

Research article

Mapping French terms in a Belgian guideline on heart failure to international classifications and nomenclatures: the devil is in the detail

Jamoulle M, Cardillo E, Roumier J, Warnier M, Stichele RV. Mapping French terms in a Belgian guideline on heart failure to international classifications and nomenclatures: the devil is in the detail. *Inform Prim Care*. 2014;21(4):189–198.

<http://dx.doi.org/10.14236/jhi.v21i4.66>

Copyright © 2014 The Author(s). Published by BCS, The Chartered Institute for IT under Creative Commons license <http://creativecommons.org/licenses/by/4.0/>

Author address for correspondence:

Marc Jamoulle
Department of General Practice
Liège University
Place du 20 Août
4000 Liège, Belgium
Email: marc.jamoulle@doct.ulg.be

Accepted September 2014

Marc Jamoulle

Department of General Practice, Liège University, Liège, Belgium

Elena Cardillo

Fondazione Bruno Kessler, Trento, Italy

Joseph Roumier

Centre d'Excellence en Technologies de l'Information et de la Communication, Charleroi, Belgium

Maxime Warnier

University of Toulouse, Toulouse, France

Robert Vander Stichele

University of Ghent, Ghent, Belgium

ABSTRACT

Introduction With growing sophistication of eHealth platforms, medical information is increasingly shared across patients, health care providers, institutions and across borders. This implies more stringent demands on the quality of data entry at the point-of-care. Non-native English-speaking general practitioners (GPs) experience difficulties in interacting with international classification systems and nomenclatures to facilitate the secondary use of their data and to ensure semantic interoperability.

Aim To identify words and phrases pertaining to the heart failure domain and to explore the difficulties in mapping to corresponding concepts in ICPC-2, ICD-10, SNOMED-CT and UMLS.

Methods The medical concepts in a Belgian guideline for GPs in its French version were extracted manually and coded first in ICPC-2, then ICD-10 by a physician, an expert in classification systems. In addition, mappings were sought with SNOMED-CT and UMLS concepts, using the UMLS SNOMED-CT browser.

Results We identified 143 words and phrases, of which 128 referred to a single concept (1-to-1 mapping), while 15 referred to two or more concepts (1-to-*n* mapping to ICPC rubrics or to the other nomenclatures). In the guideline, words or phrases were often too general for specific mapping to a code or term. Marked discrepancy between semantic tags and types was found.

Conclusion This article shows the variability of the various international classifications and nomenclatures, the need for structured guidelines with more attention to precise wording and the need for classification expertise embedded in sophisticated terminological resources. End users need support to perform their clinical work in their own language, while still assuring standardised and semantic interoperable medical registration. Collaboration between computational linguists, knowledge engineers, health informaticians and domain experts is needed.

INTRODUCTION

Speakers of different languages have different ways in representing the world; such differences lie not only in terms, but also in the concepts themselves.¹ The medical language is no exception. The translation of English medical terminological resources to other languages could be difficult.

The general practitioner (GP) is an important node of the information flow in health care. He is at the crossing point between patient and health care worlds and is also a focal point in the life cycle of clinical information. The family doctor has to manage lay terms as well as professional terms in the native language of the patient.² Moreover, as patients are considering the Internet and social networking as information sources³ for health care, mappings to multilingual lay terms are inescapable.⁴

For GPs, guidelines offer coherent, comprehensive and consensual recommendations which at least, if designed rigorously, minimise the potential harms to patients.⁵

In this article, we report the terminological analysis of a guideline for GPs in Belgium (a West European bilingual country), produced in 2012, according to explicit guideline development methods.⁶ The topic of the guideline was 'heart failure', a prevalent chronic disease. The guideline was published in French and Dutch, with identical structure, content and text size. The analysis of the words and phrases used in this guideline provided the material for the MERITERM consortium (meriterm.org), a multidisciplinary research group who has proposed a hybrid methodology for the creation of a multilingual reference terminology for the primary care domain in Europe, by combining the application of the International Standards Organisation's (ISO's) standards for multilingual terminologies and the use of Semantic Web technologies for connecting this resource to linguistic resources as well as to medical nomenclatures, classifications and thesauri.^{7,8} The long-term objective is to build an end-user terminology that supports not only coding activities for medical registration, but also for information retrieval, quality assurance and creation of information through epidemiological research⁹ considering multilingualism and semantic interoperability.¹⁰

Ontologies and semantic integration technologies are seen as tools to manage knowledge in health care, but require further research and development as well as experimental work.¹¹ Ontology development should be preceded by thorough descriptive terminological analysis of the real-life language of the persons who will ultimately use the ontology.¹² By analysing the terminological content of a clinical guideline and making an attempt to link the extracted words and phrases to a broad range of relevant nomenclatures and classifications, we provide a glimpse of the complexity of language, machine terminology, medical semantics and ontologies.

