TECHNICAL REPORT

ISO/TR 12300

First edition
2014-11-15

Health informatics — Principles of mapping between terminological systems

Informatique de santé — Cartographie des terminologies de classifications

Reference number
ISO/TR 12300:2014(E)

ISO copyright office
Case postale 56 • CH-1211 Geneva 20
Tel. + 41 22 749 01 11
Fax + 41 22 749 09 47
E-mail copyright@iso.org
Web www.iso.org

Published in Switzerland

Contents

Page

Foreword

ISO (the International Organization for Standardization) is a worldwide federation of national standards bodies (ISO member bodies). The work of preparing International Standards is normally carried out through ISO technical committees. Each member body interested in a subject for which a technical committee has been established has the right to be represented on that committee. International organizations, governmental and non-governmental, in liaison with ISO, also take part in the work. ISO collaborates closely with the International Electrotechnical Commission (IEC) on all matters of electrotechnical standardization.

The procedures used to develop this document and those intended for its further maintenance are described in the ISO/IEC Directives, Part 1. In particular the different approval criteria needed for the different types of ISO documents should be noted. This document was drafted in accordance with the editorial rules of the ISO/IEC Directives, Part 2 (see www.iso.org/directives).

Attention is drawn to the possibility that some of the elements of this document may be the subject of patent rights. ISO shall not be held responsible for identifying any or all such patent rights. Details of any patent rights identified during the development of the document will be in the Introduction and/or on the ISO list of patent declarations received (see www.iso.org/patents).

Any trade name used in this document is information given for the convenience of users and does not constitute an endorsement.

For an explanation on the meaning of ISO specific terms and expressions related to conformity assessment, as well as information about ISO's adherence to the WTO principles in the Technical Barriers to Trade (TBT), see the following URL: Foreword — Supplementary information.

The committee responsible for this document is ISO/TC 215, *Health informatics*.

Introduction

The benefits of data sharing and reuse are well known. One of the key principles underpinning health informatics is that data should be collected once and reused to the greatest extent possible.

Mapping is the process of associating concepts from one terminological resource to concepts in another terminological resource and defining their equivalence in accordance with a documented rationale and a given purpose. The terminological resources can be related (different versions of the same resource) or completely different resources. The process of mapping identifies whether there is a relationship between the concepts and, if so, the level of meaning expressed by that relationship. It is a way to integrate different terminological resources used for different purposes and where a bridge between them is required for interoperability and that bridge can be built through mapping. Thus, different data sources can be compared and linked to enable the data to be exchanged between information systems, compared over time, or aggregated for different purposes. The end product (deliverable) of the process is a set of individual maps (relationships) between two terminological resources that defines the cardinality and degree of equivalence between concepts and rule set structures and enables the automated translation between the terminological resources.

As an example in health care, data collected for communicating information about direct patient care (using clinical terminologies) can be reused for statistical and administrative reporting of morbidity data (using clinical classifications) by transforming the terminological representations into classification representations.

Terminological resources include all mechanisms for representation of data including terminologies, classifications, and code systems.

Quality maps are always built for a purpose. Skilled mapping personnel are required to ensure the quality and integrity of map development and mapping rules. The development of rules (either paper-based or computer algorithms) that support conversion of data are crucial to standardize the process and create logical maps that a computer can use repeatedly to consistently convert data from one form to another.

This Technical Report provides guidance for organizations charged with creating or applying maps to meet their business needs. It identifies issues and discusses both the potential in and the limitations of applying the map. This Technical Report also establishes and harmonizes the basic principles for developing, maintaining, and using maps and gives guidelines for good practice that underpin the mapping process. This Technical Report does not provide information or guidance on processes required to produce a map in any given situation nor the intellectual property rights of those who own the various terminologies or classifications.

There is a lack of common understanding of the need for mapping between terminological resources, the process of mapping, and requirements for computational functionality in the mapped relationships between the different terminological resources used in health care. Thus, documenting the general principles that underpin the mapping process are essential to good decision making and governance. These will provide guidance about good practice, will support convergence of international knowledge, standardize processes, structure, and approach to the development of infrastructure and tools supporting the mapping process.

There are broadly three core reasons to map data from one code system to another through a map. These include the following:

— support interoperability (information sharing between systems and organizations);

— reuse of data collected for one purpose to meet another purpose (secondary use);

— convert from an older, no longer relevant to purpose terminological resource to a new alternative representation.

Information sharing might require information collected in the local system to be converted to a "common language" such as that represented by international standard terminological resources such as SNOMED CT or ICD. The common language should be agreed upon in order for computer systems to communicate effectively. Any information in local systems not in the common language should be translated (mapped) into the common language and when information is received from others, it should be converted from the common language into the language of the local system.

The increasing use of terminologies to collect data supporting direct patient care has enabled the reuse of this data for other purposes. Data collected for secondary use are generally aggregated and collected through classifications. Secondary use includes, but is not limited to, reusing the information for the following purposes:

a) funding;

b) statistical aggregation and reporting (morbidity and mortality);

c) providing a research basis for evidence-based medicine;

d) measuring quality and safety of care;

e) health planning or setting health policy;

f) monitoring resource utilization;

g) public health surveillance.

Reusing the data through mapping reduces the need for recollection of data, thereby simplifying the administrative burden of data collection, although it should be understood that the administrative burden might increase overall due to the maintenance of mapping when continuing to use multiple code systems and maps. Facilitating the automation between various terminological systems used in health care reduces the costs of providing care and improves the quality of the data and the timeliness (availability).

Decisions on whether or not to map or whether to move from, for example, a classification of clinical information to a more precise clinical terminology needs to be based upon a wide range of factors including the ability to accurately represent meaning, the need to represent information in a manner suited to purpose in each use environment, including the need to aggregate and compare data over time. There are also significant costs and skills associated with mapping. The difference between a "once off" map table to meet a singular, conversion process and the decision to use maps as a long-term mechanism to support reporting and analysis need to be understood by those making decisions on these infrastructure approaches recognizing all of the benefits, requirements, and costs which might include the following:

— decision makers in government, healthcare authorities, and healthcare facilities;

— developers, implementers and managers of health information systems, clinical information systems and clinical decision support systems;

— classification and terminology communities of practice;

— all users of clinical data, such as health statisticians, researchers, public health agencies, health insurance providers, health risk organizations, data analysts, and data managers;

In this Technical Report,

— mapping refers to the establishment of semantic comparability between terminological resources (these resources include terminologies, classifications and other code systems),

— the term "concept" is applied throughout this Technical Report to represent a "unit of thought" expressed in a terminology (it should also be noted that some terminological systems do not explicitly represent concepts, but rather terms, i.e. meaning cannot be assumed explicitly by the code or terms used), and

— the term “terminological resource” is applied throughout this Technical Report to collectively mean either a classification or a terminology used to classify or encode data in healthcare.

Examples are drawn from the published literature on mapping to illustrate key concepts and enhance understanding. However, full and complete understanding of the principles and guidelines requires some background knowledge of the coding of healthcare data, the various terminological systems used, and the many uses of the coded data.

Health informatics — Principles of mapping between terminological systems

1 Scope

This Technical Report provides guidance for organizations charged with creating or applying maps to meet their business needs. It explains the risks inherent in the mapping process and discusses the issues that need to be considered in the development, maintenance, and use of maps in health care. This Technical Report also identifies variations in process, precision, and administration when mapping for different purposes and in different environments.

Importantly, this Technical Report establishes and harmonizes the basic principles for developing, maintaining, and using maps and gives guidelines for good practice that underpin the mapping process. Terminological resources includes terminologies, classifications, and code systems used in the regulatory environment as it relates to healthcare and reporting requirements in healthcare.

This Technical Report is general in nature and does not describe the specific methods applied in the mapping process nor does it describe maps between databases and data sets, even though many of the principles stated here will apply to those types of maps. This Technical Report does not include consideration of the intellectual property rights and expectations of the owners of terminologies or classifications. It is the responsibility of the mapper and process to ensure that these legal rights are protected and acknowledged as part of the mapping processes.

2 Terms and definitions

Where there are terms used in this document that are not defined in this clause, they are considered to be generic to the English language and not specific to this document. Additional definitions and terms can be found at the international health informatics Standards Knowledge Management Tool and Glossary website www.skmtglossary.org.

Terms are presented alphabetically in logical groups and each definition is best understood through an understanding of the whole family of terms to which it belongs.

2.1 General

2.1.1
auto-matching
computational mapping task undertaken using an algorithm based upon the relationship between concepts

Note 1 to entry: Separate files of concept content from different coding systems are compared using an algorithm to determine whether there are concepts which match each other; that is, whether each coding system has content in common

[SOURCE: National eHealth Transition Authority — Australia (NEHTA), 2005]

2.1.2
cross map
see *map* (2.1.11)

2.1.3
cross map target
see *map target* (2.1.15)

2.1.4
data aggregation
process by which information is collected, manipulated, and expressed in summary form

Note 1 to entry: Data aggregation is primarily performed for reporting purposes, policy development, health service management, research, statistical analysis, and population health studies.

[SOURCE: ISO/TS 18308:2004]

2.1.5
cardinality
number of times a data element can repeat within an individual occurrence or object view

EXAMPLE A person can have one date of birth (cardinality 1) but *n* addresses in a lifetime (cardinality of *n* = many)

2.1.6
custodian
one who guards and protects or maintains property or records

2.1.7
equivalence
condition of being equal or the same in value, worth, or function

Note 1 to entry: In terminological systems, two concepts are (semantically) equivalent if their domain of meaning overlap and their semantic definitions are interpreted as identical.

Note 2 to entry: In the context of terminological resources, equivalence and semantic equivalence are often considered as synonyms.

[SOURCE: Oxford Dictionary (2013), modified]

2.1.8
human mapping
use of human knowledge and skill to build maps between concepts and/or terms in different coding systems

Note 1 to entry: Each map is built singly and individually. The process requires examination of each and every concept in each coding system within the scope of the map. Informed judgements or decisions are made about the shared meaning of concepts. Some electronic or computational tools are used but only in support of work process. The use of tools might still require manual oversight to determine equivalence of meaning.

