specific treatment factors, thereby bridging the gap between epidemiological modeling and personalized healthcare decision-making.

3 Methodology

This section presents the methodology employed in the design and development of the COViDRO ontology. Various methodologies exist in the literature for ontology design, including METHONTOLOGY [55], DILIGENT [56], NeOn [57], UPON [58], and YAMO [59], among others. The design approach for COViDRO draws inspiration from the YAMO and NEON methodologies, which provide a systematic and step-by-step process for developing a large-scale faceted ontology with formal definitions.

The design process of the COViDRO ontology comprises ten distinct steps, denoted as S1 to S10, as illustrated in Figure 1. The figure has been created using yEd graph editor [60]. The following sections provide a detailed description of each step in the COViDRO ontology development process.

S1: Deriving the purpose: In this initial step, the purpose and scope of the COViDRO ontology are clearly defined. The primary objective of COViDRO is to provide a comprehensive representation of *PRADiCT*. Additionally, COViDRO enables different organizations, such as government agencies, healthcare institutions, researchers, data publishers, and news agencies, to annotate and describe COVID-19 information effectively. By establishing a clear objective, the COViDRO ontology sets the foundation for its development and utilization in addressing the challenges posed by the COVID-19 pandemic.

S2: Development of competency questions: In this step, the purpose defined in S1 is further elaborated by formulating a set of competency questions (CQ I to CQ X) that focus on all the aspects of *PRADiCT*. These competency questions serve as a guide to ensure that the ontology captures the necessary information and relationships required to address these specific areas. Some examples of these competency questions have been provided in Table 2.

S3: Terminology extraction: This systematic approach to terminology extraction, a.k.a. knowledge extraction, ensures that COViDRO is enriched with a robust set of terms, enhancing its effectiveness in representing and analyzing various aspects of COVID-19 treatment. By utilizing authoritative sources and incorporating a variety of data types, this process contributes to the completeness and reliability of the extracted terminology.

The primary focus during the development of COViDRO is to extract terminology and their definition related to COVID-19 treatment options based on the *PRADiCT* framework. This process will further help in defining the class in the ontology. Knowledge extraction process involves systematically gathering information from diverse authoritative sources and data types, ensuring that the ontology is comprehensive, reliable, and reflective of the latest advancements in COVID-19 treatment.

Table 1: Summary of related ontologies. O1-O13 signifies the ontology number. NA: Not Applicable

Ontology No.	Ontology [Ref. No.]	Coverage	Concept Overlap with COVIDRO	PRADiCT Factors Represented	Personalized Treatment Recommendations
O1	COVID-19 Surveillance Ontology [26]	COVID-19 cases, respiratory illnesses	Epidemiology, symptoms, patient data	Clinical findings (diagnosis, symptoms)	No
O2	CIDO-COVID-19 [27]	Disease, diagnosis, symptoms, therapy, prevention	Disease pathology, treatment strategies	Clinical findings, treatment procedures	No
O3	COVIDCRFRAPID [28]	WHO's case record form, semantic data model	COVID-19 case reports, patient symptoms	Clinical findings (diagnosis, symptoms)	No
O4	DRUGS4COVID19 [29]	Drug-virus relationships, medication effects	Pharmacology, medication data	Drug interactions, adverse effects	No
O5	ROC Ontology [30]	Effectiveness and impact of government responses	Policy measures, interventions	None	No
O6	COVID-19 Ontology [31]	Text mining, data harmonization, drug repurposing	Epidemiology, virology, molecular interactions	Clinical findings, molecular mechanisms	No
O7	CODO [32]	Epidemiological data, case tracking	Patient demographics, symptoms, test results	Patient risk factors, clinical findings	No
O8	DrOn [33]	Drug products, ingredients, biological activity	Pharmacology, drug classification	Drug interactions	No
O9	CIRO [34]	COVID-19 infection risks, contact tracing	Risk factors, disease spread	Patient risk factors	No
O10	DINTO [35]	Drug-drug interactions and mechanisms	Pharmacology, molecular mechanisms	Drug interactions	No
O11	ODAE [36]	Drug adverse events, effects on different populations	Adverse effects, medical conditions	Adverse effects, patient risk factors	No
O12	Knowledge4COVID-19 [37]	Drug-drug interactions, drug side effects, treatment toxicities, semantic annotations	Pharmacology, adverse events, medical knowledge graphs	Drug interactions, adverse effects, molecular mechanisms	No
O13	**COVIDRO** [-]	**COVID-19 therapeutics, patient risk factors, adverse effects, drug interactions**	NA	**All PRADiCT factors covered**	**Yes**

Table 2: Examples of competency questions (CQs) related to COVID-19 therapeutics.

CQ No.	Competency Questions
CQ I.	Which therapeutics are recommended for COVID-19 patients with specific underlying health conditions?
CQ II.	What are the potential adverse effects of specific COVID-19 drugs?
CQ III.	Identify the drug interactions that may occur with a particular COVID-19 therapeutic.
CQ IV.	What are the risk factors associated with severe COVID-19 infection?
CQ V.	How does the choice of COVID-19 therapeutics vary based on patient characteristics?
CQ VI.	Which drugs should be avoided due to known interactions with commonly prescribed COVID-19 medications?
CQ VII.	What are the potential drug-drug interactions between COVID-19 therapeutics and medications commonly used to manage other health conditions?
CQ VIII.	What are the recommended therapeutic options for COVID-19 patients with a history of allergies or hypersensitivity to certain medications?
CQ IX.	How do pre-existing conditions, such as diabetes, cardiovascular diseases, and respiratory disorders, impact the choice and efficacy of COVID-19 treatment?
CQ X.	What are the most effective treatment strategies for COVID-19 patients with mild symptoms?

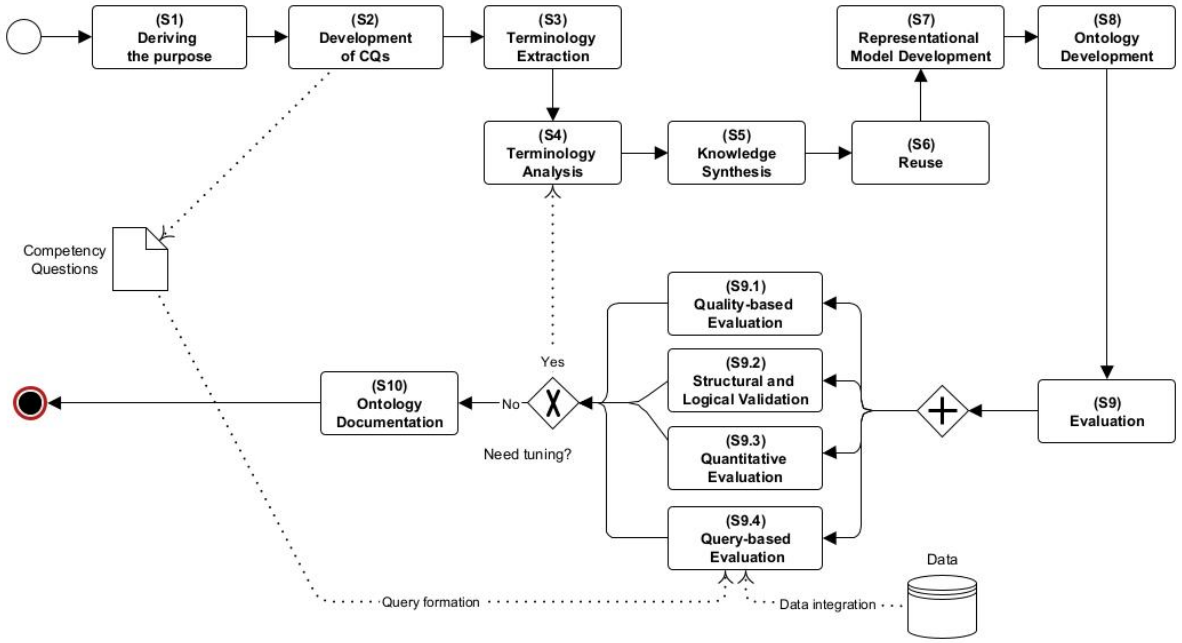


Fig. 1: COViDRO ontology development methodology. Solid arrows represent the sequential flow of steps in the ontology development process. Dashed arrows signify iterative feedback loops, where evaluation outcomes may necessitate revisiting earlier stages for refinement and tuning. Additionally, dashed lines related to competency questions (CQs) indicate its role in SPARQL query formation. Dashed lines related to data integration highlight its role in incorporating data inside the model for SPARQL query-based evaluation

To ensure thorough and systematic terminology extraction, several authoritative sources were selected, including the World Health Organization (WHO) [61], the U.S. Food and Drug Administration (FDA) [62], NIH COVID-19 Treatment Guidelines [63], and the Centers for Disease Control and Prevention (CDC) [64]. Clinical trials from ClinicalTrials.gov [65] and the International Clinical Trials Registry Platform (ICTRP) [66] were also reviewed to incorporate the latest findings on COVID-19 treatments. In addition, existing ontology repositories and portals, such as BioPortal [80], EMBL-EBI [81], and various COVID-19-related databases (e.g., [82,83,84]), were consulted. A detailed description of these sources and the information they provide is available in Appendix, section 9.

We also examined related ontologies (O1-O12), as discussed in the related work, to ensure compatibility with established frameworks. Large-scale observational studies were gathered from repositories like the COVID-19 Data Portal [85]. Furthermore, medical literature from peer-reviewed journals has been reviewed to ensure the inclusion of up-to-date and high-quality research.

The literature search has been carried out using the following steps:

- **Keywords and Search Terms:** “COVID-19 treatment”, “antiviral therapy”, “COVID-19 therapeutics”, “COVID-19 clinical guidelines”, “COVID-19 patient outcomes”, “COVID-19 drug interactions”, “adverse effects”, and “risk factors”.
- **Inclusion Criteria:** Peer-reviewed medical literature, clinical guidelines from reputable health organizations, and real-world patient data from published case studies.
- **Exclusion Criteria:** Preprints without peer review, non-medical sources, and opinion-based publications lacking empirical data.
- **Databases Searched:** PubMed[1], Scopus[2], Web of Science[3], Google Scholar[4], WHO and CDC guidelines (see Appendix, section 9 for details).

The terminology extraction process involved three main steps:

1. Extracting terminology related to drug efficacy, adverse effects, patient risk factors, and clinical procedures.

[1] https://pubmed.ncbi.nlm.nih.gov/
[2] https://www.scopus.com/
[3] https://mjl.clarivate.com/search-results
[4] https://scholar.google.com/

2. Organizing terms into specific categories, such as Pharmaceutical Preparation, drug-drug interactions in COVID-19 therapeutics, risk factors, adverse effects, and others (further refined in S4 and S5).
3. Iterative reviews and refinements to ensure completeness and accuracy, are carried out through multiple round discussions among authors.

During the terminology extraction, discrepancies across different sources, particularly regarding treatment efficacy and adverse reactions, were observed. These conflicts were resolved by giving precedence to guidelines from authoritative sources like WHO, NIH, and FDA over other studies (see Appendix, section 9 for details).

The extracted terminology serves as the foundational elements of COViDRO, enabling the ontology to represent and analyze COVID-19 treatment options comprehensively. These terms are crucial for supporting informed decision-making by healthcare professionals and researchers. Examples of extracted terminology include: "Medication", "Diagnosis", "Drug interaction", "Symptom", "Obesity", "Risk factor", "SARS-CoV-2 Antiviral Drug", "Age factor", "Pregnancy", "Underlying health condition", "Smoking", "Ritonavir-Boosted Nirmatrelvir", "Molnupiravir", "Tocilizumab", "Immune modulator", "High exposure risk occupation", "Moderate risk patient", "Ventilation", "Adverse effect", "Vital sign", "Monoclonal antibody", "Supplemental oxygen", "BMI", "Blood pressure", "SpO2", and others.

S4: Terminology analysis: In this step, we analyze the terms extracted in S3 by breaking down complex and compound concepts into their fundamental entities. We carefully examine each term to determine whether it should be classified as a class or a property based on its inherent definition and properties, as outlined in Section 4 of the COViDRO ontology. This analysis ensures that each term is appropriately categorized within the ontology, facilitating accurate representation and organization of *PRADiCT*.

S5: Knowledge Synthesis: In this step, we establish associations between the extracted concepts from S4 and synthesize the knowledge to create a comprehensive class hierarchy within the COViDRO ontology. By analyzing the semantic relationships and properties of each term, we organize the concepts into a structured hierarchy. This hierarchy captures the interconnections between various categories, allowing for a systematic representation of COVID-19-related information. Figure 2 illustrates an excerpt of the class hierarchy within COViDRO. Each category is further expanded with specific subcategories, providing a clear and organized framework for understanding and navigating the COVID-19 domain within the ontology.

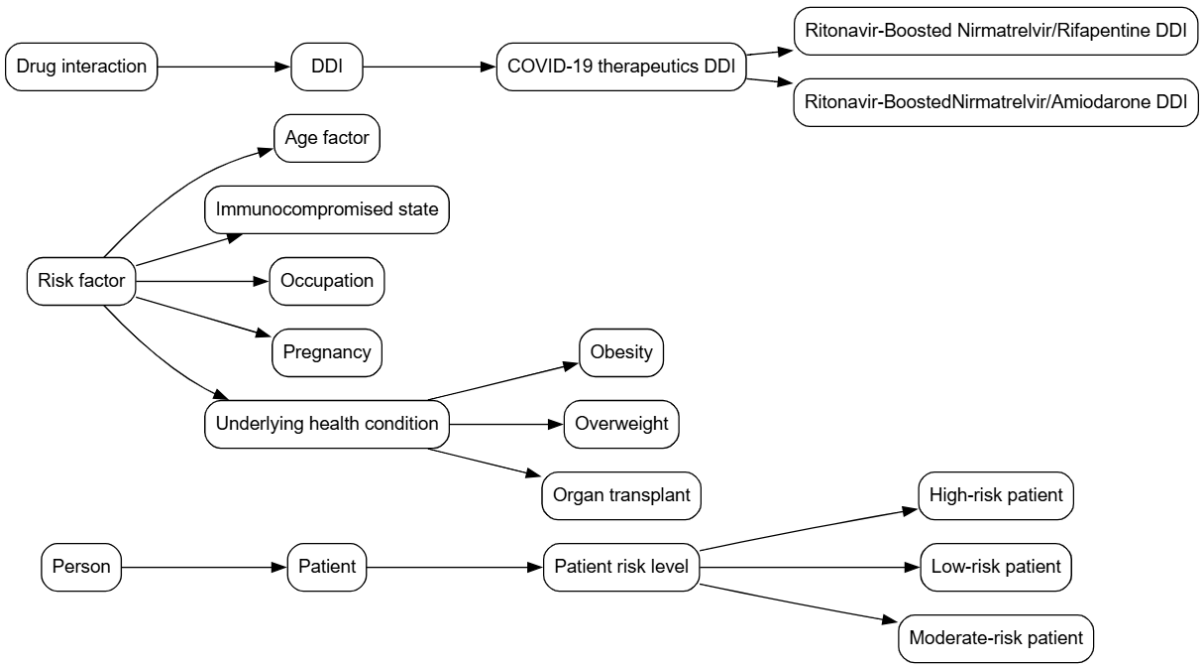


Fig. 2: Partial class hierarchy.

S6: Reuse: In the development of COViDRO, we follow the best practice of reusing pre-existing vocabulary and terms in ontology construction [92]. By leveraging well-known vocabularies i.e., SCHEMA, FOAF, ORG, SNOMED-CT; existing ontologies i.e., CODO, OBO ontologies (OGMS, SYMP, OAE, DINTO, NCIT) and others (see Table 3), we promote data consistency and enhance the ontology's compatibility with other datasets. For example, we repurpose terms from well-known vocabularies SCHEMA to represent commonly used concepts like "`Patient`", while FOAF is used to represent "`Agent`" entities such as "`Person`". This approach enables COViDRO to become a linked data vocabulary, facilitating seamless integration with other datasets that utilize the same

vocabulary. By reusing and standardizing terms, COViDRO promotes data interoperability and enhances the overall utility of ontology within the broader ecosystem of COVID-19-related information. A partial overview of the reused concepts is provided in Table 3, while Table 5 lists the reused vocabularies alongside their corresponding prefixes.

Table 3: Excerpts of concepts with reused vocabulary/ontology in COVIDRO. For some ontologies, the term "OBO" in brackets indicates that they are part of the OBO Foundry Ontologies.

