

SNOMED Clinical Terms UK Edition UK concept History Substitutions Table

Published October 2018

A SNOMED CT data derivative

Reduces query execution false negatives occurring due to concept inactivation

Information and technology
for better health and care

Distribution:

Within biannual UK Clinical Extension of SNOMED CT.

Document Control:

The controlled copy of this document is maintained in the NHS Digital corporate network. Any copies of this document held outside that area, in whatever format (e.g. paper, email attachment), are considered to have passed out of control and should be checked for currency and validity.

Purpose of this Document

This document provides information relating to the History Substitution Table issued alongside the UK Edition of SNOMED Clinical Terms®.

Contents

1	Introduction	3
1.1	Document Scope	3
1.2	Product Scope	3
1.3	Audience	3
1.4	Product Status	3
1.5	Product History	4
1.6	Enquiries	4
2	Background	5
2.1	Concept inactivation in SNOMED CT	5
2.2	Effect of concept inactivation on reporting	6
2.3	SNOMED CT historical relationships	6
3	UK Substitutions Table	7
3.1	Scope: Current UK Edition	7
4	Anticipated use	8
5	Design	9
5.1	Periodic regeneration of the table	9
5.2	Source data	10
5.3	Cardinality: Multiple substitutes for one inactive concept	11
5.4	Content in the table derived from SNOMED CT Descriptions	11
5.5	Expected history patterns	14
5.6	History paths	15
5.7	Foreseeable types of error relating to the history path	16
5.8	Filters	16
5.9	TABLE SCHEMA	17
6	Release file format	23
7	Production methods	24
8	Quality management	25
8.1	Safety status	25
8.2	Responsibilities	26
9	Related documentation	26

1. Introduction

This document describes the SNOMED CT Concept History Substitutions Table ('Substitutions Table'), a computed derivative of the UK Clinical Edition of SNOMED CT. The product aims to facilitate correct processing of data involving SNOMED CT concept identifiers with 'inactive' status. The use cases for the product therefore overlap with those of the existing Query Table product; the Substitutions Table is one of the inputs to the computation used to produce the Query Table. The technical report '[SNOMED CT: Inactive Content](#)' provides further detail on the issue and use cases.

1.1 Document Scope

This document provides an outline of:

- the rationale for creating the derivative product
- the intended use cases

The document does not discuss clinical governance, safety or acceptance testing issues.

1.2 Product Scope

The Substitution Table product addresses concept inactivation issues only in relation to concepts found in the SNOMED Clinical Terms® International and UK Clinical Editions.

It does **not** currently cover similar issues relating to concepts from SNOMED Clinical Terms® UK Drug Edition.

1.3 Audience

Suppliers and clinical users of clinical information systems already based upon, or that process SNOMED CT, or are planning future systems with these capabilities.

Secondary users of data derived from SNOMED CT systems, including large data repositories, health networks, research units, research networks, CSUs and Commissioning Groups, Public Health Observatories

NHS programmes and organisations contracting or commissioning related activity involving either suppliers or users listed above.

1.4 Product Status

This document plus the table it describes constitute a product at the 'Supported Product' stage as defined within the IReS Terminology Product Development Lifecycle (link [here](#)) and is distributed by NHS Digital as the UK National Release Centre (UK NRC) for SNOMED CT.

'Supported Product' status means:

1. Both the release format specification of the product and the method of its content preparation shall remain fixed indefinitely unless a significant safety risk is identified that cannot be mitigated without changing them. Where changes are deemed necessary to improve a product then a formal consultation procedure will be undertaken which may

include some or all parts of the product development process and may include an option for parallel running (i.e. support for both existing and new specification).

2. UK NRC commits to continue to support, maintain and publish the product against that fixed specification indefinitely, subject to the considerations above or proper product termination

3. Quality assurance may be ongoing but the product is approved for deployment in live clinical systems, subject to standard safety assessment procedures associated with deployment of any product into a live environment

4. The commitment to release against a stable specification does not preclude continued parallel evolution of the specification and consequent development of improved variants which may or may not be considered as new products

1.5 Product History

Approval to prepare and release the SNOMED CT Concept History Substitutions Table as a product was obtained from the UK Edition Committee on 6th April 2011. The first release occurred on 1st October 2011, as a Technology Preview; the Edition Committee approved it becoming a 'Supported Product' in August 2018.

1.6 Enquiries

Content error reports, or requests for clarification or additional functionality, should be made to information.standards@nhs.net with 'SNOMED CT Substitutions Table' in the subject line.

2. Background

2.1 Concept inactivation in SNOMED CT

SNOMED CT, like Clinical Terms Version 3, is a dynamic terminology: content can be added, edited or made inactive where necessary. Each release of SNOMED CT correspondingly partitions all concepts within into either '**Active**' or '**Inactive**'. In RF2, this initial binary partitioning is achieved through the *active* field in the *sct2_Concept* table; RF1 releases employ a *ConceptStatus* metadata value on each concept within the *sct1_Concept* table.

'**Active**' currently further subdivides into 'current' or 'pending move', whilst '**Inactive**' has seven subdivisions: 'retired (no reason)', 'duplicate', 'erroneous', 'ambiguous', 'outdated', 'limited', and 'moved'. In RF2, this further partitioning is represented through the membership of the 900000000000489007|Concept inactivation indicator attribute value reference set (foundation metadata concept)|. In RF1, different *ConceptStatus* values appear in the *sct1_Concept* table; whether the concept is fundamentally 'Active' or 'Inactive' is never explicitly recorded in RF1 but rather only implied through this *ConceptStatus* value.

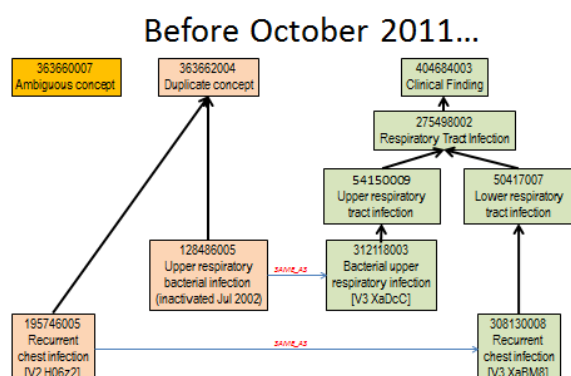
Refer to the SNOMED CT Technical Implementation Guide¹ for detail on concept status.

