An Ontology to Bridge the Clinical Management of Patients and Public Health Responses for Strengthening Infectious Disease Surveillance: A Design Science Study

Sachiko Lim, Paul Johannesson

Submitted to: JMIR Formative Research
on: October 16, 2023

Table of Contents

An Ontology to Bridge the Clinical Management of Patients and Public Health Responses for Strengthening Infectious Disease Surveillance: A Design Science Study

Sachiko Lim[1] MSc; Paul Johannesson[1]

[1]Stockholm University Kista SE

Corresponding Author:
Sachiko Lim MSc
Stockholm University
Nodhuset, Borgarfjordsgatan 12
Kista
SE

Abstract

Background: Novel surveillance approaches using digital technologies, including the Internet of Things (IoT), have evolved, enhancing traditional infectious disease surveillance systems by enabling real-time detection of outbreaks and reaching a wider population. However, disparate, heterogenous infectious disease surveillance systems often operate in silos due to a lack of interoperability. As a life-changing clinical use case, the COVID-19 pandemic has manifested that a lack of interoperability can severely inhibit public health responses to emerging infectious diseases. Interoperability is thus critical for building a robust ecosystem of infectious disease surveillance and enhancing preparedness for future outbreaks. The primary enabler for semantic interoperability is ontology.

Objective: This paper aims to design the IoT-based management of infectious disease ontology (IoT-MIDO) to enhance data sharing and integration of data collected from IoT-driven patient health monitoring, clinical management of individual patients, and disparate heterogeneous infectious disease surveillance.

Methods: The (IoT-MIDO) was developed using the Basic Formal Ontology (BFO) as the top-level ontology. We reused the classes from existing BFO-based ontologies as much as possible to maximize the interoperability with other BFO-based ontologies and databases that rely on them. We formulated the competency questions as requirements for the ontology to achieve the intended goals.

Results: We demonstrate five use cases using the simplified ontological models to show the potential applications of IoT-MIDO: 1) IoT-driven patient monitoring, risk assessment, early warning, and risk management; 2) clinical patient management of infectious diseases; 3) epidemic risk analysis for timely response at the public health level; 4) infectious disease surveillance; and 5) transforming patient information into surveillance information.

Conclusions: The development of the IoT-MIDO was driven by competency questions. Being able to answer all of the formulated competency questions, we successfully demonstrated that our ontology has the potential to facilitate data sharing and integration for orchestrating IoT-driven patient health monitoring in the context of an infectious disease epidemic, clinical patient management, infectious disease surveillance, and epidemic risk analysis. The novelty and uniqueness of the ontological model lie in building a bridge to link IoT-based individual patient monitoring and early warning based on patient risk assessment to infectious disease epidemic surveillance at the public health level. The ontology can also serve as a starting point to enable potential decision support systems to support public health organisations and practitioners.

(JMIR Preprints 16/10/2023:53711)
DOI: https://doi.org/10.2196/preprints.53711

Only make the preprint title and abstract visible.
No, I do not wish to publish my submitted manuscript as a preprint.

2) If accepted for publication in a JMIR journal, would you like the PDF to be visible to the public?

✓ **Yes, please make my accepted manuscript PDF available to anyone at any time (Recommended).**
Yes, but please make my accepted manuscript PDF available only to logged-in users; I understand that the title and abstract will remain v
Yes, but only make the title and abstract visible (see Important note, above). I understand that if I later pay to participate in <a href="http

Original Manuscript

Original Paper

An Ontology to Bridge the Clinical Management of Patients and Public Health Responses for Strengthening Infectious Disease Surveillance: A Design Science Study

Sachiko Lim[1], Paul Johannesson[1]
[1] Department of Computer and Systems Sciences, Stockholm University, Kista, Stockholm, Sweden

Corresponding Author:
Sachiko Lim
Department of Computer and Systems Sciences,
Stockholm University,
Nodhuset, Borgarfjordsgatan 12, 164 07 Kista, Sweden
Email:sachiko@dsv.su.se
Phone: +46 (0)760968462

Abstract

Background:
Novel surveillance approaches using digital technologies, including the Internet of Things (IoT), have evolved, enhancing traditional infectious disease surveillance systems by enabling real-time detection of outbreaks and reaching a wider population. However, disparate, heterogenous infectious disease surveillance systems often operate in silos due to a lack of interoperability. As a life-changing clinical use case, the COVID-19 pandemic has manifested that a lack of interoperability can severely inhibit public health responses to emerging infectious diseases. Interoperability is thus critical for building a robust ecosystem of infectious disease surveillance and enhancing preparedness for future outbreaks. The primary enabler for semantic interoperability is ontology.

Objective: This paper aims to design the IoT-based management of infectious disease ontology (IoT-MIDO) to enhance data sharing and integration of data collected from IoT-driven patient health monitoring, clinical management of individual patients, and disparate heterogeneous infectious disease surveillance.

Methods: The ontology modeling approach was chosen for its semantic richness in knowledge representation, flexibility, ease of extensibility, and capability for knowledge inference and reasoning. The IoT-MIDO was developed using the Basic Formal Ontology (BFO) as the top-level ontology. We reused the classes from existing BFO-based ontologies as much as possible to maximize the interoperability with other BFO-based ontologies and databases that rely on them. We formulated the competency questions as requirements for the ontology to achieve the intended goals.

Results: We designed an ontology to integrate data from heterogeneous sources, including IoT-driven patient monitoring, clinical management of individual patients, and infectious disease surveillance systems. This integration aims to facilitate the collaboration between clinical care and public health domains. We also demonstrate five use cases using the simplified ontological models to show the potential applications of IoT-MIDO: 1) IoT-driven patient monitoring, risk assessment, early warning, and risk management; 2) clinical patient management of infectious diseases; 3) epidemic risk analysis for timely response at the public health level; 4) infectious disease surveillance; and 5) transforming patient information into surveillance information.

Conclusions:
The development of the IoT-MIDO was driven by competency questions. Being able to answer all the formulated competency questions, we successfully demonstrated that our ontology has the potential to facilitate data sharing and integration for orchestrating IoT-driven patient health monitoring in the context of an infectious disease epidemic, clinical patient management, infectious disease surveillance, and epidemic risk analysis. The novelty and uniqueness of the ontology lie in building a bridge to link IoT-based individual patient monitoring and early warning based on patient risk assessment to infectious disease epidemic surveillance at the public health level. The ontology can also serve as a starting point to enable potential decision support systems, providing actionable insights to support public health organizations and practitioners in making informed decisions in a timely manner.

Keywords: infectious disease; ontology; IoT; infectious disease surveillance; patient monitoring; epidemic management; risk analysis; early warning; data integration; semantic interoperability

Introduction

Overall global health has improved over the past 30 years with the steady decline of age-standardized disability-adjusted life-year (DALY) rates [1]. Although the burden of infectious diseases remains high among children younger than ten years old, the global disease trend has seen a shift from communicable to non-communicable diseases [1]. Nevertheless, (re)emerging infectious diseases remain global health threats due to climate change, increased wildlife-livestock-human interface associated with urbanization, globalization of transport and human movement [2]. As we all have witnessed during the COVID-19 pandemic, the rapid and transnational spread of (re)emerging infectious diseases can have catastrophic consequences, including a significant loss of human lives, social disruption, and economic instability.

The COVID-19 pandemic has manifested a global vulnerability to newly emerging infectious diseases. Insufficient epidemic preparedness and delayed responses have exacerbated the spread of SARS-CoV-2 and triggered a significant increase in incidence and mortality. In Italy, for example, it has been claimed that the delayed implementation of the lockdown accounted for a substantial proportion of hospital admissions and deaths [3]. According to a study, if lockdown had been implemented one week earlier, Italy could have averted 60% of cases, 48% of ICU admissions, and 44% of deaths at the time of the study [3]. Similarly, another study indicates that if control measures had been implemented just 1–2 weeks earlier, the United States could have avoided 56.5% of reported cases and 54.0% of reported deaths at the time of the study [4].

Moreover, many countries implemented response measures when the local health systems were already becoming overstretched [5]. The prolonged period of health system disruption potentially caused increases in the incidence and mortality of other diseases due to core health service disruptions [6], [7]. Thus, early detection is crucial for preventing and responding to emerging infectious disease outbreaks [8]. It demands robust public health surveillance systems to inform effective outbreak management [8]. The main goals of the surveillance are: 1) understanding the disease burden and epidemiology, 2) monitoring disease trends, 3) identifying and early warning of public health threats (e.g., epidemics of emerging infectious diseases), 4) assessing risks and prioritizing diseases, 5) disseminating surveillance data to stakeholders and 6) planning, implementing, monitoring and evaluating public health response measures for disease control, elimination, and eradication.

Traditional infectious disease surveillance is often divided into active and passive surveillance. In active surveillance systems, health department staff proactively contact physicians, laboratories, healthcare providers, or the general population to collect information about diseases [9]. In passive surveillance systems, which are the most common type of surveillance, medical professionals report cases and deaths to the public health agency according to a list of reportable diseases [9]. While active surveillance is likely to provide complete and more accurate data than passive surveillance, the method is more expensive and labor-intensive. On the other hand, passive surveillance is incomplete and subject to underreporting and delays between event occurrences and notifications [9]. Another major limitation of traditional surveillance systems is that they cannot detect an outbreak in real time.

With advancements in information technologies and the digital revolution, novel surveillance approaches driven by IoT have evolved, enhancing traditional infectious disease surveillance systems by enabling real-time detection of outbreaks and reaching a wider population [10]. The IoT creates an ecosystem that connects people and objects through the internet, allowing them to collect and transmit data over a highly distributed network via embedded sensors. The IoT has opened new opportunities to improve public health through enhancing disease surveillance and assisting healthcare in transitioning to a proactive P4 (predictive, preventive, personalized, and participatory) medicine [10], [11]. Innovative IoT-based methods, such as participatory surveillance and digital surveillance, have been used in disease surveillance.

Participatory surveillance that leverages digital connectivity (e.g., mobile apps) requires the direct involvement of system users who voluntarily provide the information needed for informing public health actions. The data collected from each user are aggregated and analyzed for public health purposes. Although there are event-based participatory surveillance systems, many have been used to perform syndromic surveillance. The approach aims to monitor disease indicators in (near) real-time for earlier detection of and response to outbreaks to reduce morbidity and mortality [12], [13].

The primary advantages of participatory surveillance systems are four-fold [14]: 1) enable large-scale and population-based monitoring at a low cost, 2) enable engagement with populations that are hard to reach by traditional surveillance systems due to geographical constraints or social and economic situations, 3) allow for the rapid two-way communication between health authorities and system users for public health messaging and education to promote disease prevention and control activities and 4) provide flexible data systems and user interfaces, which allows health authorities to modify data elements to be collected (e.g., adding new symptoms of an emerging infectious disease) and disseminates information in near real-time. The app-based technology reduced delays in contact tracing and demonstrated the potential for preventing up to 80% of all transmissions [15]. On the other hand, a significant challenge is recruiting and retaining a representative sample of an at-risk population [14]. The approach also lacks the specificity of a laboratory test to confirm a pathogen, while it can achieve high sensitivity if the surveillance coverage is sufficiently high [14].

Digital public health surveillance uses publicly available user-contributed data collected outside conventional public health surveillance channels. Thus, such data are not generated primarily for infectious disease surveillance [16], [17].

As mentioned above, each surveillance system has its strengths and limitations, and they complement each other. A hybrid system integrating traditional and novel surveillance approaches is likely to be the most promising option in the Big Data era [18]. Informing and coordinating effective and timely outbreak management requires seamless information flows between disparate heterogeneous surveillance systems. However, they often operate in silos due to a lack of interoperability. Interoperability is defined as "the ability of disparate computer systems or software to exchange data in an efficient and meaningful way [19]". As a life-changing clinical use case, the COVID-19 pandemic has manifested that a lack of interoperability can severely inhibit public health responses to emerging infectious diseases [19]. Interoperability is thus critical for building a robust ecosystem of infectious disease surveillance and enhancing preparedness for future outbreaks.

According to the Healthcare Information and Management Systems Society, interoperability consists of four levels: functional (level 1), structural (level 2), semantic (level 3) and organizational (level 4) [20]. This study focuses on the semantic level of interoperability (aka semantic interoperability) to aim for seamless data sharing and integration in an IoT-enhanced surveillance ecosystem that includes various heterogeneous data sources. To achieve semantic interoperability, both data and their unambiguous and shared meaning need to be conveyed to the receiving systems such that they interpret and process the data correctly [21].

