Leading healthcare terminology, worldwide

Guidelines for Translation of SNOMED CT

3.0

Publication date: 2022-03-14

Web version link: http://snomed.org/translate

SNOMED CT document library: http://snomed.org/doc

This PDF document was generated from the web version on the publication date shown above. Any changes made to the web pages since that date will not appear in the PDF. See the web version of this document for recent updates.

Table of Contents

SNOMED International

Leading healthcare terminology, worldwide

Web browsable version: http://snomed.org/translate

SNOMED CT Document Library: http://snomed.org/doc

1 Introduction

SNOMED CT (Systematized Nomenclature of Medicine Clinical Terms) is a comprehensive clinical **terminology** used to code, query and analyze health data. The terminology system is maintained and distributed by SNOMED International, the trading name of the International Health Terminology Standards Development Organisation (IHTSDO), established in 2007.

SNOMED CT provides a basis on which health care organisations can plan and document health care processes, conduct outcome research, analyze health care quality and costs, and develop effective treatment recommendations. It is the result of the merger of SNOMED RT (Reference Terminology) and Clinical Terms Version 3. The terminology includes the concepts, terms and relationships necessary to accurately represent clinical information across the healthcare system. SNOMED CT contains more than **350,000** active **concepts**. Each concept can be represented by several **descriptions**. There are three types of Descriptions: Fully Specified Name (FSN), Synonym (Syn) and Definition. The concepts are arranged in concept classes with different meanings called hierarchies, covering information required in a health record such as symptoms and signs, disorders, operations, treatments, medication, administrative elements, etc.

1.1 Purpose and scope of this document

The present guidelines are based on the SNOMED International Editorial Guide (http://snomed.org/eg) and are intended for translators of SNOMED CT. They have been developed by the members of the Translation User Group and summarise the main editorial principles of SNOMED International. They can be used as a starting point for the development of national guidelines for the translation of SNOMED CT from English into the national languages of the SNOMED CT member countries. Additional information on editorial rules can be found in the Confluence editorial templates (https://confluence.ihtsdotools.org/display/SCTEMPLATES/SCT+Modeling+Templates+and+description+patterns).

The present document is to be used in conjunction with the current version of the *Guidelines for Management of Translation of SNOMED CT* (see SNOMED CT Document Library https://confluence.ihtsdotools.org/display/DOC/SNOMED+CT+Document+Library) which contain recommendations for managing a translation project. Suggestions for amendments or additions to these guidelines should be sent to info@snomed.org.

It is essential that those involved in translation of SNOMED CT concepts and in the verification, validation and approval processes of new national descriptions are familiar with the terminological principles on which SNOMED CT is based. It is equally important that the translation work complies with the SNOMED CT Editorial Guide and that translators are aware of issues such as how to select terms and lexical variants, how to use the appropriate translation technique, and how to ensure linguistic consistency.

The present document also contains general recommendations on the steps involved in a translation workflow process, as well as information on source documents and references that are useful during translation.

1.1.1 Assessment of translation quality

Translation project owners (TPOs) should ensure that translations comply with the principles on which SNOMED CT was originally based: **understandability**, **reproducibility** and **usefulness** (URU). **Understandable** means that the meaning of a concept can be understood by most healthcare providers without reference to private or inaccessible information; **reproducible** means that several users apply the concept to the same situations; **useful** means that the concept has a practical value for the users that is self-evident or can be easily explained. Furthermore, the information contained in the translated concepts must be semantically equivalent to the information contained in the core source terminology (International Release). SNOMED International has developed a document that describes how to assess the quality of the translation and the degree of consistency. The document entitled *A methodology and toolkit for evaluating SNOMED CT translation quality* (2012) outlines and defines a set of requirements or “quality characteristics” (QCs) and related metrics.

There are three types of QCs for assessing translation quality:

- Structural QCs, which cover the management and organization of the translation project;

- Process QCs, which refer to the activities during the actual translation process;
- Result QCs, which refer to the translation result in the target language.

In this document, a shortlist of (9) QCs is drawn up and for each QC quality score (what and how to measure, how to evaluate, etc.) there are sample questionnaires (to support the evaluation) as well as a general “rating”.

It is recommended that TPOs refer to Annex A of the current document and to the Methodology and Toolkit document to ensure that they include all quality metrics identified as **mandatory** for use in their project and quality plans. Metrics for measuring translation quality can be reformulated as contractual clauses in formal agreements between a TPO and a Translation Service Provider (TSP) on the services to be provided and the level of service.

1.1.2 Change requests

Requests about inconsistencies, errors or proposed changes to the core terminology (International Release) should be submitted through Content Request Service (CRS) (https://confluence.ihtsdotools.org/display/SCTCR/Accessing+CRS).

1.2 Target group

The target group of this document are all persons directly involved in the translation of SNOMED CT into another language, i.e. translators, reviewers, domain experts, validators as well as managers and members of an editorial board or equivalent group of experts who define the linguistic and terminological guidelines for the specific translation of the target language.

2 An introduction to terminological principles

Understanding of terminology and specialized knowledge representation is a key factor for successful scientific translation. SNOMED CT translations must therefore be based on **terminological principles** developed by Terminology Science (TS). This discipline is a branch of linguistics that deals with specialized vocabulary. It focuses on the study of lexical units (words) with a specialized meaning called **terms**. Terms are analyzed from the point of view of their origin, formal structure, meaning and their functional characteristics. Since terms are used to denote **concepts**, TS also deals with the formation and development of concepts and relationships between concepts and their classification in **concept systems**. Furthermore, TS formulates the principles of concept **definition** and the evaluation of existing definitions. In view of the fact that the properties and functioning of terms depend to a large extent on their lexical environment, it is common practice to consider not individual terms but the entire terminology used in a particular field of knowledge (also called subject area) as the main subject of TS.

In the following, we will give a summary description of concepts and terms, concept systems, and definitions. Please see the international terminology standard ISO 1087-1:2000 for a systemic description of these notions.

2.1 Concepts and terms

The **semiotic triangle** (Figure 1) introduced by Ogden and Richards (1923/1989) is of central importance in terminology work. This triangle is a model of how symbols are related to the objects they represent. It consists of 3 elements: concept ("**Thought or Reference**"), designation ("**Symbol**") and object ("**Referent**"). A **concept** is the mental image of concrete or abstract objects in reality and consists of characteristics based on properties observed in an object or in its relationship to other objects. Concepts are represented by a **designation** (symbol) such as words (in general language), **terms** (in specialized language) or other perceivable notations (e.g. graphics, numbers, etc.). In SNOMED CT the concepts are represented by so-called **Descriptions** (see Section 3.4).

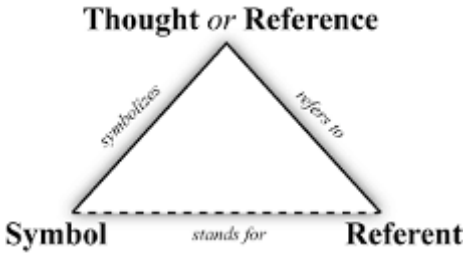

Figure 1 – The semiotic triangle (Odgen and Richards 1923/1989)

The dotted line between the referent (object) and the symbol (description) indicates that there is no direct connection between the two, as the object must be conceptualised before it can be rendered. There is also no one-to-one (i.e. one term refers to one concept) but a many-to-one relation between concept and term (different terms refer to one and the same concept). Examples are the English terms "typhoid fever" and "infection by Salmonella Typhi" for the concept |typhoid fever|. This is called **synonymy**, which is a frequent phenomenon in clinical terminology. Conversely, one term can refer to different concepts (i.e. one and the same term has different meanings), which is called **polysemy**. For example, the term "drug" refers, depending on the context, to the following concepts: (1) |substance used for diagnostic/treatment/ prevention purposes|; (2) |chemical substance used as narcotic or hallucinogen|; (3) |substance used in dyeing or chemical operations (obsolete)|.

2.2 Concept systems

The main goal of terminology work is to clarify and standardize concepts and terms to facilitate unambiguous communication in natural (human) language. Concepts are related to other concepts (conceptual relations) and form networks or **concept systems**, which are extremely useful in many contexts: they provide an overview of a subject area and make it possible to delimit and define concepts, to form new descriptions, to evaluate existing and competing descriptions, and to structure systematic representations of a given domain. Since concept systems are important tools for analysing and translating concepts, they play a crucial role in terminology work.

According to ISO 1087-1:2000 **conceptual relationships** can be divided into hierarchical relationships (generic is-a relationships), partitive relationships (whole-part relationships) and associative (non-hierarchical) relationships. In SNOMED CT generic and attribute relationships are used to define concepts in a formal language that is understandable to both humans and computers (Figure 2).

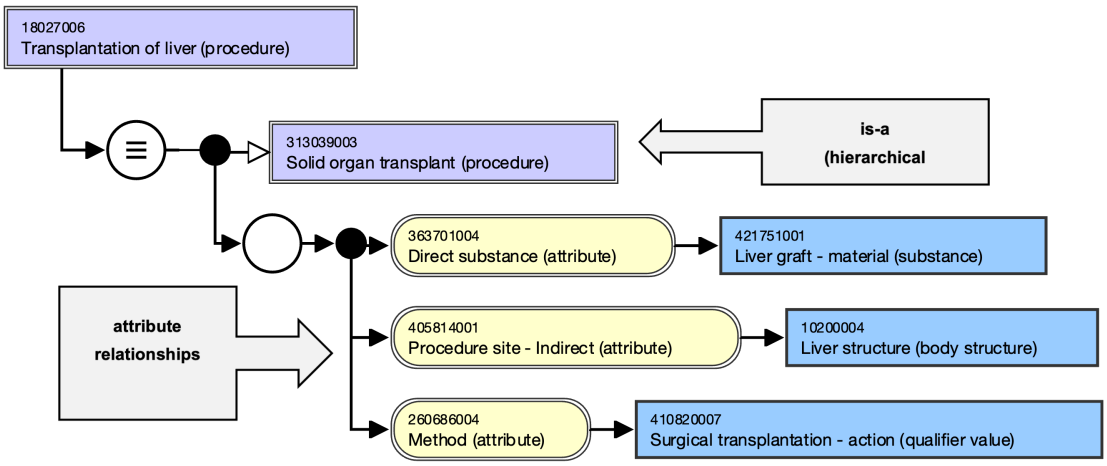

Figure 2 – Hierarchical and attribute relationships in SNOMED CT; example with the stated view of concept 18027006 |Transplantation of liver (procedure)| (International Edition 2021-07-31)

2.3 Definitions

In terminology work, the definition of a concept is extremely important. A definition describes the characteristics of a concept to delimit it from neighboring concepts. Definitions are therefore essential for the clear assignment of terms to the underlying concepts. There are different types of definitions (analytical, encyclopaedical). Definitions can be written in natural language or in a (semi-)formal language.

