AEO: a realism-based biomedical ontology for the representation of adverse events

Yongqun He[1], Zuoshuang Xiang[1], Sirarat Sarntivijai[1], Luca Toldo[2], Werner Ceusters[3]

[1] University of Michigan Medical School, Ann Arbor, MI 48105, USA. [2] Merck KGaA, 250, Frankfurterstrasse, 64293 Darmstadt, Germany, [3] State University of New York at Buffalo, NY, USA

Abstract. The Adverse Event Ontology (AEO) is a realism-based biomedical ontology for adverse events. Currently AEO has 384 representational units annotated by means of terms including 369 AEO-specific terms and 115 terms from existing feeder-ontologies. In AEO, an adverse event is defined as a pathological bodily process that is induced by a medical intervention. This requirement for causal association between an adverse event and a medical intervention clearly distinguishes our approach from other approaches according to which any untoward phenomenon observed to have appeared in a mere temporal relation with some medical intervention becomes reported as an 'adverse event'. We label such phenomena as being the subject of 'adverse event hypotheses'.

Keywords: Adverse event, AEO, Adverse Event Ontology

1 Introduction

While medical interventions such as drug administrations, vaccinations, use of medical devices, and uptakes of nutritional products (e.g., infant formulae) are applied with the goal of producing positive effects, they might induce un-wanted adverse reactions (i.e. side effects). An ideal medical intervention should have high efficacy and minimal unwanted reactions, using a minimal dosis in case of substance administrations. It is however well known that any substance (even water) might give rise to un-wanted adverse reactions, if administered at the wrong dosis.

Adverse event related morbidity and mortality have become a major public health issue. To better organize adverse event information, different sorts of systems such as COSTAR (http://purl.bioontology.org/ontology/CST), MedDRA (http://purl.bioontology.org/ontology/MDR), the Common Terminology Criteria for Adverse Events (CTCAE; http://purl.bioontology.org/ontology/CTCAE), and the WHO Adverse Reaction Terminology (WHO; http://purl.bioontology.org/ontology/WHO) have been developed. These systems are typically constructed as controlled vocabularies, terminologies or classification systems and differ from biomedical ontologies most of which are consensus-based controlled vocabularies of terms and relations with associated definitions, which are

logically formulated to promote automated reasoning. Bosquet et al., for instance, have shown that terminological reasoning improves the performance of both data mining [1] and data access [2] in pharmacovigilance databases, and have done preliminary work toward the proposal of a categorial structure for adverse drug reactions (ADRs) [3]. However, although logically formulated definitions and axioms have the capacity to produce *valid* reasoning in deductive logic-based reasoning systems, they do not guarantee *sound* reasoning. Typical for prevailing paradigms in biomedical ontology design is concept-orientation which lacks a formal method to relate representational units to that in reality about which they are representations and these representations are therefore more vulnerable for mistakes that lead to unsound reasoning [4].

The Adverse Event Ontology (AEO), in contrast, is a realism-based effort to formally define adverse event and its associated terms using the framework of the OBO Foundry ontology design [5]. In this report, we present our current development of AEO, thereby distinguishing it from another recent effort to generate an Adverse Event Reporting Ontology (AERO).

2 Methods

2.1 AEO development principles

The development of AEO follows the OBO Foundry principles such as openness, collaboration, and use of a common shared syntax [5] in addition to the principles of Ontological Realism [6]. AEO is thus aligned with the Basic Formal Ontology (BFO) [7] and the Relation Ontology (RO) [8]. The AEO development follows many guidelines provided by Ceusters et al. [9] in generating ontological representations of adverse events on the basis of inspecting the sorts of particulars that are involved when an adverse event comes into existence.

2.2 AEO editing and access

The Web Ontology Language (OWL) is used as AEO's representation language. AEO is edited using Protégé 4 Ontology Editor (http://protege.stanford.edu). For ontology reuse, OntoFox (http://ontofox.hegroup.org/) [10] was used to extract ontology terms from external ontologies and import into AEO. New AEO unique identifiers were generated for those adverse event-specific terms.

The latest AEO, although not completely curated in terms of the principles mentioned earlier, is available for public view and download at http://sourceforge.net/projects/aeo/. AEO has been submitted to NCBO BioPortal for public visualization and querying: http://purl.bioontology.org/ontology/AEO.