Concepts	**Reused Vocabulary**
Clinical Finding	CODO
COVID-19 Diagnosis	CODO
Mild and Very Mild COVID-19	CODO
Moderate COVID-19	CODO
Severe COVID-19	CODO
Sign	SNOMED CT
Vital Sign	OGMS (OBO)
Status	CODO
Symptom	SYMP (OBO)
Adverse Effect	OAE (OBO)
COVIDRO_DrugInteraction	DINTO (OBO)
COVIDRO_DDI	DINTO (OBO)
Medications	NCIT (OBO)
COVIDRO_Antibiotic	SNOMED CT
COVIDRO_Antifungal	SNOMED CT
COVIDRO_Antiviral	SNOMED CT
COVIDRO_Anticoagulant	SNOMED CT
COVIDRO_Vasodilator	SNOMED CT
Agents	FOAF
Person	FOAF
Organization	ORG
Observational Findings	CODO
Risk Factor	CODO
Age Factor	CODO
Diseases	OGMS (OBO)
Disorder	OGMS (OBO)
Occupation	SCHEMA
Pregnancy	SNOMED CT

S7: Representational model development: We concentrate on developing a representational model that effectively captures and organizes the domain knowledge derived from the previous steps. The goal is to create a clear and illustrative framework that defines the various classes, properties, and their relationships within COViDRO. The representational model, depicted in Figure 3 (see, Section 4), serves as a comprehensive summary of the COViDRO ontology, showcasing its extensive coverage and providing a visual representation of the ontology's structure and components. Through the design of the representational model, COViDRO achieves a coherent and structured representation of *PRADiCT*.

S8: Ontology development: Following S7, we proceed to develop the COViDRO ontology, which is expressed in OWL-DL, a description logic-based ontology language. The Protégé ontology editor [68], developed at Stanford University, has been used as an ontology editing environment for designing the ontology, along with additional plugins such as Pellet Reasoner [70], OntoDebug [69], and SPARQL [71], as described in the subsequent sections. Further details of COViDRO ontology have been provided in Section 4. The formal notation of the COViDRO ontology is provided below.

Formally, the COViDRO ontology, denoted as E, can be represented as:

$$E = (IE, CoE, RE, PE)$$

where:

– IE represents a set of individuals (instances) in the ontology.
– CoE denotes a collection of concepts (classes).
– RE signifies a set of hierarchical relationships between concepts.
– PE represents a set of properties associated with the ontology.

This notation explicitly lists the fundamental components of the ontology.

Individuals and Concepts: Concepts (CoE): The set CoE consists of concepts that define classifications for individuals (IE). If $Co \in CoE$ is a concept, an individual belonging to this concept is represented as IE_{Co}. If there are multiple individuals within Co, they can be denoted as $IE_1^C o, IE_2^C o, \ldots$.

Hierarchy Relations (RE): Hierarchical relationships (RE) define the subsumption structure of concepts and properties in the ontology. These relationships are represented as:

$$RE = \{(Co_1, Co_2) \mid Co_1 \subseteq Co_2, \quad \forall Co_1, Co_2 \in CoE\}$$

where Co_1 is a subclass of Co_2, inheriting all its properties and attributes. For example, the concept `Patient` is a subclass of `Person`, meaning that all characteristics of `Person` apply to `Patient` as well.

Properties (PE): Properties within the ontology are categorized as follows:

– **Data Properties (DPE)**: A data property associates an individual with a specific value and is represented as:

$$(IE_{Co}, dp, v)$$

where IE_{Co} is an individual belonging to concept Co, dp is a data property, and v is the corresponding value (e.g., numerical or textual). For instance, for the concept `VitalSign`, an instance such as `vital0001` and `vital0002` may have a data property `SpO2` with values:

$$(\text{vital0001}, \text{SpO2}, 92), \quad (\text{vital0002}, \text{SpO2}, 98)$$

– **Object Properties (OPE)**: Object properties define relationships between individuals and are represented as:

$$(IE_1^{Co}, op, IE_2^{Co})$$

where IE_1^{Co} and IE_2^{Co} are individuals, and op is an object property linking them. For example, in the COViDRO ontology, the concept `PresentMedicalHistory` (PMH001) and `Medication` (Ivermectin) are linked via the object property `ongoingMedication`:

$$(\text{PMH001}, \text{ongoingMedication}, \text{Ivermectin})$$

COViDRO ontology maintains a self-contained structure to ensure interoperability. If a concept $Co \in CoE$ exists in an external ontology z, then the relationships and properties defined in E hold within z as well. That is:

$$(IE_z^{Co}, dp, v) \in DPE \quad \text{and} \quad (IE_1^{Co}, op, IE_2^{Co}) \in OPE$$

where z represents any compatible ontology integrated with COViDRO.

S9: Evaluation: The evaluation phase is critical in ensuring the consistency, correctness, and effectiveness of COViDRO, as discussed in Section 6. A multi-tier evaluation process is conducted, focusing on quality assessment, structural integrity, logical soundness, semantic richness, and functional performance. First, a quality-based evaluation is performed using the Ontology Pitfall Scanner (OOPS!) [79]. This evaluation has been conducted in two steps; in simple evaluation, COViDRO alone is evaluated to identify and resolve common modeling errors, such as undefined relations, missing disjoint axioms, and insufficient annotations. Additionally in advanced evaluation, COViDRO is benchmarked against 12 other related ontologies (O1–O12), mentioned in Section 2, through a comparative pitfall-based evaluation matrix, assessing its structural and usability dimensions.

Structural and logical validation is carried out by utilizing OntoDebug [69] for structural integrity checks and the Pellet reasoner [70] for logical consistency verification, ensuring the ontology remains free from faulty axioms and contradictions. This phase focuses solely on COViDRO to confirm its internal logical soundness without external comparisons. A quantitative evaluation has been conducted using OntoMetrics to assess COViDRO's structural complexity and semantic richness. This step involved comparing COViDRO with other ontologies (O1-O12) based on metrics such as axioms, class count, inheritance richness, and relation richness.

After running the Pellet reasoner, the inferred ontology is generated and downloaded for SPARQL-based evaluation. Finally, based on predefined competency questions (CQ1–CQ10), a query-based assessment is conducted on inferred model, using Apache Jena-Fuseki, testing COViDRO's capability to retrieve clinically relevant information (as detailed in Section 6.4). This step is conducted exclusively on COViDRO, emphasizing its effective expansion beyond the base ontology CODO and its impact on enhancing semantic reasoning. The evaluation examines how COViDRO extends CODO's capabilities by enabling more precise knowledge representation, supporting advanced inferencing, and facilitating structured data retrieval for personalized therapeutic decision-making.

The overall methodology of COViDRO evaluation has been illustrated in Table 4, detailing the tools used, key evaluation steps, and outcomes. S9.1–S9.4 represents the multi-tiered approach for COViDRO evaluation (also illustrated in Fig. 1).

Table 4: Summary of evaluation methodology for COViDRO

Sl. No.	Tools Used	Description	Key Steps
S9.1	OOPS! Scanner	Quality-based evaluation to detect modeling errors such as undefined relations, missing disjoint axioms, and insufficient annotations.	– Simple evaluation: Identify and resolve pitfalls (critical, important, minor) – Advanced evaluation: Benchmark against 12 COVID-19 ontologies (O1-O12) using pitfall-based evaluation matrix – Compare structural, functional, and usability dimensions
S9.2	OntoDebug, Pellet Reasoner	Structural and logical validation to ensure the ontology is free from modeling errors and contradictions.	– Use OntoDebug for structural debugging and fault detection – Use Pellet Reasoner for logical consistency checking and inference validation
S9.3	OntoMetrics	Quantitative evaluation assessing structural complexity and semantic richness.	– Measure base metrics (axioms, logical axioms, class count, object/data properties) – Evaluate schema metrics (attribute richness, inheritance richness, relation richness) – Analyze knowledge base metrics (average population, class richness) – Compare with other ontologies (O1-O12)
S9.4	Apache Jena-Fuseki	Query-based evaluation testing COViDRO's knowledge retrieval capacity.	– Derive SPARQL queries from competency questions (CQ1-CQ10) – Populate ontology with test data – Execute queries to retrieve therapeutic recommendations, adverse effects, and drug interactions

In cases where inconsistencies or unexpected results arise, the ontology is iteratively refined by revisiting S4 until all validation criteria are satisfied. This iterative step has been illustrated using a dashed arrow in Fig. 1.

S10: Ontology Documentation: During this step, our main focus is to create comprehensive and well-organized documentation for the COViDRO ontology, making it easier for the scientific community to understand, disseminate, and reuse the information. To achieve this, we utilize the WIDOCO (WIzard for DOCumenting Ontologies) tool [73], which automates the process of generating ontology documentation by extracting and processing essential information from the ontology's metadata, annotations, and axioms. The generated documentation provides detailed insights into various aspects of the ontology, including classes, properties, individuals, and ontology metrics. Additionally, it encompasses explanations about the introduction, overview, description, and visualization of the ontology. This information has been compiled and presented in a user-friendly web interface in the form of an HTML page.

4 COViDRO Ontology

This section provides a comprehensive overview of the COViDRO ontology, encompassing its essential classes, properties, and exemplar individuals. The current version of the ontology can be obtained from: https://w3id.org/codo/1.4. It is worth noting that COViDRO brought in a total of 135 classes, 32 object properties, and 15 data properties offering a wide range of concepts that can be applied across diverse healthcare settings. Figure 3 presents a high-level overview of COViDRO, illustrating its primary classes and object properties. Due to space constraints, not all classes and properties, are depicted. The figure also showcases the integration of external vocabularies through their respective prefixes. For example, the `foaf:Person` class in the figure indicates that the "Person" class has been reused from the FOAF vocabulary. Detailed descriptions of COViDRO's classes and properties are provided in the following subsections.

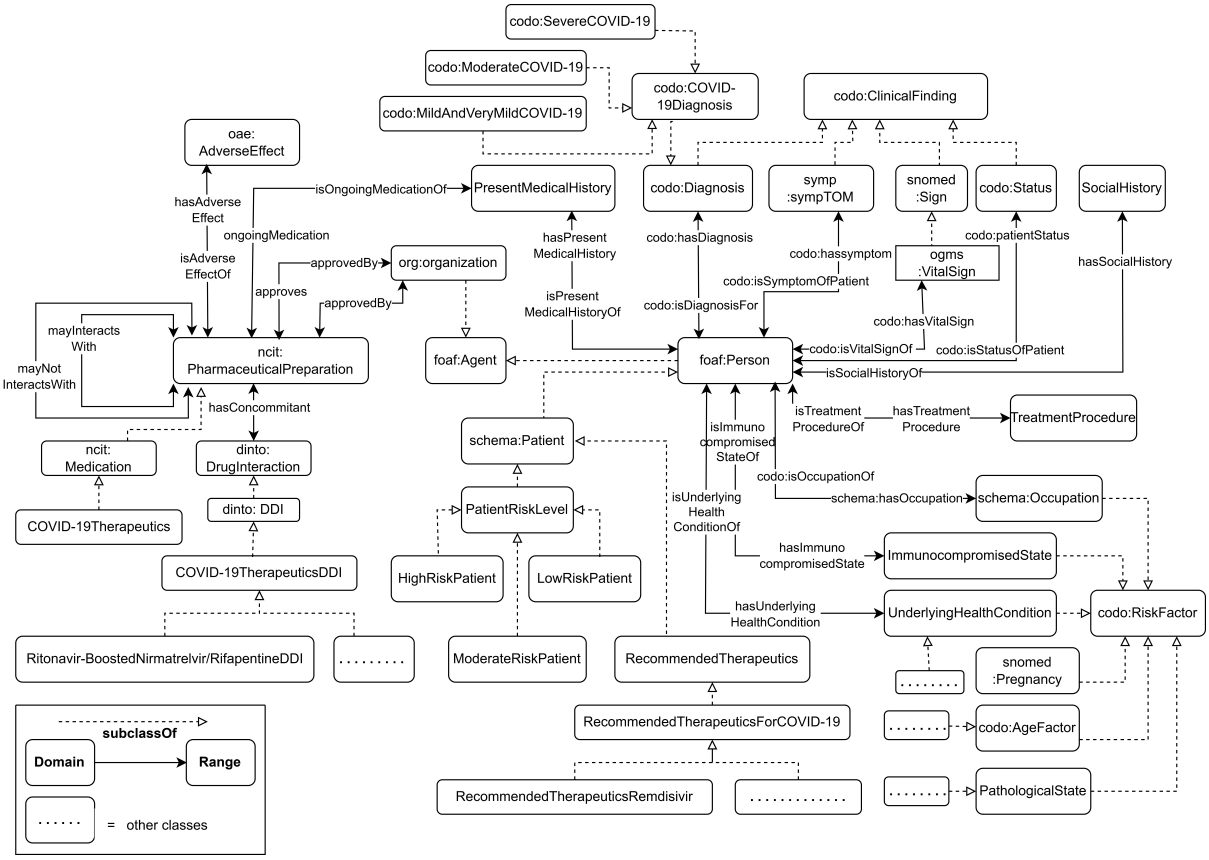


Fig. 3: Overview of COViDRO, illustrating its classes and object properties (entity relations). The reused vocabularies and their corresponding prefixes are listed in Table 5.

Table 5: List of vocabularies and prefixes reused in COViDRO.

Vocabulary	**Prefix**
FOAF	http://xmlns.com/foaf/0.1/
ORG	http://www.w3.org/ns/org
SCHEMA	http://schema.org/
CODO	http://w3id.org/codo
SNOMED CT	http://snomed.info/id/
OBO	http://purl.obolibrary.org/obo/
OGMS	http://purl.obolibrary.org/obo/OGMS
SYMP	http://purl.obolibrary.org/obo/SYMP
OAE	http://purl.obolibrary.org/obo/OAE
DINTO	http://purl.obolibrary.org/obo/DINTO
NCIT	http://purl.obolibrary.org/obo/NCIT

4.1 Classes

COViDRO, as a comprehensive ontology, is designed to address the critical need for accurate and well-structured knowledge concerning COVID-19 treatment options. The COViDRO model has 135 classes, each representing a set of individuals that share common properties.

The COViDRO ontology is tailored to depict *PRADiCT* through a well-defined set of main classes, including `DrugInteraction`, `RiskFactor`, `Person`, `AdverseEffect`, `PharmaceuticalPreparation`, and `TreatmentProcedure`, all interconnected by explicit relationships. These relationships are meticulously defined to ensure accurate representation and effective data retrieval.

Within the ontology, the main classes have extended subclasses, each precisely specified with relevant axioms. For instance, `DrugInteraction` is further refined by the subclass `DDI` or Drug-drug Interaction, which elaborates on the concept of drug-drug interactions and the influence of one drug on another's disposition and effects. This subclass is defined with a qualified cardinality (∀`hasConcommitant` exactly 2 `PharmaceuticalPreparation`), specifying that instances of `DDI` must involve exactly two relationships with instances of `PharmaceuticalPreparation`. Figure 4 illustrates an excerpt of the class hierarchy for the class `DrugInteraction`.

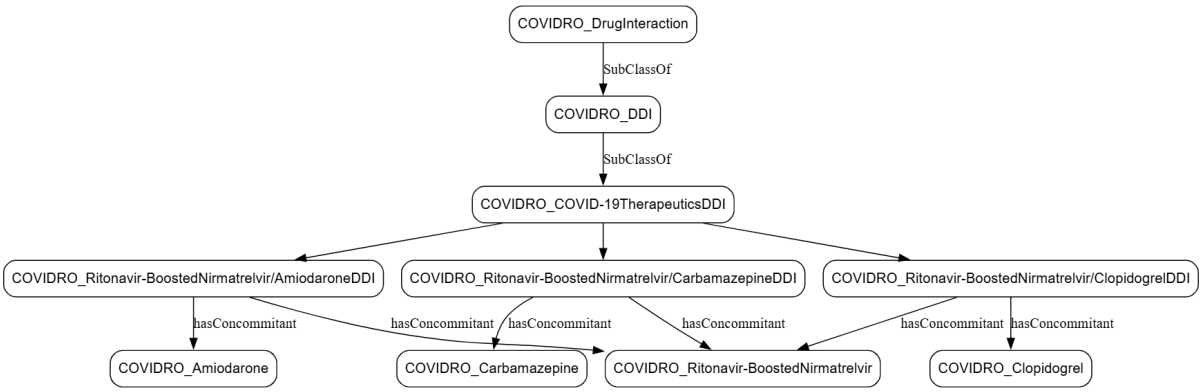


Fig. 4: An excerpt of drug-drug interactions (DDIs) in the COVIDRO ontology. The hierarchy illustrates subclass relationships among drug interactions, with labeled edges indicating the `hasConcommitant` property, representing the pharmaceutical preparations involved in each interaction.

Another significant class is `AdverseEffect`, which aptly captures undesirable outcomes resulting from medical treatments, interventions, medications, or substance exposures. Through the object property `hasAdverseEffect`, this ontology establishes a clear link between a `PharmaceuticalPreparation` and an `AdverseEffect`, facilitating the detailed description of adverse effects associated with specific pharmaceutical products.

When it comes to therapeutic interventions, the class `RecommendedTherapeutics` proves invaluable, representing treatments specifically suggested for various medical conditions. This class offers flexibility in therapeutic options, considering an individual's medical history and the severity of their condition. Under this umbrella, the subclass `RecommendedTherapeuticsForCOVID19` further narrows down therapeutic choices for COVID-19 patients. One of its exemplifying subclasses, `RecommendedTherapeuticsAnakinra`, outlines precise conditions and restrictions for administering the therapeutic drug `Anakinra` to hospitalized adults with pneumonia who require supplemental oxygen and are at risk of severe respiratory failure. One example of a formal definition of such subclass can be found in Eq. 2 of Table 7 and Table 11. Figure 5 illustrates an excerpt of class hierarchy for the class `RecommendedTherapeutics`.

