



orphanet

October 2014

# ICD-10 coding rules for rare diseases

Procedural document

[www.orpha.net](http://www.orpha.net)

# Table of contents

<b>Table of contents.....</b>	<b>2</b>
<b>General points .....</b>	<b>3</b>
I. Purpose.....	3
II. Reference documents.....	3
III. Range of application .....	3
IV. Definitions.....	3
<b>Attributing the codes .....</b>	<b>4</b>
I. ICD-10 reference version .....	4
II. Default rule: priority to any ICD-10 mention .....	4
III. Default rule: use single codes .....	4
IV. Double coding by the dagger-and-asterisk system .....	4
V. Entities representable by a set of ICD-10 codes.....	5
VI. Entities representable by a three-character ICD-10 code .....	6
VII. Entities not mentioned in ICD-10.....	7
<b>Typing the codings .....</b>	<b>8</b>
I. Typing the matching between concepts .....	8
II. Typing the ICD-10 term level.....	9
III. Validating the codings.....	9
<b>Using the codings.....</b>	<b>9</b>
I. Coding indication and retrieval of medical records .....	9
II. Coding indication only .....	10
III. Codes not to be used .....	10
<b>Decisions for specific groups of diseases .....</b>	<b>10</b>
I. Tumours .....	10
II. Secreting tumours .....	11
III. Cancer-predisposing syndromes.....	11
IV. Mitochondrial diseases.....	11
V. Chromosomal microdeletions and microduplications .....	12
VI. Cerebellar malformations .....	12
VII. Joubert syndrome, with or without other involvement .....	12
VIII. Spinocerebellar ataxias .....	12
IX. Hereditary sensory and autonomic neuropathies .....	12
X. Congenital bile acid synthesis defects.....	12
XI. Congenital disorders of glycosylation .....	12
XII. Susceptibility to infections.....	12
XIII. Distal arthrogryposes.....	12

# General points

## I. Purpose

This document aims to define how rare diseases in the Orphanet database are attributed a code in the World Health Organization's *International Classification of Diseases*, 10th edition (ICD-10).

## II. Reference documents

### *International Statistical Classification of Diseases and Related Health Problems*

- **10th Revision:** searchable on the WHO online browser for the 2010 version in English, the URL is: <http://apps.who.int/classifications/icd10/browse/2010/en>
- **List of Official ICD-10 Updates:** PDF files with the ICD-10 updates endorsed by the WHO over the years. URL: <http://www.who.int/classifications/icd/icd10updates/en/>
- **Linearization principles:** a description of the rules followed to establish the linearization rules applied to Orphanet's classifications of rare diseases

## III. Range of application

The ICD-10 coding of Orphanet diseases is managed by a dedicated information scientist, under responsibility of the scientific direction.

## IV. Definitions

The ICD-10 defines three levels of terms associated with ICD-10 codes to represent their range of application: *main terms*, *inclusion terms*, *index terms*.

A **main term** is the primary identification of an ICD-10 code. It is displayed in bold letters beside the code in the ICD-10 tabulated list.

An **inclusion term** is a term associated to an ICD-10 code to precise its range and use. It is displayed in the ICD-10 tabulated list in plain letters under the code and main term. It can be a synonym of the main name, a subtype of the disease represented by main name, or a specific disease subsumed under a code aiming to represent a group of diseases as a whole.

An **index term** has the same purpose than an inclusion term but is found only in the ICD-10 alphabetic index. It is not displayed in the ICD-10 tabulated list.

# Attributing the codes

## I. I. ICD-10 reference version

The coding of rare diseases in Orphanet by ICD-10 codes is based on the 2010 online version of the ICD-10, but takes into account the official WHO updates endorsed in 2011 and 2012.

## II. Default rule: priority to any ICD-10 mention

Any explicit mention from the ICD-10 has priority over internal decisions, even if the ICD-10 dispositions are deemed to be inaccurate.

E.g. *Progeria* is an inclusion term of *E34.8 Other specified endocrine disorders*. Therefore *ORPHA740 Hutchinson-Gilford progeria syndrome* must be coded E34.8, even if it is not correct to describe it as an endocrine disease.

## III. Default rule: use single codes

The general rule is to represent diseases by a single ICD-10 code, whether the published ICD-10 mentions the disease or not.