3 Results

3.1 AEO statistics

Currently AEO has 484 representational units, annotated by means of 369 terms with specific AEO identifiers, and 115 terms imported from existing ontologies (Table 1). This ontology development design avoids regeneration of new ontology terms that are not in the scope of the adverse event domain and supports efficient ontology reuse on the condition that the feeder ontologies are based on the same principles. Existing ontologies are used in two different ways in AEO: one is to import the whole ontology (here BFO and RO), and the other is to import individual terms from existing ontologies. The OntoFox method is a newly developed approach to make individual term importing easy and standardized [10], although additional steps are required to make sure that the definitions for these terms in the feeder-ontologies correspond to the intended referents in AEO.

Table 1. Summary of ontology terms in AEO or imported from existing ontologies.

Ontology Names	Classes	Object properties	Total
AEO (Adverse Event Ontology)	368	1	369
BFO (Basic Formal Ontology)	39	0	39
RO (Relation Ontology)	6	25	31
IAO (Information Artifact Ontology)	2	0	2
OBI (Ontology for Biomedical Investigations)	8	3	11
OGMS (Ontology for General Medical Science)	5	0	5
VO (Vaccine Ontology)	19	3	22
NCBITaxon (NCBI Taxonomy)	5	0	5
Total	452	32	484

Fig. 1 lists key terms in AEO. Based on the *adverse event* definition, AEO required the term *medical intervention*, which currently includes four subclasses: *vaccination* (imported from VO), *drug administration*, *medical device usage*, and *nutritional product usage*. Each of these medical interventions can induce corresponding adverse events, e.g., *vaccine adverse event* (Fig. 1). Each adverse event may have different outcomes. For example, a symptom (e.g., rash) is a common outcome of an adverse event.

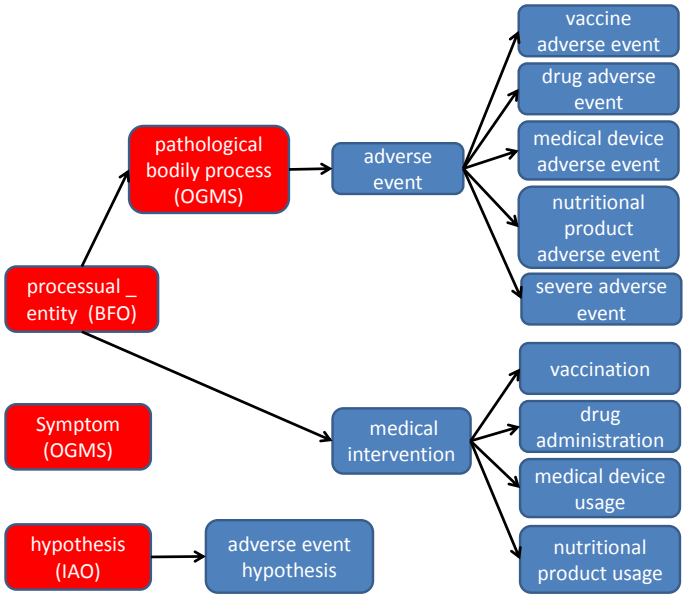

Fig. 1. Key terms in AEO.

3.2 Logical definition of 'adverse event' in AEO

The term 'adverse event' may mean different things in different settings [9]. In AEO, an adverse event is defined as "a pathological bodily process that is induced by a medical intervention." As defined in OGMS (http://code.google.com/p/ogms/), a pathological bodily process (OGMS_0000060) is a bodily process that is clinically abnormal. This definition fits well with adverse event and thus is chosen as the parent term of *adverse event* in AEO.

The "induced" in the AEO 'adverse event' definition indicates the existence of a causal chain. A medical intervention is a process in which several independent continuants (e.g., anatomical parts of human body) participate in a variety of ways and of which other processes are parts in which these or other independent continuants participate. Some independent continuants existed already before the intervention started (e.g. cells and molecules of the patient), others are created (e.g. molecular complexes formed by bodily molecules and drugs) or modified (e.g. opening and closing of membrane channels, folding of proteins) through processes that are part of the intervention or bodily processes that come into existence in response to the creation or modification of these continuants. After the intervention, there are still bodily processes going on in which at least one of the independent continuants just mentioned participates and further independent continuants are created. The term "induced" means that there is at least one chain of processes that starts with some process that is part of the intervention and ends with a pathological bodily process, the chain being further such that for each process within it (except the first one) there is at least one independent continuant that participated or was created in the process immediately preceding it. Note that we are not saying that there is one

such independent continuant that participates in the entire chain, but rather something like this:

P1: C1, C2, C3
P2: C2, C4, C5
P3: C5, C6, …

Just temporal precedence is not enough because that would allow for chains of processes in which there is a pair that does not "share" at least one continuant.

