

Evaluation of automated term groupings for detecting anaphylactic shock signals for drugs

Julien Souvignet¹, MS, Gunnar Declerck¹, PhD, Béatrice Trombert^{1,2}, MD, PhD, Jean Marie Rodrigues^{1,2,3}, MD, PhD, Marie-Christine Jaulent¹, PhD and Cédric Bousquet^{1,2}, PharmD, PhD

¹ INSERM U872, Eq. 20, Paris, France

² University of Saint Etienne, Department of Public Health and Medical Informatics, Saint-Etienne, France

³ WHO FIC Collaborative Centre for International Classifications in French Language, Paris, France

Abstract

Signal detection in pharmacovigilance should take into account all terms related to a medical concept rather than a single term. We built an OWL-DL file with formal definitions of MedDRA and SNOMED-CT concepts and performed two queries, Query 1 and 2, to retrieve narrow and broad terms within the Standard MedDRA Query (SMQ) related to 'anaphylactic shock' and the terms from the High Level Term (HLT) grouping related to 'anaphylaxis'. We compared values of the EB05 (EBGM) statistical test for disproportionality with 50 active ingredients randomly selected in the public version of the FDA pharmacovigilance database. Coefficient of correlation was $R^2 = 1.00$ between Query 1 and HLT; $R^2 = 0.98$ between Query 1 and SMQ narrow; $R^2 = 0.89$ between Query 2 and SMQ Narrow+Broad. Generating automated groupings of terms for signal detection is feasible but requires additional efforts in modeling MedDRA terms in order to improve precision and recall of these groupings.

Introduction

The main objective of pharmacovigilance is to reduce drug-related risks. All adverse drug reactions (ADR) are not known at the time of commercialization and this may lead to improper care of the patient. The continuous development of new drugs requires an early detection of their unknown adverse effects¹. These discoveries may lead to suspension or withdrawal of drugs-treatments. A constant and sustained post-marketing surveillance process of ADRs is therefore essential².

Reporting of ADRs observed by health professionals, as well as a continuous analysis of case reports by regulatory authorities and pharmaceutical industry, is a necessary step towards drug-related risks reduction. Analysis of reported ADRs can be carried out by a manual expert review, but such process becomes more and more difficult at a human level due to the large amount of information to analyze³. Drawing expert's attention on relevant combinations of drug-adverse reaction pairs in pharmacovigilance databases is necessary. To this end, different automated methods have been developed to supplement qualitative clinical methods⁴.

ADRs in case reports are usually coded with the MedDRA®* terminology⁵ (Medical Dictionary for Drug Regulatory Activities) and stored in databases that constitute knowledge on suspected ADRs. Signal detection in pharmacovigilance should take into account all terms related to a medical concept rather than a single term^{6,7}. For instance, if a given drug is suspected to cause acute renal failure, using the MedDRA term 'Renal failure acute' is generally not sufficient for the algorithms to detect a signal. When selecting case reports it is recommended to add related MedDRA terms such as 'Renal impairment', 'Blood creatinine abnormal' or 'Dialysis' in order to have a broader scope. Several authors have studied the impact of grouping terms before signal detection with different outcomes^{8,9,10}.

We assume that it is possible to generate groups of MedDRA terms using knowledge engineering methods to represent a given clinical condition¹¹. A prerequisite to perform such groups by terminological reasoning (logical inferences based on semantic content) is that formal representations of the semantics of terms are available¹². To that

* MedDRA® is a registered trademark of the International Federation of Pharmaceutical Manufacturers and Associations.

aim, we have developed an OWL-DL (Web Ontology Language – Description Logic) file with formal definition of ADRs (named *OntoADR*[†]) in order to support semantic query-based generation of groups of terms relating to similar medical conditions.

The goal of the present study is to assess the efficiency of this DL-query based MedDRA terms grouping method for statistical research of signals in pharmacovigilance databases. Anaphylactic shock topic was selected for the following reasons. First this topic has a related HLT (High Level Term) and can be associated with both a narrow and broad part of a SMQ (Standardized MedDRA Query), so multiple queries were built to retrieve respectively terms within the narrow and broad part of the SMQ. Second our grouping and the SMQ share common terms but present a higher number of terms that are present only in our grouping or in the SMQ. While the interpretation of high correlation in statistical measure would be trivial with comparable groupings, explaining such correlation among groupings that present a degree of dissimilarity was more challenging and could provide deeper understanding of signal detection with large groups of terms compared to single preferred terms. We performed this evaluation on the US Food and Drug Administration's (FDA) public database¹⁴ and we used Standardized MedDRA Queries as gold standard¹⁵.