In particular, when the ICD-10 does not mention the disease and therefore needs to be interpreted, the most significant involvement must be selected, meaning the one:

- corresponding to the most severely affected body system
- most determining for the prognosis
- whose specialist is most likely to be relied on for disease management.

The selection of the most significant involvement should generally be consistent with the Orphanet linearization for rare diseases.

There are nevertheless a number of exceptions that are described below.

## IV. Double coding by the dagger-and-asterisk system

The system of a main code with a dagger associated to a secondary code with an asterisk has been introduced in ICD-9 and maintained in ICD-10 to represent several cases when two approaches are useful. The ICD-10 user manual lists the following uses of this system:

- local manifestation of a generalized disease, especially infections
- functional activity (and consequences) of endocrine tumours
- the organic cause of a mental or behavioural disorder
- a toxic or pharmacologic cause of disease
- a traumatic cause of disease

Such double codes are allowed only when the possibility is explicitly afforded by the ICD-10. A secondary asterisk code can be used only in association with a primary dagger code.

The dagger-and-asterisk system is used when relevant to code Orphanet entities.

E.g. *ORPHA137586 Herpes simplex virus keratitis* is coded by the association of:

- *B00.5+ Herpesviral ocular disease*
- *H19.1\* Herpesviral keratitis and keratoconjunctivitis*

## V. Entities representable by a set of ICD-10 codes

It may happen that the ICD-10 provides several contextual codes when Orphanet only has a general entity. This is not infrequent especially with infectious diseases, which have for historical and practical reasons very detailed ICD-10 codes, whereas they are less detailed in Orphanet.

Coding an instance of such diseases properly requires additional information compared with the range of Orphanet entry. In such cases, it is useful to provide the whole set of codes that can be used to represent the disease at large, according to the context.

However, this possibility of multiple coding should normally be used only when the union of all codes represents the disease accurately. There is one important exception: the coding of tumours. Codes for tumours always lack specificity because they are intended only to indicate location and behaviour in a very general way. Rare tumours often have to be represented with a set of general codes.

Most often (but not always), the set of ICD-10 codes will be a series of four character-codes under a heading which is too generic compared with the Orphanet entity.

E.g. *ORPHA31202 Melioidosis* is represented by the following ICD-10 codes:

- *A24.1 Acute and fulminating melioidosis*
- *A24.2 Subacute and chronic melioidosis*
- *A24.3 Other melioidosis*
- *A24.4 Melioidosis, unspecified*

Coding by the whole upper code A24 would result in a loss of specificity because A24 is described *Glanders and melioidosis* (A24.0 is used to refer specifically to glanders) and therefore subsumes melioidosis with another disease.

Coding with *A24.4 Melioidosis, unspecified* would exclude all cases that in practice are specified, and suggest a single coding possibility, whereas additional information must in fact be taken into account to code the case properly.

By contrast, associating the *A24.1*, *A24.2*, *A24.3* and *A24.4* codes will allow to retrieve effectively cases of melioidosis as a whole from medical records.

This representation of a single Orphanet entity by a defined set of ICD-10 codes defines for typing

purposes the BTNT type of coding (see later).

This possibility for multiple coding is extended to groups of diseases, when they can be represented accurately and comprehensively.

E.g. *ORPHA1047 Sideroblastic anemia* is represented by the following ICD-10 codes:

- *D64.0 Hereditary sideroblastic anaemia*
- *D64.1 Secondary sideroblastic anaemia due to disease*
- *D64.2 Secondary sideroblastic anaemia due to drugs and toxins*
- *D64.3 Other sideroblastic anaemias*

Using *D64 Other anaemias* and would be very imprecise.

## VI. Entities representable by a three-character ICD-10 code

Some entities have a good match with a three character code, with specific manifestations further represented in ICD-10 by subdivision into four-character codes. For instance, *ORPHA49 Cystic fibrosis* matches E84 Cystic fibrosis, which is further subdivided as follows:

*E84 Cystic fibrosis (Incl.: mucoviscidosis)*

- *E84.0 Cystic fibrosis with pulmonary manifestations*
- *E84.1 Cystic fibrosis with intestinal manifestations*
  - *Distal intestinal obstruction syndrome*
  - *Meconium ileus in cystic fibrosis+ (P75\*)*
  - *Excl.: meconium obstruction in cases where cystic fibrosis is known not to be present (P76.0)*
- *E84.8 Cystic fibrosis with other manifestations*
  - *Cystic fibrosis with combined manifestations*
- *E84.9 Cystic fibrosis, unspecified*

In such instances, the entity is to be coded in the Orphanet database by the whole set of possible four-character codes, rather than by the single three-character code. The rationale for this choice is to direct coders using Orphanet as a reference towards one of the four-character codes that they must actually use.