An alternative definition for 'adverse event' would be to assign it as a child term of ogms:sign, which has the textual definition of "A quality of a patient, a material entity that is part of a patient, or a processual entity that a patient participates in, any one of which is observed in a physical examination and is deemed by the clinician to be of clinical significance." Although this appears to cover different adverse events, this ogms:sign definition is too broad since all adverse events are processes. At the same time, it is too narrow because there are adverse events that are not observed. The definition of sign clearly states "is observed in a physical examination", instead of "CAN BE observed".

3.3 Key entities associated with AEO adverse events

For some particular to qualify as an instance of adverse event as defined in AEO the following key entities must exist:

(1) #1: a medical intervention (e.g., vaccination, drug administration)
(2) #2: a patient
(3) t1: the time at which the medical intervention is given to the patient
(4) #3: a clinically abnormal process (e.g., a fever process) which is an instance of adverse event if causally related to #1
(5) t2: the time at which the clinically abnormal process happens

These elements can be modeled in the adverse event design pattern of Fig. 2. Basically, both *adverse event* and *medical intervention* are subclasses of *processual_entity* (BFO). Instances of these two processes occur each at a specific time (Whenever continuants are involved, we need to specify a time [8]). The corresponding causal relation between the referents of these two process terms is represented using the object property term *induced_by* in AEO. The meaning of this relation term is illustrated in the section above. Such a relation term is not available in RO or any other ontologies. It is noted that the OBI term *process is result of* (OBI_1110060) is for direct causality and does not fit in our modeling of possible of indirect causality.

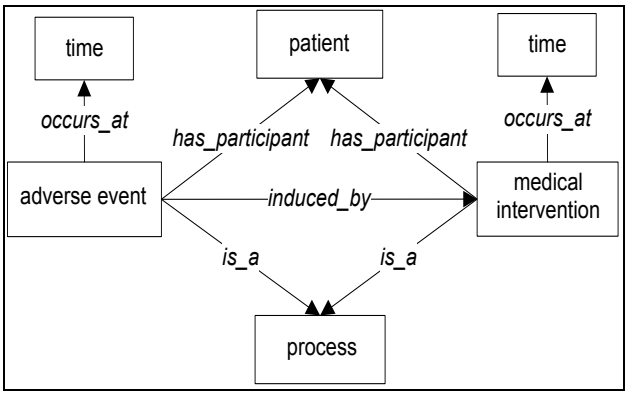

Fig. 2. Basic AEO adverse event design pattern.

Fig. 2 introduces the basic adverse event at the class level. In clinical cases, instance level modeling can be generated. For example, a specific vaccination process carried out on a particular patient is an *instance_of* vaccination which is a medical intervention. To illustrate this and other important points, an example is provided in the next section. It is also noted that the time at which a medical intervention is given to a patient is always earlier than the time at which an adverse event occurs, i.e., t1 earlier t2 (this one can be made more precise in the context of some guideline, e.g., t1 less-than-4-days-earlier-than t2).

An adverse event can have different outcomes, including a symptom (e.g., fever) and another process (e.g., bacterial infection). AEO uses symptom-related terms (e.g., fever generation) from other existing ontologies such as the Gene Ontology (GO).

3.4 Adverse events vs. adverse event hypotheses

The requirement of a causal relation between an adverse event and a medical intervention in AEO is an important point. Rehan et al provides physician's guide to how to assess causal relations of adverse events induced by drug administration [11]. The causal requirement is different from that in concept of adverse event in existing adverse event reporting systems (*e.g.*, USA VAERS: http://vaers.hhs.gov/). Since what is reported as 'adverse event' in these systems may not be truly induced by a medical intervention (e.g. vaccination), these adverse event reporting systems contain rather references to pathological processes that happened in a specific timeframe after a medical intervention, some of which might be indeed adverse events in AEO sense.