[SOURCE: National eHealth Transition Authority — Australia (NEHTA): 2005, modified]

2.1.9
individual map
cross map
index from one term to another, sometimes using rules that allow translation from one representation to another indicating degree of equivalence

Note 1 to entry: Entry in a map which indicates how to translate from an individual source concept to a target concept. The term map is often used to indicate a table of individual map entries. It is for this reason that the individual and map tables are being differentiated.

Note 2 to entry: The use of this term is often used in ways which are confusing. It is essential to always make it clear whether you are referring to an individual map or a map table (or set).

Note 3 to entry: In SNOMED CT, each individual map is represented as a row or group of rows in a map Reference Set. It links a single map source concept code (e.g. SNOMED CT Concept ID) to one or more codes in a map target (e.g. ICD Code).

Note 4 to entry: A map is often computable and is the outcome of the mapping process.

2.1.10
map
See *individual map* (2.1.9) or *map table* (2.1.11)

2.1.11
map table
map reference set (in SNOMED CT)
map set
group of individual maps used to convert a range of entries from source to target code system

2.1.12
mapping
process of defining a relationship between concepts in one coding system to concepts in another coding system, in accordance with a documented rationale, for a given purpose

Note 1 to entry: Quality mapping will produce a usable map table, be a reproducible and understandable process.

2.1.13
map source
<mapping>
synonym: source
terminology, coding scheme, or classification used as the starting point for map production

Note 1 to entry: Map source is used as a term which can apply to an individual map source (a single code or term), as well as to the coding system. To differentiate between these, individual map source should be used when referring to a single term/concept.

2.1.14
map target
synonym: target (in a map), target scheme
terminology, coding scheme, or classification to which some or all of the concepts in another terminology, coding system, or classification (the map source) are mapped

Note 1 to entry: Map target is used as a term which can apply to an individual map source (a single code or term), as well as to the coding system. To differentiate between these, individual map target should be used when referring to a single term/concept.

Note 2 to entry: In SNOMED CT, some map targets might be derived from two or more associated statements and in these cases; the combination can be expressed as a set of associated rules. Each map target is represented as a row in the map table with each individual map target appearing at least once in the map reference set used to define the map table.

[SOURCE: International Health Terminology Standards Development Organization (IHTSDO) – to be non SNOMED/US specific, modified]

2.1.15
ontology
organization of concept for which a rational argument can be made

[SOURCE: ISO 17117]

2.1.16
scenario
description of high-level business activities defining process and requirements

Note 1 to entry: People often refer to scenarios as use cases, but in the context of this Technical Report, a use case is considered to be a more technical approach as defined in ISO/IEC 19501:2005.

2.1.17
semantic correspondence
measure of similarity between two concepts

Note 1 to entry: This term measures and indicates whether two concepts are semantically equivalent (thereby achieving equivalence) or not.

2.1.18
semantic equivalence
condition of being equal or the same in meaning

Note 1 to entry: In terminological systems, two concepts are semantically equivalent if their domain of meaning overlap and their semantic definitions are interpreted as identical.

Note 2 to entry: In the context of terminological resources, equivalence and semantic equivalence are often considered as synonyms.

2.2 Terminological resources

2.2.1
classification
exhaustive set of mutually exclusive categories to aggregate data at a pre-prescribed level of specialization for a specific purpose

[SOURCE: ISO 17115:2006]

Note 1 to entry: Classifications include a place, though not always specific, for all concepts required for the specific purpose of the classification. They include broad catch all categories and 'unspecified' sections to capture those concepts, where it is not possible or practical, for the purpose to be more specific.

2.2.2
classification
<healthcare>
synonym of statistical classification, statistical healthcare classification, healthcare classification system, and healthcare classification.

EXAMPLE International Statistical Classification of Diseases and Related Health Problems, 10th Revision (ICD-10), International Standard Classification of Occupations.

Note 1 to entry: Classifications include a place, though not always specific, for all concepts required for the specific purpose of the classification. They include broad catch all categories and 'unspecified' sections to capture those concepts, where it is not possible or practical, for the purpose to be more specific.

2.2.3
coding system
combination of a set of concepts, a set of code values, and at least one code scheme mapping code values to coded concepts

[SOURCE: ISO 17115:2006]

EXAMPLE Country codes.

Note 1 to entry: Coded concepts are typically represented by terms but can have other representation. Code values are typically numeric or alphanumeric or a mixture of these, e.g. ICD Code J44.5.

2.2.4
coding system
<information technology>
collection of rules that maps the elements in one set, the "coded set" onto the elements in a second set.

[SOURCE: ISO 2382-4]

2.2.5
coding system
<health messaging>
combination of a set of code meanings and a set of code values, based on a coding scheme.

[SOURCE: Health Level 7 (HL7)]

2.2.6
concept
unit of knowledge created by a unique combination of characteristics

[SOURCE: ISO 1087-1:2000]

Note 1 to entry: A concept can be represented using one or more terms, pictures, icons, or sounds.

Note 2 to entry: Informally, the term "concept" is often used when what is meant is "concept representation". However, this leads to confusion when precise meanings are required. Concepts arise out of human individual and social conceptualizations of the world around them. Concept representations are artefacts constructed of symbols (from ISO 17115).

Note 3 to entry: Concept representations are not necessarily bound to particular languages. However, they are influenced by the social or cultural context of use often leading to different categorizations (from ISO 1087-1:2000).

[SOURCE: http://www.nlm.nih.gov/research/umls/meta2.html]

2.2.7
context
related conditions and situations that provide a useful understanding and meaning of a subject

[SOURCE: ISO/TR 17119:2005]

2.2.8
term
linguistic representation of a concept

Note 1 to entry: A term can contain symbols and have variants, e.g. different forms of spelling.

2.2.9
terminology
structured, human readable, and machine-readable representation of concepts

Note 1 to entry: This includes the relationship of the terminology to the specifications for organizing, communicating, and interpreting such a set of concepts.

[SOURCE: ISO 17115:2007, modified (used by IHTSDO)]

2.2.10
terminology
<healthcare> a terminology that is designed for use in computer systems. The term vocabulary or health or medical language is used to indicate the broader idea of linguistic representation without the specification of computability.

[SOURCE: ISO 17115:2007, modified (used by IHTSDO)]

2.2.11
terminological resource
controlled set of terms in healthcare

Note 1 to entry: Usually designed and controlled for use with computers for specific healthcare purpose, such as data entry, aggregation, retrieval, and analysis.

2.2.12
vocabulary
synonyms: health language, medical language
sum or a stock of words employed by a language, group, individual work, or in a field of knowledge

Note 1 to entry: In health informatics, computable vocabularies, including terms, concept identifiers, etc. are referred to as "terminologies".

[SOURCE: Merriam-Webster's Dictionary]

3 Symbols and abbreviations

DRG	Diagnosis Related Group
ICD	International Classification of Diseases
ICPC	International Classification of Primary Care
IHTSDO	International Health Terminology Standards Development Organization
NANDA	North American Nursing Diagnosis Association — producers of the International Nursing Diagnosis: Definitions and Classification
SNOMED CT	SNOMED Clinical Terms — a clinical terminology owned and maintained by IHTSDO
WHO	World Health Organization

4 Summary of key issues for decision makers

4.1 General

The core requirements for production and/or maintenance of a map table need to be evaluated to determine whether the production of a map table is warranted. If mapping is undertaken, consideration needs to be given on whether/when to convert to the alternative terminology resource and reduce the need for processes to maintain the map table, as well as the relationship and impact of the purpose, risks, and costs to the process of map development.

Though there are many issues and requirements when developing or using a map table, this section of the Technical Report highlights major elements affecting decision making on whether to develop and use a map table and what type of map table to build and the requirements for building and maintaining that map.

4.2 Purpose of the map

A map can only be effectively built if the purpose of that map is clearly defined and, preferably, for a single purpose.

The purpose of a map may be defined using different criteria, for example:

Use case	Healthcare domain
Interoperability	Clinical information exchange
Interoperability	Epidemiological reporting
Reuse	Activity-based funding — mapping from ICD to Diagnosis Related Group (DRG) — a classification to classification mapping

Use case	Healthcare domain
Reuse	Workload review — mapping from clinical record to support workload analysis
Conversion	Admission system — mapping from legacy systems to newly implemented system where the codes used are different

Each domain and use might need different map solutions to represent information in the manner required.

There can be more than one business case for a mapping project with the possibility of multiple end products being produced. For example, in a project that maps SNOMED CT to the International Classification of Diseases — Version 9 Clinical Modification (ICD-9-CM), there may be one map for all SNOMED CT clinical finding concepts to ICD-9-CM and another map of active SNOMED CT clinical finding concepts (only) to ICD-9-CM. In one scenario, all concepts may be needed if legacy databases contain codes that have since been inactivated (retired).

In the second scenario, only active concepts might be needed for mapping current patient data.[6] This reference relates to ICD-9-CM but the same is true of any of the International Classification of Diseases (ICD) family, including ICD-10. Another example is that a map from SNOMED CT to ICD for the purpose of morbidity reporting might need to include the rules related to code allocation relevant to Diagnosis Related Group allocation. These rules are often largely fiscal in their purpose and might not be appropriate for other purposes.

There are also cases where a map is used to support direct patient care. The requirements and implications to care for maps with this use case need to be carefully considered to reduce risks.

If a map is expected to support more than one purpose, there are risks which include the following:

— inaccuracy: accuracy of the map can be compromised as decisions made on the target concept, to which an individual source concept is mapped, is influenced by the purpose;

— cost of development/maintenance: the development of a single map table intended to suit multiple purposes will be more expensive to build and maintain as there are likely to be more decisions required on which target concept to select whenever the match is not precise. Each such decision requires clinical adjudication and documentation of the basis for the decision so that the map can be consistently built and maintained over time. These processes require specialist personnel and governance, each of which is costly.

