LEADING ARTICLE

Drug Safety 2005; 28 (1): 19-34
0114-5916/05/0001-0019/$34.95/0

Appraisal of the MedDRA Conceptual Structure for Describing and Grouping Adverse Drug Reactions

Cédric Bousquet,[1,2] *Georges Lagier*,[3] *Agnès Lillo-Le Louët*,[1] *Christine Le Beller*,[1] *Alain Venot*[4] and *Marie-Christine Jaulent*[2]

1 Centre Régional de Pharmacovigilance, Hôpital Européen Georges Pompidou, Assistance Publique-Hôpitaux de Paris, Paris, France
2 Santé Publique et Informatique Médicale, INSERM ERM202, UFR Broussais-Hôtel-Dieu, Paris, France
3 Centre Régional de Pharmacovigilance, Hôpital Fernand Widal, Assistance Publique-Hôpitaux de Paris, Paris, France
4 Laboratoire d'Informatique Médicale et Bioinformatique (LIM & BIO), UFR de Santé, médecine et biologie humaine Léonard de Vinci, Bobigny, France

Abstract

Computerised queries in spontaneous reporting systems for pharmacovigilance require reliable and reproducible coding of adverse drug reactions (ADRs). The aim of the Medical Dictionary for Regulatory Activities (MedDRA) terminology is to provide an internationally approved classification for efficient communication of ADR data between countries. Several studies have evaluated the domain completeness of MedDRA and whether encoded terms are coherent with physicians' original verbatim descriptions of the ADR.

MedDRA terms are organised into five levels: system organ class (SOC), high level group terms (HLGTs), high level terms (HLTs), preferred terms (PTs) and low level terms (LLTs). Although terms may belong to different SOCs, no PT is related to more than one HLT within the same SOC. This hierarchical property ensures that terms cannot be counted twice in statistical studies, though it does not allow appropriate semantic grouping of PTs. For this purpose, special search categories (SSCs) [collections of PTs assembled from various SOCs] have been introduced in MedDRA to group terms with similar meanings. However, only a small number of categories are currently available and the criteria used to construct these categories have not been clarified.

The objective of this work is to determine whether MedDRA contains the structural and terminological properties to group semantically linked adverse events in order to improve the performance of spontaneous reporting systems.

Rossi Mori classifies terminological systems in three categories: first-generation systems, which represent terms as strings; second-generation systems, which dissect terminological phrases into a set of simpler terms; and third-generation systems, which provide advanced features to automatically retrieve the position of new terms in the classification and group sets of meaning-related terms.

We applied Cimino's desiderata to show that MedDRA is not compatible with the properties of third-generation systems. Consequently, no tool can help for the automated positioning of new terms inside the hierarchy and SSCs have to be entered manually rather than automatically using the MedDRA files. One solution could be to link MedDRA to a third-generation system. This would allow the current MedDRA structure to be kept to ensure that end users have a common view on the same data and the addition of new computational properties to MedDRA.

Standardisation and classification of medical language data are dependent on the use of controlled medical vocabularies. In the pharmacovigilance setting, adverse drug reactions (ADRs) are coded using terminological systems such as the WHO Adverse Reaction Terminology (WHO-ART), Coding Symbols for a Thesaurus of Adverse Reaction Terms (COSTART) and Hoechst Adverse Reactions Terminology System (HARTS). The Medical Dictionary for Drug Regulatory Affairs (MEDDRA), derived from the Adverse Drug Reactions Online Information Tracking (ADROIT) terminology, was accepted as the basis for development of a new terminology at the fourth International Conference on Harmonisation for the Registration of Pharmaceuticals for Human Use (ICH 4) in July 1997. MEDDRA was extensively modified to become MedDRA[1], the Medical Dictionary for Regulatory Activities project.[1] MedDRA is now maintained, developed and distributed by the Maintenance and Support Services Organization (MSSO).

The purpose of MedDRA is to describe the steps of drug development and regulatory issues about human exposure. It includes terms for the description of ADRs, indications, signs and symptoms, family history, investigations and surgical procedures. It is used in the US FDA Adverse Event Reporting System (AERS) and Vaccine Adverse Event Reporting System (VAERS) databases, the European Medicines Agency (EMEA) Eudrawatch system and the Japanese prescription-event monitoring system. Since January 2003, all electronic exchange of case reports of ADRs in the EU, Iceland, Liechtenstein and Norway should be coded with MedDRA.

The computational representation of terminological resources for medical records and medical statistics is an active field of research in the medical informatics community. Most items in the patient record are described using medical terms. Physicians can naturally understand the meaning of terms and exchange this meaning with colleagues because of their knowledge of medicine. When information is transferred from the paper-based patient record to the electronic patient record, the meaning of the information may be lost in different computational operations. Retaining the meaning of information in automated systems relies on a formal representation of medical information. Formal representation consists of a logical language to perform inference reasoning; for example, "every chronic arthritis is an arthritis".

The European Committee for Standardization[2] defined a European pre-standard for the categorical structures of concept systems. The IMIA WG6 is the International Medical Informatics Association's working group on natural language, classification and concept representation. At the 1997 meeting of the IMIA, terminological issues and solutions were discussed. Cimino proposed 'desiderata' for controlled medical vocabularies in the 21st century.[3] The most important feature of a terminological system is domain completeness: it implies that every term necessary for describing a knowledge domain may be available for the end user. Chute et al.[4] evaluated the domain completeness of

1 MedDRA is a registered trade mark owned by the International Federation of Pharmaceutical Manufacturers Associations.

major vocabularies by selecting medical language data from a sample of medical records. Terms not found in a controlled vocabulary had to be added. However, Cimino argued that the "add more terms until we are done" approach to terminological work is no longer sufficient.[3]

Much work has been done on ADRs to evaluate the domain completeness of MedDRA to retrieve the original verbatim term used by the physician in the patient chart. However, to date, the structural properties of MedDRA have not been assessed. Our objective is to determine the extent to which the structure of MedDRA complies with Cimino's desiderata. Automated signal detection is a new field in pharmacovigilance. The aim is to identify unknown relationships between drugs and adverse effects using a disproportionality measure. Our long-term objective is to improve the performance of current signal detection methods by grouping terms with similar meaning.

In this article, we investigate whether MedDRA has the necessary terminological requirement to automatically infer relevant groupings of terms. The current terminological systems and standards in medicine, including a description of MedDRA, and the three generations of terminological systems are reviewed. We present Cimino's criteria for new terminological systems and propose these as a methodological basis for the appraisal of MedDRA. Finally, we discuss issues related to the structure of MedDRA and argue that MedDRA has the qualities of a classification system but lacks some properties of clinical nomenclature in medicine. A new methodological approach is proposed to overcome these limitations.

1. Terminological Systems in Medicine

1.1 Classification Versus Nomenclature

A classification is a rigid structure designed to put an item in one (and only one) class of the terminology. The advantage is that the classification is unique and reproducible. However, one of the limitations of classification systems arises when an item can belong to more than one class. For example, *Staphylococcus aureus* sepsis belongs to the sepsis class (a pathological process) and the infectious disease class.

The objective of a nomenclature is to enumerate all items that appear in a knowledge field. At the most basic level, it is a list of names. To capture clinical terms in the electronic patient record, clinical nomenclatures have been introduced to allow fine-grained concept representation. For example, as surgeons need accurate localisation of lesions, a nomenclature of surgical procedures must describe all topographical areas in great detail.