In this document concepts are referred to as either *active* or *inactive*.

NB although commonly synonymous with 'inactive', 'retired' is treated in this document as a reserved word for use only to refer to the 'retired (no reason)' concept status.

Over time and between releases of SNOMED CT, therefore, individual concepts may change their status: an active concept may become inactive (and very rarely also vice versa). A concept that was previously 'active-current' may first become 'active-pending move' and later 'inactive-moved'. Other status transitions are also possible.

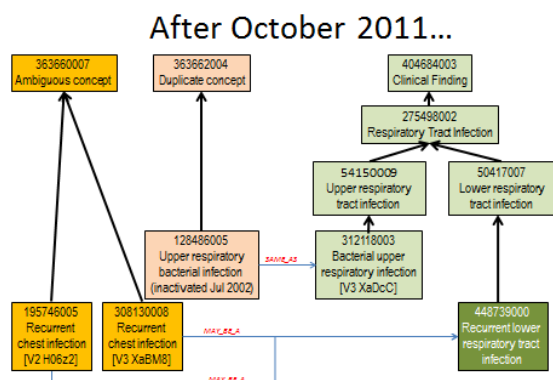
Whenever a concept moves from 'active' to *any* flavour of 'inactive' it also loses all its modelled attributes on inactivation, including any IS_A relationships that previously positioned the concept correctly within SNOMED CT's primary taxonomic hierarchy; in RF1 all inactive concepts exist instead in a separate, very flat, taxonomic tree. In RF2, they have no taxonomic ancestors or descendants at all (thus, similar to Retired concepts in CTV3).



For example, until October 2011, [308130008|Recurrent chest infection](#) was an active concept in the main taxonomy, classified below [275498002|Respiratory tract infection](#) as shown in Figure 1. Note that [195746005|Recurrent chest infection](#) also existed, but was already inactive (being a *duplicate* of the active [308130008](#)) and, accordingly, was **not** classified as a descendent of [275498002|Respiratory tract infection](#)

¹ Publically available from the SNOMED International Document Library
<https://confluence.ihtsdotools.org/display/DOC/Release+Documents>

Figure 1: Position of 308130008 prior to October 2011



From October 2011, 308130008|Recurrent chest infection| also became inactive: ‘chest’ often means the entire thorax whilst, in use, ‘chest infection’ *usually* means more specifically the lungs. Both 308130008 and 195746005 are therefore now (re-)classified as ‘ambiguous’, and moved outside the main taxonomy: neither is a subtype of 275498002|Respiratory tract infection|.

Figure 2 : Position of 308130008 after October 2011

2.2 Effect of concept inactivation on reporting

The taxonomic relocation of inactive content has significant consequences for any concept retrieval activity that is reliant on the taxonomy, such as decision support and population reporting:

Where a SNOMED CT system has data recorded using codes that have since become inactive, data processing will behave predictably but inappropriately in most retrieval or reporting contexts *unless* the SNOMED CT history mechanism is used appropriately. Query specifications that calculate their data extraction codeset at runtime, by reference to SNOMED CT’s taxonomy alone (the black arrows in Figures 1 and 2) will - in most cases - no longer return appropriate data items once the relevant codes have become inactive.

Clinically, this will manifest as new ‘false negatives’: patients are not returned by a particular query when the clinical expectation, based on the coded data, is that they should be. In the example of Figures 1 and 2, patients originally recorded prior to October 2011 as suffering from 308130008|Recurrent chest infection| will no longer be returned by a search for all patients suffering from any kind of 275498002|Respiratory tract infection| if that query is run against any version of SNOMED CT released after October 2011.

To alleviate this problem, NHS Digital publishes two products: a ‘table of active substitutions’ within the release of the UK Clinical Edition of SNOMED CT, and the SNOMED CT Query Table (currently as a standalone TRUD Subpack).

2.3 SNOMED CT historical relationships

In addition to demonstrating how inactive concepts move out of the main taxonomy, Figure 1 and Figure 2 (above) also show that inactive concepts may be associated with alternative active concepts, within the main taxonomy, by means of non-taxonomic ‘historical relationships/associations’. In the figures, these associations are represented by the blue lines. In Figure 1, for example, the inactive 195746005 has a SAME_AS association² with 308130008, indicating that the two concepts are semantically equivalent.

² In RF1, these associations were represented as flavours of a special class of relationship within the sct1_relationships table; in RF2 the same associations are represented instead by means of those Historical

3 UK Substitutions Table

The Substitutions Table is a pre-computed derivative of these SNOMED CT historical associations. For each inactive concept in the SNOMED CT release from which it derives, the Substitutions Table lists one or more active concepts to which it has been associated via SNOMED CT historical relationships/associations.

[Where an inactive concept exists but for which no active substitute is derivable through the History Mechanism, then the substitutions table relists the inactive concept as its own substitute. The table also includes the Fully Specified Names for the inactive concept and each of its candidate active substitutes. See section 3 for further details.]

The substitutions table provides other metadata about the inactive concept and its substitute(s):

- Identifying the substitution candidate concept and its current status (four attributes)
- Indicating the certainty of the substitution (three attributes)
- Further analysing the semantic characteristics of both concepts (seven attributes)

In theory, the generation of this substitution table could be achieved by any enterprise or individual providing they had access to all the necessary information and standards. However, the NHS Digital provides this table as a precomputed result so that only one enterprise must invest effort in its design preparation and distribution, whilst seeking to assess and improve its usefulness to stakeholders.

Given the design set out in section 3 and the Editorial Policy for the UK Edition of SNOMED CT (all concepts ever released must appear in all subsequent releases as either active or inactive) and despite the filters outlined in section 3.8 the result is that:

For any concept ever released as part of the UK Edition there will either be an active concept in the most recent release or there will be one or more entries in the substitution table.

In the rare case where an active concept becomes first inactive and then at a later date active again, that concept will only appear in editions of the substitutions tables that accompany a release of the UK Edition of SNOMED CT within which that concept is inactive.

SNOMED International documentation on the standard for recording the attested history for a concept is set out in the SNOMED CT Technical Implementation Guide . See also section 7 of this document.

