COVID-19 Related Pathology

Communicating SARS-CoV-2/COVID-19 test results to General Practice: Mapping PMIP-EDIFACT 'standardised descriptions' to SNOMED CT

3rd June 2020

Information and technology
for better health and care

Contents

Change Log:

Version	Date	Summary of Changes
1.0	3rd June 20	First Draft – material taken from "COVID-19 Related Pathology Guidance for naming, coding and processing SARS-COV-2 Pathology Tests"

Purpose

The purpose of this document is to describe the mapping support for communicating SARS-CoV-2/COVID-19 test results to General Practice and to support the guidance document "COVID-19 Related Pathology Guidance for naming, coding and processing SARS-COV-2 Pathology Tests". This document therefore includes a description of:

- The syntax of the temporary coding scheme for representing each result in the PMIP-EDIFACT message.

- A description of the publication format for the mapping files

- A description of the development, maintenance and governance processes associated with the temporary coding scheme and mapping files.

Initial releases of the mapping data were released in the format described in Appendix 1. In order to improve maintenance and governance process the format of the mapping file has changed from earlier releases, and will, for now, use the SNOMED CT simple map reference set pattern.

Background

The national standard for conveying the results of SARS-COV-2 pathology tests to General Practice is the PMIP-EDIFACT message which uses a tightly-managed set of Read codes (the Pathology Bounded Code List – PBCL). Strategic changes to terminology and pathology standards over recent years mean that the membership of the PBCL cannot be extended, and therefore an alternative mechanism is needed if new pathology tests (in this case SARS-CoV-2/COVID-19 tests) are to be communicated using the PMIP-EDIFACT message. Longer term strategic solutions are under development but incomplete. Therefore a tactical approach, based on the communication of a 'standardised description' code proxy in a suitable text field of the PMIP-EDIFACT message (and its subsequent translation into a corresponding SNOMED CT code for storage and display in General Practice systems) is being used.

Source (communication) scheme

The source scheme is a set of 'standardised descriptions' developed specifically for this purpose. This will allow parsing and processing of results by EPR message handlers. Technically the 'standardised descriptions' conform to the following rules:

- no more than 30 characters long (currently none exceed 15 characters)

- containing only characters from the UNOC character set

Each standardised description includes a segment summarising the test name (14 characters) and a segment indicating the value (1 character).

- All standardised descriptions have an 11 character prefix 'SARS-CoV-2-'

- Tests names are distinguished by the next 3 characters.

- Result values or value types are indicated by the final character. The characters used and their intended corresponding expansions or interpretations are:

Final Character	Expansion
Y	POSITIVE
N	NEGATIVE
Q	INDETERMINATE
U	UNKNOWN
c	ARBITRARY CONCENTRATION

An example 'Standardised description' is **SARS-CoV-2-ORGY**. This can be broken down into:

SARS-CoV-2-ORGY	The initial opening 11 characters
SARS-CoV-2-**ORG**Y	The three character test name identifier
SARS-CoV-2-ORG**Y**	The single character result identifier

Note: as the source scheme only functions as a map source scheme and has no stand-alone value, its members will only be published in the 'scheme value' fields of the mapping files, and will not be separately published. For internal purposes the scheme is referred to as:

NHS SARS-CoV-2 laboratory results standardised descriptions code set

Target (storage and analysis) scheme

The target scheme is SNOMED CT, with SNOMED CT ConceptIds and DescriptionIds being the identifier of each target entity in the respective mapping files.

- o Combining the DescriptionId with other fields from the corresponding row of the Descriptions table allow identification of the corresponding Term (for record storage and human-readable display)

- o Combining the corresponding ConceptId with data from the Relationships table[1] allows identification of the class of concept communicated, and thus whether the concept communicated requires a separate value for interpretation:

  - o Target concepts that are transitively kinds of 404684003 | Clinical finding (finding) should be interpretable without the need for any additional 'value'

---

[1] Or a transitive derivative as described at
https://confluence.ihtsdotools.org/display/DOCRELFMT/Transitive+closure+file

      o Target concepts that are transitively kinds of 363787002 | Observable entity (observable entity) require an accompanying value for interpretation.

      o (this distinction reproduces the ExpectValue field of the file described in Appendix 1).

As examples:

The source 'Standardised description' **SARS-CoV-2-ORGY** is mapped to the target SNOMED CT ConceptId **1240581000000104** (with a preferred term of SARS-CoV-2 (severe acute respiratory syndrome coronavirus 2) detection result *positive)*. Transitive analysis reveals that **1240581000000104** is a kind of 404684003 | Clinical finding (finding)

This code includes a statement that the test result is **positive**, and therefore has no need for any other value representation to be sent. *Indeed, if there is a separate value sent this must be regarded as an error, since there is no way of confirming that this additional value is consistent or inconsistent with that declared in a statement using the finding.*

Nature of the maps

Purpose

The maps are produced to translate **FROM** 'SARS-CoV-2 standardised descriptions' (entered into each relevant PMIP-EDIFACT message) **TO** SNOMED CT ConceptIds and DescriptionIds (and thus specific terms) for entry into GP system records.

