Cross-Border Challenges in Informatics with a Focus on Disease Surveillance and Utilising Big-Data 15
L. Stoicu-Tivadar et al. (Eds.)

doi:10.3233/978-1-61499-389-6-15

Ontologies to capture Adverse Events Following Immunisation (AEFI) from real world health data

Harshana LIYANAGE[a] and Simon de LUSIGNAN[a,1]
[a] *Department of Health Care Management & Policy, University of Surrey, UK*

Abstract. Immunisation is an important part of health care and adverse events following immunisation (AEFI) are relatively rare. AEFI can be detected through long term follow up of a cohort or from looking for signals from real world, routine data; from different health systems using a variety of clinical coding systems. Mapping these is a challenging aspect of integrating data across borders. Ontological representations of clinical concepts provide a method to map similar concepts, in this case AEFI across different coding systems. We describe a method using ontologies to be flag definite, probable or possible cases. We use Guillain-Barre syndrome (GBS) as an AEFI to illustrate this method, and the Brighton collaboration's case definition of GBS as the gold standard. Our method can be used to flag definite, probable or possible cases of GBS. Whilst there has been much research into the use of ontologies in immunisation these have focussed on database interrogation; where ours looks to identify varying signal strength.

Keywords. Biomedical Ontologies, Immunisation, Drug-Related Side Effects, Adverse Reactions, Computerized Medical Records Systems.

Introduction

Monitoring burden of vaccine preventable disease and vaccines' benefits and risk are essential elements of modern public health surveillance. Investigators are moving towards using "big data" [1]. International surveillance is likely to become more important with increased globalisation [2]. Building a common data model is important for international surveillance [3]. However, we also need automated methods that "understand" the semantics of these data [4]; ontologies formally define the semantic relationships between data items and offer the allure of enabling machine processing [5, 6].

Detection of Adverse Events Following Immunization (AEFIs) is vital if we are to quantify the benefits of vaccination. The Vaccine Adverse Event Reporting System (VAERS) is an example of a national passive surveillance method used for detecting adverse events in USA. VAERS allows direct reporting by members of the public and utilised automated methods for classifying adverse events reported to it [7]. However, such reporting systems vary between countries, collecting data in non-standard ways.

[1] Corresponding Author.

Vaccine informatics focuses on development and use of bioinformatics methods during preclinical, clinical and post-licensing stages of vaccine development and deployment [8]. This branch of informatics has largely focussed on the digitalisation of a well regulated process and ensuring complete datasets are collected and extracted from data sources. Whilst useful, we propose a much more agile approach.

The Brighton Collaboration has developed case definitions required to define AEFI; these include Guillain-Barre syndrome (GBS) [9]. GBS is rare, and very large datasets are needed to detect it [10]; a further challenge is that GBS is a clinically heterogeneous disorder. We use GBS as an AEFI to illustrate our method, and the Brighton collaboration's definition of GBS as the gold standard (Table 1). Our ontology might be used to detect the signals associated with definite or possible cases of GBS.

Table 1. Brighton Collaboration Guillain-Barré syndrome (GBS) case definition (CSF=cerebrospinal fluid)

Diagnostic certainty	Clinical criteria
Level 1	Bilateral and flaccid weakness of the limbs *and* Decreased or absent deep tendon reflexes in weak limbs *and* Monophasic illness pattern, with interval between onset and nadir of weakness between 12 hours and 28 days, and subsequent clinical plateau *and* Electrophysiologic findings consistent with GBS *and* Cytoalbuminologic dissociation (ie, elevation of CSF protein level above laboratory normal value, and CSF total white cell count < 50 cells/μL) *and* Absence of an identified alternative diagnosis for weakness
Level 2	Bilateral and flaccid weakness of the limbs *and* Decreased or absent deep tendon reflexes in weak limbs *and* Monophasic illness pattern, with interval between onset and nadir of weakness between 12 hours and 28 days, and subsequent clinical plateau *and* CSF total white cell count <50 cells/μL (with or without CSF protein elevation above laboratory reference range) or, if CSF not collected or results not available, electrophysiologic studies consistent with GBS *and* Absence of identified alternative diagnosis for weakness
Level 3	Bilateral and flaccid weakness of the limbs *and* Decreased or absent deep tendon reflexes in weak limbs *and* Monophasic illness pattern, with interval between onset and nadir of weakness between 12 hours and 28 days, and subsequent clinical plateau *and* Absence of identified alternative diagnosis for weakness
Level 4a	GBS based on a clear statement from a treating neurologist that GBS is the diagnosis being present in the medical record and there being no contradictory information

1. Vaccine related ontology development

Vaccination programmes are generally implemented in a global scale; while data related to vaccine coverage, benefits and risks are generated and managed at national level. This introduces a diversity of data, which is a challenge for global vaccination monitoring [11]. Ontologies developed to detect AEFI must cope with this complexity.