3.1 Scope: Current UK Edition

The substitutions table is specific to inactive concepts; it does not offer substitutions for other SNOMED CT components such as SNOMED CT descriptions.

The table is specific to the UK Edition and to a particular Release of the UK Edition: the table only includes substitutions for those concepts published as inactive in the related UK Edition

Association cRefsets listed below 900000000000522004|Historical association reference set (foundation metadata concept)| in the RF2 metadata hierarchy.

release, whether as part of the International Edition, or the biannual releases of the UK Clinical and Drug Extensions.

The table is produced and distributed only with each ‘UK Combined Release’³ and not with each release of the UK Drug Extension. Concepts that become inactivated within a monthly release of the UK Drug Extension will not appear in the table until the next biannual Combined Release.

This document anticipates and mitigates the effects of the UK Edition being released in Release Format 2⁴ (RF2), and the refactoring of the algorithm this entailed to derive the substitutions table from slightly different source material.

Limited status concepts used in the UK

Prior to the April 2010 release of the UK Edition of SNOMED CT, the set of active concepts included those with ‘limited’ status (ConceptStatus=6). Since April 2010, SNOMED International has reclassified limited status as a subtype of inactive, with the effect that (like all other inactive content) all existing limited status content is now moved outside the main SNOMED CT taxonomy.

UK implementers are believed to have made substantial prior and persistent use of limited status concepts. In view of this, the UK Edition from April 2010 onward implemented a workaround: For each newly inactivated limited status (6) concept, a new corresponding active concept was created back within the main hierarchy but using a different ConceptID and with current status (ConceptStatus=0)⁵. The substitution table therefore provides, for each such limited status concept, its corresponding current workaround substitute.

4 Anticipated use

The anticipated and intended use of the substitution table is to allow implementers of the UK Edition to identify, for each of their design artefacts which use the SNOMED CT UK Edition, the possible updates needed to these such that they use only active concepts.

Relevant design artefacts include but are not limited to:

- Bindings between data items and individual terminology concepts (forms designs, message specifications etc.)
- Subsets of Concepts or Terms, however used
- Query definitions
- Knowledge tags
- Decision support rules

³ The UK Combined Release comprises the International Core (from SNOMED International) along with the UK Clinical Extension, but not the UK Drug Extension

⁴ SNOMED CT Release Format 2 is described in SNOMED CT Technical Implementation Guide and other SNOMED International documents

⁵ As well as Status = 0 (current) there could also be cases of Status 11 i.e. pending move.

Where the substitution table indicates that the authors of SNOMED CT have, upon concept inactivation, selected only one substitute concept, then most implementers are anticipated to adopt that substitute.

Where the substitution table includes more than one entry for an inactive concept then the additional metadata derived from the descriptions may help determine which alternative (if any) to use.

One potential mode of use could permit many design artefacts to remain physically unchanged, and for the substitutions table to be introduced into the runtime architecture in order to achieve a *logical* update. For example, where many data entry templates have previously been authored but now contain references to inactive concepts, the operation to write-commit new data entry items to an EPR repository could be passed through the substitutions table so that the inactive concepts were substituted with their active counterparts (though, if subsequent querying architectures used the Query Table, this logical substitution operation would be unnecessary at write time, since the Query Table would function to achieve the same net 'dynamic update' effect but at read time).

Although the Substitutions and Query Table artefacts insulate design artefacts from some changes to SNOMED content, they can not protect from all. Re-evaluation of some artefact designs may still be a necessity each time one or more substitutions are made, or the need may be less stringent. The owner of any design artefact must however determine whether and when any change is still required, subject to SNOMED CT license conditions. NHS Digital takes no responsibility for judging how the substitutions are to be used or for the resultant revised design artefacts (such as those in the list of artefact types above).

As the substitutions table is regenerated and released with each release of the UK Edition, each release of the table is specific to a release of the UK Edition. Implementers should not use a version of the substitutions table together with earlier or later versions of the UK Edition.

5 Design

The design of the table is outlined in this section. Where only one active substitute candidate exists for an inactive concept, only one row appears in the substitutions table (NB this includes ambiguous concepts where only one alternative was specified by SNOMED CT authors). Where more than one candidate exists, the table has one row for each candidate (see section 3.3).

5.1 Periodic regeneration of the table

The table is regenerated for each scheduled biannual combined release based on the latest available relationships and concept status included in the UK Edition. Some substitutions will remain the same in perpetuity, others may change once or more with successive releases.

For any 'out-of-cycle' release of the UK Edition a decision whether to regenerate and publish the table will be made by the NHS Digital.

5.1.1 Relationship to SNOMED International release of the international core

The SNOMED International core release precedes the UK Edition release by three months; in the interval between the release of the International Edition and the UK Edition there will typically be history relationships in the international core which will therefore not also become represented in the substitution table until the first subsequent release of the UK Edition. For implementers who use *only* the UK Edition this time offset has limited or no material consequence.

Exceptionally, where critical errors exist in the International Edition and have been corrected, the NHS Digital would consider making an out-of-cycle release of the UK Edition – including potentially of the History Substitutions table - so as to more immediately propagate such corrections.

5.2 Source data

The substitution table is prepared solely by use of the attested history relationships or associations held in SNOMED CT⁶ plus, where required, the membership of the UK subset 29511000000134 (Former Limited-Status Concepts Imported from Core). The Concepts and Descriptions tables are then used to retrieve attribute values and to compute flag values included in the table.

The table does not embellish or contradict the UK Edition of SNOMED CT (apart from the minor exception set out below in section 3.2.1) Consequently the table is a reflection of the decisions of the authors of SNOMED CT, both the UK authors and those who work directly on SNOMED International's International Edition.

This substitution table has been prepared using expertise and experience of the SNOMED CT History Mechanism including the movement of concepts between UK Extension and the international Core, and the UK-specific workaround to the mass inactivation of all limited status concepts that occurred from the April 2010 release.

The NHS Digital has conformed to the SNOMED CT standard in the production of the table so far as is possible, but the documented standard does not deal explicitly with every variant of possible movement of content between extensions in combination with each permitted editorial option. NHS Digital has therefore interpreted the standard in order to produce the resultant substitution table.