The aim of this article was to explore the semantic difficulties in mapping French terms from a general practice guideline on heart failure to relevant concepts in English-based international nomenclatures and classifications.

Our research question was: what are the barriers to the process of mapping clinical terms in clinical guidelines in a language other than English (i.e. French) to concepts available in English language classifications and nomenclatures?

Method

The identification of the clinical terms relevant for patient care within the text of the heart failure guideline was carried out manually by a French-speaking family doctor with expertise in medical classification and terminology.

The selected words were then classified according to an ad hoc grid (disease, function, objective finding, process general, risk, symptom and therapeutics process). Words and phrases related to 'objective findings' and 'therapeutic process' were excluded to remove technical names of medical procedures and medications. Acronyms and names of clinical scales and their domain values were also excluded, as these were taken from standardised terminologies from the beginning. Included French words and phrases were entered into a database (a multilingual reference terminology) and considered as preferred term of a concept, with an informal definition, to indicate the intended sense, in case of polysemy (words having more than one sense). This was necessary to assure correct mapping to concepts in international classifications, but also to further identify the correct preferred terms in other languages¹³ and in the Dutch version of the guideline.

Here, report mapping of the selected concepts pertaining to French words and phrases, on the one hand, and concepts in international nomenclatures and classifications on the other hand. After three revisions, all the results were arranged in a database showing the mapping between the terms in French and the concepts in the four international classifications and nomenclatures (e.g. Box 1).

The first step was to semantically map the selected terms to ICPC-2,¹⁴ a classification specific to general practice and primary care. In a second step, the identified ICPC-2 codes were mapped to ICD-10,¹⁵ the main classification for encoding diseases in secondary care. For each mapped concept in ICPC-2 and ICD-10, we retrieved the term and its code. Furthermore, mappings to SNOMED-CT,¹⁶ the nomenclature and reference terminology for coding in electronic health records, and to UMLS,¹⁷ the main terminology used for indexing knowledge in medicine, were sought. For SNOMED-CT, we retrieved the corresponding concept, the code and the fully specified name, including the 'semantic tag'.¹⁸ For UMLS, we retrieved the corresponding term, the code, the 'semantic type', and the definition (if available).¹⁹ For SNOMED-CT, it was not possible to retrieve a definition, as a formal definition of concepts is not provided in SNOMED-CT.

We used specific online browsers for each of the four classifications and nomenclatures described above to extract the mappings (see Box 2).

Codes and semantic values of the different terms and their meanings were then subjected to a comprehensive critical study.^{20,21}

Box 1 Column labels and row content of the database with as example: the term cardiac insufficiency and its correspondences in ICPC-2, ICD-10, SNOMED-CT and UMLS

Column label (short)	Column label (long)	Row content
Term (French)	Term used in French for selected concept in French guideline	Décompensation cardiaque
Term (English)	English translation of the term	Cardiac insufficiency
Type	Type of the concept (diagnosis, symptoms and risk)	Diagnosis
ICPC code	Code of the corresponding rubric or chapter in ICPC-2	K77
ICPC-2 rubric	Label of the corresponding rubric in ICPC-2	Heart failure
ICD-10 code	Code of the corresponding concept in ICD-10	i50
ICD-10 label	Label of the corresponding category in ICD-10	Heart failure
SNOMED-CT ID	Corresponding identifier of the corresponding concept in SNOMED-CT	825890014
SNOMED-CT FSN	Fully specified name in SNOMED-CT (with its semantic tag)	Heart failure (disorder)
UMLS CUI	Concept unique identifier (CUI) of the corresponding concept in UMLS	C0018801
UMLS concept name	Concept name in UMLS (with its semantic type)	Heart failure
UMLS Def_Source	Source of the definition in UMLS if any	CSP/PT
UMLS definition	Definition of the concept if any in UMLS	Inability of the heart to pump blood at an adequate rate to fill tissue metabolic requirements or the ability to do so only at an elevated filling pressure.

Box 2 Four classifications and terminologies browsers available online

- **ICPC-2e (en) v4.2beta browser** accessible through the Web page of the World Health Organisation (WHO) collaborating centre for the Family of International Classifications in the Netherlands. <http://icpc.who-fig.nl/browser.aspx> (web site of the WICC ICPC update group), which allows also to find the correct transcoding code to ICD-10.
- **ICD online browser** WHO. ICD-10 Version: 2010. International Statistical Classification of Diseases and Related Health Problems 10th Revision. accessible through: <http://apps.who.int/classifications/icd10/browse/2010/en>
- **SNOMED CT browser** of the UMLS terminology services, a service of the U.S. National Library of Medicine | National Institutes of Health <https://uts.nlm.nih.gov//snomedctBrowser.html> allowing to find the corresponding meaning in SNOMED-CT 2011.
- **Metathesaurus browser** of the UMLS terminology services, a service of the U.S. National Library of Medicine | National Institutes of Health <https://uts.nlm.nih.gov//metathesaurus.html> allowing to find the correspondence between the SNOMED-CT concept and the CUI of UMLS (linked to the SNOMED-CT browser).