The use of a three-character code remains possible when there is no further subdivision in ICD-10, e.g. *ORPHA924 Acanthosis nigricans* represented by *L83 Acanthosis nigricans*, because there are no four-character codes for this disease.

## VII. Entities not mentioned in ICD-10

Many rare diseases are not mentioned at all in ICD-10, even as an index term. The coder must therefore interpret the ICD-10 to find the most appropriate representation.

The selection of the appropriate code is largely contextual. However, two kinds of criteria can most often be followed:

- the position of the entity within the Orphanet classifications of rare diseases
- the linearization selected for the entity

Codes can be considered as heritable by default throughout Orphanet classifications. This means that if a disease is mentioned in ICD-10, its coding will be inherited by all its subtypes, unless the ICD-10 explicitly tells otherwise

E.g. ORPHA685 Hereditary spastic paraplegia is coded G11.4 Hereditary spastic paraplegia. The code G11.4 is also used for the subentities :

- *ORPHA102012 Pure hereditary spastic paraplegia*
  - *ORPHA100980 Autosomal dominant pure spastic paraplegia*
  - *ORPHA100982 Autosomal recessive pure spastic paraplegia*
  - *ORPHA320332 X-linked pure spastic paraplegia*
- *ORPHA102013 Complex hereditary spastic paraplegia*
  - *ORPHA98888 X-linked complex spastic paraplegia*
  - *ORPHA100979 Autosomal dominant complex spastic paraplegia*
  - *ORPHA100981 Autosomal recessive complex spastic paraplegia*
  - *ORPHA320360 Maternally-inherited spastic paraplegia*
- *ORPHA320335 Pure or complex hereditary spastic paraplegia*
  - *ORPHA320342 Pure or complex autosomal dominant spastic paraplegia*
  - *ORPHA320346 Pure or complex autosomal recessive spastic paraplegia*
  - *ORPHA320350 Pure or complex X-linked spastic paraplegia*

and all subtypes of hereditary spastic paraplegia identified by numbers, further down in the hierarchy.

By extension, if a group of diseases can be coded by a precise ICD-10 code (e.g. *ORPHA98473 Muscular dystrophy* exactly matches *G71.0 Muscular dystrophy*), all subordinate entities can be presumed to be coded the same. This rule must however be mitigated by the fact that many rare disease entities are classified under several parents. It is then necessary to choose which parent in the Orphanet classification has priority: this is done according to the Orphanet linearization.

Structural groups in the Orphanet classification that are used only for clarity, without their own content, must be excluded from the heritability rule. This covers especially the groups named "Syndrome with [some sign or symptome]".



E.g. *ORPHA98683 Syndrome* with a symptomatic strabismus tells too little about the subordinated entities to warrant their coding as strabismus. The codings of the subordinated entities will be various, e.g. :

- *ORPHA480 Kearns-Sayre syndrome* → *H49.8 Other paralytic strabismus* (as inclusion term)
- *ORPHA908 Fragile X syndrome* → *Q99.2 Fragile X chromosome*
- *ORPHA861 Treacher-Collins syndrome* → *Q75.4 Mandibulofacial dysostosis* (as inclusion term)
- *ORPHA199 Cornelia de Lange syndrome* → *Q87.8 Congenital malformation syndromes predominantly associated with short stature* (as inclusion term)
- *ORPHA506 Leigh syndrome* → *Q31.8 Other specified degenerative diseases of nervous system* (as inclusion term)
- etc.

## Typing the codings

The Orphanet database uses the following annotating system for entries to precise the nature of the correspondance with the associated ICD code (or the several associated codes).