5.2.1 Time-offset for releases between UK Drug Extension and the UK combined release

There can be a time-offset between the release of the 'UK Combined Release'⁷ and the release of the UK Drug Extension. The table is built with substitutions for the whole of the UK Combined Release and the UK Drug Extension.

NB The UK Drug Extension is released more frequently than the UK Combined Release (the substitution table being released only as part of the latter).

⁶ Relationships with CharacteristicType=2 under RF1, or the equivalent encoding within the refsets found below 900000000000522004[Historical association reference set (foundation metadata concept)] in the RF2 metadata hierarchy

⁷ The UK Combined Release comprises the International Core (from SNOMED International) along with the UK Clinical Extension, but not the UK Drug Extension

The released content spanning both the UK Combined Release and the UK Drug Extension will, at the time of the release of the UK Combined Release, always have one or more entries in the substitutions table for all concepts in an inactive state.

5.3 Cardinality: Multiple substitutes for one inactive concept

While many substitutions indicate a single active replacement, this is not always the case.

The authors of SNOMED CT make existing concepts inactive for a variety of reasons, including perceived ambiguity. When inactivating a concept as ambiguous, authors will *usually* point from the inactive concept to two or more candidate active substitutes⁸. In this situation each candidate substitute concept will be unambiguous in its own right, but which is more likely to be the preferred substitute will depend on the particular original context-of-use of the now inactive concept. The substitutions table does not suggest any default or ranked preference as to which of the candidate concept substitutions to use. In some data retrieval settings, especially those where false negatives are more serious than false positives, it may be appropriate to return the original concept as though it had all possible interpretations.

SNOMED CT implementers must make their own judgements on which candidate substitute concepts to use where more than one is listed, based on their knowledge of the original exact context-of-use.

5.4 Content in the table derived from SNOMED CT Descriptions

The substitutions table includes content derived from SNOMED CT Descriptions associated with the inactive concept and its candidate substitutes. This information is included partly to increase the direct human readability of the table, and partly to support typical workflows for validating suggested substitutes or choosing between multiple candidates.

5.4.1 Source of the Description data

For each entry in the table the NEWCONCEPT has the Description of a current Fully Specified Name [FSN] taken from the Descriptions table.

For the OLDCONCEPT the FSN Description is taken from the Concept table and is augmented by attribute data to indicate whether there has ever been more than one FSN for that OLDCONCEPTID.

The rationale for this approach is that where there are multiple FSNs previously existing: that the one in the Concept Table is more likely to have a semantic tag than any previously inactivated FSNs which are held in the Descriptions Table.

5.4.2 Flags relating to Terms, Semantic Tags, and Top Level Hierarchies

Data about the respective Fully Specified Names [FSNs] of the old and new concept is included to help users judge which candidate substitution is (are) appropriate to use, by

⁸ there are many ambiguous concepts with only one substitute designated by the authors

direct inspection of the data in the table and so including settings where a full back-end SNOMED terminology service is not available.

Two flags are included which relate to the Fully Specified Name. They indicate respectively:

- whether the term of the OLDCONCEPT and the NEWCONCEPT are identical (the term string excluding the semantic tag)
- whether the semantic tag is identical

A further flag relates to the Top Level Hierarchy [TLH] of each of OLDCONCEPT and the NEWCONCEPT. In this design, the TLH is deduced from their respective Semantic Tags of the FSNs; not by computation of IS-A relationships in the prior release files (see below). Whilst this approach is not as rigorous as one computed via relationships, it is adequate for the purpose of this table.

An example of the move of a concept between TLHs would be a 'finding' being remodelled as a 'situation with explicit context'.

5.4.2.1 Semantic Tags which were changed en masse

Since SNOMED CT was first published, many concepts have changed the Semantic Tag assigned to them in their Fully Specified Name. In many cases, this change signals a significant 'taxonomic category change' for the concept and thus a significant shift in its meaning. However, *some* patterns of Semantic Tag change exist that were systematically applied and where the change should not be interpreted as indicating a change of meaning of the concept.

The following are the predominant examples of such 'no change' patterns:

- no tag > (some particular tag)
- (context-dependent category) > (situation)
- (finding) > (situation)

It remains for the user of the table to consider these specific cases where the tags are not identical.

Further information about semantic tags is included in Appendix C

5.4.3 Valid Combinations of flags relating to FSNs

The set of valid combinations of flag values for the Descriptions and Top Level Hierarchies are shown in Table 1 below, along with an interpretation of what each combination signifies.

Table 1: Valid combinations of flag value for FSN Descriptions

Flags			
TLHIDENTICALFLAG	FSNTAGLESSIDENTICALFLAG	FSNTAGIDENTICALFLAG	Interpretation
1	0	1	The terms are not the same but are under the same top level hierarchy, with the same semantic tag
1	0	0	The terms and semantic tags are not the same however they are from the same top level hierarchy
1	1	0	The term is the same and from the same top level hierarchy, but with different semantic tags
1	1	1	The terms and tags are the same and from the same top level hierarchy This will apply to all cases where there is no active substitution identified, so the record points back to the same concept
0	0	0	The terms and tags are not the same and are not from the same top level hierarchy
0	1	0	The terms are the same but from a different top level hierarchy NB There are cases where this does not indicate any change in meaning but simply a correction to the modelling
0	any	1	Invalid

5.4.4 Missing Semantic Tags

There may be cases where the SNOMED CT FSN of the OLDCONCEPT either has no Semantic Tag, or where the tag fails to be machine identified by the algorithm used in the production of this table. In either case the flag TagIdenticalFlag may wrongly indicate a non-identical tag. This is deemed a low risk in the context of the intended use for the table.

5.4.5 Flag for multiple FSNs for one inactive concept

There are inactive concepts which have had more than one FSN through time. The design does not set out to exhaustively represent all the past and current FSN combinations for any pair of concepts in a record within this table.

Instead an attribute 'OLDCONCEPTFSNTAGCOUNT' indicates the number of distinct Semantic Tags for a given inactive concept. This treats the absence of a semantic tag on a FSN as a countable value. An inactive concept with two FSNs, one of which has no Semantic Tag will have a value of 2 for this attribute.

5.5 Expected history patterns

Common patterns

The history patterns found for inactive concepts fall into the following general patterns:

1. An inactive concept has no historical associations declared for it; such inactivation typically occurs when the meaning is impossible to determine, or when the meaning should not be represented in SNOMED CT at all.

