Towards the Integration of Abnormality in Diseases

Yuki Yamagata, Kouji Kozaki
Department of Knowledge Science
I.S.I.R, Osaka University
8-1 Mihogaoka, Ibaraki, Osaka, Japan
yamagata@ei.sanken.osaka-u.ac.jp

Takeshi Imai, Kazuhiko Ohe
Department of Medical Informatics
Graduate School of Medicine
The University of Tokyo
7-3-1, Hongo, Bunkyo-Ku, Tokyo, Japan

Riichiro Mizoguchi
Research Center for Service Science
Japan Advanced Institute of Science and Technology
1-1 Asahidai, Nomi, Ishikawa, Japan

***Abstract*—Knowledge of abnormalities is important for understanding diseases. A number of resources provide terms related to abnormalities in biomedicine. In this paper, we investigate the differences in the hierarchical structure of these biomedical resources and discuss issues of reuse and integration with regard to abnormal states based on ontological theory. Then, we show a solution for integrating them by linking abnormal states in our abnormality ontology to other resources according to the meaning of the concepts at each level.**

Keywords—ontology; abnormality; disease

I. INTRODUCTION

In order to understand diseases, one must first capture the abnormal states in diseases adequately. They are observed as symptoms by patients or as signs by clinicians in clinical findings. Clinical test data can also provide evidence for the existence of abnormal states. In addition, in basic research, through analysis of disease models of animals, many researchers make efforts to understand how causative agents are related in the etiological process. Moreover, abundant knowledge about abnormalities is available in scientific articles. The above observation shows that abnormality is a key factor for capturing diseases, assisting in the interoperability between basic research and clinical medicine for integrating a wide variety of knowledge across domains. We have been involved in the development of a disease ontology [1]. As part of this research, we have focused on abnormal states in the definition of diseases and have rigorously systematized an abnormality ontology [2, 3].

A number of terminologies and standard vocabularies have been developed for many years, and recently, ontologies have also been constructed in the biomedical domain. They offer useful data, and some of these terms include abnormality concepts. In order to make efficient use of these resources, we need a solution that ensures interoperability between abnormality knowledge across domains. To achieve this, first, one has to elucidate one's own perspectives of the resources before using them and to make the meanings of concepts explicit.

In this paper, we discuss differences of existing biomedical resources based on ontological engineering theory with respect to abnormalities. Next, we investigate the relationships between abnormal states in our abnormality ontology and corresponding terms in biomedical resources. By mapping concepts from our abnormality ontology to other biomedical concepts, we can establish mutually complementary relationships between biomedical concepts in different levels, which will contribute to ensuring interoperability across biomedical resources. Our goal is to provide not only a theory for a better understanding of abnormal states but also useful information for clinical practice. We are planning to integrate abnormalities in the definition of diseases in the Department of Cardiovascular Medicine and other medical departments at the University of Tokyo Hospital in our ontology with external biomedical resources at each level of meaning as a concrete example.

This paper is organized as follows. In Section 2, we introduce our ontology of abnormalities. In Section 3, we examine the characteristics of biomedical resources, discuss the problem with them based on ontological theory and propose a solution for integration with respect to abnormal states. Finally, in Section 4, we present concluding remarks and a give an outline of future work.

II. ABNORMALITY ONTOLOGY

This section provides an overview of our abnormality ontology.

A. Three-Layer Ontological Model of Abnormal States

In order to develop an *is-a* hierarchical tree of abnormalities, it is important to conceptualize them from a consistent viewpoint. To this end, we have been developing an abnormality ontology [2, 3] having a three-layer structure:

- Level 1: Generic abnormal states
- Level 2: Object-dependent abnormal states
- Level 3: Specific disease-dependent abnormal states

The top-level categories define very basic and generic concepts, for example, "small in area," "hypofunction," etc., which are commonly used not only in clinical medicine but also in other domains. Level 2 concepts are dependent on objects. In the lower level of the tree, concepts are designed to represent abnormalities at specific human organ / tissue / cell levels. For example, by specializing "small in area" at Level 1, " narrowed cross-sectional area of tube ", where the cross-sectional area of a tubular structure has become narrowed, is defined at Level 2, and this is further specialized in the definitions "vascular stenosis" (blood vessel-dependent), "arterial stenosis", "coronary

artery stenosis" (coronary artery-dependent), and so on. Level 3 concepts are captured as specific disease-dependent (context-dependent) abnormal states. For example, "coronary artery stenosis" at level 2 is defined as a constituent of ischemic heart disease at Level 3. In our ontological approach, common concepts can be kept distinct from specific ones and can be appropriately defined according to their context.