### I. Typing the matching between concepts

- **Exact match** (shorthand: **E**): the Orphanet entity and the ICD-10 code have the same range of application, they describe the same pathological entity.
- **Narrow term to broad term** (shorthand: **NTBT**): the Orphanet entity has a narrower range than the ICD-10 code used to represent it.
- **Broad term to narrow term** (shorthand: **BTNT**): the Orphanet entity has a broader range than the ICD-10 code used to represent it. In practice, this type of coding is only used when an Orphanet entity (most often groups of diseases or infectious diseases) can be exactly represented by a set of ICD-10 codes. (See before.)
- **Not yet decided/unable to decide** (shorthand: **ND**): reserved for for complex cases when the mapping cannot be qualified by any of the preceding labels.
- **Wrong match** (shorthand: **W**): reserved for coding errors with correction pending.

Matchings are done on a semantic rather than syntactic basis; they relate to the concept associated to the Orphanet entity, not with the name used to identify that concept.

A discrepancy of names does not preclude an exact match, if the ranges of the Orphanet entity and the ICD-10 code are intended to be the same.



E. g. *ORPHA 98878 Hemophilia A* is coded by *D66 Hereditary factor VIII deficiency*, and this is an exact match.

Conversely, similar names in ICD-10 and Orphanet do not automatically point to an exact match: the concepts must be screened in their hierarchical environment to check whether they are actually the same.

E.g. *ORPHA98503 Motor neuron disease* is not an exact match of *G12.2 Motor neuron disease*, even if the labels are the same. The ICD-10 entity excludes inherited spinal muscular atrophies, whereas the Orphanet entry includes them.

## II. Typing the ICD-10 term level

The level of the ICD terms associated with the ICD-10 code that match the Orphanet entity is also precised by the following annotations:

- **Specific code:** the Orphanet entity is matched by an ICD main term. This is normally associated with exact matches.
- **Inclusion term:** the Orphanet entity is not matched at ICD main term level, but only at the ICD inclusion term level.
- **Index term:** the Orphanet entity is not matched at ICD main term or inclusion term level, but only at the ICD index term level.
- **Attributed term:** the Orphanet entity has no matching term at all in ICD-10. Interpreting the ICD-10 is required to assign a code.

## III. Validating the codings

Once one or several codes are assigned to a disease, the coder must specify in the database whether the codings are **validated** or **not yet validated**.

**Validated** means that the coding is regarded as sure.

**Not yet validated** means that the coding is provisional and must be checked by a medical expert.

# Using the codings

## I. Coding indication and retrieval of medical records

Codes associated to a disease by an exact match (typed as *E*) can be used:

- to check which ICD-10 code must be used to represent the disease in medical records
- conversely, to retrieve cases of the disease from medical records (for instance to

establish statistics) using the ICD-10 code. Since the relation between the disease and the code is exact, such information can be retrieved in an easy, precise and reliable way.

Codes associated to a disease by a broad-to-narrow relation (typed as *BTNT*) also allow both those uses, though less directly:

- they give a hint towards the appropriate ICD-10 codes available for representing the disease in medical records. However, additional information is required to choose which code must precisely be selected.
- cases of the disease can be retrieved from medical records (for instance to establish statistics) putting together the full set of selected ICD-10 codes. BTNT mappings are allowed in the database only when this.

## II. Coding indication only

When the relation of the code to the disease is of a narrow-to-broad type (*NTBT*), the code can be used as a coding hint only. It is not possible in those instances to retrieve precise disease data from medical records using ICD-10 codes, because the code's scope extends to other, if related, conditions.

## III. Codes not to be used

Codes typed as wrong (*W*) must evidently not be used at all. They are hidden to users and only kept in the database to be later corrected by its managing team.

# Decisions for specific groups of diseases

## I. Tumours

Orphanet entities for tumours are a systemic challenge to ICD-10 coding, because their definition is primarily based on morphology, while ICD-10 uses tumoral behaviour (malignant, benign, uncertain or unknown) then topography as defining criteria, and uses the additional *ICD-O International Classification of Diseases for Oncology* to represent morphology.

The ICD-10 coding of Orphanet tumour entities will first depend of their behaviour, then on their topography.

If the tumour is usually benign, with only rare cases of malignant transformation, the code for the benign tumour can be used alone.

If the tumour has a high potential for malignant transformation, the code for the malignant tumour can be used alone.

In other instances, using two or three codes to represent possible behaviours (malignant, benign, uncertain or unknown) may be more helpful. This is sometimes explicitly warranted by ICD-10.