The data stored in such an adverse event reporting system is typically used to generate hypotheses about whether there is causality involved between what is reported as adverse events and medical interventions. Such a hypothesis, represented by the term *adverse event hypothesis* in AEO, becomes critical when a dramatically large amount of cases are reported following the same medical intervention. Therefore, adverse event reporting is not an end. To find potential safety problem is an ultimate goal of reporting adverse events. This is one reason why AEO aims to represent not only the adverse event hypothesis, but also the final causal association.

In a clinical setting, a clinician or a patient reports an event, which eventually may be proven to be caused by the medical intervention. However, when it is proven, such an event does not "become" an AEO adverse event. It was an instance thereof from the very beginning.

3.5 Comparison with other adverse event representation systems

Many other adverse event representation systems have been developed. Ceusters et al. provides an excellent survey and summary on different types of adverse event representation [9]. Here we particularly compare our AEO approach with the representation model for adverse drug reactions (ADRs) provided by Bosquet et al [3].

Edwards et al define an adverse drug reaction as "An appreciably harmful or unpleasant reaction, resulting from an intervention related to the use of a medicinal product, which predicts hazard from future administration and warrants prevention or specific treatment, or alteration of the dosage regimen, or withdrawal of the product" [3, 12]. The problem with this definition is that it is not specified, for instance, for whom the reaction is unpleasant (appreciation can be different for the patient, his caregivers and his relatives) and that it is prone to, so we assume, unwanted interpretations. Imagine a patient that took an oral overdose of some medicinal product and therefore is subjected to gastric suction to remove what is left in the stomach. Due to erroneous manipulation of the suction device, the patient develops a gastric bleeding. Clearly, this intervention is related to the use of a medicinal product, but it would be wrong to state, although in line with Edwards' definition, that this gastric bleeding is an adverse drug reaction.

Bosquet et al generated an ADR model that contains 19 semantic categories, and the categorical structure consists of 8 semantic categories within that model. Sixteen semantic links are described in their ontology. The set of minimal constraints are 4: an ADR should be classified as a disorder, an accident, an investigation, or a syndrome. A structural disorder is defined by at least one location and one morphology. A functional disorder is defined by at least one abnormal function. There are at least one semantic link is_related_to and one semantic category "Drug".

The work by Bosquet et al largely differs from ours. First, their ontology is based on categorial design, while AEO is based on OBO foundry ontology design. Second, their approach does not model time dependency between a drug administration and an adverse event. Third, a causal relation between a drug administration and an adverse event is not clearly specified in their system, although it can be assumed to be the case under some interpretation of 'resulting from' in their definition.

3.6 Use case study: vaccine-induced adverse events.

In the USA, more than 10 million vaccines per year are administrated to children less than 1 year old, usually between 2 and 6 months of age. At this age, infants are at greatest risk for many medical adverse events such as high fevers, seizures, and sudden infant death syndrome.

Fig. 3 provides an example of how AEO represents a specific vaccine-induced fever adverse event. In this example, Bob (a patient) was vaccinated with an Afluria flu vaccination at time t1, and then had a fever at time t2. Since it is notified in the vaccine instruction that fever generation is an expected adverse event and Bob was in good health before the vaccination, Bob's fever generation is considered as an adverse event induced by the vaccination process. The term *fever generation* is imported from the Gene Ontology (GO).

