

The Ontology of Vaccine Adverse Events (OVAE) and its usage in representing and analyzing vaccine adverse events

Erica Marcos¹, Yongqun He^{2*}

¹ College of Literature, Science, and the Arts, University of Michigan, Ann Arbor, MI 48109, USA.

² Unit for Laboratory Animal Medicine, Department of Microbiology and Immunology, Center for Computational Medicine, and Comprehensive Cancer Center, University of Michigan, Ann Arbor, MI 48109, USA.

ABSTRACT

Licensed human vaccines can induce various adverse events in vaccinated patients. Many known vaccine adverse events (VAEs) have been recorded in the package inserts of commercial vaccine products. To better represent and analyse VAEs, we developed the Ontology of Vaccine Adverse Events (OVAE) as an extension of the Ontology of Adverse Events (OAE) and the Vaccine Ontology (VO). OVAE has been used to represent and classify the adverse events recorded in package insert documents of commercial vaccines licensed by the USA Food and Drug Administration (FDA). OVAE currently includes over 1100 terms, including 87 distinct types of VAEs associated with 63 human vaccines licensed in the USA. Specific VAE occurrence rates associated with different age groups have been recorded in OVAE. SPARQL scripts were developed to query and analyse the OVAE knowledge base data. The top 10 vaccines accompanying with the highest numbers of VAEs and the top 10 VAEs most frequently observed among vaccines were identified and analysed. Different VAE occurrences in different age groups were also analysed. The ontological representation and analysis of the VAE data associated with licensed human vaccines improves the classification and understanding of vaccine-specific VAEs which supports rational VAE prevention and treatment and benefits public health.

1 INTRODUCTION

Many licensed vaccines exist to protect against a variety of diseases and infections. They are extremely useful in decreasing infection prevalence in human populations. Due to the public health benefits of vaccines, their coverage has been increasing in recent years. Each vaccine often induces different types of adverse events. As vaccine usage increases, the risk of adverse events proportionally increases (Cunha, Dorea, Marques, & Leao, 2013). Many known vaccine adverse events (VAEs) have been recorded in the package inserts of commercial vaccine products. There is also a need to predict probabilities of different adverse events arising in different individuals, which can potentially lead to a decline in the risk of developing an adverse event.

Two existing ontologies are closely related to the VAE studies. The Ontology of Adverse Events (OAE) is a community-based biomedical ontology in the area of adverse events (He, Xiang, Sarntivijai, Toldo, & Ceusters, 2011; Sarntivijai et al., 2012). OAE defines an ‘adverse event’ as a pathological bodily process that occurs after a medical intervention (*e.g.*, vaccination, drug administration). The OAE ‘adverse event’ is a subclass of the ontol-

ogy term ‘pathological bodily process’ defined in the Ontology of General Medicine Science (OGMS) (<http://code.google.com/p/ogms/>). To be consistent with most practice uses of the term, OAE does not assume a causal relation between an ‘adverse event’ and a medical intervention. OAE has defined over 2,000 types of adverse events that are commonly found in different medical interventions. The community-based Vaccine Ontology (VO) represents various vaccines, vaccine components, and vaccinations (He et al., 2009; Lin & He, 2012). Both OAE and VO are OBO Foundry candidate ontologies and are developed by following the OBO Foundry principles (Smith et al., 2007).

OAE has been shown to significantly increase the power of analysis of case report data from the Vaccine Adverse Event Reporting System (VAERS) (Sarntivijai, et al., 2012). However, there has been no published paper to analyze commonly known VAEs recorded in the package insert documents of FDA licensed vaccines. Compared to the often noisy data creating difficulty in identifying the causality, the adverse events recorded in the official package inserts are known adverse events to specific vaccines.

To better represent various VAEs and support vaccine safety study, we developed the Ontology of Vaccine Adverse Events (OVAE) as an extension of the biomedical ontologies OAE and VO. In this paper, we introduce the basic framework of the OVAE and how OVAE is used to represent and analyze all adverse events reported in the product package inserts of commercial vaccines currently used in the USA market.

2 METHODS

2.1 OVAE ontology generation

Following VO and OAE, OVAE is also edited with the Web Ontology Language (OWL2) format (<http://www.w3.org/TR/owl-guide/>). FDA-licensed human vaccines represented in VO were imported to OVAE using the tool OntoFox (Xiang, Courtot, Brinkman, Ruttenberg, & He, 2010). Those adverse event terms reported in the package inserts of FDA licensed human vaccines were also imported to the OVAE using OntoFox.