In healthcare settings, reference terminologies or terminology standards have played a significant role in facilitating data standardization and providing semantic interoperability. International Classification of Diseases and Related Health Problems (ICD) is the global terminology standard designed to promote international comparability in classifying diseases, injuries, and causes of death. ICD is also used to standardize the reporting and monitoring of health conditions [22]. The systematized nomenclature of medicine clinical terms (SNOMED CT) is an international clinical reference terminology for facilitating the electronic exchange of clinical health information consistently. SNOMED CT provides the ability to create compositional concepts that combine multiple concepts to form a more detailed representation of a clinical problem statement [23]. It can also be mapped to external coding systems such as ICD-10 to promote semantic interoperability. In nursing, NANDA International (NANDA-I), the Nursing Interventions Classification (NIC), and the Nursing Outcomes Classification (NOC) are three major terminologies that have been used to describe nursing judgments, treatments, and nursing-sensitive patient outcomes [21]. The International Classification for Nursing Practice (ICNP) has also been developed to represent the dynamic nature of nursing practices and their cultural variations [24].

Moreover, various technical interoperability standards have been developed. For example, the Health Level 7 Fast Healthcare Interoperability Resources (HL7/FHIR) is a medical information standard created by HL7 to enable RESTful data exchange [25]. The Observational Medical Outcomes Partnership (OMOP) Common Data Model (CDM) has been designed to standardize the structure and content of observational data [26]. Clinical Data Interchange Standards Consortium (CDISC) Foundational Standards have been developed to support non-clinical and the lifecycle of the clinical research process from planning, data collection, exchange, management, and analysis to reporting of the findings derived from clinical trials [27]. Although various robust reference terminologies and standards are available in clinical medicine, further coordination across standards is necessary to avoid creating standard-specific silos [28].

Another enabler for semantic interoperability is ontology, which refers to "a formal, explicit specification of a shared conceptualization [29]”. Ontologies are machine-interpretable and provide the ability to reconcile the meaning of data held across heterogeneous data sources. Data senders and receivers share a common understanding of the meanings of the data exchanged [30], [31]. The main strengths of using an ontology are that it provides flexible and technology-agnostic methods for data sharing and integration, expresses relationships between concepts, and enables reasoning [30]. Ontologies are similar to reference terminologies in that they both systematically represent a domain of interest. However, ontologies are more expressive than terminologies, providing richer semantic relationships by representing concepts, their relationships, and axioms [32]. They, thus, serve as a basis for knowledge graphs and also support semantic reasonings.

Enormous efforts have been devoted to developing ontologies to enable data sharing, integration, and analysis for infectious disease surveillance and response. Infectious Disease Ontology (IDO) Core, which was released in 2010, is based on Basic Formal Ontology (BFO) and covers entities and relations relevant to infectious diseases in general. It also includes terms for population-level processes (e.g., infection incidence, epidemic, pandemic) [33]. IDO Core serves as a hub from which extensions based on pathogen type (i.e., virus, bacteria, fungi, and parasite) are developed, IDO Virus (VIDO), IDO Bacteria, IDO Fungus, and IDO Parasite [33]. Each of those extensions is further partitioned into pathogen-specific ontologies such as IDO-Dengue Fever (IDODEN) [34], IDO-HIV(IDOHIV), and IDO-influenza (IDOFLU). To facilitate sharing, integrating, and analyzing COVID-19 data, three new IDO extensions have recently been developed [33], namely Virus Infectious Disease Ontology (VIDO)[35], the Coronavirus Infectious Disease Ontology (CIDO) [36] and IDO-COVID-19, which is an extension of CIDO [33]. The Apollo Structured Vocabulary (Apollo-SV), which is also based on BFO, provides a standardized representation for configurations and output of epidemic simulators [37]. It aims to aid in locating and accessing a simulator, understanding its characteristics, performing analyses, and analyzing outputs to inform policy or decisions about disease control.

To our knowledge, however, no ontology exists that supports IoT-enhanced infectious disease surveillance, risk analysis, and early warning of infectious diseases at individual and public health levels. The aim of this paper is to design an infectious disease ontology that can support data sharing and integration of data collected from IoT-driven patient health monitoring, clinical management of individual patients, and disparate heterogeneous infectious disease surveillance.
The envisaged ontology, called IoT-MIDO, may aid prompt, timely, and concerted responses to infectious disease outbreaks with the effective allocation of limited resources. The novelty and uniqueness of the ontology lies in incorporating IoT-related concepts and concepts relevant to infectious disease surveillance and management. This will facilitate semantic interoperability between IoT-based individual patient monitoring and infectious disease management at the public health level. It could thus ease barriers to bringing benefits to individual as well as population health, which are often seen in isolation due to the historical dichotomization of clinical medicine and public health.

Methods

Objectives of the ontology

We designed an ontology to enhance the collaboration between IoT-driven patient health monitoring, clinical management of individual patients and infectious disease surveillance. The overall goals of the ontology are to enable the sharing and integration of data collected from disparate heterogeneous surveillance systems and to support risk analysis and early warning for better patient management and triage as well as for early response to infectious disease epidemics. As requirements for the ontology to achieve the intended goals, we formulated the competency questions described in Table 1. Competency questions (CQs) specify functional requirements for an ontology and are used to evaluate whether the ontology fulfills the elicited requirements [38]. The use of CQs has been proposed in several ontology engineering methodologies such as the Tropos methodology [39] and the Neon methodology framework [40].

Table 1. Competency questions to elicit and evaluate requirements for IoT-MIDO.

	IoT-driven patient monitoring, risk assessment, early warning, and patient management
CQ1	What information is gathered that can be used for early warning of patients' health risks?
CQ2	Which measurements are monitored using IoT devices?
CQ3	What are the main types of recommendations for patient management that can be utilized for patient triage according to the detected events (i.e., anomalies in early warning score)?
CQ4	How is it possible to perform contact tracing activity based on patient information?
	Clinical management of infectious diseases
CQ5	What is the outcome of an infectious disease process?
CQ6	Which treatment provides prophylaxis for the patient?
CQ7	Which treatment is used for the disease process?
CQ8	Which vaccine is used to mitigate the disease process?
CQ9	Which symptoms does a person playing the role of a symptomatic infectious agent carrier develop?
CQ10	Which laboratory tests are used for diagnosing a patient?
CQ11	Which risk factors can increase the risk of contracting an infectious disease?
	Epidemic risk analysis
CQ12	What is the risk score for infectious disease epidemic (i.e., a score to predict epidemic severity) and the associated risk level?
CQ13	What infectious disease control strategies has a country implemented?
CQ14	What is the strictness of a country's response to an infectious disease epidemic?
	Infectious disease surveillance
CQ15	On which infectious disease does a country conduct surveillance?
CQ16	What is an epidemic threshold to determine an infectious disease epidemic?
CQ17	What testing strategy does a country have?
CQ18	What types of disease surveillance data are collected for infectious disease surveillance?
CQ19	What kind of population-based statistics are computed?

CQ20	For which infectious diseases is the contact tracing activity performed?
CQ21	To which surveillance system are disease surveillance data reported?
CQ22	What is a case definition (i.e., a set of standard criteria for identifying cases to monitor the trend of the infectious disease under investigation) of an infectious disease?
	Transforming individual patient information into surveillance information
CQ23	What patient information is integrated into case-based surveillance data?

Design of the ontology

The IoT-based management of infectious disease ontology (IoT-MIDO) was developed using the Basic Formal Ontology (BFO) as the top-level ontology. BFO is a domain-independent upper-level ontology created to provide a common top-level structure for enhancing semantic interoperability across different domain ontologies [41]. This facilitates information sharing with multiple ontologies, built upon the BFO by making data smarter by adding both computer and human interpretable semantics to the raw data. Moreover, when developing the IoT-MIDO ontology, we reused the classes from existing BFO-based ontologies as much as possible to maximize interoperability with other BFO-based ontologies and databases that rely on them. Instead of introducing all of the classes and their definitions and properties, we demonstrate five use cases in the results section to show the potential usage of IoT-MIDO, using the simplified ontological models in a similar way to that done by the authors of IDODEN [34]. The complete ontology model and the definitions of all the classes are available in Media Appendix 1.

Table 2 shows the lists of ontology prefix classes that we reused in the use cases. When describing the following use cases, classes are written in italics starting with an upper case (e.g., *IoTStream*), and associations are written in italics starting with a lower case (e.g., *generatedBy*).

Table 2. List of ontology prefixes in which classes are reused and imported into use cases.

Ontology prefix	Ontology Full Name
FOAF	The Friend Of A Friend ontology [42]
IDOMAL	Malaria Ontology [43]
IoT-Stream	A Lightweight Ontology for IoT Data Streams [44]
NCIT	NCI Thesaurus OBO Edition [45]
SOSA	The Sensor, Observation, Sample, and Actuator ontology [46]
UO	The Units Ontology [47]

geo	GeoSPARQL Ontology [48]
CODO	The COviD-19 Ontology for cases and patient information [49]
EFO	Experimental Factor Ontology [50]
IDO	Infectious Disease Ontology [51]
SYMP	Symptom Ontology [52]
LABO	clinical LABoratory Ontology [53]
OGMS	Ontology for General Medical Science [54]
VO	Vaccine Ontology [55]
ICDO	International Classification of Disease Ontology [56]
GENEPIO	Genomic Epidemiology Ontology [57]
APOLLO_SV	Apollo Structured Vocabulary [37]
TRANS	Pathogen Transmission Ontology [58]
OBI	Ontology for Biomedical Investigations [59]
IDOBRU	Brucellosis Ontology [60]
iot4phm	IoT for Patient Health Monitoring Ontology [61]

Ethical considerations

This study did not collect any primary or secondary data. Therefore, it was not necessary to obtain approval from the Institutional Review Board (IRB). The authors, however, recognize the importance of responsible data handling. Furthermore, informed consent was not applicable, as the study did not directly involve human subjects.

Disclosure of generative AI usage

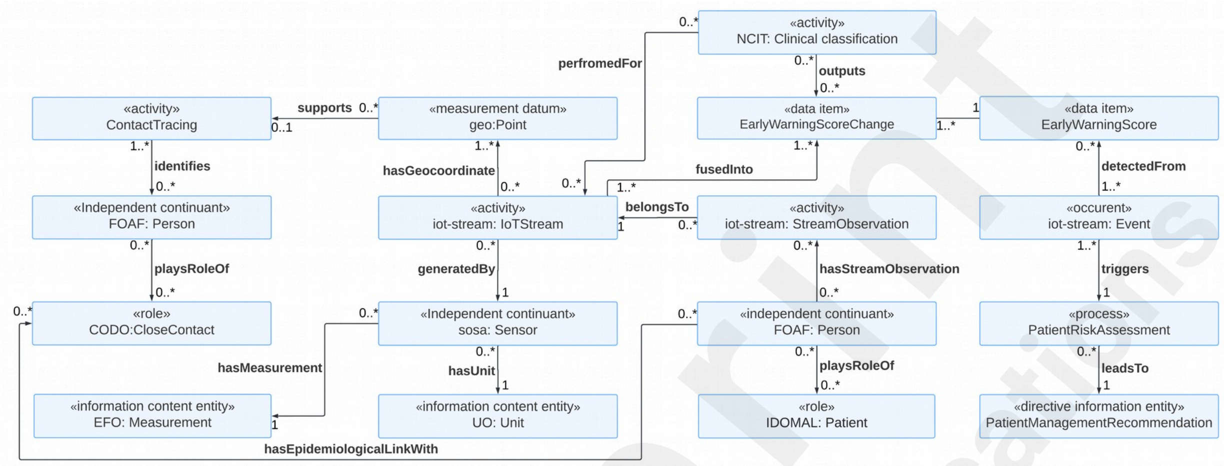
We used generative AI only for checking and improving formulations. We did not use it for creating the content of the study, including the ontology and use cases. .

Results

Use Case 1: IoT-driven patient monitoring, risk assessment, early warning, and risk management

The first use case is remote monitoring of patients' vital signs and other health information using IoT devices to detect their health risks proactively (Figure 1). We reused our IoT for patient health monitoring (IoT4PHM) ontology with modifications to create IoT-MIDO [61]. The IoT4PHM was built on the IoT-Stream ontology, consisting of four original concepts: IoTStream, StreamObservation, Analytics, and Event. In addition, the ontology is linked with six concepts imported from external ontologies: qoi:Quality, iot-lite:Service, sosa:Sensor, qu:QuanityKind, qu:Unit, and geo:Point. We developed the IoT4PHM by adding three classes to the IoT-Stream ontology: Patient, UnderlyingHealthcondition, and PatientManagement. The goal of the ontology was to facilitate data integration and sharing, knowledge representation, reasoning, and computer-assisted data analysis to enable IoT-based patient health monitoring and management for the prevention, early detection, and mitigation of patient deterioration. We first modified the IoT4PHM ontology to better conform to BFO by importing classes from BFO-based ontologies (e.g., replacing the qu:QuantityKind class with the Measurement class from EFO).

Figure 1. An ontology model for IoT-powered remote monitoring of a patient's vital signs, health risk assessment, early warning, and proactive health risk management.