Background

FDA AERS

The FDA's Adverse Event Reporting System (AERS)¹⁴ is the official database for spontaneous reports of adverse drug reactions in the United States. This database consists of more than 2 million reports submitted by manufacturers (by regulatory mandate) and by clinicians and patients (through the MedWatch program¹⁵).

The data structure of AERS consists of 7 data sets: patient demographic and administrative information, drug/biologic information, patient outcomes, report sources, drug therapy start and end dates, indications for use/diagnosis and adverse events which are coded with MedDRA.

MedDRA

MedDRA is a terminology used by regulatory authorities and the biopharmaceutical industry to code information in ADR reports including ADRs/AEs (whether diagnoses, signs, symptoms, etc.), indications, medical and social history, investigations, and medical and surgical procedures¹⁶. MedDRA provides a standard terminology with a hierarchy of terms, organized by System Organ Class (SOC), divided into High-Level Group Terms (HLGT), High-Level Terms (HLT), Preferred Terms (PT) and Lowest Level Terms (LLT).

Identifying clinically related terms in MedDRA is not an easy task as those terms might exist in different locations in the hierarchy. The original MedDRA hierarchy already offers HLT groupings, sets of several medically related PTs within the same SOC. But it was recognized that HLTs are not always sufficient to represent clinical conditions involving several organs (e.g., kidney, liver, cardiovascular and respiratory systems)¹¹. This led to the development of SMQs¹² that combine terms from multiple SOCs.

SMQs are groupings of MedDRA terms, that relate to a defined medical condition or area of interest and which are intended to aid in case identification. Within a SMQ narrow terms help users to identify case reports that are highly likely to represent the condition of interest, and broad terms other case reports that may be related to a given medical condition but lack of specificity (e.g., clinical findings or results of investigations observed in these medical conditions but also in other conditions). A broad search with a SMQ includes both the narrow and broad terms.

HLT and SMQs are constructed manually by expert consensus and can be reused as a standard to allow international comparison between drugs. However they do not cover all medical conditions that may be related to a drug or may not have the specificity required. For example there is a SMQ for 'gastrointestinal bleeding' but not for 'upper gastrointestinal bleeding'. Such a grouping can be requested to MSSO to be added in a future version, but there is no way to get them quickly.

OntoADR

OntoADR^{†13} is an OWL-DL file with formal definitions of Adverse Drug Reactions that is being developed to support logic queries and to perform terminological reasoning for MedDRA terms grouping. Concepts are defined

[†] Access to *OntoADR* is currently not public due to right restrictions with the terms of use of MedDRA® and SNOMED-CT®.

with semantic properties corresponding to relations used in the medical domain as defined in Systematized Nomenclature of Medicine – Clinical Terms (SNOMED-CT®) clinical terminology. Twenty-six relations were selected from SNOMED-CT, among which: *hasFindingSite*, which specifies the body site affected by a condition; *hasAssociatedMorphology*, which describes the morphologic changes seen at the tissue or cellular level that are characteristic features of a disease; or *hasOccurrence*, which refers to the onset or period of life during which a condition first presents. To define MedDRA concepts in OntoADR, we used UMLS[‡] (Unified Medical Language System) metathesaurus mappings with SNOMED-CT. When MedDRA concepts could not be mapped with a SNOMED-CT concept, its formal definition was achieved manually by knowledge engineers and pharmacovigilance experts. Through OntoADR, MedDRA concepts are thus defined by sets of properties corresponding to a decomposition of their medical meaning, and can be grouped together using queries.

Signal Detection

Several statistical methods for signal detection in pharmacovigilance have been proposed by researchers: Dumouchel for the Food and Drug Administration with the Empirical Bayes Geometric Mean (EBGM)¹⁷, Bate for World Health Organization (WHO) with the Information component (IC)¹⁸ or Evans for the Medicines and Healthcare products Regulatory Agency (MHRA) with the Proportionate Reporting Ratio (PRR)¹⁹. The calculation of these indicators is based on the number of observed cases to be significantly greater than the number of expected cases.

Methods

FDA AERS

Input data for this study were taken from the public release of the FDA's AERS database, which covers the period from the first quarter of 2004 to the end of 2010.

Prior to analysis, all drug names coded with free text were cleaned-up using text mining approach. Adverse events were coded with preferred terms (PTs) but over the years, MedDRA evolution has caused some PT to be demoted into LLT, so unification in preferred terms had to be made. Duplicate reports and follow-ups were also deleted in order to keep the most recent case number (a numerical id describing a case report in FDA AERS database).