As of 11 May 2013, our ontology has 21,669 abnormal states constituted of 6,302 diseases across 13 medical departments, and among them, clinicians have currently refined concepts of 9,985 abnormal states constituted of 1,602 diseases across five major departments.

B. *Representation of Abnormal State*

In medicine, abnormal states are interpreted from the diverse perspectives of specialists such as clinicians, pathologists, biologists, geneticists, and so on, and correspondingly a variety of representations of abnormal states are used. Therefore, we classified the abnormal states into three categories: a property[1] (e.g., hypertension), a qualitative representation (e.g., blood pressure is high), and a quantitative representation (e.g., blood pressure 180 mmHg). In previous work, we proposed a Property-Attribute interoperable representation framework for abnormal states [2, 3] on the basis YAMATO [4].

We captured all abnormal states as properties[2] represented by a tuple: <Property (P), Property Value (Vp)>, e.g., <stenosis, true>. We specified the property by decomposing it into a tuple: <Attribute (A), Attribute Value (V)>. The Attribute Value can be either a Qualitative Value (Vql) or a Quantitative Value (Vqt). For example, "arterial stenosis" is decomposed into <cross-sectional area (A), small (Vql)> as a qualitative representation, or <cross-sectional area (A), 5mm^2 (Vql)> as a quantitative representation. Then, we introduce "Object" to identify the target object, and we represent an abnormal state as a triple: <Object (O), Attribute (A), Attribute Value (V)>. This is the basic form in our representation model of abnormalities. In addition, we introduce "Sub-Object (SO)" as an advanced representation for what will be focused on. For example, in the case of "hyperglycemia", since the glucose concentration (A) means the ratio of the focused object (SO) relative to the whole mixture (O), the representation of "hyperglycemia" is a quadruple, <blood (O), glucose (SO), concentration (A), high (V)>. In another case of an advanced representation, "Colonic polyposis" is described as <colon (O), polyp (SO), number (A), many (V)>. Our model can deal with both clinical test data and abnormal states in the definition of diseases. The clinical test data can be represented in the form <Object (O), Attribute (A), Quantitative Value (Vqt)> (OAVqt), which can be converted into a property representation form <Object (O), Property (A), Property Value (Vp)> (OPVp) via a qualitative representation form. For example, in terms of the state of hypertension[3], our model ensures interoperability among the forms <blood (O), pressure (A), 180mmHg (Vqt)>, <blood pressure, high>, and <hypertension, true>. Therefore, our model realizes interoperability between test data and abnormal states in the definition of diseases. In practice, we found that more complicated terms, such as adding modifier words (e.g., transient hypertension), and three types of compound words (e.g., "blood pressure" as an OA type: < blood (O), pressure (A)>). We have developed a guideline and deal with them as a variation of data representation.

III. UTILIZATION OF BIOMEDICAL RESOURCES

In this section, we investigate the hierarchy of "coronary artery stenosis" in existing biomedical resources and perform a comparison between our ontology and existing biomedical resources to make them interoperable with each other with respect to abnormalities.

A. *Charestictics of biomedical resources*

1) *SNOMED-CT and MeSH Terminologies*

a) SNOMED-CT: Systematized Nomenclature of Medicine-Clinical Terms (SNOMED-CT) is a clinical terminology first developed by the College of American Pathologists (CAPs) and is currently maintained by the International Health Terminology Standards Development Organization (IHTSDO) [5]. The major characteristic of SNOMED-CT is a large collection of clinical terms that contain more than 310,000 concepts. It is widely used as an international standard vocabulary. SNOMED-CT has a hierarchical structure. The root concept of the hierarchy is named "SNOMED-CT," and there are 19 top-level categories, including Clinical finding/disorder, Body structure, Organism, Substance, and other things important for clinical health. Most of the abnormal states are included in the category Clinical finding/disorder. In the hierarchical tree, concepts are linked by *is-a* relationships. SNOMED-CT allows multiple inheritances. In addition, one concept can have relationships other than *is-a*, like "finding site," "method," "clinical course" and so on, to connect concepts between different categories, e.g., "occlusion of artery *has finding site* arterial structure."