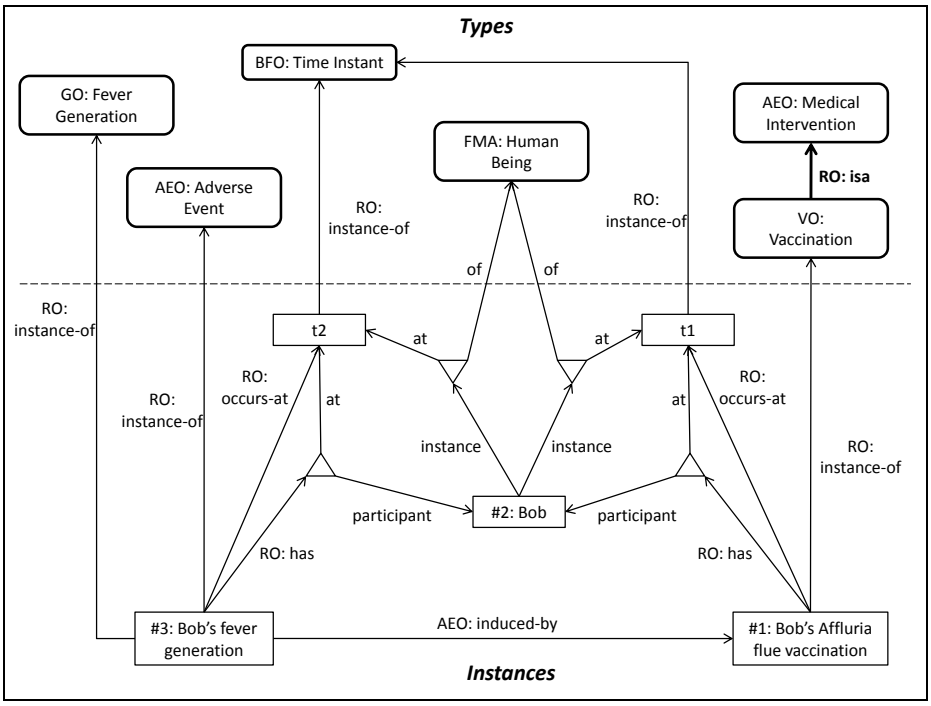

Fig. 3. Modeling of vaccination-induced fever adverse event in AEO.

The Brighton Collaboration (https://brightoncollaboration.org/) is a global research network that set vaccine safety research standards and does not either assume a cause-and-effect relation. According to the Brighton Collaboration, fever is defined as an elevation of body temperature above the normal [13]. Similar to other Brighton Collaboration definitions, the fever definition itself defines a clinical entity without inference of a causal relation to a given exposure. Therefore, the time interval from immunization until onset of the event cannot be part of the definition itself [13]. However, since AEO assumes such a cause-and-effect relation, this time interval is an important study topic in the AEO representation of an influenza vaccination and a fever vaccine adverse event. Therefore, we argue that AEO and those domain-specific adverse event ontologies aligned with AEO represent a knowledgebase of adverse events caused by medical interventions, where the data stored in regular adverse event reporting systems contain many random (coincident) and false positive events that are not induced by medical interventions.

4 Conclusion and Discussion

Adverse events endanger patients' safety and result in considerable extra healthcare costs. A community-based ontological representation of adverse events is crucial for improving adverse event research. Since adverse events are directly associated with public health, extensive efforts have been taken worldwide to represent and analyze adverse events. For example, the EU-funded the Patient Safety through Intelligent Procedures in medication (PSIP) project aims to develop innovative tools for generating and providing relevant knowledge to healthcare professionals and patients for ADE prevention. Another relevant project funded by EU is the European Public Warning System (EU-ALERT). The French VigiTermes project is a powerful application that automates potential adverse event detection by identification of statistical and semantic links between drugs, treatments and induced pathologies or symptoms. The EU funded ReMINE project uses an adverse event ontology to manage patient safety risks in hospital settings [14]. These projects are currently under way with a focus on using ontologies in order to facilitate identification of drug related adverse events, combining ontologies with information extraction and also applying ontologies to hospital data. The advent of AEO provides an opportunity for the adverse event research community to work together towards ontology-based adverse event information representation and data analysis.

To monitor and study these adverse events, many vaccine and drug adverse event reporting systems have been established to collect information about adverse events that occur after the administration of licensed vaccines. The examples of national vaccine safety surveillance programs include the Vaccine Adverse Event Reporting System (VAERS) in the USA (http://vaers.hhs.gov/about/index) and the Adverse Events Following Immunization Reporting program by the Public Health Canada (http://www.phac-aspc.gc.ca/im/aefi-form-eng.php). These reported data contains coincidental events and those truly caused by vaccines. In our view, an ontological representation using AEO will provide a unified and machine-readable representation of various adverse events and support more advanced adverse event data analysis.

AEO is different from the new Adverse Event Ontology Reporting Ontology (AERO) (http://purl.obolibrary.org/obo/aero.owl). AERO focuses on clinical adverse event reporting. These adverse events do not assume causal relation between a reported adverse event and a medical intervention (e.g., vaccination). AERO focuses on clinical instance data. In contrast, AEO specifies the causal relation and focuses on a class level classification.