To perform signal detection, we randomly selected 50 active ingredients from the 500 most frequent drugs present in FDA case reports (see table 1).

Table 1. List of 50 randomly selected active ingredients.

ACETAMINOPHEN	CLARITHROMYCIN	FAMOTIDINE	METFORMIN	RAMIPRIL
ACYCLOVIR	CLINDAMYCIN	FENTANYL	METHADONE	RIBAVIRIN
ALENDRONATE	CODEINE	FLUDARABINE	METOPROLOL	RISPERIDONE
ATENOLOL	CYTARABINE	FUROSEMIDE	METRONIDAZOLE	SPIRONOLACTONE
ATORVASTATIN	DEXAMETHASONE	GABAPENTIN	NIFEDIPINE	TEMAZEPAM
AZATHIOPRINE	DIAZEPAM	GLIMEPIRIDE	OLANZAPINE	TERAZOSIN
AZITHROMYCIN	DOXORUBICIN	IBUPROFEN	PAROXETINE	THALIDOMIDE
BACLOFEN	ENALAPRIL	INFLIXIMAB	PHENOBARBITAL	THEOPHYLLINE
BISOPROLOL	ESOMEPRAZOLE	IRINOTECAN	PRAVASTATIN	TRAZODONE
CETUXIMAB	ETOPOSIDE	LOPERAMIDE	PROPOFOL	ZIDOVUDINE

MedDRA groupings used as gold standard

We used MedDRA version 14.1 in English Language available on 1 September 2011 from the MedDRA Maintenance and Support Services Organization (MSSO) Web site¹⁷. As term grouping reference (gold standard) for our topic 'Anaphylactic shock' we selected *HLT Anaphylactic Responses* and *SMQ Anaphylactic/anaphylactoid shock conditions*. This is a sub-SMQ from the 'Shock' SMQ that also contains other sub-SMQ such as 'Toxic-septic shock conditions' or 'Hypovolaemic shock conditions'. The 'Shock' SMQ has inclusion criteria (e.g., organ failure

[‡] The UMLS is a set of files and software developed by the NLM (U.S. National Library of Medicine) that brings together many health and biomedical vocabularies and standards (including MedDRA and SNOMED-CT) to enable interoperability between computer systems. <http://www.nlm.nih.gov/research/umls/>

terms and terms containing the words ‘anuria’ or ‘hypoperfusion’) and exclusion criteria (e.g., electrical shock and traumatic shock terms). This SMQ has some specific terms (Narrow) and less specific terms (Broad) (see Table 3).

OntoADR queries

We used OntoADR (November 2011 build). Two queries were developed to match the safety topic: ‘Anaphylactic shock’. The first one, named Query 1, is a basic query targeting pure anaphylaxis criteria and no restriction on the ‘shock’ character.

hasDefinitionalManifestation some 'Anaphylaxis' (Query 1)

This query aims to replicate the HLT and focuses only on the manifestation and not on the ‘shock’ property.

Query 2 is a more SMQ-like query, also targeting cardiovascular/respiratory/hepatic system affection with acute and shock or failure character.

hasDefinitionalManifestation some 'Anaphylaxis'
 OR (
 (**hasFindingSite** some 'Structure of cardiovascular system'
 OR **hasFindingSite** some 'Structure of respiratory system'
 OR **hasFindingSite** some 'Kidney structure')
 AND **hasClinicalCourse** some 'Sudden onset AND/OR short duration'
 AND (**hasDefinitionalManifestation** some 'Shock'
 OR **hasDefinitionalManifestation** some 'Failure')
) (Query 2)

Signal Detection

Multiple statistical tests are used for pharmacovigilance analyses to identify signals of drug-associated adverse reactions that are significantly reported more frequently than expected. All are based on 4 numerical values involving all drugs and all adverse reactions in a pharmacovigilance database (see Table 2).

Table 2. The four algebraic values used for statistical test in a database.

	ADR or ADR group	Other reactions	Total
Drug of interest	<i>a</i>	<i>b</i>	<i>a+b</i>
Other drugs	<i>c</i>	<i>d</i>	<i>c+d</i>
Total	<i>a+c</i>	<i>b+d</i>	

Using these values, statistical tests estimate expected reporting frequencies for each couple (drug - adverse reaction) and determinate a value for a signal.