The substitution table will still contain a row for all such concepts, but NEWCONCEPTID will be equal to OLDCONCEPTID

2. An inactive concept which has exactly one historical association declared for it, pointing at some other concept (though not necessarily to an active concept).
3. An inactive ambiguous concept possesses several historical associations pointing at several different concepts.

Where historical associations do exist, there is the possibility that any concept pointed at is in another extension, where it is itself inactive. Consequently, to find its substitutes any algorithm to trace the historical associations must be recursive. Under certain circumstances it may also be required to suppress certain substitutions.

'Former' limited status concepts

From January 2010, and following a consultation with its international user community, SNOMED International determined that all concepts with status 'Limited' (i.e. status 6) should be classified as a type of inactive where formerly they had been a type of active concept. A set of policies relating to this change were introduced from April 2010, primarily intended to defer the impact within the UK realm. These policies lead to more a complex history pattern for limited status concepts.

A concept of 'limited' status (status 6) may have none, or one, or many candidate substitutes referenced by historical associations and these can be any combination of active and inactive. A selection of 'Former' limited status concepts were at one time included in a subset which was then used to give precedence to which candidates should be included in the substitution table. The expected patterns are therefore:

4. A limited status concept which has – in addition to any candidate substitutions to other inactive concepts - exactly one candidate that is both current and also included in the UK former limited status concepts subset (UKFLSCS)

5. A limited status concept which has - in addition to any candidate substitutions to other inactive concepts - exactly one candidate that is current, but none that are both current and also included in the UK former limited status concepts subset (UKFLSCS)
6. A limited status concept which only has candidate substitutions to other inactive concepts in UKFLSCS

In each of cases (2,3,4,5 and 6) above there is the possibility that any concept pointed at is in another extension and is itself inactive.

5.5.1 Patterns where the history points at one or more active concepts

Where there are inactive concepts for which one or more traversals of the historical associations leads to an active concept then these are included in the substitution table, but none where the substitution is inactive.

5.5.2 Patterns where the history points only at inactive concept(s)

Where there are inactive concepts for which all possible recursions over the historical associations still point only to other inactive concepts, then the entry (or entries) in the substitution table will point to these inactive substitution concepts.

5.5.3 Patterns where the history points to neither active nor inactive concept(s)

Where there are concepts for which no historical associations point to either active or inactive concepts, then the entry in the substitution table points back to itself. This should ensure that each and every inactive concept in the release has one or more entry in the substitution table.

5.6 History paths

A 'history path' results from recursively following historical associations between two or more concepts.

For each row, the substitution table includes a record of a history path between the two concepts. This information is provided without any specific utility in mind but in anticipation that it could be valuable in audit processes.

As the set of historical associations in SNOMED CT can form a network, there can be multiple history paths between one 'OLDCONCEPTID' and the listed 'NEWCONCEPTID'.

For each pairing of OLDCONCEPTID and NEWCONCEPTID just one history path is included in the table, this does not imply that other paths do not exist. Precedence as to which path is included in the row is given if a path does include any concept with ambiguous status. This retains the conservative nature of the substitution table: showing all cases where ambiguity has been encountered at all.

5.6.1 Traversal of ambiguous concepts within a history path

Any particular history path may include a concept which has status 'ambiguous' and this may inform the user of the substitution table whether that substitution is appropriate for them to

use in any particular context of use. The table includes the 'ISAMBIGUOUS' flag which indicates whether one or more ambiguous concepts were included in the particular history path included in the record.

Meaning of the ISAMBIGUOUS flag:

The ISAMBIGUOUS flag is set:

- When the history path as recorded in the table includes one or more ambiguous concept(s) whether this is the OLDCONCEPTID, the NEWCONCEPTID, or one of the concepts traversed in the history path to get between these
- Whenever there is actually more than one distinct substitute listed in the table, regardless of whether any history path involved an ambiguous concept or traversed a MAY_BE history relationship.
- When other data suggests that there should be more than one distinct substitute in the table (even if only one is actually listed e.g. the history path involves an ambiguous concept but this has only one MAY_BE pointer to a substitute).

5.7 Foreseeable types of error relating to the history path

The foreseeable types of error are noted, these should be dealt with in the ways set out in section 6.

5.7.1 Unexpected history patterns, errors and omissions in the history assertions

In a very small number of cases unexpected patterns have been found in the history record; UK NRC request that these are referred to the Information Standards Service Desk⁹ for their consideration. Possible actions would include the alteration of the history record for the offending concepts by the authors of SNOMED CT or an adjustment to the processing algorithm used in the production of the substitution table .

5.7.2 Errors or ambiguities in the semantics of the history relationships

Any questions about either errors or ambiguities in the semantics of the history mechanism should be raised via the Information Standards Service Desk which could then lead to it being progressed via NHS Digital as the UK NRC to SNOMED International, but recognising that the substitution table is an additional contribution to the toolkit for implementers, and that the semantics of the history mechanism are long established.

5.8 Filters

In special cases, the substitution table will contain only a filtered subset of all the available substitutions that can be determined from the SNOMED CT data. These special cases are detailed here:-

⁹ information.standards@nhs.net

For any concept with more than one substitute (whether active or inactive) and where at least one is also in the Former Limited Status Subset (Subset ID 29511000000134), then only those substitutes that are also in that subset are listed (whether active or inactive).

NB The result can differ if a filter is applied before, or after other parts of the processing algorithm.

The filters applied relate only to the processing of concepts, none are driven directly by the Descriptions.

5.9 Table Schema

The format of the table is a simple Tab delimited file. This is the format prepared for initial release as Technology Preview (May 2011).