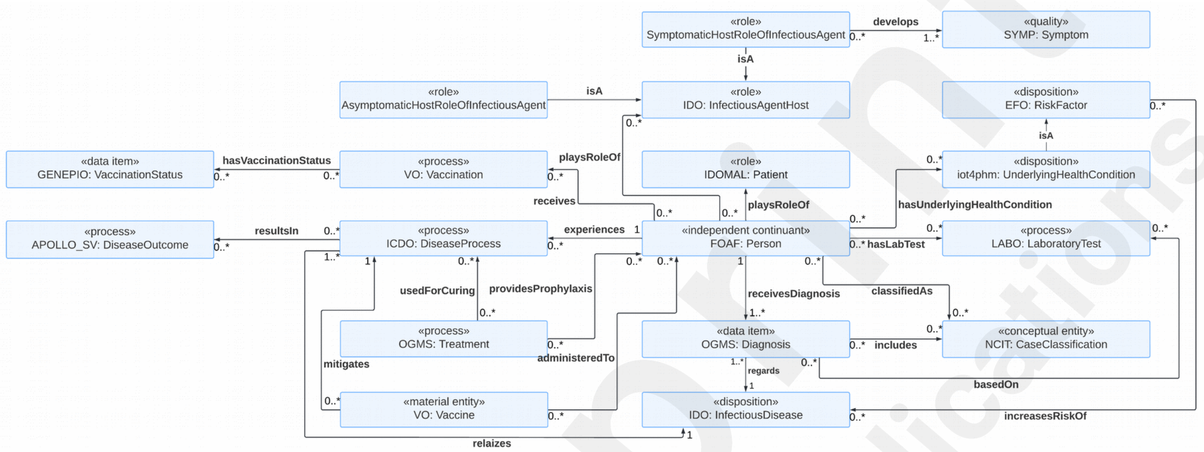
Use Case 2: Clinical patient management of infectious diseases

The second use case is clinical patient management of infectious diseases (Figure 2). When a *Person* who has acquired an infection seeks healthcare services, they start to play the role of a *Patient*. The *Patient* receives a *Diagnosis* of *InfectiousDisease* and is classified into *CaseClassification* imported from the Ontology for General Medical Science (OGMS) [54]. The subclasses of *CaseClassification* are *ConfirmedCase, ProbableCase, Suspec*tedCase, and *Negative*. The *ConfirmedCase* further has a subclass of *LaboratoryConfirmedCase*, a case confirmed by one or more of the laboratory methods that conform to the laboratory criteria included in the case definition. The *Diagnosis* includes information on *CaseClassification*. The latter only specifies whether a person has a particular infectious disease, while *Diagnosis* is a more comprehensive summary of patients' medical conditions. The *Diagnosis* is based on the *LaboratoryTes*t imported from the clinical LABoratory Ontology (LABO) [53]. The class includes information on the type of laboratory test the *Patient* has undergone.

An example of a laboratory test is a nucleic acid amplification test, such as a reverse transcription-polymerase chain reaction (RT-PCR), an antigen test, and an antibody test in the case of testing for SARS-CoV-2. When a *Person* is identified as a case of a particular infectious disease, they start to play the role of *InfectiousAgentHost,* which has two child classes: *AsymptomaticHostRoleOfInfectiousAgent* and *SymptomaticHostRoleOfInfectiousAgent*. In some infections such as SARS-CoV-2, a patient who never has symptoms associated with an infection (i.e., an asymptomatic patient) can still transmit an infectious agent to others. The *Person* who plays the role of *SymptomaticHostRoleOfInfectiousAgent* has the *developes* association with the *Symptom*.

The *Patient* experiences the *DiseaseProcess,* which realizes the *InfectiousDisease*. The *Vaccine* administered to the *Person* may mitigate the *DiseaseProcess*. The *Treatment* may either *provideprophylaxis* to the *Person* or is *usedForCuring* the *DiseaseProcess* that the *Person* experiences. The *DiseaseProcess* has the *resultsIn* association with the *DiseaseOutcome* such as death, convalescence, or long-term sequelae. The *Person* receives *Vaccination* using the *Vaccine*, which results in having *VaccinationStatus*.

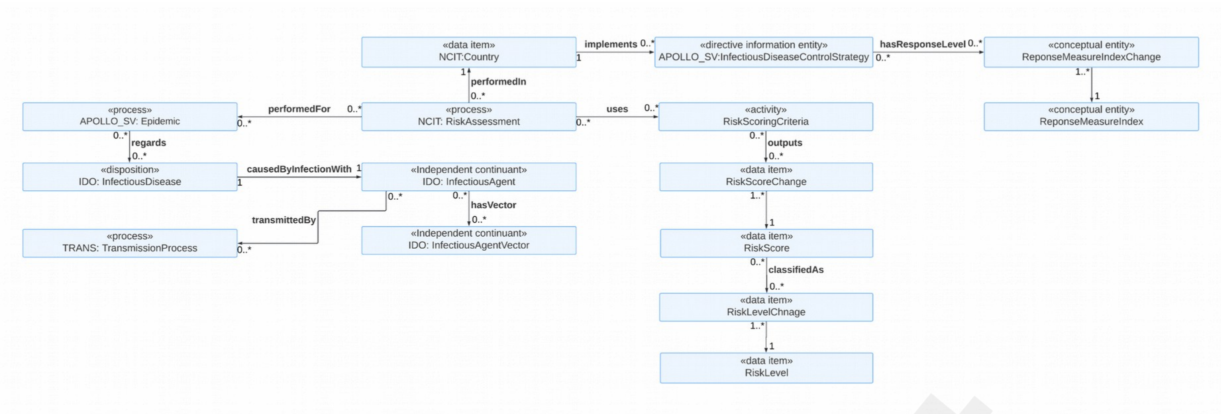
Figure 2. An ontology model for clinical management of infectious diseases in healthcare settings, supporting the seamless integration of data on vaccination status, laboratory test results, diagnosis, treatment, and underlying health conditions that may affect the course and outcome of an infectious disease.



Use Case 3: Epidemic risk analysis for timely response at the public health level

The third use case is epidemic risk assessment at the public health level (Figure 3). Risk assessment is essential for informing evidence-based public health decision-making about preparedness for and response to an infectious disease epidemic [62]. We thus added the *RiskAssessment* to handle information on risk assessment for the *Epidemic* regarding the *InfectiousDisease*. The information may include geographical scope (i.e., national, subnational, local/community-level), the time of the risk assessment performed, population group (e.g., general population and vulnerable population), and the person or organization who has performed the risk assessment. Various risk-scoring criteria can be used for risk assessment. It is informative to understand which risk scoring criteria are implemented for risk assessment. The *RiskScoringCriteria* handles information on the criteria's source, version, and developer. The *RiskScore* represents the overall risk score for an outbreak of an infectious disease in a specific geographical area. Since risk scores may change over time, we also add the *RiskScoreChange* class to track the risk trajectories.

Figure 3: An ontology model for autonomously assessing infectious disease epidemic risk at the national level, with the goal of improving preparedness and promoting timely response.



In addition, we included the *RiskLevel* class because risk level (e.g., high, moderate, low) is often determined based on the risk score and is frequently used for risk communications (e.g., risk assessment reports). Because risk scores can change over time, the risk level also changes accordingly. The *RiskLevelChange* is thus also included. An infectious disease epidemic's risk depends on the infectious agent's transmission mode. In our ontology, the *InfectiousAgent* is transmitted by the *TransmissionProcess*. For vector-borne infectious diseases, the *InfectiousAgent* has the *InfectiousVector.* Based on the risk assessment outputs, the *Country* may implement *InfectiousDiseaseControlStrategy*, which has a response level according to the *ResponseMeasureIndex*. The *ResponseMeasureIndexChange* is included to assess the changes in implemented infectious control strategies over time.

There are various ways of operationalizing risk assessment. For example, one possible way is to use risk score criteria that are proposed by the World Health Organization (WHO) or European Centre for Disease Prevention and Control (ECDC) [62], [63] (Figure 4). Both scoring criteria output the overall risk score based on impact score and likelihood/probability scores. For the WHO criteria, the impact is determined by three factors: vulnerability assessment, severity assessment, and coping capacity assessment. Another possible way to operationalize is to use the criteria suggested by Lesmanawati et al. [64] (Figure 5). The scoring criteria of their risk analysis framework, called EpiRisk, provide rapid risk prediction based on country-specific risk scores computed by summing disease and country risk scores. The disease risk score has eight parameters (i.e., the type of pathogen, basic reproductive number, mode of transmission, the occurrence of asymptomatic transmission, case fatality rate, therapy/drug availability, and vaccine availability). On the other hand, the country risk score has seven parameters, including the proportion of health expenditure to the country's GDP, the state of peace (assessed by the peace index proposed by the Peace Institute), the type of country border, population density, physician density and hospital beds per 1000 individuals.

Figure 4. Implementation example using ECDC rapid risk assessment methodology [63].

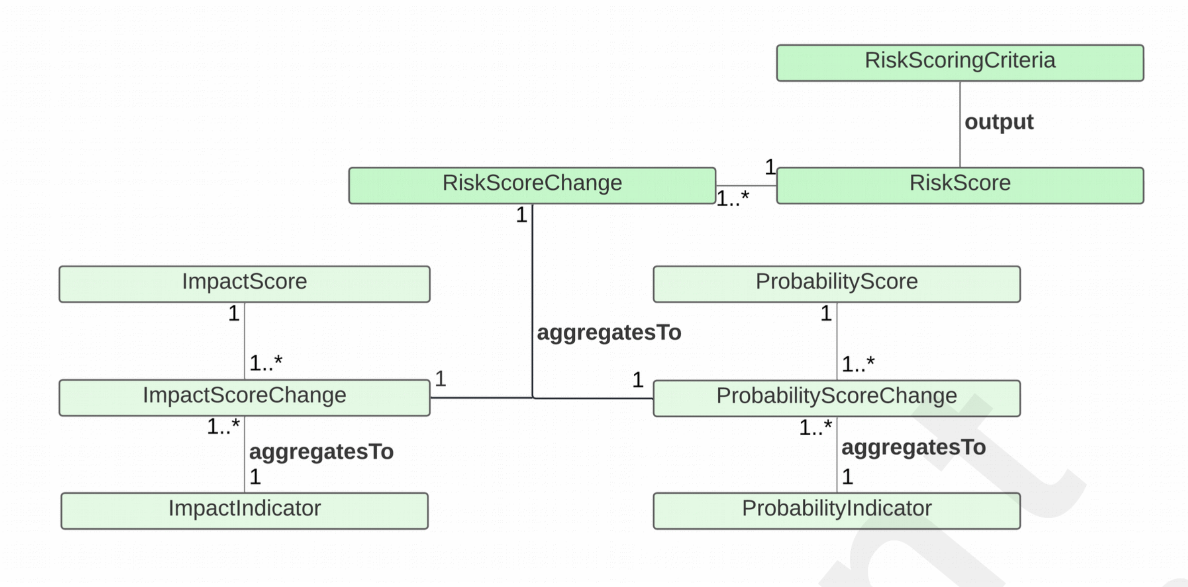
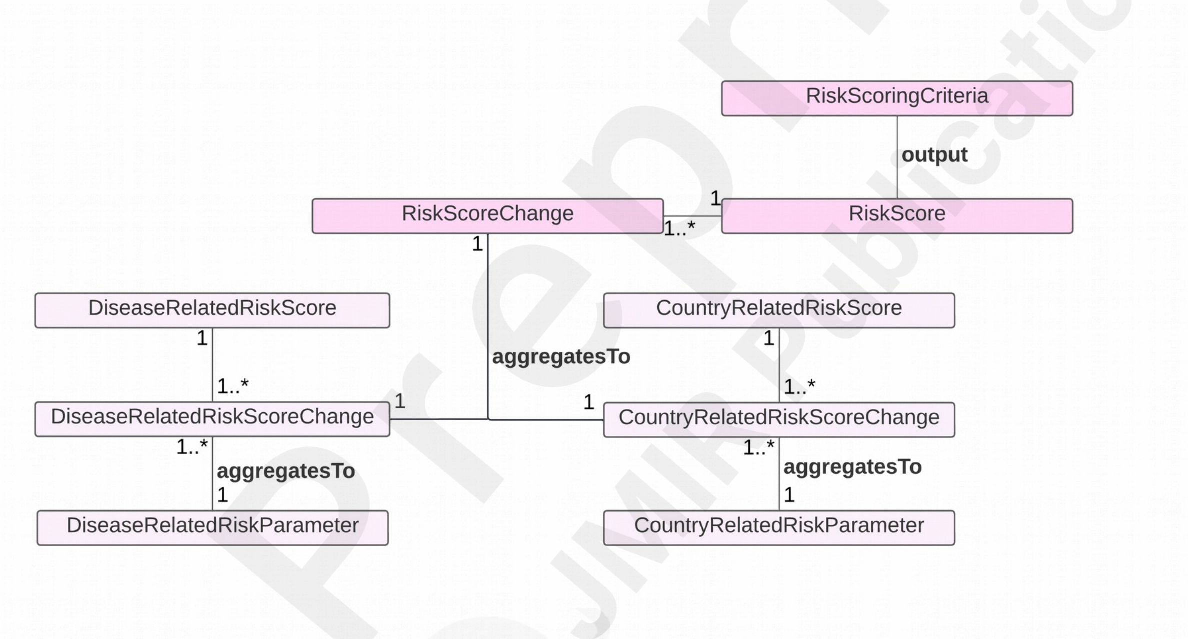


Figure 5. Implementation example of risk assessment based on the proposed framework by Lesmanawati et al. [64].