We implemented current data mining algorithms (PRR, ROR, Yule-Q, IC and EBGM) and we selected EBGM because it is the algorithm recommended by the FDA. Each algorithm for signal detection has a metric, to test if a signal is detected. For EBGM, we used a criterion: the EB05 metric had to be greater than or equal to a threshold value of 2. EB05 is a lower one-sided 95% confidence limit of EBGM.

For every 50 active ingredients we selected, we calculate EB05 values for every group of term (HLT, SMQ, Query 1 and 2) and compared them.

To evaluate the proportion of variability in the data set, we use the coefficient of determination R^2 , which is the correlation coefficient squared. We estimated if there was a linear relation ($y = ax + b$) or even equality ($y = x$) between signal values for SMQs and our grouping. R^2 is a statistical value giving some information about the goodness of fit of a model. The coefficient of determination ranges from 0 to 1: an R^2 of 1.0 indicates that the regression line perfectly fits the data.

Results

Anaphylactic shock							
HLT Anaphylactic responses				OntoADR Query 1			
Type	MedDRA Label	in Query 1?	in Query 2?	MedDRA Label	in HLT?	in SMQ (N)?	in SMQ (N+B)?
HLT	Anaphylactic reaction	Yes	Yes	Anaphylactic reaction	Yes	Yes	Yes
HLT	Anaphylactic shock	Yes	Yes	Anaphylactic shock	Yes	Yes	Yes
HLT	Anaphylactic transfusion reaction	Yes	Yes	Anaphylactic transfusion reaction	Yes	Yes	Yes
HLT	Anaphylactoid reaction	Yes	Yes	Anaphylactoid reaction	Yes	Yes	Yes
HLT	Anaphylactoid shock	Yes	Yes	Anaphylactoid shock	Yes	Yes	Yes
HLT	Anaphylactoid syndrome of pregnancy	Yes	Yes	Anaphylactoid syndrome of pregnancy	Yes	No	No
HLT	First use syndrome	Yes	Yes	First use syndrome	Yes	No	No
TOTAL HLT (/7)		7 (100%)	7 (100%)	TOTAL Query 1 (/7)	7 (100%)	5 (71%)	5 (71%)
SMQ Anaphylactic/Anaphylactoid shock conditions				OntoADR Query 2			
Type	MedDRA Label	in Query 1?	in Query 2?	MedDRA Label	in HLT?	in SMQ (N)?	in SMQ (N+B)?
SMQ Narrow	Anaphylactic reaction	Yes	Yes	Acute prerenal failure	No	No	Yes
SMQ Narrow	Anaphylactic shock	Yes	Yes	Acute pulmonary oedema	No	No	No
SMQ Narrow	Anaphylactic transfusion reaction	Yes	Yes	Acute respiratory failure	No	No	Yes
SMQ Narrow	Anaphylactoid reaction	Yes	Yes	Anaphylactic reaction	Yes	Yes	Yes
SMQ Narrow	Anaphylactoid shock	Yes	Yes	Anaphylactic shock	Yes	Yes	Yes
SMQ Narrow	Circulatory collapse	No	No	Anaphylactic transfusion reaction	Yes	Yes	Yes
SMQ Narrow	Shock	No	Yes	Anaphylactoid reaction	Yes	Yes	Yes
TOTAL SMQ Narrow (/7)		5 (71%)	6 (86%)	Anaphylactoid shock	Yes	Yes	Yes
SMQ Broad	Acute prerenal failure	No	Yes	Anaphylactoid syndrome of pregnancy	Yes	No	No
SMQ Broad	Acute respiratory failure	No	Yes	Cardiac failure acute	No	No	No
SMQ Broad	Anuria	No	No	Cardiogenic shock	No	No	No
SMQ Broad	Blood pressure immeasurable	No	No	Cor pulmonale acute	No	No	No
SMQ Broad	Cerebral hypoperfusion	No	No	Endotoxic shock	No	No	No
SMQ Broad	Grey syndrome neonatal	No	No	First use syndrome	Yes	No	No
SMQ Broad	Hepatic congestion	No	No	Hepatorenal failure	No	No	Yes
SMQ Broad	Hepatojugular reflux	No	No	Hypovolaemic shock	No	No	No
SMQ Broad	Hepatorenal failure	No	Yes	Neurogenic shock	No	No	No
SMQ Broad	Hypoperfusion	No	No	Peripheral circulatory failure	No	No	No
SMQ Broad	Jugular vein distension	No	No	Renal failure acute	No	No	Yes
SMQ Broad	Multi-organ failure	No	No	Septic shock	No	No	No
SMQ Broad	Myocardial depression	No	No	Shock	No	Yes	Yes
SMQ Broad	Neonatal anuria	No	No	Shock haemorrhagic	No	No	No
SMQ Broad	Neonatal multi-organ failure	No	No	Toxic shock syndrome	No	No	No
SMQ Broad	Neonatal respiratory failure	No	No	Toxic shock syndrome staphylococcal	No	No	No
SMQ Broad	Organ failure	No	No	Toxic shock syndrome streptococcal	No	No	No
SMQ Broad	Propofol infusion syndrome	No	No	Traumatic shock	No	No	No
SMQ Broad	Renal failure	No	No	TOTAL Query 2 (/26)			
SMQ Broad	Renal failure acute	No	Yes	7 (27%)	6 (23%)	10 (38%)	
SMQ Broad	Renal failure neonatal	No	No				
SMQ Broad	Respiratory failure	No	No				
TOTAL SMQ Narrow+Broad (/29)		5 (17%)	10 (34%)				

