Health & Social Care Information Centre — hscic

Document filename: ICD-10 5th Edition: Tables of Coding Equivalences (Analysis) File Specification			
Directorate / Programme	Architecture, Standards and Innovation / Clinical Classifications Service	**Project**	N/A
Document Reference		CCS-2014-356	
Project Manager	N/A	**Status**	Draft
Owner	Lynn Bracewell	**Version**	V1.0
Author	Andrew Brooke	**Version issue date**	15/07/2015

ICD-10 5th Edition: Tables of Coding Equivalences (Analysis) File Specification

Document Management

Revision History

Version	Date	Summary of Changes
V0.1	15-7-2014	Updated to reflect ICD-10 5th Edition
V0.2	04-9-2014	Updated following V0.1 review
V0.3	26-03-2015	Updated to reflect previous ISAS comments about mapping to HRG's
V0.4	15-06-2015	Updated template headings and dates to reflect 2016 release
V0.5	14-07-2015	Version sent for approval
V1.0	15-07-2015	Approved version

Reviewers

This document must be reviewed by the following people: author to indicate reviewers

Reviewer name	Title / Responsibility	Date	Version
Hazel Brear	Senior Classifications Manager	19-Sep-2014	V0.2
Adrian Wilkins	Technical Services Workstream Lead	16-Jul-2014	V0.1
Tim Ireland	Development Team Manager	14-Jul-2015	V0.4
Ruth Toner	Classifications Development and QA Specialist	16-Jul-2014 04-Sep-2014 06-Jul-2015	V0.1 V0.2 V0.4

Approved by

This document must be approved by the following people: author to indicate approvers

Name	Signature	Title	Date	Version
Lynn Bracewell	Lynn Bracewell	Head of Terminology and Classifications development, Architecture, Standards and Innovation	15-07-2015	V1.0

Glossary of Terms

Term / Abbreviation	What it stands for
category	A category is a three-character code and rubric/description.
subcategory	A subcategory is a four-character code and rubric/description. A decimal

	point may be shown between the third and fourth character.
	Most three-character categories are subdivided into four-character subcategories.
	Insertion of the character 'X' into the fourth field of codes where only three characters exist is mandatory, so the codes are of a standard length for data processing and validation. This provides a standard 'filler' character.
	Codes with a format of AnnX are included in references to 4 character codes in this specification.
	Many systems, but not all systems, automatically insert the filler 'X', therefore all users of coded clinical data should be aware of this requirement.
(sub)category	Refers to both category and subcategory in this document
rubric	The text associated with a (sub)category. Also known as title or code description
Field name	The official name for a field within a record
Char	Character (A letter (e.g., a, A, b, and B), number (0, 1, 2, or 3) etc.
DSV	Delimiter Separated Values (DSV) format
	Specifically, the variant is "Tab Separated Values" (the delimiter being a tab character).
WHO	World Health Organisation
HES	Hospital Episode Statistics
TRUD	Technology Reference data Update Distribution
	https://isd.hscic.gov.uk/trud3/user/guest/group/0/home

Document Control:

The controlled copy of this document is maintained in the HSCIC corporate network. Any copies of this document held outside of 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.

Contents

1 Introduction

1.1 Purpose

The ICD-10 Tables of Coding Equivalences (Analysis) (ToCE) are designed to assist with the comparison of data collected according to ICD-10 5th Edition with that collected according to the ICD-10 4th Edition. These files will act as an invaluable user-friendly reference for all ICD-10 users.

1.2 Audience

Information analysts wishing to compare data recorded using the ICD-10 4th Edition with that recorded using the ICD-10 5th Edition and visa versa.

1.3 Out of scope

The HSCIC National Casemix Office (NCO) is responsible for the mapping of ICD-10 codes to Healthcare Resource Groups (HRG) for use in associated groupers. The ICD-10 updates will be incorporated into the HRGs to ensure the codes are accepted in the grouper(s) to be used from April 2016 onwards. See www.hscic.gov.uk/casemix for details about HRGs.

The NCO is responsible for the mapping of ICD-10 codes to HRG's for use in associated groupers and use the ToCE to inform this work.

2 Background

2.1 ICD-10

The International Statistical Classification of Diseases and Related Health Problems, Tenth Revision, (ICD-10) is a comprehensive classification of causes of morbidity and mortality. It contains an alphanumeric listing of codes for diseases, injuries, external causes of injury and poisonings, and factors influencing health status and contact with health.

It is maintained and published in three Volumes by the World Health Organisation. See http://www.who.int/classifications/icd/en/

2.2 Tables of Coding Equivalences (ToCE)

The ToCE issued with ICD-10 5th Edition is the source of equivalences between ICD-10 5th Edition and ICD-10 4th Edition.

