Medical terminology and discourse

Joost Buysschaert

1 Introduction

The translation of medical texts poses several challenges. An obvious difficulty for medical translators is the understanding of the original text, as not all translators have had extensive medical training and the topic may be very technical. Yet, as Vandaele (2001a) has pointed out, it is often sufficient for the translators to familiarise themselves with a number of core concepts of the subject at hand. A second difficulty related to the more technical documents, such as research papers, is constituted by the syntactic complexity that is typical of ‘special languages’ (also called ‘specialised languages’, ‘languages for special/specific purposes’, ‘LSPs’ - cf. Cabré, Sager and DeCesaris 1999: 228). However, by far the most challenging aspect of medical source texts is their abundance of medical terminology. In many cases, the brunt of the translation work consists of establishing appropriate target equivalents for the technical terms.

The present chapter explores several aspects of this terminological challenge and starts with an analysis and illustration of the remarkably rich terminological variation in the medical domain. Next, the possible solutions of term normalisation or standardisation are discussed, with an overview of the main competing efforts in this respect. Term variation can, however, be functional and this is illustrated in a section on discourse issues, which illustrates how medical communication functions in a variety of discourse settings, necessitating different terminological choices. The chapter goes on to discuss the role that language technology can play in the management of medical terminology and closes with some general hints on the types of sources that may be useful in researching the meaning, usage and translation of medical terms.

2 Terminological variation as a major challenge

Not only is terminology abundant in medical texts; it is also characterised by sometimes confusing variation, with different terms referring to the same concept or the same term referring to different concepts. The present section illustrates this, starting with examples of synonymy in medical terminology. Next, the importance of usage differences is discussed,
followed by an illustration of polysemy, homonymy, the frequent use of abbreviations, variation in spelling, and the occasional confusion between hyponyms and hypernyms. The search for translation equivalents can also be hampered by false friends, outdated variants and lexical gaps. Each of these obstacles will be illustrated and its impact on medical translation discussed.

2.1 Synonymy

Medical dictionaries often mention one or several synonyms for the same medical term. An extreme but by no means exceptional example may be refractory epilepsy, which is also known as drug refractory epilepsy, drug resistant epilepsy, medication resistant epilepsy, pharmaco-resistant epilepsy and intractable epilepsy. There are various reasons why synonymy is common in medical terminology. The first one is the existence of a term with a Latin or Greek root alongside a term in the vernacular, both of which are in use: e.g. os sacrum - sacred bone; coryza – common cold; epistaxis – nosebleed; pertussis – whooping cough. The second reason is the competition between eponymous terms and descriptive terms, particularly in the case of names for diseases and syndromes. Eponymous terms are named after the person(s) who first discovered or described the condition (e.g. Caffey-Silverman syndrome) or after the place where the condition was (first) encountered (e.g. Lassa fever; an eponym of this kind is also called a geonym). In many cases, eponymous terms also have a descriptive synonym; in the case of Caffey-Silverman syndrome, the corresponding descriptive label is infantile cortical hyperostosis. Some medical terminologists favour the descriptive terms because of their self-explanatory nature; others point out that the eponymous terms are easier to remember and, in the case of reference to persons, give credit to those who brought the condition to the fore. The issue has been heavily debated but the synonymy persists (see Whitworth 2007 versus Woywodt and Matteson 2007; see also Cappuzzo 2008). Synonymy may also arise when there are different schools of medicine with their own preferences, as also described by Bowker and Hawkins (2006), who name conceptual, linguistic and social factors as causes of medical term variation.

In choosing a preferred synonym, usage differences come into consideration (as discussed below), as does frequency of use in (recent) authoritative sources. Major languages will have respected medical journals that can be queried for clarification, or authoritative medical dictionaries and normative term collections (of the types discussed below under the heading of normalisation and standardisation).