* To whom correspondence should be addressed:
yongqunh@med.umich.edu

New OVAE-specific terms were generated with IDs containing the prefix “OGSF_” followed by seven auto-incremental digital numbers and edited using the Protégé 4.2 OWL ontology editor (<http://protege.stanford.edu/>).

2.2 Data source of known VAEs

The official FDA website that provides supporting documents of licensed vaccines was the primary data source (FDA, 2013). A PDF version of a package insert document is available for almost every vaccine in the data source. The PDF document includes a section called “Adverse Reactions” that contains text descriptions of known vaccine adverse events associated with the vaccinated population.

2.3 Data collection and formatting to ontology

Based on the OVAE framework and the adverse event description in the package inserts, a design pattern was first generated to lay out the relations between different ontology classes, properties, terms and data types. The design pattern was used to form an MS Excel template for collection of individual adverse events for different vaccines. All the data in each package insert were manually examined and input to the Excel worksheet. Following the manual data collection and annotation, the program Ontorat (<http://ontorat.hegroup.org>) was used to transform the Excel file data to the OVAE ontology format (Xiang, Lin, & He, 2012).

2.4 VAE data analysis

To identify specific OAE or VO hierarchical structure among a list of terms, OntoFox was first used to extract the input OAE or VO terms and all associated terms required for proper hierarchical assertion and inference. The output OWL files were then visualized using a Protégé OWL editor. The SPARQL Protocol and RDF Query Language (SPARQL) is a W3C recommended language to query OWL RDF triple store (“SPARQL query language for RDF,”). After OVAE was also deposited in the Hegrup RDF triple store, SPARQL was used for querying the OVAE knowledgebase from the RDF triple store to address a list of scientific questions.

2.5 OVAE ontology websites and license

The OVAE source code is available in a Google Code website: <http://code.google.com/p/ovae>. The OVAE project website is: <http://www.violinet.org/ovae>. OVAE has been deposited to the BioPortal project of the National Center of Biomedical Ontology (NCBO) (<http://bioportal.bioontology.org/ontologies/3227>). OVAE is also deposited in the Ontobee linked data server (<http://www.ontobee.org/browser/index.php?o=OVAE>) (Xiang, Mungall, Ruttenberg, & He, 2011). The OVAE source is freely available under the Apache License 2.0.

3 RESULTS

3.1 OVAE system design and statistics

The goal of current OVAE development is to generate an ontology-based VAE knowledgebase that represents known adverse events (AEs) associated with licensed vaccines. Such a knowledgebase incorporates the OAE terms of AEs together with the vaccine information defined in the VO. As the primary developer of the OAE and VO, we argue that OAE is not appropriate or responsible for representing various AEs specific for any particular medical intervention including vaccination due to the following reasons. First, OAE emphasizes the representation of various AEs general for most medical interventions and related topics (*e.g.*, methods for analysis of the causal relation between AEs and medical interventions, and factors affecting the causality analysis). Currently OAE is already large and contains over 3,000 terms. It is expected that many more AE terms will be added to OAE. Therefore, it is ideal to make OAE focused and as concise as possible. Secondly, AE researchers related to specific medical intervention domains may have more domain-specific demands and requests. For example, VAE researchers would like to link AEs to different vaccines. The vaccine (or drug) researchers may not be interested in drug (or vaccine) specific AEs. As a relatively independent domain, VAEs have been focuses of many vaccine researchers and groups. Independent from drug AEs, clinical VAEs are reported to vaccine-specific VAERS system in the USA (Varricchio et al., 2004). Meanwhile, the Vaccine Ontology (VO) is not suitable for representing complex VAE data. VO has been focused on classification of various vaccines, including licensed vaccines, vaccines in clinical trials, and vaccines only verified in laboratory animal models. VO also represents various types of vaccine components (*e.g.*, vaccine antigens, adjuvants, and vectors), vaccine attributes (*e.g.*, vaccine organism viability and virulence), vaccination methods, and other concise and closely related vaccine information. The inclusion of complex VAE information to VO would make VO imbalanced and lack of focus. Due to these reasons, we generated the VAE-specific OVAE. Since both OAE and VO use the Basic Formal Ontology (BFO) (<http://www.ifomis.org/bfo>) as the top level class, the alignments between OVAE, OAE, and VO are easy and straightforward.