1.2 The MedDRA Classification System

MedDRA terms are organised in five levels (figure 1): system organ class (SOC), high level group term (HLGT), high level term (HLT), preferred term (PT) and low level term (LLT). The preferred level of analysis is the PT. PTs are classified according to SOCs, which are clinical conditions (medical procedures or patient history) relevant for the pre- and post-marketing of drugs, and SOCs used to describe diseases or findings. PTs are grouped in HLTs and then in HLGTs in the same SOC. There were 26

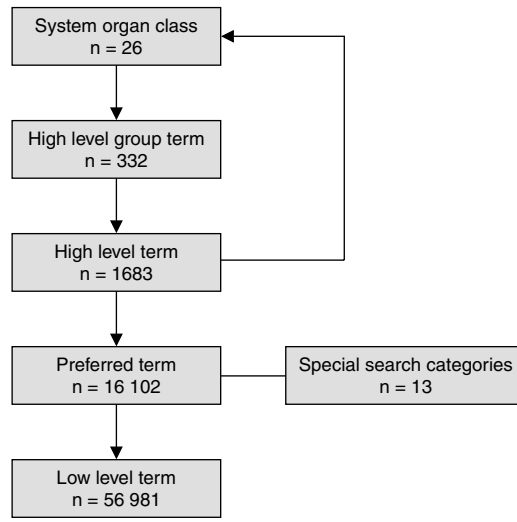

Fig. 1. The Medical Dictionary for Regulatory Activities (MedDRA) classification system comprises five levels: system organ class, high level group term, high level term, preferred term and low level term. The number of terms in each level relates to MedDRA version 5.1.

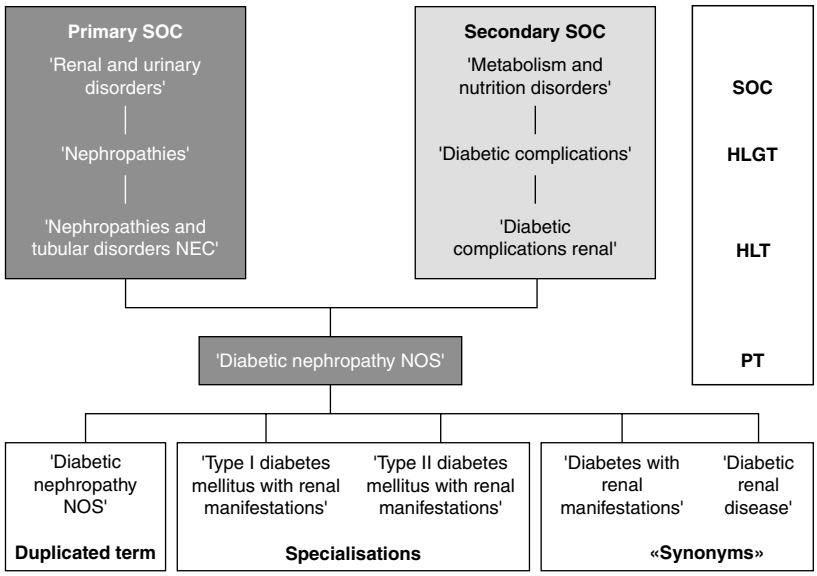

Fig. 2. Diabetic nephropathy in the Medical Dictionary for Regulatory Activities (MedDRA). **HLGT** = high level group term; **HLT** = high level term; **NEC** = not elsewhere classified; **NOS** = not otherwise specified; **PT** = preferred term; **SOC** = system organ class.

SOCs, 332 HLGTs, 1683 HLTs, 16 102 PTs and 56 981 LLTs in MedDRA version 5.1.

LLTs are used for data entry and can be true synonyms. For example, the LLTs 'Diabetes with renal manifestations' and 'Diabetic renal disease' are synonyms of the PT 'Diabetes nephropathy' (figure 2). LLTs may be more granular terms to capture the physician's verbatim term. For example, the LLT 'Type II diabetes mellitus with renal manifestations' is one kind of 'Diabetes nephropathy'. All PTs are duplicated at the LLT level.

Each PT is linked to a primary SOC and can be linked to other secondary SOCs. For example, the primary SOC for 'Diabetic nephropathy' is 'Renal and urinary disorders'. The secondary SOC is 'Metabolism and nutrition disorders'. 'Nephropathies' and 'Diabetic complications' are HLGTs. 'Nephropathies and tubular disorders NEC' and 'Diabetic complications renal' are HLTs.

Not elsewhere classified (NEC) terms have been introduced at the HLT level to indicate a group of PTs that do not fit a given category in the MedDRA hierarchy. They are provided as a set of terms that do not share common semantic features. The only common property is that they do not belong to well specified categories in MedDRA.

A special search category (SSC) is a set of PTs that relate to a common clinical situation but are not necessarily hierarchically related. For example, the 'Haemorrhage' SSC comprises hundreds of terms including all bleeding or haemorrhagic conditions such as 'Gastric ulcer haemorrhage', 'Haemarthrosis' and 'Blood urine present'. MedDRA SSCs are entered manually after reviewing all PTs to evaluate whether they belong to a SSC. Only searching for 'Haemorrhage' is not sufficient because this condition may be expressed by many other terms that do not contain this string, for example 'Haemoptysis', 'Epistaxis', 'Ecchymosis' or 'Menorrhagia'.

1.3 Major Terminological Systems in Medicine

Terminological systems include the International Classification of Diseases (ICD),[5] the Medical Subject Heading (MeSH) developed by the National Library of Medicine (NLM) for the coding of medical literature in the Medline database,

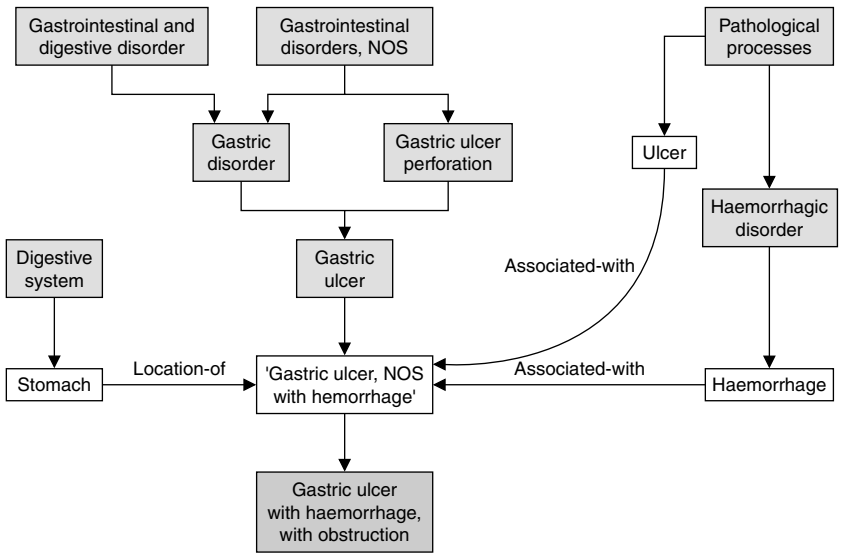

Fig. 3. Definition of 'Gastric ulcer, NOS with haemorrhage' in the Unified Medical Language System (UMLS) semantic network. This concept is associated by semantic links to other concepts: 'stomach', 'haemorrhage' and 'ulcer'. **NOS** = not otherwise specified. Grey boxes = hierarchical relations; white boxes = semantic definition of 'Gastric ulcer, NOS with haemorrhage' (other related concepts and relations in SNOMED International).

and the Systematised Nomenclature of Medicine (SNOMED).[6] The Unified Medical Language System (UMLS)[7] proposes a semantic network developed by the NLM to link terms from 60 controlled vocabularies and provide semantic definitions of terms. For example, the MedDRA term 'Gastric ulcer haemorrhage' is described by a UMLS concept 'Gastric ulcer, NOS (not otherwise specified) with haemorrhage' and is linked to controlled vocabularies such as MeSH, ICD-10 and SNOMED. Figure 3 shows semantic relations between UMLS concepts in the semantic network. For example 'Gastric ulcer, NOS with haemorrhage' is linked to the stomach concept by the 'location-of' semantic relation.