2.2 Usage differences

The existence of synonyms does not guarantee that the variants are mutually interchangeable in all discursive contexts. Some synonyms are appropriate as informal lay terms in doctor-patient exchanges but would be replaced by more technical names in research articles. An example is the term belly versus abdomen. Languages differ in the number of lay terms they offer. In French, for example, many medical terms that are used in formal discourse are also used as lay terms; whereas English has collarbone alongside clavicle, French only has clavicule. Dutch, on the contrary, has many lay terms. When translating a patient-oriented text from French into Dutch, the translator will have to avoid the
systematic use of Dutch cognates of the French terms in contexts where lay alternatives may be more appropriate.

Apart from differences in register (technical versus lay language), translators also encounter regional usage differences. For example, British English uses noradrenaline, whereas American English calls the same hormone norepinephrine. Similarly, differences exist between European French and Canadian French and between Netherlandic Dutch and Belgian (Flemish) Dutch, and translators have to be aware of these differences so that they can adapt the text for their target readership. The task becomes especially difficult when the readership is mixed, as with the Dutch translations commissioned by the European Medicines Agency, which are meant to be used in the Netherlands as well as Belgium. In some cases, it may be advisable to resort to alternative solutions such as the use of slashes or parentheses in order to include variants in the text.

2.3 Polysemy and homonymy

Polysemy occurs when a given term has more than one meaning and the meanings are related. Polysemy is fairly common in non-technical language but should be avoided in special languages when possible. However, several examples of polysemy are found in medical terminology. A well-known case is the term drug, which may refer to a therapeutic agent but also to a stimulating or depressing substance that is potentially addictive. In the case of a polysemic expression in the source text, it is important to ascertain if the target language also uses one term for both meanings, or whether it uses distinct names.

Homonymy refers to those cases where terms share pronunciation and spelling but have different meanings that are not related in an obvious way. A medical example is the term callus, which may refer to the formation of extra tissue at the site of a bone fracture but may also mean ‘hardened skin’. There is likely to be a common etymological origin for the two meanings but the link between them is more distant than in the case of polysemy. Ideally, specialised languages should again avoid this kind of ambiguity but it may be argued that the language of medicine has evolved into a combination of various special languages (in the case of callus, the competing special languages are those of orthopaedics and dermatology). It is again the role of the translator to ascertain if the ambiguity persists in the target language, or whether a choice must be made between distinct equivalents.

2.4 Abbreviations

The use of abbreviations is very common in medical texts. Abbreviations may be regarded as synonyms of the full forms that they stand for. In a number of cases, the abbreviations have become more common than the full forms, like AIDS for acquired immune deficiency syndrome. However, as the domain of medicine is a combination of many subdisciplines, some abbreviations have several different meanings. The online mediLexicon gives 16 possible meanings for the abbreviation OCM, ranging from obstructive cardiomyopathy to ovine choriomammotrophin.

Finding an equivalent abbreviation in a target language often requires a detour via the full form. In cardiology, CPR stands for cardiopulmonary resuscitation. The full term is translated into French as réanimation cardio-pulmonaire, and a further search for the French term reveals that the abbreviation RCP is in use. The German equivalent is Herz-Lungen-Wiederbelebung, which is abbreviated to HLW.
It is not uncommon for English abbreviations to be used in other languages even though the full forms of the terms are different. The English abbreviation ACTH for adrenocorticotropic hormone frequently appears in French texts, despite the fact that the full French term is hormone corticotrope. English acronyms, i.e. abbreviations that are pronounced as a word, are often used unchanged in other languages when they are (relatively) recent coinages (though pronunciation and spelling may be adapted to the target language). AIDS is an example that has already been mentioned; the acronym is also used in many other languages, including German (AIDS), Dutch (aids) and Swedish (aids); French and Spanish use sida (short for syndrome d’immunodéficience acquise resp. síndrome de inmunodeficiencia adquirida). A similar example is SARS for severe acute respiratory syndrome.