4.3 Map use frequency

Maps can be developed as a one off to convert from one system to another or used ongoing for translation or reporting. Details of one-time maps versus maps maintained for ongoing use are provided in 5.3. Decision makers need to be aware of the risks associated with ongoing use of maps (maintained maps). These risks include the following:

— Accuracy: Every time a map is used, the potential to decrease the quality and accuracy of the information represented exists. Each time a map is produced, the documentation and rules upon which that map is based needs to be followed and re-evaluated. If a map is intended for single use, this ongoing management is not required;

— Outsourcing of a single-use map to convert from old data to the new data offers the opportunity to obtain required skills for a limited time and specified budget. This makes governance a simpler process but also requires oversight to ensure that the requirements of the map and conversion meet the requirements of all stakeholders;

— Cost: The cost of building and maintaining a map over the long term is expensive. Consideration should be given to conversion to the target system, if this is practical and meets data collection needs, as this will be much more cost effective over the long term.

4.4 Building the map

Building the map requires definition of the scope of what will be included, use, and structure of the resulting map. The building of a map requires clear governance, oversight of relevant users, and terminological, as well as clinical expertise. Maintaining a map requires consistency and reproducibility. To achieve this, quality personnel are needed as is clear and precise documentation of the process and the decisions made during the mapping process.

Tools which can assist with mapping, such as software which matches terms used to describe concepts in two different terminological resources, can be very helpful to speed up the process of mapping. Computer-based tools can also be necessary in order to browse the resources in order to support manual verification or building of each individual map.

Risk Factors

a) Accuracy: Documentation of issues and how each issue is resolved is required to support consistent accuracy of the individual map entries. It will not be unusual to revisit rules or editorial guidelines and adjust the maps accordingly, until development is completed.

b) Clinical Risk: To be clinically safe to use for clinical communication or recording, the meaning intended by the original data should not be changed when mapped and used in the target resource. Application of risk management strategies to minimize and quantify the risk is essential in this environment and will impact the personnel needed and costs of building the map. Risk management strategies should include the following:

   1) organizational risk, risk to patients, risk to staff;

   2) consideration of roles and responsibilities and capacity to deliver, including areas of expertise;

   3) procedures for variations to scope;

   4) analysis of exposure;

   5) delivery, including scheduling budget and staff/consultant management;

   6) communication protocols, including procedures to respond to errors/unforeseen circumstances;

   7) documentation and record management;

   8) staff development, disaster recovery plan, compliance, and development review.

c) Personnel: Suitably qualified personnel are essential to efficiently and accurately build a map. Domain experts might also be required to provide guidance as part of the clinical review or quality assurance process. Details of the skills needed are identified later in this Technical Report, but difficulties in finding these people should not be underestimated. There are also skills and knowledge required for appropriate implementation and use of maps.

d) Cost: Quality documentation and skilled personnel will reduce the cost of developing and maintaining the map but will not be inexpensive.

e) Tools: It is essential that the task to be performed by any tools used is clearly specified to ensure that the tool is actually able to deliver what is required and to minimize manual intervention. The cost of tools and training to use those tools should be included in the overall cost of the project.

f) Use outside purpose: Identify risks associated with using the map outside of its intended purpose and scenarios when it could be used in such ways and develop counter measures. It is important to ensure clarity and intent is maintained.

g) Misinterpretation: Will the use of a map introduce any risk of loss of meaning? “Code drift” can occur when mapping between terminological resources where the lack of an equivalent code results in the choice of a map result which either increases or decreases the results in specific categories to which the concept is assigned.

EXAMPLE 1 A map between the SNOMED CT concept of stress fracture of metacarpal bone (441558003) can be mapped to ICD-10-am Stress fracture, not elsewhere classified Hand (M84.34).

This ICD-10-am code represents stress fractures of the metacarpal bone but also of fingers, joints between these bones and the carpus. The meaning intended has drifted to a less specific concept.

EXAMPLE 2 A map between the SNOMED CT concept of prolonged fever (248435007) does not map precisely to a condition in ICD-10-am This term is not represented in the index to the code system; however, it is possible to code persistent fever (R50.8 — Other specified fever). If this map was established, prolonged fever would be mapped to a code which includes fever with chills, fever with rigors, as well as persistent fever. The meaning has been changed, though slightly. This is code drift.

4.4.1 Use of 'local' terminological resources in maps

Most of the common terminological resources, such as SNOMED CT, LOINC, ICD family of classifications, and WHO (World Health Organization), ISO, or national value domains, have established governance, development, and maintenance procedures which are transparent and well publicized, particularly when they are used in nationally accepted and defined data collections.

However, local or regional resources, such as those used to display information in a specific situation, are usually designed to suit local requirements for encoding healthcare data, such as problem lists or lists of routinely performed interventions. These local resources might need to be mapped to a common language target such as SNOMED CT. Abbreviations used in a data collection or display situation should be approached with caution. Preference is to avoid usage of abbreviations. It is strongly recommended for the mapping activity to expend all abbreviations from both terminologies. That will not prevent using recognized abbreviations in the target terminology that might have multiple descriptions, therefore, might allow using a preferred term that is an abbreviation when the main concept description is all expended and is unambiguous. In that case, only non-ambiguous abbreviations should be used, preferably those of international use and endorsement or at least of national approval. Ambiguity needs to be tested outside the individual specific instance of a department or specialization as once abbreviations are included in a clinical record, their initial context might be less clear and with that, their meaning. An example of this issue was identified recently in an emergency department where clinicians were using the abbreviation ACS which to some meant acute coronary syndrome and to others, altered conscious state. Only when a review of the text in their records was undertaken did this problem become clear.[21]

The creation and use of legacy "local coding systems" should be discouraged and phased out. These systems sometimes do not show a standardized concept and some of these legacy systems represent neither the characteristics of a classification nor those of a coded terminology system.

Local systems are generally developed for pragmatic reasons and are not nationally endorsed, though they might be a locally modified version of an endorsed coding system. The local terminological resource might or might not be robust or well managed or appear within any structural context other than an alphabetical listing. There might be no formal governance or maintenance mechanisms and their development might not conform to any standard principles. They might reflect colloquial or informal terms which might be imprecise, inaccurate, and potentially misleading. Local coding systems arise through need and are often a quick solution to an immediate problem. However, they are a major factor as to why many healthcare organizations have data interoperability problems between their systems. [15]

Though there might be issues of quality with the content or use of code systems in a local system, this does not necessarily mean that local coding systems should not be used in mapping projects. Indeed, these data often require maps to enable migration to a more robust, comparable, terminological resource. The point being made is that the reason for using the local resource and a detailed explanation of how the system was developed, as well as the current status of the local resource, should be transparent and specified in the documentation accompanying the map so that all stakeholders can judge for themselves the veracity of the source terms and the relevance of the map to their business needs.

4.5 Commitment and adequate resourcing

If our health care systems are to achieve the collect once, use often goal inherent in so many of our electronic health record systems, then mapping will be an essential requirement to facilitate the migration of healthcare data from legacy systems to electronic health records or facilitate automated classification assignment for administrative and reporting purposes. Most other maps between terminological resources, however, should be a short-term solution for clinical data management since creation, versioning, and maintenance of the resulting maps is a resource intensive process.

Accordingly, it is crucial for sponsors to consider why the map is needed, what problem the map is designed to solve, and the intended life span of the map. Use cases should show the requirements for maps from the specific source resource to the target resource. Business cases should demonstrate the resources required. Committing to using a mapping means the sponsor is committing to long-term maintenance and that the associated costs and processes for governance and staffing infrastructure, tooling, and vendor software management when a mapping is used are high and on-going. Consider too, that each time either of the underlying resources applied in the map are updated, the entire map should be updated, tested, and re-released.

When using or investing in the development of a map, it is vital that there is clear understanding of the tasks required and the complexity of the task for development and on-going maintenance of the map. Sponsors who require a map to meet a given purpose need to understand that the development of that map is a costly and time-consuming process and that significant specialist skills will be needed. In the context of clinical care, a map between resources is far more complex than the simpler versions used in the administrative data collection systems used almost exclusively in the 20th century.

4.6 Challenges in achieving automated classification assignment on the basis of a map

As mentioned previously, the increasing use of clinical terminologies for encoding data at the point of care will lead to the need for maps that assist with automating or semi-automating the process of classification of healthcare data for administrative and reporting purposes (otherwise known as automated classification assignment). There are existing mappings between SNOMED CT and ICD-10; however, these are intended only to provide guidance to a human coder for the assignment of an ICD-10 code to a SNOMED CT-encoded clinical record. The value of these existing maps is that they can quickly provide guidance to the human coder, based on the knowledge of clinical and mapping experts that is implicit in the suggested codes. The disadvantage of the existing maps is that every new code assignment requires manual review and this generates what is sometimes called “double coding”. This occurs when a particular diagnosis should first be manually coded into a SNOMED CT concept and then later, manually coded into the appropriate ICD-10 code (including the application of morbidity coding rules). It would be more efficient if a code could be assigned only once, using a highly granular clinical terminology, and then all subsequent classifications of the case could be generated in an automated (or perhaps semi-automated) way.

Below are two explanations often challenged when building a software system that can perform the entire classification assignment on the basis of mapped codes without human intervention. These examples relate to a SNOMED CT to ICD-10 map.

EXAMPLE 1 Correct ICD code assignment depends on the entire record. A SNOMED CT encoded clinical record is very likely to have multiple SNOMED CT codes that describe the clinical findings, diagnoses, and procedures that were part of a given healthcare episode. These codes, in turn, will be embedded within the information architecture of the electronic health record, which will carry the dates, times, record headers, and other contextual information that might influence the way a particular SNOMED CT code is to be interpreted. All of this information is to be taken into account in order to assign the correct ICD-10 code. This requirement is true for all use cases for representation of the source concept, the map, and the resultant information.

In the instance of converting SNOMED CT information, the recording of "hypertension of pregnancy" in a field of conditions confirmed not present. This is useful information for morbidity coding in ICD-10 but requires interpretation of the field in which the information was contained.