Properties

In order to reproduce the most valuable fields of the original mapping table (see Appendix 1), the maps will be distributed a complementary pair of simple map reference sets:

- COVID-19 test result communication to general practice (concept) simple map reference set

- COVID-19 test result communication to general practice (description) simple map reference set

Cardinality: Each map will have a cardinality of 1:1

- **ONE** 'standardised description' maps to **ONE** SNOMED CT DescriptionId.

- **ONE** 'standardised description' maps to **ONE** SNOMED CT ConceptId.

Direction: The intended use/direction of the maps is **FROM** each 'standardised description' **TO** each SNOMED CT ConceptId and DescriptionId

Loss/gain of meaning and degree of equivalence: There is no loss or gain in meaning when mapping from each standardised description to each SNOMED CT ConceptId or DescriptionId. The degree of equivalence between each 'standardised description' and each corresponding SNOMED CT ConceptId or DescriptionId is comparable to the SNOMED CT map correlation value *447557004 | Exact match map from SNOMED CT source code to target code (foundation metadata concept).*

Heuristics: no heuristics are required for this map. The intention is simply to use a standard alphanumeric string for communication in the PMIP-EDIFACT message with 1:1 translation TO a SNOMED CT ConceptId and DescriptionId.

Mapping file nature and publication

In order to make best use of established terminology team quality assurance processes and technical tooling, the map is now maintained and published as a pair of SNOMED CT simple map reference sets.

Mapping file names

- COVID-19 test result communication to general practice (concept) simple map reference set

- COVID-19 test result communication to general practice (description) simple map reference set

Mapping file format

Consistent with other simple map reference sets, the format of the simple map mapping files is defined thus:

Field	Data type	Purpose
id	UUID	A 128 bit unsigned Integer, uniquely identifying this reference set member.
effectiveTime	Time	The inclusive date or time at which this version of the identified reference set member became the current version.
active	Boolean	The state of the identified reference set member as at the specified effectiveTime .
		If active = 1 (true) the reference set member is part of the current version of the set, if active = 0 (false) the reference set member is not part of the current version of the set.
moduleId	SCTID	Identifies the SNOMED CT module that contains this reference set member as at the specified effectiveTime .
		This will be 999000021000000109	SNOMED CT United Kingdom clinical extension reference set module (core metadata concept)
refsetId	SCTID	Identifies the reference set to which this reference set member belongs
referencedComponentId	SCTID	A reference to the SNOMED CT component to be included in the reference set. This will be the DescriptionId or ConceptId for each relevant SARS-CoV-2/COVID-19 SNOMED CT test code.
mapTarget	String	The equivalent code in the other terminology, classification or code system. This is the 'standardised description' corresponding to each SARS-CoV-2/COVID-19 SNOMED CT test code.

The two fields that hold the codes from the mapping schemes are:

referencedComponentId

This field (typed to carry a SCTID) will contain each ConceptId or DescriptionId

Mapping file	Field contents
COVID-19 test result communication to general practice (concept) simple map reference set	ConceptId
COVID-19 test result communication to general practice (description) simple map reference set	DescriptionId

mapTarget

This field (typed to carry a string) will contain each 'standardised description string'.

Note on field names and the mapping direction.

As repeatedly stated above, the maps are **FROM** the standardised description **TO** the SNOMED CT components. The SNOMED CT simple map reference set pattern has been chosen because of its established nature in the release and UK tooling (rather than create a brand new reference set pattern). This has the benefit of making the maps easy to maintain and publish, but (1) creates the need for two mapping files and (2) may result in paradoxical field names. If possible referencedComponentId will be labelled as 900000000000505001 | Map target | in the reference set descriptor file for both reference sets. If this is not possible then users should still **regard the referencedComponentId as the 'Target' of each map**, even if it is labelled as 900000000000500006 | Map source concept |. Likewise, each mapTarget will be labelled as 900000000000499002 | Scheme value | in the reference set descriptor file for both reference sets. Users should therefore **regard the mapTarget as the 'Source' of each map**.

Publication

Each mapping file will be published along with each release of the UK edition of SNOMED CT, and will therefore be released with the same frequency. Any requirement for 'out of cycle' changes (either to the maps themselves or to components referenced within the maps) will need to be considered on its merit.

Licensing

In keeping with other similar products, the maps will be published under OGL & SNOMED CT licensing arrangements.

Interpretation and use of mapping file data

Files should be processed in the same way as all other standard SNOMED CT RF2 files, including the need for filtering of inactive rows from snapshot files.

Detailed guidance on the use of the maps (**FROM** 'standardised descriptions' **TO** SNOMED CT) is given in the document "COVID-19 Related Pathology Guidance for naming, coding and processing SARS-COV-2 Pathology Tests", particularly *section 4.3. Primary Care EPR Result Processing*.