2.5 Variation in spelling

There are a number of differences between British and American spelling in English medical terms which the translators have to take into account if their target language is English. Common cases include (with UK spelling followed by US spelling):

- ae versus e: anaemia – anemia
- oe versus e: oesophagus – esophagus
- -re versus -er: titre – titer
- -our versus -or: tumour – tumor

In other languages, too, there may be variation in the degree to which originally Latin spellings have been adapted to spelling rules in the vernacular versions. The difficult Late Latin spelling diarrhoea (a term of Greek origin) survives in British English, loses an -o- in American spelling, becomes diarrhée in French, loses its h and is shortened to diarré in Danish; and is spelt diarrée in Dutch. Sometimes vernacular spellings are used alongside original spellings in the same text, for example when the Latin name of a condition is added in parentheses. A Dutch source may refer to artrose (arthrosis), where the vernacular form uses -t- and Latin -th-.

Some languages have also introduced spelling reforms in recent decades (amongst them German and Dutch). Older sources that are still available will use older spellings and even newer sources may not have fully adapted. A particularly thorny issue are the spelling rules for compounds and eponyms introduced in the 1995 Dutch spelling reform. The official rules prescribe downsyndroom but the alternative Down-syndroom has also remained in use. In 2011, a team of medical terminologists around the editor of the well-known medical dictionary Pinkhof opined that the older spellings should remain a valid alternative in medical texts (Van Everdingen and Van den Eerenbeemt 2012). The authoritative journal Nederlands Tijdschrift voor Geneeskunde, on the contrary, chose to adhere to the new official rules.

2.6 Hyponyms and hypernyms

Another common issue when translating medical terminology is that common subtypes of illnesses (hyponyms) are sometimes confused with main types of the condition (hypernym). Many sources list variola major as a synonym of smallpox, whereas the term
only refers to the more common form of smallpox. The translator has the difficult task to decide whether it is acceptable to go along with this blurring of fine distinctions or to distinguish more rigorously between hyponyms and hypernyms in the translation.

2.7 False friends

Some terms may appear similar but have different meanings in other languages. The word *surface* is used in English as well as in French, but the *anterior surface* of an intraocular lens is not *la surface antérieure* but *la face antérieure*. Similarly, a *median section* is not *une section médiane* but *une coupe médiane*. A better-known example of false friends between French and English are the words *preservatives* (substances used to prevent decay) versus *préservatifs* (condoms).

2.8 Diachronic variation

Translators should be aware that terms they find in (especially printed) dictionaries may have been superseded by new terms. This diachronic variation may have different reasons: new normalisation efforts, changes in classification reflecting new research, or the avoidance of misnomers after research has proved that the old name gives an inaccurate description. The term *Lyme arthritis*, for example, is no longer in use for *Lyme disease*, because it is an inaccurate, or at best only a partial, description of the condition. An example of successive changes in classification together with an attempt at normalisation can be found in the types of epileptic seizures. Although the term *grand mal* was replaced a long time ago by the newer classification *tonic-clonic seizure*, the old term of French origin is still common.

2.9 Lexical gaps

Rogers (2015) aptly describes the phenomenon of lexical gaps (or terminological gaps in terminology) as a common phenomenon in specialised translation. A lexical/terminological gap occurs when there is no equivalent of the source term in the target language. In the case of medicine, this gap typically occurs when a newly emerged concept receives its name in only one language, now most commonly English (*stent* or *regenerative medicine* are examples). When such a new term occurs only once in a translation, a provisional solution can be a paraphrase or a circumlocution. However, once the concept becomes established, it becomes necessary to find a fixed equivalent in the target language. Authors writing on the topic in their native languages, or translators, will then feel the need to come up with suggestions of their own, a process which is known as secondary term formation (Sager 1990: 80). There are three common methods of secondary term formation. The first is to simply accept the source term as a loan term, typically with minor adjustments to reflect the requirements of the target language (*stent*, originally an eponym, has been copied to many languages; German uses a capital letter [*Stent*]. Spanish adjusts the term slightly to *estent*). The second method is called loan translation, where component parts of the source term are translated separately and then joined together (which is the obvious choice for *regenerative medicine*). The third method of secondary term formation, which is rarely applied, is the creation of an entirely new term not based on the source; a notable example is *taaislijmziekte*, a neologism coined as the Dutch equivalent of *cystic fibrosis*. 
Resorting to loan translation can be risky in the case of multiword units. For a term like *total laryngectomy* a French translator may be tempted to opt for *laryngectomie complète*, whereas the phrase *laryngectomie totale* is the more common translation.