EXAMPLE 2 ICD-10 requires attention to inclusion and exclusion criteria ICD-10 requires attention to WHO (ICD-10, Volume 2 Disease Index) and National Guidance or coding standards, all rules and conventions at chapter, block, and individual code level, including inclusion and exclusion criteria, as well as the requirements established for the use case of the classification, such as mortality, morbidity, or diagnosis representation. Initially, in 1994, ICD-10 was released internationally as a diagnostic classification for general epidemiology, health management, and clinical use, such as national analysis of public health and general health of the population, as well as for reimbursement purposes, resource allocation, health quality, and in support of the creation of health guidelines. ICD-10 would be used to classify disorders and health problems to store and retrieve these encoded data for diagnostic information for clinical, epidemiology, quality purposes, and the population of national and WHO mortality and morbidity statistics. Thus, creating the opportunity to analyse, interpret, and compare mortality and morbidity data collected from different countries using the same classification system. In 2004, ICD-10 2nd edition amended new guidelines for recording and coding of clinical data and some practical aspects of clinical coding. ICD-10 was not created to encode every health problem leading to a health services encounter. However, ICD-10 does provide a wide variety of signs, symptoms, abnormal findings, health complaints, and social situations with explicit context, which could stand in place of a diagnosis.

To avoid double-counting, exclusion criteria are applied. For example, the ICD-10 code D62 Acute post haemorrhagic anaemia has an exclusion criterion for "congenital anaemia from foetal blood loss", which should be coded instead as P61.3. If a SNOMED CT-encoded record for a newborn contains a SNOMED CT code for the clinical finding of "anaemia due to acute blood loss", selecting the right ICD-10 code requires examination of the standards, inclusion and exclusion criteria, and interpretation of the context of a newborn record.

The challenge for those attempting to automate the coding process on the basis of a map is to create a rule-based knowledge structure within the map that can be used by intelligent software to properly check all the inclusion and exclusion criteria required by the ICD-10 classification. Automated mapping can be achieved when, after consideration of the inclusion and exclusion criteria, the map result has a cardinality of one-to-one. More often than not, however, maps have a cardinality of one-to-many and thus, human intervention is required to validate the decision.

It is also important to consider who will maintain the maps required and what skills and knowledge will be required to maintain both the map and the implementation of the map into information systems. Also, relevant is the impact of standards for the use of terminological resources upon maps, such as modifications to coding rules and guidelines, which might not change the actual concept but change the way it is intended to be used in a given context.

4.7 Context is crucial to map production

Subclause 4.6 indicates the need for clear understanding of the record structure within which concepts are represented when automating conversion from one code system to another through a map. This clause looks at context more generally, in particular, the need to understand how context and use case change the intended meaning. This should be understood and addressed when creating a map. The understanding of context should consider the intended meaning by the clinician at the time of collection. Concepts derive meaning through the place they sit within the hierarchy and the relationships they might have with other terms within that terminological resource.[14] Context often acquires a meaning through use (often unintended) but based upon the interpretation of terms available to clinicians at the time of making record entries. Mapping decisions might also need to take into account this prior usage data (what a code has been used for in contrast to its intended meaning on creation and publication).

Context is not only relevant when automating code conversion for mapping purposes; however, it becomes a real issue when mapping from local coding systems, such as problem lists or lists of procedures in which the concepts are generally not systematically structured; the only structure might be an alphabetical listing. Without context, a concept can be ambiguous and thus, potentially misleading and individual mapping personnel might interpret the concept differently on the basis of their existing knowledge and skill level. For example, a concept listed as 'raised blood pressure' in a problem list could be linked to a clinical finding or a disorder in SNOMED CT depending on the context. Does the concept mean an occasion of high blood pressure or does it mean the ongoing medical condition?[14]

Context is crucial for maps built for the purpose of enabling automated code assignment from a source to target terminological resource and is largely why the process may never be fully automated and human intervention is required. For example, if "hypertension of pregnancy" is recorded in a SNOMED CT encoded record, and also elsewhere in this record there is a laboratory finding of proteinuria during the pregnancy, the presence of this laboratory finding in the record changes the ICD code that should be assigned. Some ICD code assignments are also dependent on knowing the patient's age and sex and any automated rules built to assist automated code assignment need to take context and coding conventions and rules into account.

The key message relevant to context is that mapping and automated code assignment are extremely difficult to do in a context vacuum, as context influences meaning and accurate, consistent, and credible code assignment.[18]

4.8 Implementing a map table

How a map table is integrated and used is dependent on the purpose of the map, the scenario, users' requirements, and local software implementation and coding /classification expertise.[7] Good practice in the implementation of a map requires that users and implementers understand the purpose for which the map was intended to be used. If used for other purposes, it can result in misrepresentation of clinical concepts. It is bad practice to use a map for purposes other than that intended and this can result in misrepresentation of clinical concepts. For example, maps cannot be used to add specificity to healthcare data that was captured at a more generic level.[17] Implementers will often embed the content of the maps into existing applications and workflow processes. Improper use of the maps could lead to user frustration and unexpected, potentially unsafe outcomes.

4.9 Decision making matrix

The matrix below summarizes the major issues for decision makers and should be used to assess the quality, benefits, and risks of building and using a map table.

Quality indicators are used to identify the accuracy of the map table resulting from the process used. This might be assessed by considering whether quality measures for validation have been put in place and the likelihood of these identifying any errors. High, medium, or low is used in the sample provided in Table 1 to indicate whether this characteristic delivers high-quality, medium-quality, or low-quality maps.

Cost indicators are used to assist in identifying whether a given characteristic represents a high, medium, or low cost of mapping. The cost of maps is dependent upon the build/maintenance frequency, team size and construction, and tools available.

Clinical risk indicator is used to provide guidance on whether a specific characteristic impacts safe clinical use of the map. High risk means there will be change of clinical meaning which, if used in clinical record keeping or communication, has the potential to cause harm to the patient.

Table 1 — Map characterists and risks

Map characteristic	Quality indicator	Cost indicator	Clinical risk	Comment
One-time map (e.g. for conversion from one system to another)	—	Med to Low	Potential risk if historical data changes meaning	There might be some clinical risk if the conversion of data changes the meaning in the historical patient record.
Permanent map	—	High to Med	—	
— Indicates that the characteristic does not impact this indicator.				

Table 1 *(continued)*

Map characteristic	Quality indicator	Cost indicator	Clinical risk	Comment
Single-purpose map	High	—	Low in the long term	Though the quality of the map might vary, once converted, the data will no longer require a map and is therefore likely to produce higher quality data in the long term.
Multipurpose map	Med to Low	High to Med	High	There are risks that the map will not consistently represent concepts in a manner suited to all purposes. It is more difficult to develop a multipurpose map and therefore, more costly to develop and maintain.
Business requirement/ scenarios well defined and within purpose	High	—	—	The specification of quality scenarios for purpose not only improves the quality of the map as it is developed and used, but also supports the definition of the relevant usage of the data in a clinical environment.
Clear definition of purpose	High to Med	—	High impact (not always negative)	A clearly defined purpose improves the quality of the map and reduces clinical risk.
Clear understanding of how map will be used	High to Med	—	High impact (not always negative)	Understanding how the map will be used has the potential to improve the decision making made when building the map.
Known maintenance cycle and map lifecycle	Med	High	—	The lifecycle of source and target updates impacts the frequency of map update, which impact potential errors and cost. The more frequent the maintenance cycle, the more expensive and potentially lower quality the map.
Skills available: Required skills for development and maintenance are available and appropriate	High	Medium	Low to High depending upon whether the map is being used in a clinical environment	If these skills are not available, the map quality will be reduced and the cost of the project is likely to increase.
— Indicates that the characteristic does not impact this indicator.				

Table 1 *(continued)*

Map characteristic	Quality indicator	Cost indicator	Clinical risk	Comment
Local system additions or extensions are required	Med to High	Low to Medium	Low	Local systems, as well as national centres, bear costs associated with each release and the maintenance of maps required. If using a nationally maintained terminology or classification natively within local systems, it is likely to be cheaper and more accurate than maintaining a totally separate code system (terminological system).
Detailed methodology of decisions made when maps are maintained and applied when building and maintaining the map	High	Low	Med to High if decisions are not consistent over time, the meaning will also be inconsistent and might represent risk in clinical use	This is a quality support activity and a cost-reduction activity as it reduces the need to re-decide how to manage conflicts when mapping.
Source and Target concepts express the clinical idea without ambiguity, to the level of detail required	High	Low	—	Where concepts in both mapping systems are clearly defined, consistency of the map will be easier to maintain.
Source is more detailed than the target	High		High	There is considerable clinical risk in using the target information for direct patient care where the target system is less detailed than the source.
— Indicates that the characteristic does not impact this indicator.				

5 Principles of mapping

5.1 Overview

In this Clause, the key issues previously highlighted and summarized succinctly for decision makers are extended. To understand these issues, it is essential to understand the principles of the mapping process and the utility of the maps. Decision making is based upon this knowledge and assessment of the following:

— purpose of the map and its business utility;

— clinical safety issues which might result from changing the way in which a concept is represented;

— anticipated costs required to build and maintain the map and the systems which use it;

— skills required to build and maintain a map.

This Technical Report identifies and explains each of these concepts and supports evaluation of requirements and risks for any given purpose.

5.2 Terminologies vs classifications

It is common to question why healthcare is moving from classification-based representation to computable terminologies. To understand this change and the relevant place of maps in these changes, it is important to understand broadly the difference between a terminology and a classification, as these terms are used often. The description offered here is high level but sufficient to understand the needs for change. The differences are represented here by SNOMED CT (as the terminology) and ICD-based codes as the classification. The principles are true to varying degrees between all terminologies and classifications.

Classifications have certain common attributes which support reporting and aggregation of data to meet statistical and administrative data collections, which are part of national minimum data sets and required reporting formats. Well-behaving terminologies, as defined by Cimino, represent precise meaning and have computable relationships which allow logic to be applied, which supports automated clinical decision making and precise clinical communication. These differences are summarized in Table 2. Individual instances of terminologies and classifications might vary from these principles in different degrees but the principles are generally true in healthcare.