b) MeSH: Medical Subject Headings (MeSH) is a thesaurus of medical terms developed by the National Library of Medicine and is used for indexing biomedical articles [6]. The 2014 MeSH contains 27,149 terms as descriptors (MeSH headings) and 219,000 synonyms as entry terms. A hierarchical tree called "MeSH tree" is organized into 16 categories, e.g., anatomy, organisms, diseases, chemicals, and drugs. Most of the concepts of abnormal states are classified into the diseases category. MeSH also allows multiple inheritances.

2) PATO and HPO Phenotype Ontologies: In biomedical domains, scientists observe entities through experiments or clinical findings to capture abnormal states. Therefore, repre-

[1] The property discussed here is based on ontological theory, not Web Ontology Language (OWL) *property* which means a link (relation) between two nodes.

[2] A state derived from or associated with a property (P) is defined as: "a temporal entity derived by a time-indexed property (P) in which the bearer participates". A property has a bearer, while a state has the bearer as a participant.

[3] Our model can describe the mechanism "hypertension" more specificity by using other factors; e.g., accumulation of atheroma in arterial wall <arterial wall (O), atheroma (SO), quantity (A), much (V) >→decreased elasticity of arterial wall <arterial wall (O), elasticity (A), low (V) >→hypertension.

sentations of knowledge about abnormalities are usually given as phenotypes.

Phenotypic Quality Ontology (PATO) is an ontology of phenotypic qualities for annotating biological phenotypes across species [7]. PATO provides qualities to describe phenotypic information, such as size, color, weight and so on, and qualities have abnormal states as subclasses of each quality; for example, size has lower concepts like "increased/decreased length," "increased/decreased volume," "dwarf-like", in addition to length, area, and volume. PATO is widely used as a standard vocabulary for biological measurement in biocommunities. PATO also plays a role as a reference ontology for databases of specific species, such as *Drosophila* [8].

The Human Phenotype Ontology (HPO) is an ontology for human specific phenotypes [9]. It has been developed using phenotypic information from the human genetic diseases database Online Mendelian Inheritance in Man (OMIM) [10] and medical articles, and provides over 10,000 terms. HPO classifies human phenotypes into three categories: 1) mode of inheritance, 2) onset and clinical course, and 3) phenotypic abnormalities. Abnormal state concepts are mainly subclasses of 3) phenotypic abnormalities.

3) LOINC: Logical Observation Identifiers, Name and Codes (LOINC) was developed by the Regenstrief Institute as a universal code for exchanging clinical test data [11]. It is widely used in 185 countries. LOINC provides six fields (component (analyte), property, timing of the measurement, sample type, scale type, and method). By combining the contents of each field with colons, the name of a clinical test can be described. For example, "a test for glucose tolerance about after 2 hours serum glucose for 100g oral" is represented by "GLUCOSE^2H POST 100 G GLUCOSE PO:MCNC:PT:SER/PLAS:QN."

B. *Ontological Issues and Solution with Regard to Abnormality*

1) PATO and HPO: Both PATO and HPO representations are based on property representations in accordance with the upper ontology Basic Formal Ontology (BFO) [12]. Therefore, one could say that these representations are similar to our property representation of abnormal states. However, in PATO and HPO, the relationship between Property (P) and Attribute (A) & Value (V) is unspecified, which leads to a lack of interoperability with clinical test data. The data values are mostly quantitative, and clinical test data are represented in terms of A and V. To achieve interoperability with clinical test data, not only the OP form but also the OAV form is required. Our model can deal with both representation forms and can achieve interoperability between OP and OAV, which will be of great assistance in medical practice

The next issue is that PATO does not differentiate Property (P) from Attribute (A) in the same way as BFO does, which results in multiple inheritance from two parents in the PATO hierarchical tree: one is abnormal states, and the other is property (quality). For example, "decreased area" in PATO has two super-classes, "decreased size" and "area" (Fig. 1). Such multiple inheritance makes things complicated, and it is difficult to understand the underlying meaning of concepts. If computers integrate PATO concepts and other resources in a naïve way, inappropriate consequences might be derived unexpectedly. For example, HPO provides the concept of "coronary artery stenosis," and an incorrect relation such as "coronary stenosis *is-a* area" might be derived by a naïve combination of PATO and HPO. "Area" is not an abnormal state. In our ontology, parameters such as pressure, area, concentration, and so on, are properly dealt with as an attribute (A), so that there is no possibility of deriving inappropriate consequences such as the one shown above.