Definitions play a key role in the translation of SNOMED CT (see Section 4). Descriptions cannot be translated literally but their meaning (i.e. their conceptual content) must first be clarified. This is the only way to ensure that the meaning of the description is correctly reflected in the target language. In SNOMED CT concepts are defined in different ways (see Section 3.2).

3 SNOMED CT as a health terminology

The following sections outline the main features of the structure of SNOMED CT and particular aspects of interest to translators. More detailed information on the structure of SNOMED CT can be found in the SNOMED CT Guides (https://confluence.ihtsdotools.org/display/DOC/Guides).

3.1 Concept systems and ontologies

Conceptual relationships can be coded in formal structures or so-called **ontologies**. The construction of ontologies is based on conceptual networks derived from **expert knowledge**. The aim is to represent domain-specific knowledge (i.e. the conceptualization of a particular domain) in a formal, computer-readable and consensus-based way in accordance with criteria agreed upon by a community of experts. The elements of knowledge are defined according to a set of properties and linked via explicit relationships of different kinds (both hierarchical and non-hierarchical such as partitive, cause-effect, etc.). Since the concepts and relations are represented formally, the information can be processed electronically, allowing different kinds of operations such as data retrieval and standardized data exchange.

SNOMED CT is a **terminological ontology** that combines the terminological approach of concept system creation described in Section 2.2 with knowledge-based ontology design. It is developed by domain experts for domain experts as knowledge resource for information exchange for use in various applications. The system uses a relationship-based representation of concepts (Concept 1 <relation> Concept 2) by means of the formal ontology language Description Logic (e.g. *Common Cold* <causative agent> *Virus*).

3.2 SNOMED CT definitions

In SNOMED CT concepts are defined in three different ways.

(1) The meaning of a SNOMED CT concept is expressed in a human-readable form by its Fully Specified Name (FSN). This is a Description that represents the meaning of a concept in a way that is unambiguous and independent of the context in which it is used (for a detailed description, see Section 3.4).

(2) Each concept also has a formal concept definition that provides a computer-processable representation of the meaning of the concept. Conceptual relationships are generated in two ways, either by human users proposing a new concept or automatically by the SNOMED CT OWL description logic classifier. The manually created representation is referred to as “stated view” (Figure 3).

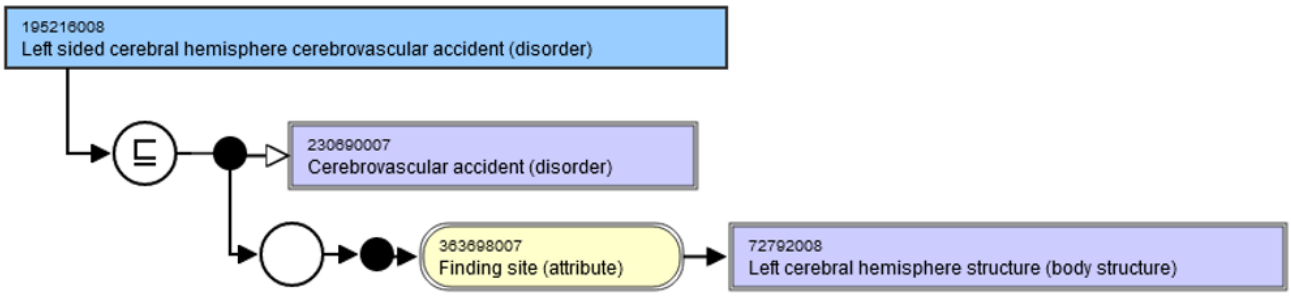

Figure 3 – Diagrammatic representation of the relationships as “stated view” of the concept 195216008 |Left sided cerebral hemisphere cerebrovascular accident (disorder)| (International Edition 2021-07-31)

The automatically generated representation is referred to as “inferred view” (Figure 4).

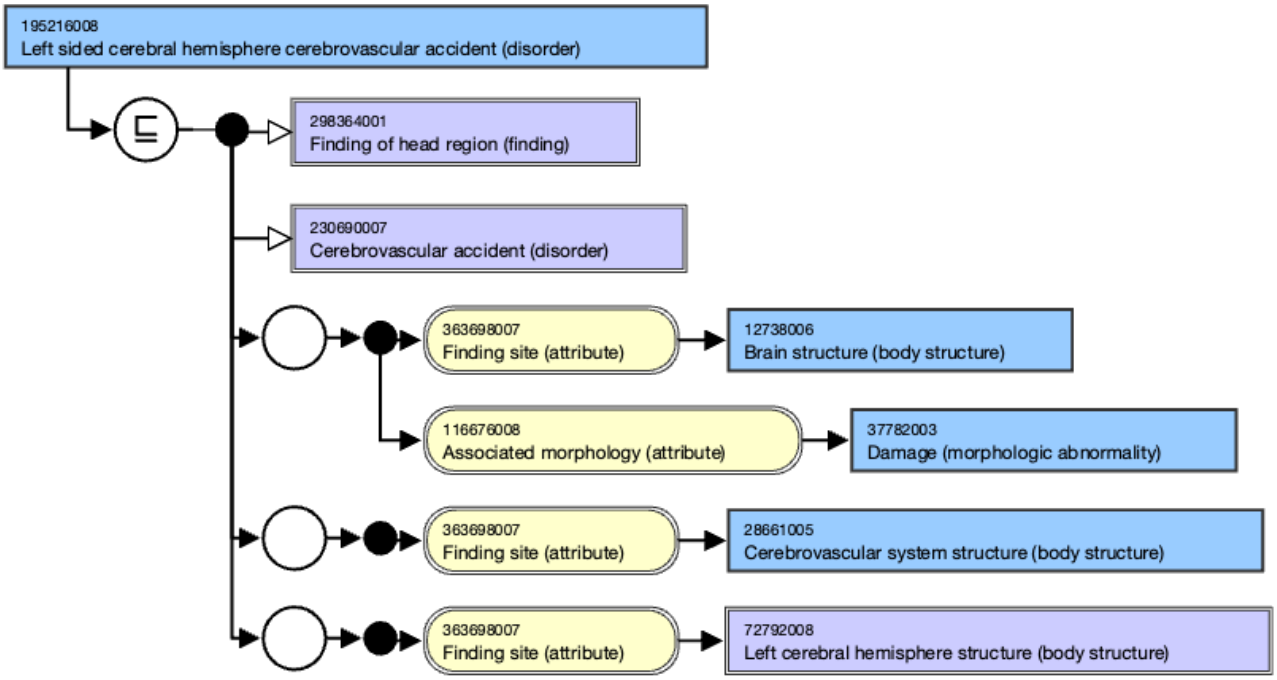

Figure 4 – Diagrammatic representation of the relationships as "inferred view" of the concept 195216008 |Left sided cerebral hemisphere cerebrovascular accident (disorder)| (International Edition 2021-07-31)

For translators it is of outmost importance to look also at the formal definition of the concept as SNOMED CT translations should be concept-based (see Section 4). The formal definition can help to understand the meaning and to increase translation consistency.

The formal rules of description logic can also be used to support meaning-based retrieval from records containing SNOMED CT expressions or concepts (https://confluence.ihtsdotools.org/display/DOCGLOSS/description+logic).

(3) Textual definitions that describe the meaning of a concept in natural language (see Section 3.4 and Figure 5). Only a minority of concepts have a textual definition.

Allergic disposition (finding)

SCTID: 609328004, Defined, Active. Descendants Count: 1326 concepts.

United States of America English language reference set

Term	Acceptability (US)
F ☆ Allergic disposition (finding)	Preferred
S ★ Allergic disposition	Preferred
S ✓ Allergy	Acceptable
D ★ A propensity to developing a pathological immune process generally directed towards a foreign antigen, which results in tissue injury. It is most often applied to type I hypersensitivity but other hypersensitivity types especially type IV (e.g. allergic contact dermatitis) may be involved. Revised nomenclature for allergy for global use:Report of the Nomenclature Review Committee of the World Allergy Organization, October 2003.	Preferred

Figure 5 – Illustration of the three SNOMED CT Description types in the US English language refset of concept 609328004 |Allergic disposition (finding)| (International Edition 2021-07-31) F - Fully Specified Name; S - Synonym; D - Definition

3.3 The hierarchical and multi-axial structure of SNOMED CT

SNOMED CT contains more than 350,000 concepts from **clinical and non-clinical** concept fields or **domains** (e.g. Body Structure, Clinical Finding, Procedure, Substance, Event, Environment or geographical location, Physical object, etc.). These concepts are arranged in 19 top-level **hierarchies** according to the domain to which they belong (Figure 6).

Figure 6 – The 19 SNOMED CT top level hierarchies as displayed in the SNOMED International browser (https://browser.ihtsdotools.org) taxonomy view

At the top of the SNOMED CT hierarchy is the **root concept** (|SNOMED CT concept|), which represents the terminology itself. All other concepts are derived from this root concept through at least one sequence of is-a relationship. This means that the root concept is a **supertype** of all other concepts and all other concepts are subtypes of the root concept. The direct subtypes of the root concept are called “Top Level Concepts”; each of these Top Level Concepts, together with its many subtype descendants, forms a branch of the hierarchy that contains similar types of concepts and also names the branches of the subtype hierarchy. For example, the intermediate concept 788951001 |Hemorrhage of digestive system (disorder)|, which has 17 children, is itself a type of |Disease| which is a type of |Clinical finding|. The further down the hierarchy, the more granular, or in other words more precise, a concept is. This principle is referred to as **granularity** (Figure 7).