## 3 Normalisation and standardisation

### 3.1 The rationale behind standardisation

The proliferation of terms that denote the same medical concepts has led to various attempts at establishing standards or norms for use within a particular context. The efforts have resulted in coded term collections presented in various formats, with tree diagrams, thesauri, controlled vocabularies and ontologies amongst the most popular ones. In what follows, the neutral label ‘term collections’ will be used for all of them, as there can be confusion as to the exact meaning of the denominations that have been used to refer to the various term repositories (Vanopstal et al. 2011). What they have in common is that they try to establish preferred terms for each concept and that they assign a code (for example a number, or a combination of letters and digits) to each concept. The code serves as a unique identifier for the concept and binds all the intralingual variants that refer to the same concept and even the translations (if supplied). The overview that follows introduces some of the best-known medical term collections that aim at standardisation, and indicates how these collections can be used by a medical translator in finding a preferred term among synonyms, in finding definitions and sometimes in finding translations.

### 3.2 Selected examples of term collections

One of the most popular early initiatives of standardisation was the *Nomina Anatomica*, which listed standard names in Latin for human anatomical parts. It started off in 1895 with a version now known as the *Nomina Anatomica Basiliensia*; subsequent versions include the *Jena Nomina Anatomica* of 1933 and the *Paris Nomina Anatomica* of 1955. It was widely used between 1956 and 1989, when it was replaced by the *Terminologia Anatomica* (TA). The TA covers more than 8,000 concepts and is available in Latin and (American) English; it was supplemented with the *Terminologia Histologica* in 2005. The collection can be consulted online.

Another classification with historical roots is the *International Classification of Diseases* (ICD). Its precursor, the *International List of Causes of Death*, was adopted by the International Statistical Institute in 1893. As the title suggests, the list was intended for use in death certificates. The World Health Organisation took over the management of the list in 1948, and renamed it as *International Classification of Diseases* (ICD, World Health Organisation 2016). The ICD became the international standard for reporting diseases and health conditions. The 10th version, known as ICD-10, was released in 1990 and has been in use in over 100 countries. It is available in 42 languages, and the English and French versions can be consulted online. Meanwhile, ICD-11 has also become available and is expected to be implemented by WHO member states by January 2022.

A valuable medical vocabulary resource of much more recent origin, provided by the United States National Library of Medicine, is the *Medical Subject Headings*, abbreviated as MeSH. It is structured as a hierarchically organised term collection, intended for indexing and cataloguing biomedical information. MeSH was started in the 1960s as an indexing tool for the bibliography *Index Medicus* and is now used in particular to index
and search publications in the extensive MEDLINE/PUBmed database. Its terms are called ‘headings’ or ‘descriptors’ and their unique IDs have been used for other purposes as well, including the structuring of medical libraries. The MeSH is constantly updated and is freely available via the MeSH Browser. Translators can use the MeSH to find reliable (American) English medical terms. Apart from the ‘descriptors’ (i.e. the preferred terms), the Browser also supplies ‘scope notes’, which are in most cases definitions, as well as ‘entry terms’, which are synonyms or closely related terms.

The MeSH has been translated into several languages but the translations do not always cover the complete collection or do not keep track of the annual updates. Some of the translations have been criticised for having an English bias (Vandaele 2001b: 113) but they nevertheless offer a useful starting point. The translations can be found via the Unified Medical Language System (UMLS) or via Babelmesh, which offers a dozen languages. The French version is updated annually by Inserm, the Institut national de la santé et de la recherche médicale, and is also available via Le MeSH bilingue.