Table 3. Results of both query grouping and comparison with the content of HLT and SMQ used as gold standard.

Table 3 describes the result of terms grouping by performing Query 1 and Query 2 in OntoADR. On the left side are presented HLT and SMQ terms used as gold standard, and on the right side, terms from Query 1 and 2.

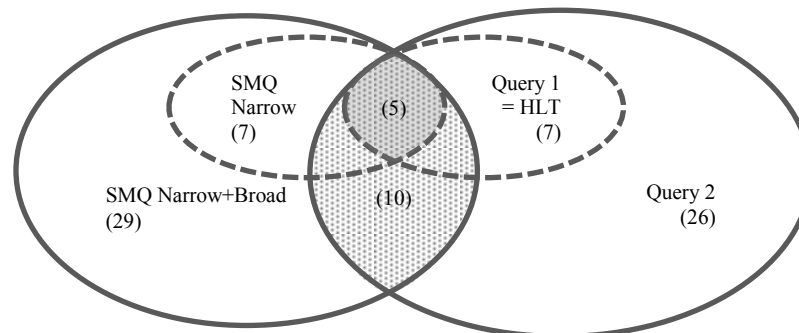


Figure 1. Venn-diagram representing group of terms, their intersections and their cardinal numbers.

For easier comparison, MedDRA terms common to or absent in other groupings are presented in table 3. Intersections of group of terms are illustrated in Figure 1. The content of Query 1 and HLT were identical. Two preferred terms present in SMQ narrow were absent from Query 1 ('Anaphylactoid syndrome of pregnancy' and 'First use syndrome'). Query 2 could retrieve an additional preferred term within the narrow part of the SMQ ('Shock'). With Query 1 no preferred terms were found within the broad part of the SMQ while Query 2 was able to propose four additional preferred terms related to the broad part ('Acute prerenal failure', 'Acute respiratory failure', 'Hepatorenal failure' and 'Renal failure acute'). Query 2 identified 14 additional terms that were not present in the SMQ neither the HLT (e.g., 'Acute pulmonary oedema', 'Cardiac failure acute', 'Cardiogenic shock', etc.).

Table 4 shows recall, precision and F-Measure for term-grouping, and also signal R^2 for each query vs. gold standard. Terms within the HLT and Query 1 were identical and both precision and recall were good (71.4%) for Query 1 versus the SMQ narrow as few additional terms were retrieved by the query. In the same time, the coefficient of determination for the signal was excellent (0.98). Precision and Recall were lower (34.5% and 38.5%) with Query 2 vs. SMQ Narrow+Broad as several terms absent from the SMQ were retrieved (e.g., 'Acute pulmonary oedema', 'Cardiac failure acute'). But, in terms of signal detection, R^2 is very good (0.89).

Table 4. Recall, Precision, F-measure for grouping and signal R^2 for each query.

Query 1	SMQ N	SMQ N+B	HLT	Query 2	SMQ N	SMQ N+B	HLT
Recall	71,4%	17,2%	100,0%	Recall	85,7%	34,5%	100,0%
Precision	71,4%	71,4%	100,0%	Precision	23,1%	38,5%	26,9%
F-measure	71,4%	27,8%	100,0%	F-measure	36,4%	36,4%	42,4%
Signal R^2	0.98	0.17	1.0	Signal R^2	0.51	0.89	0.42

Reminder: Query 1 tends to be closer to HLT (and SMQ Narrow) while Query 2 aims to approximate SMQ Narrow+Broad.