Table I gives more details on the 'Gastric ulcer, NOS with haemorrhage' UMLS concept. UMLS provides synonyms (e.g. 'Bleeding gastric ulcer'), terms from other languages (e.g. 'Ulcere gastrique hemorragique'), terms from other sources (e.g. 'Stomach ulcer haemorrhage' in COSTART) and other related concepts (e.g. 'Haemorrhage, stomach' from the SNOMED nomenclature).

2. The Three Generations of Terminology Systems

Rossi Mori et al.[8] define three generations of terminological systems according to concept representation systems. This typology is based on the 'model for representation of semantics in medicine' project of the European Committee for Standardization.[2] As medicine is organised into various domains of expertise (surgery, obstetrics, psychiatry etc.), each element of a domain is described with a limited number of semantic categories. The European Committee for Standardization introduced a categorical structure to organise high-level concepts within each domain. For example, the term *Streptococcus pneumoniae* can be attached to the semantic category 'living agent' in the infectious diseases domain.

In the following subsections, we examine the maintenance of terminological systems, properties of data retrieval systems and difficulties with implementation in the three generations of terminological systems.

Adding new terms:

Table I. Details of the Unified Medical Language System concept 'Gastric ulcer, NOS with haemorrhage'

CUI
C0151982
Semantic type
Disease or syndrome
Synonyms
Bleeding gastric ulcer
Gastric ulcer haemorrhagic
Ulcer stomach with haemorrhage
Other languages
Ulcere gastrique hemorragique (French)
Magengeschwuer blutend (German)
Ulcera gastrica sanguinante (Italian)…
Other sources
Stomach ulcer haemorrhage (COSTART)
Bleeding gastric ulcer (Read codes)
Gastric ulcer, NOS with haemorrhage (SNOMED Intl 1998)
Gastric ulcer haemorrhagic (WHO-ART)…
Other related concepts
Haemorrhage (SNOMED Intl 1998) [relation: associated-with]
Stomach (SNOMED Intl 1998) [relation: location-of]
Ulcer (SNOMED Intl 1998) [relation: associated-with]
Related and possibly synonymous concepts
Peptic ulcer haemorrhage
Gastric ulcer with melena…

COSTART = Coding Symbols for a Thesaurus of Adverse Reaction Terms; **CUI** = concept unique identifier; **Intl** = international; **NOS** = not otherwise specified; **SNOMED** = Systematised Nomenclature of Medicine; **WHO-ART** = WHO Adverse Reaction Terminology.

- What happens when new concepts have to be added to the terminological system?
- Does the terminological system provide tools to specify the position of the new concept in the system and detect inconsistencies or redundancies in the system?

Data retrieval:

- Does the terminological system include features to help refine user queries using differentiating criteria, for example all diseases caused by *S. pneumoniae*?
- Does the terminological system include features to help refine user queries using semantic relatedness, for example, all diseases defined by haemorrhagic conditions in the upper digestive tract?

Software requirements:

- Is the system end-user friendly and does it require special software?

2.1 First Generation: Paper-Based Systems

First-generation terminological systems such as the ICD are based on textual descriptions of concepts. No categorical structure is provided and concepts are designated by codes and strings. These traditional systems are paper based but can be electronically available to allow the storage, transmission and retrieval of strings and codes attached to the concepts. This type of vocabulary has a fixed and usually unique hierarchy devoted to a single application. The vocabulary cannot be dynamically increased and users rely on a predefined list of expressions.

Data retrieval is restricted to the exact matching of strings and codes. For example, the concept gastric ulcer haemorrhage can be coded in the ICD-9 CM in 'Diseases of the digestive system' with the codes 531.0 (acute form) or 531.4 (chronic form) [figure 4].

In the ICD, the asterisk (*) and dagger (†) approach is used to retrieve a diagnosis when the statement contains information about both an underlying disease and a manifestation in a particular organ or site. For example, tuberculosis of the spinal column may be coded as a dagger code A18.0† (Chapter I – Certain infectious and parasitic diseases) and as an asterisk code M49.0* (Chapter XIII – Diseases of the musculoskeletal system and connective tissue).

The vocabulary can be installed as a flat file. Tools help users to find codes or navigate in the classification. Queries are limited to predefined classes and differentiating criteria are not allowed.

2.2 Second Generation: Compositional Systems

Second-generation systems are compositional systems, such as SNOMED, that are built using a categorical structure and a cross thesaurus. The categorical structure is composed of a set of metaterm descriptors to describe a concept in a domain of expertise. For example, <morphology>, <living agent> and <topography> are descriptors for the categorical structure of infectious diseases. Gastric ulcer haemorrhage is an example of a 'molecular'

(520-579) Diseases of the digestive system
…

530 Diseases of esophagus
…

531 Gastric ulcer

531.0 Acute gastric ulcer with haemorrhage
531.1 Acute gastric ulcer with perforation
531.2 Acute gastric ulcer with haemorrhage & perforation
531.3 Acute gastric ulcer without mention of haemorrhage or perforation
531.4 Chronic or unspecified gastric ulcer with haemorrhage
531.5 Chronic or unspecified gastric ulcer with perforation
531.6 Chronic or unspecified gastric ulcer with haemorrhage & perforation
531.7 Chronic gastric ulcer without mention of haemorrhage or perforation
531.9 Gastric ulcer, unspecified as acute or chronic, without mention of haemorrhage or perforation

532 Duodenal ulcer
…

Fig. 4. Description of gastric ulcer haemorrhage in the ninth edition of the International Classification of Diseases (ICD-9 CM) classification.

terminological phrase that may be dissected into basic units. Terms that cannot be further dissected are 'atomic' terms, for example, ulcer, stomach, haemorrhage. The cross thesaurus is a multiaxial thesaurus that provides the atomic terms to enter descriptors from the categorical structure.

Gastric ulcer haemorrhage should be described by the following dissection:

disease that

has aspect: haemorrhage

has location: stomach

has cause: ulcer

Relationships such as 'has aspect' and 'has cause' are semantic links. These descriptions comply with the multiaxial representation of SNOMED (figure 5). This type of vocabulary can be easily extended by end users by adding new atomic terms or by creating new complex terms according to the compositional properties of this system.