The attribute sequence in the table is

OLDCONCEPTID
OLDCONCEPTSTATUS
NEWCONCEPTID
NEWCONCEPTSTATUS
PATH
ISAMBIGUOUS
ITERATIONS
OLDCONCEPTFSN
OLDCONCEPTFSNTAGCOUNT
NEWCONCEPTFSN
NEWCONCEPTFSNSTATUS
TLHIDENTICALFLAG
FSNTAGLESSIDENTICALFLAG
FSNTAGIDENTICALFLAG

5.9.1 OLDCONCEPTSTATUS

Content

The status of the old concept

Format

VARCHAR(18) or BIGINT(20)

Permissible Values

Either:

1 – 11 integer values exclusively in line with SNOMED CT Technical Reference Guide and Release Format 1 (RF1)

Or: (to future proof the attribute specification)

Upon adoption of RF2, the metadata concept identifier for the status e.g. 900000000000486000 for 'limited component'

Notes

Refer to SNOMED CT Technical Reference Guide

5.9.2 NEWCONCEPTID**Content**

SNOMED CT ConceptID

Format

VARCHAR(18) or BIGINT(20)

Permissible Values

Same as for *OLDCONCEPTID*

5.9.3 NEWCONCEPTSTATUS**Content**

The status of the new active concept

Format

VARCHAR(18) or BIGINT(20)

Permissible Values

See *OLDCONCEPTSTATUS*

5.9.4 PATH**Content**

The set of historical associations navigated to get from the *OLDCONCEPTID* concept to *NEWCONCEPTID*

Format

VARCHAR(255)

Notes

THIS column contains a string which states each of the concepts which are in the history path between the inactive concept and the target substitute, and for each of these it also indicates the status of the concept at the time of table generation

Syntax

- Each successive concept is preceded by '>'
- Following '>' is the ConceptID of the concept which is included in a multi-step path between OLDCONCEPTID and NEWCONCEPTID
- Immediately following the ConceptID is a ':' separator
- Following the separator is the concept status value

This pattern is repeated for each intermediate concept in the specific history path without limit apart from field length.

There is a theoretical worst case maximum of eleven concepts in the path (based on 22 chars i.e. max ConceptID length of 18 char for each Path element plus 4 chars each for status and separators) – though to date no PATH longer than 5 items has been encountered.

Note: after adoption of RF2, concept status reported in the PATH value *remains* encoded using RF1 'magic values' (0...11)

Sample Data

>641171000000100:6>681291000000101:6

>402211000000102:4

>424271000000109:6>402581000000106:6>407731000000100:6>432211000000107:6>449171000000101:4

>212946001:6

>467211000000103:4

>471561000000100:2

>212991005:6>157747005:6>212991005:6

>444141000000106:2

5.9.5 ISAMBIGUOUS

Content

nAry flag (though currently binary) set if one or more history path used to arrive at the target indicates ambiguity (as set out in Appendix B)

Format

CHAR(1) or TINYINT (not BOOL; later versions may be nAry)

Permissible Values

0 (no ambiguity encountered)

1 (ambiguity encountered)

Notes

This flag indicates if any from the whole documented range of ambiguities has been identified

5.9.6 ITERATIONS

Content

TINYINT(4)

Format

String (max 2 char)

Permissible Values

-1|0|1|2|...

Notes

Indicates the number of links that were recursively followed beyond the single initial history relationship in a history path between the old and replacement concepts.

The value range is from -1 upwards with a practical limit but not theoretical upper limit.

The -1 value indicates that no replacement was found which met the inclusion criteria, consequently the non-active *OLDCONCEPTID* is mapped to itself as the *NEWCONCEPTID*

0 Indicates that just one relationship was traversed to get between the *OLDCONCEPTID* and the *NEWCONCEPTID* (ie there were no iterations of the relationship traversal)

1 indicates that one iteration was performed so one intermediate ConceptID is included in the PATH attribute. Values higher than one further required recursions.

Generally for all values except -1 this value indicates how many conceptIDs should be found in the history path between *OLDCONCEPTID* and *NEWCONCEPTID*.

5.9.7 OLDCONCEPTFSN

Content

The Fully Specified Name associated with the *OLDCONCEPT* as recorded in the current release

Format

VARCHAR (255) CHARACTER SET utf8 COLLATE utf8_bin

Permissible Values

String constructed from the SNOMED CT character set

Notes

The FSN for the OLDCONCEPT will be taken from the Concept Table

5.9.8 OLDCONCEPTFSNTAGCOUNT

Content

Count of the number of distinct semantic tags found on any FSN for the OLDCONCEPT

Format

TINYINT

Permissible Values

≥ 1

Notes

This is calculated from the Descriptions Table. Note that absence of a tag at all is to be counted.

5.9.9 NEWCONCEPTFSN

Content

The Fully Specified Name associated with the NEWCONCEPT as recorded in the current release

Format

VARCHAR (255) CHARACTER SET utf8 COLLATE utf8_bin

Permissible Values

String constructed from the SNOMED CT character set

Notes

Where the NEWCONCEPT is an active concept this FSN will be taken from the Descriptions Table where the DescriptionType = 3 and the DescriptionStatus indicates it is the active FSN for that Concept

Where the NEWCONCEPT is inactive then the FSN will be taken from the Concept Table

5.9.10 NEWCONCEPTFSNSTATUS

Content

The status of the FSN of the new concept

Format

VARCHAR(18) or BIGINT(20)

Permissible Values

For any active NEWCONCEPT any one of

0 – Current
6 – Limited
11 - Pending move

For any inactive NEWCONCEPT

= any of the statuses other than those listed above

Upon adoption of RF2, the concept identifier for the appropriate description status e.g. 900000000000494007 for 'Inappropriate component'.

5.9.11 TLHIDENTICALFLAG

Content

Flag which is set if the Top Level Hierarchy to which both the OLDCONCEPT and the NEWCONCEPT is identical

Format

BOOL

Permissible Values

0 – Top Level Hierarchies are not identical
1 – Top Level Hierarchies are identical

Notes

The Top Level Hierarchy may be identical even If the tag on the FSN is not.

This flag will identify moves between Top Level Hierarchies; these are likely to entail more radical impacts than if the tag alone has changed.

NB Some TLH moves do not indicate radical change of meaning such as finding > situation with explicit context

5.9.12 FSNTAGLESSIDENTICALFLAG

Content

Flag which is set if the FSN of the OLDCONCEPT and of the NEWCONCEPT are identical when the Semantic Tag has been removed

Format

BOOL

Permissible Values

0 - Terms without tags are not identical
1 – Terms without tags are identical

Notes

This value cannot be set to 0 if the flag `FSN_IdenticalFlag` has value =1

As this is generated with reliance on string manipulation there may be a small residue of cases where the stripping off of the tag is not successfully executed. Consequently the data in this attribute cannot be assumed to be 100% exact, yet is still useful.