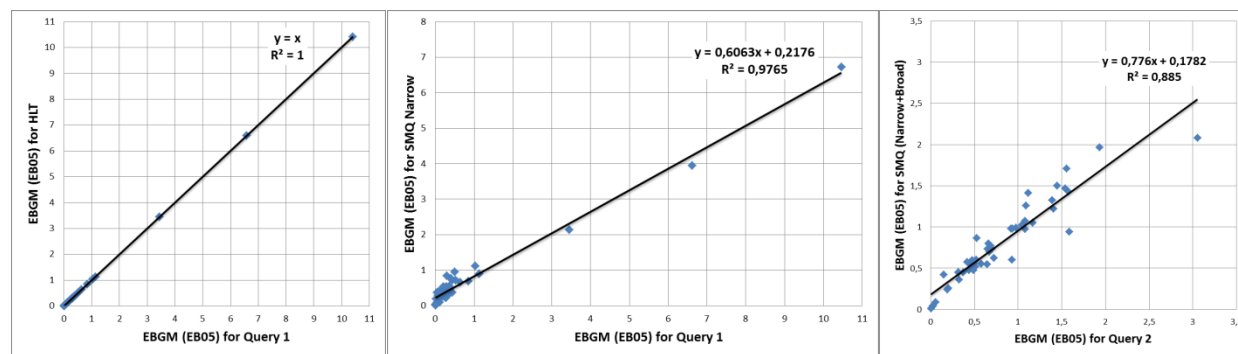


Figure 2. Results for signal detection for each query vs. SMQs used as gold standard.

Figure 2 illustrates how EB05 values are correlated between each grouping. Each dot represents the EB05 value of an active ingredient with a group of terms (x and y coordinate).

Discussion

Results of signal detection

As can be seen in the graphs of Figure 2, results of EB05 with our queries are highly correlated with measures of EB05 using the SMQ. This linear relationship is indicative that low (respectively high) measures of EB05 using the SMQ are related to low (respectively high) measures of EB05 when using our groupings. However the model fits

more with $y = ax + b$ than $y = x$ (intercept of the line with the axes was not the origin and slope of the line was different from 1.0) thus inducing different measures of EB05 with both groupings. Although the correlation is a predictive model of EB05 with SMQ knowing EB05 with our groupings, the interpretation of this correlation as an explicative model is difficult (i.e., it is tough to explain how measures of EB05 with our groupings can explain measures of EB05 with the SMQ). However we consider that results of high correlation were not due to chance and propose below an interpretation of the findings (i.e., why results of signal detection are highly correlated despite several terms are different in both groupings). We also replicated the results on other safety topics such as ‘Upper Gastrointestinal Hemorrhage’ or ‘Neutropenia’, with also very good coefficient of determination R^2 for the signal. The ability to retrieve similar findings with other safety topics pleads against the hypothesis that such finding was caused by chance for anaphylactic shock.

Building of OWL queries

Before choosing our querying strategies, we tried to use a strict definitional query, making a restriction both on ‘Shock’ and ‘Anaphylaxis’ on the *hasDefinitionalManifestation* semantic axis. But such a query only returns the MedDRA PTs: ‘Anaphylactic shock’ and ‘Anaphylactoid shock’. If we want the query to catch also anaphylactic reactions terms (and not only shocks terms), as it is the case in the SMQ ‘Anaphylactic/anaphylactoid shock conditions’ taken as gold standard (or even in HLT ‘Anaphylactic responses’), we have to delete the restriction on ‘Shock’ (cf. Query 1). And if we want the query to catch also anaphylactoid terms (and not only anaphylactic terms), we have also to delete the restriction on ‘Shock’ (cf. Query 2).

Some of the terms of the SMQ that are not returned by those different queries could be caught via an extension of query 2 (suppression or lessening of some of the initial restrictions). But the main drawback of such a procedure is that it generates a lot of noise. For instance, the PT ‘Circulatory collapse’ of the SMQ ‘Anaphylactic/anaphylactoid shock conditions’ can be caught by query 2 if the restrictions on the ‘Shock’ and ‘Failure’ characters are suppressed. But this suppression makes literally exploding the number of terms returned by the query (more than 80 terms) and therefore decreases dramatically precision. If the grouping is further reduced by a manual selection of safety topic relevant terms, this drawback is partially attenuated. But if it is not the case, such consequence is much more problematic, because only wrong signals will be detected (that is: signal that do not match the adverse drug event targeted by the safety topic). The same remark applies for PTs of the SMQ such as ‘Organ failure’ and ‘Multi-organ failure’ that could be returned by query 2 modulo the suppression of the restriction on the anatomical location axis; and for PTs such as ‘Renal failure’, ‘Respiratory failure’ that could be returned by query 2 modulo the suppression of the restriction on the clinical course axis (‘acute’ character).