Delving into pharmaceutical products, the class `PharmaceuticalPreparation` encompasses a wide range of medicinal products developed by pharmaceutical companies for patient use. This includes medications, drugs, vaccines, and other healthcare products. Figure 6 shows an excerpt of the class hierarchy for the class `PharmaceuticalPreparation`. The subclass `Medication` further elucidates substances used to prevent, treat, or manage diseases in humans and animals, playing a pivotal role in healthcare. Moreover, subclasses for example `COVID-19Therapeutics`, `CardiovascularMedication`, `RespiratoryMedication`, and `AntiDiabeticMedication` provide granular insights into COVID-19 management and treatments for cardiovascular disease, respiratory ailments, and diabetes, respectively.

The COViDRO ontology also recognizes the importance of risk factors by incorporating the class `RiskFactor`. This class effectively captures attributes, characteristics, or exposures that heighten the likelihood of specific diseases or health conditions. It is thoughtfully divided into several informative subclasses like `UnderlyingHealthCondition`, `Pregnancy`, `Occupation`, `ImmunocompromisedState`, and `AgeFactor`; each shedding light on the impact of different risk factors on COVID-19 outcomes. Figure 7 shows an excerpt of the class hierarchy for the class `RiskFactor`. All class hierarchy diagrams presented in this work are generated using the Graphviz tool [67].

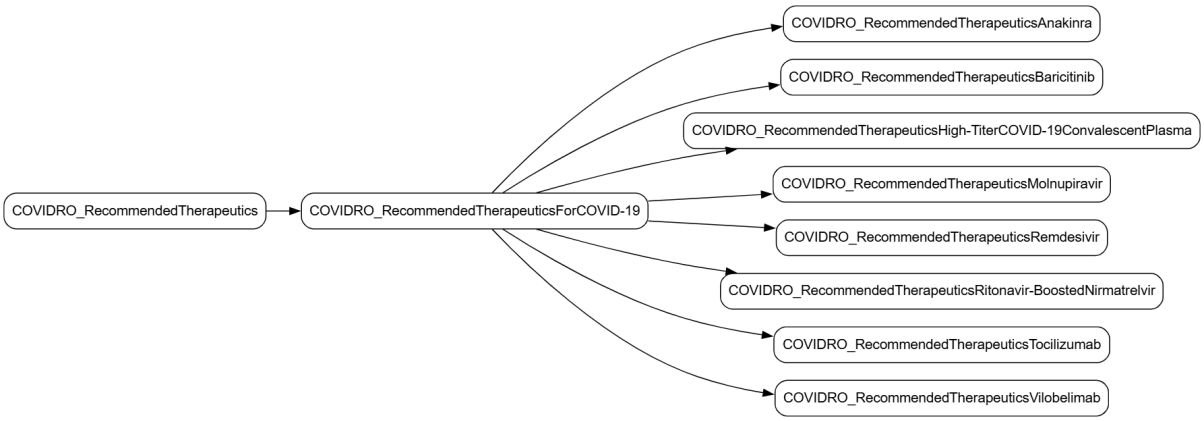


Fig. 5: An excerpt of class hierarchy for the class Recommended Therapeutics.

4.2 Properties

COViDRO incorporates both object properties and data properties to establish relationships and capture specific attributes within the ontology. Object properties facilitate connections between entities, whether they belong to the same or different classes. There are 32 object properties in COViDRO, such as `hasDiagnosis`, `mayInteractsWith`, `hasSymptom`, `ongoingMedication`, `hasVitalSign`, `hasAdverseEffect`, and among others. Each object property is associated with the specified domain and range classes, defining the relationship between different entities. For instance, the object property `hasDiagnosis` connects entities from the `Person` class to the `Diagnosis` class, indicating the relationship between a person and a diagnosis.

On the other hand, data properties are used to link entities with their corresponding property values in COViDRO, capturing specific attributes associated with the entities. COViDRO comprises 15 data properties, each linked to their respective domain classes and range data types. Examples of data properties in COViDRO include `BMI`, `dose`, `drugCode`, `route`, `strength`, and more. These data properties provide essential information about the characteristics or measurements associated with the entities within the ontology. By incorporating both object properties and data properties, COViDRO creates a comprehensive framework for representing and analyzing COVID-19-related information with rich connections and detailed attributes. An excerpt of Object property and data property hierarchy from the Protégé interface is provided in Figures 8a and 8b respectively.

5 Role of Defined Classes in COViDRO Design: Leveraging Description Logic (DL)

At the core of COViDRO's design is the "defined classes" [92], which assume a pivotal role in the representation and organization of domain knowledge. Within the COViDRO model, a total of 45 distinct defined classes have been established. These defined classes can be accessible from COViDRO vocabulary (`https://w3id.org/codo/1.4`). These defined classes are constructed using OWL-DL [75]. These classes serve as precise formal representations of specific concepts, accompanied by necessary and sufficient conditions for membership classification. Figure 9 presents an illustrative excerpt of a class hierarchy and its corresponding defined class, as exhibited within the Protégé interface. The notion of defined classes assumes exceptional significance, primarily due to its role in enabling COViDRO to perform automated reasoning and inference. This empowerment grants the ontology the capability to assess diverse patient attributes and subsequently propose suitable therapeutic interventions based on *PRADiCT*. For this purpose, COViDRO introduces six distinctive scenarios, each formulated as defined classes. These scenarios encapsulate pertinent patient factors that guide the recommendations for treatment. These six scenarios, coupled with their corresponding equation numbers as listed in Table no. 11, encompass the following:

1. By Risk Level: This situation stratifies patients into categories based on their risk level, encompassing high-risk (Eq. 1 in Table 6), low-risk, or moderate-risk classifications. These categories serve as the foundation for tailored treatment recommendations contingent upon the individual's risk profile. DL underpinned the creation of the `HighRiskPatient` class, formalizing conditions related to underlying health conditions, pregnancy, occupation, social history, and age. The `HighRiskPatient` class encompasses individuals with elevated risk levels. Leveraging DL-driven reasoning, the system automatically identifies patients meeting these criteria, equipping healthcare professionals with insights into individuals at heightened risk due to diverse factors. The `HighRiskPatient` class can be described as a subset of individuals belonging to the `foaf:Person` class, characterized by specific conditions. These conditions encompass instances of having an underlying health condition linked to `COVIDRO_Pregnant` via the `COVIDRO_hasUnderlyingHealthCondition` property or exhibiting the `COVIDRO_isPregnant` property set to `true`.

Fig. 6: An excerpt of the class hierarchy for the class Pharmaceutical preparation.

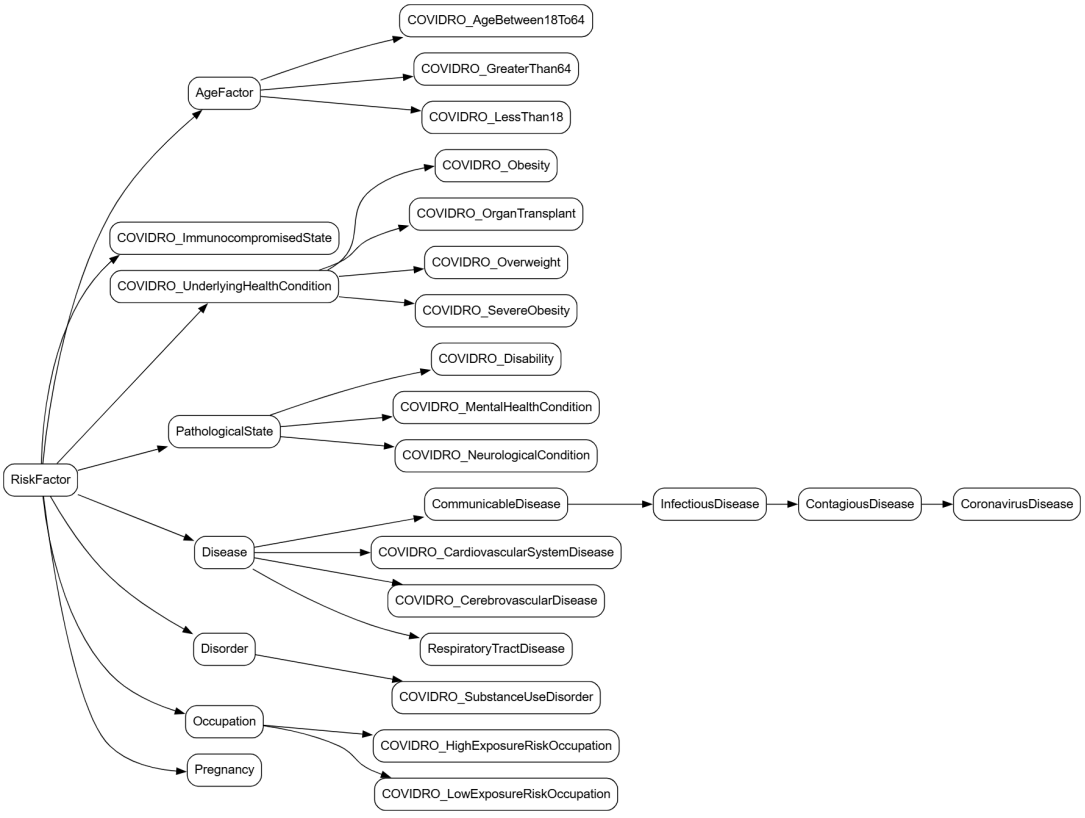
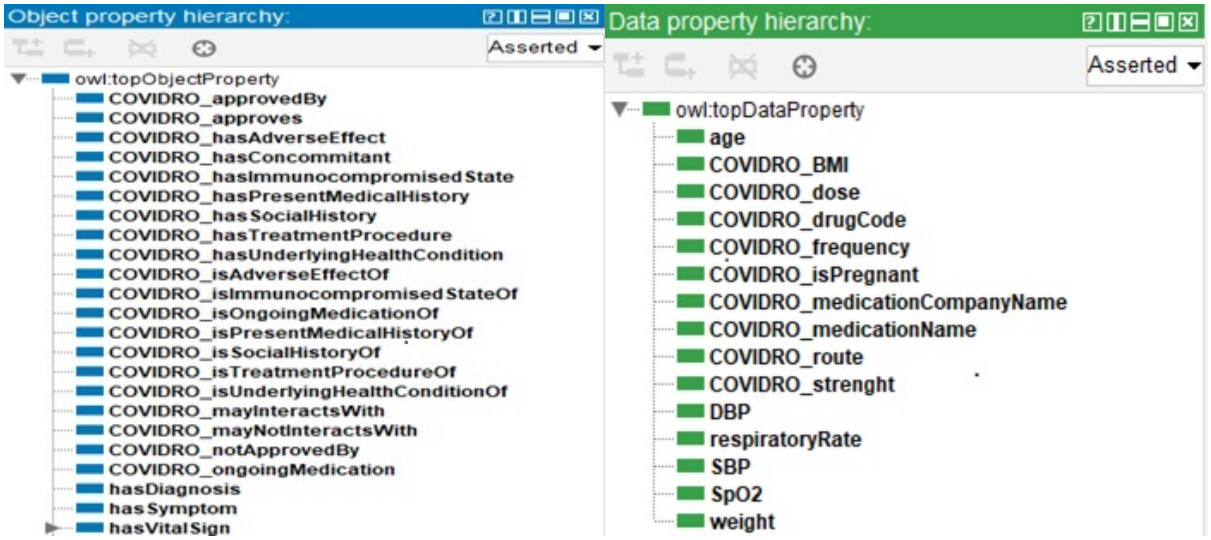


Fig. 7: An excerpt of the class hierarchy for the class Risk factor.



(a) An excerpt of COViDRO's object property hierarchy. (b) An excerpt of COViDRO data property hierarchy.

Fig. 8: A partial view of COViDRO object property and data property hierarchy

Furthermore, membership in this class extends to individuals with specific occupations falling within the `COVIDRO_HighExposureRiskOccupation` category, as indicated by the property `schema:hasOccupation`. Additionally, individuals with an immunocompromised state, denoted by the property `COVIDRO_-hasImmunocompromisedState` associated with `COVIDRO_ImmunocompromisedState`, are also encompassed within the `HighRiskPatient` class. Inclusion criteria extend to individuals with underlying health conditions, coupled with various conditions represented by the `COVIDRO_UnderlyingHealthCondition` property. Specific social history attributes, such as `COVIDRO_Alcohol`, `COVIDRO_PhysicalInactivity`, or `COVIDRO_Smoking`, further classify individuals as HighRiskPatient. Finally, individuals aged 65 or older, denoted by the `age` property with a decimal value greater than or equal to 65, are also embraced within the `HighRiskPatient` class. The `HighRiskPatient` class encapsulates individuals under the `foaf:Person` category who meet these conditions, indicating a higher risk profile in diverse medical contexts. In Table 6, DL representation captures the equivalence between the `HighRiskPatient` class and the collection of individuals delineated by the `foaf:Person` class who adhere to stipulated conditions involving diverse properties and values.

Table 6: DL Representation of the HighRiskPatient Class: Formalization of risk factors, including underlying health conditions, pregnancy, occupation, immunocompromised state, social history attributes, and age, for stratifying individuals into high-risk categories based on COVID-19 vulnerability

Component	**Description**
Classes	`HighRiskPatient, foaf:Person, COVIDRO_UnderlyingHealthCondition, COVIDRO_ImmunocompromisedState`
Properties	`COVIDRO_hasUnderlyingHealthCondition, COVIDRO_isPregnant, schema:hasOccupation, COVIDRO_hasImmunocompromisedState, COVIDRO_hasSocialHistory`
TBox	`HighRiskPatient` $\equiv$ `foaf:Person` $\sqcap$ `(((`$\exists$`COVIDRO_hasUnderlyingHealthCondition.COVIDRO_Pregnant)` $\sqcup$`(`$\exists$`COVIDRO_isPregnant.true))` $\sqcup$`(`$\exists$`schema:hasOccupation.COVIDRO_HighExposureRiskOccupation)` $\sqcup$ `(`$\exists$`COVIDRO_hasImmunocompromisedState.COVIDRO_ImmunocompromisedState)` $\sqcup$ `(`$\exists$`COVIDRO_hasUnderlyingHealthCondition.COVIDRO_UnderlyingHealthCondition)` $\sqcup$ `(`$\exists$`COVIDRO_hasSocialHistory.COVIDRO_Alcohol)` $\sqcup$ `(`$\exists$`COVIDRO_hasSocialHistory.COVIDRO_PhysicalInactivity)` $\sqcup$ `(`$\exists$`COVIDRO_hasSocialHistory.COVIDRO_Smoking)` $\sqcup$ `(`$\exists$`age.xsd:decimal[`$\geq$` 65]))` .....(Eq. 1)

2. By Drug Interaction: This scenario scrutinizes the patient's medical history, encompassing ongoing medication, to identify potential interactions with the chosen COVID-19 therapeutic regimen. Such insights are instrumental in selecting appropriate therapeutics that do not conflict with the patient's existing medication. DL facilitated the formulation of the `RecommendedTherapeuticsRitonavir-BoostedNirmatrelvir` class, character-

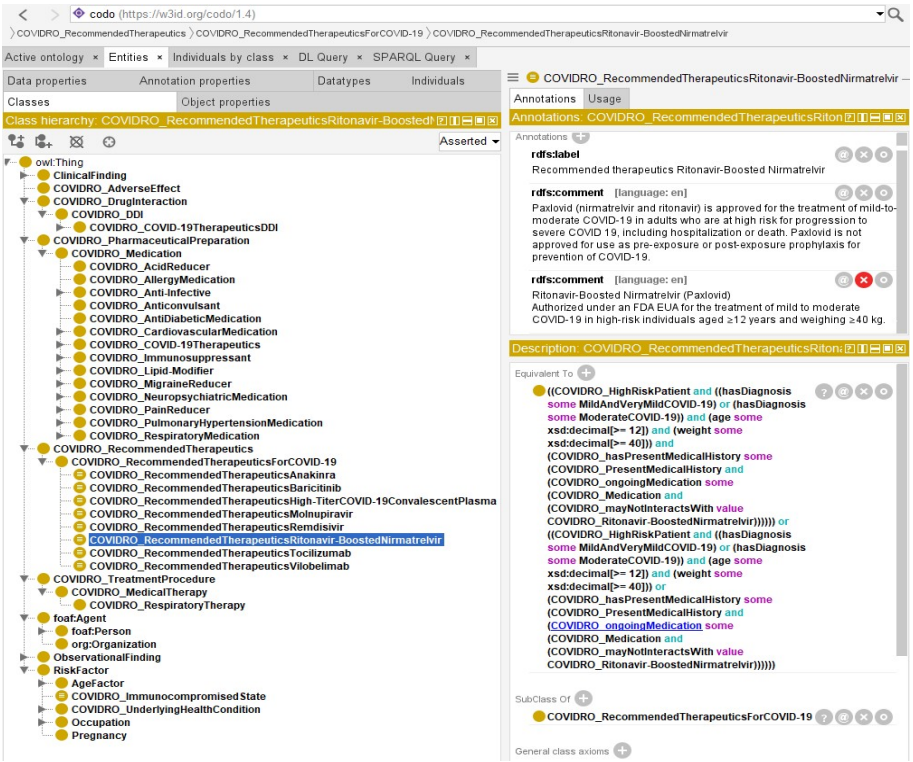


Fig. 9: An excerpt of COViDRO class hierarchy and a defined class.

ized by intricate conditions involving diagnosis, age, weight, medical history, and ongoing medication interactions (refer to Eq. 2 in Table 7).