The structure of second-generation systems processes the semantics of terms by clustering the terminological phrases according to criteria. Specialised software is needed to make queries that exploit the properties of compositional definitions of terms.

2.3 Third Generation: Formal Systems

Using a formal approach for the definition of terms is an active field of research in terminology. The formal models constitute third generation systems such as Generalized Architecture for Lan-

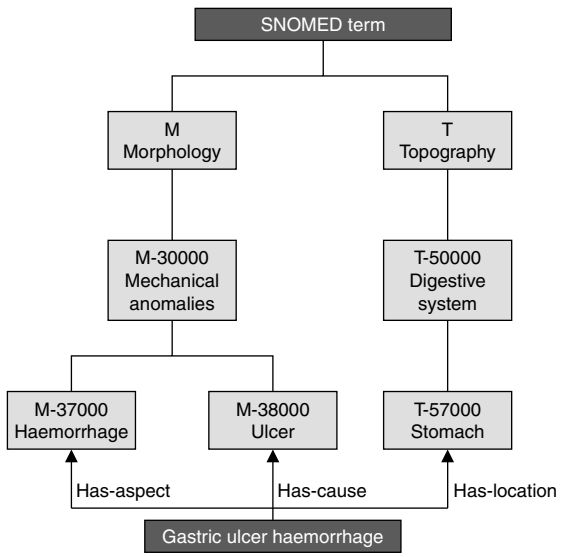

Fig. 5. Gastric ulcer haemorrhage in the systematised nomenclature of medicine (SNOMED). In SNOMED, a composite term can be represented by dissection on different axes. Gastric ulcer haemorrhage is linked to the topography axis (stomach) and the morphology axis (haemorrhage, ulcer) of SNOMED.

guages, Encyclopedias and Nomenclatures (GALEN).[9] In formal taxonomic systems, terms are described using a logical language and the position of terms in the hierarchy is found by computing subsumption relations. For example, 'Gastric ulcer haemorrhage' is a 'haemorrhage' caused by an 'ulcer' localised in the 'stomach'. The following example illustrates how a declarative form of logic statements may be applied to define the meaning of 'Gastric ulcer haemorrhage' (GUH) and describe its properties.

The linguistic expression 'GUH is a haemorrhage' is written 'haemorrhage (GUH)'.

The expression 'any GUH is a haemorrhage' is represented as (equation 1):

$$\forall \text{ x GUH (x)} \Rightarrow \text{haemorrhage (x)}$$

where $\forall$ is the universal quantifier and $\Rightarrow$ the logical operator 'implies'.

The expression 'any GUH is localised in the stomach' is written as follows (equation 2):

$$\forall \text{ x GUH (x)} \Rightarrow \exists \text{ y stomach (y)} \wedge \text{located-in (x,y)}$$

where $\exists$ is the existential quantifier and, $\wedge$ the logical operator 'and'.

The expression 'any GUH is caused by an ulcer' is (equation 3):

$$\forall \text{ x GUH (x)} \Rightarrow \exists \text{ z ulcer (z)} \wedge \text{caused-by (x,z)}$$

Equations 1–3 are sufficient and necessary conditions to describe GUH. This can be represented with the logical operator $\Leftrightarrow$ (if and only if) [equation 4]:

$$\forall \text{ x GUH (x)} \Leftrightarrow \text{haemorrhage (x)}$$
$$\wedge \exists \text{ y stomach (y)} \wedge \text{located-in (x,y)}$$
$$\wedge \exists \text{ z ulcer (z)} \wedge \text{caused-by (x,z)}$$

This type of formula is very hard to read. Sowa[10] proposed a more convenient notation called a conceptual graph in which concept nodes are represented by boxes and concept relationships are denoted by circles (figure 6).

A new concept can be automatically added to the hierarchy by computing subsumption relations. For example, GUH is a child of 'gastric ulcer' (GU) and 'gastric haemorrhage' (GH) [equation 5; equation 6].

$$\forall \text{ x GUH (x)} \Rightarrow \text{GH (x)}$$

$$\forall \text{ x GUH (x)} \Rightarrow \text{GU (x)}$$

Data retrieval benefits from this multi-inheritance mechanism. For instance, queries in both GU and GH categories retrieve the GUH concept.

Third-generation systems are more powerful but require highly specialised software to process formal definitions. Terminology servers implement services available to client applications through an application-programming interface.[9] Editors are available to define formal concepts and inference engines are necessary to perform formal reasoning on these definitions. However, the formal definition of terms is achieved after intensive knowledge acquisition and modelisation. Good technical skills are mandatory to take advantage of these tools. Table II summarises the main procedures to add new terms and retrieve data, as well as software requirements for the types of language used in the three generations of systems.

3. Cimino's Description of Terminology Requirements

The development of controlled vocabularies for terminology systems requires a standardised approach to provide quality for end users. The standardisation process has identified some common desiderata, which have been reviewed by Cimino.[3]

- Vocabulary Content: The first and most important desideratum is domain completeness because users always need more content. One approach is to add terms to the terminology as they

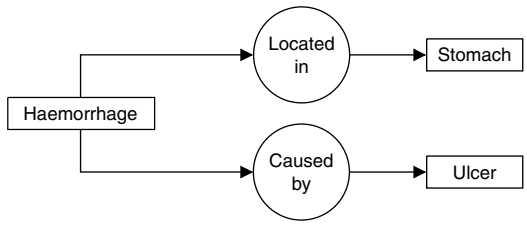

Fig. 6. Conceptual graph of 'Gastric ulcer haemorrhage'. Concepts are represented by boxes and relationships by circles. 'Gastric ulcer haemorrhage' is a haemorrhage that is located in the stomach and caused by inflammation.

Table II. Main features according to generation types of languages

Features	First generation: paper-based systems	Second generation: compositional systems	Third generation: formal systems
Adding new terms	Predefined list of allowed expressions	Creation of new terms by composition of atomic terms	Addition of new concepts by subsumption
Data retrieval	Predefined classes and no differentiating criteria	Clustering of terminological phrases according to criteria	Clustering of terms by subsumption
Software requirements	Flat file	Specialised software	High: terminology servers, concept editors

are encountered (first generation). This means that developers add complex expressions rather than preparing the ground for a compositional approach. The compositional approach is based on second- and third-generation systems. For Cimino, a compositional and formal approach is mandatory to avoid the addition of atomic terms in an exponential way.