Results of terms groupings

MedDRA terms returned by Query 1 match exactly the content of the HLT taken as gold standard (see Table 3/Figure 1). This result confirms the hypothesis that the modeling of MedDRA terms through methods of knowledge engineering and DL-queries allows to automatically realize groups of terms similar to manually grouped terms in this terminology. However Queries 1 and 2 were not sufficient to catch the terms of the SMQ. This suggests that a selection of case reports in a database would be different depending on whether we use the SMQ or a Query.

The MedDRA SMQs contain terms that allow consideration of approximate encodings. For example the PT ‘Shock’ is introduced in the narrow part of the SMQ but is not present in the HLT. In this case the term ‘shock’ has a more general scope than the medical condition anaphylactic shock because the causative factor is left without further specification. Other examples are the PTs ‘respiratory failure’ and ‘renal failure’ which are not selected in Query 2 because of imprecision about their course; the query catches terms such as ‘acute respiratory failure’ and ‘acute renal failure’ that add an extra level of information on course. According to the SMQ documentation “Terms representing chronic conditions were generally excluded”. Anaphylactic shock is a phenomenon of limited duration and terms qualified as “acute” should be preferred which is not necessarily the case when coding.

There are several kinds of shock that can be classified according to etiology. Compared to the SMQ, Query 2 adds 14 supplementary terms related to other causes of shock:

- Hypovolemic (PTs ‘hypovolaemic shock’, ‘shock hemorrhagic’): rapid fluid loss (usually blood)
- Traumatic (PT ‘traumatic shock’): reaction to injury
- Cardiogenic (PTs ‘Acute pulmonary oedema’, ‘Cardiac failure acute’, ‘Cardiogenic shock’, ‘Cor pulmonale acute’): decreased pumping ability of the heart

- Septic (PTs ‘Endotoxic shock’, ‘Septic shock’, ‘Toxic shock syndrome’, ‘Toxic shock syndrome staphylococcal’, ‘Toxic shock syndrome streptococcal’): severe infection and sepsis (usually caused by endotoxin-producing gram-negative bacilli)
- Neurogenic (PT ‘Neurogenic shock’): injury to the spinal cord

In order to improve specificity it would be useful to distinguish between terms that may be related to drugs (e.g., anaphylactic shocks) and terms that are clearly not related to drugs such as septic, neurogenic and traumatic shocks. Hypovolemic and cardiogenic shocks may be related to drugs but are not the consequence of an allergic reaction. However such a distinction is difficult to objectify in a query because the way MedDRA terms are defined in OntoADR does not allow to attend such a level of semantic precision. The MedDRA term ‘anaphylactic shock’ is not defined in OntoADR as potentially caused by drugs. Conversely, the MedDRA term ‘septic shock’ is not defined in OntoADR as generally not caused by drugs. Such kind of medical knowledge lacks in OntoADR as it lacks in SNOMED-CT or in most of current biomedical ontologies.

Perspectives

In another work, Kadoyama²⁰ studied the statistical signal of hypersensitivity with anticancer drugs using the FDA database. Hypersensitivity is a wider medical condition than anaphylaxis, as it includes severe anaphylactic reactions, but also mild reaction such as flushing and itching. The authors used the hypersensitivity terms from the National Cancer Institute - Common Terminology Criteria for Adverse Events (NCI-CTCAE) terminology and mappings to corresponding MedDRA LLTs. We plan to extend our current queries to hypersensitivity using OntoADR on anticancer drugs.

Our study focuses on a single safety topic and we plan to make such analysis on other safety topics. This will allow us to evaluate how groupings compare to single preferred terms in signal detection. A safety signal is only a starting point – something to draw the attention of a pharmacovigilance professional and a prompt to explore further a possible drug-event causal association. The actual value of groupings is their ability to gather cases of interest, and the querying method within OntoADR is promising to enable fast generation of groups of terms in order to select case reports in pharmacovigilance databases. So, we plan to make comparison between cases/data retrieved by the queries and cases retrieved by the SMQ in terms of the ability of the user to make a scientific assessment of the potential of an association between an event and a drug.

Also, the use of OWL-DL queries by pharmacovigilance professionals seems impractical. This is why we are currently developing a user interface to facilitate queries and selection of terms. A first effort of this kind is already available in the tool PharmARTS²¹ which is used to represent queries and their results.