5.9.13 FSNTAGIDENTICALFLAG

Content

Flag which is set if the FSN of the OLDCONCEPT and of the NEWCONCEPT have the identical Semantic Tag

Format

BOOL

Permissible Values

0 - Tags from the two Descriptions are not identical

1 – Tags from the two Descriptions are identical

Notes

As this is generated with reliance on string manipulation there may be a small residue of cases where the stripping off of the tag is not successfully executed. Consequently the data in this attribute cannot be assumed to be 100% exact, yet is still useful.

5.9.14 Table length estimates

The initial Technical Preview version of the table contained >200,000 rows of which >30,000 were flagged as ISAMBIGUOUS.

6 Release file format

The table and this document are released as a file which will follow prevalent file naming conventions for the UK NRC:

`xres_HistorySubstitutionTable_Concepts_GB1000000_YYYYMMDD.txt`
if derived from RF2 release data

`doc_UKSnomedCTHistorySubstitutionTable_Current-en-GB_GB1000000_YYYYMMDD.pdf`

7 Production methods

The exact methods of production for the substitution table are not set out in the body of this note but Appendix B contains detail of the methods of detecting and flagging what are deemed to be concepts which are ambiguous.

The general production method consists of:

Treating all concepts as inactive if they currently have a status other than any of:

- Current (coded as status 0)
- Pending Move (coded as status 11)

Concepts with a status “Limited” were only considered active until the January 2010 SNOMED CT International Release. In the International Release these are now considered to be inactive. However, for the purposes of the UK it is valuable to provide entries in the substitution for these Limited status concepts.

Preparation of the substitution table relies on recursively following the historical associations added into the terminology for concepts which are inactivated *either* by SNOMED International as part of the International Core SNOMED CT *or* by NHS Digital as part of the UK Edition.

This leads to a table which contains:

- For each inactive concept which has one or more active substitutes: All active substitutes but none which are inactive AND for any concept with more than one substitute, none which are NOT in the Former Limited Status Subset (Subset ID 29511000000134)
- For an inactive concept with no substitutes whether active or otherwise: An entry pointing to this concept itself.
- The substitution table contains no replacement concepts which are themselves inactive in the UK Edition, except where there is no replacement which is active.

Processing which generates the table will continue to use all the SNOMED CT data. This permits concept reinstatement to an active status on the rare occasion it is found to be necessary. This guards against effects of a concept reinstatement; from one of the inactive statuses to one of the active ones. Reinstatement is legitimate within the SNOMED CT standard.

Under RF1, substitutions are retrieved from historical relationships where an extant relationship is one of

- 168666000 SAME AS
- 149016008 MAY BE A
- 370124000 REPLACED BY

Also

- 159083000 WAS_A relationships are used for historical mapping, but only for ‘Limited’ status concepts
1. 384598002 MOVED_FROM relationships are used (inversely) for concepts with Moved status
 2. Where no active substitute is identified but one or more inactive substitutes can be: one or more entries pointing to the inactive substitute concept(s).

Under RF2, the same flavours of association are extracted from the various Historical Association refsets.

After completion of all the processing steps relating to the concept history as set out above, the final steps in production append the Description data as set out in section 3.9.1, and to then compute and insert the flag values for the flags relating to Description strings. It is recognised that there will be a proportion of exceptions where the parsing methods used to isolate the Semantic Tags may fail. This will be addressed as part of the quality management of the table.

The computation and comparison of the top level hierarchy for each of the concepts is then used to set the flag value of TLHIDENTICALFLAG. This requires for any inactive concept the use of the component history table to find the hierarchy to which it was related at the point immediately prior to its (most recent) inactivation.

8 Quality management

The standard method of requesting review of SNOMED CT content should be used where it is thought that the substitutions presented are inappropriate. The Information Standards Service Desk should be contacted at information.standards@nhs.net.

Quality issues may relate to:

- Technical mechanism of production of the table
- Basis for inclusion or exclusion of substitutes
- The Authorship of SNOMED CT content

Any requests or suggestions relating to the format of the table may also be via this route.

8.1 Safety status

As a computed derivative of data drawn from the primary SNOMED release, this product has not been subject to any separate formal clinical safety assessment.

By providing in this document an extensive description of the table and how its contents are generated it is anticipated that a safety assessment will be adequately supported via documentation and inspection of the table.

Feedback from those who reviewed the product during the Technical Preview stage up till November 2011 found no deviations from their own calculated substitutions, and the table has been used extensively since both in-house (to support the maintenance of all published refset) and externally.

For the most part, since the content of what is presented in the table is entirely derived from the content of SNOMED CT, any general statements about the safety for use of SNOMED CT itself may be inherited by this product with the following exceptions:

- The appropriateness of applying filters to exclude some substitutions where these have been deemed inappropriate
- The inclusion of only one history path per record where in some cases there are multiple history paths

- The policy of regenerating and releasing the product only for certain out-of-cycle releases

8.2 Responsibilities

Editorial responsibility for content of the table and artefacts lies with the UK Edition Committee.

Responsibility for the production, quality and distribution of the product and artefacts lies with NHS Digital as the UK NRC.

The UK IReS Strategy Board is responsible for providing strategic leadership for all Terminology and Classification products, services and operations in NHS Digital as well as maintaining other UK national and international commitments, agreeing and defining expectations and delivering oversight to the NHS Digital operational units.

9 Related documentation

The Release Notes from the UK Edition should be read in conjunction with this document if a deeper understanding of the SNOMED CT content is needed. These notes in turn point the reader to the major items of SNOMED International documentation of SNOMED CT such as the Technical Implementation Guide and the Technical Reference Guide.

The Editorial Policy, as documented by either NHS Digital or by SNOMED International, guides the authors of SNOMED CT content and is relevant to the substitutions which they add into the terminology for inactive content.

A sample of the table data is included as Appendix A.

APPENDIX A SAMPLE TABLE

The embedded sample from the table is illustrative of the data format.

This embedded Excel spreadsheet should open when double clicked.