Results

In total, 283 words and phrases were selected from the French version of the Belgian guideline. Eight entries were excluded because they were acronyms (e.g. 'AIT Attaque Ischémique Transitoire' [TIA for transient ischaemic attack]) or names and domain values of clinical scales (e.g. New York Heart Classification with four domain values). We excluded 132 words and phrases pertaining to objective findings and therapeutic process. We retained 143 words and phrases.

Out of these 143 words and phrases, 128 referred to a single concept (1-to-1 mapping); an additional 15 words and phrases were so general that eight of them needed to be represented by two more specific concepts and six by three concepts, to be able to make a bridge to ICPC or other nomenclatures in a 1–*n* mapping. The final reference terminology database had 160 lines. (The file is available on <http://meriterm.org/K77.xlsx>.)

Another eight of the 128 selected terms were so general and unspecific that it was not possible to find any correspondence in the ICPC-2 classification, while for an additional eight terms only a partial correspondence with ICPC-2 was found at the level of the chapter. For these very generic terms, equivalents could be found in the SNOMED-CT and UMLS, but with discrepancies in the choice of semantic tags or types (Table 1).

In addition, five of the 15 terms which could not be represented with a single concept pertained to general but related cardiovascular concepts (slightly different in the label, but without consistent difference in meaning): cardiomyopathy, coronaropathy, coronary ischaemia, myocardial ischaemia and coronary heart disease (Table 2).

Six general concepts needed to be split over two distinct or more concepts in ICPC, based on insulin dependent/non-insulin dependent; acute/chronic or male/female distinctions (Table 3).

Table 1 Terms so general that they could not or only partially be mapped to the ICP-2 or ICD 10 classifications – the dash indicates inability to map to concepts in the system

Term (French)	Term (English)	ICPC-2 code	ICD-10 code	ICD-10 label	SNOMED-CT FSN (semantic tag)	UMLS concept name (semantic type)
Affection thoracique	Thoracic disease	–	–	–	Disorder of thorax (disorder)	Thoracic Diseases (disease or syndrome)
Besoin spirituel	Spiritual need	–	–	–	Spiritual need of patient (observable entity)	Spiritual need (finding)
Cancer	Cancer	–	C00–C97	Malignant neoplasms	Malignant neoplastic disease (disorder)	Malignant Neoplasms (neoplastic process)
Fonctions cognitives	Cognitive functions	–	–	–	Cognitive functions (observable entity)	Cognitive functions (mental process)
Filtration glomerulaire	Glomerular filtration	–	–	–	Glomerular filtration, function (observable entity)	Glomerular filtration (organ or tissue function)
Résistance vasculaire systémique	Systemic vascular resistance	–	–	–	Systemic vascular resistance (observable entity)	Total peripheral resistance (physiologic function)
Risque	Risk	–	–	–	Risk of (contextual qualifier) (qualifier value)	Risk (qualitative concept)
Tolérance à l'effort	Exercise tolerance	–	–	–	Exercise tolerance (observable entity)	Exercise Tolerance (clinical attribute)
Anémie	Anaemia	B	D64.9	Anaemia, unspecified	Anaemia (disorder)	Anaemia (disease or syndrome)
Maladie du foie	Liver disease	D	K70–K77	Diseases of liver	Disorder of liver (disorder)	Liver diseases (disease or syndrome)
Maladie de cœur	Heart disease	K	I51.9	Heart disease, unspecified	Heart disease (disorder)	Heart disease (disease or syndrome)
Maladies vasculaires périphériques	Peripheral vascular diseases	K	I70–I79	Diseases of arteries, arterioles and capillaries	Peripheral vascular disease (disorder)	Peripheral vascular diseases (disease or syndrome)
Fracture	Fracture	L	–	–	Fracture of bone (disorder)	Fracture (disease or syndrome)
Maladie du poumon	Lung disease	R	–	–	Disorder of lung (disorder)	Lung diseases (disease or syndrome)
Maladie de la thyroïde	Thyroid disease	T	E00–E07	Disorders of thyroid gland	Disorder of thyroid gland (disorder)	Thyroid diseases (disease or syndrome)
Maladie du rein	Kidney disease	U	–	–	Kidney disease (disorder)	Kidney diseases (disease or syndrome)

Table 2 Some very generic terms found in the French guidelines and their mappings to ICP-C-2 with due correspondences