Updates to the ToCE are released in conjunction with updates to the ICD-10 classification. Feedback from users may merit earlier change to the data file and potentially clarification to the specification.

Building products that state the equivalences between two versions requires an element of expert coding judgement. It is always the case in such circumstances that use of the equivalence tables will precipitate queries. It is expected that this may, ultimately, reveal areas where change is warranted. This is recognised as an unavoidable and essential part of the development process.

3 WHO Copyright

The World Health Organization (WHO) is the copyright holder of ICD-10 which is used under licence for United Kingdom government purposes.

ICD-10 codes, terms and text used by permission of the WHO, from: International Statistical Classification of Diseases and Related Health Problems, Tenth Revision (ICD-10) Volumes 1-3. Geneva, World Health Organization, 1992-2016.

4 WHO updates to ICD-10

ICD-10 5th Edition includes the WHO cumulative changes made between January 2011 and January 2016. ICD-10 updates were obtained from the WHO as PDF documents describing revisions to the classification. These were used to crosscheck updates to the 5th Edition Maintenance and release.

The Clinical Classifications Service of the HSCIC is responsible for maintaining and issuing the ICD-10 Tables of Coding Equivalences.

The files are updated and released to support new versions of ICD-10.

5 ICD-10 Tables of Coding Equivalences

5.1 Construction of the Tables

The list of equivalences is constructed by classification experts with an excellent knowledge of the ICD-10 Classification. Using their clinical coding skills, they compare the changes made in the ICD-10 5th Edition, Tabular list (changes to code descriptions, includes and excludes notes) and the Alphabetical Index (changes to index trails) to the ICD-10 4th Edition and apply the ICD-10 classification rules, conventions and National Clinical Coding Standards.

5.2 Mappings

The ToCE provides analysts with backward and forward equivalence mapping tables as a guide to the codes that should have been used for the same diagnosis in both versions of ICD-10.

The mappings provide:

One to one mapping – where a single ICD-10 code maps to a single ICD-10 code. The code may be the same or different.

One to many mapping – where a single ICD-10 code may map to a number of possible ICD-10 codes or combinations of ICD-10 codes.

One to NONE mapping – where a new concept has been introduced in ICD-10 5th Edition which cannot be classified in ICD-10 4th Edition or when an existing concept in ICD-10 4th Edition is not represented in ICD-10 5th Edition.

Many to one mapping – where a previously un-subdivided three-character code in ICD-10 4th Edition is expanded in ICD-10 5th Edition.

The majority of equivalence mappings are one to one. However several of the maps are complex and therefore special notations, as described in the following tables, have been used to construct the code combinations.

Important Note

The coding of clinical data is governed by classification rules, conventions and standards contained in the Clinical Coding Standards. Analysts should seek clinical coding advice, as necessary, in respect of complex analyses and mappings to ensure valid outcomes.

Notation	Description
+	Indicates that a combination of two or more ICD-10 codes are equivalent [a + b = a and b together] to the source code.
Or	Indicates that an ICD-10 code will map to alternative codes [a or b = either a or b]
+/or	Indicates that a combination of ICD-10 codes or individual ICD-10 codes are equivalent [a +/or b = a + b together or either a or b] to the source code
NONE	Indicates that there are no equivalent ICD-10 codes
()	Used to group combinations of ICD-10 codes and distinguish between alternatives [(a + b) or (c + d) = a and b together or c and d together]
Xnn-	The hyphen is used at four-character code level to denote that the source code maps to each individual target code within the whole category, e.g. **A03.-** includes **A03.0**, **A03.1**, **A03.2**, **A03.3**, **A03.8** and **A03.9**
Xnn.n – **Xnn.n** *or* **Xnn.-** *or* **Xnn –** **Xnn**	A range describes a range of categories. It may be open or closed. A range may be closed, where an upper bound is provided. A closed range includes the categories of both bounds, plus any of their subcategories. e.g. **I84.7-I84.9** includes **I84.7**, **I84.8**, **I84.9**. "Choose any code (inclusive) between **I84.7** and **I84.9**" A range may be open, where no upper bound is provided. An open range includes any subcategories prefixed by the range. A range implies a choice of any item from that range. e.g. **N18.-** includes **N18.1**, **N18.2**, **N18.3**, **N18.4**, **N18.5**, **N18.9** "Choose any subcategory of **N18**" The following example is provided: <table><tr><td>**U80.8**</td><td>Agent resistant to other penicillin-related antibiotic</td><td>**U82.2-U82.9**</td></tr></table> This indicates that each individual code between and including **U82.2** to **U82.9** are equivalent to **U80.8**, i.e. **U82.2**, **U82.8** and **U82.9** are equivalent to **U80.8.**