E.g. Thymoma is represented in ICD-10 by two index terms :

- *Thymoma (benign)* is ascribed to the code *D15.0 Benign neoplasm of other and unspecified intrathoracic organs:Thymus*
- *Thymoma - malignant* is ascribed to the code *C37 Malignant neoplasm of thymus*

Most ICD-10 codes for rare tumours are established by the information network *RARECARENet - Surveillance of Rare Cancers in Europe* as part of their and Orphanet's collaboration.

## II. Secreting tumours

Some secreting tumours can be coded both as a tumour and as the endocrine disorder caused by their secretion. In such cases, both codes are added into the database, with the mapping typed *ND*.

For instance, *ORPHA2965 Prolactinoma* is coded both *D35.2 Benign neoplasm: Pituitary gland* and *E22.1 Hyperprolactinaemia*.

## III. Cancer-predisposing syndromes

No explicit representation is available in ICD-10. They are coded like the relevant cancer, with the mapping typed *ND*.

## IV. Mitochondrial diseases

A few are mentioned as inclusion terms (e.g. Leigh syndrome, Kearns-Sayre syndrome), but most of them are not, and are generally problematic because of their multisystemic involvement.

The only relatively generic code for a mitochondrial disease is *G71.3 Mitochondrial myopathy, not elsewhere classified*. It can be used by default for all mitochondrial myopathies non explicitly mentioned in ICD-10.

In other instances, the code must be attributed according to the predominant involvement. Using the following is recommended:

- predominantly encephalopathic: use *G31.8 Other specified degenerative diseases of nervous system*
- encephalomyopathies with cerebral malformations (Fukuyama muscular dystrophy, muscle-eye-brain disease...) : use *Q04.8 Other specified malformations of brain*
- predominantly cardiomyopathic: *I42.2 Other hypertrophic cardiomyopathy*
- clinics dominated by metabolic anomalies: use *E88.8 Other specified metabolic disorders*
- no clear predominance: use *E88.8 Other specified metabolic disorders*

## V. Chromosomal microdeletions and microduplications

For all of them not explicitly mentioned in ICD-10, use:

- *Q92.3 Minor partial trisomy for microduplications*
- *Q93.5 Other deletions of part of a chromosome for microdeletions*

## VI. Cerebellar malformations

Not mentioned in ICD-10. They must by necessity be coded with *Q04.8 Other specified congenital malformations of brain*, however incorrect it is to refer to the cerebellum as “brain”.

## VII. Joubert syndrome, with or without other involvement

All forms are to be coded *Q04.3 Other reduction deformities of brain*.

## VIII. Spinocerebellar ataxias

The ICD-10 structure under *G11 Hereditary ataxia* is based on the age of onset:

- *G11.0 Congenital nonprogressive ataxia*
- *G11.1 Early-onset cerebellar ataxia* (Note: Onset usually before the age of 20)
- *G11.2 Late-onset cerebellar ataxia* (Note: Onset after before the age of 20)

The various types in Orphanet are defined mostly genetically, the age of onset can be variable and the genotype-phenotype correlations are not good.

## IX. Hereditary sensory and autonomic neuropathies

To be coded *G60.8 Other hereditary and idiopathic neuropathies*.

## X. Congenital bile acid synthesis defects

To be coded *K76.8 Other specified diseases of liver*.

## XI. Congenital disorders of glycosylation

To be coded *E77.8 Other disorders of glycoprotein metabolism*.

## XII. Susceptibility to infections

To be coded as immunodeficiencies, not as infections. Often appropriately coded by *D84.8 Other specified immunodeficiencies*.

## XIII. Distal arthrogyroses

To be coded *Q68.8 Other specified congenital musculoskeletal deformities*.

Please note that all data presented in this report are available for download at [Orphadata](#)

For any questions or comments, please contact us: [contact.orphanet@inserm.fr](mailto:contact.orphanet@inserm.fr)

Editors-in-chief: Ana Rath ● Editor: Bertrand Bellet ● Visual design: Julie Christ

*The correct form when quoting this document is:*

« Orphanet - ICD-10 coding rules for rare diseases », Orphanet procedural document, *Reports collection*, September 2014,  
<http://www.orpha.net/orphacom/cahiers/docs/GB/ICDcodingRulesProcedure.pdf>