Term (French)	Term (English)	ICPC-2 code	ICPC-2 rubric	ICD-10 code	SNOMED-CT FSN (semantic tag)	UMLS concept name (semantic type)
Cardiomyopathie	Cardiomyopathy	K84	Heart disease other	I42	Cardiomyopathy (disorder)	Cardiomyopathies (disease or syndrome)
		K73	Congenital anomaly cardiovascular	I42.4	Congenital hypertrophy of cardiac ventricle (disorder)	Congenital hypertrophy of cardiac ventricle (congenital abnormality / disease or syndrome)
		K76	Ischaemic heart disease with or without angina	I25.5	Ischaemic congestive cardiomyopathy (disorder)	Ischaemic congestive cardiomyopathy (disease or syndrome)
Coronaropathie	Coronaropathy	K74	Ischaemic heart disease without angina	I25	Coronary arteritis (disorder)	Coronary arteritis (disease or syndrome)
		K75	Acute myocardial infarction	I25	Acute myocardial infarction (disorder)	Acute myocardial infarction (disease or syndrome)
		K76	Ischaemic heart disease with or without angina	I25	Chronic ischaemic heart disease (disorder)	Chronic myocardial ischaemia (disease or syndrome)
		K74	Ischaemic heart disease without angina	I25	Coronary arteritis (disorder)	Coronary arteritis (disease or syndrome)
		K76	Acute myocardial infarction	I25	Chronic ischaemic heart disease (disorder)	Chronic myocardial ischaemia (disease or syndrome)
Ischémie coronaire	Coronary ischaemia	K74	Ischaemic heart disease without angina	I25	Coronary arteritis (disorder)	Coronary arteritis (disease or syndrome)
		K76	Acute myocardial infarction	I25	Chronic ischaemic heart disease (disorder)	Chronic myocardial ischaemia (disease or syndrome)
Ischémie myocardique	Myocardic ischaemia	K74	Ischaemic heart disease without angina	I25.9	Myocardial ischaemia (disorder)	Myocardial ischaemia (disease or syndrome)
		K75	Acute myocardial infarction	I25	Acute myocardial infarction (disorder)	Acute myocardial infarction (disease or syndrome)
		K74	Ischaemic heart disease without angina	I25	Coronary arteritis (disorder)	Coronary arteritis (disease or syndrome)
Pathologie de l'artère coronaire	Coronary artery pathology	K74	Ischaemic heart disease without angina	I25	Coronary arteritis (disorder)	Coronary arteritis (disease or syndrome)
		K75	Acute myocardial infarction	I25	Acute myocardial infarction (disorder)	Acute myocardial infarction (disease or syndrome)
		K76	Ischaemic heart disease with or without angina	I25	Chronic ischaemic heart disease (disorder)	Chronic myocardial ischaemia (disease or syndrome)

Table 3 Six terms of the French guideline causing problems of specification – the dash indicates inability to bridge to concepts in the system

Term (French)	Term (English)	ICPC-2 Code	ICPC-2 Rubric	ICD-10 Code	ICD 10 Label	SNOMED-CT FSN (semantic tag)	UMLS Concept Name (semantic type)
Diabète	Diabetes	T89	Diabetes insulin dependent	E10	Insulin-dependent diabetes mellitus	Diabetes mellitus (disorder)	Diabetes mellitus (disease or syndrome)
		T90	Diabetes non-insulin dependent	E11	Non-insulin-dependent diabetes mellitus	Diabetes mellitus (disorder)	Diabetes mellitus (disease or syndrome)
Consommation d'alcool excessive	Excessive alcohol consumption	P15	Chronic alcohol abuse	F10.1	Harmful use of alcohol	Alcohol abuse (disorder)	Alcohol abuse (mental or behavioural dysfunction)
		P16	Acute alcohol abuse	F10.0	Acute alcohol intoxication	Alcohol intoxication (disorder)	Acute alcoholic intoxication (mental or behavioural dysfunction)
Maladies sexuellement transmissibles	Sexually transmitted diseases	Y	Male Genital	A50–A64	Infections with a predominantly sexual mode of transmission	Sexually transmitted infectious disease (disorder)	Sexually transmitted diseases (disease or syndrome)
		X	Female Genital	A50–A64	Infections with a predominantly sexual mode of transmission	Sexually transmitted infectious disease (disorder)	Sexually transmitted diseases (disease or syndrome)
Problèmes sexuels	Sexual problems	P08	Sexual fulfillment reduced	F52.1	Sexual aversion and lack of sexual enjoyment	—	Sexual fulfillment reduced (sign or symptom)
		P	Psychological	—	—	History of – sexual problem – female (situation)	H/O: sexual problem – female (finding)
		P	Psychological	—	—	History of – male sex function problem (situation)	H/O: male sex function problem (finding)
		P	Psychological	—	—	Abnormal sexual function (finding)	Sexual dysfunction (finding)
Syphilis	Syphilis	Y07	Impotence not otherwise specified	N48.4	Impotence of organic origin	Impotence (disorder)	Erectile dysfunction (disease or syndrome)
		Y70	Syphilis male	A52.0	Cardiovascular syphilis	Syphilis (disorder)	Syphilis (disease or syndrome)
		X70	Syphilis female	A52.0	Cardiovascular syphilis	Syphilis (disorder)	Syphilis (disease or syndrome)
Trouble du rythme cardiaque	Heart rhythm disorders	K84	Heart disease other	I45.9	Conduction disorder and unspecified	Conduction disorder of the heart (disorder)	Conduction disorder of the heart (disease or syndrome) – Cardiac arrhythmia (pathologic function)