As an extension of OAE and VO, OVAE targets for not only importing related terms from these two ontologies but also including many OVAE-specific terms. The primary data source for generating vaccine-specific AE ontology terms in current OVAE is the official vaccine package inserts available in the USA FDA website (FDA, 2013). Each official vaccine package insert document provided by the USA FDA includes a section called

“Adverse Reactions”. The results provided in the section were obtained from carefully designed clinical trials with randomized controls and worldwide post-marketing experience. Therefore, the VAE information provides basic known VAEs that are likely to occur after an administration of a specific vaccine in a human vaccinee. Based on the officially documented information, OVAE includes many OVAE-specific terms, for example, ‘Afluria-associated pain AE’ to define a pain AE specific for Afluria-vaccinated patients. As shown in detail later in the paper, the generation of these new terms allows the inclusion of more detailed information about these VAEs, for example, the VAE occurrences in human vaccinee populations in different age groups.

Table 1 lists the OVAE statistics as of May 1, 2013. OVAE used the most recent BFO 2.0 Graz version (<http://purl.obolibrary.org/obo/bfo.owl>) as the top level ontology. Since BFO 2.0 is not yet finalized, some relation terms (e.g., ‘part of’ or BFO_0000050) are still used in OVAE but do not necessarily comply with the most recent BFO 2.0. During the process of importing many AE or vaccine-related terms from OAE and VO to OVAE, many terms from other existing ontologies, including OGMS, Ontology for Biomedical Investigation (OBI) (Brinkman et al., 2010), Phenotypic Quality Ontology (PATO) (“PATO - Phenotypic Quality Ontology,”), and Information Artifact Ontology (IAO) (<http://code.google.com/p/information-artifact-ontology/>), have also been imported to OVAE (Table 1). To maintain the ontology asserted and inferred hierarchies and support intact reasoning capability, the OntoFox software was used for external term importing (Xiang, et al., 2010). In summary, OVAE includes 1,199 terms that contains 652 OVAE specific terms (with “OVAE_” prefix). In addition, OVAE including all 113 terms from the BFO version 2.0, 315 VO terms, 105 OAE terms, 3 OBI terms, 3 IAO terms, and 2 OGMS terms (Table 1). By referencing the vaccine package insert data, OVAE represents 87 distinct AEs associated with 63 licensed human vaccines.

Table 1. Summary of ontology terms in OVAE.

Ontology Names	Classes	Object		Total
		properties	properties	
OVAE	650	1	1	652
BFO	35	78	0	113
OBI	2	1	0	3
PATO	7	0	0	7
IAO	3	0	0	3
OAE	105	0	0	105
OGMS	2	0	0	2
VO	307	5	3	315
Total	1110	85	4	1199

3.2 OVAE design pattern of representing VAE

The general design pattern of representing a VAE in OVAE is shown in Fig. 1. Specifically, a licensed vaccine, manufactured by a company and having specific quality (e.g., using inactivated vaccine organism), is targeted to immunize a human vaccinee against infection of a microbial pathogen. A particular vaccination route (e.g., intramuscular route) is specified. A specific VAE (e.g., Afluria-associated injection-site pain adverse event) occurs in a human vaccinee and after (*preceded_by*) a vaccination. The human vaccinee, having a specific age (defined via a datatype) at the time of vaccination, is part of the population of human vaccinees using this vaccine. The VAE occurrence is defined as a frequency of an adverse event associated with the administration of a vaccine in a vaccinee population. The new object property term ‘has VAE occurrence’ is defined in OVAE to specify a VAE occurrence (xsd:decimal datatype) in a human vaccinee population that has been individually vaccinated with a specific vaccine during a specific time period. To simplify the representation of axioms linking vaccine adverse event and human vaccinee population, OVAE generates a shortcut relation ‘occurs in population’ (Fig. 1).

The vaccine attributes and vaccination details are imported from VO. Their inclusion in the design pattern is due to their possible contribution to the VAE determination. For example, a live attenuated vaccine and a killed inactivated vaccine may in general induce different types or levels of VAEs, which can be analyzed by statistical analysis (Sarntivijai, et al., 2012).