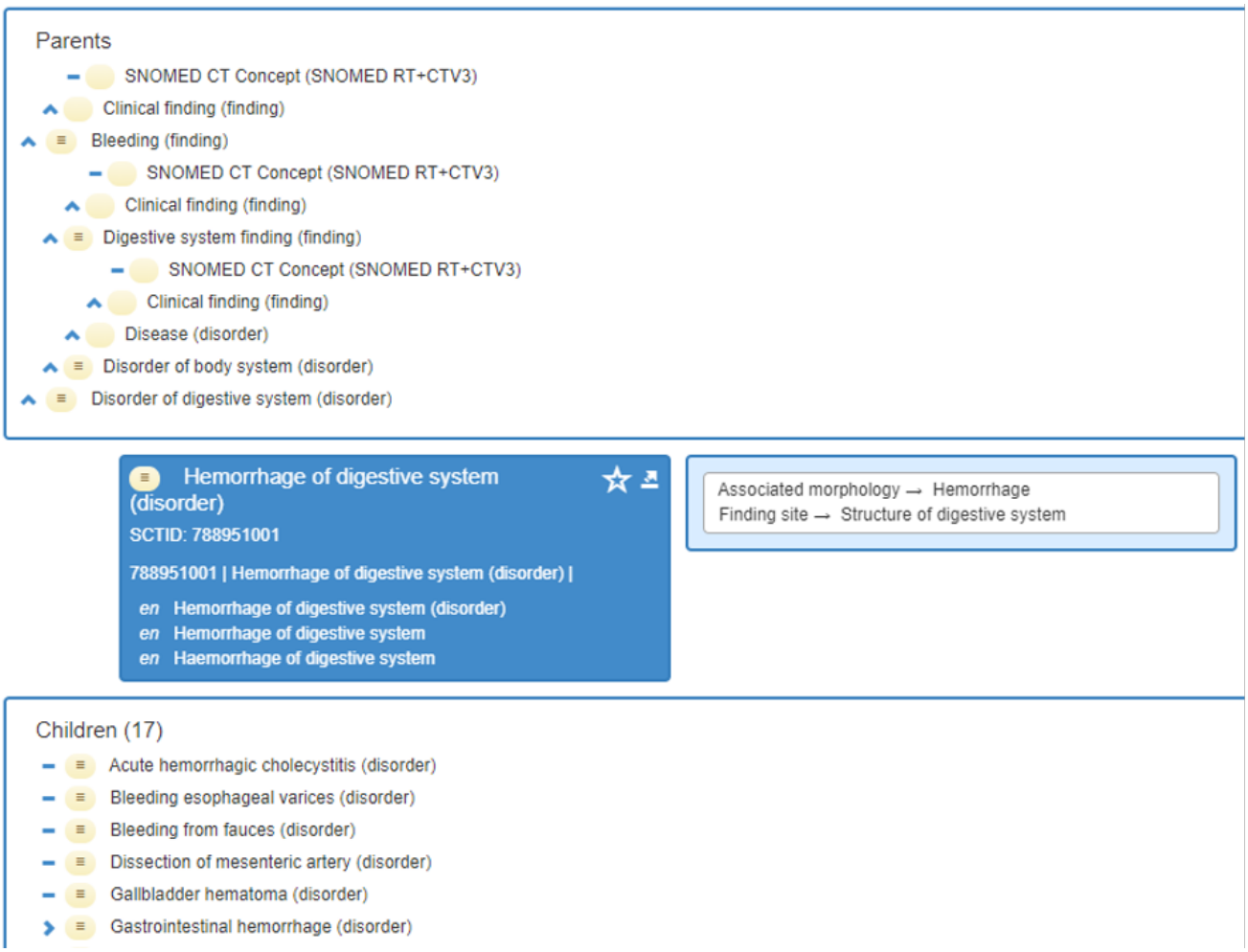

Figure 7– Granularity (SNOMED International browser (https://browser.ihtsdotools.org) summary view, International Edition 2021-07-31)

The top rectangle in Figure 7 shows the hierarchy view up to the root concept for each parent |Bleeding (finding)| and |Disorder of digestive system (disorder)| of the concept 788951001 |Hemorrhage of digestive system (disorder)|. The lower rectangle shows some of the 17 child concepts, while the middle rectangle shows the concept Descriptions. The rectangle to the right shows the concept's relationships.

As many clinical concepts are multidimensional by nature, concepts can have more than one parent. This creates a polyhierarchical structure, which increases the amount of information provided. As the formal representation language of SNOMED CT uses both hierarchical and attributive relationships, the multidimensional character of concepts can be expressed accordingly. Figure 8 shows the polyhierarchical structure of SNOMED CT concepts.

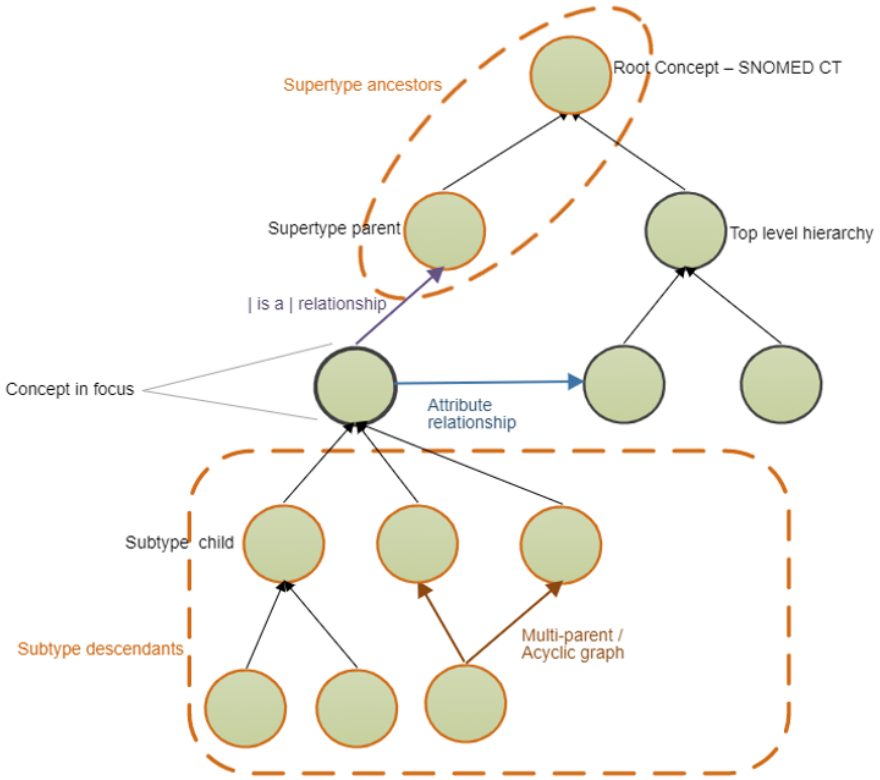

Figure 8 – Polyhierarchical structure of concepts (as shown in SNOMED International Starter Guide)

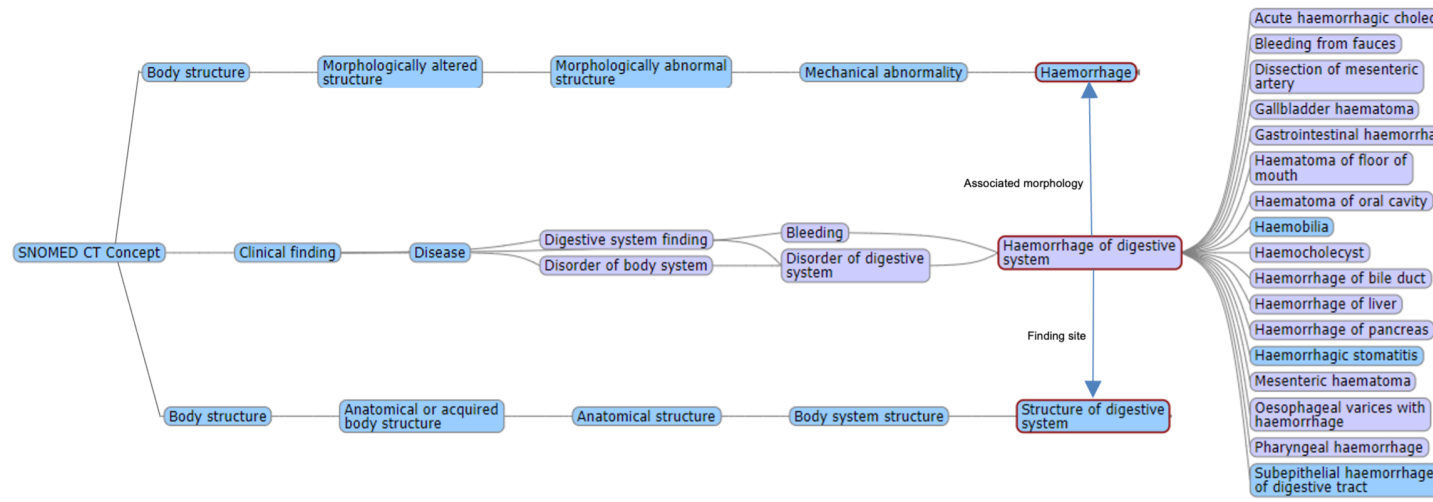

Figure 9 – Full polyhierarchical view of concept 788951001 |Hemorrhage of digestive system (disorder)| (International Edition 2021-07-31) using Ontoserver Shrimp browser (https://ontoserver.csiro.au/shrimp)

3.4 Description types

Each concept is represented at least by two types of Descriptions: one **Fully Specified Name (FSN)** and **Synonyms (SYN),** one of which is marked as the **Preferred Term (PT)** (Figure 10). The presence of a Description of type definition is optional.

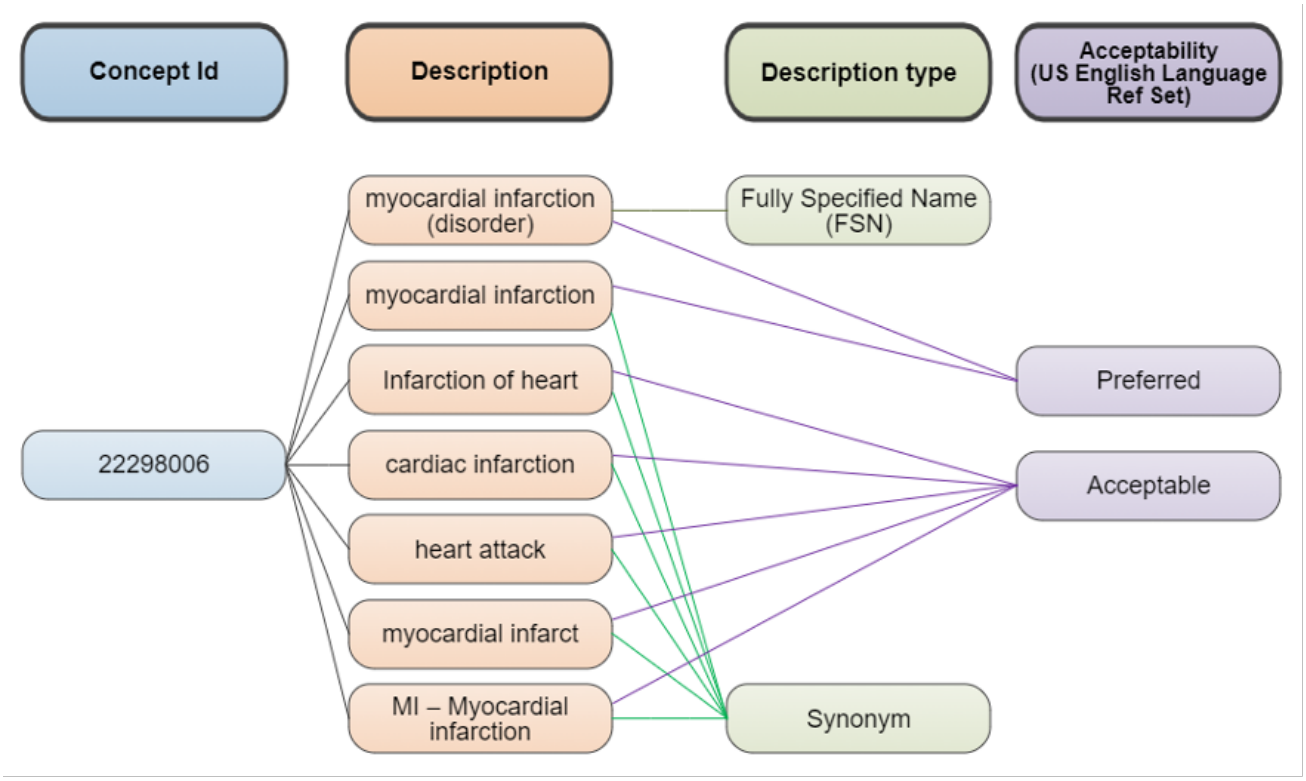

Figure 10 – SNOMED CT descriptions

The FSN is "a description that represents the meaning of a concept in a way that is unambiguous and independent of the context in which it is used" (SNOMED International 2020). Therefore, an FSN can also be considered as a kind of definition or explanation as it describes the content of the concept as explicit as possible (see Section 3.2). If a synonym, the logical definition or the logical definition or the text definition conflicts with the FSN, the FSN is the gold standard or so-called "Source of Truth". The FSN should be able to stand alone as complete, unambiguous & comprehensible.