Another collection that has established itself as an important player since 1999 is SNOMED CT (with CT meaning Clinical Terms), not in the least because it has been the nomenclature of choice in some of the eHealth initiatives (Cangioli et al. 2016). SNOMED CT also has the advantage of being the largest collection of its kind, covering more than 300,000 unique concepts and many medical specialties, and of being computer processable. SNOMED CT aims to provide the core general terminology for electronic health records. The collection, originally in English only, can be accessed via the SNOMED CT International Browser. A search for a medical term yields a concept code, a description, hypernyms, hyponyms and other relationships; in the case of synonyms some will be labelled as ‘preferred’, others as ‘acceptable’. Several translation projects are in progress to make the collection available in other languages; the SNOMED CT International Browser gives access to some of them.

SNOMED CT can also be consulted via the SNOMED CT browser of UTS. UTS stands for UMLS Terminology Services, where UMLS is short for Unified Medical Language System; it is one of the many useful resources provided by the United States National Library of Medicine. The UTS also contains the Metathesaurus Browser. A search for an English medical term in the Browser offers a concept code, definitions and in some cases also translations (from the MeSH translations).

3.3 Interoperability of standards

As explained, there are many competing standards in the field of medicine, each with their own proponents. The account above is limited to the historically best-known collections and to those that cover medicine in general, leaving out other nomenclatures that focus on specific medical areas. HeTOP is an interesting portal that lists many more sources and allows simultaneous searches in several of them.

Some other attempts have been made to link systems. One example is I-MAGIC, which searches for a matching ICD-10 code starting from a SNOMED CT term. For example, typing in the term *acute necrotizing pancreatitis* first yields the SNOMED CT code for this concept (which is 7881005); clicking on ‘Get ICD codes’ shows the code of the same concept in ICD 10 (K85.91) and reveals that ICD 10 calls the condition *acute pancreatitis with uninfected necrosis, unspecified*. Technical committees of the ISO and the European Union are also developing standards that aim to promote compatibility and interoperability between healthcare systems (ISO/TC 215 and CEN/TC 251 in particular).
3.4 Standardisation in drug licensing documents

Another type of terminological standardisation is the one imposed on texts required for drug licensing purposes. A typical example are the templates used by the European Medicines Agency\(^\text{21}\) for the documents needed to submit a licensing application. The templates are available in all of EU’s official languages and standardise the way in which documents like Summaries of Product Characteristics or Package Leaflets have to be formulated and translated. Translation of the licensing documents is mandatory for the medicinal products in order to obtain marketing authorisation in all EU member states. This not only means that, for example, all the headings in a Summary of Product Characteristics have standard translations; it stands to reason that terms used in these headings should also be translated in the same way in the rest of the text. If a term such as *excipients* was translated in one way in the heading and in another way in the body of the text, this might cause confusion.

4 Discourse issues

When choosing appropriate medical terminology, it is important to know the target readership (or audience) of the text. Medical communication can function in a variety of discourse settings, each with its own requirements. This section will examine the following settings: communication between health professionals; communication between specialists in a subfield; communication between health professionals and patients or the general public; mixed scenarios; and interpreting doctor-patient exchanges.

4.1 Communication between health professionals

Much of the medical translation work is intended for domain specialists, i.e. readers who are themselves doctors or medical professionals. Examples include documents needed to obtain marketing authorisation for medicines, as already hinted above; research papers and their abstracts, which are also occasionally submitted for translation; and medical reports, including discharge summaries, operative reports or medical imaging reports, some of which need translation when submitted for insurance claims. The density of technical terms in these texts is striking.

Communication between doctors or other health professionals will naturally favour the more technical terms over the lay alternatives. However, given the challenges of term variation discussed above, translators should be aware of their clients’ preferences even among the technical terms. It is not unusual for clients to criticise translations simply because their own preferred terms have not been used. In such cases, it may be helpful to compile a list of terms as found in the source texts, supply them with suggested equivalents in the target language and have the client vet the equivalents either prior to translation or prior to revision. Term extraction software (like Multiterm Extract,\(^\text{22}\) SynchroTerm\(^\text{23}\) or SketchEngine,\(^\text{24}\) among several others) can be helpful in identifying technical terms in the source texts.