Table 2 — Comparison of classification and terminology attributes

Attribute	Classification (example ICD-10)	Terminology (example SNOMED CT)
Supplementary concepts, such as — not elsewhere classified — not otherwise specified	Yes Example: A28.8 Other specified zoonotic bacterial diseases, not elsewhere classified	Should not include these concepts as the intention is to represent precise meaning which is often covered by selection of a concept which is less specific. For example: choice of Asthma, rather than a specific type of asthma.
Representation of "catch all" codes such as — other	Yes A05.8 Other specified bacterial food-borne intoxications	No A terminology should not include these concepts as the intention is to represent precise meaning. If there are concepts which need to be described which are not in the terminology, this should be discussed with your National Release Centre.
Represent more than one concept with a single code	Yes, they can but they do not have to (mixed functionality) J11.0 Influenza with pneumonia, virus not identified	Yes The description, logic, and concept model provides definitional information about a concept which might include the causative agent, as well as the condition.
All inclusive – a place to represent anything in the area for which the classification or terminology is used	Yes R53 Malaise and fatigue, includes asthenia not otherwise specified, debility — chronic and not otherwise specified, general physical deterioration, lethargy, and tiredness	No SNOMED CT is not complete, but has the capacity to represent any concept needed. It is being constantly reviewed and expanded as required.
NOTE There are concepts in SNOMED CT today with the words not otherwise specified; these are the result of loading of ICD concepts some time ago. These concepts are being more appropriately modelled and represented and are not to be used.		

Table 2 *(continued)*

Attribute	Classification (example ICD-10)	Terminology (example SNOMED CT)
Computable logical ontological structure and relationships. The structure and relationships of a concept define the concept. Concepts are defined written descriptions, coding rules and guidance, and the place of the code within the classification.	No ICD-10 has hierarchical relationships which are not strictly ontologically-based and though they represent a child/parent relationship within the classification, it is not possible to determine anything more than this from the code system.	Yes SNOMED CT has a definitional structure which supports a computer identifying relationships between concepts and is therefore able to determine equivalence and subsumption relationships and meaning of a concept.
Logical, computable relationships such as — Is a — has finding site — has clinical course — has associated morphology	No — limited The only relationships in ICD-10 are those of parent and child or those created through dual codes such morphology codes. The code system, therefore, provides these relationships through usage not inherently through the classification structures themselves	Yes SNOMED CT has full relationship specification, including a wide range of definitional relationships such as those suggested here. There are many more types of relationships than these and new relationship types can be established when needed to accurately represent meaning.
A concept might have multiple parents.	No — ICD, Yes ICPC 2	Yes SNOMED CT concepts will have as many parent concepts as needed to define the concept.
NOTE There are concepts in SNOMED CT today with the words not otherwise specified; these are the result of loading of ICD concepts some time ago. These concepts are being more appropriately modelled and represented and are not to be used.		

Table 2 *(continued)*

Attribute	Classification (example ICD-10)	Terminology (example SNOMED CT)
A concept has multiple descriptions from which one can be chosen as the preferred description or be used as the commonly used terms for clinicians or other defined users.	No In ICD, each code has one specific description to be used. I21.9 Acute myocardial infarction, unspecified.	Yes SNOMED CT includes many descriptions of a concept, often in alternative languages. Some of these descriptions might be appropriate for clinical use in the record or might suit display for patients. For example: Myocardial infarction (disorder) Myocardial infarction (preferred term) Infarction of heart Cardiac infarction Heart attack (could be preferred term for patient representation)
A concept is defined by attribute-value pairs that allow subsumption.	No For example: In ICD-10, the single hierarchy does provide a concept of parent-child relationships. However, the fact that Acute transmural myocardial infarction of anterior wall (I21.0) is a type of Acute myocardial infarction (I21), can only be determined by assumption using the code not through attributes of the concept.	Yes SNOMED CT may use a wide range of attribute-value pairs. Acute myocardial infarction of anterior wall (54329005) has a defining attribute of IS A Acute myocardial infarction (57054005) These definitions used throughout SNOMED CT support computer calculation of parent/child relationships, including subsumption.
NOTE There are concepts in SNOMED CT today with the words not otherwise specified; these are the result of loading of ICD concepts some time ago. These concepts are being more appropriately modelled and represented and are not to be used.		

The differences between terminologies and classifications do not mean that one is better than the other, simply that one might suit a given business purpose better than the other.

EXAMPLE Direct patient care requires precision and the relationships in a terminology support automated decision making, while consistent aggregation and coding rules support quality statistical and administrative purposes.

Terminologies and classifications sometimes use more than one concept to represent a meaning. In terminologies, this practice (post-coordination) has governance rules which support computability. These rules might be highly complex to implement in computer systems and should not be assumed to be achievable without confirmation with software providers. Classifications also use multiple codes to indicate meaning, such as the use of a malignant body part code in conjunction with a morphology code to indicate the type of cancer. If a map has requirements to work with these dual (or more) code representations, this will add significantly to the complexity of map development, as well as the complexity and costs of implementation.

5.3 One-time or maintained maps

Maps are used in many ways, but fall into the following two broad groups:

a) one-time maps;

b) permanent (maintained) maps.

One-time maps are used when changing from one terminological resource to another, as in when converting data from an old system to a new one where the ”old” code system will no longer be used.

Permanent maps are used when both source and target resources are to be kept in production (actively used).

EXAMPLE 1 Clinical data collection requires both the specificity and the utility of a clinical terminology (including the knowledge inherent in the relationships of that terminology), but this data are also needed by others in a format suited to comparison with pre-existing classified data or more aggregated or statistical analysis there will need to be a map between the clinical terminology and the classification. In this example, the data source is the detailed clinical terminology which is mapped to the classification system for secondary use.

EXAMPLE 2 Clinical data collected in a local system using a local terminology developed to meet local requirements needs to be converted to the national terminology in order to send information to other care providers (interoperability). If this map is used to support the ongoing use of the local terminology, this would be a permanent map.

In these circumstances, it might or might not be practical to use the shared national terminology in your local system, in which case, conversion to that terminology might be a solution (one-time map). Where this is not practical or, as in the first example, there is a need for data at more than one level of specificity, a map is needed (permanent map).

5.4 Business case for mapping

Mapping is a way to integrate different terminological resources used for different purposes; a bridge between them is required for interoperability and that bridge can be built through the use of a purpose-specific map. Thus, different data sources can be compared and linked to enable the data to be exchanged between information systems.

The end product (deliverable) of the process is a map table between two terminological resources that defines the cardinality and degree of equivalence between concepts and any rule used to determine the result of the map (such as age/sex differences in the map) and enables the automated translation between the coding systems.

A common example of the use of a map in health care is when data are collected for communicating information about direct patient care (using clinical terminologies) which can be reused for statistical and administrative reporting of morbidity data (using clinical classifications) by transforming the terminological representations into classification representations.

Map tables are always built for a purpose. Skilled mapping personnel are required to ensure the quality and integrity of map development and mapping rules. The development of rules (either paper-based or computer algorithms) that support conversion of data are crucial to standardize the process and create logical maps that a computer can use repeatedly to consistently convert data from one form to another.

It is important to understand the business case for mapping and the processes of mapping, as well as the requirements for computational functionality in the mapped relationships between the different coding systems used in health care when deciding whether to map and what type of map to develop. The general principles that underpin the mapping process and providing guidance about good practice supports convergence of international knowledge, standardizes processes, and facilitates a consistent structure and approach to the development of infrastructure and tools to support the mapping process are essential to good decision making and governance for mapping processes.

Scenarios should be clearly articulated to, and agreed upon by, all stakeholders prior to designing the requirements of the maps themselves. The following sections are scenarios which require a map to meet a given purpose.

5.5 Why are maps used?

5.5.1 Mapping for health statistics

The increasing use of terminologies to collect data supporting direct patient care has the potential to make that data available for other purposes (reuse or secondary use). Data collected for statistical purposes are generally less specific than that required for direct clinical care and are therefore, often aggregated. Classifications are often used for this purpose.

Example of the difference in specificity between a terminology and a classification.
Source SNOMED CT   Clinical use: 88776002 AC Globulin deficiency factor
Target ICD-10   Secondary use: D68.2 hereditary deficiency of other clotting factor
The target code includes more than 10 different hereditary clotting factor deficiencies, each quite rare. This is sufficient detail for statistical purposes but insufficient for clinical research or direct patient care.

Secondary use includes, but is not limited to, reusing the information for the following purposes:

— funding;

— statistical aggregation and reporting (morbidity and mortality);

— providing a research basis for evidence-based medicine;

— measuring quality and safety of care;

— health planning or setting health policy;

— monitoring resource utilization;

— public health surveillance.

Where data are captured administratively, there might be no need to use terminologies as their greatest utility relates to clinical data collection and interchange. Information collection should fit purpose at the point of collection therefore, collection of data in classifications is appropriate, if that classification content and use best fits the original collection purpose.

Where it is necessary to collect data using a clinical terminology AND that data are required for reuse to be aggregated and classified for utility in other usage areas, a map will be required. The direction of this map is from the clinical terminology to the classification and the goal is to identify the classification concepts that best fit the reuse requirements. This is likely to mean that a single target classification might represent multiple clinical terminology concepts.

Maps enable clinicians to collect data of value to the patient's care and for this data to automatically be converted and aggregated into statistical classifications which are more useful for the purposes defined above. One does not replace the other, each are useful for the purpose for which they are intended.

Facilitating the automation between terminological resources used in healthcare reduces the costs of data collection and improves the quality and the timeliness (availability) of that data.

Users of maps for this purpose need to be wary as this type of mapping can often be the most difficult to use as it is not generally possible to meet the statistical requirements, for example, for financial and epidemiological purposes in a single map.

5.5.2 Maps for the migration/preservation of legacy data

When terminological resources are changed or updated or a resource is planned, implemented, or mandated for use, maps are developed to

a) preserve and enable continued use (querying) of legacy data or

b) facilitate migration from the superseded resource to the new resource (such maps might be a temporary solution until complete migration is achieved or they might be a more permanent solution, that is until the legacy data are no longer useful).

These maps require the following:

a) Scenarios to describe how the map tables (forward or reversed) between terminological resources should be used and might need to be approached differently depending on their intended end use;

b) The rules and data structure of each of the terminological resources be understood;

c) That the meaning of the concept in the source resource should be equivalent to the meaning of the concept in the target resource including all rules and convention of the target resource and the relevance of inclusions and exclusions, as well as aggregations for each individual map record. This is often difficult and sometimes impossible to achieve as terminological resources are each developed for very different use cases;

d) Where applicable, representation and constraints of compositional grammar need to be taken into account for each individual map. This is particularly true when using SNOMED CT.