A FSN is composed of a term and a "semantic tag" between parenthesis at its end (e.g. |Myocardial biopsy (procedure)|). The tag indicates the hierarchy to which the concept belongs (e.g. procedure, disorder, organism, etc.). Each concept has a FSN, which is unique in SNOMED CT in a given language, even in cases of FSNs with identical terms that refer to concepts belonging to different categories. For example, |Hematoma (morphologic abnormality)| is the FSN that represents what the pathologist sees under the microscope, whereas |Hematoma (disorder)| is the FSN that indicates the clinical diagnosis of a hematoma. This FSN is not intended for display in clinical records, but is used to clarify the hierarchical relationships of concepts.

A FSN is only used when the user needs to be able to distinguish between hierarchies (which is not necessary in most use cases for healthcare professionals as they will use a reference set or a single hierarchy as value set for their EHR fields).

Every concept must have at least one active fully specified name in US English. This is the point of reference for the meaning of concepts in the SNOMED CT International Edition. Language reference sets must include a single preferred fully specified name for each concept in a language context. For concepts that are part of an extension, the preferred fully specified name in a language specified by that extension may be the point of reference (https://confluence.ihtsdotools.org/display/DOCGLOSS/FSN).

A *synonym* is "a word or phrase that expresses the meaning of a SNOMED CT concept in a specific language" (SNOMED International 2020). Each concept may have multiple synonyms, one of which is marked as "preferred" to support consistent recording of the underlying concept. The Preferred Term (PT) is "the term deemed to be the most clinically appropriate way of expressing a concept in the specified language context" (SNOMED International 2020). It is the recommended description to display a concept. The PT should be

unique within its own hierarchy and should not contain homonyms. The other synonyms are marked as 'Acceptable'. Synonyms are useful for retrieval purposes.

Each national edition of SNOMED CT must include at least one Fully Specified Name and one Preferred Term to represent each concept. Each Description has a unique numeric Description Identifier in addition to a Description Type.

Further information on concept Descriptions can be found in the SNOMED CT Starter Guide (http://snomed.org/sg).

A textual description or definition (see Section 2.3) is also a Description Type, but it is not mandatory. It is provided for a limited number of concepts, where there is a requirement for additional detail, such as specifying the alignment of a SNOMED CT concept with a specific clinical definition of a condition. For example, the concept 22649008 |Photodermatitis (disorder)| is defined as "An abnormal inflammatory skin condition resulting from exposure to ultraviolet light, most commonly sunlight. May result from phototoxic or photoallergic reactions or both" (Table 1).

Text definitions can, but do not have to be translated. SNOMED International intends to introduce more Description types in the future and make them configurable. For example, a Description type could be configured for search terms that include Descriptions (i.e. terms) that are inappropriate for display and data exchange but useful for clinicians to quickly retrieve them at the data collection interface. Examples are Descriptions containing spelling errors, Descriptions that are not correct but commonly used in clinical practice (e.g. "uterine fibroma" for "uterine leiomyoma") or polysemous acronyms (e.g. "IRA" which stands both for "acute renal insufficiency" and "acute respiratory insufficiency".

Description type	**Description**
Fully Specified Name	Photodermatitis (disorder)
Synonym (Preferred)	Photodermatitis
Synonym (Acceptable)	Photosensitivity dermatitis
Text definition	An abnormal inflammatory skin condition resulting from exposure to ultraviolet light, most commonly sunlight. May result from phototoxic or photoallergic reactions or both.

Table 1 – SNOMED CT description types for the concept 22649008 |photodermatitis (disease)| in the Great Britain English language reference set

For translation purposes the three types of Definitions can be useful (see Section 4). Professional translators without medical expertise will have greater difficulty understanding all SNOMED CT by their formal relationships. Therefore, natural language definitions and textual sources that provide information on the contextual use of a concept are important additional tools.

3.5 SNOMED CT logical definitions, attributes and relationships

In the SNOMED CT ontology, each concept is logically defined through its relationships to other concepts. For the definition of a concept, attribute-value pairs are used in addition to the hierarchical "is a" relationships. An attribute represents a characteristic of the meaning of a concept to which a value is assigned.
The attributes that can be applied depend on the concept model. The concept model defines which attribute value-pairs can be applied to which (sub)hierarchy of concepts and how those attribute-value pairs need to be grouped together to ensure correct unambiguous understanding of the concept definition (see the Editorial Guide for a detailed description of concept models). For example, a procedure may have a method, and a disorder may have an etiology, but a procedure cannot have an etiology, and disorder cannot have a method. Defining characteristics represent the values of a range of relevant attributes. Depending on the nature of the concept, there are more than 100 approved attributes to be used in the concept model. The permitted range of values for an attribute depends on the rules specified in the concept model.

All SNOMED CT concepts are linked at least by one is-a relationship to the immediate superordinate(s) concept(s) and by zero to many attributive relationships. Both types of relationships together form the (formal) definition of a concept. Relationships are not limited to concepts and can also refer to numerical values in products.

There are two types of defined concepts: fully defined and primitive concepts.

A fully (or sufficiently) defined concept has at least one sufficient definition that distinguishes it from any concepts or expressions that are neither equivalent to, nor subtypes of, the defined concept.

For example, the concept 74400008 |Appendicitis (disorder)| is sufficiently defined by the following definition because any concept for which these defining relationships are true, is either the disorder appendicitis or a subtype of appendicitis.

74400008 |Appendicitis (disorder)|

=== 64572001 |Disease (disorder)| : {

116676008 |associated morphology| = 23583003 |inflammation| ,

363698007 |finding site| = 66754008 |appendix structure|

}

A primitive concept has a concept definition that is not sufficient to computably distinguish it from other concepts (https://confluence.ihtsdotools.org/display/DOCGLOSS/primitive+concept).

For example, the concept 5596004 |atypical appendicitis (disorder)| is primitive because the following definition is not sufficient to distinguish atypical appendicitis from its parent concept <<<74400008 |appendicitis (disorder) nor it sibling:

See below the definition |Atypical appendicitis|, its parent |Appendicitis| and its sibling |Catarrhal appendicitis|. These three concepts share the same SNOMED CT logical definition in the stated view. The clinical information not represented in the SNOMED CT logical definition of |Atypical appendicitis| and |Catarrhal appendicitis| is *emphasized*, making them primitive concepts.

74400008 |Appendicitis (disorder)|

<<< 64572001 |Disease (disorder)| : {

116676008 |associated morphology| = 23583003 |inflammation|,

363698007 |finding site| = 66754008 |appendix structure|

5596004 |*Atypical* appendicitis (disorder)|

<<< 64572001 |Disease (disorder)| : {

116676008 |associated morphology| = 23583003 |inflammation|,

363698007 |finding site| = 66754008 |appendix structure|

8744003 |*Catarrhal* appendicitis (disorder)|

<<< 64572001 |Disease (disorder)| : {

116676008 |associated morphology| = 23583003 |inflammation|,

363698007 |finding site| = 66754008 |appendix structure|

3.6 SNOMED International Editorial Guide and Confluence Templates

SNOMED International has numerous resources to promote the understanding and use of SNOMED CT. These documents are available in the Document Library, which is a collection of SNOMED CT documents, including introductory material, practical guides, technical specifications and reference material (https://confluence.ihtsdotools.org/display/DOC/SNOMED+CT+Document+Library).

The Editorial Guide provides detailed information on the rules under which the international content of SNOMED CT is authored. It describes the scope, hierarchies, authoring principles, style guidelines and concept model rules of SNOMED CT for each domain (http://snomed.org/eg).

For example, when translating the Body structure hierarchy, translators are presented with a large number of concepts represented by terms containing the words “structure”, “part of”, “entire” and “all”, which should be interpreted with caution (for more details see Appendix B of the Editorial Guide; https://confluence.ihtsdotools.org/display/DOCEG/Anatomical+concept+model).

4 Translating SNOMED CT

An example is the translation of |Pericardium| to Norwegian (Figure 11). Medical dictionaries and encyclopedias list “perikard” and “hjertesekk” (literally “heart sac”) as synonyms and valid translations. For most practical uses, this would be true. However, the translator should keep in mind the fine granularity of the SNOMED CT concept hierarchy, in which the |Pericardium| is a superordinate concept that includes several subordinate concepts, including the |Pericardial sac|. Considering the concepts in the hierarchy, the proper translation of |Pericardium| is “perikard”, while “hjertesekk” is equivalent to |Pericardial sac|.