It is possible to reconcile AEO and AERO in a future time. While the events included in AERO for a specific medical intervention may be larger than the true adverse events caused by this intervention, AEO has more depth and targets for representation of a knowledgebase of adverse events truly cased by medical interventions. How to find out the cause-and-effect relation from the reported adverse events in adverse event reporting systems is often a challenge. It will surely benefit

the public health and has been a critical research topic ever since such an adverse event reporting system is invented.

Many efforts are required to improve AEO. For example, it is important to link AEO to other adverse event representation systems, including MedDRA and WHO-ART, for better adverse event data representation and knowledgebase establishment, although caution is here required because of the lack of formal rigor in these systems [15]. It will be challenging and rewarding to predict and identify temporal causal relations of adverse events using informatics approaches (e.g., statistical algorithms, and literature mining). The drug adverse events are often affected by the genetic background (e.g. SNPs) of the patient. The intricate drug-patient and drug-drug interactions are crucial to determine the final adverse event outcomes. Some adverse events happen due to cross-interactions between drug and non-drugs (e.g. grapefruit). Sometimes, an adverse event emerges when a drug is removed. It would be ideal to model these interactions in AEO with a purpose to understand the fundamental adverse event mechanisms.

Acknowledgments. Development of AEO is supported by NIH grant 1R01AI081062.

References

1. Bousquet, C., et al., *Implementation of automated signal generation in pharmacovigilance using a knowledge-based approach.* Int J Med Inform, 2005. **74**(7-8): p. 563-71.
2. Alecu, I., et al., *PharmARTS: terminology web services for drug safety data coding and retrieval.* Stud Health Technol Inform, 2007. **129**(Pt 1): p. 699-704.
3. Bousquet, C., et al., *Semantic categories and relations for modelling adverse drug reactions towards a categorial structure for pharmacovigilance.* AMIA Annu Symp Proc, 2008: p. 61-5.
4. Klein, G.O. and B. Smith, *Concept Systems and Ontologies: Recommendations for Basic Terminology.* Transactions of the Japanese Society for Artificial Intelligence, 2010. **25**: p. 433-441.
5. Smith, B., et al., *The OBO Foundry: coordinated evolution of ontologies to support biomedical data integration.* Nat Biotechnol, 2007. **25**(11): p. 1251-5.
6. Smith, B. and W. Ceusters, *Ontological Realism as a Methodology for Coordinated Evolution of Scientific Ontologies.* Applied Ontology, 2010. **5**(3-4): p. 139-188.
7. Bittner, T., Smith, B. , *Normalizing Medical Ontologies Using Basic Formal Ontology.* Tagungsband der 49. Jahrestagung der Deutschen Gesellschaft für Medizinische Informatik, Biometrie und Epidemiologie, 2004.
8. Smith, B., et al., *Relations in biomedical ontologies.* Genome Biology, 2005. **6**(5): p. R46.
9. Ceusters, W., et al., *An evolutionary approach to the representation of adverse events.* Stud Health Technol Inform, 2009. **150**: p. 537-41.
10. Xiang, Z., et al., *OntoFox: web-based support for ontology reuse.* BMC Res Notes, 2010. **3**: p. 175.
11. Rehan, H.S., D. Chopra, and A.K. Kakkar, *Physician's guide to pharmacovigilance: terminology and causality assessment.* Eur J Intern Med, 2009. **20**(1): p. 3-8.
12. Edwards, I.R. and J.K. Aronson, *Adverse drug reactions: definitions, diagnosis, and management.* Lancet, 2000. **356**(9237): p. 1255-9.
13. Michael Marcy, S., et al., *Fever as an adverse event following immunization: case definition and guidelines of data collection, analysis, and presentation.* Vaccine, 2004. **22**(5-6): p. 551-6.
14. Ceusters, W., et al., *An Evolutionary Approach to Realism-Based Adverse Event Representations.* Methods of Information in Medicine, 2011. **50**(1): p. 62-73.
15. Bousquet, C., et al., *Appraisal of the MedDRA conceptual structure for describing and grouping adverse drug reactions.* Drug Safety, 2005. **28**(1): p. 19-34.