Evaluation of automated term groupings for detecting upper gastrointestinal bleeding signals for drugs

Julien Souvignet, Gunnar Declerck, Marie-Christine Jaulent and Cédric Bousquet

INSERM U872, Eq. 20, Paris, France
Department of Public Health and Medical Informatics, CHU University of Saint Etienne, France

Introduction

Case reports are usually coded with the MedDRA® terminology (Medical Dictionary for Drug Regulatory Activities) and stored in databases that constitute putative knowledge on adverse drug reactions (ADRs). Groupings of MedDRA terms already exist and are called Standardized MedDRA Queries (SMQ) [1]. They are made manually by experts, but do not cover all medical conditions or may not have the required specificity to cover a safety topic. To support the generation of new SMQs in an automated way, we have developed an OWL-DL (Web Ontology Language – Description Logic) representation of MedDRA named OntoADR. The goal of the present study is to compare results between the SMQ and our DL-query based MedDRA terms grouping method by performing an evaluation on the ‘Upper gastrointestinal bleeding’ safety topic.

1. Method

OntoADR contains formal definitions of ADRs based on MedDRA terms and semantic properties from SNOMED-CT® (Systematized Nomenclature of Medicine – Clinical Terms) such as: hasFindingSite (body site affected by a condition) or hasAssociatedMorphology (morphologic changes of a disease). To define OntoADR concepts, we used UMLS (Unified Medical Language System) metathesaurus to extract mappings between MedDRA and SNOMED-CT with a methodology already described in [2]. When the formal definition couldn’t be built with this method, it was achieved manually by experts. Two queries were developed to match the safety topic. Query 1 targets hemorrhage in the upper gastrointestinal tract structure.

\[
\text{hasFindingSite some 'Upper gastrointestinal tract structure'}
\ AND \ \text{hasAssociatedMorphology some 'Hemorrhage'} \quad \text{(Query 1)}
\]

Query 2 aims to add investigations and findings:

\[
\begin{align*}
\text{\{Query 1\}} & \ OR \\
\text{Interprets some 'Occult blood screening'} & \ AND \ \text{hasInterpretation some 'Positive'} \ OR \\
\text{Interprets some 'Evaluation of stool specimen'} & \ AND \ \text{hasAssociatedMorphology some 'Hemorrhage'}
\end{align*}
\quad \text{(Query 2)}
\]

To build the reference grouping (gold standard), we manually selected terms related to the upper part of the gastrointestinal tract from ‘Gastrointestinal hemorrhage’ SMQ (27 out of 50).
2. Results

The content of Query 1 and SMQ selected terms was similar (recall: 74.1%; precision: 83.3%). Seven terms present in SMQ were absent from Query 1 but five of them where caught in Query 2 (recall: 92.6%; precision: 86.2%) (see Figure 1). Both queries also proposed four additional terms.

3. Discussion

Almost all MedDRA terms of the SMQ grouping taken as gold standard were returned by Query 2. This result confirms the hypothesis that the modeling of MedDRA terms allows to automatically generate lists of terms comparable to manually grouped terms.

Two MedDRA terms were not retrieved by our queries: ‘Duodenal operation’ which refers to a procedure possibly related to upper gastrointestinal bleeding treatment; ‘Ulcer haemorrhage’ which is not semantically defined as a gastro-intestinal related ulcer in MedDRA. And amongst the four additional terms retrieved by the two queries (see Figure 1), only ‘Erosive duodenitis’ describes a potential ADR. Other terms are not a priori related to drugs.

We are currently developing a user interface to assist pharmacovigilance professionals in the creation of such query-based MedDRA terms groupings.

4. Acknowledgments

This work was supported by funding from the European project PROTECT (http://www.imi-protect.eu/). Grant agreement N°115004.

References