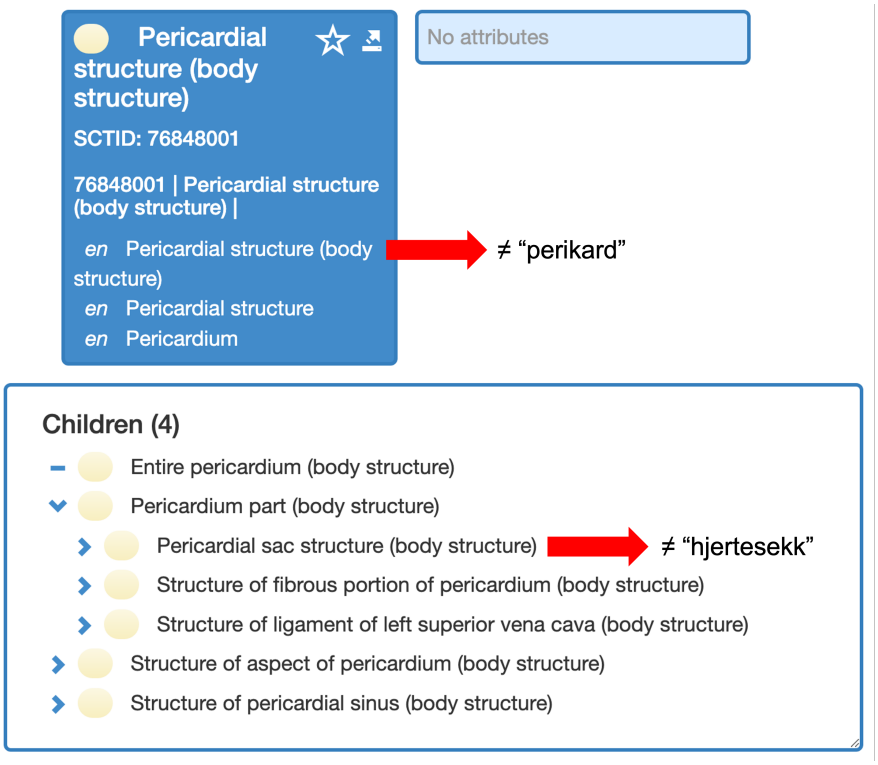

Figure 11 – Example of concept-based vs literal translation approach

4.1 Basic approach

In practice, this means that a number of basic principles must be followed with regard to the morpho-syntactic structure of terms to ensure consistency, but in cases where there are serious conflicts with everyday clinical language, clinical use should have priority.

The following sections address the main issues that translators may face. The goal of any translation is to achieve consistent, usable, and clinically acceptable translation equivalents in the target language.

4.1.1 The principle of concept based translation

The translation of complex terms requires a relatively high level of domain knowledge to ensure that each translation (term or phrase) is an accurate representation of the original concept. Therefore, the translation team must carefully check that the meaning of each source concept is clearly understood, including the hierarchical position of the concept and its relationship to other concepts. In Dutch, for example, the literal translation of the English term “arthrosis” without prior analysis of its meaning (“disorder of one or more joints”) generates the incorrect term “arthrosis”(which is the Dutch term for |Osteoarthritis|) instead of the correct concept-based translation "arthropathie".

It is recommended to follow the translation workflow steps illustrated in Figure 12:

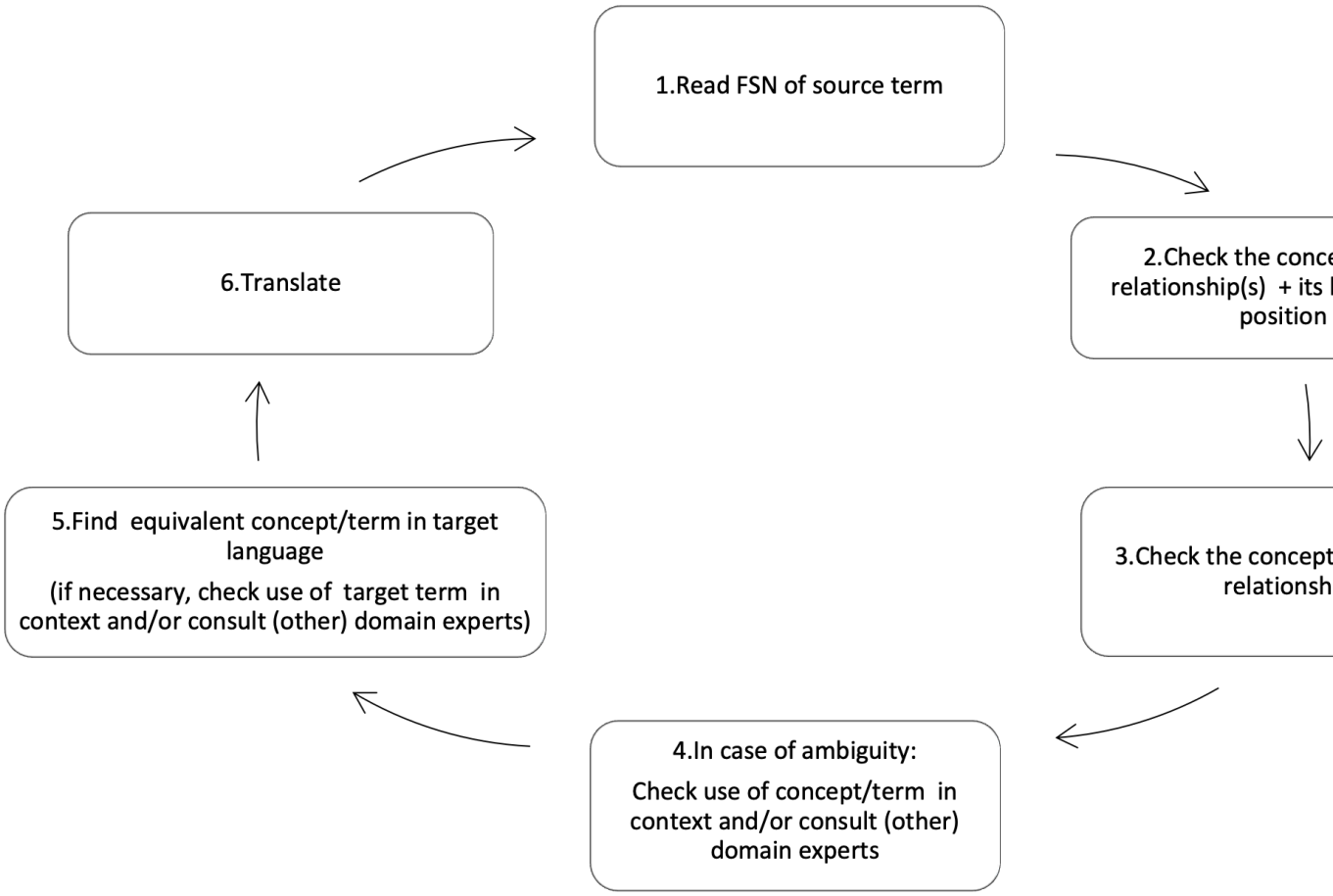

Figure 12 – Recommended translation workflow steps

4.2 General linguistic principles

The following paragraphs deal with general aspects of translation that are relevant for all languages.

For each target language, national linguistic guidelines, including syntactic, morphological and orthographic rules, must be provided to describe the specific rules of the target language.

4.2.1 Language for Specific Purposes

Since the target group of the national edition of SNOMED CT are professionals, high quality language for specific purposes terms should prevail. It is recommended to establish some preliminary general principles for the choice of lexical variants before starting the translation work.

In Danish and German, for example, it is common to use pure Latin or Greek in the field of anatomy, while diagnostic terms or terms describing procedures are often **hybrids** or mixed forms (so called **neoclassical compounds** such as *gastroenteritis* and *mammography*), or, in some cases, also general language terms.

Colloquial expressions (medical slang words such as *blue pipes* for veins or *cabbage* for a *coronary artery bypass graft* or CABG) should be avoided; if needed, they can be added as synonyms.

Recommendations or guidelines issued by a national language council, authority or body should generally be followed. However, it may be necessary to formulate specific morphological and spelling rules that apply to the terminology of the target language. Specific practices related to the construction of medical terms published in national medical journals should also be taken into account.

4.2.2 Ambiguities in the source language

As mentioned in the previous sections, all translations should start with a thorough analysis of the conceptual meaning of the FSN and the concept's formal definition.

It is possible that the conceptual meaning is not correctly reflected by the formal definition in the International Edition. In this case, it is recommended that the NRC reports the error to SNOMED International via the Content Request Service (CRS) (http://snomed.org/crs). Licensed organisations from non-member countries can report directly through CRS. Issues can also be raised for discussion in the JIRA SNOMED CT Translation Feedback Group prior to submitting a CRS request to see how the concept's meaning has been interpreted in other countries. The correct translation can then be produced after a response is received.

4.2.3 Selection of the right term

The term (i.e. a word, a group of words or compound, or an expression) is the essential component of communication in a specialised subject field. The primary meaning in a language for specific purposes is therefore conveyed in its specialised vocabulary, which should consist of terms that meet the following requirements:

- **Unambiguity**: the term should be understood in the same way by any professional that uses it.
- **Linguistic accuracy**: a term should conform to the language-specific morphological, syntactic, orthographical and phonotactic conventions.
- **Transparency/motivation**: a term should be systematic and self-explanatory, i.e. it should reflect the characteristics of the underlying concept.
- **International recognizability**: terms based on Latin and Greek word elements may be preferred; e.g. universal scientific terms such as “hepatocystis”, “nephritis”, “appendectomy”, “hepatitis”, etc.).
- **Psychological acceptability**: term usability should be taken into account whenever possible (in accordance with the rules of standard language, and not slang).
- **Systematicity and consistency**: semantically similar concepts should be expressed by terms with similar morphological and syntactical structures. It may be interesting to consider the use of translation templates to enhance consistency of translation across similar concepts (see Section 4.3).
- **Conciseness**: a term should be concise and in keeping with the principle of linguistic One should be aware that descriptions of the type FSN and Synonym are technically limited to 255 characters (definitions are limited to 4096 characters).

In terminological practice, it is not possible to meet all requirements simultaneously in any language. Psychological acceptance, in particular, often hampers compliance with some of the principles. For example, commonly used eponyms such as “Apgar Score” or “Down's Syndrome” contradict the search for transparency and also challenge the principles of international recognition. Also the simultaneous use of different term types for the same concept (e.g. terms consisting of noun + noun vs. noun + prepositional phrase vs. noun + adjective combinations) challenge systematicity and consistency (e.g. “incision of brain” vs “intracerebral incision”).

The creation of language-specific rules for the translation of SNOMED CT should therefore be carried out with the greatest possible thoroughness and care and should take into account not only linguistic and terminological aspects but also psychological ones. The needs and expectations of the users remain the primary concern and clinical acceptance of the terminology is of crucial importance.

The national translation guidelines should be based on the general editorial framework of SNOMED CT and, where necessary, be adapted to the needs of the respective national language.

4.2.4 Translation of the Fully Specified Name

FSNs may or may not be translated: there are both arguments for and against the FSN translation.

Arguments in favor are:

- Translating the FSN helps identifying duplicate concepts
- Translating the FSN provides local terminologists with a description that is unique, unambiguous and independent from context
- The stringent wording of the translated FSN allows more options for the PT (e.g. use of acronyms)

Arguments against are:

- The FSN is the source of truth
- The FSN should be immutable, which does not allow for inaccurate translation
- The editor may not allow multiple FSNs for countries with multiple languages
- The translation of the FSN requires more resources

The decision whether or not to translate the FSN depends on the individual NRCs.

4.2.5 Concept equivalence

Concept equivalence means that a word or phrase means exactly the same thing in the source and target languages. Determining equivalence requires a deep understanding of the source language and culture. A literal (word-for-word) translation is in many cases wrong, because you first need to understand the meaning behind the words.