Table 4 Number of terms of the French guidelines not mapped to the classifications or nomenclatures

Not in UMLS (1)	Systemic vascular resistance	Drug induced ankle oedema
Drug induced ankle oedema	Risk	Heart valve disease
Not in SNOMED (2)	Exercise tolerance	Lung disease
Drug induced ankle oedema	Not in ICD (24)	Referral to cardiologist
Sexual fulfilment reduced	Thoracic disease	Intercourse
Not in ICPC (8)	Spiritual need	Systemic vascular resistance
Spiritual need	Functional capacity	Magnetic resonance
Cancer	Cognitive functions	Risk
Functional capacity	Coronary angiography	Chest X-ray
Cognitive functions	Cardiac echography	Cognitive status
Glomerular filtration	Glomerular filtration	Nutritional status
	Fracture	Fluid balance
	Weight measure	Exercise tolerance
		Isotopic ventriculography

As shown in Tables 1–4, eight terms could not be mapped to ICPC-2, 24 to ICD-10, 1 to UMLS and 2 to SNOMED-CT, and only one term, ‘drug induced ankle oedema’, did not map to three classifications/nomenclatures (UMLS, SNOMED-CT and ICD-10).

While looking for semantic mappings, we noticed that heterogeneity is the rule. Even if SNOMED-CT disorders (semantic tags) are often related to UMLS disease or syndrome (semantic type) and UMLS finding and sign, and symptoms recover more or less what is called symptoms in ICPC, no such semantic tag as symptom exists in SNOMED-CT.

We did observe several discrepancies in the semantic values of the two classifications and the two nomenclatures. T06 (loss of appetite) is a symptom in ICPC-2 as well as in ICD-10 (classified in Chapter R), a finding in SNOMED-CT and a disease or syndrome in UMLS (as indicated by the specific browser). The absolute distribution of the retrieved semantic meanings seems to correspond roughly at first glance: we found 80 UMLS diseases or syndrome, 96 SNOMED-CT disorders, 89 ICD10 disease and 76 ICPC-2 diagnoses (Table 5).

However, 41 entries pertaining to ICPC-2 symptoms have corresponding entries in SNOMED-CT (except one) as: 27 findings, 11 disorders and two observable entities. The same ICPC-2 entries are distributed in eight different semantic types in UMLS: 22 sign and symptoms, seven findings and five others (Figure 1).

Hence, observed barriers to mapping French words and phrases were problems of polysemy, homonymy and the use of overly general concepts (hypernymy) which could either not be represented or needed to be related to several more specific concepts (hyponymy) in the different classifications and nomenclatures. In addition, we observed differences in semantic tags and types between the different systems, illustrating differences in world of reference.

This expert-driven review of the meaning of words and phrases from a single guideline in a specific field of general practice, and the mapping of these French words and phrases to concepts in international classifications and nomenclatures proved to be a labour-intensive process.

DISCUSSION

By examining the terminological content of a French guideline for GPs about heart failure, we were able to identify 128 words and phrases that could be related to a single concept, with a 1-to-1 mapping to international classifications, and 15 that needed to be related to two or more concepts, to be able to map (1-to-*n* mapping). In total, 160 concepts were mapped to (1) two international medical classifications (ICPC-2 and ICD-10) and (2) two international nomenclatures SNOMED-CT (for medical registration) and UMLS (for indexing and information retrieval).

A barrier observed was the very low specificity of words and phrases used in the narrative version of the guideline. A first step was to choose the intended implicit sense of each observed word and phrase for mapping to the international classifications. Often several quasi-synonyms were used for very general concepts that needed to be represented in the classifications or nomenclatures by more specific concepts.

In this article, we attempted the mapping to classifications for all observed terms, but it is clear that authors of guidelines will have to make choices in the fuzzy vocabularies and carefully select well-defined concepts at the right level of abstraction and also select the most accurate representation of the concept by a preferred term in their own language. In further terminological work, additional meanings of selected terms in the given language and synonyms for the concept should be dealt with in linguistic resources capable of dealing with the richness and fuzziness of natural language(s).

The attempt to map to several international classifications and nomenclatures was supported using the ICPC classification from general practice as a grid for access to the more granular ICD classification, and ultimately to the two much more specific nomenclatures. The use of the SNOMED-CT to UMLS mapping was useful to retrieve international definitions of concepts (not available in SNOMED-CT). For a number of concepts, this approach was not working, and ICPC (and even ICD) needed to be bypassed.