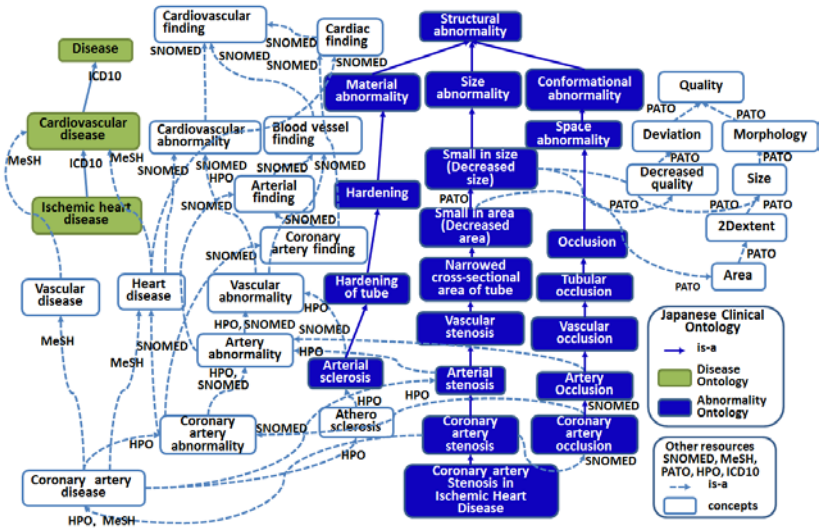

Fig. 1 *Is-a* hierarchy of coronary artery stenosis.

With respect to the intrinsic nature of abnormal states, states can change, and their essential characteristic is understood with respect to how the corresponding attribute (A) has the value (V), which can change as time goes. Then, one can find that the authentic "*is-a*" relation should be "decreased area *is-a* decreased size" rather than "decreased area *is-a* area", and our model adopts the former.

Next, we discuss ontological issues in HPO. The HPO hierarchy is based primarily on where the abnormal states occur (e.g., arterial stenosis *is-a* abnormalities of peripheral arteries). Therefore, it is difficult to know the *is-a* relation of "stenosis" comprehensively. In the HPO tree, *ad hoc* creation of correct *is-a* relationships is found, such as "renal artery stenosis *is-a* arterial stenosis." To make matters worse, HPO has no generic concepts at the upper level nor any information about what attribute takes what value, and we cannot capture the essential property of the state as "stenosis." Furthermore, specialization using the part-whole relation appears in the *is-a* hierarchy, e.g., <abnormality of the vasculature *is-a* abnormality of the cardiovascular system>, which is misleading.

With respect to abnormal states, it is important to capture the identity of the state by revealing "where and how its intrinsic attribute takes a particular abnormal value." On the basis of such understanding, our ontology adopts a single *is-a* relationship to inherit the intrinsic nature of abnormal states. For example, all lower classes of "small in area" inherit the property of "the area (A) is small (V)," such as" narrowed cross-sectional area of tube," "vascular stenosis," and so on.

In practice, however, only a single hierarchical structure seems incompatible with various perspectives in biomedicine. In order to support different perspectives, we developed a technology for dynamically generating an on-demand classification hierarchy [13]. For instance, if researchers want to know the classification of abnormal states in terms of anatomical structure, by using the *is-a* relation of Object (O) for each abnormal state, a subclassification according to anatomical structure abnormalities can be generated (Fig. 2). The partonomy of organs can also be used for generating a hierarchy, which is similar to the part-of-whole relationship of anatomical structure: e.g., "atrium abnormality" is classified as a subcategory of "cardiac abnormality."