4.2.5.1 Cultural variations

Cultural differences must always be taken into account when translating. The question of conceptual equivalence or the lack of such equivalence must therefore be given special attention.

Due to its historical development, part of the source language terminology of SNOMED CT is based on British or American administrative structures or clinical practices that are not necessarily internationally valid. This is currently the case with regard to subhierarchies such as Social context, Regimen/therapy, Environment or Geographical location, Substance and Pharmaceutical/Biological product hierarchies. In the case of drug and vaccine concepts, for example, the pharmaceutical form of drugs may vary from country to country. At the same time, local regulations may dictate which drugs should be included in the national edition. The pharmaceutical/ biological products subhierarchy is therefore a typical example of a hierarchy that requires a complete overhaul involving local pharmacologists to adapt to local needs, and preferably to define a national extension.

Although it is theoretically possible to find national equivalents (e.g. paraphrases) that make such culture-dependent (i.e. not locally anchored) concepts understandable, the local applicability of such culture-dependent concepts is questionable. One example is the US concept of |veteran hospital|, a special kind of medical facility where former US soldiers receive medical care. It may therefore be necessary to develop new subhierarchies within a national edition in order to do justice to culture-specific concepts. In summary, it is advisable to identify the subhierarchies with concepts that are not nationally valid before the national translation process begins. The decision on how to deal with these concepts should then be made locally. Not translating these concepts is a viable option. If it is felt that these concepts should be local to the UK or US, SNOMED International can be asked to move them to these extensions.

4.2.5.2 False friends

False friends are word pairs from two languages that are similar in writing (orthographic) or pronunciation (phonetic) but that differ significantly in meaning. The words can either be related in origin but developed differently or have a purely coincidental similarity. False friends can lead to an unrecognized false translation when translated literally. Examples are the English term “ventilator” (meaning a device used to maintain an airflow into and out of the lungs of a patient who is unable to breathe) and the German term “Ventilator” (meaning fan; the correct German equivalent in the clinical context for “ventilator” is “Beatmungsgerät”).

Similar problems arise when a term in one language covers a broader concept than in the other language. An example are **holonyms**, i.e. terms that denote a whole whose part is designated by another term (e.g. “hand" in relation to "finger”). Conversely, **meronyms** denote a component or a part of something (e.g. “finger” in relation to “hand”). Translation problems arise when there is no 1:1 correspondence between source and target language. For example, the English terms “drug” and “substance” are holonyms without equivalents in some other languages. Depending on the specific term, the translator has to look for alternative equivalents such as meronyms (e.g. Danish “præparat eller medicament” as the equivalent of “drug” and “misbruger stoffer” as the equivalent of “misuses drugs”) or to use the conjunction en/of (e.g. in Dutch “medicatie en/of drugs”) to emphasize that both are included in the concept’s meaning.

Concept-based translation makes it possible to check the equivalence between terms and concepts. It is particularly important that all team members, especially translators and reviewers, are made aware of the difference between literal and concept-based translation.

4.3 Translation techniques

There are various techniques that can be used differently for each language to translate SNOMED CT terms. Table 2 gives an overview of the most common techniques.

Translation technique	Definition	Source Description (English)	Target Description (Danish)
Equivalence or reformulation	using a different language-specific term to refer to the same concept	*heart burn*	*halsbrand*
		cardiac arrest	*hjertestop*
Literal translation	word-for-word translation	*external blind*	*ekstern blind*
		cystocele affecting pregnancy	*cystocele med indvirkning på gravidite* (“cystocele with effect on pregnancy”*)*
Borrowing	using the source term in the target language	*cardiac output*	*cardiac output*
		shaken baby syndrome	*shaken baby-syndrom*
Calque	creating or using a neologism in the target language by adopting the structure of the source language	*closed fracture of metacarpal bone*	*lukket fraktur af metakarpal-knogle* (“closed fracture of metacarpal bone”)
Amplification (description)	using a paraphrase to explain a term that has no equivalent in the target language	*battered wife*	*hustru der har været udsat for vold i hjemmet* (’“wife who has been exposed to violence in the home”)
		high birth weight	*barn heavy-for-date uanset gestationsalder* (“child heavy-for-dates regardless of gestational age”)

Table 2 – Examples of translation techniques with description of the morpho-syntactic structure of the source and target language terms

4.4 Naming conventions

When creating descriptions for concepts, certain linguistic rules must be observed. Additional rules apply to the individual hierarchies. Both the general and the hierarchy-dependent rules for the (English) International Edition are described in detail in the Editorial Guide (https://confluence.ihtsdotools.org/display/DOCEG/General+Naming+Conventions).

In a translation context, the following issues are important.

4.4.1 General conventions

4.4.1.1 Morphological patterns of descriptions

In general, the morphological patterns of Descriptions in all SNOMED CT hierarchies should be as consistent as possible. As it is crucial to take into account the morphological and syntactic features of the target language, the national guidelines provide specific instructions on the recommended Description patterns in a given language (e.g. consistent naming of diseases such as “Type 2 diabetes mellitus” vs “Diabetes type 2” in Germanic languages such as Dutch and Norwegian).

4.4.1.2 Verbs and tenses

The Editorial Guide contains a set of rules for the use of verbs and tenses (inclusive gerunds) in English to be used in the different hierarchies (e.g. the use of verbs in the past tense such as “Hand tendon ganglion excised” indicates that the procedure was performed). When translating it is important that these rules are implemented as accurately and consistently as possible in the target language. The language-specific implementation of these rules can be described in the national guidelines (e.g, how to translate the English gerund).

4.4.1.3 Word order

The word order within noun phrases should match the syntax of general language. In certain cases, however, it is necessary to deviate from the normal word order. For example, in English adjectives are used attributively (i.e. before the noun; e.g. "regular pulse"). In SNOMED CT a number of terms represent the results of tests or examinations; in these cases the adjective comes after the noun ("pulse regular" instead of "regular pulse"), which allows the results to be listed in a drop-down format in a software application. This format, in which all terms begin with the same word (in this case the Observable Unit + the Value, a combination that converts the concept into a finding), makes it easier to keep track of all possible outcomes.

The rules regarding word order and rule-based exceptions can be specified in the national guidelines.

4.4.1.4 Choice of lexical variants

Depending on language and domain (e.g. anatomy, physiology, clinical medicine) medical terms can have different forms:

- Latin/Greek terms: *diabetes mellitus*, *pes valgus*
- Hybrid terms (i.e. Latin/Greek terms morphologically and syntactically adapted to the national language): *common hepatic artery* (Latin/Greek: *arteria hepatica communis*), *arteriosclerotic retinopathy* (Latin/Greek: *retinopathia arteriosclerotica*)
- National language terms: *stomach ache, placing a patient on a bedpan, bad taste in mouth.*

Which forms should be used for the translation of SNOMED CT depends on the conventions of the respective target language and/or the domain and are to be defined in the national guidelines.

4.4.2 Naming conventions for specific hierarchies

The following sections discuss a number of specific issues relevant to translation into the different target languages. The relevance of each topic may vary depending on the target language.

4.4.2.1 Organism names (bacteria, viruses, plants, animals, etc.)

The organism hierarchy uses international taxonomic names to a large extent, which can be found in the Editorial Guide. Where this is not in contradiction with national language policies, the names of organisms should be retained as universal (international) scientific terms and should be consistent with existing spelling rules, in particular as regards capitalization of terms (e.g. “Chlamydia pneumoniae”, “Spirochaete dentium”, “Dependovirus”, etc.; it is important to note that the current version of SNOMED CT may contain incorrect lower case letters.)

The Organism hierarchy has no relation types with which to specify properties. As a result, it contains instead a large number of grouper concepts such as “gram-negative bacterium”. In English, groupers are represented by the common names "infectious agents" or "arthropod organisms". If a common name is used in English, it is recommended to use a common name in the target language as well. Most organism concepts correspond to a taxonomic entry in the Linnaean taxonomy, but some, such as gram-negative bacteria, have only a common name. In that case, you should translate to a common name in the target language.

SNOMED CT also tends to lag behind the very fast developments in the micro-organism thesauri. It contains many obsolete names as separate concepts, which have become synonymous with another concept. Such issues are quickly resolved when pointed out. The Editorial Guide lists specific resources for bacteria, viruses, parasites, and so forth (https://confluence.ihtsdotools.org/display/DOCEG/Organism+Naming+Conventions).

The Organism hierarchy has no relation types with which to specify properties. As a result, it contains instead a large number of grouper concepts such as “gram-negative bacterium”. In English, groupers are represented by the common names "infectious agents" or "arthropod organisms". If a common name is used in English, it is recommended to use a common name in the target language as well. Most organism concepts correspond to a taxonomic entry in the Linnaean taxonomy, but some, such as gram-negative bacteria, have only a common name. In that case, you should translate to a common name in the target language.

Explicit rules for the use of international taxonomic names in the Organism hierarchy and other hierarchies, as well as the use of common names as Preferred Terms in specific categories or contexts should be specified in the national guidelines.

4.4.2.2 (Biochemical) names, ingredients of medicinal products, enzyme and hormone names

A term that refers to a chemical substance in a medicinal product can be interpreted in two ways: Either it is the name of a specific component of the product (e.g. morphine, glycogen) or it is a general term for the chemical substance itself. Ingredients are listed in the SNOMED CT Pharmaceutical/biological product hierarchy, and chemical substances are listed under their generic name in the Substance hierarchy.

In the target language it may be common that different orthographic principles are used for products and substances. The specific spelling rules for chemicals and biochemical agents, ingredients, enzymes, and hormones should be defined in the national guidelines.

Whether substances and products that are not available in one's own country are translated, and if so, what principles apply in this case, should also be addressed in the national guidelines.

4.4.2.3 Foreign (loan) words and abbreviations, acronyms and initials

National translation guidelines should include instructions on the extent to which **foreign words** are accepted in the target language. Examples are the English terms “cardiac output”, “Rift Valley fever”, “spindle cells”, “bias”, etc. or the French terms “tabatière”, “plaques”, “grand mal” that are used unchanged in different other languages. Accepted foreign-language terms should in any case comply with the orthographic rules of the source language. This means, for example, that the accents in French words such as “péan”, “tabatière” and “debridement” must be taken into account when they are used in another language.

Some (sub)hierarchies (e.g. Staging and scales) can contain a large number of concepts for which there is no official translation into other languages. Examples are the tumor staging system **“**tumor-node-metastasis (TNM)” or the “Bristol Language Assessment Scales (BLADES)” score. In these cases the term of the source language is preferred.

Acronyms and initials – both foreign and national – also play an important role in translation: "AIDS" (acquired immune deficiency syndrome) is an example of an abbreviation that is accepted as such in many languages, while in French and Spanish it is "SIDA", the abbreviation for the national expressions. It is recommended that only well-known and widely used foreign or national abbreviations should be included in national editions. As a rule, abbreviations should never be used in the FSN and only rarely be part of the Preferred Term (PT) as they tend to be highly ambiguous.