A limitation of this article was that the extraction of words and phrases, the choice of intended sense and the mapping

Table 5 Distribution of the Semantic values in the terms of the French guideline

UMLS concepts (semantic types)	163	SNOMED-CT (semantic tags)	155	ICD-10 equivalent category	114	ICPC-2 equivalent category	129
Disease or syndrome	80	Disorder	96	Disease	89	Diagnosis	76
Finding	19	Finding	28	-	-	-	-
Sign or symptom	21	-	-	Symptoms, signs & findings (Chapter R)	24	Symptom/complaint	45
Pathologic function	10	Observable entity	17	-	-	-	-
Diagnostic procedure	5	Procedure	6	Factors influencing health status (Chapter Z)	1	Process	8
Mental or behavioural dysfunction	5	Situation	2	-	-	-	-
Mental process	3	Qualifier value	2	-	-	-	-
Organism attribute	2	Event	2	-	-	-	-
Organism function	2	Substance	1	-	-	-	-
Clinical attribute	2	Physical object	1	-	-	-	-
Congenital abnormality	1	-	-	-	-	-	-
Neoplastic process	2	-	-	-	-	-	-
Physiologic function	2	-	-	-	-	-	-
Anatomical abnormality	1	-	-	-	-	-	-
Functional concept	1	-	-	-	-	-	-
Hazardous or poisonous substance	1	-	-	-	-	-	-
Health care activity	1	-	-	-	-	-	-
Individual behaviour	1	-	-	-	-	-	-
Injury or poisoning	2	-	-	-	-	-	-
Medical device	1	-	-	-	-	-	-
Organ or tissue function	1	-	-	-	-	-	-

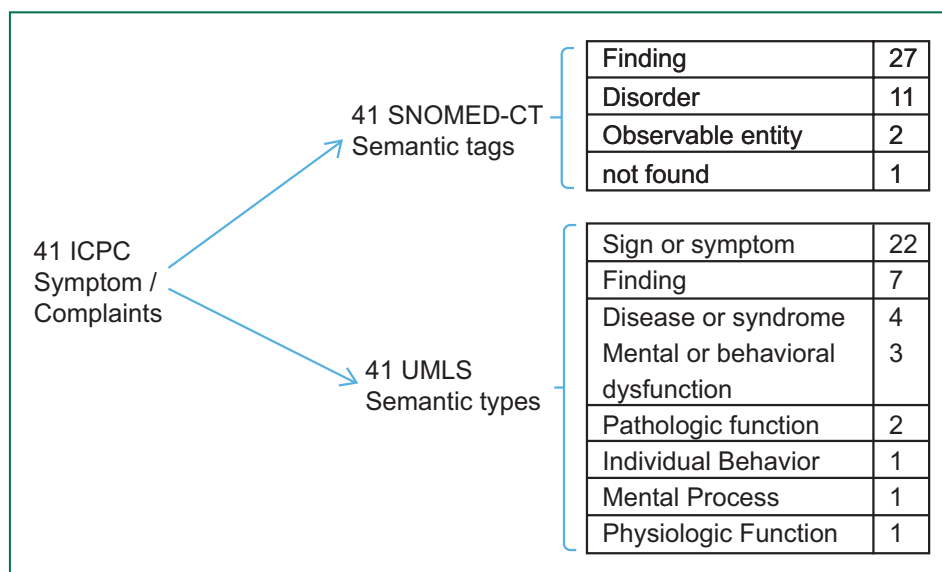


Figure 1 Semantic values mapping distribution in 41 items classed as Symptoms in ICPC-2 on 160 analysed

to international concepts was only performed by a single researcher, albeit physician and expert in terminology. The process included an element of translation, which is always a shift, not only between two languages but between two cultures.²² The chosen terms and proposed mappings reflect the bias, background, errors, omissions and misinterpretations of the coder. Validation by a consensus process, including medical translators, is needed. Consultations between representatives of multiple languages will be necessary to decide on the ultimate selection of a collection of reference concepts, suitable for a feasible mapping to international classifications and nomenclatures.