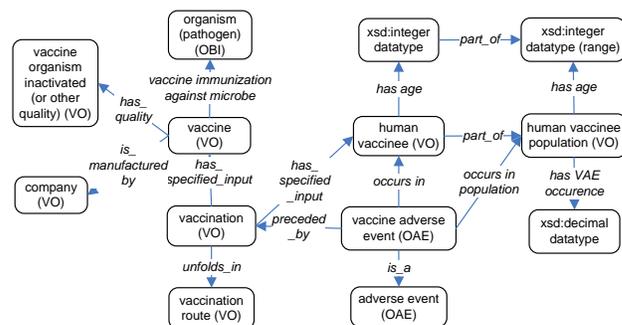


Fig. 1. OVAE design pattern of a human vaccine adverse event.

One novelty in the design pattern is the generation and application of the population term ‘human vaccinee population’ to define a VAE occurrence. In previous versions of OAE and VO, only ‘vaccinee’ and ‘human vaccinee’ (i.e., a human being administered with a vaccine) exist. However, it is incorrect to say that a specific human vaccinee has a VAE occurrence of some percentage (e.g., 10%). An occurrence is defined only for a population. The generation of the term ‘human vaccinee population’ solves the ontology modeling issue. Any particular human vaccinee is part of a human vaccinee population.

(Table 2). It is noted that the information does not dictate the severity of AEs associated with each vaccine, but instead indicates those vaccines that are licensed for human use in the USA and display the most variation in their reported AEs.

Table 2. Top 10 vaccines with the largest variety of VAE reported.

Vaccine Name	VO_ID	Total # VAE
Recombivax HB (Hepatitis B)	VO_0010737	23
Menactra (Meningitis)	VO_0000071	21
Comvax (Meningitis, Hepatitis A)	VO_0000028	19
Prevnar (<i>Streptococcus pneumoniae</i>)	VO_0000090	19
Tetanus and Diphtheria Toxoids Absorbed by MA Biological (Tetanus, Diphtheria)	VO_0000111	18
Fluarix (Influenza)	VO_0000045	15
Fluarix Quadrivalent (Influenza)	VO_0000983	15
PedvaxHIB (Meningitis)	VO_0000083	15
RabAvert (Rabies)	VO_0000094	14
Boostrix (Tetanus, Diphtheria, Pertussis)	VO_0000015	14

Note: The disease or infection being protected against is specified next to each vaccine name. The vaccines are sorted based on the VAEs recorded in their package insert documents.

Secondly, we evaluated the top VAEs that have been reported most frequently among all vaccines licensed in the USA and represented by OVAE (Table 3). Most of the top 10 frequently observed VAEs are expected, such as injection-site pain and redness, fever, and local swelling. The headache and myalgia (*i.e.*, muscle pain) AEs demonstrate two types of pain. Similar to different pains, malaise (*i.e.*, uneasiness and discomfort) and fatigue are sensory AEs. It is noted that the information does not dictate which VAEs are the most severe, but indicates which VAEs are commonly observed in currently licensed vaccines in the USA.

Table 3. Top 10 most frequently reported VAEs

AE Name	OAE_ID	Total # vaccines	%
Injection-site pain AE	OAE_0000369	43	68.3
Headache AE	OAE_0000377	39	61.9
Fever AE	OAE_0000361	34	54.0
Local swelling AE	OAE_0001139	30	47.6
Injection-site redness AE	OAE_0001546	25	40.7
Irritability AE	OAE_0001105	23	36.5
Malaise AE	OAE_0000390	21	33.3
Injection-site erythema AE	OAE_0000644	20	31.7
Myalgia AE	OAE_0000375	19	30.2
Fatigue AE	OAE_0000034	18	28.6

To better understand the top VAEs associated with licensed human vaccines, the hierarchical structure of the top

10 VAEs (Table 3) was extracted using the tool OntoFox and visualized using Protégé ontology editor (Fig. 3). The hierarchical visualization indicates that most of the top ranked VAEs belong to the behavior and neurological AE branch.