Where the map requires transformation from a 'standardised description' to a SNOMED CT ConceptId, the data provided in the *COVID-19 test result communication to general practice (concept) simple map reference set* file should be used.

Where the map requires transformation from a 'standardised description' to a SNOMED CT term for display, the data provided in the *COVID-19 test result communication to general practice (description) simple map reference set* file should be used, and the Descriptions table used to complete the join from the DescriptionId.

Governance, development and assurance processes

Authoring skills

Content changes will be made by terminology specialists with:

- SNOMED CT authoring knowledge, skills and experience

- Assessed as competent to promote new or changed SNOMED CT content for release

- Experience of local code set creation and maintenance

Content changes

Automated change validation

- Snow Owl terminology tooling task editor enforces completion of predefined metadata which remain accessible once the task is completed.

- Snow Owl terminology tooling validation rules enforce compliance with a predefined list of quality assurance rules.

Inspection and review

Peer review by two additional terminology authors following authoring using Snow Owl automated workflow.

Escalation and governance

As required, consultation will be made with the Pathology Standards Governance Board (PSGB) regarding the mapping tables or the standardised descriptions scheme. Changes and updates will also be agreed with PSGB.

Maintenance and evaluation

Reference sets & maps will be maintained in synchronisation with the primary release data and are subject to quality assurance processes to review the output and any changes occurring because of updates to the UK Edition of SNOMED CT and the International Edition.

Any maintenance changes will be reviewed by an author, and for reference set/mapping changes discussed with and approved by the reference set owner. Any other changes would be appropriately escalated.

Version control

Version control for each reference set is managed in line with SNOMED CT release file naming conventions. The VersionDate element is mandatory for all release file types. Its format is an 8-digit number in the pattern "YYYYMMDD", in compliance with the ISO-8601 standard.

Version control for the standardised descriptions scheme is managed by the requirement for each import file to have the mandatory field Code System Version completed.

Appendix 1: Original mapping table

This appendix describes the original mapping file used in testing and is for information only. The original mapping file has been replaced by the SNOMED CT simple map reference set approach to improve publication and governance processes.

The original mapping tables included only active rows. Each release of the mapping file was named as follows, including a YYYYMMDD date stamp:

`covid_lab_sdsctmap_uk_YYYYMMDD.txt`

The mapping tables were released on TRUD and Delen but the latest file at the time of writing is below.

Original mapping file schema

The original mapping file was presented as a TAB delimited file with rows terminated by CR/LF combination. The first row contained the relevant field names. The mapping file used the following format/schema:

Column	Length	Type / Pattern	Database type	Note
LabId	16	STRING	BINARY (16)	Text string 'Standardised description' in 16 characters of fewer
ConceptId	18	SCTID	VARCHAR (18) or BIGINT(20)	SNOMED ConceptID
DescriptionId	18	SCTID	VARCHAR(18) or BIGINT(20)	SNOMED DescriptionID
Term	255	STRING	VARCHAR(255)	SNOMED CT preferred synonym for this concept (included for ease of reading)
ExpectValue	1	0|1	BOOLEAN	0 = Separate value not expected, 1 = Separate value expected

Original mapping file column details

LabId

This is the 15 character 'Standardised description' that will be used to communicate the test name and the result/nature of the result. Its format is described in detail in the SG18 section above. The LabId therefore serves as the source scheme to the mapping process.

ConceptId, DescriptionId and Term

These are the standard SNOMED CT Concept and Description identifiers as well as preferred synonym terms for the target of each map, and represents the SNOMED CT-encoded test result or test name to be entered in the record as a result of the map.

ExpectValue

This boolean field helps discriminate between two types of target code.

Where ExpectValue=0, the 'value' of the test result will be stated in the 'finding' concept that is the target of the map.

For example, the source 'Standardised description' (in LabId) **SARS-CoV-2-ORGY**

is mapped to the target SNOMED CT 'finding' concept (in ConceptId, DescriptionId and Term)

1240581000000104 SARS-CoV-2 (severe acute respiratory syndrome coronavirus 2) detection result **positive**

This code includes a statement that the test result is positive, and therefore has no need for any other value representation to be sent, and so will have expectValue set to 0. *Indeed, if there is a separate value sent this must be regarded as an error, since there is no way of confirming that this additional value is consistent or inconsistent with that declared in a statement using the finding.*

By contrast, the source 'Standardised description' (in LabId) **SARS-CoV-2-IgTc**

is mapped to the target SNOMED CT 'observable entity' concept (in ConceptId, DescriptionId and Term)

1321341000000103 SARS-CoV-2 (severe acute respiratory syndrome coronavirus 2) IgG arbitrary concentration in serum

Alone this code is incomplete as a test result, and requires a value representation to be sent, and this row will have expectValue set to 1. *Indeed, if a separate value is not sent this must be regarded as an error, since there is no way that a statement made using this observable entity code can be interpreted as a result without an accompanying value.*

As noted above, it is expected that the ARBITRARY CONCENTRATION observable entity codes will be sent in Segment group 18, with values and Units of Measure as +RSL Tag +C830.