Use Case 4: Infectious disease surveillance

The fourth use case is modeling infectious disease surveillance (Figure 6). The *InfectiousDiseaseSurveillance* is a subclass of *DiseaseSurveillance*, which is a subclass of *HealthSurveillance*. The *HealthSurveillance* is performed for an *InfectiousDiseaseAgent*. The *InfectiousDiseaseSurveillance*, performed in a particular *Country*, collects *DiseaseSurveillanceData* that are reported to the *DiseaseSurveillanceSystem* and are used for computing the *PopulationBasedStatistic* based on the *Population* data of the *Country*. The *DiseaseSurveillanceData* has two subclasses: *CasedBaseDiseaseSurveillanceData* and *AggregatedDiseaseSurveillanceData*. We also included *DiseaseSurveillanceDataChange* for handling longitudinal surveillance data.

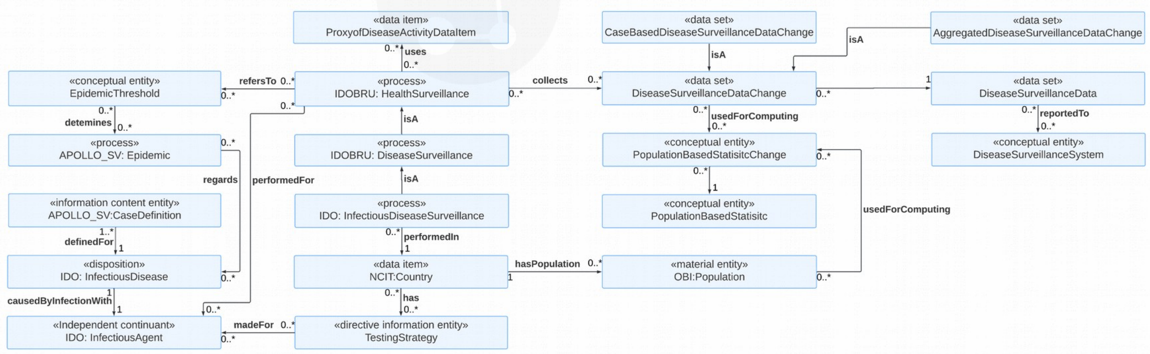
HealthSurveillance refers to the *EpidemicThreshold,* which defines the *Epidemic* regarding *IntctiousDisease*. The *CaseDefinition* is a set of standard criteria for identifying cases to monitor the trend of the *InfectiousDisease* under investigation [65]. Using uniform case definitions is essential for public health surveillance. By ensuring every case is equivalent, the number of cases and disease incidence across different time points and geographical areas can be meaningfully compared [66]. Case definitions can vary across countries, especially at an early stage of newly emerged infectious disease outbreaks [67]. Furthermore, the definitions can be modified as new evidence on infectious diseases becomes available. When analyzing and interpreting surveillance data, it is essential to know on which case definition the diagnosis of an infectious disease is based. The number of cases needs to be interpreted with caution when a new version becomes available. The *CaseDefinition* may include information on the source of the case definition and its version and has four subclasses: laboratory criteria, clinical criteria, epidemiological criteria, and diagnostic imaging criteria.

HealthSurveillance can utilize the *ProxyDiseaseActivityDataItem*, which can include data items that may provide an earlier indication of an epidemic spread than traditional epidemiological metrics such as confirmed cases or deaths [68]. Kogan et al. evaluated six digital data sources as proxies of COVID-19 activity to detect COVID-19 outbreaks as early as possible [68]:

1. Google Trends patterns for a suite of COVID-19-related terms
2. COVID-19-related Twitter activity
3. COVID-19-related clinician searches from UpToDate
4. Predictions by the global epidemic and mobility model (GLEAM), a state-of-the-art metapopulation mechanistic model
5. Anonymized and aggregated human mobility data from smartphones
6. Kinsa smart thermometer measurements

They found that increased digital data stream activity anticipates the increase in confirmed cases and deaths 2–3 weeks earlier than traditional surveillance methods. Although all metrics discussed in their study have limitations, the authors proposed using the combination of disparate health and behavioral data or early warning of increased COVID-19 activity [68]. We believe that those digital proxy data become an important asset for future infectious disease surveillance, thus they are included in our ontology.

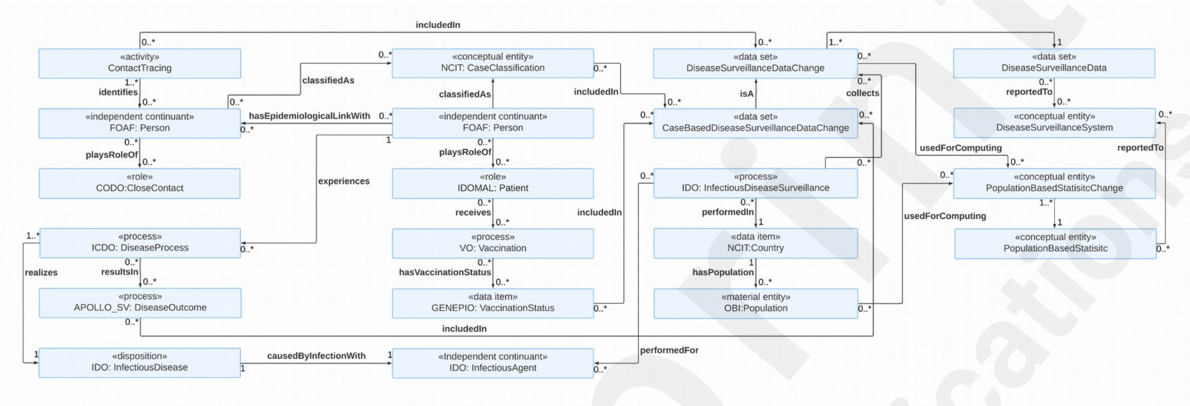
Figure 6. An ontology model for integrating dynamic data on infectious disease case definition, testing strategy, epidemic threshold, and statistics such as morbidity and mortality to track the spread of the disease at the national level.



Use Case 5: Ontological model of transforming patient information into surveillance information

The last use case is transforming patient information into surveillance information (Figure 7). Individual patient information such as *VaccinationStatus*, *CaseClassification*, and DiseaseOutcome are included in the *CaseBasedSurveillanceDataChange*. The information on *ContactTracing* may also be included in the *DiseaseSurveillanceDataChange*. This contributes to transforming individual patient information into datasets that can be used for infectious disease surveillance. The rest of the parts in Figure 5 have already been described in the use cases mentioned above.

Figure 7. An ontology model for aggregating individual patient data into surveillance information to provide actionable insights for informed decision-making and timely public health interventions.



Answers to competency questions

IoT-driven patient monitoring, risk assessment, early warning, and patient management

The answer to CQ1 is that the *EarlyWarningScore* class handles the information on the absolute value of the early warning score that is computed based on a particular *ClinicalClassification* scheme. An anomaly in the early warning score is detected as an *Event* that warns patients of health risks. A type of measurement that an IoT device is monitoring can be found in the *Measurement* class, which has the subclasses *PhysiologicalMeasurement* and *EnvironmentalMeasurement*. This answers CQ2. The answer to CQ3 is that the *PatientManagementRecommendation* has three subclasses: *HealthEducation, ReferralToHealthcar*e, and *EmergencyAlert* depending on the severity of the detected *Event* (e.g., an early warning score). The classification can serve as triage which helps ensure the effective allocation of limited resources to those who need it most and prevent overburdening healthcare systems. The answer to CQ4 is that an *IoTStream* generated by the *Sensor* has a *hasGeocoordinate* relationship with the *Point* class. The information on the geolocation handled in the *Point* class can support the *ContactTracing* activity by using positioning systems.

Clinical management of infectious diseases

The answer to CQ5 is that the outcome of an infectious disease process that a *Patient* experience is handled in the *DiseaseOutcome,* which can have instances of convalescence, death, ICU admission, and hospitalization. The answer to CQ7 and CQ8 is that the *Treatment* class has a *provideProphylaxis* association with the *Patient* class and has the *usedForCuring* association with the *DiseaseProcess*. Thus, the information on medications used for treating an infectious disease and prophylaxis can be handled using our ontology. The answer to CQ8 is that the Vaccine having a mitigated association with the *DiseaseProcess* may include information such as the identifier of the vaccine, the status code of the vaccine, the administered vaccine product, the vaccine manufacturer, the expiration date, the administered date, a performer who administered the vaccine to a person and possible side effects that can be caused by the vaccine. The answer to CQ9 is that the *Symptom* class represents symptoms developed by the symptomatic patients, which is especially important for characterizing newly emerging infectious diseases. The answer to CQ10 is that a *Person* playing the role of the *Patient* has a *hasLabTes*t association with the *LaboratoryTest* class, which has a *basedOn* association with the *Diagnosis*. The *LaboratyTest* represents the information on the laboratory test used for the patient. The *Diagnosis* includes the *CaseClassification* of the *Person,* which includes instances of confirmed case, laboratory-confirmed case, probable case, suspected case, and negative. The answer to CQ11 is that the *UnderlyingHealthCondition* class handles the information on the risk factors that pose an increased risk of an infectious disease. This information is critical to identifying high-risk groups who need to be prioritized for treatment and public health interventions.

Epidemic risk analysis

The answer to CQ12 is that the *RiskAssessment* class, which is performed for the *InfectiousAgent,* uses the *RiskScoringCriteria* to output the *RiskScore* that can be classified as a *RiskLevel* such as very high, high, medium, low, or very low. The answer to CQ13 is that the *Country* class has the *implements* association with the *InfectiousDiseaseControlStrategy* so that it allows the retrieval of information on how a country responds to an infectious disease epidemic at a given time. This information is important for evaluating the effectiveness of the control strategies in reducing the transmission of an infectious disease. The answer to CQ14 is that the *InfectiousDiseaseControlStrategy* has the *hasResponseLevel* association with the *ResponseMeasureIndex* class, which handles the information on the score showing the strictness of government responses. The *ResponseMeasureIndex* has a subclass of *OxfordCOVID-19GovenmentSringencyIndex*, which systematically collects data related to closure and containment, health and economic policy from over 180 countries and territories, which allows comparisons of policy responses within and across countries over time [69]. If another stringency index is preferred, a user can add it as a subclass of the *ResponseMeasureIndex* on demand. To track the change in stringency index scores, we introduced the *ResponseMeasureIndexChange* class in the model.

Infectious disease surveillance

The answer to CQ15 is that *HealthSurveillance*, which is the upper class of *InfectiousDiseaseSurveillance*, has a *performedIn* association with the *Country* and has a *performedFor* association with the *InfectiousAgent*. Thus, it is possible to model multiple infectious disease surveillances for different infectious agents that have been or are currently underway in a particular country. The answer to CQ16 is that *HealthSurveillance* has a *refersTo* association with the *EpidemicThreshold*. The *EpidemicThreshold* has a *determine* association with the *Epidemic* class and has a *thresholdFor* association with the *InfectiousDisease*. For example, crossing the epidemic threshold of 10% positive influenza laboratory tests indicates increased influenza activity and thus the start of the seasonal epidemic [70]. The answer to CQ17 is that *Country* has an association with the *SamplingStrategySpecification,* which has a subclass of *TestingStrategy*. Knowing that case counts are reported under each testing strategy is vital since the changes in testing strategies significantly influence case counts, and thus the epidemic dynamics need to be interpreted with caution when there are changes in testing strategies [71], [72]. The answer to CQ18 is that *DiseaseSurveillanceData* has two subclasses: *CaseBasedInfectiousDiseaseSurveillanceData* and *AggregatedInfectiousDiseaseData*. Thus, both types of surveillance data can be represented in each subclass since some countries report case-based data while others submit only aggregated data. The answer to CQ19 is that the *PopulationBasedStatistic* can have subclasses as necessary, such as *CaseNotificationRate, DeathRate, HospitalAdmissionRate, HospitalBedOccupancyRate, TestPositivityRate, TestingRate, VaccinationUptake, EffectiveReproductiveNumber, Seroprevalence,* and *CaseFatalityRate*. The *Country* has a *Population,* which has an *useForComputing* association with *PopulationBasedStatistic* since the (sub)population data are necessary for computing population-based statistics. The answer to CQ20 is that *ContactTracing* has a *performedFor* association with the *InfectiousDisease* class. Thus, contact tracing activities for multiple infectious diseases can be modeled simultaneously. The answer to CQ21 is that the *DiseaseSurveillanceData* has a *reportedTo* association with *DiseaseSurveillanceSystems* since several infectious disease surveillance systems collect different types of disease surveillance data that may be combined for analysis, and knowledge of the prominence of data are thus needed when referencing to the raw data. The answer to CQ22 is that the *CaseDefinition* has a *definedFor* association with the *InfectiousDisease* and has a *basedOn* association with the *CaseClassification*. The case definition is a set of criteria for systematically counting cases and thus is indispensable for infectious disease surveillance. The *CaseDefinition* has four subclasses: *LaboratoryCriteria, ClinicalCriteria, EpidemiologicalCriteria and DiagnosticImagingCriteria*.