Through DL-powered reasoning, healthcare professionals can swiftly assess the potential interactions between a patient's ongoing medication and the recommended therapeutic, offering valuable guidance toward treatment options that harmonize with the current medications. The recommendation class `RecommendedTherapeuti-csRitonavir-BoostedNirmatrelvir` encompasses two primary conditions. In the first condition, `RecommendedTherapeuticsRitonavir-BoostedNirmatrelvir` class can be described as a subset of individuals belonging to the `COVIDRO_HighRiskPatient` class and meets specific criteria: they must be diagnosed with `Mild-AndVeryMild-COVID-19` or `ModerateCOVID-19`, be aged 12 years or older, and have a weight of 40 or higher.

Furthermore, they need to possess both a present medical history and ongoing medication, specifically, some `COVIDRO_Medication`, which should not have interactions with `COVIDRO_Ritonavir-BoostedNirmatrelvir`. This condition covers individuals who meet these requirements, and also whether there could be interactions between their ongoing medication and `COVI-DRO_Ritonavir-BoostedNirmatrelvir`. The inclusion of the second condition accounts for cases where the patient's present medical history lacks any ongoing medication, eliminating the possibility of interactions with `COVIDRO_Ritonavir-BoostedNirmatrelvir`. In such cases, the patient would be recommended the drug `Ritonavir-BoostedNirmatrelvir` if they satisfy the conditions of diagnosis, age, and weight.

3. By COVID-19 Diagnosis: This situation categorizes patients based on their COVID-19 diagnosis, differentiating between mild, moderate, and severe (Eq. 3 in Table 8) cases. The ontology is capable of suggesting tailored therapeutics contingent upon the severity of the patient's condition. DL paved the way for the creation of the `SevereCOVID-19` class, reliant on symptomatology and vital sign measurements. Through DL-driven reasoning, healthcare professionals can discern cases where patients exhibit severe symptoms and specific vital sign measurements, leading to insights into the gravity of COVID-19 cases and the formulation of suitable treatment strategies. The class `SevereCOVID-19` consists of individuals falling under the category `COVID-19Diagnosis` and meeting specific conditions. These conditions include being a `foaf:Person` who displays a combination of symptoms associated with COVID-19, namely, either experiencing symptoms like `COVIDRO_-BreathingWithoutDifficulty`, `COVIDRO_Cough`, `COVIDRO_Fever`, `COVIDRO_Pneumonia`, `COVIDRO_SoreThroat`, or `COVIDRO_URTI`, or presenting with both `COVIDRO_Pneumonia` and specific vital sign measurements.

These measurements involve the `mostRecentVitalSign` which includes instances of `VitalSign` where either the diastolic blood pressure (DBP) is less than 60 or the systolic blood pressure (SBP) is less than 90, and the oxygen saturation (SpO2) is less than 90. Individuals satisfying these conditions are considered to have `SevereCOVID-19`. In simpler terms, `SevereCOVID-19` denotes instances of COVID-19 diagnosis linked to individu-

Table 7: DL Representation of the RecommendedTherapeuticsRitonavir BoostedNirmatrelvir Class: Formalization of conditions involving diagnosis, age, weight, medical history, and ongoing medication to evaluate potential interactions with Ritonavir-Boosted Nirmatrelvir, aiding in the selection of appropriate therapeutics for COVID-19 patients.

Component	Description
Classes	RecommendedTherapeuticsRitonavir-BoostedNirmatrelvir, COVIDRO_HighRiskPatient, MildAndVeryMildCOVID-19, ModerateCOVID-19, COVIDRO_PresentMedicalHistory, COVIDRO_Medication, COVIDRO_Ritonavir-BoostedNirmatrelvir
Properties	hasDiagnosis, age, weight, COVIDRO_ongoingMedication, COVIDRO_hasPresentMedicalHistory, COVIDRO_mayNotInteractsWith
TBox	RecommendedTherapeuticsRitonavir-BoostedNirmatrelvir ≡ ((COVIDRO_HighRiskPatient ⊓ ((∃hasDiagnosis.MildAndVeryMildCOVID-19) ⊔ (∃hasDiagnosis.ModerateCOVID-19)) ⊓ (∃age.xsd:decimal[≥ 12]) ⊓ (∃weight.xsd:decimal[≥ 40])) ⊓ (∃COVIDRO_hasPresentMedicalHistory.(COVIDRO_PresentMedicalHistory ⊓ (∃COVIDRO_ongoingMedication.(COVIDRO_Medication ⊓ (∃COVIDRO_mayNotInteractsWith.COVIDRO_Ritonavir-BoostedNirmatrelvir)))))) ⊔ ((COVIDRO_HighRiskPatient ⊓ ((∃hasDiagnosis.MildAndVeryMildCOVID-19) ⊔ (∃hasDiagnosis.ModerateCOVID-19)) ⊓ (∃age.xsd:decimal[≥ 12]) ⊓ (∃weight.xsd:decimal[≥ 40])) ⊓ (∃COVIDRO_hasPresentMedicalHistory.(COVIDRO_PresentMedicalHistory ⊓ (∃COVIDRO_ongoingMedication.(COVIDRO_Medication ⊓ (∃COVIDRO_mayNotInteractsWith.COVIDRO_Ritonavir-BoostedNirmatrelvir))))) .....(Eq. 2)

als experiencing specific symptoms, such as difficulty in breathing, cough, fever, pneumonia, sore throat, or URTI [77]. Moreover, the presence of pneumonia symptoms is crucial. These individuals should also possess recent vital sign measurements indicating low diastolic blood pressure (DBP), or low systolic blood pressure (SBP) accompanied by low oxygen saturation (SpO2).

Table 8: DL Representation of the SevereCOVID-19 Class: Formalization of conditions based on symptomatology and vital sign measurements, including diastolic blood pressure (DBP), systolic blood pressure (SBP), and oxygen saturation (SpO2), to classify individuals with severe COVID-19 cases and support tailored therapeutic recommendations

Component	Description
Classes	SevereCOVID-19, COVID-19Diagnosis, foaf:Person, VitalSign
Properties	hasSymptom, mostRecentVitalSign, DBP, SBP, SpO2
TBox	SevereCOVID-19 ≡ COVID-19Diagnosis ⊓ (∃isDiagnosisFor.(foaf:Person ⊓ ((((∃hasSymptom.COVIDRO_BreathingWithoutDifficulty ⊓ (∃hasSymptom.COVIDRO_Cough)) ⊔ (∃hasSymptom.COVIDRO_Fever) ⊔ (∃hasSymptom.COVIDRO_Pneumonia) ⊔ (∃hasSymptom.COVIDRO_SoreThroat) ⊔ (∃hasSymptom.COVIDRO_URTI)) ⊔ (∃hasSymptom.COVIDRO_Pneumonia) ⊓ (∃mostRecentVitalSign.((VitalSign ⊓ (∃DBP.xsd:integer[< 60])) ⊔ ((VitalSign ⊓ (∃SBP.xsd:integer[< 90])) ⊓ (∃SpO2.xsd:integer[< 90])))))) .....(Eq. 3)

4. By Risk Factors: This scenario evaluates various risk factors that might impact treatment determinations, encompassing age considerations (Eq. 4 in Table 9), immunocompromised status, pregnancy, occupation (e.g., high exposure or low exposure), underlying health conditions (e.g., obesity) (Eq. 5 in Table 10), diseases, disorders, organ transplants, and mental/neurological conditions. DL facilitated the establishment of the "AgeBetween18To64" and "Obesity" classes, capturing individuals within specific age ranges and BMI values. DL-driven reasoning furnishes insights into patient demographics (age) and health attributes (obesity), aiding in the customization of treatment recommendations and the comprehension of potential risk factors. The concept of "AgeBetween18To64" is defined as encompassing all individuals represented by the "foaf:Person" class whose age property is a decimal value greater than or equal to 18 and less than 65. In simpler terms, it includes all individuals aged 18 or older but not yet 65 [76].

The concept of "Obesity" is defined as encompassing all individuals within the "foaf:Person" class who possess a recorded vital sign associated with body mass index (BMI), characterized by the "hasVitalSign" property. This recorded vital sign should be linked to the "VitalSign" class and feature a BMI value represented

Table 9: DL Representation of the AgeBetween18To64 Class: Formalization of individuals within the foaf:Person class whose age is between 18 and 64, supporting patient classification based on age-related risk factors.

Component	Description
Classes	`foaf:Person, AgeBetween18To64`
Properties	`age`
TBox	`AgeBetween18To64` $\equiv$ `foaf:Person` $\sqcap$ ($\exists$`age.xsd:decimal[`$\geq$ `18, <` $<$ `65])` .....(Eq. 4)

as an integer within the range of 30 and 40. In simpler terms, "`Obesity`" covers all individuals with a BMI between 30 and 40, indicating their classification as obese [78].

Table 10: DL Representation of the Obesity Class: Definition of individuals within the foaf:Person class possessing a Body Mass Index (BMI) between 30 and 40, facilitating identification of obesity as a risk factor in patient evaluation.

Component	Description
Classes	`foaf:Person, Obesity, VitalSign`
Properties	`hasVitalSign, COVIDRO_BMI`
TBox	`Obesity` $\equiv$ `foaf:Person` $\sqcap$ ($\exists$`hasVitalSign.(VitalSign` $\sqcap$ ($\exists$`COVIDRO_BMI.xsd:integer[`$\geq$ `30,` $<$ `40])))` .....(Eq. 5)

5. By Treatment Procedure: This situation addresses various treatment procedures, including medical therapy and respiratory therapy, to provide suitable therapeutic recommendations for specific medical interventions.

6. By Drug Adverse Effects: This scenario identifies potential adverse effects of specific COVID-19 drugs (e.g., anaphylaxis, dysgeusia, and nausea), facilitating healthcare professionals in understanding and managing adverse reactions. COViDRO synergizes the insights extracted from all six scenarios to deliver well-matched COVID-19 therapeutic recommendations. A concrete example of the formal definition of this recommendation class can be located in Eq. 2 in Table 7. By amalgamating all the factors of *PRADiCT*, COViDRO's RecommendedTherapeutics class becomes an invaluable tool for informed decision-making.

Across each scenario, DL plays a critical role in structuring data into formal classes, properties, and axioms. This organized structure underpins automated reasoning, identifying individuals or cases adhering to predefined criteria. As a result, healthcare professionals gain insights to facilitate informed choices, personalized treatments, and nuanced comprehension of patient contexts. By leveraging DL-driven ontologies, healthcare practitioners efficiently navigate complex scenarios, drawing meaningful insights from data. Through the interplay of defined classes and corresponding axioms, COViDRO scrutinizes patient data, considering diverse scenarios to furnish personalized treatment recommendations based on individual patient attributes. The ontology's capacity for automated reasoning can empower healthcare professionals to make informed decisions, ultimately enhancing patient care within the context of the COVID-19 pandemic.

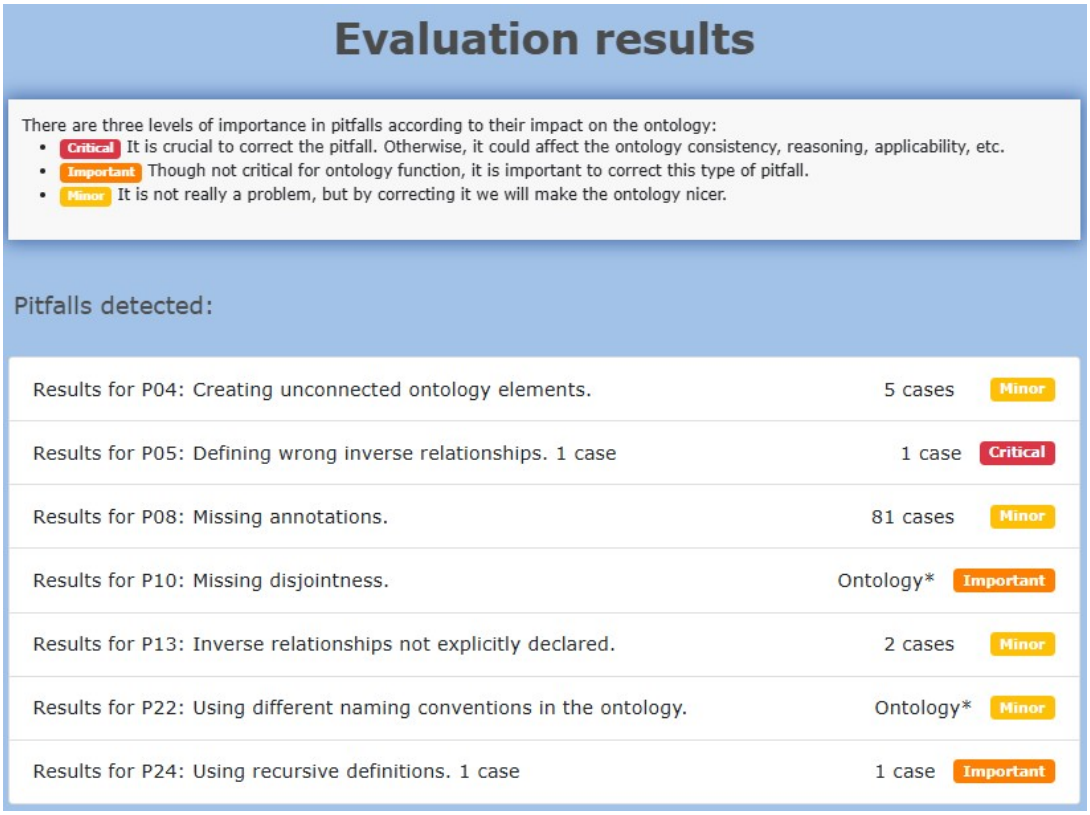
Table 11 presents a partial overview of different situations within the COViDRO framework, their associated defined classes (provides a representative sample rather than an exhaustive list of 45 defined classes), the role of DL in formalizing conditions, and their application in healthcare decision-making. Table 11 also includes relevant DL expressions, equation numbers, and references for the defined classes' concepts.

6 Result evaluation of COViDRO

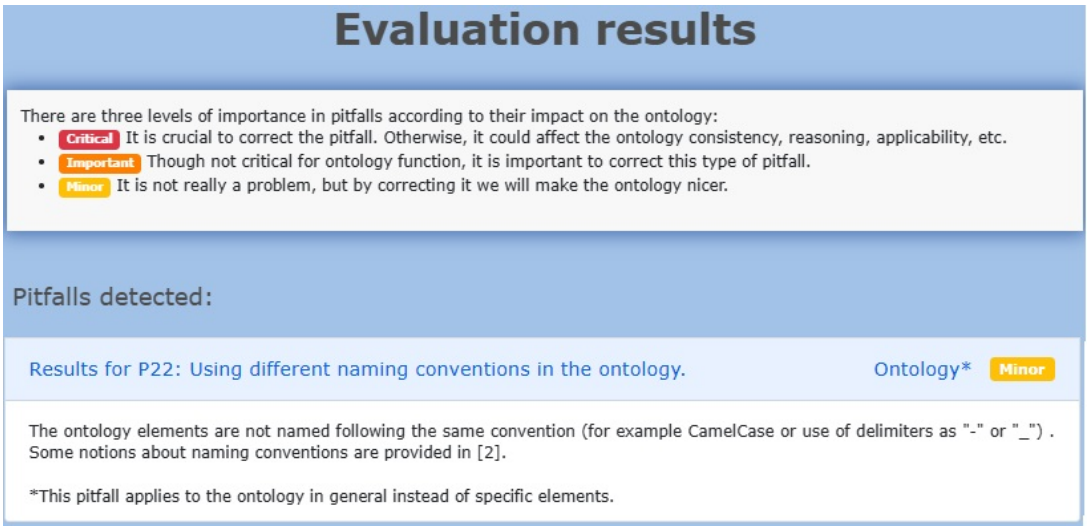
6.1 Quality-based Evaluation

The OOPS! Scanner has played a crucial role in evaluating the quality, structural integrity, and comprehensiveness of the COViDRO ontology. This assessment is conducted through a two-tier evaluation: (1) simple evaluation and (2) advanced evaluation, as discussed S9 in Section 3 and illustrated in S9.1 of Table 4.