4.4.2.4 Eponyms

Eponyms are an essential part of medical terminology and clinical language. While most eponyms are used internationally, there are some language-specific eponyms that can be difficult to translate. Examples include the disease names "Wegener's syndrome" and "Reiter's disease", which are controversial because they were associated with the Nazi regime. Therefore, an eponym can disappear from language use quite quickly, while a descriptive term retains its meaning, regardless of whether it is used frequently or not.

Another disadvantage of eponyms is that they do not meet the essential requirements of well-formed terms such as clarity, linguistic correctness and transparency. However, due to their frequent occurrence in clinical practice, eponyms are also used in SNOMED CT. The national guidelines should include rules to help the translator translate eponyms correctly (e.g. systematically check whether the eponym exists in the target language or whether another (language-specific) eponym or term is used instead).

4.4.2.5 Case significance

Case Sensitivity Indicator	Meaning	Examples
cl	First letter of the description may or may not be capitalized while the case of the rest of the description cannot be changed.	• Family history of Prader-Willi syndrome (situation) • Born in Australia (finding) • Neonatal jaundice with Dubin-Johnson syndrome (disorder)
CS	Cannot change any case in the description Changing case may change the meaning of the term or is not commonly used	• Down syndrome • English as a second language (finding) • pH measurement (procedure) • mm (qualifier value)
ci	Entire description may be lower or upper case Changing case does not change the meaning of the term	• Fracture of tibia (disorder) • Blood compatibility test (procedure) • Bite of fish (event) • Floor mat (physical event)

Table 3 - Case sensitivity rules (https://confluence.ihtsdotools.org/display/DOCEG/Case+Significance)

If the case of the initial letter in a source or target language description is significant, i.e. the term must begin with an uppercase letter or a lowercase letter, this is taken into account in the SNOMED CT source and target language description files in the Case Sensitivity indicator metadata ((https://confluence.ihtsdotools.org/display/DOCEG/Case+Significance). When translating, it is important that the value of this description field is correctly chosen in the target language description considering the term used in that language description and regardless of the case sensitivity indicator value used in the English description.

4.4.2.6 Punctuation, characters and numerals

4.4.3 Translation templates

To improve the consistency of translation of concepts of the same type, the use of translation templates is of particular interest. Translation templates, similar to the SNOMED International Editorial templates (https://confluence.ihtsdotools.org/display/SCTEMPLATES), are recommendations for naming, word order, use of the noun

or adjective form, etc., which together provide a framework for constructing the FSN and synonyms of certain subhierarchies (i.e. subhierarchies with common features) of the terminology.

An example of a translation template is the one described below, which was used in the Belgian-French translation process. The target concepts of this translation template were the concepts representing a bronchus of a specific (lobe of) lung. The FSN structure of these concepts is represented in the form of the English source template as follows: "Structure of bronchus of [laterality] [localisation] lobe".

When this English template is decomposed, it becomes clear that it is derived from the translation of two (subfamilies of) nested concepts: the concept 955009 |Bronchial structure (body structure)|, which represents a bronchus, and the concepts representing the lobes of the lung (for which a first template had already been created).

The notion of 'nested concept' is used to describe a word or phrase that is part of an FSN but is itself also a SNOMED CT concept. It is important to recognise nested concepts when translating SNOMED CT content, even if no templates are used, as the translation of the more pre-coordinated concepts should be coherent with the translation of these 'parent' nested concepts. This is especially true for disease and procedure concepts that refer to anatomical parts in their FSNs.

In the above example, the potential [laterality] values = (empty), right, left refer to the lung laterality and the potential [localisation] values = (empty), upper, middle, lower refer to the position of the lung lobe.

For this family of concepts the following Belgian-French translation template is adopted:

PT: "bronche lobaire [localization, new French anatomical nomenclature] [laterality, Fr] "

Syn1: "structure de la bronche lobaire [localization, new French anatomical nomenclature] [laterality, Fr]"

Syn 2: "bronche lobaire [localization, old French anatomical nomenclature] [laterality, Fr] "

Syn 3**:** "bronche du lobe [localization, new French anatomical nomenclature] du poumon [laterality, Fr]"

The template is based on the Belgian-French translation of the concept 39582006 |Lobar bronchus structure (body structure)|, which, in turn, depends on the translation of the concepts 955009 |Bronchial structure (body structure)| and 31094006 |Structure of lobe of lung (body structure)|.

Adopting this template to the concept 11339004 |Structure of bronchus of right upper lobe (body structure)|, the following translations are generated:

PT: "bronche lobaire supérieure droite"

Syn 1: "structure de la bronche lobaire supérieure droite"

Syn 2: "bronche lobaire crâniale droite"

Syn 3: "bronche du lobe supérieur du poumon droît"

The development of translation templates takes some time, but has the following advantages:

- Translation templates force authors to consider families of concepts rather than isolated concepts in translation, which often leads to a better understanding of the medical reality represented or not represented by those concepts and reduces the risk of misinterpretation of concept meaning. This can be particularly useful for primitive concepts or hierarchies without a concept model.
- They ensure that sibling concepts of the same "family" are translated in a consistent way, even if their translation is distributed among different translators ("horizontal homogeneity" of translation). This is very useful for large families with several hundred concepts such as "neoplasm of [organ]".
- They ensure that the translation of more pre-coordinated concepts matches the translation of the "nested" concepts contained in their FSN ("vertical homogeneity" of translation). In reality, one has to work in such a way as to build a series of overlapping inverted pyramids, from simple concepts to the more pre-coordinated terms, and create a hierarchy of dependent templates.
- The meaning of words that are not SNOMED CT concepts can, if possible, be set in context. This is interesting not only for translation, but also for improvement of the Core terminology, as it may lead to the inclusion of new SNOMED CT attributes or attribute values to further define primitive concepts.

- The templates can be implemented in a dedicated computer programme to provide translators with automatic translation suggestions for concepts, thus reducing the translation time for batches with concept families.
- Templates can also support quality assurance, as they can be used to retroactively align old translations with new templates.
- When a link between the translation template and the concepts to which it has been applied is stored in a database, translation templates ensure that all dependent concepts to be reviewed can be traced back if the national translation of a nested concept needs to be corrected or if the FSN of a nested concept changes in SNOMED International and the change has an impact on the understanding of the concept in the target language.

Research is being conducted to investigate the possibility of storing the relationships between translation templates and their output descriptions in the local language in Refset format.

5 Sources of information

In addition, NRCs should create and maintain a document that specifies which sources are authoritative and should accompany the translation process so that all translators involved have a common basis for determining which sources are acceptable. Translators and other professionals involved in the translation process should also have access to a range of reliable and recognised sources of information in their own language. Both external and internal sources can be used. For example, the Dutch NRC uses the following external sources:

- Pinkhof Geneeskundig woordenboek (Pinkhof medical dictionary)
- Terminologia anatomica (https://www.anatomicalterms.info/)
- the same thesauri for organisms as given by SNOMED International in the Editorial Guide
- Orphanet (the quality of translation can be questionable, but it is often the only source for a rare disease)
- Farmacotherapeutisch kompas (for substances)
- the archive of the *Nederlands Tijdschrift voor Geneeskunde* (NTvG) (Dutch Journal of Medicine)

The following internal sources are used:

- a self-created Java API that suggests possible translations based on the concept definition (only for concepts that meet certain criteria, e.g. are fully defined)
- a spell checker that incorporates the Pinkhof Geneeskundig woordenboek (Pinkhof medical dictionary)
- indexed comments logged for each previous translation
- guidelines agreed during the translation project, both general and specific phrases
- recorded discussions on proposed guidelines (so that the rationale for a guideline can be looked up or it can be checked whether a guideline was considered but rejected)
- previous translations, for consistency

For reasons of lack of quality, the Dutch translation of ICD-10 or ICPC is deliberately not used.

In addition to sources in the local language, English sources can also be useful in determining the exact nature and meaning of a term.

Wherever possible, selected internal working documents, textbooks, reference works, etc. should be directly accessible in electronic form for anyone to find information on a particular concept by means of text examples, definitions or expressions containing a particular word or phrase.

Machine translation systems such as Google Translate and DeepL or online translation tools such as eTranslation (https://ec.europa.eu/cefdigital/wiki/display/CEFDIGITAL/eTranslation) can be a potential help, but should be used with caution. Currently, the accuracy of translation is not yet sufficient to translate clinical content with the exactness required from a clinical risk perspective. However, MT can be helpful in providing an initial translation that can then be reviewed; or it can be limited to template concepts.

In the following sections relevant existing electronic information sources and internet references (Section 5.1) are listed, followed by an overview of sources that can be created by means of corpus tools or automated suggestions (Section 5.2). It is advisable to check the validity and quality of each source, especially the national editions of international sources.

5.1 Electronic information sources and internet references

Section 4.4.2.1 Organism names (bacteria, viruses, plants, animals, etc.) already contains a list of specific resources for bacteria, viruses, parasites, etc. (https://confluence.ihtsdotools.org/display/DOCEG/Organism+naming+conventions). The following additional sources may also be helpful:

Ready access to already approved, translated terms

- It should be possible for translators and other persons involved in the translation process to refer to previously approved terms containing similar constructions and/or word combinations.

Online dictionaries and databases

- *Dictionnaire médical de l'Académie de Médecine* – version 2022 (http://dictionnaire.academie-medecine.fr/index.php)
- *Oxford Concise Medical Dictionary* (10de edition). (2020). Oxford University Press. (https://doi.org/10.1093/acref/9780198836612.001.0001)
- IATE (Interactive Terminology for Europe) (https://iate.europa.eu/home)
- WordNet (http://globalwordnet.org/resources/wordnets-in-the-world/)
- info (ATI) (https://www.anatomicalterms.info/)

National corpora

Internal working documents such as:

- national guidelines for translation
- overview of the principle decisions of the Editorial Board or a similar competent body
- lists of examples of translated terms and/or corrected terms representing specific semantic or morphosyntactic features

Textbooks on the different areas of clinical practice

Reference files or books such as:

- national versions of medical dictionaries or lexicons
- national version of the chemistry nomenclature
- national version of *Nomina Anatomica, Terminologia Anatomica* or other nomenclatures
- national version of the *International Classification of Diseases* (ICD-11)
- national versions of other classifications

Explorative resources

These sources should be used with care and serve as inspiration, but not as an authoritative source as they contain, for example, synonyms that are not entirely interchangeable or outdated terminology and the like. Examples are:

- Wikipedia
- Pubmed (https://pubmed.ncbi.nlm.nih.gov/)
- MeSH (https://www.ncbi.nlm.nih.gov/mesh/)
- Orphanet (https://www.orpha.net/consor/cgi-bin/index.php?lng=EN)
- OMIM® (Online Mendelian Inheritance in Man®) (https://www.omim.org/)
- INN (International Nonproprietary Names) (https://www.who.int/teams/health-product-and-policy-standards/inn)
- com: to search in multiple English dictionaries (https://www.onelook.com/reverse-dictionary.shtml)

Medical publications

- electronic versions of articles from renowned national medical journals (these can be valuable references for the use of specific terms and abbreviations of foreign and local origin. It should be noted, however, that the expressions chosen in these journals may not always be linguistically correct or consistent if they are foreign language expressions. One should also consider the publication date, as terminology may have changed following new discoveries or techniques).
- clinical guidelines and quality assessment documents
- medical protocols
- national standardised guidelines for procedures

Recommended internet references

- A list of valid internet references with useful information should also be established. The recommended resources can be found in the Editorial Guide (see list of webpages at the beginning of chapter 5). Some relevant taxonomies or nomenclatures in English may be available in the target language.