Transforming individual patient information into surveillance information

The answer to CQ23 is that the *CaseClassification*, the *VaccinationStatus*, and the *DiseaseOutcome* classes have an *includedIn* association with the *CaseBasedDiseaseSurveillanceData*. Thus, individual patients' data on case classification, vaccination status, and disease outcomes can be integrated seamlessly into surveillance data through the *CaseBasedDiseaseSurveillanceData* class.

Therefore, the IoT-MIDO successfully addresses all of the CQs and can potentially be used for 1) IoT-driven remote patient health monitoring, risk assessment, early warning, and patient management in the context of infectious disease outbreaks, 2) clinical management of infectious diseases, 3) epidemic risk analysis, 4) infectious disease surveillance and 5) transforming individual patient information into surveillance information.

Discussion

Principal findings

The aim of our work was to design an infectious disease ontology that can support semantic data integration from disparate heterogeneous sources, including IoT-driven patient monitoring, clinical management, and infectious disease surveillance systems, and interlinking them. For this end, we developed the IoT-MIDO ontology, and by addressing competency questions and presenting five use cases, we demonstrated the potential of this ontology to enhance data interoperability, integration, and analysis across healthcare and public health domains and to assist healthcare practitioners and public health decision-makers in interpreting data available and their semantic relationships. The IoT-MIDO ontology can also serve as a starting point for enabling automated reasoning to drive actionable insights and informed decision-making to improve the health of individual patients as well as populations at large.

Comparing IoT-MIDO and OMOP CDM

Similar to our work, the Observational Medical Outcomes Partnership Common Data Model (OMOP CDM) can be used to support data integration from disparate sources in the context of infectious diseases. The model was developed by The Observational Health Data Sciences and Informatics (OHDSI) as an open-source common data standard to store observational health data [73]. The OMOP CDM effectively leverages a relational database schema to represent structured tabular data. Through integration with standardized vocabularies, it facilitates data analysis across multiple data sources and meaningfully compares and reproduces results from various observational studies [73].

However, in contrast to OMOP CDM, we have chosen an ontology modeling approach for knowledge representation because of the following capabilities. Ontologies offer more comprehensive and explicit knowledge representation, enriching data with richer semantic contexts and meanings, and modeling relationships between concepts in both human- and machine-interpretable formats [74]. Their expressiveness provides capabilities for automated reasoning, inferencing, and more precise querying over data.

Furthermore, ontologies provide greater flexibility in modeling complex relationships beyond mere hierarchies, accommodating any data formats, including structured, unstructured and semi-structured data [75]. While their flexibility enables more seamless data integration, their ease of extensibility allows for adaptation to the dynamic growth of data.

Ontologies are compatible with the World Wide Web Consortium (W3C) standards for the Semantic Web. The W3C endorses the use of International Resource Identifiers (IRIs) [76] and Uniform Resource Identifiers (URIs) [77]and URIs allows data from distributed and heterogeneous systems and databases to be linked together on the Web of Linked Data. This enables the establishment of semantic links between IoT-MIDO and existing specific infectious disease ontologies such as IDODEN [34] and IDO-COVID-19 [33], addressing the major issue of siloed information across disconnected systems, which prevents a comprehensive understanding of public health data.

Compatibility with other BFO-based ontologies

There are many well-established BFO-based ontologies existing, such as VIDO [35], CIDO [36] and IDO-COVID-19 [33], which are all ultimately extended from IDO [51]. IoT-MIDO adopts terms from those established ontologies, such as *Infectious disease* and *Infectious agent,* and *infectious disease surveillance* from IDO. This provides the possibility of easily extending IoT-MIDO further, according to the needs of ontology users to create semantic relationships with other BFO-based ontology concepts, for example, to characterize virus species and to overlapping or concurrently occurring multiple infectious disease epidemics.

The novelty and uniqueness characterizing our ontology lie in adopting concepts from IoT-Stream [44], such as *IoT stream* and *stream observation*. Incorporating these concepts, together with concepts related to infectious disease surveillance, control strategy, and response measure, which are both newly created or adopted from existing ontologies, builds a bridge to link IoT-based individual patient monitoring and early warning based on patient risk assessment to infectious disease epidemic surveillance at the public health level.

Strengths, limitations, and future directions

This study introduced IoT-MIDO, making a crucial initial step in addressing the long-standing issue of information silos caused by the historical segmentation between clinical medicine and public health, as well as the lack of interoperability across disparate systems. We made efforts to minimize the creation of new ontology concepts, opting instead to reuse existing domain-specific ontology concepts, particularly those based on BFO, to enhance semantic interoperability with external ontologies. Furthermore, digital data collected through IoT-driven systems provide information, including patients' real-time physical and mental health statuses, lifestyles and socio-demographics [78]. Such data are underutilized in the context of infectious disease surveillance, yet they can enhance existing surveillance systems by integrating with data obtained using conventional surveillance approaches. This integration can contribute to the development of more comprehensive epidemiological profiles at public health level.

However, two primary limitations must be acknowledged. Firstly, the ontology has yet to undergo formalization using dedicated ontology development tools, such as Protégé. We recognize this shortfall and intend to rectify it by formalizing the ontology in the Web Ontology Language (OWL) using Protégé in future iterations. Secondly, our evaluation of the ontology has been primarily focused on its ability to answer competency questions. We acknowledge the need for iterative empirical evaluation using live datasets to assess the capabilities of the ontology to integrate individual patient and surveillance data. Such evaluations are crucial not only for validating its practical utility but also for facilitating continuous enhancements to its quality over time.

These limitations highlight areas for future research and development. Moreover, several studies have proposed ontology-driven approach to ensure data compliance [79], [80], [81]. For instance, Debruyne et al. proposed an ontology, an extension of the provenance ontology called PROVO, to represent collected informed consent and its changes over time [79]. While many of existing studies predominantly focus on verifying compliance with policies regulations during the data processing stage, their ontology model is employed to assess potential compliance issues at dataset creation stage before the data undergoes processing.

Inspired by these studies, we plan to expand our ontology to include entities representing health data security, privacy, informed consent, and compliance with the legal requirements such as GDPR in future versions. Incorporating these entities is essential for safeguarding patients' sensitive information and ensuring compliance with legal requirements and/or contractual agreements. Enabling compliance verification at the dataset creation stage saves time and resources that would otherwise be spent on post hoc compliance analysis during data processing. Additionally, it strengthens data security and privacy measures by identifying and addressing vulnerabilities or compliance gaps early in the data lifecycle.

Conclusions

The primary aim of IoT-MIDO is to address the issue of information silos between clinical medicine and public health and enhance interoperability across different systems. Demonstrating IoT-MIDO's capability to answer diverse competency questions highlights its potential as a tool for achieving this goal. The seamless integration and sharing of clinical and public health data is particularly important during the rise of emerging infectious disease outbreaks when knowledge and evidence about infectious disease agents, transmission routes, and disease profiles are still scarce. This integration can accelerate the understanding of comprehensive and consistent epidemiological profiles and facilitate the effective and timely planning and execution of preventive and control measures.

As health data volume and complexity grow with increasing IoT applications in healthcare, standardizing such disparate heterogeneous data is essential for the effective and efficient use and integration of such data in infectious disease surveillance. Ontologies like IoT-MIDO help systematically represent knowledge to make disparate heterogeneous data ready for integration and comparable for further analyses, enabling informed and proactive interventions at both individual and population levels. IoT-MIDO's interoperability with existing ontologies, especially BFO-based ones like IDO, also facilitates seamless data exchange and the creation of comprehensive surveillance ecosystems.

Moreover, while the ontology standardizes the way of data sharing, it also provides flexibility to meet local needs, for example, by adding country-specific concepts and reference data such as codes while semantically linking them to global concepts or reference data.

Looking ahead, IoT-MIDO could underpin advanced decision support systems and predictive analytics tools in healthcare. These systems, leveraging computational semantic and logical reasoning, can assist in accurate diagnoses and optimal treatment decisions while enabling proactive surveillance and early detection of public health threats.

In conclusion, IoT-MIDO contributes to using IoT technologies to enhance healthcare delivery and modernize infectious disease surveillance, by providing means to bridge individual patient data with epidemiological data. By enabling the integration of these data from multiple domains, the ontology enhances interoperability, thereby advancing the use of IoT technologies in providing personalized preventive measures and care for individual patients as well as improving preparedness and response to infectious disease epidemics.

Authors’ Contributions

SL conceived the idea for the study. SL developed the ontology, and PJ acted as a reviewer of successive versions of the ontology. After SL drafted the manuscript, PJ reviewed and improved the content of the paper. PJ approved the final version of the manuscript for submission.

Conflicts of Interest

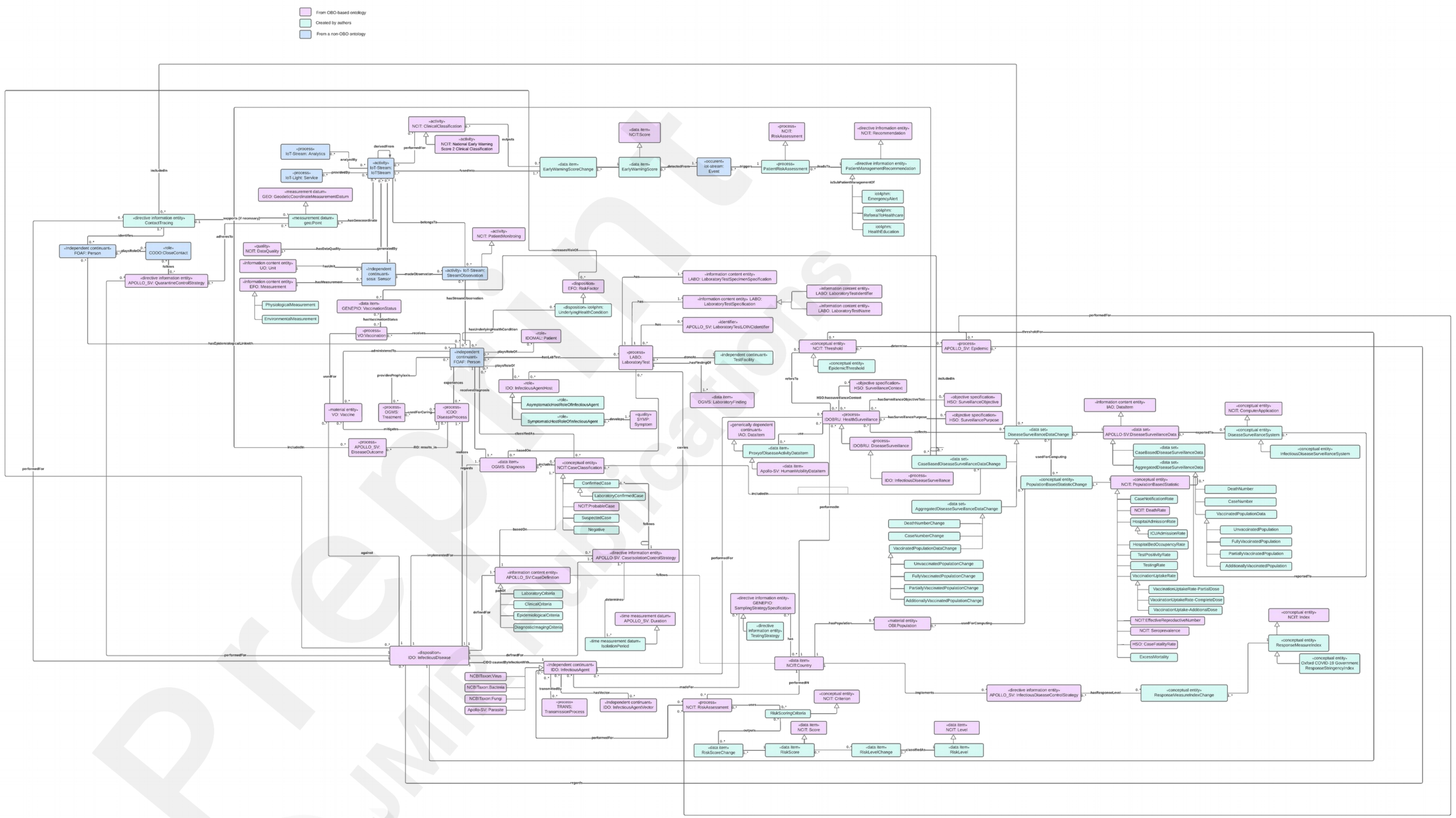
None declared.