In the simple evaluation, pitfalls are identified within COViDRO alone. The scanner detected various pitfalls categorized as "`critical`", "`important`", and "`minor`" as illustrated in Figure 10a. These pitfalls include unconnected ontology elements, wrong inverse properties, missing annotations and disjoint axioms, undeclared inverse relationships, using different naming conventions, and using recursive definitions. Guided by the scanner's evaluation report, significant efforts are undertaken to resolve critical and important issues, resulting in substantial improvements to the model. However, one minor pitfall persists, as shown in Figure 10b. For instance, the



(a) OOPS! Evaluation Results before modification



(b) OOPS! Evaluation Results after modification

Fig. 10: Comparison of evaluation results before and after modification

Table 11: A partial overview of COVIDRO Situation Analysis, Defined Classes, and DL Role

Situation	Description	Defined Class	DL Role	DL Expression	Application	DL Expression Equation No.	Defined Class' Concept Ref No.
By Risk Level	Differentiates patients based on risk level: high, low, or moderate.	High risk patient	Formalizes risk conditions	Defined by conditions involving health, pregnancy, occupation, age, and more.	Tailored treatment recommendations based on risk profile.	1	[5]
By Drug Interaction	Analyzes patient's medical history for drug interactions.	Recommended therapeutics Ritonavir-Boosted Nirmatrelvir	Ensures safe drug usage	Conditions involving diagnosis, age, weight, medical history, and interactions.	Identifies safe therapeutics considering medication interactions.	2	[9]
By COVID-19 Diagnosis	Categorizes patients by COVID-19 severity.	Severe COVID-19	Identifies severe cases	Instances with severe symptoms and vital sign patterns.	Recommends specific therapeutics for severe cases.	3	[64,77]
By Risk Factors	Assesses risk factors: age, immunity, occupation, obesity, etc.	Age between 18 to 64, Obesity	Profiles risk factors	Conditions involving age ranges and BMI values.	Considers demographics and health conditions for tailored treatments.	4, 5	[63,65] [76,78]
By Treatment Procedure	Addresses treatment procedures: medical, respiratory therapy.	-	Guides treatment choice	-	Recommends therapies suitable for specific medical interventions.	-	-
By Drug Adverse Effects	Identifies adverse effects of COVID-19 drugs.	-	Alerts potential side effects	-	Informs healthcare professionals about drug-related risks.	-	-

scanner highlights a minor pitfall concerning naming conventions for entities and properties, where the ontology adheres to CapitalizeEachWord for classes and camelCase for properties.

To further benchmark COViDRO against other related ontologies, an advanced evaluation has been conducted. This involves a comparative analysis based on structural, functional, and usability dimensions, as detailed in [79] and further illustrated in Table 12. The table shows the 41 pitfalls across the three dimensions and specifies their definitions. The full pitfalls catalog can be accessed at (http://oops.linkeddata.es/catalogue.jsp).

The evaluation includes twelve related ontologies (O1-O12) alongside COViDRO (O13), as shown in Table 13. The "x" symbol denotes the absence of a pitfall, while "Yes" indicates the presence of a pitfall where the occurrence is not quantifiable. The numerical values indicate the exact count of occurrences for quantifiable pitfalls within a given ontology.

Unlike some other ontologies (e.g., O3, O7, O10), which show issues related to inferencing (P11, P13) or modeling decisions (P07, P24), COViDRO effectively avoids these problems. This confirms its strong reasoning support and well-defined structure. Additionally, COViDRO does not suffer from ontology language pitfalls (P34, P38), reinforcing its adherence to best practices in conceptual modeling. COViDRO does not exhibit pitfalls related to application context (P36, P37, P38) or requirement completeness (P04), aligning with ontologies such as O1, O5, O8, and O9. This suggests that COViDRO successfully meets domain-specific requirements and provides a well-structured knowledge representation.

COViDRO maintains high ontology clarity (P08 is absent) and ontology understanding (P08, P11, P13 are absent), meaning it avoids ambiguities commonly found in other ontologies. However, Pitfall P22 (naming inconsistencies) is present, likely due to minor variations in terminology from reused external sources. This is similar to O1, O3, O7, and O12, which also exhibit naming inconsistencies. COViDRO demonstrates strong

Table 12: Classification of pitfalls based on ontology evaluation dimensions with pitfall definitions

Dimension/Category	Pitfalls and Definitions
Structural Dimension	
Modelling Decisions	– P02. Creating synonyms as classes – P03. Creating the relationship "is" instead of using "subclassOf", "instanceOf" or "sameIndividual" (Creating the relationship "is" instead of using "rdfs:subClassOf", "rdf:type" or "owl:sameAs") – P07. Merging different concepts in the same class – P21. Using a miscellaneous class – P24. Using recursive definitions – P25. Defining a relationship as inverse to itself – P26. Defining inverse relationships for symmetric ones – P33. Creating trivial property chains (single property)
Wrong Inference	– P05. Defining incorrect inverse relationships – P06. Including cycles in class hierarchy – P19. Defining multiple domains/ranges for properties – P27. Defining incorrect equivalent properties – P28. Defining incorrect symmetric relationships – P29. Defining incorrect transitive relationships – P31. Defining incorrect equivalent classes
No Inference	– P11. Missing domain/range in properties – P12. Not declaring equivalent properties – P13. Not declaring inverse relationships – P30. Not declaring equivalent classes
Ontology Language	– P34. Untyped class – P35. Untyped property – P38. Missing OWL ontology declaration
Functional Dimension	
Real World Modelling	– P04. Creating unconnected ontology elements – P10. Missing disjointness declarations
Requirements Completeness	– P04. (defined earlier) – P09. Missing domain information
Application Context	– P36. URI contains file extension – P37. Ontology not web-accessible – P38. (defined earlier) – P39. Ambiguous namespace – P40. Namespace hijacking
Usability Dimension	
Ontology Clarity	– P08. Missing annotations – P22. Inconsistent naming conventions
Ontology Understanding	– P02., P07., P08., P11., P12., P13., P37. (defined earlier) – P20. Misusing ontology annotations – P32. Multiple classes with same label
Ontology Metadata	– P38. (defined earlier) – P41. No license declared

reasoning support, well-structured modeling, and clarity, making it a robust ontology for healthcare knowledge representation.

Table 13: Quality-based evaluation of existing COVID-19 ontologies pitfall with COViDRO. O1: COVID-19 Surveillance Ontology, O2: CIDO-COVID-19, O3: COVIDCRFRAPID, O4: DRUGS4COVID19, O5: ROC Ontology, O6: COVID-19 Ontology, O7: CODO, O8: DrOn, O9: CIRO, O10: DINTO, O11: ODAE, O12: Knowledge4COVID-19, O13: COVIDRO. "Yes" indicates the presence of a pitfall where the occurrence is not quantifiable, numeric value represents the exact count of occurrences for quantifiable pitfalls and "x" confirms that the pitfall was checked but not found.

Pitfall		**Ontologies**												
		O1	**O2**	**O3**	**O4**	**O5**	**O6**	**O7**	**O8**	**O9**	**O10**	**O11**	**O12**	**O13**
Structural Dimension														
Modelling Decision	P02	x	x	x	x	x	x	x	x	x	x	x	x	x
	P03	x	x	x	x	x	x	x	x	x	x	x	x	x
	P07	x	x	5	x	x	x	x	x	x	x	x	x	x
	P21	x	x	10	x	x	x	x	x	x	x	x	x	x
	P24	x	x	1	x	x	x	1	3	x	x	x	x	x
	P25	x	x	x	x	x	x	x	x	x	x	x	x	x
	P26	x	x	x	x	x	x	x	x	x	x	x	x	x
	P33	x	x	x	x	x	x	x	x	x	x	x	x	x
Wrong inference	P05	x	x	x	x	x	x	x	x	x	x	x	x	x
	P06	x	x	x	x	x	x	x	x	x	x	x	x	x
	P19	x	x	x	x	x	x	x	x	x	x	x	1	x
	P27	x	x	x	x	x	x	x	x	x	x	x	x	x
	P28	x	x	x	x	x	x	x	x	x	x	x	x	x
	P29	x	x	x	x	x	x	x	x	x	x	x	x	x
	P31	x	x	x	x	x	x	x	x	x	x	x	2	x
No inference	P11	2	x	11	33	x	x	39	x	55	x	x	x	x
	P12	x	x	x	x	x	x	x	x	x	x	x	x	x
	P13	35	x	5	5	x	x	34	x	x	6	77	37	x
	P30	x	x	4	x	x	x	1	x	x	x	x	x	x
Ontology Language	P34	x	x	1	x	x	x	7	x	x	x	x	x	x
	P35	x	x	x	x	x	x	x	x	x	x	x	x	x
	P38	x	x	x	x	x	x	x	x	x	x	x	x	x
Functional Dimension														
Real World Modeling	P04	4	x	3	2	x	x	1	x	x	x	x	1	x
	P10	Yes	x	Yes	Yes	x	x	Yes	Yes	x	Yes	x	Yes	x
Requirement Completeness	P04	4	x	3	2	x	x	1	x	x	x	x	1	x
	P09	x	x	x	x	x	x	x	x	x	x	x	x	x
Application Context	P36	x	Yes	x	x	x	Yes	x	x	x	x	x	x	x
	P37	x	Yes	x	x	Yes	Yes	x	x	x	x	x	x	x
	P38	x	x	x	Yes	x	x	x	x	x	x	x	x	x
	P39	x	x	x	x	x	x	x	x	x	x	x	x	x
	P40	x	x	x	x	x	x	x	x	x	x	x	x	x
Usability-Profiling Dimension														
Ontology Clarity	P08	236	x	365	x	x	x	52	102	x	17	x	105	x
	P22	Yes	x	Yes	x	x	x	Yes	x	x	x	x	Yes	Yes
Ontology Understanding	P02	x	x	x	x	x	x	x	x	x	x	x	x	x
	P07	x	x	5	x	x	x	x	x	x	x	x	x	x
	P08	236	x	365	x	x	x	52	102	x	17	x	105	x
	P11	2	x	11	33	x	x	39	x	55	x	x	x	x
	P12	x	x	x	x	x	x	x	x	x	x	x	x	x
	P13	35	x	5	5	x	x	34	x	x	6	77	37	x
	P20	x	x	3	x	x	x	2	x	x	x	x	x	x
	P32	x	x	40	x	x	x	x	x	x	x	x	x	x
	P37	x	Yes	x	x	Yes	Yes	x	x	x	x	x	x	x
Ontology Metadata	P38	x	x	x	x	x	x	x	x	x	x	x	x	x
	P41	Yes	x	x	x	x	x	x	x	Yes	x	Yes	Yes	x

A summary of the comparative results has been provided in Figure 11, where the x-axis represents different ontology frameworks (O1 to O13), and the y-axis indicates the pitfall count for each ontology. The distribution highlights significant variations in pitfalls across ontologies, with O3 and O11 exhibiting the highest counts, notably influenced by specific pitfalls such as P08 and P13 respectively. In contrast, ontologies like O2, O5, O6, and O13 show inconsiderable pitfalls, indicating a relatively cleaner structure. O9 and O11 display moderate pitfall counts, with a noticeable presence of P11 and P13. The stacked bar representation further reveals the diverse nature of pitfalls affecting each ontology, emphasizing areas requiring improvement for enhanced quality and interoperability. While some ontologies (O1, O3, O12) require major improvements, COVIDRO has fewer pitfalls, making it a better candidate for real-world applications.

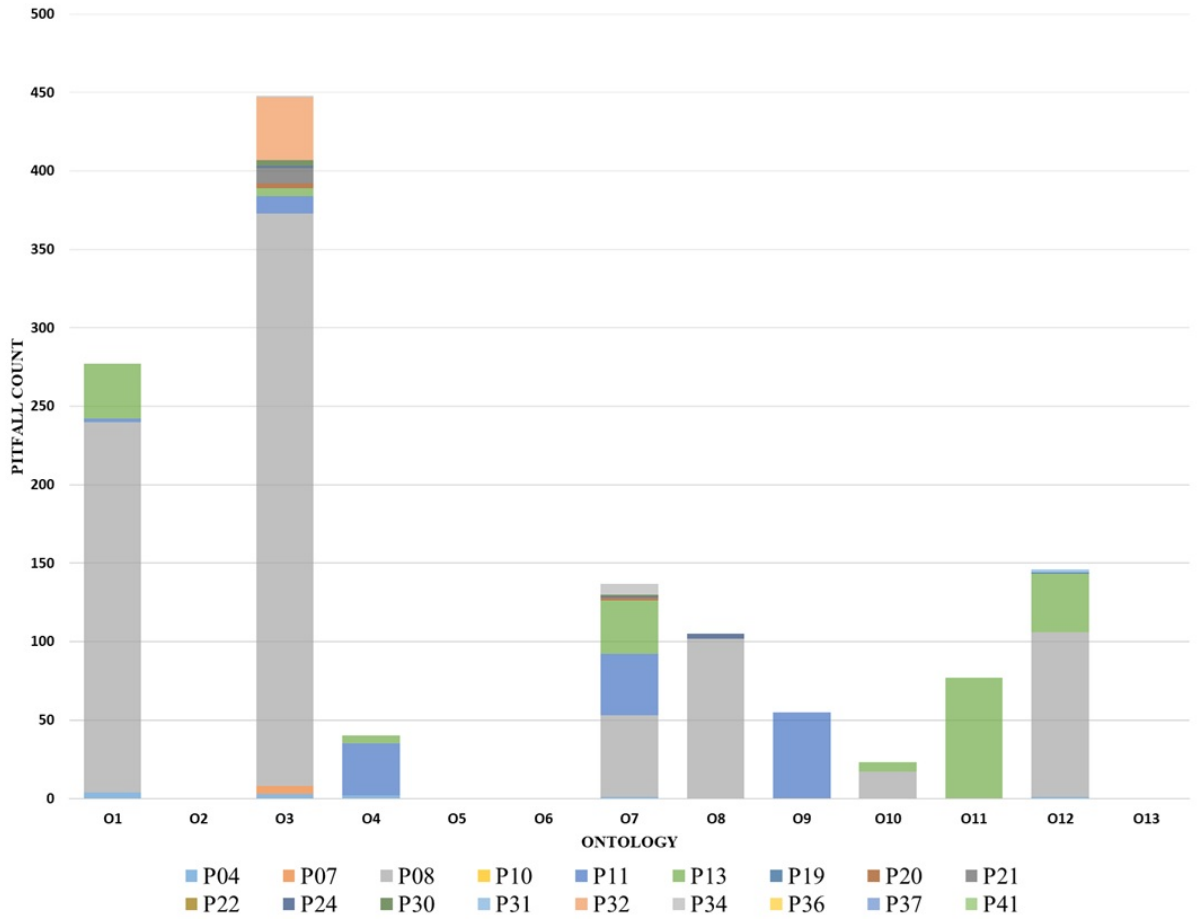


Fig. 11: OOPS! scanner results for quality-based evaluation of existing COVID-19 ontologies with COViDRO.

Despite minor pitfalls, the evaluation confirms that COViDRO achieves a high level of completeness and usability. The structured quality assessment and comparative benchmarking validate the ontology's effectiveness in supporting open healthcare metadata and interoperability within the domain.

6.2 Structural and Logical Validation

To ensure the correctness, structural integrity, and logical soundness of COViDRO, we conducted a rigorous evaluation using OntoDebug and Pellet Reasoner (as discussed S9 in Section 3 and illustrated in S9.2 of Table 4) within COViDRO alone. Each tool played a distinct role in verifying different aspects of the ontology's quality. At this stage, we focused solely on COViDRO rather than including other related ontologies because our primary objective is to refine and validate its internal structural integrity and logical consistency before benchmarking it against external models.

i. Structural Integrity and Debugging with OntoDebug : To detect and resolve potential modeling errors, we employed OntoDebug [69], a test-driven ontology debugging tool available as a Protégé plug-in. OntoDebug facilitates fault detection and repair by leveraging interactive debugging techniques, ensuring that the

ontology adheres to its intended design. Through debugging, the tool ensured that no faulty axioms were present, thereby confirming the structural integrity, coherence, and correctness of COViDRO. The evaluation process is visually demonstrated in Figure 12.

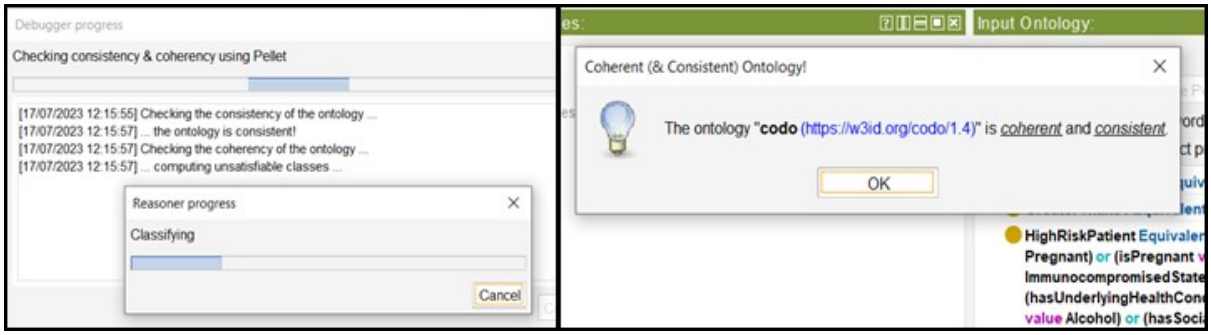


Fig. 12: Structural Integrity and Debugging with OntoDebug.

ii. Logical Consistency Checking and Inference with Pellet Reasoner : In addition to debugging, we used the Pellet Reasoner [70], an OWL-DL reasoner, to perform logical consistency checking on COViDRO. Pellet systematically examined the ABox and TBox reasoning, ensuring that the ontology adheres to formal logical constraints without contradictions. The reasoner detected no inconsistencies or errors, affirming the high level of correctness in the model.