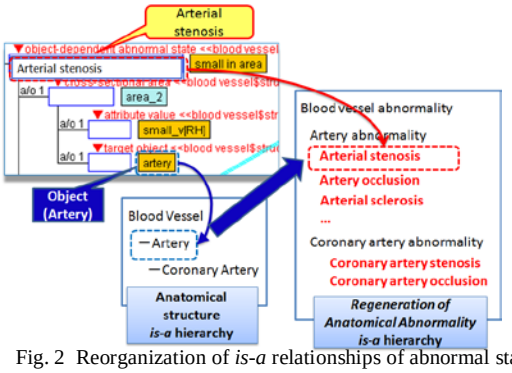

Fig. 2 Reorganization of *is-a* relationships of abnormal states in terms of anatomical structure.

Another critical issue with HPO is that HPO does not differentiate abnormal states from diseases and organize them in one hierarchy, which may be misleading and confusing in clinical practice. In HPO, the parent of "coronary stenosis" is considered to be not an abnormal state but a disease, namely, "coronary disease." However, as shown in Fig. 1, "coronary disease" has three parents: "arterial stenosis," "coronary artery abnormality", and "atherosclerosis," which seem to be abnormal states. Therefore, there is no guarantee that a computer can derive the correct answer to the question of whether "coronary artery stenosis" is a disease or an abnormal state.

Our ontology enables us to make a distinction between generic abnormal states, object-dependent ones, and disease-specific ones with a unified representation and allows us to specialize concepts to the required granularity with consistency. We conceptualize a disease as an entity represented in terms of abnormal states, and we deal with abnormal states and diseases as different entities. Therefore, we can focus on the intrinsic nature of the states themselves from one viewpoint and can develop an ontology from the viewpoint of state, without mixing up the viewpoint of disease.

2) SNOMED-CT and MeSH: The development of SNOMED-CT and MeSH started before ontological engineering had become mature, and thus they have some problems ontologically [14]. SNOMED-CT and MeSH do not have any formal upper ontology. This allows for multiple inheritance that results in complicated relations. Such a structure may lead to inconsistencies in daily clinical practice, since the organization of concepts is not principled, and hence relationships among concepts lack consistency. As shown in Fig. 1, in SNOMED and MeSH, abnormal states and diseases are mixed up in the hierarchy, as is the case in HPO. In MeSH, the upper class "coronary stenosis" is a disease, namely, "coronary disease," and at an even higher level, diseases are defined, like "heart disease" and "cardiovascular disease." In SNOMED, the upper concept of "coronary artery stenosis" is a "coronary occlusion," and furthermore, at even higher levels, there are various concepts, such as "heart disease," "vascular (blood vessel) finding," "disorder of cardiovascular system," "disorder of soft tissue" and so on.

SNOMED-CT terms have been used for electronic health record (EHR) systems. Because SNOMED-CT does not differentiate abnormal states from diseases, serious clinical problems may occur. For example, imagine a case where a clinician examines an angiographic image and, by using the SNOMED-CT term "coronary artery stenosis," records it for evidence of "abnormal states" in the EHR system. Due to inheriting properties from one of the upper concepts in SNOMED-CT, namely, "disease," there is a possibility of deriving an erroneous consequence that the "abnormal state" is a "disease." Our ontology makes a distinction between diseases, disease-dependent abnormal states, and disease-independent abnormal states. Therefore, there is no possibility of deriving inappropriate consequences, which is important for computer processing. In order to manage high-level clinical knowledge in EHR, we need a reliable method for representation, and our ontological model provides us with high reliability and sophisticated technology for realizing interoperability between heterogeneous pieces of clinical information.

3) LOINC: LOINC was excluded from the above-mentioned comparative research, because LOINC does not deal with concepts related to abnormal states.

LOINC provides the form O (So) A and is useful for interoperability among various clinical test data. However, LOINC does not have Value (V). To realize interoperability between clinical test data and abnormal states, a Quantitative Value (Vqt) is needed in the representation form. Our model can deal with quantitative data in the OAV form, and, therefore, we can transform it into the OP form of abnormal states. As a result, our model has the ability to maintain interoperability between clinical test data and abnormal states in diseases.