Data Availability

The link to the IoT-MIDO model and additional information, including definitions of the entities and their attributes is provided in Multimedia Appendix 1.

Multimedia Appendix 1

The entire model of the IoT-MIDO and the definition table of the ontology classes are available via a link to IoT-MIDO files.



References

[1] GBD 2019 Diseases and Injuries Collaborators, “Global burden of 369 diseases and injuries in 204 countries and territories, 1990-2019: a systematic analysis for the Global Burden of Disease Study 2019,” *Lancet*, vol. 396, no. 10258, pp. 1204–1222, Oct. 2020, doi: 10.1016/S0140-6736(20)30925-9.

[2] S. Villa *et al.*, “The COVID-19 pandemic preparedness ... or lack thereof: from China to Italy.,” *Glob Health Med*, vol. 2, no. 2, pp. 73–77, Apr. 2020, doi: 10.35772/ghm.2020.01016.

[3] R. Palladino, J. Bollon, L. Ragazzoni, and F. Barone-Adesi, “Excess Deaths and Hospital Admissions for COVID-19 Due to a Late Implementation of the Lockdown in Italy.,” *Int J Environ Res Public Health*, vol. 17, no. 16, Aug. 2020, doi: 10.3390/ijerph17165644.

[4] S. Pei, S. Kandula, and J. Shaman, “Differential effects of intervention timing on COVID-19 spread in the United States.,” *Sci Adv*, vol. 6, no. 49, Dec. 2020, doi: 10.1126/sciadv.abd6370.

[5] J. Wise, “Covid-19: Global response was too slow and leadership absent, report finds,” *BMJ*, vol. 373, p. n1234, May 2021, doi: 10.1136/bmj.n1234.

[6] L. Cilloni *et al.*, “The potential impact of the COVID-19 pandemic on the tuberculosis epidemic a modelling analysis.,” *EClinicalMedicine*, vol. 28, p. 100603, Nov. 2020, doi: 10.1016/j.eclinm.2020.100603.

[7] E. Sherrard-Smith *et al.*, “The potential public health consequences of COVID-19 on malaria in Africa,” *Nat Med*, vol. 26, no. 9, pp. 1411–1416, 2020, doi: 10.1038/s41591-020-1025-y.

[8] L. de Vries, M. Koopmans, A. Morton, and P. van Baal, “The economics of improving global infectious disease surveillance.,” *BMJ Glob Health*, vol. 6, no. 9, Sep. 2021, doi: 10.1136/bmjgh-2021-006597.

[9] J. Murray and A. L. Cohen, “Infectious Disease Surveillance,” *International Encyclopedia of Public Health*, pp. 222–229, 2017, doi: 10.1016/B978-0-12-803678-5.00517-8.

[10] J. T. Kelly, K. L. Campbell, E. Gong, and P. Scuffham, “The Internet of Things: Impact and Implications for Health Care Delivery,” *J Med Internet Res*, vol. 22, no. 11, p. e20135, Nov. 2020, doi: 10.2196/20135.

[11] L. Hood and M. Flores, “A personal view on systems medicine and the emergence of proactive P4 medicine: predictive, preventive, personalized and participatory,” *N Biotechnol*, vol. 29, no. 6, pp. 613–624, Sep. 2012, doi: 10.1016/J.NBT.2012.03.004.

[12] J. W. Buehler, R. S. Hopkins, J. M. Overhage, D. M. Sosin, and V. Tong, “Framework for evaluating public health surveillance systems for early detection of outbreaks: recommendations from the CDC Working Group.,” *MMWR Recomm Rep*, vol. 53, no. RR-5, pp. 1–11, May 2004.

[13] K. J. Henning, “Overview of Syndromic Surveillance: What is Syndromic Surveillance?,” *MMWR Morb Mortal Wkly Repr*, vol. 53, no. suppl, pp. 5–11, Sep. 2004, Accessed: Jan. 23, 2022. [Online]. Available: https://www.cdc.gov/mmwr/preview/mmwrhtml/su5301a3.htm

[14] M. S. Smolinski, A. W. Crawley, J. M. Olsen, T. Jayaraman, and M. Libel, “Participatory Disease Surveillance: Engaging Communities Directly in Reporting, Monitoring, and

Responding to Health Threats.,” *JMIR Public Health Surveill*, vol. 3, no. 4, p. e62, Oct. 2017, doi: 10.2196/publichealth.7540.
[15] M. E. Kretzschmar, G. Rozhnova, M. C. J. Bootsma, M. van Boven, J. H. H. M. van de Wijgert, and M. J. M. Bonten, “Impact of delays on effectiveness of contact tracing strategies for COVID-19: a modelling study,” *Lancet Public Health*, vol. 5, no. 8, pp. e452–e459, Aug. 2020, doi: 10.1016/S2468-2667(20)30157-2.
[16] Z. Shakeri Hossein Abad *et al.*, “Digital public health surveillance: a systematic scoping review.,” *NPJ Digit Med*, vol. 4, no. 1, p. 41, Mar. 2021, doi: 10.1038/s41746-021-00407-6.
[17] P. Kostkova *et al.*, “Data and Digital Solutions to Support Surveillance Strategies in the Context of the COVID-19 Pandemic,” *Front Digit Health*, vol. 3, 2021, doi: 10.3389/fdgth.2021.707902.
[18] L. Li, D. Novillo-Ortiz, N. Azzopardi-Muscat, and P. Kostkova, “Digital Data Sources and Their Impact on People’s Health: A Systematic Review of Systematic Reviews,” *Front Public Health*, vol. 9, p. 645260, May 2021, doi: 10.3389/fpubh.2021.645260.
[19] M. Prosperi, J. S. Min, J. Bian, and F. Modave, “Big data hurdles in precision medicine and precision public health.,” *BMC Med Inform Decis Mak*, vol. 18, no. 1, p. 139, Dec. 2018, doi: 10.1186/s12911-018-0719-2.
[20] The Healthcare Information and Management Systems Society (HIMSS), “Interoperability in Healthcare | HIMSS.” Accessed: Oct. 25, 2021. [Online]. Available: https://www.himss.org/resources/interoperability-healthcare
[21] J. M. Alyea, B. E. Dixon, J. Bowie, and A. S. Kanter, “Chapter 9 - Standardizing Health-Care Data Across an Enterprise,” in *Health Information Exchange*, B. E. Dixon, Ed., Academic Press, 2016, pp. 137–148. doi: https://doi.org/10.1016/B978-0-12-803135-3.00009-8.
[22] World Health Organization, “ICD-11 implementation or transition guide,” 2019.
[23] P. L. Elkin *et al.*, “Evaluation of the content coverage of SNOMED CT: ability of SNOMED clinical terms to represent clinical problem lists,” in *Mayo Clinic Proceedings*, Elsevier, 2006, pp. 741–748.
[24] A. Coenen, “The international classification for nursing practice (ICNP®) Programme: advancing a unifying framework for nursing,” *Online J Issues Nurs*, vol. 3, 2003.
[25] D. Bender and K. Sartipi, “HL7 FHIR: An Agile and RESTful approach to healthcare information exchange,” in *Proceedings of the 26th IEEE International Symposium on Computer-Based Medical Systems*, 2013, pp. 326–331. doi: 10.1109/CBMS.2013.6627810.
[26] The Observational Health Data Sciences and Informatics, “Data Standardization – OHDSI.” Accessed: Jan. 04, 2023. [Online]. Available: https://www.ohdsi.org/data-standardization/
[27] F. E. Wood and M. J. Fitzsimmons, “Clinical Data Interchange Standards Consortium (CDISC) Standards and Their Implementation in a Clinical Data Management System,” *Drug Inf J*, vol. 35, no. 3, pp. 853–862, Jul. 2001, doi: 10.1177/009286150103500323.
[28] R. Facile *et al.*, “Use of Clinical Data Interchange Standards Consortium (CDISC) Standards for Real-world Data: Expert Perspectives From a Qualitative Delphi Survey,” *JMIR Med Inform*, vol. 10, no. 1, p. e30363, 2022, doi: 10.2196/30363.
[29] R. Studer, V. R. Benjamins, and D. Fensel, “Knowledge engineering: Principles and methods,” *Data Knowl Eng*, vol. 25, no. 1–2, pp. 161–197, Mar. 1998, doi:

10.1016/S0169-023X(97)00056-6.
[30] H. Liyanage, P. Krause, and S. de Lusignan, “Using ontologies to improve semantic interoperability in health data,” *BMJ Health & Care Informatics*, vol. 22, no. 2, p. 309, Apr. 2015, doi: 10.14236/jhi.v22i2.159.
[31] S. Heiler, “Semantic Interoperability,” *ACM Comput. Surv.*, vol. 27, no. 2, pp. 271–273, Jun. 1995, doi: 10.1145/210376.210392.
[32] R. A. Greenes, “Chapter 2 - A Brief History of Clinical Decision Support: Technical, Social, Cultural, Economic, and Governmental Perspectives,” in *Clinical Decision Support (Second Edition)*, R. A. Greenes, Ed., Oxford: Academic Press, 2014, pp. 49–109. doi: https://doi.org/10.1016/B978-0-12-398476-0.00002-6.
[33] S. Babcock, J. Beverley, L. G. Cowell, and B. Smith, “The Infectious Disease Ontology in the age of COVID-19,” *J Biomed Semantics*, vol. 12, no. 1, p. 13, 2021, doi: 10.1186/s13326-021-00245-1.
[34] E. Mitraka, P. Topalis, V. Dritsou, E. Dialynas, and C. Louis, “Describing the breakbone fever: IDODEN, an ontology for dengue fever.,” *PLoS Negl Trop Dis*, vol. 9, no. 2, p. e0003479, Feb. 2015, doi: 10.1371/journal.pntd.0003479.
[35] J. Beverley *et al.*, “Coordinating Coronavirus Research: The Virus Infectious Disease Ontology”.
[36] Y. He *et al.*, “CIDO, a community-based ontology for coronavirus disease knowledge and data integration, sharing, and analysis,” *Sci Data*, vol. 7, no. 1, p. 181, 2020, doi: 10.1038/s41597-020-0523-6.
[37] W. R. Hogan *et al.*, “The Apollo Structured Vocabulary: an OWL2 ontology of phenomena in infectious disease epidemiology and population biology for use in epidemic simulation,” *J Biomed Semantics*, vol. 7, no. 1, p. 50, 2016, doi: 10.1186/s13326-016-0092-y.
[38] M. Uschold and M. Gruninger, “Ontologies: principles, methods and applications,” *Knowl Eng Rev*, vol. 11, no. 2, pp. 93–136, 1996, doi: 10.1017/S0269888900007797.
[39] P. C. B. Fernandes, R. S. S. Guizzardi, and G. Guizzardi, “Using Goal Modeling to Capture Competency Questions in Ontology-based Systems,” *Journal of Information and Data Management*, vol. 2, no. 3, p. 527, Oct. 2011, doi: 10.5753/jidm.2011.1425.
[40] M. C. Suárez-Figueroa, A. Gómez-Pérez, and M. Fernández-López, “The NeOn Methodology framework: A scenario-based methodology for ontology development,” *Appl Ontol*, vol. 10, pp. 107–145, 2015, doi: 10.3233/AO-150145.
[41] R. Arp, B. Smith, and A. D. Spear, *Building ontologies with basic formal ontology*. Mit Press, 2015.
[42] M. Graves, A. Constabaris, and D. Brickley, “FOAF: Connecting People on the Semantic Web,” *Cat Classif Q*, vol. 43, no. 3–4, pp. 191–202, Apr. 2007, doi: 10.1300/J104v43n03_10.
[43] P. Topalis *et al.*, “IDOMAL: An ontology for malaria,” *Malar J*, vol. 9, no. 1, pp. 1–11, 2010, doi: 10.1186/1475-2875-9-230.
[44] T. Elsaleh, S. Enshaeifar, R. Rezvani, S. T. Acton, V. Janeiko, and M. Bermudez-Edo, “IoT-Stream: A Lightweight Ontology for Internet of Things Data Streams and Its Use with Data Analytics and Event Detection Services,” *Sensors (Basel)*, vol. 20, no. 4, p. 953, Feb. 2020, doi: 10.3390/s20040953.
[45] J. Balhoff, M. Brush, N. Vasilevsky, and M. Haendel, “NCI Thesaurus OBO Edition.”