Furthermore, Pellet's classification and inference capabilities validated the hierarchical structure of COViDRO. An exemplary inference result from the ontology is depicted in Figure 13. By integrating OntoDebug

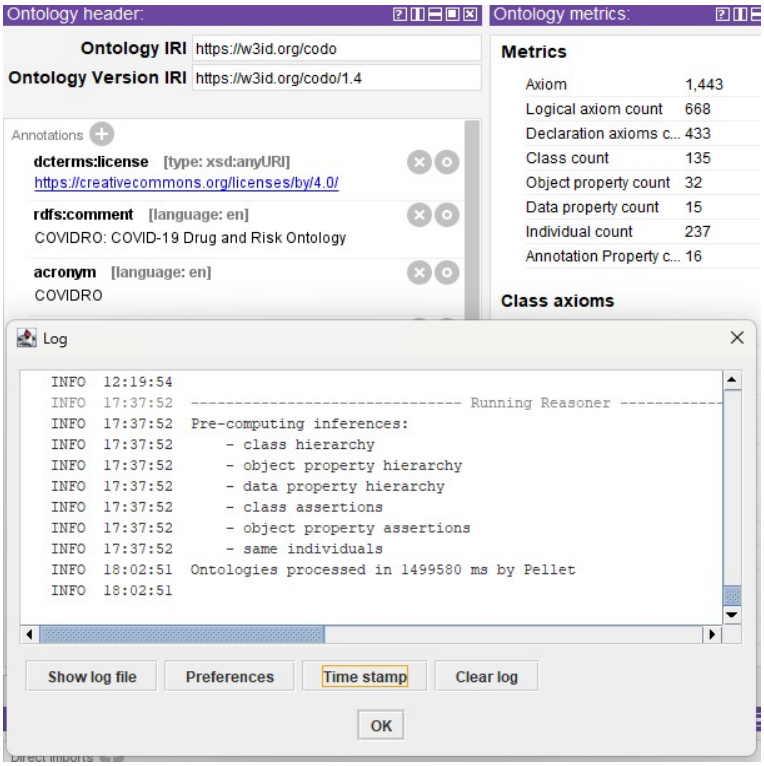


Fig. 13: Logical Consistency Checking and Inference with Pellet Reasoner.

for structural debugging and Pellet for logical validation, we ensured that COViDRO is a coherent, consistent, and well-structured ontology. The combined use of interactive debugging and automated reasoning guarantees that

the ontology is free from modeling errors, logically sound, and capable of supporting reliable knowledge-based applications.

6.3 Quantitative Evaluation

To assess the structural complexity and semantic richness of COViDRO, we employed the OntoMetrics tool [74], a metric-based tool for evaluating ontologies using predefined statistical measures (as discussed S9 in Section 3 and illustrated in S9.3 of Table 4). This tool facilitates the quantitative assessment of ontologies by categorizing features into different metric groups, such as Schema, Instance, Base, Graphs, and Individual Axioms. This quantitative evaluation of COViDRO involves measuring key attributes within the ontology, including node structure, depth, breadth, and the total number of levels. These metrics provide insights into the complexity, connectivity, and efficiency of the ontology's representation of healthcare-related knowledge. However, this method does not identify inconsistencies or anomalies within the ontology, which has been fulfilled by the previous step 6.2.

Table 14 provides a comparative evaluation of COViDRO with other ontologies (O1-O13) based on the OntoMetrics framework.

Table 14: Evaluation of Ontologies Based on Various Metrics. NA values indicate metrics unavailable due to parsing errors during validation.

Metric Category	**Metric**	**Ontologies (O1-O13)**												
		O1	O2	O3	O4	O5	O6	O7	O8	O9	O10	O11	O12	O13
Base Metrics	Axioms	161	366	2,791	241	NA	38,987	1,942	32,36,696	531	6,08,214	77,544	965	1,443
	Logical Axioms	28	152	866	32	NA	2,622	899	14,98,183	211	3,07,380	15,772	285	668
	Class Count	32	165	396	11	NA	2,270	90	7,29,094	89	28,178	3,673	119	135
	Object Property	0	13	6	17	NA	9	73	130	34	73	130	37	32
	Data Property	0	0	7	20	NA	1	50	1	3	17	18	49	15
	Individuals	0	0	85	0	NA	6	271	27	14	19,898	550	0	237
Class Axioms	Subclass Axioms	28	148	765	10	NA	2,611	74	14,97,760	113	1,61,988	14,573	69	127
	Equivalent Class	0	3	0	0	NA	7	9	79	18	11,592	315	45	45
Schema Metrics	Attribute Richness	0	0	0.018	1.818	NA	0.0004	0.556	0.000004	0.034	6.71	0.0049	0.412	0.111
	Inheritance Richness	0.875	0.897	1.932	0.909	NA	1.150	0.822	2.05	1.270	5.75	3.968	0.580	0.941
	Relation Richness	0	0.103	0.008	0.630	NA	0.007	0.526	0.00015	0.315	1	0.030	0.543	0.377
Knowledge Base	Avg. Population	0	0	0.215	0	NA	0.0026	3.011	0.000037	0.157	0.71	0.150	0	1.756
	Class Richness	0	0	0.051	0	NA	0	0.311	0.000026	0.045	0	0	0	0.326

The analysis is divided into the following metric categories:

- **Base Metrics**: Base metrics capture fundamental ontology attributes: Axioms (1,443) in COViDRO indicate a moderate structural complexity compared to O8 (32,36,696 axioms) and O10 (6,08,214 axioms), but significantly higher than smaller ontologies like O1, O2, O4, O9 and O12. Logical Axioms (668) define logical relationships within the ontology, contributing to its expressiveness. Class Count (135) reflects the overall conceptual coverage of the ontology. While COViDRO is compact, it includes a diverse range of healthcare-related concepts. Object Properties (32) and Data Properties (15) suggest a balanced representation of conceptual relationships and attribute-based data properties. Individual Count (237) is relatively low, indicating that COViDRO is more schema-focused than instance-heavy ontologies like O10 (19,898 individuals).
- **Class Axioms**: Subclass Axioms (127) highlight the depth and hierarchical complexity of the ontology, demonstrating a well-structured taxonomy. Equivalent Class Axioms (45) suggest some degree of semantic overlap, which enhances concept alignment.
- **Schema Metrics**: Schema metrics evaluate the design principles and information distribution within the ontology. Attribute Richness (0.111) reflects a lower attribute-to-class ratio, suggesting that COViDRO primarily focuses on conceptual relationships rather than complex data attributes. Inheritance Richness (0.941) indicates a well-balanced class hierarchy, showing consistent depth across levels. Relation Richness (0.377)

suggests a moderate diversity in relationships, indicating that COViDRO leverages both object and data properties effectively.

– **Knowledge Base Metrics**: Average Population (1.756) signifies a reasonable distribution of individuals across classes, higher than many smaller ontologies but lower than instance-heavy ones. Class Richness (0.326) measures the distribution of instances across classes. COViDRO's value suggests a moderate level of population within defined categories, ensuring a well-structured but not overly instance-reliant framework.

Ontology O5 has not been parsed by the OntoMetric tool; therefore, all values for O5 are marked as "NA". The results of the OntoMetrics evaluation of COViDRO (O13) against other related ontologies are displayed in Figure 14. X-Axis (Horizontal) represents different ontologies (O1 to O13) evaluated using the OntoMetrics tool. Y-Axis

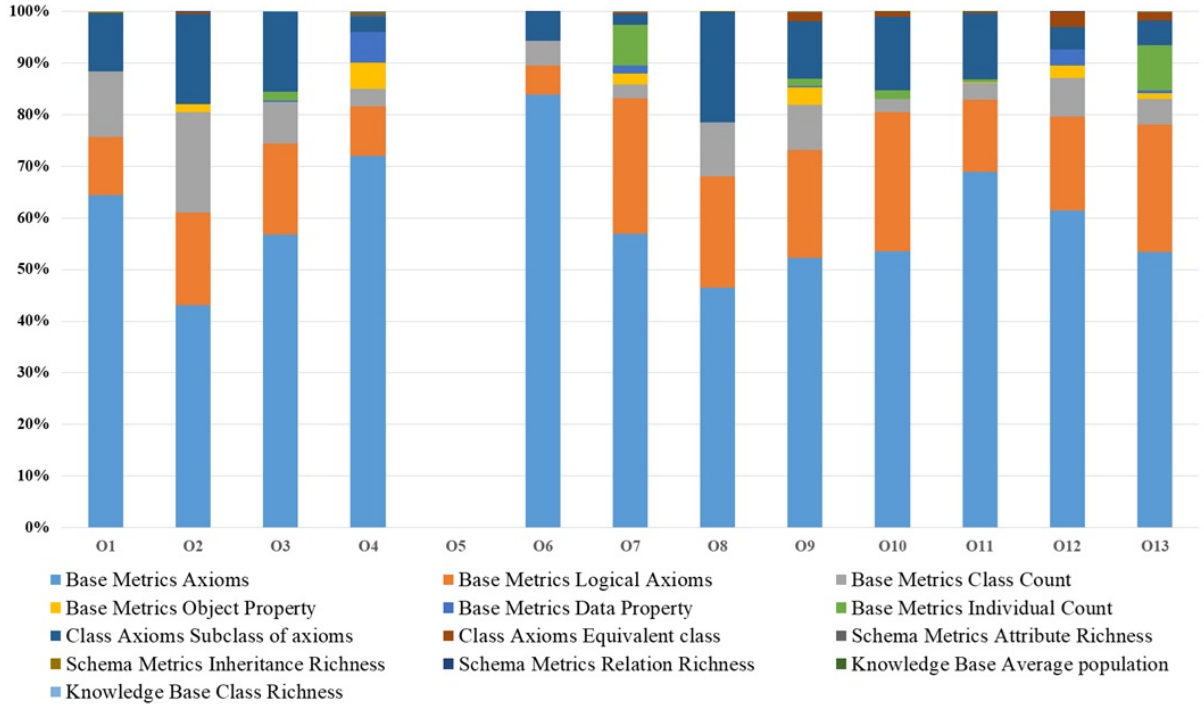


Fig. 14: OntoMetrics results for Quantity-based evaluation of existing COVID-19 ontologies with COViDRO.

(Vertical) displays percentage values (0 percent to 100 percent), representing the proportionate contribution of various OntoMetrics categories within each ontology.

Each bar in the stacked bar chart corresponds to an ontology (O1 to O13) and consists of multiple color-coded segments, each representing a different OntoMetrics category. The height of each segment reflects the relative contribution of the respective metric in the overall ontology structure. Since the metrics in Table 14 are on different scales as they include raw counts (e.g., number of axioms, class count) and ratio-based metrics (e.g., inheritance richness, relation richness). To create Figure 14, we normalized these values to a percentage scale to allow for meaningful visual comparison across different ontologies. Specifically, the values for each ontology were first summed to obtain a total metric score per ontology. Then, each metric was converted into a percentage by dividing its value by the total sum and multiplying by 100. This method ensures that all ontologies have a total contribution of 100%, allowing for direct comparison.

For example, in Ontology O1, the base metric axioms (161), logical axioms (28), class count (32), subclass axioms (28), and inheritance richness (0.875) contributed significantly to the total sum of metrics (249.875). When scaled, their respective proportions were 64.45% (axioms), 11.21% (logical axioms), and 12.81% (class count). A similar process was applied to each ontology, ensuring that metrics with different absolute scales are visually comparable in Figure 14.

COVIDRO (O13) demonstrates higher class richness and average population (light blue and dark green segments). This suggests that instances (real-world entities) are well-distributed across classes, making it a more comprehensive ontology for real-world applications than many other COVID-19 ontologies. O13 contains a high proportion of logical axioms (orange), indicating robust logical consistency and well-defined relationships among concepts. Compared to ontologies like O1, O4, O6 and O12, which have lower logical expressiveness, O13 provides a richer semantic structure for knowledge representation.

O13 has a well-balanced object property (yellow) and data property (sky blue) representation, unlike some ontologies (e.g., O3 and O6). This ensures that COVIDRO effectively integrates both concept-to-concept and concept-to-data relationships, making it more suitable for real-world healthcare applications. As compare to O1, O2, O4 and O9, COVIDRO has a high percentage of base metric axioms (light blue), indicating a strong foundational structure. However, unlike O1, and O2, which lack schema richness, O13 ensures that its foundational structure is complemented by meaningful schema attributes and class axioms. O13 includes a balanced number of subclass axioms (dark blue) and equivalent class axioms (brown). This indicates that the ontology effectively captures hierarchical depth while minimizing redundancy, unlike some ontologies (e.g., O1, O2, O4 and O9) that either overemphasize subclassing or lack proper class alignment, making those ontologies less expressive and harder to integrate.

COVIDRO outperforms many existing COVID-19 ontologies by maintaining a strong balance between logical structure, hierarchical richness, relation diversity, and instance representation. Its high class richness and structured schema make it a more effective model for real-world healthcare data management, setting it apart from other ontologies that may be either too rigid or lack sufficient expressiveness (O6, O8).

6.4 Query-based Evaluation

For the query-based evaluation of COViDRO, we selected only CODO for comparison, as it serves as the foundational base ontology (as discussed S9 in Section 3 and illustrated in S9.4 of Table 4). Our goal is to test COViDRO's ability to enhance and extend CODO's epidemiological framework by incorporating personalized therapeutic decision-making. At this stage, we have only targeted the base ontology, CODO, and have not included other related ontologies to maintain a focused evaluation of COViDRO's semantic reasoning capabilities.

Following the execution of the Pellet reasoner in Section 6.2 (ii), the inferred ontology is generated and downloaded for SPARQL-based evaluation [71]. Since various COViDRO classes were constructed using Description Logic (DL), they contain asserted axioms that contribute to inferred knowledge. To assess this inferred knowledge, we executed SPARQL queries post-reasoning using Apache Jena-Fuseki [72], which supports SPARQL 1.1.

To evaluate the ontology's effectiveness and semantic model, we derived SPARQL queries from the competency questions (CQs) outlined in Section 3. The ontology has pre-populated with test data to facilitate query execution. Successful retrieval of expected results confirmed COViDRO's ability to fulfill its intended objectives. Table 15 provides an overview of the SPARQL queries, their coverage areas, the competency questions addressed, and their practical implications.

Figures 15a, 15b, and 15c illustrate specific SPARQL queries corresponding to CQ I, CQ II, and CQ III. For example, Figure 15a, demonstrates a SPARQL query that retrieves recommended therapeutics based on patient attributes such as age (66 years), SpO2 level (98), underlying condition (cancer), primary symptoms (cough, sore throat, and normal breathing), and ongoing medication (Aspirin). The query identifies Ritonavir-Boosted Nirmatrelvir and Molnupiravir as suitable therapeutic options, ensuring they do not interact negatively with Aspirin. COVIDRO's ontology defines a `COVIDRO_RecommendedTherapeuticsRitonavir-BoostedNirmatrelvir` class, which includes logical constraints as illustrated and discussed in Table 7, section 5. This rule ensures that any patient meeting these conditions is recommended for Ritonavir-Boosted Nirmatrelvir. The Pellet reasoner successfully inferred this rule, allowing SPARQL queries to retrieve suitable therapeutics based on a patient's condition. Additionally, the reasoner can identify multiple individuals who qualify for the same or alternative treatments based on available patient data.

Furthermore, COViDRO enables healthcare professionals to assess the potential adverse effects of specific COVID-19 drugs, such as Ritonavir-Boosted Nirmatrelvir (Figure 15b), by retrieving relevant information, including adverse effects like anaphylaxis, dysgeusia, and nausea. Moreover, healthcare professionals can determine potential drug interactions with Ritonavir-Boosted Nirmatrelvir (Figure 15c) by employing the corresponding SPARQL query. These functionalities offered by COViDRO empower healthcare professionals to make informed decisions on treatment plans.

The foundational role of CODO in developing COViDRO is essential to highlight, as CODO primarily focuses on the epidemiological aspects of COVID-19, including virus transmission, patient records, and test results (see, Section 2). Several key classes and properties from CODO have been reused and extended in COViDRO to ensure semantic interoperability between the two ontologies. CODO provides a structured representation of epidemiological concepts crucial for modeling patient health status and risk factors. Key components incorporated into COViDRO include classes such as `Diagnosis`, `VitalSign`, `RiskFactor`, and `Status`, object properties like `hasDiagnosis`, `hasSymptom`, `hasVitalSign`, `patientStatus`, and `mostRecentVitalSign`, along with data properties such as `age`, `SBP`, `DBP`, `respiratoryRate`, `SpO2`, and `weight` (illustrated using the CODO prefix in Figure 3).