- Concept Orientation: The concept orientation requirement means that terms must correspond to at least one meaning (non-vagueness) and no more than one meaning (non-ambiguity). In addition, a concept should not correspond to more than one term in the terminology (additional terms that reference this concept should be recognised as synonyms). Although one concept can convey different meanings according to the context, it must possess only one definition in the context defined by the controlled vocabulary.
- Concept Permanence: The concept permanence requirement is a corollary of the precedent. It implies that once a concept is created, its meaning is inviolate. Its PT can be marked as inactive, but it must be kept to ensure adequate behaviour when data have been coded with an older version of the vocabulary.
- Non-Semantic Concept Identifiers: A concept must have a unique identifier to allow each term to be mapped with a single concept. This non-semantic concept identifier should be independent of the meaning of the concept and its position in the hierarchy.
- Polyhierarchy: Most available vocabularies are strict hierarchies. Vocabulary developers disagree about whether a single concept should be described according to a single hierarchy or to multiple hierarchies. Cimino favours the latter because medical concepts can often be described according to multiple views or body organs.
- Formal Definitions: Semantic definitions of terms should be available in operational form on the computer. This implies that developers have to provide formal definitions of terms with knowledge representation and inference features.
- Rejection of «NEC» Terms: A concept should be defined in a unique way that is independent of the presence or absence of related terms (elsewhere classified) because this kind of definition groups together miscellaneous terms that do not share a common definition at the semantic level.
- Multiple Granularities: The multiple granularity feature allows all users to express concepts with sufficient details according to the task. For example, a surgeon needs a very accurate description of a patient's lesions whereas this is not the case for a general practitioner in charge of overall healthcare.
- Multiple Consistent Views: Different views of the same terminology should be available according to the requirements of the user. For example, the structure of a hierarchy of diagnoses differs if it is built to express an anatomical perspective or an aetiological perspective.
- Context Representation: A vocabulary should be usable independently of the context in which it was first created.
- Graceful Evolution: Evolution of the vocabulary should favour the users rather than the developers of the terminology by providing understandable reasons for changes.
- Recognised Redundancy: If a concept is represented by two different names, recognition that

both terms are in fact synonyms should be easy and automatic.

4. MedDRA Structure Evaluation

We performed an appraisal of the MedDRA structure from three different viewpoints. First, we applied Cimino's criteria to test whether MedDRA complied with the desiderata of a 'terminology in the 21st century'.[3] Second, we studied how MedDRA was classified in the three generation systems framework described by Rossi Mori et al.[8] Third, we evaluated the distance between ADRs with close meanings in the gastrointestinal SOC.

4.1 Compliance of MedDRA to Cimino's Desiderata

The results of our evaluation according to Cimino's desiderata are presented in table III.

The analysis indicated that MedDRA has better domain completeness than previous ADR dictionaries such as WHO-ART because more terms are available for coding ADRs. We did not quantitatively assess the performance of MedDRA to accurately code verbatim terms.

MedDRA is concept oriented through the use of codes as non-semantic identifiers. Terms no longer used in the terminology are not deleted but rather marked as obsolete to allow compatibility with earlier versions of the vocabulary.

Table III. Appraisal of the Medical Dictionary for Regulatory Activities using Cimino's desiderata

Desiderata	Excellent	Fair	Poor	None
Vocabulary content	x			
Concept orientation	x			
Concept permanence	x			
Non-semantic concept identifiers	x			
Polyhierarchy			x	
Formal definitions				x
Rejection of 'not elsewhere classified' terms				x
Multiple granularities		x		
Multiple consistent views			x	
Context representation			x	
Graceful evolution	x			
Recognised redundancy				x

We found that because of the structural design of MedDRA, a PT cannot have more than one HLT in the same SOC; no polyhierarchy exists at this level. For example, there is a missing relationship between the 'Gastric ulcer haemorrhage' PT and the 'Gastric and esophageal haemorrhages' HLT (figure 7). Polyhierarchy is allowed across another SOC. The 'Gastric ulcers and perforations' HLT is linked to the 'Vascular haemorrhagic disorder' HLGT in the 'Vascular disorder' SOC. However, no automated mechanism is available for end users to check that terms with the same meaning are in other parts of the terminology.

MedDRA does not provide formal definitions of terms. Contrary to formal representation models, the ability to express fine clinical details is impaired because MedDRA is limited to terms that are already present in the classification. Although NEC terms in MedDRA are now limited to HLTs, their presence introduces an additional difficulty to exploit classification using automated tools.

MedDRA allows multiple granularities but the depth of the hierarchy is arbitrarily restricted to five levels. Vocabularies based on second- and third-generation systems can organise terms of the hierarchy with an arbitrary number of levels.

MedDRA is organised according to organ classes that are inappropriate for compositional descriptions of terms. It provides multiple consistent views on data through the multi-SOC structure but these views are not based on a parallel hierarchy of defining axes such as in the SNOMED nomenclature. As shown in figure 2, the multi-SOC feature allows 'Diabetic nephropathy' to be linked to the 'Renal and urinary disorder' and 'Nutrition and metabolism disorder' SOCs. In second-generation systems, true multi-axial structure is achieved through the use of dissections. For example, in SNOMED, 'Diabetic nephropathy' is described using the 'Kidney' concept in the topography axis and the 'Diabetes mellitus' concept in the disease axis.

In addition, MedDRA has no context representation. Although SOCs may help to identify the semantic category of a term, MedDRA does not include a feature to ensure that only terms relevant to a

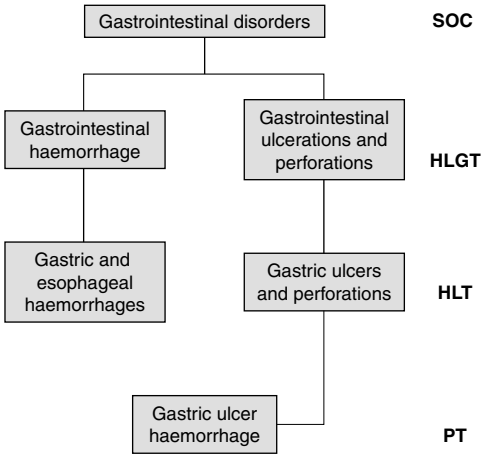

Fig. 7. The Medical Dictionary for Regulatory Activities (MedDRA) terms can be linked to only one high level term (HLT) inside the same system organ class (SOC). For example, 'Gastric ulcer haemorrhage' belongs to the 'Gastric ulcer and perforations' HLT but not to the 'Gastric and esophageal haemorrhage' HLT. **HLGT** = high level group term; **PT** = preferred term.

given data field are available to the end user to code this field. All terms are available to enter the indication for a drug even though these may designate an adverse effect and not be relevant in the context of drug indication.

One main problem is related to maintenance of the MedDRA terminology. Graceful evolution can be completed because of medical evaluation of the introduction of new terms. However, MedDRA does not include an automatic mechanism to recognise redundancy or locate the position of new concepts in the hierarchy.

4.2 Generation System Membership

MedDRA is a first-generation system. End users must rely on a predefined list of allowed expressions provided by the MSSO. Data retrieval is constrained by predefined classes (HLT, HLGT, SOC). A flat file format may be sufficient to implement the vocabulary. However, some editors are already proposing tools to enhance user-friendly navigation inside the MedDRA structure.

4.3 Position of Medically Related Terms

'Haemorrhage' is a characteristic of 'Gastric ulcer haemorrhage' in common with other terms such as 'Esophageal haemorrhage' or 'Esophagitis haemorrhagic'. Figure 8 shows the dispersion of preferred terms (in italic type) in the 'Gastrointestinal and disorders' SOC that express an upper digestive haemorrhagical medical condition.

5. Discussion

5.1 Current Limitations of MedDRA

Cimino's desiderata have been revised and extended by experts in clinical terminology to define the ISO/DTS 17117 standard.[11] In phase I of the Genetic Modification Clinical Research Information System (GEMCRIS) development, this standard was used to evaluate vocabularies (MedDRA, SNOMED, National Cancer Institute Enterprise Vocabulary Services [NCI-EVS], ICD and MeSH) for the online adverse event reporting to the National Institutes of Health (NIH) and the FDA.[12] SNOMED had the most features recommended by ISO/DTS 17117. MedDRA was recognised as an operational vocabulary for regulatory purposes. Authors of the study decided to use a mapping between MedDRA and SNOMED that is half completed.