Other Coding Conventions	Description
Decimal point	Precedes a fourth character including three-character codes suffixed with an "X"
X-Suffix	An "X" filler is used to fill the fourth position at the end of undivided three-character category codes. e.g. **I48** is represented as **I48.X** in ICD-10 4th edition as it has no descendants. **I48.X** becomes **I48** in the ICD-10 5th Edition as it gains 6 subdivisions - **I48.0** to **I48.9**
D and A Suffix	"D" suffix denotes "dagger" usage; an "A" suffix denotes "asterisk" usage. For information about dagger and asterisk codes see: *ICD10_Edition5_CodesAndTitlesAndMetadataFileSpecification_GB_20160401* in the NHS ICD-10 5th Edition Data Files - available for download from TRUD.

5.3 Retired codes

Codes which exist in ICD-10 4th Edition but are retired from ICD-10 5th Edition are listed as source codes in the forward equivalence mapping table and have ICD-10 5th Edition target codes provided or are marked as "NONE". Retired codes do not appear as source codes in the backward equivalence table; however they may still be listed as an equivalent target for a source code that is present in ICD-10 5th Edition in the backwards equivalence table.

For the list of changes to codes including retired codes see: *ICD10_Edition5_SummaryOfChanges_GB_20160401.pdf* in the NHS ICD-10 5th Edition Data Files - available for download from TRUD.

5.4 Technical description

This section describes the files and their content.

The ICD-10 5th Edition Tables of Coding Equivalences (Analysis) files are distributed at: https://isd.hscic.gov.uk/trud3/user/guest/group/0/home

Forward and backward ToCE are provided in four separate text files as follows:

- *ICD10_Edition5_TableOfCodingEquivalencesForward_GB_20160401.txt*
- *ICD10_Edition5_TableOfCodingEquivalencesBackward_GB_20160401.txt*
- *ICD10_Edition5_TableOfCodingEquivalencesWithDescriptionsForward_GB_20160401.txt*
- *ICD10_Edition5_TableOfCodingEquivalencesWithDescriptionsBackward_GB_20160401.txt*

The first and second files are aimed at technical analysts and statisticians that require machine-readable files to manipulate large volumes of data. These users should seek clinical coding input as advised at Important Note at **section 6.2** above, as necessary.

The third and fourth files contain code descriptions and an additional field to filter the equivalence maps.

5.4.1 Field List for files without code descriptions

Numbe	Name	Null	Description
1	CODE	No	ICD-10 Code of the *source* category*
2	MAP	No	Map to the *target* revision
3	NOTES	Yes	Explanatory Notes
4	DIFFERENT_TARGET	No	False if the map describes the same category in the *target* revision, true if otherwise.

*Category is used generically to mean three, four or five-character codes in this document.

5.4.2 CODE

Form	See description
Null?	No

Description

This field will carry the ICD-10 code identifying the category in the source revision of the data.

5.4.3 MAP

Form	See description
Null?	No

Description

This field contains the map information. This data is used to specify the possible expected output when moving between versions of the ICD-10 data.

5.4.4 NOTES

Form	See description
Null?	Yes

Description

This field contains notes about this mapping. In practice, this has been used to provide a list of categories associated to the source categories, usually categories with the complementary USAGE_UK (daggers + asterisks).

e.g.

D63.0A

"Associated dagger code **C00-D48**"

5.4.5 DIFFERENT_TARGET

Form	"true" or "false"
Null?	No

This field expresses whether the target mapping for this row differs from the original source.

5.5 Map grammar

5.5.1 Vector

Each map file describes a single *vector*, a quantity with both origin, and direction, or in this case, a source and a target.

In this context, we are moving between one version of the ICD-10 data and another, the two versions concerned here are ICD-10 4th Edition and ICD-10 5th Edition.

5.5.2 Source

The source revision is the starting point of the map. The codes in the CODE column are drawn from this revision.

e.g. If your existing data is encoded using ICD-10 4th Edition, you should use a map which has this version as its source.

5.5.3 Target

The target version is the point which the map describes as its output. The codes in the MAP column are drawn from this version.

e.g. If the intention is to encode ICD-10 4th Edition data using ICD-10 5th Edition codes, the use of a map that has the ICD-10 5th Edition as its target is required.