Some readers might think, "Why not reuse other resources such as clinical terminology or existing ontologies?", because it seems that by combining existing resources, many abnormal states would be easily covered. Such good candidates for reuse concerning abnormal states are PATO, HPO, LOINC, and

SNOMED-CT. As we illustrated in the problem above, just reutilizing existing resources cannot integrate all knowledge about abnormal states. It is too difficult to find all inappropriate usages of *is-a* and resulting misclassifications from the huge and heterogeneous system of concepts and to make an effort to modify them. Furthermore, since each resource has its own viewpoint, integrating them into a unified perspective must be a hard task because it necessarily requires comparative analysis and validation of accuracy in integrated concepts.

Unified theoretical considerations have resulted in interoperability between various representation forms, which will enable us to establish a computer-understandable model for abnormal states. We need more sophisticated organization of related representations, including quantitative and qualitative data and knowledge at higher levels of abstraction about abnormal states in the definition of diseases to exploit all of them in a consistent manner. Our model is the first one to make such exploitation possible, and will be of great assistance in medical practicee.

The differences of resources are summarized in Table 1.

TABLE I. COMPARISON OF BIOMEDICAL RESOURCES

Name	LOINC	SNOMED-CT	MeSH	PATO	HP	Our Ontology
Contents	Data code system for clinical test	Controlled vocabulary for clinical terms	Controlled vocabulary for medical literature	Phenotype ontology	Human Phenotype ontology	Abnormality ontology
Number of Concepts	46,000	310,000	27,000	2000	10,000	21,000
is-a tree	N	Y	Y	Y	Y	Y
Class inheritance	N	multiple	multiple	multiple	multiple	single
Generic abnormal states	N	N	N	Y	N	Y
Anatomical-dependent abnormal states	N	Y	Y	N	Y	Y
Disease-dependent abnormal states	N	Y	N	N	N	Y
Upper ontology	N	N	N	BFO	BFO	YAMATO
OP/ OAV formalism	O (So) A	N	N	OP	OP	OPV O (So)AV
Open source	Y	N	Y	Y	Y	Plan to open

C. *Integration of biomedical abnormal states*

1) Integration of abnormal states

As illustrated in Section 2, our ontology provides three levels of abnormal states from generic to disease-specific.

Level 1 in our ontology defines generic concepts corresponding to the PATO concepts, and our Level 1 concepts can be mapped to related PATO concepts (Fig. 3). The lower Level 2 concepts are human anatomical structure-dependent abnormal states, which correspond to the HPO concepts. By creating links between Level 2 concepts and HPO concepts, it will be possible to navigate from the HPO concepts to the upper generic concepts of PATO. Level 3 provides disease-specific abnormal states, such as "myocardial ischemia in ischemic heart disease," "chest pain in angina pectoris", and so on. In the revised version 11 of the International Classification of Diseases (ICD), diseases contain information of "causal properties" [15], and therefore, we are planning to map our Level 3 concepts to the corresponding concepts in the ICD. Level 3 abnormal states are described in the causal chains of diseases [1]. By mapping our disease concepts of disease ontology to the ICD, ICD users can understand the causal relationships of the abnormal states in diseases. Our ontology can also allow users to navigate related concepts in other resources, such as HPO, PATO, etc.

Mapping our ontology to other resources for integrating various data related to abnormalities will bring benefits to the users of other resources, too. First, one can find concepts from generic to specialized terms easily by referring to the single *is-a* tree in our abnormality ontology. For example, although HPO does not care about consistent *is-a* relationships in terms of "stenosis", by referring to "arterial stenosis" at Level 2 in our ontology through mapping, HPO users can get the *is-a* relationships: "arterial stenosis *is-a* vascular stenosis *is-a* narrowed cross-sectional area of tube *is-a* small in area." Since "small area" is linked to a PATO concept, via our ontology, users might find orthologous concepts of other species. Specifically, human phenotypes can be linked to the phenotypes of model organisms, e.g., mouse, rat, etc., if the set composed of Attribute (A) and Value (V) are identical, and the Object (O) has structural similarity. PATO2YAMATO aims to integrate phenotype descriptions residing in differently structured comparison contexts [16]. By applying PATO2YAMATO, mapping of concepts across species and integrating knowledge from various species may be possible.