Huntley and Veverka[13] performed a preliminary test of MedDRA for the US FDA that included a review of MedDRA by medical informatics experts. These experts approved the use of MedDRA as a starting point for a medical terminology but noted that MedDRA did not have a truly multi-axial structure. Although this point was briefly mentioned, the consequence of this has not been addressed in MedDRA evaluations.

As mentioned previously, MedDRA is a first-generation system. It adds flexibility because of multi-SOC and SSCs but these features comply with the first-generation properties. Multi-SOC is not synonymous with multi-axial. Second-generation systems offer multiple points of view depending on a set of dissections that organise the metathesaurus according to different axes in a compositional way.

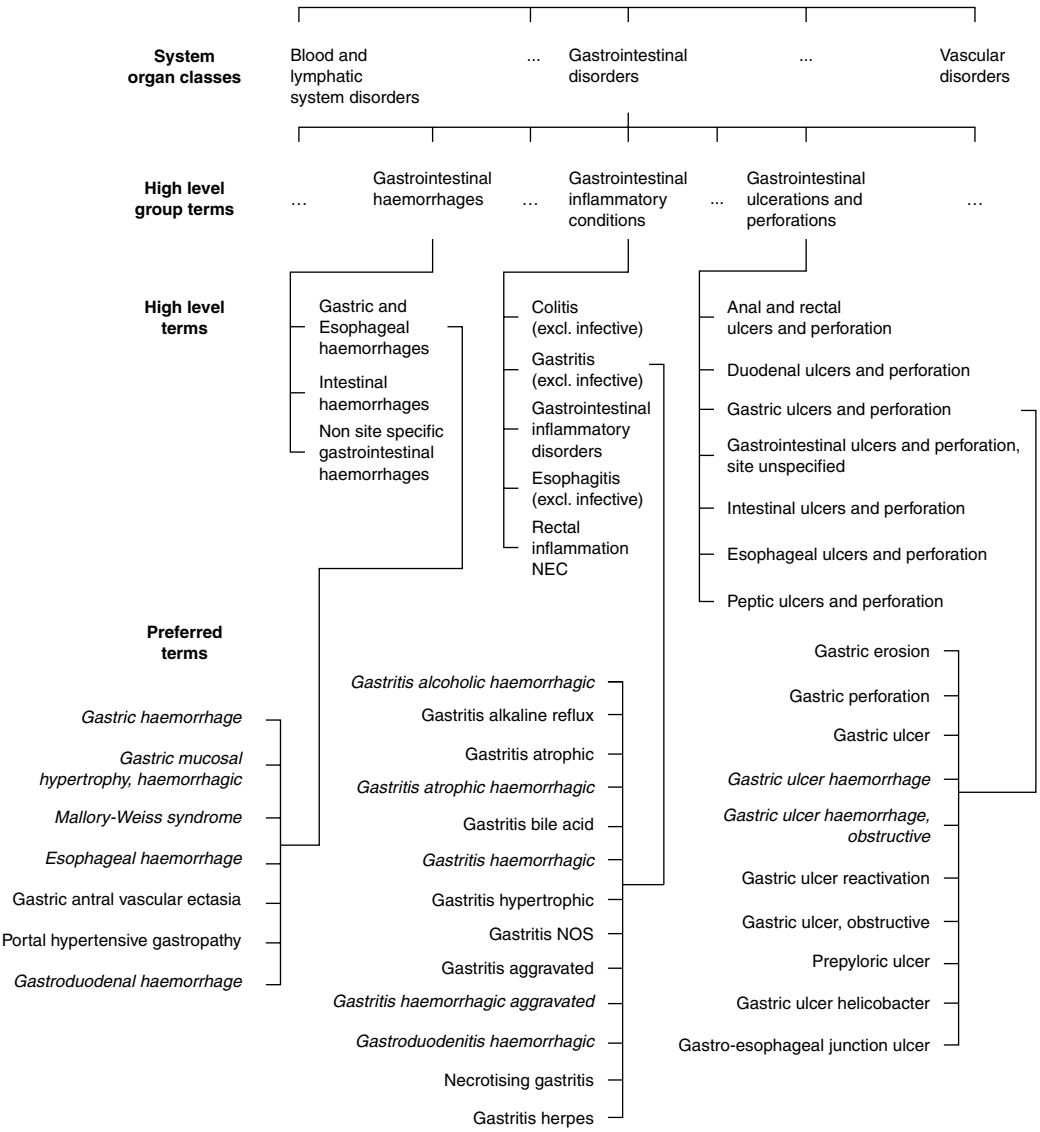

Fig. 8. Portion of the gastrointestinal system organ class in the Medical Dictionary for Regulatory Activities (MedDRA). Low level terms are not shown. Preferred terms in italic type express an upper digestive haemorrhagical medical condition. **Excl** = excluding; **NEC** = not elsewhere classified; **NOS** = not otherwise specified.

Moreover, there is no formal representation of terms in MedDRA such as in third-generation systems.

Brown and Douglas[14] used a data set containing all ADR reports for a single drug from the UK Medicines Control Agency with MedDRA version 2.2. The data from the gastrointestinal SOC highlighted upper gastrointestinal bleeding. However, the number of upper gastrointestinal bleeding events was underestimated when identification of adverse events was restricted to the 'Gastrointestinal haemorrhage' HLGT. It failed to retrieve a case report of 'Gastric ulcer haemorrhage' in the 'Gastrointestinal

ulcerations and perforations' HLGT. Brown and Douglas pointed out this limitation of MedDRA to advise users to carefully check the results provided by the selection of terms attached to a given high-level category.

This limitation is explained by a characteristic of first-generation systems. MedDRA obeys the taxonomic principle that one child has only one parent in the hierarchy. This principle is important for statistical purposes since it avoids counting the same item several times, but it does not support polyhierarchy.

We argue that Brown and Douglas's outcomes are a result of the absence of polyhierarchy inside the same SOC. For example, in the gastrointestinal system, various items showing a haemorrhagic condition are present under different HLGTs (figure 8). Assigning a PT to a unique HLGT leads to misclassification. For example, 'Gastric ulcer haemorrhage' is present in the 'Gastrointestinal ulcerations and perforations' HLGT and not in the 'Gastrointestinal haemorrhage' HLGT (figure 7). Inconsistencies in the hierarchy, such as positioning of a concept in the wrong class or missing links could be identified by computing subsumption relations in a formal representation of the MedDRA structure.

5.2 MedDRA is a Specialised Terminology for Drug Regulatory Activities

5.2.1 Other Multi-Purpose Medical Terminologies Are Available

The purpose of UMLS is to provide a link between different terminologies. Since the recent inclusion of MedDRA in the UMLS, each MedDRA term is linked to a UMLS concept. Properties of third-generation systems may be added to first-generation systems by using these types of links. In an earlier study, we tried to use links of the semantic network to automatically extract semantic definitions of MedDRA terms, but we had only partial success using this strategy.