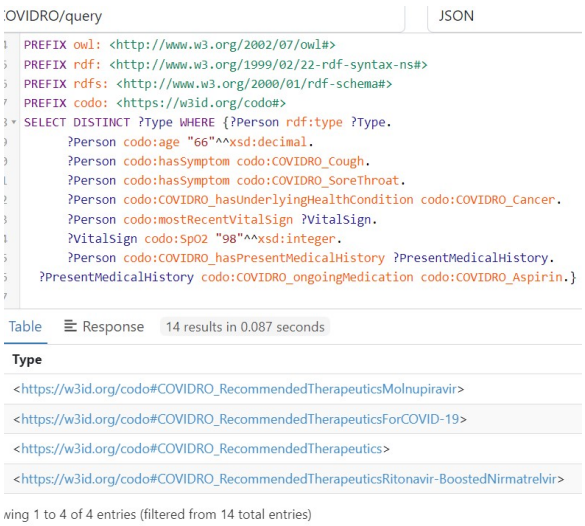
While CODO effectively models epidemiological data, it cannot represent personalized treatment and therapeutic decision-making. To address this gap, COViDRO extends CODO by integrating patient-specific treatment factors described by the PRADiCT framework. COViDRO includes 135 classes, 32 object properties, 15 data properties (as discussed in Section 4), and 45 DL-defined classes that enhance reasoning capabilities through formalized knowledge representation (as discussed in Section 5). The seamless integration of CODO and COViDRO is evident in SPARQL queries, for example, in Figure 15a, which demonstrates how both ontologies work together

Table 15: Overview of SPARQL queries with their respective coverage areas, competency questions, result descriptions, and practical implications

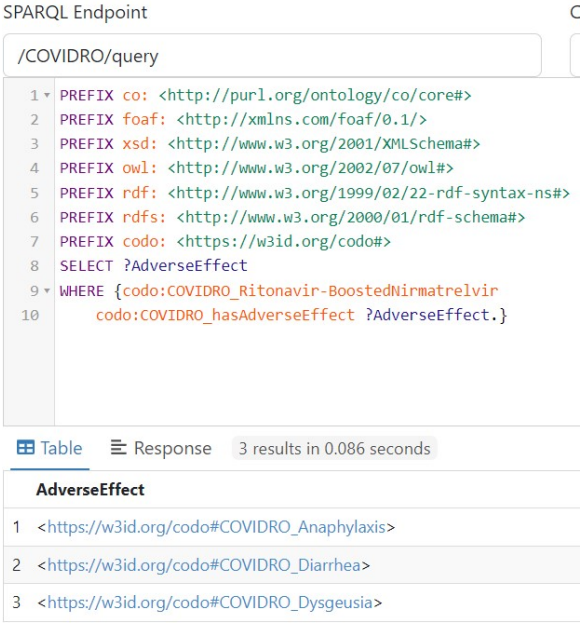
Coverage	Competency Question (CQ)	SPARQL Query	Query Result Description	Practical Implication
Recommended Therapeutics	CQ I. Which therapeutics are recommended for COVID-19 patients with specific underlying health conditions?	`SELECT DISTINCT ?Type WHERE { ?Person rdf:type ?Type. ?Person codo:age "66"^xsd:decimal. ?Person codo:hasSymptom codo:COVIDRO_Cough. ?Person codo:hasSymptom codo:COVIDRO_SoreThroat. ?Person codo:COVIDRO_hasUnderlyingHealthCondition codo:COVIDRO_Cancer. ?Person codo:mostRecentVitalsign ?Vitalsign. ?Vitalsign codo:SpO2 "98"^xsd:integer. ?Person codo:COVIDRO_hasPresentMedicalHistory ?PresentMedicalHistory. ?PresentMedicalHistory codo:COVIDRO_ongoingMedication codo:COVIDRO_Aspirin. }`	Ritonavir-Boosted Nirmatreivir and Mohupiravir are appropriate recommendations for COVID-19 patients with specific characteristics (age 66, symptoms cough/sore throat, cancer condition, SpO2 98, ongoing Aspirin medication). These therapeutics don't interact with Aspirin.	Healthcare professionals can confidently prescribe these therapeutics to COVID-19 patients with specific characteristics, considering age, symptoms, health conditions, SpO2 level, and current medications.
Adverse Effect	CQ II. What are the potential adverse effects of specific COVID-19 drugs?	`SELECT ?AdverseEffect WHERE { codo:COVIDRO_Ritonavir-BoostedNirmatreivir codo:COVIDRO_hasAdverseEffect ?AdverseEffect.}`	Retrieves adverse effects like anaphylaxis, dysgeusia, and nausea associated with Ritonavir-Boosted Nirmatreivir.	Healthcare professionals can be aware of potential adverse effects when prescribing COVID-19 therapeutics.
Drug Interaction	CQ III. Identify drug interactions with particular COVID-19 therapeutics.	`SELECT ?InteractsWith ?NotInteractsWith WHERE { codo:COVIDRO_Ritonavir-BoostedNirmatreivir codo:COVIDRO_mayInteractsWith ?InteractsWith. codo:COVIDRO_Ritonavir-BoostedNirmatreivir codo:COVIDRO_mayNotInteractsWith ?NotInteractsWith.}`	Identifies potential drug interactions with Ritonavir-Boosted Nirmatreivir.	Professionals can avoid adverse effects/reduced efficacy by checking interactions before prescribing COVID-19 treatments.

Prefix Used:

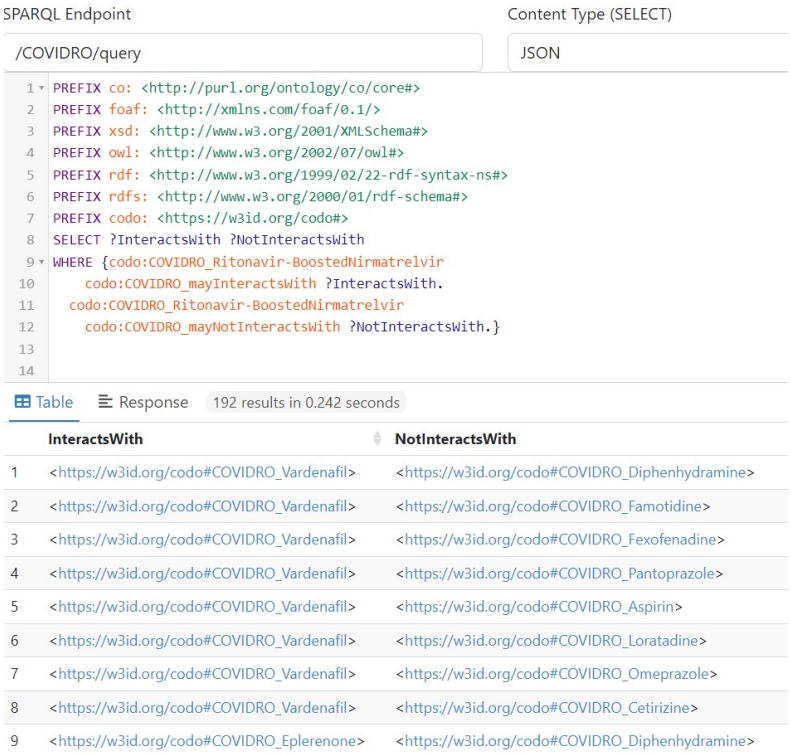
```
PREFIX foaf: <http://xmlns.com/foaf/0.1/>
PREFIX xsd: <http://www.w3.org/2001/XMLSchema# >
PREFIX owl: <http://www.w3.org/2002/07/owl# >
PREFIX rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns# >
PREFIX rdfs: <http://www.w3.org/2000/01/rdf-schema# >
PREFIX codo: <https://w3id.org/codo/>
```



(a) SPARQL query retrieves the therapeutics recommended for COVID-19 patients with specific underlying health conditions.



(b) SPARQL query retrieves the potential adverse effects of specific COVID-19 drugs.



(c) SPARQL query retrieves the drug interactions that may occur with a particular COVID-19 therapeutic.

Fig. 15: SPARQL queries demonstrating the retrieval of therapeutics, adverse effects, and drug interactions related to COVID-19 treatments.

to provide therapeutic recommendations based on patient-specific conditions. In this query, CODO-based concepts such as `codo:age`, `codo:hasSymptoms`, `codo:mostRecentVitalSign`, and `codo:SpO2` are reused, while COViDRO introduces new concepts like `COVIDRO_hasUnderlyingHealthCondition`, `COVIDRO_hasPresentMedicalHistory`, `COVIDRO_ongoingMedication`, and other to support therapeutic decision-making.

By leveraging formal descriptions of DL-defined classes, COViDRO enables inference-driven therapeutic recommendations tailored to individual health conditions. As illustrated in SPARQL query, Figure 15a, COViDRO enables DL-based inferencing to identify eligible patients for specific COVID-19 therapeutics, ensure recommended drugs do not interact negatively with ongoing medications, and provide evidence-based treatment suggestions. The extension can facilitate a broader scope of query-driven decision support, allowing healthcare professionals to make data-informed clinical decisions.

While additional queries are executed to further evaluate COViDRO, they are not included due to space constraints. However, all queries yield the expected results, reinforcing the ontology's semantic validation.

7 Discussion and Limitation

COViDRO highlights its robustness as a structured and semantically rich ontology designed to support healthcare decision-making, particularly in the context of COVID-19 treatment recommendations, drug interactions, and patient risk assessments. The multi-tier evaluation framework applied to COViDRO confirms its strengths in quality assurance, logical consistency, structural complexity, and query performance.

Compared to other related ontologies, COViDRO demonstrates a more comprehensive coverage of clinical and pharmaceutical knowledge, ensuring its practical utility for healthcare professionals and researchers (as discussed in Section 6). The quality-based evaluation using the Ontology Pitfall Scanner (OOPS!) identified significantly fewer pitfalls in COViDRO compared to other related ontologies. While some existing ontologies exhibit critical modeling errors such as undefined inverse relations, missing disjoint axioms, and using recursive definitions, COViDRO successfully mitigates these issues, with only a minor pitfall detected (see, Section 6.1).

Furthermore, benchmarking against 12 other COVID-19 ontologies revealed that COViDRO excels in structural clarity, avoiding redundancies and inconsistencies often found in other models (see, Table 13). Its well-defined property constraints and logical organization contribute to its overall reliability in knowledge representation. Structural and logical validation performed using OntoDebug and the Pellet reasoner further confirmed that COViDRO is free from contradictions (see, Section 6.2), ensuring its suitability for automated reasoning.

Quantitative evaluation using OntoMetric provided additional insights into COViDRO's complexity and richness (see, Section 6.3). The evaluation shows that COViDRO has a well-structured schema with moderate complexity and strong hierarchical depth. COViDRO class richness (0.325926) and inheritance richness (0.940741) values show its well-balanced design aspect (see, Table 14). A key aspect of COViDRO's evaluation was its query-based performance, which tested its ability to retrieve clinically relevant information using SPARQL queries in Apache Jena-Fuseki (see, Section 6.4). The ontology was assessed on its ability to answer predefined competency questions related to therapeutic recommendations, adverse effects, and drug interactions. The ontology also enhances the reasoning capabilities of CODO, its base model, by incorporating additional semantic relationships that improve the accuracy of drug interaction alerts and personalized treatment recommendations.

In comparison with existing ontologies, COViDRO is the only framework that fully integrates all five PRADiCT components (see, Table 1). While some ontologies focus on specific aspects, such as drug interactions (O10: DINTO) or Patient risk factors (O9: CIRO), they do not provide a holistic representation of COVID-19 therapeutics. Furthermore, the ontology adheres to best practices by reusing well-established vocabularies such as SNOMED CT, OBO Foundry ontologies, and CODO (see. S6 of Section 3), thereby ensuring interoperability with other healthcare datasets.

Despite these advancements, COViDRO has certain limitations. One key limitation is the potential for incomplete coverage of all possible drug interactions, risk factors, and treatment options, given the constantly evolving landscape of COVID-19 research. While efforts have been made to ensure accuracy, ongoing updates, and refinements may be required to address emerging knowledge gaps and maintain the ontology's relevance. The effectiveness of COViDRO is also dependent on the quality and availability of data sources utilized for integration. Inaccuracies or gaps in data could impact the ontology's ability to provide precise recommendations and decision support, highlighting the need for continuous data validation and improvement.

8 Conclusion and future directions

The development of COViDRO represents a significant advancement in utilizing ontologies to enhance evidence-based decision-making in the fight against COVID-19. By offering a structured framework for patient categorization and personalized therapeutic recommendations, COViDRO bridges the gap between complex clinical data and actionable insights. Its standardized, machine-readable knowledge base fosters collaboration among healthcare professionals and researchers, while its ability to automate reasoning and tailor recommendations to individual patient needs highlights its value as an innovative tool.

While the current version of COViDRO demonstrates considerable potential, there are ample opportunities for further improvement. Expanding the ontology to include a broader range of COVID-19 therapeutics and their interactions is a critical next step, ensuring its relevance amid evolving treatment protocols and emerging therapies. Integrating real-world data from diverse patient populations will also refine its recommendations, accounting for variables such as comorbidities, demographics, and regional healthcare disparities.

Future efforts should prioritize collaboration with healthcare professionals to validate and refine the ontology, ensuring alignment with current medical standards and practical clinical needs. Interdisciplinary partnerships with experts in epidemiology, pharmacology, and data science will further enhance COViDRO's utility, potentially leading to complementary tools and broader applications in addressing other infectious diseases or public health crises.

To maintain its relevance in the rapidly changing COVID-19 landscape, COViDRO must continuously incorporate the latest research findings, such as insights into long COVID [93], new variants, and updated therapeutic guidelines. Expanding its scope to include broader risk factors—such as mental health conditions, social determinants of health, and environmental influences—will strengthen its ability to address the pandemic's multifaceted impacts [94].

Finally, increasing COViDRO's visibility and accessibility within the healthcare and research communities will be key to its long-term success. This could involve creating user-friendly web interfaces, offering SPARQL-endpoint, and publishing case studies that demonstrate its practical applications.

By addressing these priorities, COViDRO can continue to evolve as a dynamic, impactful resource for both healthcare providers and researchers, ensuring its relevance and effectiveness in combating future outbreaks and preparing for healthcare challenges.

9 Appendix: Source Information

The following sources provide valuable data on COVID-19 treatment guidelines, drug interactions, side effects, and other essential information for the development of COViDRO.

(a) World Health Organization (WHO): https://covid19.who.int/

– **Information Provided:** COVID-19 treatment guidelines, vaccination updates, country-specific data, and health recommendations. It provides authoritative, global guidelines and recommendations on COVID-19 treatment and prevention.

(b) NIH COVID-19 Treatment Guidelines: https://www.covid19treatmentguidelines.nih.gov/

– **Information Provided:** Detailed information on the treatment of COVID-19, including antiviral therapy, immunotherapy, and clinical management strategies.

(c) U.S. Food and Drug Administration (FDA): https://www.fda.gov/drugs/emergency-preparedness-drugs/coronavirus-covid-19-drugs

– **Information Provided:** Information on FDA-approved drugs for COVID-19, emergency use authorizations, and drug safety alerts. This is a key resource for information on authorized treatments and their regulatory status.

(d) Centers for Disease Control and Prevention (CDC): https://www.cdc.gov/coronavirus/2019-ncov/your-health/treatments-for-severe-illness.html

– **Information Provided:** Guidelines for treating severe COVID-19 illness, including recommendations for hospitalized patients.

(e) Infectious Disease Society of America (IDSA): https://www.idsociety.org/practice-guideline/covid-19-guideline-treatment-and-management/

– **Information Provided:** This resource provides guidelines on evidence-based recommendations for COVID-19 treatment. This includes guidance on pharmacologic treatments, infection prevention, and diagnostics.

(f) COVID-19 Drug Interactions: https://www.covid19-druginteractions.org/

– **Information Provided:** Drug interaction database for COVID-19 treatments, including information on the safe use of antiviral therapies. This resource provides an extensive database of drug interactions related to COVID-19 therapeutics, critical for ensuring patient safety.

(g) ClinicalTrials.gov: https://www.clinicaltrials.gov/

– **Information Provided:** Clinical trial data, including ongoing studies on COVID-19 treatments, vaccines, and drug interactions. This resource provides access to clinical trial results, helping to stay updated on experimental and emerging treatments.

(h) International Clinical Trials Registry Platform (ICTRP): https://www.who.int/clinical-trials-registry-platform

- **Information Provided:** Global registry of clinical trials, including COVID-19-related studies.

(i) NCBO BioPortal: https://bioportal.bioontology.org/

- **Information Provided:** Repository of biomedical ontologies, including COVID-19-related ontologies.

(j) EMBL-EBI Ontology Lookup Service (OLS): https://www.ebi.ac.uk/ols4

- **Information Provided:** Comprehensive ontology lookup service, including biomedical ontologies.

(k) COVID-19 Data Portal: https://www.covid19dataportal.org/

- **Information Provided:** Large-scale observational studies, patient data, and COVID-19 research datasets. This resource provides real-world data on COVID-19 outcomes, treatment effectiveness, and patient characteristics.

(l) IndiaFightsCorona - COVID-19 in India: https://www.mygov.in/covid-19

- **Information Provided:** Data and updates on COVID-19 in India, including vaccination, treatment, and government policies. This resource provides country-specific data and information on India's response to the pandemic.