Clinical guidelines are another, now increasingly common example of medical communication aimed at health professionals. The guidelines offer evidence-based advice on a wide range of topics. Among the best-known are the Cochrane Systematic Reviews.\(^\text{25}\)
Further guidelines of various origins can be consulted via Guidelines International Network. There is a demand for the translation of these authoritative guidelines and the correct rendering of technical terms can again become an issue. An interesting account of the challenges encountered during the translation of the Cochrane reviews (or their abstracts) is found in von Elm et al. (2013).

4.2 Communication between specialists in a subfield

As explained above, medicine combines many subdisciplines and each may well have specific terms that are not in wider use. The subdiscipline of medical imaging may be quoted as an example, as this field uses many specific abbreviations. Translators may be called upon to translate imaging reports in the context of insurance claims, when they are likely to struggle with these abbreviations. The first point of enquiry is the meaning of the abbreviations in their source language; for English, the Royal College of Radiologists has a patient-oriented site which explains a number of common abbreviations, but the actual number of abbreviations in use is much larger. Translation of medical reports is best regarded as a specialist area and translators who want to work in this area will need to develop their own termbases with abbreviations, their explanations and equivalents in the target language.

4.3 Communication between health professionals and patients or the general public

The Internet has an abundance of patient-oriented sites; even before it came into being, patients were addressed directly through medication leaflets, variously called patient package inserts, patient information leaflets or package leaflets. The question of how to make these leaflets patient-friendly has been a long-standing issue (cf. Kenny et al. 1998 or Gal and Prigat 2005, among many other publications; see also the European Commission’s Guideline on the readability of the labelling and package leaflet of medicinal products for human use (2009). Problems that have been identified range from print size to the risk of misinterpretation, but the issue of clear medical terminology has inevitably been also part of the debate.

More recently, online patient-oriented medical guidelines have begun to appear, and they face the same problems of readability as medical leaflets. Åhfeldt et al. (2006) offer a literature review on patient-friendly documentation in general, not just leaflets, illustrating many of the challenges. Specific difficulties of translating lay-friendly medical information are discussed in Askehave and Zethsen (2011), amongst other publications. Wermuth and Verplaetsé (2019: 99–104) describe the transition from technical medical documentation to lay documentation as a form of intralingual translation.

The website Multilingual Glossary of Technical and Popular Medical Terms offers suggestions for lay equivalents to medical terms in nine European Languages. However, it is important to keep in mind that when patients are diagnosed with a particular disease or condition, they are likely to remember the term, even if it is technical, and there will be no need to replace it with a vague lay alternative in the patient leaflet or patient guideline. For example, the Glossary offers serious skin disease as a lay alternative for pemphigus; this could be used as an explanatory description in some contexts, but pemphigus can be retained in leaflets aimed at patients who suffer from the condition.
4.4 Mixed scenarios

The term eHealth (sometimes spelt e-Health) has been a buzzword for some time now and refers to electronic support in healthcare practice. One important aim has been to make patients’ records available via online networks, for the benefit of other caregivers as well as for the patient. The electronic health record (EHR) or electronic medical record (EMR) can then serve a double purpose: fast communication between health staff as well as ‘patient empowerment’, as the patients also have access to their medical information. However, this double purpose presents its own set of challenges including the aforementioned term variation. Grön and Bertels (2018: 42) have noted that EHRs mix specialist terms with lay terms and even informal terms. Specialist terms might be unsuitable for the patients and limit their agency in making medical decisions; informal terms may be unsuitable for health professionals who might not be familiar with the expressions used. Avoiding these problems would need large termbases listing all the variants and making terms clickable in the records so that synonyms can be provided instantly. Similarly, if EHRs are meant to be available across language borders (so that a Spanish doctor can consult a Danish holiday maker’s record), it would be useful if translations could likewise be supplied via clicking on problem terms. To make this happen, substantial work must be carried out by medical terminologists worldwide. Similar efforts have already been undertaken to support multilingual searching (Diekema 2012); an example of a cross-lingual search facility developed for PubPsych, a search engine in psychology, is discussed in España-Bonet et al. (2019). The facility also uses MeSH terminology.