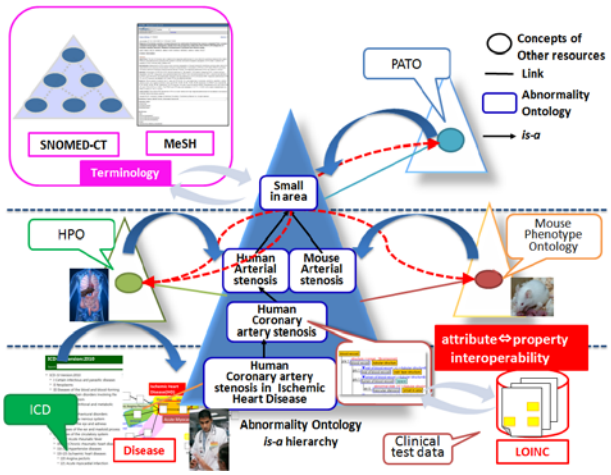

Fig. 3. Integration of abnormal states in biomedicine.

2) Integration of components of abnormal states

We are also planning to link the components of the <Object (O), Attribute (A), Value (V)> representation of abnormal states to other resources. We will try to connect the Object (O) to concepts in FMA (The Foundational Model of Anatomy ontology) [17].

3) Integration of biomedical articles

By mapping MeSH terms, it will be possible to retrieve biomedical articles related to abnormal states or diseases. If we obtain new findings of the constituents of diseases, we can add new relationships to the causal chains of the diseases, which might contribute to the elucidation of the etiological mechanism. Our mapping is also useful for MeSH users to understand how their research subjects are involved in various abnormal states in the human body in diseases. This will contribute to the progress of biomedical research.

D. *Concluding Remarks*

A large volume of data and concepts related to abnormal states is currently available in existing resources. However, there are no resources that cover all levels of abnormal states

from generic to disease-specific. We performed a comparison between our ontology and existing resources, and identified the issue that the heterogeneous levels of meanings in the different resources and multiple perspectives prevented us from reusing and integrating them. This motivated us to develop an abnormality ontology from the generic level to the disease-specific level.

Our medical ontology project started seven years ago. Since then, it has been refined and revised several times by discussion with both ontologists and clinicians. Our ontology will play a role in proper guidelines for giving an ontological point of view in various controlled vocabularies, and will lead to the development of a consistent hierarchical structure with a unified representation. Mapping our ontology with other resources at each level of meanings will contribute to ensuring interoperability across biomedical resources. Since our approach allows users to navigate from generic concepts to specific concepts in other domains by following links, it offers complementary information. We are currently applying the concepts of abnormal states in the definition of diseases in the Department of Cardiovascular Medicine and several other departments at The University of Tokyo Hospital in our ontology and mapping them to external biomedical resources, and this work is expected to be completed in the near future. Level 1 generic concepts have previously been developed by ontologists, and Level 3 disease-specific concepts were also described. Currently, we are developing and enriching Level 2 concepts to link each Level 3 concept to upper-level common concepts. By developing Level 2, we will be able to find more commonalities across diseases. For example, in cardiovascular medicine, "increased blood creatine kinase (CK) concentration in acute myocardial infarction" is defined by a clinician at Level 3. Next, "increased blood CK" at Level 2 is defined and mapped to "elevated serum creatine phosphokinase" in HPO. Then, the generic concept "increased concentration" is defined at Level 1 in our ontology, which is mapped to "increased concentration" in PATO. We can find commonalities with other concepts in the definition of diseases in other medical departments in our ontology. For example, "increased blood CK concentration in muscular dystrophy" used in the Neurology Department has commonality with "increased blood CK concentration," and moreover, "increased blood cholesterol concentration in hyperlipidemia" has commonality with the same generic concept, "increased concentration." We are planning to map the disease "acute myocardial infarction" in our disease ontology to "acute myocardial infarction" in the ICD. In our ontology, diseases are defined as causal relationships of abnormal states, and from this mapping, ICD users will be able to know the causal relationship, "decreased blood flow" causes "myocardial ischemia," which results in "myocardial necrosis" in acute myocardial infarction. This causal relationship is also available to users of HPO by following the linked concept of abnormal states. With respect to "myocardial ischemia," by mapping this to the MeSH term "myocardial ischemia," we can collect articles that are related to the concept.