References

1. Filip R, Gheorghita Puscaselu R, Anchidin-Norocel L, et al (2022) Global Challenges to Public Health Care Systems during the COVID-19 Pandemic: A Review of Pandemic Measures and Problems. J Pers Med 12:1295. https://doi.org/10.3390/jpm12081295
2. Cascella M, Rajnik M, Aleem A, et al (2022) Features, Evaluation, and Treatment of Coronavirus (COVID-19). https://www.ncbi.nlm.nih.gov/books/NBK554776/
3. Kumar D, Trivedi N (2021) Disease-drug and drug-drug interaction in COVID-19: Risk and assessment. Biomed Pharmacother 139:111642. https://doi.org/10.1016/j.biopha.2021.111642
4. Alshammari TM (2016) Drug safety: The concept, inception and its importance in patients' health. Saudi Pharm J 24:405–412. https://doi.org/10.1016/j.jsps.2014.04.008
5. People with Certain Medical Conditions. (2023). Centers for Disease Control and Prevention https://www.cdc.gov/coronavirus/2019-ncov/need-extra-precautions/people-with-medical-conditions.html
6. Sedlakova J, Daniore P, Horn Wintsch A, et al (2023) Challenges and best practices for digital unstructured data enrichment in health research: A systematic narrative review. PLOS Digit Health 2:e0000347. https://doi.org/10.1371/journal.pdig.0000347.
7. Remdesivir. COVID-19 Treatment Guidelines https://www.covid19treatmentguidelines.nih.gov/therapies/antivirals-including-antibody-products/remdesivir/.
8. Gandhi Z, Mansuri Z, Bansod S (2020) Potential Interactions of Remdesivir with Pulmonary Drugs: a Covid-19 Perspective. SN Compr Clin Med 2:1707–1708. https://doi.org/0.1007/s42399-020-00462-2
9. Research, C. for D. E. and. Coronavirus (COVID-19) | Drugs. FDA (2023).https://www.fda.gov/drugs/emergency-preparedness-drugs/coronavirus-covid-19-drugs
10. Borst W (1997) Construction of Engineering Ontologies. PhD thesis, Institute for Telematica and Information Technology, University of Twente.
11. Studer R, Benjamins VR, Fensel D (1998) Knowledge engineering: Principles and methods. Data Knowl Eng 25:161–197. https://doi.org/10.1016/S0169-023X(97)00056-6.
12. Hogan A, Blomqvist E, Cochez M, et al (2022) Knowledge Graphs. ACM Comput Surv 54:1–37.https://doi.org/10.1145/3447772
13. Pavlopoulos GA, Secrier M, Moschopoulos CN, et al (2011) Using graph theory to analyze biological networks. BioData Mining 4:10. https://doi.org/10.1186/1756-0381-4-10
14. Huajun Chen, Zhaohui Wu, Heng Wang, Yuxin Mao (2006) RDF/RDFS-based Relational Database Integration. In: 22nd International Conference on Data Engineering (ICDE'06). IEEE, Atlanta, GA, USA, pp 94–94 https://doi.org/10.1109/ICDE.2006.127
15. Cox SJD, Gonzalez-Beltran AN, Magagna B, Marinescu M-C (2021) Ten simple rules for making a vocabulary FAIR. PLOS Computational Biology 17:e1009041. https://doi.org/10.1371/journal.pcbi.1009041
16. Garijo D, Poveda-Villalón M (2020) Best Practices for Implementing FAIR Vocabularies and Ontologies on the Web. In: Cota G, Daquino M, Pozzato GL (eds) Studies on the Semantic Web. IOS Press https://doi.org/10.3233/SSW200034
17. Haendel MA, Chute CG, Robinson PN (2018) Classification, Ontology, and Precision Medicine. N Engl J Med 379:1452–1462. https://doi.org/10.1056/NEJMra1615014

18. Fonou-Dombeu JV, Achary T, Genders E, et al (2021) COVIDonto: An Ontology Model for Acquisition and Sharing of COVID-19 Data. In: Attiogbé C, Ben Yahia S (eds) Model and Data Engineering. Springer International Publishing, Cham, pp 227–240
19. Baader F, Horrocks I, Sattler U (2005) Description Logics as Ontology Languages for the Semantic Web. In: Hutter D, Stephan W (eds) Mechanizing Mathematical Reasoning. Springer Berlin Heidelberg, Berlin, Heidelberg, pp 228–248
20. Calvanese D, Lenzerini M, Nardi D (1998) Description Logics for Conceptual Data Modeling. In: Chomicki J, Saake G (eds) Logics for Databases and Information Systems. Springer US, Boston, MA, pp 229–263 https://doi.org/10.1007/978 − 1 − 4615 − 5643 − 5_8
21. Miotto R, Wang F, Wang S, et al (2017) Deep learning for healthcare: review, opportunities and challenges. Brief Bioinform 19:1236–1246. https://doi.org/10.1093/bib/bbx044
22. Jean, S.-S., Lee, P.-I. & Hsueh, P.-R. TJean S-S, Lee P-I, Hsueh P-R (2020) Treatment options for COVID-19: The reality and challenges. J Microbiol Immunol Infect 53:436–443. https://doi.org/10.1016/j.jmii.2020.03.034
23. Ontologies: Science Data Sharing | Learn Science at Scitable. https://www.nature.com/scitable/topicpage/ontologies-scientific-data-sharing-made-easy-77972/.
24. Description Logics - an overview | ScienceDirect Topics. https://www.sciencedirect.com/topics/computer-science/description-logics.
25. Horrocks I (2005) OWL: A Description Logic Based Ontology Language. In: Van Beek P (ed) Principles and Practice of Constraint Programming - CP 2005. Springer Berlin Heidelberg, Berlin, Heidelberg, pp 5–8
26. de Lusignan S, Liyanage H, McGagh D, et al (2020) COVID-19 Surveillance in a Primary Care Sentinel Network: In-Pandemic Development of an Application Ontology. JMIR Public Health Surveill 6:e21434. https://doi.org/10.2196/21434
27. He Y, Yu H, Ong E, et al (2020) CIDO, a community-based ontology for coronavirus disease knowledge and data integration, sharing, and analysis. Sci Data 7:181. https://doi.org/10.1038/s41597-020-0523-6
28. WHO COVID-19 Rapid Version CRF semantic data model - Summary | NCBO BioPortal. https://bioportal.bioontology.org/ontologies/COVIDCRFRAPID.
29. Badenes-Olmedo C, Chaves-Fraga D, Poveda-VillalÓn M, et al (2020) Drugs4Covid: Drug-driven Knowledge Exploitation based on Scientific Publications http://arxiv.org/abs/2012.01953(2020)
30. Al Qundus J (2021) ROC: An Ontology for Country Responses towards COVID-19 https://api.semanticscholar.org/CorpusID:233240680
31. Sargsyan A, Kodamullil AT, Baksi S, et al (2021) The COVID-19 Ontology. Bioinformatics 36:5703–5705. https://doi.org/10.1093/bioinformatics/btaa1057
32. Dutta B, DeBellis M (2020) CODO: An Ontology for Collection and Analysis of Covid-19 Data: Proceedings of the 12th International Joint Conference on Knowledge Discovery, Knowledge Engineering and Knowledge Management 76–85. https://doi.org/10.5220/0010112500760085
33. The Drug Ontology < Ontology Lookup Service < EMBL-EBI. https://www.ebi.ac.uk/ols/ontologies/dron.
34. CIRO: COVID-19 infection risk ontology | PLOS ONE. https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0282291.
35. DINTO: Using OWL Ontologies and SWRL Rules to Infer Drug-Drug Interactions and Their Mechanisms - PubMed. https://pubmed.ncbi.nlm.nih.gov/26147071/.
36. Yu H, Nysak S, Garg N, et al (2019) ODAE: Ontology-based systematic representation and analysis of drug adverse events and its usage in study of adverse events given different patient age and disease conditions. BMC Bioinformatics 20:199 https://doi.org/10.1186/s12859-019-2729-1
37. Sakor A, Jozashoori S, Niazmand E, et al (2023) Knowledge4COVID-19: A semantic-based approach for constructing a COVID-19 related knowledge graph from various sources and analyzing treatments' toxicities. Journal of Web Semantics 75:100760. https://doi.org/https://doi.org/10.1016/j.websem.2022.100760
38. Schema.org - Schema.org. https://schema.org/. Accessed 18 Jul 2023
39. About: FOAF. https://dbpedia.org/page/FOAF. Accessed 18 Jul 2023
40. The Organization Ontology. https://www.w3.org/TR/vocab-org/. Accessed 6 Feb 2025.
41. SNOMED CT | NCBO BioPortal. https://bioportal.bioontology.org/ontologies/SNOMEDCT. Accessed 6 Feb 2025.
42. OBO Foundry. https://obofoundry.org/principles/fp-000-summary.html. Accessed 19 Jul 2023.
43. Bain D, Dutta B (2024) Systematic Analysis of COVID-19 Ontologies. In: Garoufallou E, Sartori F (eds) Metadata and Semantic Research. Springer Nature Switzerland, Cham, pp 74–91. https://doi.org/10.1007/978 − 3 − 031 − 65990 − 4_7
44. Project Ping | NCBO BioPortal. https://bioportal.bioontology.org/projects/Ping. Accessed 19 Jul 2023.
45. Keloth VK, Zhou S, Lindemann L, et al (2020) Mining Concepts for a COVID Interface Terminology for Annotation of EHRs. In: 2020 IEEE International Conference on Big Data (Big Data). pp 3753–3760
46. Qundus JA, Schäfermeier R, Karam N, et al (2021) ROC: An Ontology for Country Responses towards COVID-19
47. Wu H, Zhong Y, Tian Y, et al (2021) Automatic diagnosis of COVID-19 infection based on ontology reasoning. BMC Med Inform Decis Mak 21:271. https://doi.org/10.1186/s12911-021-01629-0

48. González-Eras A, Santos RD, Aguilar J, Lopez A (2022) Ontological engineering for the definition of a COVID-19 pandemic ontology. Inform Med Unlocked 28:100816. https://doi.org/10.1016/j.imu.2021.100816
49. IDSA Guidelines on the Treatment and Management of Patients with COVID-19. https://www.idsociety.org/practice-guideline/covid-19-guideline-treatment-and-management/. Accessed 19 Jul 2023
50. Wang X, Liang T, Li J, et al (2021) Artificial Intelligence-Empowered Chatbot for Effective COVID-19 Information Delivery to Older Adults. IJEHMC 12:1–18. https://doi.org/10.4018/IJEHMC.293285
51. Toumia S, Mejri A (2021) An Ontology-based Approach For Conformance Checking of Decision Mining Rules
52. Silega N, Varén E, Varén A, et al (2022) Exploiting an Ontological Model to Study COVID-19 Contagion Chains in Sustainable Smart Cities. Information 13:40. https://doi.org/10.3390/info13010040
53. Giunchiglia F, Bocca S, Fumagalli M, et al (2021) iTelos – Purpose Driven Knowledge Graph Generation
54. Lin AY, Yamagata Y, Duncan WD, et al A Community Effort for COVID-19 Ontology Harmonization. 6
55. Fernández-López M, Gomez-Perez A, Juristo N (1997) METHONTOLOGY: from ontological art towards ontological engineering. Engineering Workshop on Ontological Engineering (AAAI97) https://api.semanticscholar.org/CorpusID:10550105
56. Vrandečić D, Pinto HS, Dr. Tempich C, Sure-Vetter Y (2005) The DILIGENT knowledge processes. J Knowledge Management 9:85–96. https://doi.org/10.1108/13673270510622474
57. Suárez-Figueroa MC, Gomez-Perez A, Fernández-López M (2012) The NeOn Methodology for Ontology Engineering. In: Ontology Engineering in A Networked World. pp 9–34 https://doi.org/10.1007/978 − 3 − 642 − 24794 − 1_2.
58. De Nicola A, Missikoff M, Navigli R (2005) A Proposal for a Unified Process for Ontology Building: UPON. pp 655–664 https://doi.org/10.1007/11546924_64.
59. Dutta B, Chatterjee U, Madalli D (2015) YAMO: Yet Another Methodology for large-scale faceted Ontology construction. Journal of Knowledge Management 19:6–24. https://doi.org/10.1108/JKM-10-2014-0439
60. experts yWorks the diagramming yEd Graph Editor. In: YWorks Diagramming Experts. https://www.yworks.com/products/yed. Accessed 16 January 2023.
61. Therapeutics and COVID-19: Living guideline, 13 January 2023.https://www.who.int/publications-detail-redirect/WHO-2019-nCoV-therapeutics-2023.1.
62. Commissioner, O. of the. Know Your Treatment Options for COVID-19. FDA (2023). https://www.fda.gov/consumers/consumer-updates/know-your-treatment-options-covid-19
63. Information on COVID-19 Treatment, Prevention and Research. COVID-19 Treatment Guidelines https://www.covid19treatmentguidelines.nih.gov/.
64. COVID-19 and Your Health. Centers for Disease Control and Prevention. (2020). https://www.cdc.gov/coronavirus/2019-ncov/your-health/treatments-for-severe-illness.html
65. Home | Beta ClinicalTrials.gov. https://www.clinicaltrials.gov/.
66. International Clinical Trials Registry Platform (ICTRP). https://www.who.int/clinical-trials-registry-platform.
67. Ellson J, Gansner ER, Koutsofios E, et al (2004) Graphviz and Dynagraph — Static and Dynamic Graph Drawing Tools. In: Jünger M, Mutzel P (eds) Graph Drawing Software. Springer Berlin Heidelberg, Berlin, Heidelberg, pp 127–148
68. The protégé project: a look back and a look forward: AI Matters: Vol 1, No 4. https://dl.acm.org/doi/10.1145/2757001.2757003.
69. OntoDebug - Interactive Ontology Debugging in Protégé. http://isbi.aau.at/ontodebug/.
70. Sirin E, Parsia B, Cuenca Grau B, et al (2007) Pellet: A Practical OWL-DL Reasoner https://doi.org/10.2139/ssrn.3199351 (2007).
71. SPARQL 1.1 Query Language. https://www.w3.org/TR/sparql11-query/.
72. Apache Jena - Apache Jena Fuseki. https://jena.apache.org/documentation/fuseki2/. Accessed 13 Jun 2023
73. WIDOCO Tutorial. https://dgarijo.github.io/Widoco/doc/tutorial/.
74. OntoMetrics. https://ontometrics.informatik.uni-rostock.de/ontologymetrics/. Accessed 12 Jun 2024
75. OWL - Semantic Web Standards. https://www.w3.org/OWL/.
76. Age. National Institutes of Health (NIH). (2022).https://www.nih.gov/nih-style-guide/age
77. COVID-19 symptoms and severity. https://www.who.int/westernpacific/emergencies/covid-19/information/asymptomatic-covid-19.
78. Defining Adult Overweight and Obesity. Centers for Disease Control and Prevention.(2022). https://www.cdc.gov/obesity/basics/adult-defining.html
79. Poveda-Villalón M, Gómez-Pérez A, Suárez-Figueroa MC (2014) OOPS! (OntOlogy Pitfall Scanner!): An Online Tool for Ontology Evaluation. International Journal on Semantic Web and Information Systems 10:7–34. https://doi.org/https://doi.org/10.4018/ijswis.2014040102
80. Welcome to the NCBO BioPortal | NCBO BioPortal. https://bioportal.bioontology.org/.
81. Ontology Lookup Service (OLS). https://www.ebi.ac.uk/ols4.
82. Home | Department of Economic Affairs | Ministry of Finance | Government of India. https://dea.gov.in/.
83. World Health Organization (WHO). https://www.who.int.
84. IndiaFightsCorona COVID-19 in India, Vaccination, Dashboard, Corona Virus Tracker | mygov.in. https://www.mygov.in/covid-19.

85. COVID-19 Data Portal - accelerating scientific research through data. https://www.covid19dataportal.org/.
86. A. Setiawan F, Murdani D, Riana F, Dwimawati E (2021) COPOMBOCY: A COVID-19 Pandemic Ontology Model of Bogor City. In: Proceedings of the 2021 International Conference on Computer, Control, Informatics and Its Applications. ACM, Virtual/online conference Indonesia, pp 86–90
87. Sharma S, Jain S (2023) CovidO: an ontology for COVID-19 metadata. J Supercomput. https://doi.org/10.1007/s11227-023-05509-4
88. Abhilash CB, Mahesh K (2022) Ontology-Based Interestingness in COVID-19 Data. In: Garoufallou E, Ovalle-Perandones M-A, Vlachidis A (eds) Metadata and Semantic Research. Springer International Publishing, Cham, pp 322–335
89. Patel A, Debnath NC (2022) Development of the $InBan_CIDO$ Ontology by Reusing the Concepts Along with Detecting Overlapping Information. In: Inventive Computation and Information Technologies. Springer, pp 349–359
90. Calvanese D, Lenzerini M, Nardi D (1998) Description Logics for Conceptual Data Modeling. In: Chomicki J, Saake G (eds). Springer US, Boston, MA, pp 229–263
91. Calvanese D, De Giacomo G, Nardi D, Lenzerini M (2001) Reasoning in expressive description logics. In: Handbook of automated reasoning. Elsevier Science Publishers B. V., NLD, pp 1581–1634
92. Noy NF, McGuinness DL (2001) Ontology Development 101: A Guide to Creating Your First Ontology
93. Aiyegbusi OL, Hughes SE, Turner G, et al (2021) Symptoms, complications and management of long COVID: a review. J R Soc Med 114:428–442. https://doi.org/doi.org/10.1177/01410768211032850
94. Kirkbride JB, Anglin DM, Colman I, et al (2024) The social determinants of mental health and disorder: evidence, prevention and recommendations. World Psychiatry 23:58–90. https://doi.org/doi.org/10.1002/wps.21160