Strang et al.[15] performed a semi-quantitative evaluation of several terminologies including ICD, UMLS, MedDRA and SNOMED for the indexing requirements of therapeutic systems. For this particular purpose, SNOMED seemed to be the most appropriate coding system, partly because of its compositional properties. Mapping of MedDRA concepts to SNOMED is currently underway.[12]

5.2.2 MedDRA Cannot Represent the Whole Summary of Product Characteristics Information

It is proposed that MedDRA could be used for coding indications if applicable in the electronic transmission of individual case safety reports.[16] In 1996, Brown and Clark[17] evaluated the capability of MedDRA to represent medicinal product data sheet information. Although the exact and acceptable matching of terms was the highest for adverse effects (89.7%), the representation of terms in the indication section was only 62.7%. As for ADRs, we argue that the addition of more terms in MedDRA will not resolve this problem.

The semantic elements of the summary of product characteristics (SPCs) described in natural language are often richer than the labels used to describe pathologies in MedDRA. Indications and contraindications of drugs must rely on a conceptual structure to describe non-terminological information (e.g. 'child under twelve'). For example, the indications label includes information on efficacy of the drug for a disease, diseases for which the drug is not indicated, type of patient (e.g. infant, elderly, pregnant), type of action (e.g. preventive, curative, symptomatic, diagnostic procedure), severity and disease progression.[18]

5.3 MedDRA is Not a Clinical Nomenclature

Terms may be organised according to different structures including classification systems, thesaurus and clinical nomenclatures. Each structure presents advantages and drawbacks, and the choice depends on whether the purpose is statistical study or coding of patient data.[19] MedDRA is not a 'clinical nomenclature' because its purpose is not to capture all details of the patient chart. Continuing with the 'classification versus nomenclature' controversy, Ingenerf and Giere asserted that Cimino's desiderata are not valid for classification-oriented vocabularies.[20] The description of clinical data with a sufficient atomic-level granularity implies a formal representation of terms within a nomencla-

ture. However, we need a strict classification structure that provides a set of disjunctive classes to assign each term to one (and only one) purpose-dependent predefined class and that is reproducible. MedDRA belongs to this last category.

MedDRA could be used for statistical purposes and linked to a third-generation system in order to define MedDRA terms at the semantic level. An example of such a strategy is the two-layered approach of the Classification Commune des Actes Médicaux (CCAM) terminology for coding of surgical procedures in France.[21] CCAM terms are available to end users through a multi-axial classification system. The conceptual representation developed with the GALEN server adds a semantic to CCAM terms, allowing computer processing of patient records coded with CCAM.

5.4 MedDRA Improves Accuracy of Adverse Drug Reaction Descriptions

The section 'term selection: points to consider' in MedDRA[22] is representative of some present limitations of MedDRA. Section 3.5, 'body site vs event specificity', states that some terms describe both body site location and event information but others give only event information. For example, it is possible to select 'rash on face' but 'rash on chest' is not part of the terminology.

However, a reference terminology is built for the purpose of representing medical data in the patient chart. In pharmacovigilance, detail requirements are not the same as in the clinical setting. Physicians need very accurate descriptions of the patient's characteristics and problems. For example, just coding 'stenosis of the left coronary artery' may be sufficient for the general practitioner but does not provide sufficient details for the cardiologist. On the other hand, 'stenosis of the left coronary artery' provides too much detail for the purpose of coding ADRs.

Achieving domain completeness is an important issue but this feature must not be confused with the perfect translation of the original verbatim or physician description of the adverse effect. Topography is not always an important issue for coding adverse events. For example, providing the localisation of an oedema may indicate the severity and physiopathological state of the event. Oedema of the eyelids is not comparable to an oedema of the legs. As an adverse event, the localisation of a rash is generally irrelevant. Actually, the purpose of MedDRA is to give sufficient detail for the description of classes of adverse effects such as rash and not to reflect the clinically relevant description of a medical event as stated by the physician in the patient chart.

5.5 We Need Tools to Identify Related Concepts

5.5.1 Grouping of Related Terms in High-Level Categories

PTs should be grouped in some situations because they are too specific to generate a signal effectively. For example, Yokotsuka et al.[23] decided to group 'alanine aminotransferase increased', 'aspartate aminotransferase increased' and 'liver function tests abnormal NOS'. SSCs do not address important issues for the management of ADRs, drug development and regulatory processes. These SSCs should be broad enough to achieve sensitivity at the expense of specificity but the specificity of current SSCs is too low.[14]

Two new sets are currently being developed to provide more categories of PTs for end users. The MSSO has proposed MedDRA analytical groupings (MAGs).[24] MAGs group terms from different parts of the hierarchy but are not restricted to PTs. An HLT or HLGT can be included in a MAG after exclusion of one or two PTs that are not relevant in the context. A new Council for International Organizations of Medical Sciences (CIOMS) group is working on the definition of special search queries (SSQs) to identify case series according to a given safety problem.[25] MSSO and CIOMS are currently merging their efforts to provide standardised MedDRA queries (SMQs) as groups of MedDRA terms that relate to defined medical conditions or areas of interest.[26] End users will benefit from these new categories.

However, SSCs, MAGs, SSQs and SMQs must be entered manually and require the work of many

experts to retrieve all PTs or higher level terms that are relevant for a category. Even when terms are placed in a non-intuitive way, they can be retrieved anywhere if the search is based on a formal definition of terms. Experts would benefit from an automated tool to process MedDRA files. Validating sets of terms proposed by the computer is much easier and less error prone than the retrieval of all relevant terms because some terms can be found in some very odd portions of MedDRA.

5.5.2 Grouping of Terms is the Key to Better Signal Detection

The 'points to consider' document proposes a section on the purposes of using MedDRA. Half of the objectives (five of nine) are related to the grouping or aggregation of terms relevant for signal detection.[22] It includes review, comparison and frequency calculation of medically similar ADRs. Brown[27,28] has recently pointed out the need to group related terms that express the same medical condition.

Experts manually compare drugs and ADRs to find common patterns between case reports. Recent studies in the field of automated signal detection have used data mining techniques.[29] These approaches do not take advantage of a preliminary unsupervised clustering of cases using semantic knowledge. In its current state, MedDRA does not allow appropriate and automated clustering of ADRs since it is basically a first-generation system. We are currently building a formal model of MedDRA terms.[30] This type of conceptual organisation is called ontology in the knowledge acquisition and modelling field. Data mining techniques would benefit from semantic definitions of terms in such an ontology.