In this way, our approach allows users to navigate various abnormality knowledge across domains, which will create a bridge between basic research and clinical medicine. We published the causal chains in our disease ontology in the form of Linked Data (LD)[18], and have made available a browsing system that links our data to DBPedia [19] and 3D images of related anatomical parts provided by BodyParts3D [20]. Next, we are planning to map our abnormal states to other resources using LD and to scale up to applications requiring more complicated knowledge.

ACKNOWLEDGMENT

A part of this research is supported by the Japan Society for the Promotion of Science (JSPS) through its “FIRST Program”" and the Ministry of Health, Labour and Welfare, Japan. The authors are deeply grateful to Drs. Ryota Sakurai, Natsuko Ohtomo, Aki Hayashi, Takayoshi Matsumura, Satomi Terada, Kayo Waki, and other, at The University of Tokyo Hospital for describing disease ontology and assisting us with their broad clinical knowledge. We also would like to thank other team members, Drs. Yoshimasa Kawazoe, Masayuki Kajino, and Emiko Shinohara from The University of Tokyo for useful discussions related to biomedicine.

REFERENCES

[1] R. Mizoguchi, K Kozaki, H Kou, Y Yamagata, T Imai, K Waki, K Ohe, “River flow model of diseases,” in ICBO2011, 2011, pp. 63-70.

[2] Y Yamagata, K Kou, K Kozaki, T Imai, K Ohe , R Mizoguchi, “Ontological model of abnormal states and its application in the medical domain,” in ICBO2013, 2013, pp. 28-33.

[3] Y Yamagata, K Kozaki, T Imai, K Ohe , R Mizoguchi, “An ontological modeling approach for abnormal states and its application in the medical domain.,” J Biomed Semantics, 2, 5:23, 2014.

[4] R. Mizoguchi, “YAMATO: Yet another more advanced top-level ontology.” In Proceedings of the Sixth AOW 2010, 2010, pp:1-16.

[5] SNOMED-CT [http://www.ihtsdo.org/snomed-ct/].

[6] MeSH [http://www.nlm.nih.gov/mesh/meshhome.html]

[7] GV Gkoutos, EC Green, AM Mallon A Blake, S Greenaway, JM Hancock, D Davidson, “Ontologies for the description of mouse phenotypes,” Comp Funct Genomics, vol.5, pp: 545-51, 2004.

[8] G Grumbling, V Strelets, “FlyBase: anatomical data, images and queries,” Nucleic Acids Res, vol. 34, pp: D484-488, 2006.

[9] S Köhler, SC Doelken, CJ Mungall, et al, “The Human Phenotype Ontology project: linking molecular biology and disease through phenotype data,” NAR, vol. 42 (Database issue), pp.D966-74. 2014.

[10] OMIM [http://www.ncbi.nlm.nih.gov/omim]

[11] LOINC [http://loinc.org/]

[12] Grenon P, Smith B, Goldberg L, “Biodynamic ontology: applying BFO in the biomedical domain,” Stud Health Technol Inform, vol. 102, pp. 20-38, 2004.

[13] K Kozaki et al., “Dynamic Is-a Hierarchy Generation System Based on User's Viewpoint,” in procceing of JIST, LNCS 7185, pp.96-111, 2012.

[14] S Schulz, B Suntisrivaraporn, F Baader, “SNOMED CT's problem list: Ontologists' and Logicians' therapy suggestions,” Stud Health Technol Inform. vol.129(Pt 1), pp:802-806, 2007;.

[15] ICD 11 [I http://www.who.int/classifications/icd/revision/en/]

[16] H Masuya, R Mizoguchi: “An advanced strategy for integration of biological measurement data ,” in ICBO2011, 2011, pp. 79-86.

[17] C Rosse, JVL Mejino. “A reference ontology for biomedical informatics: the Foundational Model of Anatomy,” J Biomed Inform, vol. 36, pp: 478-500., 2003.

[18] K Kozaki, Y Yamagata, T Imai, K, et al., “Publishing a disease ontologies as linked data,”. in Proc. of JIST2013, 2014, pp. 110-128.

[19] DB pedia [http://dbpedia.org/]

[20] BodyParts3D, The Database Center for Life Science [http://lifesciencedb.jp/bp3d/]