OLD CONCEPT ID	OLD CONCEPT STATUS	NEW CONCEPT ID	NEW CONCEPT STATUS	PATH	IS AMBIGUOUS	ITERATIONS	OLD CONCEPT FS	OLD CONCEPT FS TAG COUNT	NEW CONCEPT FS	NEW CONCEPT FS STATUS	TL IDENTICAL FLAG	FS TAG LESS IDENTICAL FLAG	FS TAG IDENTICAL FLAG	COUNT OF OLD CONCEPTS
642061000000102	6	129118002	0	>265535003:6	1	1	Arteriography NEC (procedure)	1	Arteriography (procedure)	0	1	0	1	2
64222002	4	226646001	0		1	0	Replacement of skull tongs (procedure)	2	Removal of skull tongs with synchronous replacement (procedure)	0	1	0	1	2
64222002	4	227647000	0		1	0	Replacement of skull tongs (procedure)	2	Replacement of skull tongs (procedure)	0	1	1	1	2
64282004	4	205512009	0		1	0	Postoperative stitch abscess (disorder)	2	Postoperative stitch abscess (disorder)	0	1	1	1	2
64282004	4	238383006	0		1	0	Postoperative stitch abscess (disorder)	2	Stitch abscess (disorder)	0	1	0	1	2
64310002	4	49804002	0		1	0	Leptodactylus pentadactylus (living organism)	2	Rana catesbeiana (organism)	0	1	0	1	2
64310002	4	334873008	0		1	0	Leptodactylus pentadactylus (living organism)	2	Leptodactylus pentadactylus (organism)	0	1	0	1	2
643111000000106	6	40108008	0		1	0	Thalassemia major NEC (disorder)	1	Thalassemia (disorder)	0	1	0	1	2
643111000000106	6	75451007	0	>267520000:6	1	1	Thalassemia major NEC (disorder)	1	Thalassemia major (disorder)	0	1	0	1	2

APPENDIX B METHODS FOR FLAGGING AMBIGUOUS SUBSTITUTIONS

A flag attribute 'Ambiguous' is included in the table in section 3.9 which sets out the data specification.

The 'ambiguous' flag is set (i.e. its value = 1) whenever at least one of the following is true:

1. OLDCONCEPTID has ConceptStatus =4 (ambiguous)
2. NEWCONCEPTID has ConceptStatus =4 (ambiguous)
3. At least one concept passed through along ANY single possible path between OLDCONCEPTID and at least one candidate NEWCONCEPTID has ConceptStatus =4 (ambiguous)
4. More than one distinct active NEWCONCEPTID can be found for OLDCONCEPTID including but none of (a), (b) or (c) also apply

At least one explicitly ambiguous path

All substitution rows relating to OLDCONCEPTID will be flagged ISAMBIGUOUS if either:

1. Only one distinct active replacement NEWCONCEPTID can be identified but this can be reached via

1. Either only one path

e.g. 311189000|Air signaller (RC) - RAF (occupation)|

ONLY >413381000000105:4>311145000

2. Or via multiple different paths but all leading to the same NEWCONCEPTID

e.g. 160480004|Occupation of spouse NOS (context-dependent category)|

EITHER >568511000000107:6>266910009:4>447057006

OR >137874007:6>447057006

OR >447057006

..and where at least one of the paths so identified begins with, terminates at or passes through some ambiguous concept (ConceptStatus=4).

2. Two or more distinct active replacement NEWCONCEPTIDs (N1,N2,...) can be identified of which

1. Either all possible paths leading to all possible candidate substitutions neither

begin with nor terminate at but do pass through some ambiguous concept
(ConceptStatus=4)

e.g. 152507005|(Cremation certification) or (Stat B, C and F cremation certs)|

EITHER >26461000000101:4>270113003

OR >26461000000101:4>275656009

2. Or fewer than all but at least one of the candidate substitutions can be reached by at least one path that passes through an ambiguous concept (ConceptStatus=4)

Additionally, concepts with no 'explicitly ambiguous' paths but with more than one distinct active substitute are flagged ISAMBIGUOUS=1

These can divide between different cases:

1. Limited Status concepts where either:

1. no directly equivalent active concept exists any more, and where the original concept (before inactivation) had multiple parents. In these cases the only route back into the active concept space is via the WAS_A links, which enumerate all those parents as possible substitutes.

e.g. 165256003|Other walking aid (finding)|

At the moment this is a relatively uncommon phenomenon

2. a directly equivalent active concept exists and can be found, but it isn't a member of the Former Limited Status Concept subset and so is not given precedence over the original parent concepts, which are also listed

e.g. 147483009|Child 6 month examination NEC (procedure)|

EITHER >170251007:6>170263002|Child 6 month examination (procedure)|

OR >243788004|Child examination (procedure)|

2. Any residue of dm+d concepts for certain unbranded items that have since been diverted to a range of branded ones and which have not been resolved via editorial revision to have just one distinct active substitution during the Technology Preview product stage for the substitution table).

e.g. 331588005|Calamine lotion BP (substance)|

EITHER >14608811000001105|Calamine lotion (product)|

OR >17021711000001104|Calamine lotion (Bell,Sons & Co (Druggists) Ltd) (product)|

OR >17021811000001107|Calamine lotion (Bell,Sons & Co (Druggists) Ltd) 100 ml (product)|

OR >17021911000001102|Calamine lotion (Bell,Sons & Co (Druggists) Ltd) 200 ml (product)|

OR >17960011000001107|Calamine lotion (The Boots Company Plc) (product)|

OR >17960111000001108|Calamine lotion (The Boots Company Plc) 200 ml (product)|

APPENDIX C COMBINATIONS OF SEMANTIC TAG AND TOP LEVEL HIERARCHY

In SNOMED CT the Semantic Tags do not match 1:1 with the Top Level Hierarchies; there are multiple tags found under particular Top Level Hierarchy, there should not however be any Semantic Tags applied below more than one TLH at any point in time.

The exception to this is for inactive concepts; these are related to one of the subtypes of Inactive Concept and retain the tag they had had prior to inactivation.

Information Standards Service Desk
Tel: 0300 30 34 777
E-mail:
information.standards@nhs.net
Internet:
<https://digital.nhs.uk/snomed-ct>
Terminology and Classifications Delivery Service
NHS Digital
1 Trevelyan Square, Boar Lane, Leeds, LS1 6AE