4.5 Interpreting doctor-patient exchanges

A discourse setting that has not been discussed above is where an interpreter mediates between a doctor and a patient during a medical interview. This can be a particularly challenging experience, especially if the doctor uses technical jargon and when the cultural context of the doctor and the patient are very different. Research by Hadlow and Pitts (1991) shows that even in monolingual settings, medical staff and patients have a different understanding of common health terms. In a bilingual setting, the interpreter may try to translate technical terms using approximate lay alternatives in the target language, and risk blurring the message. Cultural considerations also play a role (cf. Kersey-Matusiak 2018). A particularly difficult problem arises when some of the topics or terms are considered taboo in the target culture, and have to be avoided through the use of paraphrases and euphemisms (Munane 2014). The International Association for Communication in Healthcare is expected to publish a position paper on healthcare communication across language barriers in 2020 (Van den Muijsenbergh, Krystallidou, and Langewitz, forthcoming).

5 Medical terminology and translation technology

There is no doubt that present-day translation technology can be very helpful to medical translators, particularly for the translation of medical terminology. Most Computer-Aided Translation (CAT) programmes offer the option to store terminology in termbases which can be automatically searched during the translation process. The technology also
allows the labelling of medical terms in the termbase so as to indicate their usage such as ‘technical’, ‘lay’, or ‘informal’, as well as to indicate regional use or clients’ preferences.

Automatic term extraction has been alluded to before; given the density of terms in a typical medical source text, it is advisable to first run a term extraction, check for potentially problematic terms and add them to the termbase before embarking on the translation. In the case of English, the Medical Subject Headings site offers a ‘MeSH on Demand’ facility, which identifies MeSH terms in submitted texts. Term extraction in medical corpora on a broader scale and involving machine learning techniques is sometimes called ‘biomedical named entity recognition’ (Krauthammer and Nenadic 2004; Campos, Matos and Oliveira 2012).

Translators increasingly make use of machine translation (MT), either for partial input during the translation process or for wholesale pre-translation followed by post-editing. MT output quality tends to vary considerably and will depend on language combinations and topic. Major MT providers like Google Translate and DeepL tend to perform reasonably well for medical translations between widely used languages and on a variety of medical topics, provided that the engines were trained on a substantial corpus of relevant bilingual material. However, medical terminology can again present a stumbling block. Typical problems that translators and post-editors have to be aware of are, a) inconsistent translations of the same term within the same text (i.e. one term used in a paragraph, a synonym used in the next paragraph or even the next sentence); b) wrong rendering of some multiword terms as the engine may translate the words separately and combine them into an expression that looks acceptable but is incorrect in the target language; and c) translation by means of a hypernym. All three types of mistake tend to be difficult to spot without relevant training. It is again advisable for translators and post-editors to use MT within a CAT programme that also allows for the use of their own termbases, as these can help to correct some of these terminological errors.

Alternatively, translators can build their own translation engines and train them on their own material, which in principle should also mean that their own terminological preferences will be picked up. However, it is difficult for such custom engines to match the overall quality of the major MT providers. Another issue is the integration of existing termbases into MT engines. In the case of phrase-based statistical MT engines, terms can be entered in the bilingual ‘phrase tables’, but it is not possible to add synonyms with the necessary usage information. If the engine is to be used for a variety of translation tasks (technical as well as patient-oriented), the method will give insufficient support. Neural MT offers even fewer opportunities for integrating terminology as it primarily focuses on fluency and in many cases prioritises natural-sounding target sentences. Although terminology could be entered in phrase tables, or a termbase could be added to the training material, Neural MT is known for regularly ignoring information supplied in this way. Recent research now aims to circumvent this obstacle (Dinu et al. 2019) but is unlikely to solve the problem of term variation if the engine is to perform translation tasks aimed at different readerships.