Accessed: Jul. 07, 2022. [Online]. Available: http://purl.obolibrary.org/obo/ncit.owl
[46] K. Janowicz, A. Haller, S. J. D. Cox, D. Le Phuoc, and M. Lefrançois, “SOSA: A lightweight ontology for sensors, observations, samples, and actuators,” *Journal of Web Semantics*, vol. 56, pp. 1–10, 2019, doi: https://doi.org/10.1016/j.websem.2018.06.003.
[47] G. V Gkoutos, P. N. Schofield, and R. Hoehndorf, “The Units Ontology: a tool for integrating units of measurement in science,” *Database*, vol. 2012, p. bas033, Jan. 2012, doi: 10.1093/database/bas033.
[48] R. Battle and D. Kolas, “Geosparql: enabling a geospatial semantic web,” *Semantic Web Journal*, vol. 3, no. 4, pp. 355–370, 2011.
[49] B. Dutta and M. DeBellis, “CODO: an ontology for collection and analysis of COVID-19 data,” *arXiv preprint arXiv:2009.01210*, 2020.
[50] J. Malone, T. F. Rayner, X. Zheng Bradley, and H. Parkinson, “Developing an application focused experimental factor ontology: embracing the OBO Community,” in *Proceedings of the Eleventh Annual Bioontologies Meeting*, Toronto, Canada, 2008.
[51] L. G. Cowell and B. Smith, “Infectious disease ontology,” in *Infectious disease informatics*, Springer, 2010, pp. 373–395.
[52] W. Kibbe, S. Fuentes, and S. Arabandi, “Symptom Ontology.” [Online]. Available: http://purl.obolibrary.org/obo/symp.owl
[53] J.-F. Ethier, A. Barton, and P. Fabry, “clinical LABoratory Ontology .” Accessed: Jan. 22, 2022. [Online]. Available: http://purl.obolibrary.org/obo/labo.owl
[54] R. Scheuermann *et al.*, “Ontology for General Medical Science .” Accessed: Jul. 08, 2022. [Online]. Available: https://www.ebi.ac.uk/ols/ontologies/ogms
[55] Y. He *et al.*, “VO: vaccine ontology,” in *The 1st International Conference on Biomedical Ontology (ICBO-2009): July*, 2009, pp. 24–26.
[56] L. Wan *et al.*, “Development of the International Classification of Diseases Ontology (ICDO) and its application for COVID-19 diagnostic data analysis,” *BMC Bioinformatics*, vol. 22, no. 6, p. 508, 2021, doi: 10.1186/s12859-021-04402-2.
[57] D. M. Dooley, E. J. Griffiths, G. Gosal, F. S. L. Brinkman, and W. W. L. Hsiao, “The Genomic Epidemiology Ontology and GEEM Ontology Reusability Platform.,” in *JOWO*, 2017.
[58] L. Schriml and E. Mitraka, “Pathogen Transmission Ontology.” [Online]. Available: http://purl.obolibrary.org/obo/trans.owl
[59] A. Bandrowski *et al.*, “The Ontology for Biomedical Investigations,” *PLoS One*, vol. 11, no. 4, p. e0154556, Apr. 2016, doi: 10.1371/journal.pone.0154556.
[60] Y. Lin, Z. Xiang, and Y. He, “Brucellosis Ontology (IDOBRU) as an extension of the Infectious Disease Ontology,” *J Biomed Semantics*, vol. 2, no. 1, p. 9, 2011, doi: 10.1186/2041-1480-2-9.
[61] S. Lim, R. Rahmani, and P. Johannesson, “Semantic Enrichment of Vital Sign Streams through Ontology-based Context Modeling using Linked Data Approach,” in *10th International Conference on Data Science, Technology and Applications, DATA 2021*, 2021, pp. 292–299.
[62] World Health Organization, “Strategic toolkit for assessing risks: a comprehensive toolkit for all-hazards health emergency risk assessment,” Geneva, 2021. Accessed: Jan. 22, 2022. [Online]. Available: https://www.who.int/publications/i/item/9789240036086
[63] European Centre for Disease Prevention and Control, “Operational guidance on rapid risk assessment methodology,” Stockholm, 2019. doi: 10.2900/104938.

[64] D. A. S. Lesmanawati, P. Veenstra, A. Moa, D. C. Adam, and C. R. MacIntyre, "A rapid risk analysis tool to prioritise response to infectious disease outbreaks," *BMJ Glob Health*, vol. 5, no. 6, p. e002327, Jun. 2020, doi: 10.1136/bmjgh-2020-002327.
[65] Centers for Disease Control and Prevention, "Principles of Epidemiology | Lesson 1 - Section 5." Accessed: Jan. 04, 2023. [Online]. Available: https://www.cdc.gov/csels/dsepd/ss1978/lesson1/section5.html
[66] A.-H. El-Gilany, "COVID-19 caseness: An epidemiologic perspective," *J Infect Public Health*, vol. 14, no. 1, pp. 61–65, 2021.
[67] A. Peralta-santos, "Assessment of COVID-19 surveillance case definitions and data reporting in the European Union," Jul. 2020.
[68] N. E. Kogan *et al.*, "An early warning approach to monitor COVID-19 activity with multiple digital traces in near real time," *Sci Adv*, vol. 7, no. 10, Mar. 2021, doi: 10.1126/SCIADV.ABD6989.
[69] T. Hale *et al.*, "A global panel database of pandemic policies (Oxford COVID-19 Government Response Tracker)," *Nat Hum Behav*, vol. 5, no. 4, pp. 529–538, 2021, doi: 10.1038/s41562-021-01079-8.
[70] European Centre for Disease Prevention and Control, "Seasonal influenza 2021−2022," in *Annual epidemiological report* , 2022.
[71] A. Terriau, J. Albertini, E. Montassier, A. Poirier, and Q. Le Bastard, "Estimating the impact of virus testing strategies on the COVID-19 case fatality rate using fixed-effects models," *Sci Rep*, vol. 11, no. 1, p. 21650, 2021, doi: 10.1038/s41598-021-01034-7.
[72] M. M. Van Gordon, K. A. McCarthy, J. L. Proctor, and B. L. Hagedorn, "Evaluating COVID-19 reporting data in the context of testing strategies across 31 low-and middle-income countries," *International Journal of Infectious Diseases*, vol. 110, pp. 341–352, 2021.
[73] J. M. Overhage, P. B. Ryan, C. G. Reich, A. G. Hartzema, and P. E. Stang, "Validation of a common data model for active safety surveillance research.," *J Am Med Inform Assoc*, vol. 19, no. 1, pp. 54–60, 2012, doi: 10.1136/amiajnl-2011-000376.
[74] M. Uschold, "Ontology and database schema: What's the difference?," *Appl Ontol*, vol. 10, no. 3–4, pp. 243–258, 2015.
[75] B. Ramis Ferrer *et al.*, "Comparing ontologies and databases: a critical review of lifecycle engineering models in manufacturing," *Knowl Inf Syst*, vol. 63, no. 6, pp. 1271–1304, 2021.
[76] M. Duerst and M. Suignard, "RFC 3987: Internationalized resource identifiers (iris)." RFC Editor, 2005.
[77] T. Berners-Lee, R. Fielding, and L. Masinter, "RFC 3986: Uniform resource identifier (uri): Generic syntax." RFC Editor, 2005.
[78] E. Katsoulakis *et al.*, "Digital twins for health: a scoping review," *NPJ Digit Med*, vol. 7, no. 1, p. 77, 2024.
[79] C. Debruyne, H. J. Pandit, D. Lewis, and D. O'Sullivan, "'Just-in-time' generation of datasets by considering structured representations of given consent for GDPR compliance," *Knowl Inf Syst*, vol. 62, no. 9, pp. 3615–3640, 2020.
[80] A. Castro, V. A. Villagra, P. Garcia, D. Rivera, and D. Toledo, "An ontological-based model to data governance for big data," *IEEE Access*, vol. 9, pp. 109943–109959, 2021.
[81] M. Palmirani, M. Martoni, A. Rossi, C. Bartolini, and L. Robaldo, "Pronto: Privacy ontology for legal compliance," in *Proc. 18th Eur. Conf. Digital Government (ECDG)*,

2018, pp. 142–151.

Supplementary Files

Untitled.
URL: http://asset.jmir.pub/assets/a58f05d0101cd8b032510d001b680466.docx

Untitled.
URL: http://asset.jmir.pub/assets/b4a9a45f4124903b1a65830c3f5f0787.docx

Untitled.
URL: http://asset.jmir.pub/assets/fd4b303bbf95bd50541688275b7891d6.docx

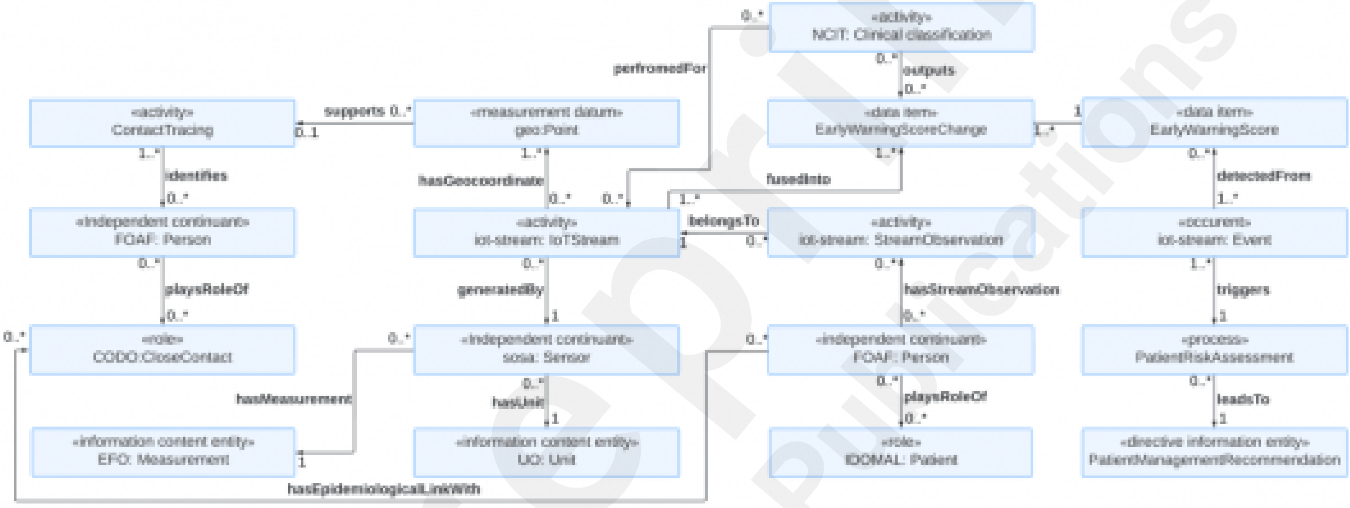
Untitled.
URL: http://asset.jmir.pub/assets/c3de60c3dd87e8d640453b2abee9b291.docx

Untitled.
URL: http://asset.jmir.pub/assets/7a22343720d8773ac216815e1a3b4530.docx

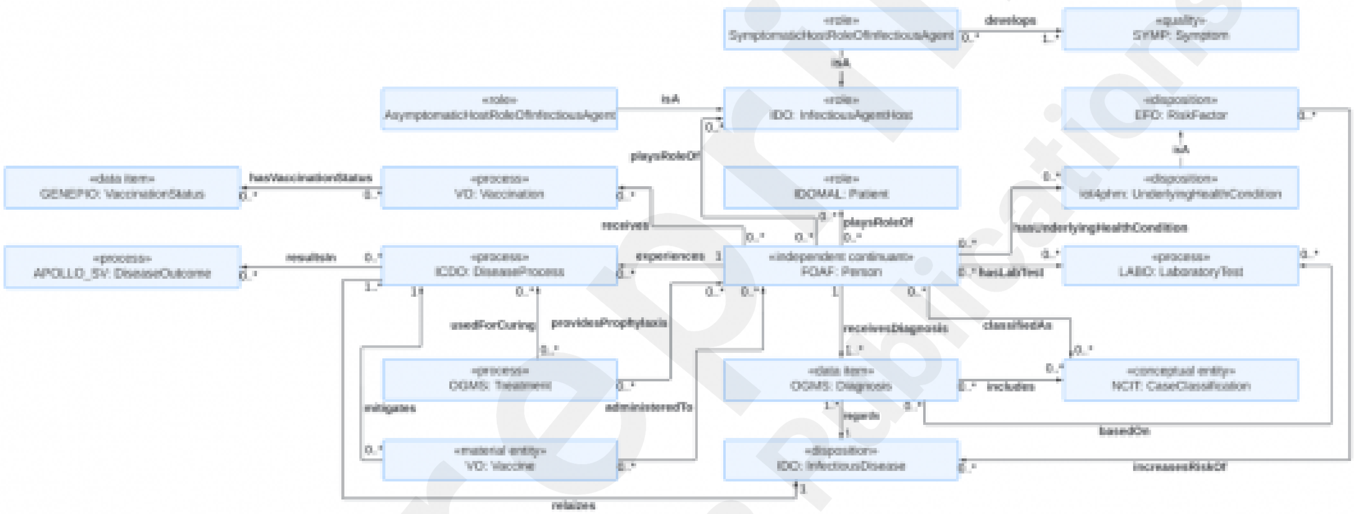
Untitled.
URL: http://asset.jmir.pub/assets/df5eaba0c3208fe4727b93e5a3960ce4.docx