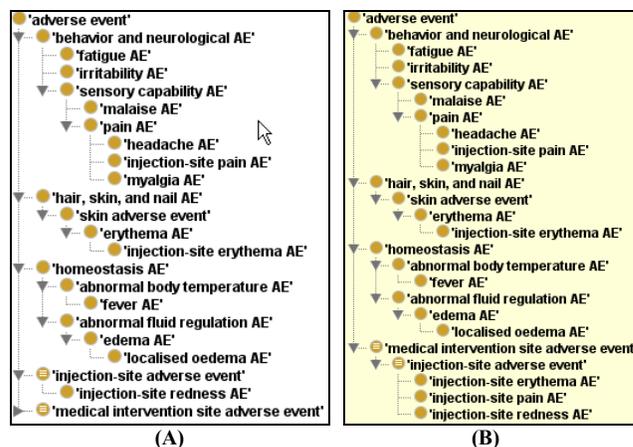


Fig. 3. Classification of top 10 AEs associated with licensed human vaccines in the US. These OAE terms have been imported to OVAE using OntoFox and visualized using Protégé OWL editor. (A) Asserted hierarchy in OAE; (B) Inferred hierarchy after reasoning.

Lastly, we compared the VAEs and VAE occurrences under different age groups. As shown in Fig. 2, the OVAE clearly represents the associations between VAEs, the VAE occurrence rates, and different ages (in years) of human vaccinee population. Our analysis can further identify which age category has a higher probability of experiencing any specific adverse events. For example, we found that *Salmonella typhi* vaccine Typhim Vi is associated with injection-site tenderness adverse events with the highest rate of 97.5% at the age group of 18-40 years old. Based on the classification of “child”, “adult”, and “child-adult” described in Section 3.3, there are 240, 160, and 177 specific VAEs in the age categories “child”, “adult”, and “child-adult”, respectively. It is also found that in general the VAE occurrences shown in the children are typically higher than those in adults. This suggests that individuals under 18 years may be more likely to experience an adverse reaction after vaccination.

4 DISCUSSION

The development of OVAE is aimed to align and reuse existing ontologies OAE and VO, and systematically represent and analyze vaccine-specific adverse events (VAEs). As demonstrated in this report, such a strategy has many advantages. First, as shown in Fig. 2, the ontological classification is easy for humans to interpret and analyze. A human can browse the hierarchical tree to quickly understand which VAEs are typically associated with a licensed vaccine. Secondly, the ontology OWL representation is also interpretable by computers and software programs. New

programs can be developed to parse and analyze the information. Thirdly, the approach of aligning OVAE with existing ontologies allows efficient integration of data presented in other ontologies (e.g., VO). Such a seamless integration makes it possible to analyze VAEs with other tools such as VO-based literature mining (Ozgun, Xiang, Radev, & He, 2011). In addition to the VAEs associated with USA licensed vaccines, the OVAE can be used to represent VAEs associated with vaccines licensed in other countries.

It is also possible to apply the OVAE framework to analyze clinical VAE data such as those case reports stored in VAERS (Varricchio, et al., 2004). For example, by comparing the reported vaccine-specific VAE cases in VAERS with the VAE occurrences reported in the package inserts and OVAE, it is easy to differentiate known VAEs and possibly new VAEs associated with the vaccine. Many differences exist in terms of the data shown in the package inserts and in VAERS database. While the data in the package inserts were typically obtained from well controlled clinical trials, clinical VAE case reports stored in VAERS came from random reports from physicians, patients, patients' parents, or other sources. The VAERS database does not indicate the total number of vaccinated human vaccinees in any given period, making it impossible to calculate exact VAE occurrences as reported in the package inserts and OVAE. However, ontological approach, together with statistical analysis, is still useful in VAERS data analysis as previously demonstrated (Sarntivijai, et al., 2012). One future research direction will be to identify novel ways to analyze VAE clinical data using OVAE.

While many AEs are common, different vaccines are associated with different AEs with various molecular mechanisms. The ontology representation of vaccine-specific AEs is a first step towards refined deep understanding of vaccine adverse events. It is also noted that the method of establishing vaccine-specific OAE extension may likely be applied for developing OAE extensions in other specified domains such as drug-associated adverse events.

ACKNOWLEDGEMENTS

This project was supported by a NIH-NIAID grant (R01AI081062). We also appreciate the work by Bin Zhao who helped clean up the ontology.

REFERENCES

Brinkman, R. R., Courtot, M., Derom, D., Fostel, J., He, Y., Lord, P., . . . Consortium, T. O. (2010). Modeling biomedical experimental processes with OBI. *Journal of Biomedical Semantics*, June 22; 21 Suppl 21:S27. doi: 10.1186/2041-1480-1181-S1181-S1187.