Acknowledgments

This work was supported by funding from the European project PROTECT Pharmacoepidemiological Research on Outcomes of Therapeutics by a European Consortium) (<http://www.imi-protect.eu/>). Grant agreement N°115004. We acknowledge Eric Sadou, Adrien Fanet and Anne Jamet who contributed to the development of OntoADR.

References

1. Meyboom RHB, Egberts ACG, Edwards IR, Hekster YA, de Koning FHP, Gribnau FWJ. Principles of signal detection in pharmacovigilance. *Drug saf* 1997;16(6):355-65.
2. Waller PC, Lee EH. Responding to drug safety issues. *Pharmacoepidemiol Drug Saf* 1999;8:535-52.
3. Edwards IR. Adverse drug reactions: finding the needle in the haystack 1997;315(7107):500.
4. Hauben M, Bate A. Decision support methods for the detection of adverse events in post-marketing data. *Drug Discov Today*. 2009 Apr;14(7-8):343-57.
5. Mozzicato P. MedDRA: an overview of the medical dictionary for regulatory activities. *Pharmaceut Med* 23:65–75
6. Hauben M, Patadia VK, Goldsmith D. What counts in data mining? *Drug Saf* 2006;29(10):827-32.
7. Brown EG (2002). Effects of coding dictionary on signal generation: a consideration of use of MedDRA compared with WHO-ART. *Drug saf*;25(6):445-52.
8. Lehman HP, Chen J, Gould AL et al. An evaluation of computer-aided disproportionality analysis for post-marketing signal detection. *Clin Pharmacol Ther* 2007;82(2):173-80.

9. Pearson RK, Hauben M, Goldsmith DI, Gould AL, Madigan D, O'Hara DJ, Reisinger SJ, Hochberg AM. Influence of the MedDRA hierarchy on pharmacovigilance data mining results. *Int J Med Inform* 2009;78(12):e97-e103.
10. Yuen N, Fram D, Vanderwall D, Almenoff J. Do Standardized MedDRA Queries Add Value to Safety Data Mining? ICPE 2008, August 17-20, 2008, Copenhagen, Denmark
11. Bousquet C, Lagier G, Lillo-Le Louët A, Le Beller C, Venot A, Jaulent MC. Appraisal of the MedDRA conceptual structure for describing and grouping adverse drug reactions. *Drug Saf.* 2005;28(1):19-34.
12. Henegar C, Bousquet C, Lillo-Le Louët A, Degoulet P, Jaulent MC. Building an ontology of adverse drug reactions for automated signal generation in pharmacovigilance. *Comput Biol Med.* 2006 Jul-Aug;36(7-8):748-67.
13. Declerck G. PROTECT WP3 – Sub-Package 6 - Novel techniques for grouping ADRs to improve signal detection - Milestone M26 - MedDRA mapping completed for all MedDRA terms relevant for the 13 selected safety topics. 2011.
14. Adverse Event Reporting System. Center for Drug Evaluation and Research, US Food and Drug Administration. Available at: <http://www.fda.gov/Drugs/GuidanceComplianceRegulatoryInformation/Surveillance/AdverseDrugEffects/default.htm>. Last accessed: 7 March 2012.
15. Kessler D. Introducing MedWatch: a new approach to reporting medication and device adverse effects and product problems. *JAMA* 1993;269:2765-8.
16. Introductory Guide MedDRA, MSSO, 2012, http://meddramsso.com/files_acrobat/intguide_15_0_English_update.pdf
17. DuMouchel W (1999) Bayesian data mining in large frequency tables, with an application to the FDA Spontaneous Reporting System (with discussion), *The American Statistician* 1999; 53:177-202.
18. Bate A, Lindquist M, Edwards IR, Olsson S, Orre R, Lansner A, De Freitas RM. A Bayesian neural network method for adverse drug reaction signal generation. *Eur J Clin Pharmacol* 1998;54(4):315-21.
19. Evans SJW, Waller P.C, Davis S. Use of proportional reporting ratios (PRRs) for signal generation from spontaneous adverse drug reaction reports. *Pharmacoepidemiol Drug Saf* 2001.;10: 483–486.
20. Kadoyama K, Miki I, Tamura T, Brown JB, Sakaeda T, Okuno Y. Adverse event profiles of 5-fluorouracil and capecitabine: data mining of the public version of the FDA Adverse Event Reporting System, AERS, and reproducibility of clinical observations. *Int J Med Sci* 2012;9(1):33-9.
21. Alecu I, Bousquet C, Degoulet P, Jaulent MC. PharmARTS: terminology web services for drug safety data coding and retrieval. *Stud Health Technol Inform* 2007;129(Pt 1):699-704.