We observed important differences between the four international terminology systems. Classifications and nomenclatures are built for different purposes, and not based on the same world of reference. UMLS Semantic types^{23,24} or SNOMED-CT Semantic tags²⁵ are different constructs. SNOMED-CT represents the views of the pathologists in a world where symptoms are non-existing and where there is only a place for disorder, finding, observable entities, morphology or structure. SNOMED-CT shares also with UMLS authors a pre-Canguilhem²⁶ view about the normal and the pathological by referring to the concept of abnormality. ICPC-2 reflects the point of view of family doctors who are close to their patients.²⁷ ICD-10 is disease-centred and an extension of a historical mortality classification, although augmented by several symptoms and procedures. Classifications and terminologies are not neutral; they reflect the world view of the

human classifier. Bowker²⁸ summed the subjective nature as 'to classify is human'. Nevertheless, despite the difficulty of reconciling these apparently so different worlds, we must focus on our common interest, the safe and effective care of the human person.²⁹

Guidelines in narrative form are difficult to implement in EHRs. In the last 20 years, major steps have been made in methods to formalise guidelines into computer usable information, often in the field of heart diseases.^{30–31} This formalisation requires the use of standardised clinical vocabularies and terminologies.³² Efforts are needed from physicians to understand the necessity to use a common standardised terminology when producing guidelines and performing medical registration. However, physicians in clinical practice also need to be helped by sophisticated linguistic and expert-based mapping support to several classifications and nomenclatures. This will require close cooperation between clinicians, knowledge engineers, computational linguists and health informaticians.

Acknowledgements

This work was supported by a grant from the Belgian EBMPPracticeNet network (www.ebmpracticenet.be) and by the institutions participating to the MERITERM consortium (www.meriterm.org): Heymans Institute of Pharmacology, University of Ghent, Ghent, Belgium (ugent.be); Fondazione Bruno Kessler, Trento, Italy (fbk.eu); Centre of Expertise in Technology for Information and Communication, Charleroi, Belgium (cetic.be).

REFERENCES

1. Eco U. Dire quasi la stessa cosa, Esperienze di traduzione. [Experiences in translation] Bompiani, 2003. PMID:12514644.
2. Cardillo E, Hernandez G and Bodenreider O. Integrating consumer-oriented vocabularies with selected professional ones from the UMLS using semantic web technologies. Szomszor M and Kostkova P (Eds). *Lecture Notes of the Institute for Computer Sciences, Social Informatics and Telecommunications Engineering*. Berlin Heidelberg: Springer, 2010. Available from: <http://mor.nlm.nih.gov:8000/pubs/pdf/2010-eHealth-ec.pdf>.
3. Jain SH. Practicing medicine in the age of Facebook. *New England Journal of Medicine* 2009;361(7):649–51. Available from: <http://dx.doi.org/10.1056/NEJMp0901277>. <http://dx.doi.org/10.1056/NEJMp0901277>.
4. Singh PM, Wight CA, Sercinoglu O, Wilson DC, Boytsov A and Raizada MN. Language preferences on websites and in Google searches for human health and food information. *Journal of Medical Internet Research* 2007;9(2):e18. doi:10.2196/jmir.9.2.e18. Available from: <http://dx.doi.org/10.2196/jmir.9.2.e18>.
5. Sonnenberg FA, Hagerty CG, Acharya J, Pickens DS and Kulikowski CA. Vocabulary requirements for implementing clinical guidelines in an electronic medical record: a case study. *Proceedings of AMIA Annual Symposium* 2005, pp. 709–13. PMID:16779132; PMCID:PMC1560696.
6. Van Royen P, Chevalier P, Dekeulenaer G, Goossens M, Koeck P, Vanhalewyn M et al. Recommandation de Bonne Pratique Insuffisance cardiaque [Guidelines on Heart Failure]. Belgium: Domus Medica. SSMG, 2011.
7. Roumier J, Jamoulle M, Romary L, Cardillo E and Vander Stichele RH. Towards a terminologies support system in primary care [Letter to the editor]. *Informatics in Primary Care* 2011;19:257–8. PMID:22828581.
8. Cardillo E, Warnier M, Roumier J, Jamoulle M and Vander Stichele RH. Using ISO and semantic web standards for creating a multilingual medical interface terminology: a use case for heart failure. *Proceedings of the 10th International Conference on Terminology and Artificial Intelligence (TIA2013) Paris, IRIT*, 2013, pp. 27–34.
9. de Lusignan S. Codes, classifications, terminologies and nomenclatures: definition, development and application in practice. *Informatics in Primary Care* 2005;13(1):65–70. PMID:15949178.
10. European Commission DE & I. ICT standards in the health sector: current situation and prospects. Special Study No. 1, 2008, p. 82. Available from: <http://goo.gl/15R8CQ>.
11. Liyanage H, Liaw ST, Kuziemyk C, Terry AL, Jones S, Soler JK et al. The evidence-base for using ontologies and semantic integration methodologies to support integrated chronic disease management in primary and ambulatory care: realist review. Contribution of the IMIA Primary Health Care Informatics WG. *Yearbook of Medical Informatics* 2013;8(1):147–54. PMID:23974562.
12. Smith B and Ceusters W. An ontology-based methodology for the migration of biomedical terminologies to electronic