Vaccine Ontology (VO) is a community based biomedical ontology in this domain. It contains more than 5000 vaccine-specific ontological terms [12]. The introduction of the Vaccine Ontology has been followed by efforts to develop ontologies to conceptualise adverse events. Ontology of Adverse Events (OAE) and Ontology of Vaccine Adverse Events (OVAE), an extension of OAE, has been developed to formally represent and analyse AEFI [13, 14]. The Adverse Event Reporting Ontology (AERO) has been introduced to standardise reporting of AEFI [15]. VO has largely been used to ensure the completeness of data capture about AEFI rather than, as we propose to systematically identify cases from clinical databases. They take little account about the granularity of the coding system, the nature of data recording, including free-text records when searching for signals of possible AEFI [6].

2. An AEFI ontology to detect GBS signals across multiple coding systems

We propose an ontology, which will be useful for formally integrating adverse event data from computerised medical records (CMR). This process takes into account the degree of specificity with which the clinical concept can be represented within the extractable CMR data; at present largely coded data [7]. The ontology will detect signals of varying specificity and sensitivity depending on the data available. We can classify the outputs as having complete, partial or no clear mapping (Table 2).

Table 2. Class of mapping possible from the ontology

Class	Mapping	Interpretation
Class 0	No clear mapping	Possible case
Class 1-3	Partial mapping	Possible/Probable case
Class 4	Complete mapping	Definite case

Class 0 will indicate no clear mapping but one or more possible symptoms or data suggestive of a *possible case* are present. Class 1 will be given when there is a diagnosis or a compatible diagnosis only, there are different codes available in the common coding systems used (Table 3).

Table 3. Representation of GBS in different coding systems

Clinical Concept	Read Code	ICD9	ICD10	ICPC
Guillain-Barré syndrome	F370.00,F370000, F370100	357.0	G61.0	N94.1

Class 2 will be assigned when there is additional supporting administrative evidence (e.g. period of admission); Class 3 will be given if the supporting clinical evidence is present. (e.g. immunoglobulin therapy) which is part of the Brighton case definition. Finally, Class 4 is a *definite case* with all aspects of mapping complete. We may be able to compare reported levels of partially mapped with completely mapped.

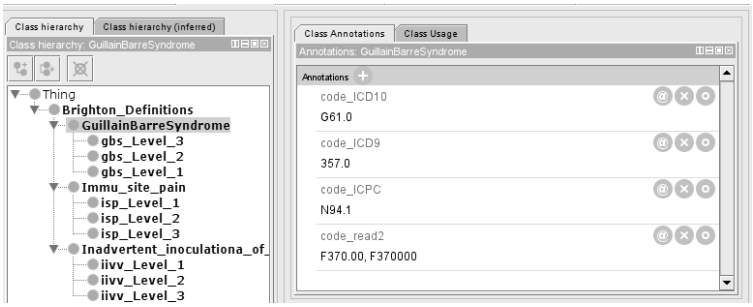

Figure 1. Representation codes for GBS in multiple coding systems

The proposed ontology has been developed (Figure 1.) using the ontology modelling tool, Protégé (http://protege.stanford.edu) [16]. The semantics of the ontology has been described according to the OWL (Web Ontology Language) specification by the W3 Consortium [17].

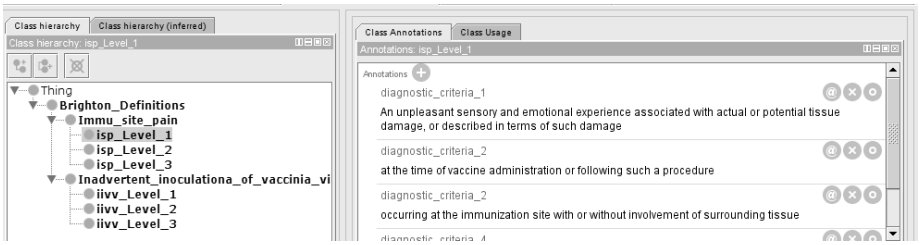

Figure 2. Representation of Brighton collaboration case definitions as ontology

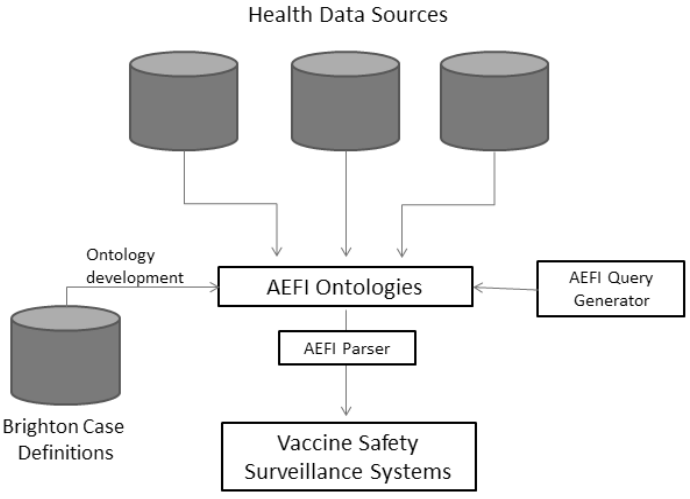

Figure 3. Parsing/extracting data from multiple coding systems using AEFI ontologies