When statistical classifications which are used for the aggregating and reporting of coded health data are updated, maps that demonstrate the relationships or links between the superseded version and the new version are developed to support data migration and historical data retrieval. In this way, multiple maps might need to be developed over time to continue to maintain historical comparability.

EXAMPLE The following is an example of migration.

When changing from ICD9 to ICD-10, it was essential that data collected in ICD9 be able to be compared to new data collected in ICD-10. Without this comparability, there would be no national statistical trends available for healthcare over the years where the changeover occurred.

A map table to represent these data includes details which explain the concepts which were deleted or added and which were expanded and made more specific. Clinical researchers performing population studies require temporal continuity for longitudinal and trend analysis. The aim is to achieve a map with the most appropriate clinical meaning.[1]

5.5.3 Maps to support historical tracking

Maps that demonstrate the relationships or links between the superseded version and the new version of a terminological resource can support not only data continuity, but also historical tracking of the changes made between versions.

The direction of these maps is usually bidirectional, which is from the source to the target and from the target to the source, which might not be as simple as reading the individual map record from one direction or the other but might require unique entries to represent map requirements in each different direction.

The aim is to provide a chronicle that assists users of the data to understand how and why the changes were made and is particularly useful for researchers performing longitudinal studies, epidemiological studies, and trend analysis to determine improvements (or not) in patient outcomes. The tracking of changes is an essential requirement for maintenance of a coding system and mapping then also becomes a maintenance mechanism in this context.

A quality clinical terminology manages historical tracking between updates not through mapping per se, but through the history mechanism in the terminology metadata or structures.

EXAMPLE 1 SNOMED CT release format 2 provides historical information to support map maintenance but all maps needs to be considered when a new release occurs.

EXAMPLE 2 Updates to ICD impact the DRG generation algorithms (which are effectively a map). Use of historical maps for ICD to DRG with a new version of ICD might result in an inappropriate DRG assignment.

Historical tracking is usually required to support a one off map if that map is to be used to compare data from historical records, while it is essential for a permanent map to have access to this type of information if the record is to be queried over time.

5.5.4 Maps to support funding mechanisms

Mapping is performed to maintain grouping logic so that casemix funding mechanisms for health services are not impacted. The logic used to identify DRG classifications is built from the ICD codes and patient demographic information, thus the systems are intrinsically linked.

Tables of equivalence were introduced to accompany historical versions of the classification system and a record from the table of equivalence is selected purely to enable correct DRG assignment. However, if classification systems change throughout different versions, the DRG group algorithms and data sets have to be changed as well. This enables complete coverage of the classification system and to ensure that the main diagnostic criteria (MDC) "drives" the appropriate DRG grouping for appropriate allocation of funding and/or support appropriate national statistics and/or data sets. The aim is to achieve a map for the most appropriate DRG assignment.

NOTE ICD-10 1st Edition from 1994 until ICD-10 2nd Edition from 2004 and in future ICD-10 and ICD11: new classification codes and their description might have different rules and conventions, as well as different inclusions and exclusions and different aggregations.

5.5.5 Maps for data continuity

Mapping can also be undertaken because different computer systems collect data in different ways with different codes but despite different system requirements, each of these systems need to share information (have semantic interoperability — accurately share meaning). The simplest and most accurate way to achieve semantic interoperability is to have all systems using the same terminology or classification for a given purpose, with a common model or information structure. However, given the history of multiple systems in healthcare and the need to function in an environment of old and new, achieving this level of standardised information representation and structure is not likely to occur quickly, simply, or immediately. Maps are often seen as the solution to this problem.

5.5.6 Maps for the integration of other coding systems in use

Mapping is performed for communication exchange between health sectors providing interoperability when different coding systems are used for different purposes.

EXAMPLE 1 International Classification of Primary Care — Version 2 (ICPC 2) is used for primary care and ICD for acute care. ICPC 2 is mapped to the ICD 10 to demonstrate the compatibility and facilitate linkage between the two classifications. The mapping is useful for conversion of data from ICPC 2 to ICD 10 and vice versa and thus provides a common language between primary care services and acute care services.[2]

EXAMPLE 2 International Nursing Diagnosis: Definitions and Classification (NANDA) taxonomy map to SNOMED CT. The scenario was developed to create maps that would enable users of SNOMED CT to interface with NANDA for point of care documentation in electronic health records. An analysis was firstly performed to ensure that all NANDA concepts were included in SNOMED CT.[6]

5.5.7 Maps to support the creation of clinical terminology local/specific use subsets

One other scenario for using a map has arisen from the use of clinical terminology as a common underlying set of concepts to represent local domain specific concepts.

EXAMPLE SNOMED CT contains more than one textual description of a concept. This feature is designed to provide access to short form, display, or interface terms. These terms can be established through the use of reference set mechanisms in which a smaller, more manageable subset of concepts or terms are arranged to suit a particular domain or use case.

If the map shows that the terminology covers the content of the resource required, then the current resource can be superseded by a reference set from the more extensive terminology. If a complete and reliable map can be done once, the local coding system can be retired. The reference set will be maintained on local and/or national level and the maintenance of the full terminology is managed by the responsible organization(s).

Mapping processes can be used to undertake a gap analysis or content coverage study. Such a study aims to identify where there are concepts in the target terminological systems which are not in the source system (and vice versa). These gaps can initially be recorded as 'no map' but also require further assessment. In some cases, such assessment might result in requests for creation of new content in the target or source. In the case of SNOMED CT, when a gap is identified, the missing concepts can be requested through a change request mechanism that each member country of IHTSDO has in place. This allow users to get concepts and codes in a timely manner for their project that can be used nationally and, if relevant, internationally.

EXAMPLE Reference [4] evaluated the ability of SNOMED CT to represent the most common problems seen at the Mayo Clinic in the United States. They found that when using the compositional nature of SNOMED CT, coverage extended to 92.3 % of the terms used commonly in medical problem lists. They recommended improvements to the target resource (SNOMED CT) through the addition of synonyms and missing modifiers which would lead to greater coverage.

Reference [25] describes a similar scenario, but with a slightly different purpose. A gap analysis is performed with the purpose of building a term in the target resource solely in order to provide a map for the source term. This type of scenario would indicate that the map is required for long-term implementation.

6 Characteristics of a quality map and mapping process

6.1 Clearly declared purpose

Every mapping project should explicitly state what the map is intended to achieve. Clearly defining the purpose allows understanding and agreement among all stakeholders. Defining the purpose leads on to defining the scope and intended audience for the mapping and the development of specific scenarios. There might be more than one purpose for the map and this might result in more than one kind of end product (more than one map).[6]

When making decisions about the target (result) of an individual map, there might be many choices and the purpose of the map will impact the single selection made. This selective nature of the process should be documented to ensure consistency throughout the map but also so that users of the map use it with awareness of the limitations that these decisions impose upon the utility of the map. A map created for fiscal reporting might not be appropriate or accurate to use for clinical or epidemiological research.

With the advent of electronic health records and the increasing use of terminologies for encoding and exchanging data at the point of care, maps from terminologies to classifications may be performed to assist with automating or semi-automating the process of classification of healthcare data for administrative and reporting purposes (otherwise known as automated code assignment). The expectation is that the development of automated coding technologies will greatly impact traditional coding practices such as those used for morbidity reporting by improving productivity, reducing costs incurred through entering data multiple times, and enhancing the accuracy of coded data.[6]

The end product is heavily dependent on its purpose and a "one size fits all" approach to mapping is not possible.[3] A map created to support funding mechanisms might apply different rules and produce a set of relationships that differ in the level of specificity from that of a mapping performed to support a clinical decision support application at the point of care. The level of specificity required at the point of care, certainly from a clinical safety (fit for purpose) aspect, will be finer-grained (more detailed) than that required for funding. It is possible that different purposes of a map might generate results between the same source and target terminological resource which produce different results. This does not mean that the maps are wrong but that the purposes are different.

Therefore, a map is always created with a specific purpose in mind and should be refined for particular scenario/s and users across healthcare settings.

Principle 1: Each map should have a (preferably single) declared purpose.

6.2 Use scenarios to define purpose and user requirements

Every mapping project should have articulated scenarios (or use cases) that originate from the business case, and through requirements gathering and refinement, these provide a framework for defining the scope, audience, purpose, methods [including guidelines, rules (or heuristics), coding systems and their versions], quality assurance plan (including validation and testing), maintenance and evaluation plan, and implementation guidelines.[6]

Scenarios should define the intended users of the map and include scenarios to explain how users are expected to interact with the map. This includes how a user selects the source of the map and the required target/s.[6]

Scenarios should be clearly articulated and agreed upon by all stakeholders prior to designing the requirements of the maps themselves.

Scenarios provide business-based examples against which to test and validate the mapping process and decision making. Collecting as much information as possible at the beginning of the mapping project provides a solid framework for the maps. Scenarios are expected to change or be refined as the mapping project progresses.[25]

6.2.1 Development of scenarios

Some examples of questions to guide the development of scenarios are the following:[7],[25]

— How will the mapped data be used? Will it be used for clinical decision support, funding, research, outcome measurements, or public health studies?

— How will the value of the legacy data be retained if migrating to a newer coding system?

— If each coding system has a different purpose, how does this impact the mapping?

— What degree of aggregation in the maps is acceptable? Will the mapped data be categorized, classified, or grouped into other data sets? To what degree will a loss of meaning be tolerated? The concept of "code drift" should be understood and the impact of this factor assessed. Code drift occurs when mapping between terminologies and/or classifications where the lack of an equivalent code might result in an increase or decrease in other categories where the clinical term might be assigned.

— How will the mapping be performed: automated methods, manual methods, or a combination of both?

— How will the data be transmitted or exchanged? What systems will be involved?

— Will there be benefits in terms of cost reduction or data entry error reduction?

— How will the mapping be maintained? What resources are required? How will updates to source and target coding systems be managed? What impact will updates have?