- health records. *Proceedings of AMIA Annual Symposium 2005*, pp. 704–8. PMID:PMC1560617.
13. Cardillo E, Warnier M, Roumier J, Jamoulle M and Vander Stichele R. Using ISO and Semantic Web standards for creating a Multilingual Medical Interface Terminology: a use case for heart failure. *Proceedings of the 10th International Conference on Terminology and Artificial Intelligence (TIA2013)*, October 28, 2013. Available from: <http://hdl.handle.net/2268/171534>.
 14. Okkes IM, Jamoulle M, Lamberts H and Bentzen N. ICPC-2-E: the electronic version of ICPC-2. Differences from the printed version and the consequences. *Family Practice* 2000;17:101–107. Available from: <http://dx.doi.org/10.1093/fampra/17.2.101>. PMID:10758069
 15. World Health Organization. *The ICD-10 Classification of Mental and Behavioral Disorders: Clinical Descriptions and Diagnostic Guidelines*. Geneva: WHO, 1992.
 16. International Health Terminology Standards Development Organization, Copenhagen. *Systematized Nomenclature of Medicine – Clinical Term (SNOMED CT)*. Available from: <http://www.ihtsdo.org/snomed-ct/>
 17. National Library of Medicine, National Institutes of Health, U.S. Department of Health and Human Services. *Unified Medical Language System (UMLS)*. Available from: <http://www.nlm.nih.gov/research/umls/>.
 18. International Health Terminology Standards Development Organization, Copenhagen. *SNOMED-CT Technical Implementation Guide 2014*. Available from: <http://www.ihtsdo.org/snomed-ct/>.
 19. Ceusters W, Rogers J, Consorti F and Rossi-Mori A. Syntactic–semantic tagging as a mediator between linguistic representations and formal models: an exercise in linking SNOMED to GALEN. *Artificial Intelligence in Medicine* 1999;15(1):5–23. Available from: [http://dx.doi.org/10.1016/S0933-3657\(98\)00043-8](http://dx.doi.org/10.1016/S0933-3657(98)00043-8).
 20. Jamoulle M. An information science eye at a general practice/family medicine dedicated guideline. Report to the EBMPPracticeNet network. 2011. Part 1 (84p.), Brussels. Available from: http://trix.docpatient.net/files/others/K77_guideline_mj_studydata_part1.pdf
 21. Jamoulle M and Roumier J. An information science eye at a general practice/family medicine dedicated guideline. Extension of data categories. Report to the EBMPPracticeNet network. 2012 Part 2 (12p.), Brussels. Available from: http://trix.docpatient.net/files/others/K77_guideline_mj_studydata_part2.pdf
 22. Eco U. *Mouse or Rat? Translation As Negotiation*. Philippines : Phoenix, 2004.
 23. Long W. Extracting diagnoses from discharge summaries. In *Proceedings of AMIA Annual Symposium 2005*, pp. 470–74. PMID:PMC1560678.
 24. Fan JW and Friedman C. Word sense disambiguation via semantic type classification. *Proceedings of AMIA Annual Symposium 2008*, pp. 177–81. PMID:PMC2655949.
 25. Gu HH, Hripcsak G, Chen Y, Morrey CP, Elhanan G, Cimino J et al. Evaluation of a UMLS auditing process of semantic type assignments. *Proceedings of AMIA Annual Symposium 2007*, pp. 294–98. PMID:18693845; PMID:PMC2655790.
 26. Canguilhem G (Ed). *The Normal and the Pathological*. New York, NY: Zone Books, 1991. PMID:1795225.
 27. Lamberts H. In het huis van de huisarts. [In the home of the home doctor]. *Verslag van Het Transitie Project*. Meditekst: Lelystad, 1991.
 28. Bowker GC and Star SL (Eds). *Sorting Things Out: Classification and Its Consequences*. Massachusetts, Cambridge, MA: MIT Press, 2000.
 29. McCracken EC, Stewart MA, Brown JB and McWhinney IR. Patient-centered care: the family practice model. *Canadian Family physician* 1983;29:2313–6. PMID:20469404; PMID:PMC2153697.
 30. Tierney WM, Overhage JM, Takesue BY, Harris LE, Murray MD, Vargo DL et al. Computerizing guidelines to improve care and patient outcomes: the example of heart failure. *JAMIA* 1995;2(5): 316–22. PMID:7496881; PMID:PMC116272
 31. Thorne C, Cardillo E, Eccher C, Montali M and Calvanese D. The VeriCliG project: extraction of computer interpretable guidelines via syntactic and semantic annotation. *Proceedings of the Workshop on Computational Semantics in Clinical Text (CSCT)* (19 March Potsdam, Germany), 2013.
 32. Sonnenberg FA, Hagerty CG, Acharya J, Pickens DS and Kulikowski CA. Vocabulary requirements for implementing clinical guidelines in an electronic medical record: a case study. *Proceedings of AMIA Annual Symposium 2005*, pp.709–13. PMID:1560696/