Figures

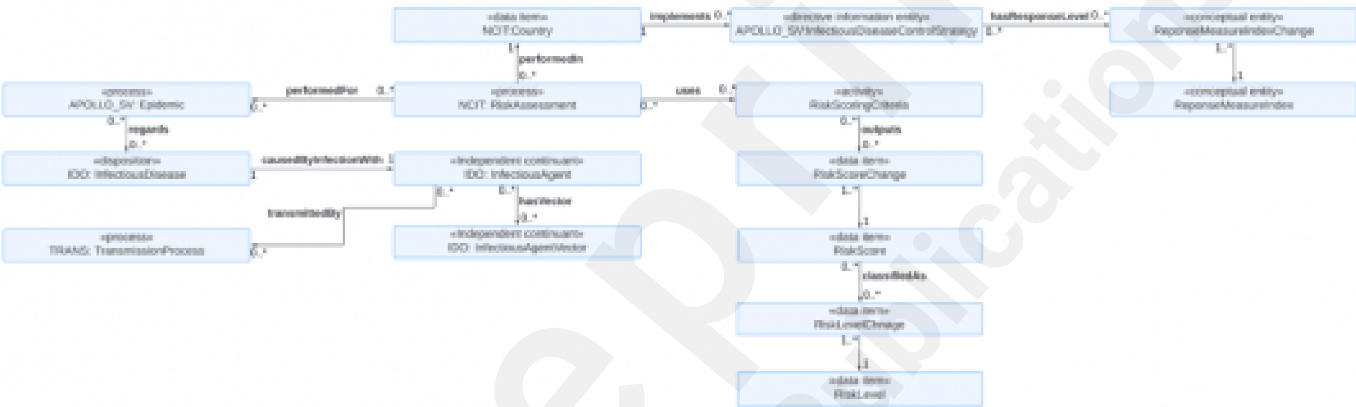
An ontology model for IoT-powered remote monitoring of a patient’s vital signs, health risk assessment, early warning, and proactive health risk management.



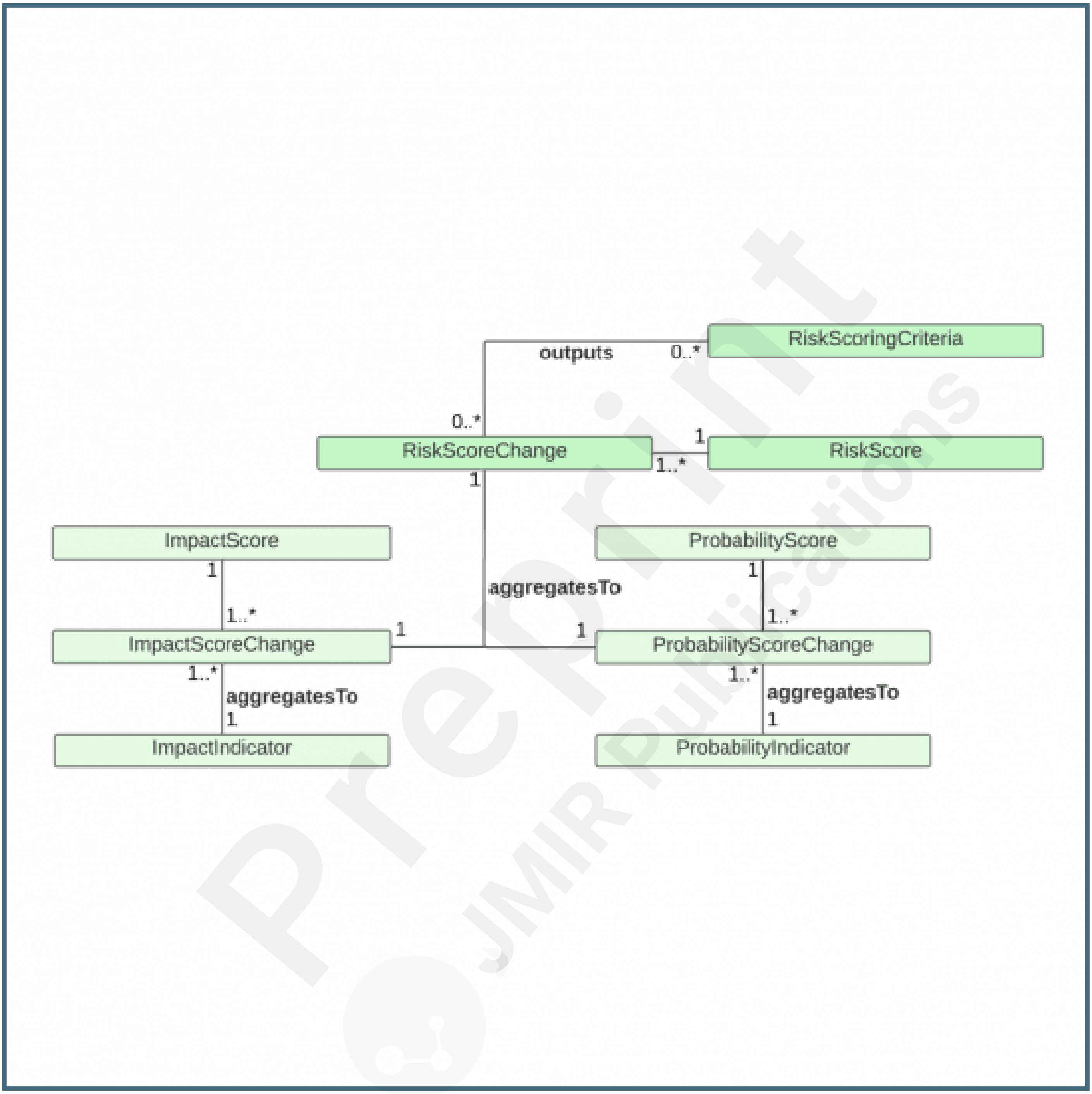
An ontology model for clinical management of infectious diseases in healthcare settings, supporting the seamless integration of data on vaccination status, laboratory test results, diagnosis, treatment, and underlying health conditions that may affect the course and outcome of an infectious disease.



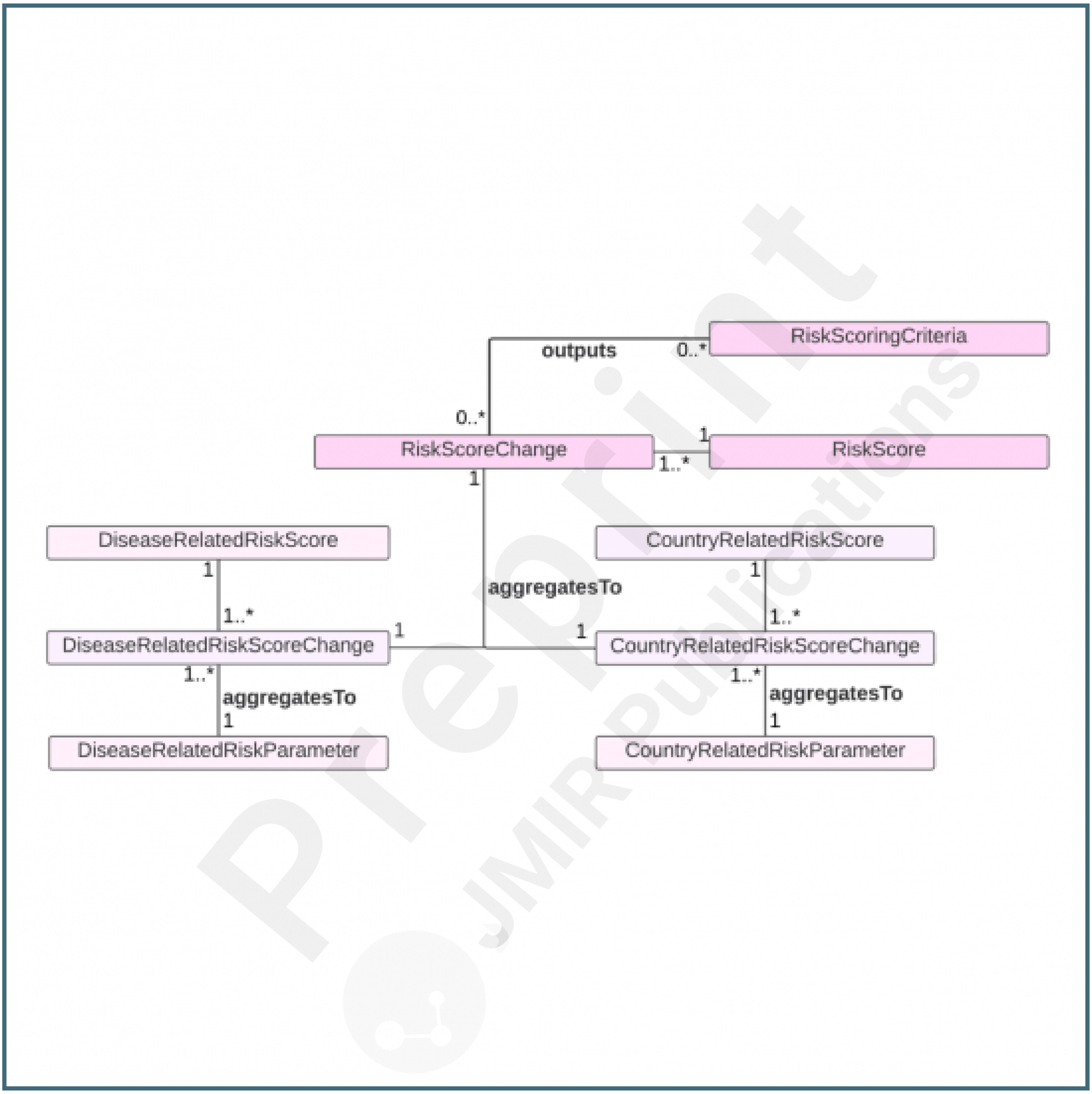
An ontology model for autonomously assessing infectious disease epidemic risk at the national level, with the goal of improving preparedness and promoting timely response.



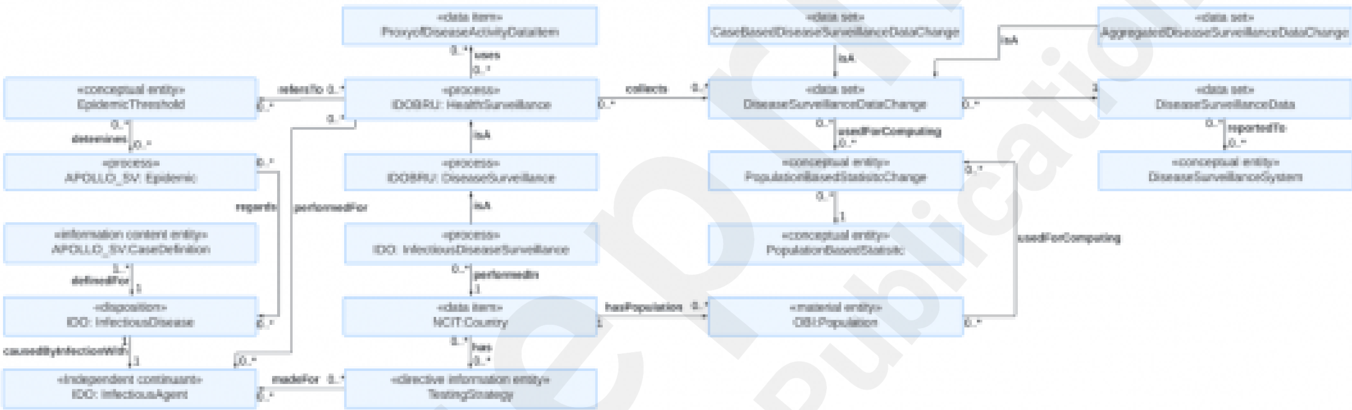
Implementation example using ECDC rapid risk assessment methodology [63].



Implementation example of risk assessment based on the proposed framework by Lesmanawati et al. [64].



An ontology model for integrating dynamic data on infectious disease case definition, testing strategy, epidemic threshold, and statistics such as morbidity and mortality to track the spread of the disease at the national level.



An ontology model for aggregating individual patient data into surveillance information to provide actionable insights for informed decision-making and timely public health interventions.