EXAMPLE

A map is needed for clinical decision support, the rules and guidelines of which are held nationally in SNOMED CT while the local system stores the required clinical concepts in reason for presentation field in a local code system. There is currently no intention to migrate the data in the local system to SNOMED CT and therefore this map will be used in an ongoing manner (a perpetual map). The local code system is used in direct patient care to represent the meaning of concepts at the same level of granularity as SNOMED CT. Where differences between the local code system and SNOMED CT are found, the local code system will be modified unless there is a need for more detail in SNOMED CT in which case, extensions will be requested of the national release centre.

The map will be implemented automatically. The map provided is to have sufficient information to convert accurately (achieving semantic equivalence) using computer system processes only. No human intervention is intended to be used.

Reason for presentation data will be extracted from the local system at triggers of pathology test order entry or medication entry. At this point, the system will convert local code into SNOMED CT using the map. The local system will look up the relevant SNOMED CT code in the relevant national clinical decision support rules table (knowledge base) which is held in the local system. The relevant table is dependent upon the action which triggered the "call" either the medication clinical decision support knowledge base or the pathology test clinical decision support knowledge base. Guidance is displayed in the local system where the clinician determines the next action. The system is to store the version of the map which provided guidance in the electronic health record as a legal record of advice provided and action taken.

The benefits of using a map are, principally, that the local system does not need to modify the codes currently used. These codes are the basis of a significant existing rule structure within the local system. Use of the map and ongoing costs of maintenance will be reviewed every two years against the cost of modifying the local system.

The map will be maintained by review of all SNOMED CT concepts in the current map to see if there are any modifications made to those concepts. These modifications will be updated in the map and released within two months of the release of the national clinical decision support rules table.

Principle 2: Scenarios are developed and articulated to define the requirements for the map table.

6.3 Machine processable format

To be of value in computer-based systems, the set of maps between two terminological resources and the rules governing their use should be in a machine-processable format to maximize the potential of automated translation and consistent application.

It is common that a format will be established while building the map which has more information than what proves to be necessary in the final map. This information might include details of the current status of each individual map, including documentation of issues. Once these have been resolved, they are retained as part of the documentation of the map but might not be included in the final computable version released for use.

Principle 3: The map table should be in a machine-processable format.

6.4 Identify the versions

Both the source and target resources should be identified by the version of that resource used in the map. Where either source or target version is updated, a new and updated map is also required to be produced. Where governance processes are maintained and applied, these maps may be considered to be "authoritative resources".

Principle 4: Identify each version of each terminological resource as a version of the map table.

6.5 Development/maintenance team skills

For the credibility of the mapping project and to instill confidence in the end product, members of the project team should have knowledge of both terminological resources involved in the mapping project (source and target) and skills in their practical application. In many cases, this will require clinical input as well.

Principle 5: Members of the project team should have knowledge of both of the terminological resource and experience in their practical application.

6.6 Conventions and rules followed

As each terminological resource has a different structure, purpose, set of conventions, and rules for application, there should be a clear indication in the scenario about the extent to which these conventions and rules should be followed to ensure reproducibility, information integrity, and fit with the purpose of the map.

EXAMPLE A map from SNOMED CT to ICD-10-am requires the application of rules used for ICD-10-am coding, such as when Chronic asthmatic bronchitis NOS (J44.8) occurs with acute lower respiratory infection (code to J44.0). Where rules such as these are applied, consideration needs to be given to potential exemptions to the rules and the implications to the structure of the map.

Principle 6: Establish the extent to which the conventions and rules of each terminological resource will be followed.

6.7 Involve custodians and users

The custodians of both terminological resources (source and target), as well as those who will use the map, should participate in the mapping project to ensure that the result accurately reflects the meaning and usage of their resources in the real world system/s.

Skilled and experienced personnel are essential to the mapping process to provide the clinical and terminological resource knowledge to build and maintain the maps and assist with implementation. The degree to which they follow the conventions and rules of each resource will always depend on the purpose of each resource and a how a map from one to the other will be defined to fit with the purpose of the map.

EXAMPLE If the mapping project is a gap analysis being performed to determine content coverage with the aim of improving the target resource, then applying the conventions of each resource might not be relevant. On the other hand, if the mapping project is to enable reimbursement, then the conventions relevant to the fiscal requirements are to be followed and the map will provide options for the user to select the most appropriate target according to reimbursement policies to drive the transformations into the correct Diagnosis Related Group (DRG) assignment or to other grouping methodologies. Similarly, in a map from a terminology to a classification for automating classification assignment, it will be necessary to reflect the rules and conventions of the classification to ensure that the output complies with statistical and reporting standards.

Involving the custodians of each terminological resource makes sense from all perspectives: political, strategic, and operational. At a minimum, both custodians should participate in defining the basic purpose and parameters of the mapping project, reviewing and verifying the maps, developing the plan for testing and validation, and devising a cost effective strategy for building, maintaining, and enhancing the map over time.[17] Consider, too, that it will be the custodians who will have a good grasp of their users' requirements and will provide informed advice that will assist in developing a credible scenario.

Principle 7: Custodians of terminological resources should be involved in mapping projects

6.8 Open and documented map production processes

The documentation supporting the map development, decision making, and maintenance should describe in detail the automated and manual methods applied in the map. Common terms used to distinguish between automated and manual methods for mapping are "auto-matching" and "human mapping".

There are many techniques which can be used. This Technical Report highlights only some of the common activities, such as auto-matching. This information is provided as an example to support decision making and should not be considered the only process available.

6.8.1 Auto-matching

Auto-matching is a computational mapping task, undertaken using an algorithm. The matching algorithm can be based on lexical, semantic, hierarchical, structural, or other characteristics of the concepts, the situation, or the individual.

Auto-matching methods are best used to assist human mapping as this is a labour-intensive process. Matching algorithms automatically identify a list of candidate maps for human mappers to consider and validate, thus reducing the time required for searching. The most commonly employed auto-matching algorithms are lexically-based (i.e. they match words, rather than actual meaning and require humans to complete and confirm the process).

Debate continues within the mapping community about whether the process of mapping can ever be fully automated. Progress is being made with mapping tools and technology. The process will at least become semi-automated for algorithmic transformation and with human review required to verify the results of each individual map.[6]

Computer algorithms used for purposes of lexical matching can effectively reduce the amount of time spent searching and mapping and transform the mapping task into one of selection and validation. Algorithms often vary in sophistication and in the number and type of features they can compare and measure. Some examples are the following:

— inflected forms, the differences between or equivalence of words in different forms such as bacteria and bacterial, and perhaps, find, finding, and found;

— ability to perform term normalization which reduces each term to its morphological uninflected form, removes plurals, and standardizes tense and sentence case;

— ability to disregard words which have little influence on matching (commonly called “stop words” because they stop the algorithm from working properly;

EXAMPLE Pain in stomach will match with stomach pain when [in] and [the] are excluded.

— phonetic searching and conversions to recognize predictable misspelling of terms contained in content files;

— ability to recognize differences in word order (gestational diabetes as opposed to diabetes, gestational);

— ability to match on similar words contained within different terms (infection and infection control);

— ability to recognize acronyms as equivalent terms (urinary tract infection and UTI).

6.8.2 Human mapping

Human mapping is the use of human knowledge and skill to build maps between concepts and/or terms in different terminological resources. Each individual map record is built singly. The process requires examination of each and every concept and coding system. Informed judgements or decisions are made about the shared meaning of concepts. Electronic or computational tools are often used to support the mapping process but are not effective in determining any equivalence of meaning.

It should be noted that the use and meaning of the words “matching” and “mapping” vary considerably between authors and may be used interchangeably. The distinction made here is on the basis of the degree of meaning.

Principle 8: The automated and manual methods applied should be transparent and documented.

6.9 Describe the direction of the map

Mapping is performed for different purposes and the direction of the map should be made explicit in the supporting documentation accompanying the map.

The map relationships between target and source concepts are usually built for mapping in one direction only; that is they are a unidirectional map from the source to the target. Forward and backward mapping is performed for different purposes and so the direction of the map should be made explicit in the supporting documentation accompanying the map.

EXAMPLE Because all maps have a direction, they are usually described in relation to their source and target, such as an ICPC 2 to SNOMED CT mapping, ICPC 2 is the source, and SNOMED CT is the target. The map from ICPC 2 goes forward to SNOMED CT.

It is not possible, in general, to simply reverse a map (i.e. apply the rules in reverse) and use it in the opposite direction. Although there will be some links that will be valid in both directions, in general, an entirely different map should be generated for reversing the direction of the map, particularly when mapping between terminological resources of differing granularity.

For example, ICPC 2 is mapped to the ICD 10 to demonstrate the compatibility between the two classifications. This mapping is bidirectional and useful for conversion of data from each system,[2] though not in every case. An illustration of the complexity of a bidirectional mapping comes from the introduction to the ICPC-2 classification: the relationship between ICPC and ICD-10 is complex. There are some concepts in both which are not exactly represented in the other. However, for most concepts, in each classification, one or more corresponding concept in the other can be mapped. Because of these complexities, quite often, conversion of a concept from one classification to the other and then re-conversion back again will not necessarily lead back to the same original code because in each direction, there might be several codes to choose from.

A mapping between versions of a terminological resource is often maintained to be bidirectional to support historical tracking/data continuity purposes, to provide a way of interpreting the data using codes from either version of the classification. When a resource is updated, the forward map reflects the inactivated concepts in the superseded version linking to the new codes in the updated version. The backward map reflects the new codes in the updated version linking back to the inactivated concepts in the superseded version. Enabling backward compatibility is particularly useful for epidemiological or longitudinal studies.

A backward map may also be required when some users do not update to the latest version of a terminological resource or when the implementation timeframes of updates to linked terminological resources are staggered, as with the case of updates to the ICD and DRG classifications. The logic to DRG (Diagnosis Related Group) classifications is built from the ICD codes, thus, the systems are intrinsically linked. Any updates to the ICD impact the DRG logic and necessitate an update to the DRG classification and map algorithm. Backward map are also required for longitudinal analysis when researchers need to know how a condition was classified over time, and what, if any, changes were made to that terminological resource.

Principle 9: Every map should describe the direction of the map.

6.10 Clearly specified cardinality

Cardinality is often used as an indicator of equivalence, though a one-to-one cardinality does not necessarily mean that the terms are equivalent. One-to-one cardinality might indicate that the only applicable map candidates might be a narrow to broad match.