Cunha, M. P., Dorea, J. G., Marques, R. C., & Leao, R. S. (2013). Vaccine Adverse Events Reported during the First Ten Years (1998-2008) after Introduction in the State of Rondonia, Brazil. *Biomed Res Int*, 2013, 853083. doi: 10.1155/2013/853083

FDA, The United States (U.S.). (2013). U.S. Food and Drug Administration. Vaccines Licensed for Immunization and Distribution in the US with Supporting Documents. URL: <http://www.fda.gov/BiologicsBloodVaccines/Vaccines/ApprovedProducts/UCM093830.htm>, accessed on April 3, 2013

He, Y., Cowell, L., Diehl, A. D., Mobley, H. L., Peters, B., Ruttenberg, A., . . . Smith, B. (2009, July 24-26). VO: Vaccine Ontology. Paper presented at the The 1st International Conference on Biomedical Ontology (ICBO 2009), Buffalo, NY, USA. Nature Precedings: <http://precedings.nature.com/documents/3552/version/1>.

He, Y., Xiang, Z., Sarntivijai, S., Toldo, L., & Ceusters, W. (2011). AEO: a realism-based biomedical ontology for the representation of adverse events. Adverse Event Representation Workshop, International Conference on Biomedical Ontologies (ICBO), University at Buffalo, NY, July 26-30, 2011. . *Proceeding of ICBO-2011*, 309 - 315.

Lin, Y., & He, Y. (2012). Ontology representation and analysis of vaccine formulation and administration and their effects on vaccine immune responses. *J Biomed Semantics*, 3(1), 17. doi: 2041-1480-3-17 [pii]10.1186/2041-1480-3-17

Ozgun, A., Xiang, Z., Radev, D. R., & He, Y. (2011). Mining of vaccine-associated IFN-gamma gene interaction networks using the Vaccine Ontology. *J Biomed Semantics*, 2 Suppl 2, S8. doi: 10.1186/2041-1480-2-S2-S82041-1480-2-S2-S8 [pii]

PATO - Phenotypic Quality Ontology. from http://obofoundry.org/wiki/index.php/PATO:Main_Page

Sarntivijai, S., Xiang, Z., Shedden, K. A., Markel, H., Omenn, G. S., Athey, B. D., & He, Y. (2012). Ontology-based combinatorial comparative analysis of adverse events associated with killed and live influenza vaccines. *PLoS One*, 7(11), e49941. doi: 10.1371/journal.pone.0049941PONE-D-12-19530 [pii]

Smith, B., Ashburner, M., Rosse, C., Bard, J., Bug, W., Ceusters, W., . . . Lewis, S. (2007). The OBO Foundry: coordinated evolution of ontologies to support biomedical data integration. *Nat Biotechnol*, 25(11), 1251-1255. doi: nbt1346 [pii]10.1038/nbt1346

SPARQL query language for RDF. from <http://www.w3.org/TR/rdf-sparql-query/>

Varricchio, F., Iskander, J., Destefano, F., Ball, R., Pless, R., Braun, M. M., & Chen, R. T. (2004). Understanding vaccine safety information from the Vaccine Adverse Event Reporting System. *Pediatr Infect Dis J*, 23(4), 287-294. doi: 00006454-200404000-00002 [pii]

Xiang, Z., Courtot, M., Brinkman, R. R., Ruttenberg, A., & He, Y. (2010). OntoFox: web-based support for ontology reuse. *BMC Res Notes*, 3, 175. doi: 1756-0500-3-175 [pii]10.1186/1756-0500-3-175

Xiang, Z., Lin, Y., & He, Y. (2012). Ontorat web server for automatic generation and annotations of new ontology terms. . *Proceedings of the International Conference on Biomedical Ontologies (ICBO), University of Graz, Graz, Austria, July 24-27, 2012.*, URL: http://ceur-ws.org/Vol-897/poster_812.pdf

Xiang, Z., Mungall, C., Ruttenberg, A., & He, Y. (2011). Ontobee: A Linked Data Server and Browser for Ontology Terms. *Proceedings of the International Conference on Biomedical Ontologies (ICBO), University at Buffalo, NY, July 26-30.* , 279-281.