5.2 Self-created corpora

Online tools

- The Sketch Engine (http://www.sketchengine.co.uk/)
- CLaRK (http://www.bultreebank.org/clark/)
- CorpusExplorer (http://www.CorpusExplorer.de)
- AntConc (http://www.antlab.sci.waseda.ac.jp/software.html)
- CoCab (http://chasen.aist-nara.ac.jp/~kaoru-ya/cocab/)
- …

Automated suggestions

- a self-created Java API that suggests possible translations based on the concept definition (only for concepts that meet certain criteria, e.g. are fully defined)
- spell checkers that incorporate national medical dictionaries
- indexed comments logged for each previous translation in an online translation environment
- guidelines agreed during the translation project, both general and specific phrases
- recorded discussions on guidelines proposals
- previous translations, for consistency
- lexicon-based translation tools such as CoMeTT, used by the Belgian CSCT (for a detailed description see Wermuth, M.-C., Walravens, M., & Lambot M.-A. (2022). Collaboration and communities of practice in the field of medical ontology translation. (P. Cadwell, F. Federici, & S. O'Brien, Eds.) *The Journal of Specialised Translation* (37), 75-98).

6 Translation process and post-translation issues

Recommendations on the various steps of the translation process can be found in the document *Guidelines for Management of Translation of SNOMED CT* (see reference in Section 1.1). This section summarises the main steps.

6.1 Translation

There should always be at least two persons involved in the initial translation – a translator and a proof-reader (to verify the initial translation). It is crucial that translators have high level linguistic skills as well as a good insight in health care. The role of the translators is to:

- translate source language concept into the target language
- proof-read the translated concept description in the target language before passing it on for further review
- raise questions so that principle decisions are made by the Editorial Board (or similar group), concept description in the target language whenever they deem necessary.

6.2 Review

Apart from proof-reading, proper review should be carried out by health and social care professionals. They act as quality assessors and their tasks are to:

- confirm that the translated description reflect the underlying concept (the description's linguistic quality, which should comply with the linguistic guidelines and general rules of the target language, should be checked by professional translators or linguists)
- approve the description that meets the requirements of good translation as in this document and national guidelines
- return unacceptable translations to the translator for correction
- if necessary, address questions on fundamental decisions to the Editorial Board
- consult with the Editorial Board with "questions of doubt"

6.3 Editing

The Editorial Board should consist of an interdisciplinary team made up of professionals with pedagogical and empirical backgrounds in medicine and nursing, linguistics and terminology, information science or technology, paramedical disciplines with sound knowledge and understanding of the English language.

The tasks of the Editorial Board are to:

- define and maintain the guidelines to which all participants involved in the translation process must adhere
- determine the validity of textbooks and references provided to translators and reviewers
- act as an advisory body for translators and reviewers
- make and develop the key decisions that are inevitably required during the translation process
- ensure that all participants are continuously updated on new decisions
- deal with particularly complicated translations and questions from translators, reviewers and others
- approve the terms that meet the requirements after revision of the "terms/cases of doubt"
- collect and register errors and related issues regarding SNOMED CT content in the International Release to be submitted through CRS by the NRC.

6.4 Progress monitoring and follow-up

A project manager and/or coordinator should be appointed to carry out a continuous assessment of the progress of the translation, general project administration and surveillance. The monitoring and follow-up roles are described in the document *Guidelines for Management of Translation of SNOMED CT* (see reference in Section 1.1).

6.5 Post-translation issues

The clinical validation of the translated descriptions plays an important role in the translation of the SNOMED CT terminology into the target language. Some descriptions reflect very complicated or rarely used concepts, and there is a risk that both translator and reviewer, despite all efforts, may have misunderstood the concept in the source terminology. There is also a risk that a particular description, however correct it may be, may be psychologically unacceptable to clinicians if they are used to employing a particular description or phrase for a particular concept. Validation should therefore be carried out by health and social care providers to ensure that the translation is useful in clinical, cultural and social settings.

In addition to the validation of the translated descriptions, a policy for the maintenance of the target language terminology and feedback to SNOMED international is required. If a concept added in a national extension is to be included in the SNOMED CT International Release it must have a FSN in English.

7 Supporting documents

SNOMED CT Document Library

- **Overviews**: https://confluence.ihtsdotools.org/display/DOC/Overviews
- **Guides**: https://confluence.ihtsdotools.org/display/DOC/Guides
  - Starter Guide: https://confluence.ihtsdotools.org/display/DOCSTART
  - Editorial Guide: https://confluence.ihtsdotools.org/display/DOCEG
- **Specifications**: https://confluence.ihtsdotools.org/display/DOC/Specifications
  - Template Syntax Specification: https://confluence.ihtsdotools.org/display/DOCSTS
- **SNOMED glossary**: https://confluence.ihtsdotools.org/display/DOCGLOSS/

National translation guidelines developed by the SNOMED International member countries:

https://www.snomed.org/our-customers/members

Documents consulted for the development of these guidelines

Cabré Castellvi, T. (1999). *Terminology: Theory, Methods and Applications.* Amsterdam/Philadelphia: John Benjamins.

COTSOES (2003). *Recommendations for Terminology Work.* Retrieved from http://www.cotsoes.org/en-publications

INFOTERM (2005). *Guidelines for Terminology Policies.* Paris: UNESCO.

International Health Terminology Standards Development Organisation (2012). *A methodology and toolkit for evaluating SNOMED CT Translation Quality*. Retrieved from https://confluence.ihtsdotools.org/download/attachments/18780052/TQA_background_20120810_v0%2015.doc?api=v2

ISO 1087-1 ((1969, 1990, 1998) 2000). *Terminology work – Vocabulary – Part 1: Theory and application.* Geneva: International Organization for Standardization.

ISO 704 ((1987, 2000) 2009). *Terminology work – Principles and methods.* Geneva: International Organization for Standardization.

Molina, L., & Albir, A.H. (2000). Translation Techniques Revisited: A Dynamic and Functionalist Approach. *Meta, XLVII*(4), 497-512.

Nuopponen, A. (2018). Terminological Concept Systems. In *Languages for Special Purposes: An International Handbook* (pp. 453-468). Berlin/Boston: Mouton De Gruyter. doi: 10.1515/9783110228014-023

Odgen, C., & Richards, I. (1923/1989). *The Meaning of Meaning: Study of the Influence of Language Upon Thought and of the Science of Symbolism* (8 ed.). Harvest/HBJ.

Pavel, S., & Nolet, D. (2001). *Handbook of Terminology.* Ottawa: Public Works and Government Services Canada.

Wüster, E. (1979/1985). *Einführung in die Allgemeine Terminologielehre und Terminologische Lexikographie.* Copenhagen Business School.

Further reading

Bodenreider, O., Smith, B., & Burgun, A. (2004). The Ontology-Epistemology Divide. A Case Study in Medical Terminology. In A. Varzi, & L. Vieu (Eds.), *Proceedings of FOIS 2004, International Conference on Formal Ontology and Information Systems.* Turin, 4-6 November 2004.

Cimino, J. (1998). Desiderata for controlled medical vocabularies in the twenty-first century. *Methods Inf Med., 37*(4-5), 394-403. Retrieved from PMID: 9865037; PMCID: PMC3415631

Madsen, B. (2004). *Handbog i begrebsarbeidje Del 1 & 2 (Handbook of concept work).* Sundhedsstyrelsen *(National Board of Health)*. Copenhagen.

Madsen, B. (2015). Concept modeling vs. data modeling in practice. In F. Steurs & H. Kockaert (Eds.), *Handbbook of Terminology* (pp. 250-275). Amsterdam/Philadelphia: John Benjamins Publishing Company.

Reynoso G.A., March A.D., Berra C.M., Strobietto R.T., Barani M., Lubatti M. et al. (2008.). Development of the Spanish Version of the Systematized Nomenclature of Medicine: Methodology and Main Issues. *Prc AMIA Symp*, (pp. 694-8).

Spackman, K. A., Dionne R., Mays R., & Role J. (2002). Role grouping as an extension to the description logic of Ontology, motivated by concept modeling in SNOMED. *Proc AMIA Symp.*, (pp. 712-6).

Spackman, K. A., & Reynoso G. (2004). Examining SNOMED from the perspective of formal ontological principles: Some preliminary analysis and observations. In U. Hahn (Ed.), *Proceedings of the KR 2004 Workshop on Formal Biomedical Knowledge Representation.* (pp. 72-80). BC, Canada: Whistler.

Toft, B. H. (2004). *Virksomhedsdatabaser – hvorfor og hvordan? En praktisk guide til terminologiarbejde* (Industriel Databases – Why and How? A Practical Guide to Terminology Work). University of Southern Denmark.

Wermuth, M.-C., Walravens, M., & Lambot M.-A. (2022). Collaboration and communities of practice in the field of medical ontology translation. (P. Cadwell, F. Federici, & S. O'Brien, Eds.) *The Journal of Specialised Translation* (37), 75-98.

Appendix A: Translation Quality Assessment

This Appendix contains a brief extract of the information contained in the IHTSDO document *A methodology and toolkit for the assessment of SNOMED CT Translation Quality* (2012) regarding the quality metrics that translation projects should include in their project and quality plans.

Table 4 shows the four (4) **structural** and **process**-related quality characteristics and associated quality metrics that have a **SMART** rating of **GREEN** are considered suitable for use by IHTSDO and mature enough for immediate use, and translation projects should consider them mandatory for use:

Component	Quality Characteristic	SMART Rating
Structure	Participants' knowledge of terminology and terminology translation processes (including also translators and reviewers competencies)	**GREEN**
Structure	Content of style guides and reference materials in the target language	**GREEN**
Process	Ongoing communication, co-operation and translation project process adjustments between the TPO and the TSP	**GREEN**
Process	Translation reviews (two-stage review process required)	**GREEN**

Table 4 – Structural and process-related quality characteristics with SMART rating of GREEN