In the current study, we propose three approaches for clustering ADRs. First, the notion of related terms is based on semantic similarity theory. In a previous study, we described a semantic distance operator.[31] This semantic distance operator could be applied in the field of ADRs to provide many clusters of closely related terms.[32] Clustering is often proposed as a preliminary step in the data mining process when new data have to be inspected with no *a priori* knowledge. Second, formal concept analysis[33] could be used to construct a new hierarchy of terms based on Galois lattices. Third, a description logic engine can process formal descriptions to determine which terms have a given characteristic or not. The result is also a hierarchy of MedDRA terms in which missing relationships between concepts are retrieved. In both cases, the new structure is a network of related concepts in which a MedDRA term can be linked to more than one category inside the same SOC. Elements of the network are grouped at any level of the hierarchy according to the defining criteria. The resulting set may be proposed to end users as a new SSC. Elements from the set as a whole can be compared with other cases in the database to search for disproportionate associations in automated signal generation tools.[30]

6. Conclusion

In this article, we have studied the structure of MedDRA according to Cimino's criteria for the evaluation of terminological systems. In its current state, MedDRA is not a 'clinical nomenclature' in the context of the three generations of terminological systems framework proposed by Rossi Mori et al.[8] The main consequence is that the position of terms in the MedDRA taxonomy cannot be computed automatically. SSCs express particular clinical conditions but have to be entered manually. We propose that a formal definition could be provided for each MedDRA term. It is possible to keep the current organisation of MedDRA terms for statistical purposes and for reproducibility between users. However, new tools need to be implemented to map MedDRA terms to their formal representation in an ontology. Formal methods designed for the maintenance of terminology systems would be convenient to detect inconsistencies in MedDRA structures and build the SSCs automatically rather than manually. Formal definitions would also be useful for processing terms in auto-encoders, automated construction of SSCs and clustering of terms in pharmacovigilance databases in the context of signal detection. Other potential applications, such as quantitative

evaluation of drug safety in clinical trial data, could benefit from the semantic grouping of terms within a formal representation of MedDRA terms.

Acknowledgements

Cédric Bousquet acknowledges the Fonds d'Etude et de Recherche du Corps Médical des Hôpitaux de Paris. The authors have no conflicts of interest that are directly relevant to the content of this study.

References

1. Brown EG, Wood L, Wood S. The medical dictionary for regulatory activities (MedDRA). Drug Saf 1999; 20 (2): 109-17
2. Medical informatics. Categorical structures of systems of concepts-model for representation of semantics. CEN TC 251 WG2, Final Draft prENV 12264
3. Cimino JJ. Desiderata for controlled medical vocabularies in the twenty-first century. Methods Inf Med 1998; 37: 394-403
4. Chute GC, Cohn SP, Campbell KE, et al. The content coverage of clinical classifications. J Am Med Inform Assoc 1996; 3: 224-33
5. WHO (1993). Manual of the international statistical classification of diseases, injuries and causes of death: tenth revision of the international classification of diseases. Geneva: World Health Organisation, 1993
6. Côté RA, Rothwell DJ, Palotay JL, et al. SNOMED International. Northfield (IL): College of American Pathologists, 1993
7. Lindberg DAB, Humphreys BL, McCray AT. The Unified Medical Language System. Methods Inf Med 1993; 32: 281-91
8. Rossi Mori A, Consorti F, Galeazzi E. Standards to support development of terminological systems for healthcare telematics. Methods Inf Med 1998; 37: 551-63
9. Rector AL, Solomon WD, Nowlan WA, et al. A terminology server for medical languages and medical information systems. Methods Inf Med 1995; 34: 147-57
10. Sowa JF. Conceptual structures: information processing in mind and machine. Reading (MA): Addison Wesley, 1984
11. International Standardization Organization. Health informatics. Controlled health vocabularies – vocabulary structure and high level indicators: ISO/DTS 17117 [online]. Available from URL: http://www.tc215wg3.nhs.uk/pages/pdf/iso17117v1.pdf [Accessed 2004 Nov 22]
12. Office of Biotechnology activities, National Institute of Health. Development of a national database of gene transfer clinical research [online]. Available from URL: http://www.orpha.net/docs/GeMCRIS.ppt [Accessed 2004 Nov 22]
13. Huntley K, Veverka MJ. The FDA's medical dictionary for drug regulatory affairs alpha test. Drug Inf J 1995; 29: 1133-43
14. Brown EG, Douglas S. Tabulation and analysis of pharmacovigilance data using the medical dictionary for regulatory activities. Pharmacoepdemiol Drug Saf 2000; 9: 479-89
15. Strang N, Cucherat M, Boissel JP. Which coding system for therapeutic information in evidence-based medicine. Comput Methods Programs Biomed 2002; 68: 73-85
16. Goldsmith D. Data elements for transmission of individual case safety reports, 2000 [online]. Available from URL: http://www.fda.gov/cder/m2/E2B/E2BSTP4_V4.4.1.pdf [Accessed 2004 Nov 22]
17. Brown EG, Clark E. Evaluation of MEDDRA in representing medicinal product information data sheet information. Pharm Med 1996; 10: 111-8
18. Duclos C, Venot A. Structured representation of drug indications: lexical and semantic analysis and object-oriented modeling. Methods Inf Med 2000; 39: 83-7
19. Ingenerf J. Taxonomic vocabularies in medicine: the intention of usage determines different established structures. In: Greenes RA, Peterson HE, Protti DJ, editors. MEDINFO'95 Proceedings, 1995 (Pt 1): 136-9
20. Ingenerf J, Giere W. Concept-oriented standardization and statistics-oriented classification: continuing the classification versus nomenclature controversy. Methods Inf Med 1998; 37: 527-39
21. Trombert-Paviot B, Rodrigues JM, Rogers JE, et al. Galen: a third generation terminology tool to support a multipurpose national coding system for surgical procedures. Int J Med Inf 2000; 58-59: 71-85
22. MedDRA term selection: points to consider, release 3.1. Application to adverse drug reactions, adverse events & medical and social history & indications [online]. Available from URL: http://www.meddramsso.com [Accessed 2003 Jun 18]
23. Yokotsuka M, Aoyama M, Kubota K. The use of a medical dictionary for regulatory activities terminology (MedDRA) in prescription-event monitoring in Japan (J-PEM). Int J Med Inf 2000; 57: 139-53
24. MSSO user group. MedDRA analytical groupings [online]. Available from URL: www.meddramsso.com/NewWeb2003/Docs/magusergroup.ppt [Accessed 2004 Nov 22]
25. New CIOMS working group: rational use of MedDRA terminology for drug safety database searches [online]. Available from URL: http://www.cioms.ch [Accessed 2003 Jun 18]
26. Brown EG. Using MedDRA: implications for risk management. Drug Saf 2004; 27 (8): 591-602
27. Brown EG. Effects of coding dictionary on signal generation: a consideration of use of MedDRA compared with WHO-ART. Drug Saf 2002; 25 (6): 445-52
28. Brown EG. Methods and pitfalls in searching drug safety databases utilising the medical dictionary for regulatory activities (MedDRA). Drug Saf 2003; 26 (3): 145-58
29. Hauben M, Zhou X. Quantitative methods in pharmacovigilance: focus on signal detection. Drug Saf 2003; 26 (3): 159-86
30. Henegar C, Bousquet C, Lillo-Le Louët A, et al. A knowledge based approach for automated signal generation in pharmacovigilance. Medinfo 2004, 626-30
31. Bousquet C, Jaulent MC, Chatellier G. Using semantic distance for the efficient coding of medical concepts. Amsterdam: IMIA, 2004: Proc AMIA Symp 2000, 96-100
32. Bousquet C, Jaulent MC, Lagier G. Using a semantic distance operator: a proposal for the clustering of ADR case reports with the MedDRA classification. Pharmacoepidemiol Drug Saf 2002; 11 Suppl. 2: S243
33. Ganter B, Wille R. Formal concept analysis. Berlin, Heidelberg: Springer Verlag, 1999

Correspondence and offprints: Dr *Cédric Bousquet*, Centre Régional de Pharmacovigilance, Hôpital Européen Georges Pompidou, 20-40 rue Leblanc, Paris, 75015, France. E-mail: cbousquet@magic.fr