5.5.4 Direction

Forward is used to note a map that proceeds along a vector that has a direction moving forward chronologically, and *backward* describes a map that has a vector that moves backward chronologically.

e.g. The file:

ICD10_Edition5_TableOfCodingEquivalencesForward_GB_20160401.txt

Describes a map that facilitates the conversion of data from the older ICD-10 4th Edition of the data, to the newer ICD-10 5th Edition.

5.6 Cross-referencing data to the Codes and Titles and Metadata file

When referring to data outside the map, care should be taken to refer to the correct revision of the data for each column.

Analysts should join the CODE column from the map file to the CODE column of the relevant main data file representing the *source* revision of the data.

e.g. For the "Forward" vector map, this is ICD-10 4th Edition data file:

ICD10_Edition4_CodesAndTitlesAndMetadata_GB_20120401.txt. This file is available from TRUD in the NHS ICD-10 4th Edition Data Files.

Analysts should strip suffixes from category codes in the MAP column and join to the relevant main data file representing the *target* revision of the data.

e.g. For the "Forward" vector map, this is:

ICD10_Edition5_CodesAndTitlesAndMetadata_GB_20160401.txt.

N16.5A   becomes   **N16.5**

N19.X   becomes   **N19**

B59.XD   becomes   **B59**

For open ranges, analysts should strip the dash character, and select rows with CODE values between the remaining string, and the remaining string plus a lowercase z character - this will include the relevant subcategories.

e.g. **C86.-**

"SELECT * FROM [data] WHERE CODE BETWEEN '**C86.**' AND **'C86.z**"

For closed ranges, Analysts should remember that all subcategories are included in the upper bound.

e.g. For **C00-D48** (the whole of ICD-10 Chapter II)

"SELECT * FROM [data] WHERE CODE BETWEEN '**C00**' AND **'D48.z'**"

5.7 "With Descriptions" format

These additional reports are provided with descriptions.

5.7.1 Field List for codes with descriptions

The fields in the actual files for this format have descriptive titles appropriate to the vector they represent.

Since these files are intended for import into Excel *(see section 7)*, we also provide column lettering.

Colum	Name	Null	Description
A	From: ICD-10 5th Edition code / From: ICD-10 4th Edition code	No	ICD-10 Code of the *source* category
B	Code Description	No	The description of the *source* category for easy reference
C	To: ICD-10 5th Edition code / From: ICD-10 4th Edition code	No	Map to the *target* revision
D	Notes	Yes	Explanatory notes
E	Target Different to Source?	No	N if the map describes the same category in the *target* revision, Y if otherwise.

F	ICD 3char/4char/5char	No	The significant number of characters in the source category code. 3, 4, or 5.

Fields are as discussed above, with the addition of the following.

5.7.2 From: ICD-10 5th Edition code / From: ICD-10 4th Edition code

This field is the *source* category code, formatted for consistency, the same way as the code references in the MAP column, that is to say:

- three-character categories with no children receive an "X" suffix
- Dagger usage categories receive a "D" suffix
- Asterisk usage categories receive an "A" suffix

5.7.3 Code Description

This field contains the description of the *source* category. The descriptions for four and five-character codes that are the product of modifying their parent category are presented with their modifier suffix. Analysts MUST NOT use these descriptions for systems implementation; they are only presented in this form for convenience.

5.7.4 Target Different to Source?

The same as the DIFFERENT_TARGET field in the "without" descriptions file, but with "Y" or "N" instead of "true" or "false".

5.7.5 ICD 3char/4char/5char

The number of significant characters for each *source* code (excluding the decimal point, and any suffixes). This is intended to assist filtering the list to only those source categories of a particular length.

6 Viewing the DSV file – advice for non-technical users

The simplest way to view these txt files available to most users is to drag the file icon into an open, empty, Excel window.

See Annex 1 of the *ICD10_Edition5_CodesAndTitlesAndMetadataFileSpecification_GB_20160401.pdf* * for instructions on how to create a more formal data import.

It is NOT recommended that you save the file from Excel because it will break the file structure. Setting the read-only flag on the file is advised.

* This is available in TRUD in the NHS ICD-10 5th Edition Data Files.

7 Enquiries

Information Standards Helpdesk:

Tel: 0300 30 34 777 weekdays from 9am to 5pm

Email: information.standards@hscic.gov.uk

For information about the Clinical Classifications Service visit website:

http://systems.hscic.gov.uk/data/clinicalcoding

For information about ICD-10 5th Edition visit website:

http://systems.hscic.gov.uk/data/clinicalcoding/codingstandards/icd10/icd10updates

For information about TRUD visit website:

https://isd.hscic.gov.uk/trud3/user/guest/group/0/home