For example, if the concept in the source is the same (no less or more than) as the concept in the target, the cardinality is one-to-one (there is one member of the match in each set). If the concept in SNOMED CT is represented as one of a much larger group (say five concepts) in ICD, cardinality would be one-to-five (1:5)

EXAMPLE Cardinality of a map does not indicate semantic equivalence. A trivial function x — 2*Y _1 gives a cardinality of 1:1 (x has one 'answer') but the values are not equivalent.

To understand the uses to which a map can be safely put, it is essential to understand the degree to which the target represents the same meaning as that of the source (cardinality). The cardinality and equivalence of each map from source to target (and target to source, if bidirectional) should be clearly denoted in the documentation supporting the map.

If the cardinality is one-to-one (i.e. the meaning is the same in the target as it was in the source), the map should be safe for clinical use. If it is not clinical, meaning may be changed, subtly or significantly, which would represent a clinical risk in direct patient care.

Cardinality may also be described as the degree of aggregation of the map. The relationships are described in Table 3.

Table 3 — Cardinality of the map

Cardinality	Relationship
One-to-one	A single source concept is linked with a single target concept or term.
One-to-many	A single source concept is linked with multiple target concepts or terms.
Many-to-one	Multiple source concepts are linked with a single target concept or term.
Many-to-many	Multiple source concepts are linked with multiple target concepts or terms.

Where the cardinality of a map is other than one-to-one, there will be a need either for

— rules to support automated decision making between the alternative target concepts (for example, if male choose code A, if female choose code B) or

— manual intervention (in this case, the map provides a set of options for the user to choose from based upon the map).

Principle 10: Cardinality of each individual map should be clearly specified.

6.11 Explicit specification of loss or gain of meaning

Because terminological resources are purpose built and have different structures, they generally do not integrate seamlessly with each other. Concepts infrequently match on a one-to-one basis. More often, there is loss or gain of meaning through one-to-many or many-to-one relationships. Some concepts cannot be mapped because the concept only exists in either the source or target. Any loss or gain of meaning should be clearly identified and advice on how to manage relationships articulated to users and implementers in the supporting documentation accompanying the mapping so that they can determine the impact on their business needs.

EXAMPLE The cardinality of each map describes the relationship between the associated concepts or terms at both ends (source and target) of the relationship link.

Figure 1 is an example of a one-to-one map from between ICD 9 CM and ICD-10. The two concepts in different classifications have the same meaning though the codes used to express them are different.

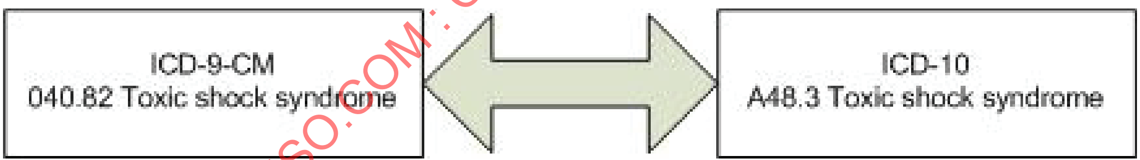

Figure 1 — Example of a one-to-one map

Figure 2 shows an example of a one-to-many map from a mapping between ICD 9 CM and ICD-10. In this instance, there is one code representing all locations of sarcoidosis while in ICD-10, there are different codes for these locations and combinations of locations.

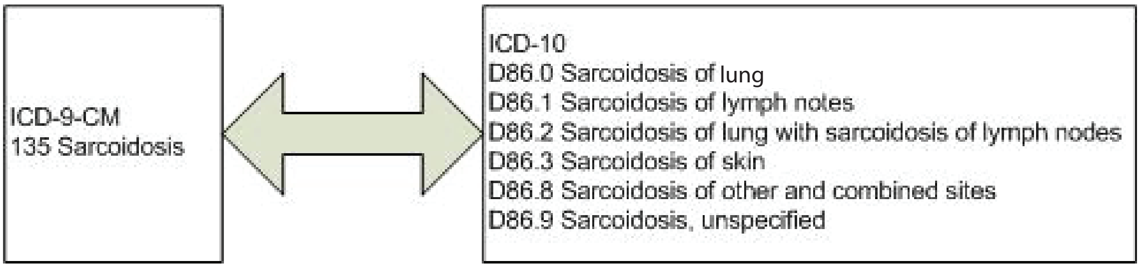

Figure 2 — Example of a one-to-many map

NOTE The maps depicted in Figure 1 and Figure 2 can also be reflected as backward maps. Figure 1 indicates both a forward and backward map with a one-to-one cardinality. With each map, there is no loss or gain of meaning. Figure 2 indicates a forward map with one-to-many cardinality and a backward map with a many-to-one cardinality. The forward map demonstrates a gain in meaning and the backward map demonstrates a loss of meaning. Both maps identify the increased specificity in ICD-10-am when compared to the ICD-9-CM.

Whether this loss or gain of meaning is important will always depend on the purpose of the map. For example, a map for the purpose of identifying the changes made to a parent classification and its country specific modification (for example, ICPC 2 PLUS to ICPC 2) may allow for less precision than a map between a terminology and a classification for the purpose of funding healthcare services.

Where a one-to-many map exists, consistent and appropriate interpretation requires a documented rules-based approach which specifies the approach taken to decision making to identify the closest match for the purpose of the map. The primary map (that which is closest in meaning or "best match") exists where there are other alternative maps available. The alternative options to the primary map (a selection of concepts is provided). These alternatives may be included as additions to the primary map (maps that should be combined to adequately reflect the meaning of the source concept). The primary map is preferred but might not always be able to be implemented. If one of the coding systems involved is a classification, then there should be some explanation about how the residual categories (the "other" and "unspecified" categories) are managed in the map.

There should always be a primary map that represents the best match. When mapping between terminological resources with concepts such as "not otherwise specified", a one-to-many mapping of the primary map will often link to a residual category as this represents the best match in meaning. Referring to Figure 2, 135 Sarcoidosis means D86.9 Sarcoidosis, unspecified. Occasionally, the mapping rule for the primary map in one-to-many mappings can be determined on clinical relevance. Referring to Figure 2 again, if the majority of cases of sarcoidosis appear in the lung, then it makes clinical sense to map 135 to D86.0 as the primary map when the purpose of the mapping is for population studies; most of the cases would be captured in these codes. Importantly, the rules applied to determine the primary map will always depend on the purpose of the mapping.

All alternative options should be represented in an order and the logic of that order explained in the supporting documentation. Which maps are selected by users will depend on the purpose of the mapping, the rule-based approach applied, and the purpose of their implementation. For example, referring to Figure 2, the primary map might have been mapped to the "unspecified" option of D86.9. However, a specialist skin clinic implementing this map might choose an alternative option of D86.3 as the code that best meets their business needs. Another option to represent alternative maps is to determine concordance based on clinical advice and frequency data and provide the percentage distribution of codes in the one-to-many map.

All additive options should also be represented in the order in which they should be applied.

The crucial point to make about applying a rules-based approach is that it serves two purposes:

a) it assists mapping personnel in making consistent decisions in determining the logic of the maps;

b) it assists users of the maps in making informed decisions about how to best link the codes in a way that meets their business requirements.[3]

Principle 11: Any loss or gain of meaning should be made explicit and risk assessed.

6.12 Demonstrate the degree of equivalence of the map

Maps or relationships between source and target are described as a measure of the degree of equivalence and the rating scale applied in determining equivalence should be described in the supporting documentation accompanying the mapping (Table 4).

Table 4 — Example of a rating scale to describe degree of equivalence

Rating	Meaning
1	Equivalence of meaning; lexical, as well as conceptual. For example, asthma and asthma; ovarian cyst and cyst of ovary.
2	Equivalence of meaning, but with synonymy. For example, ureteric calculus and ureteric stone; gall stones and cholelithiasis.
3	Source concept is broader and has a less specific meaning than the target concept/term. For example, obesity and morbid obesity; diabetes and diabetes mellitus type II.
4	Source concept is narrower and has a more specific meaning than the target concept/term. For example, feels ugly and self-image finding; acute renal failure syndrome secondary to dehydration and acute renal failure syndrome.
5	No map is possible. No concept was found in the target with some degree of equivalence (as measured by any of the other four ratings).

One illustration is the ICD-9-CM to the ICD-10-CM, tables of equivalence which have been developed to serve as a reference source. They help users navigate the complexity of equivalent meaning between two very different terminological resources. Concerning maps between terminology and classification systems, the meaning of the terminology should be understood in reference to the meaning of a classification description. The selected term in the hierarchy of SNOMED CT terminology should resemble the code and its description in the appropriate chapter of the ICD-10 system, taking into account all rules and conventions of that classification system.

Principle 12: All maps should demonstrate the degree of equivalence.

6.13 Explicit guidelines and heuristics applied in development and for implementation

Mapping begins with establishing the steps in the mapping process and then developing any guidelines and heuristics that support the scenario, the purpose, and implementation of the map. The documentation supporting mapping methodology and rationale are a core requirement to the mapping process. These are continually refined throughout the mapping process as previously unknown issues arise and are solved.

Each step in the mapping process and all guidelines and heuristics should be made explicit in the supporting documentation. All maps should meet the three requirements below, as well as the requirements to support the specific use case of the map:

a) enable all mapping personnel to interpret the definitions and create the maps in the same way.[6] Mapping is a subjective process because of the difference in skills, knowledge, and experience between mapping personnel. Establishing the mapping process, guidelines, and heuristics and applying them consistently is vital for reliable mapping;

b) assist users of the map in implementing it and making informed decisions about how to link the codes in a way that meets their business needs;[3]

c) enable all maps to be understandable and reproducible so that others not involved in the creation of the map are able to verify the map's accuracy by reproducing the map following the guidelines and heuristics.[16]

In some cases, there will be a need to follow the coding conventions or guidelines applied to each terminological resource and incorporate these rules into the mapping process to accurately identify the concepts and their degree of equivalence, but that will depend on the purpose of the mapping.

Principle 13: All mapping projects should make the guidelines and heuristics applied in developing and interpreting the maps explicit when implemented.