An AEFI mapping ontology has been developed based on Brighton Case Definitions for adverse events (Figure 2). This supports annotating code from multiple coding systems. Annotated AEFI ontologies can then be used for generating queries for

extracting AEFI related data from various health data sources; the associated AEFI parser can be used to parse the data and analyse a health data set (Figure 3).

3. Conclusion

Ontologies can be constructed that enable the consistent and reliable identification of AEFI, where there is both a complete set of coded data that map to a defined AEFI and the sematic relationships. The ontology will also identify possible and probable cases, which cannot be directly mapped to AEFI, but can be investigated further. Developing ontologies that define the relationship between clinical concept and coded health data is a small step towards the automated detection of AEFI from heterogeneous data sources.

References

[1] S.I. Hay, D.B. George, C.L. Moyes, J.S. Brownstein, Big data opportunities for global infectious disease surveillance, *PLoS Med*.**10**(4) (2013), e1001413.

[2] P. Kramarz, P.L. Lopalco, E. Huitric, L. Pastore Celentano, Vaccine-preventable diseases: the role of the European Centre for Disease Prevention and Control, *Clin Microbiol Infect*. (2013).

[3] Eurosurveillance editorial team. ECDC in collaboration with the VAESCO consortium to develop a complementary tool for vaccine safety monitoring in Europe. *Euro Surveill.* **14**(39) (2009).

[4] Seth Grimes, Unstructured Data and the 80 Percent Rule: Investigating the 80%, Clarabridge, Bridgepoints, (2008) Q3.

[5] H. Liyanage, S.T. Liaw, C. Kuziemsky, S. de Lusignan, Ontologies to improve chronic disease management research and quality improvement studies - a conceptual framework, *Stud Health Technol Inform* **192** (2013), 180-184.

[6] H. Liyanage et al., The Evidence-base for Using Ontologies and Semantic Integration Methodologies to Support Integrated Chronic Disease Management in Primary and Ambulatory Care: Realist Review, Contribution of the IMIA Primary Health Care Informatics WG, *Yearb Med Inform.***8**(1) (2013), 147-154.

[7] T. Botsis, E.J. Woo, R. Ball, Application of information retrieval approaches to case classification in the vaccine adverse event reporting system, *Drug Saf*. **36**(7) (2013), 573-82.

[8] Y. He, R. Rappuoli, A.S. De Groot, R.T. Chen, Emerging vaccine informatics, *J Biomed Biotechnol*. 2010:218590 (2013).

[9] J.J. Sejvar, K.S. Kohl, J. Gidudu et al., Brighton Collaboration GBS Working Group. Guillain-Barré syndrome and Fisher syndrome: case definitions and guidelines for collection, *Vaccine* **29**(3) (2011), 599-612.

[10] C. Fokke et al..Diagnosis of Guillain-Barre syndrome and validation of Brighton criteria, *Brain* **137**(Pt 1) (2014), 33-43.

[11] P. L. Lopalco, P. Carrillo Santisteve, Actual immunization coverage throughout Europe: are existing data sufficient?, *Clin Microbiol Infect*. doi: 10.1111/1469-0691.12430 (2013).

[12] Y. He, VO: Vaccine Ontology, *The 1st International Conference on Biomedical Ontology (ICBO 2009) (*2009).

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

[14] E. Marcos, B. Zhao, Y. He, The Ontology of Vaccine Adverse Events (OVAE) and its usage in representing and analyzing adverse events*, J Biomed Semantics* **4**(1):40 (2013).

[15] M. Courtot, R.R. Brinkman, Towards an Adverse Event Reporting Ontology. Adverse Event Representation Workshop, *International Conference on Biomedical Ontologies (ICBO),* July 26–30, 2011. NY: University at Buffalo (2011).

[16] H. Knublauch et al., The Protégé OWL Experience, Workshop on OWL: Experiences and Directions, *Fourth International Semantic Web Conference (ISWC2005),* Galway, Ireland.

[17] D.L. McGuinness, F. Van Harmelen, OWL web ontology language overview. W3C recommendation.**10** (2004):10.