6 Sources of medical terminology

There is no dearth of monolingual medical dictionaries for the major languages and bilingual dictionaries are also available for a number of language combinations, as a search via the Internet or a library catalogue will show. Traditional paper dictionaries have the advantage of being reviewed by editors and overseen by well-established publishers, but may not list recent terminology in rapidly evolving medical subdisciplines. On the other
hand, online medical glossaries and dictionaries can be of varying quality, but some can be highly relevant through their focus on very specific topics. An Internet search for a topic, adding search terms like glossary (for example: neurology lexicon OR glossary OR dictionary) often provides a wealth of resources, at least for the well-documented (especially Western) languages. Definitions can also be found via the Internet (define + search term often yields pertinent results in Google, and not just for English). Authoritative definitions for English medical terms can be found via sources like MeSH and UMLS, mentioned above. Translation suggestions can sometimes be found via Wikipedia (when there is a parallel article for the target term) or through one of the many online sites searching translation memory collections (Linguee, Glosbe, TAUS Data); sites like ProZ.com can also provide relevant information. It is always recommended to check the validity of these translation suggestions in reliable primary sources. Sites like Google Scholar or Google Books are valuable resources for confirming the use of technical terms in the target language, so is MEDLINE/PubMed. For English, BioMedSearch covers PubMed as well as medical dissertations and a variety of publications. Respected medical journals in the target language, also often available online, are likewise relevant resources for confirming a term and its use. Many of the sources mentioned can also be used to compare frequency of use of competing synonyms.

7 Conclusion

Supplying correct equivalents for the many technical terms in medical texts is a major challenge for medical translators (and interpreters). The problem is compounded by terminological variation: many medical concepts have more than one name within the same language. Although normalisation efforts exist, the different attempts tend to compete with each other. Moreover, terminological variation can be functional, for example when a lay synonym is more appropriate than the technical term in a text aimed at patients.

Future research will therefore have to concentrate on the best design of tools that reflect variation as well as functionality. The ideal tool will link all the terms referring to the same concept, in one language as well as across languages, and will in addition label variants to make clear in what discourse settings they are appropriate or preferred. This can for example lead to enriched termbases for use in a CAT tool; or to thesauri that boost search engine performance. Combining the enriched term collections with machine translation is a further challenge that will need to be addressed. For a number of well-documented languages, much essential information is already available, though it may be dispersed among various sources and will need to be aggregated. For other languages, the groundwork does not yet exist, or it needs further expansion.

Notes

3 On the topic of medical nomenclatures, clinical terminologies and coding systems, see also Wermuth and Verplaetse (2019: 90–97).
Further reading


The second section of this article deals with terminological issues as discussed in the present chapter; section 4 expands on the issue of retaining lay-friendliness in patient-oriented translations.


Section 4 of this chapter gives another account of medical nomenclatures, clinical terminologies and coding systems. Section 6 offers further illustration of the challenges involved in converting technical medical documentation to lay documentation.

ITI Medical and Pharmaceutical Network (no date) Available at: www.itimedical.co.uk

Medical translators can gain substantial support by joining dedicated networks, like the one of ITI (Institute of Translation and Interpreting). Most of the discussions on their e-group deal with terminological issues.
This is an interesting blog by the Spanish to English medical translator Emma Goldsmith. Several of its regular posts deal with terminology (type in ‘terminology’ in the search box to find previous posts).

The research group Biomettico of the University of Montreal has produced several worthwhile publications on medical terminology; references can be retrieved via the ‘Publications’ section of their site, using the keyword ‘terminology’.

Related topics
Inter- and Intralingual Translation of Medical Information, Machine Translation in Healthcare, Quality, Accessibility and Readability in Medical Translation

References
ITI Medical and Pharmaceutical Network (n.d.) Available at: www.itimedical.co.uk (Accessed: 9 August 2019).