

# Package ‘ICD10gm’

April 14, 2020

**Title** Metadata Processing for the German Modification of the ICD-10 Coding System

**Version** 1.1

**Date** 2020-04-08

**Description**

Provides convenient access to the German modification of the International Classification of Diagnoses, 10th revision (ICD-10-GM). It provides functionality to aid in the identification, specification and historisation of ICD-10 codes. Its intended use is the analysis of routinely collected data in the context of epidemiology, medical research and health services research. The underlying metadata are released by the German Institute for Medical Documentation and Information <<https://www.dimdi.de>>, and are redistributed in accordance with their license.

**Depends** R (>= 3.1.2)

**License** MIT + file LICENSE

**URL** <https://edonnachie.github.io/ICD10gm>,  
<https://doi.org/10.5281/zenodo.2542833>

**BugReports** <https://github.com/edonnachie/ICD10gm/issues>

**LazyData** true

**RoxygenNote** 7.0.2

**Imports** magrittr, dplyr, purrr, tidyr, tidyselect, stringi, rlang,  
tibble

**Suggests** testthat, knitr, rmarkdown, rvest

**VignetteBuilder** knitr

**Encoding** UTF-8

**Language** en-GB

**NeedsCompilation** no

**Author** Ewan Donnachie [aut, cre] (<<https://orcid.org/0000-0002-0668-0049>>)

**Maintainer** Ewan Donnachie <[ewan@edonnachie.net](mailto:ewan@edonnachie.net)>

**Repository** CRAN

**Date/Publication** 2020-04-14 12:20:14 UTC

## R topics documented:

get_icd_history	2
get_icd_labels	3
icd_expand	4
icd_history	5
icd_meta_blocks	6
icd_meta_chapters	7
icd_meta_codes	7
icd_meta_transition	9
icd_parse	10
icd_showchanges	12
icd_showchanges_icd3	12
is_icd_code	13
<b>Index</b>	<b>15</b>

---

get_icd_history	<i>Get ICD history metadata</i>
-----------------	---------------------------------

---

### Description

A utility function to query the `icd_meta_transition` table.

### Usage

```
get_icd_history(years = NULL, icd3 = NULL)
```

### Arguments

years	Year or years to get (numeric or character vector)
icd3	(optional) ICD codes to select (regular expression, matched exactly using <code>grep</code> )

### Details

Returns a data frame with ICD transition history, consisting of year, ICD code and label. Optional arguments allow selection of entries by year or ICD code. This is beneficial because the entire history is relatively large and rarely required in full.

### Value

data.frame, see `icd_hist`

### Examples

```
get_icd_history(years = 2009:2010, icd3 = "K52")
```

---

get_icd_labels	<i>Get or query ICD-10 labels</i>
----------------	-----------------------------------

---

### Description

A utility function to get or query [icd\\_meta\\_codes](#), returning a limited selection of ICD-10 codes and labels.

### Usage

```
get_icd_labels(year = NULL, icd3 = NULL, search = NULL, ...)
```

### Arguments

year	Year or years to get (numeric or character vector)
icd3	A character vector of three-digit ICD-10 codes to select
search	(optional) A string to search for in the label column using fuzzy matching (agrep)
...	(optional) Further arguments passed to grep when searching with icd_label

### Details

If an ICD code is provided as argument `icd3`, all corresponding codes and subcodes are returned. If a search term is provided, all codes are returned whose label matches the string approximately.

Returns a data frame with ICD metadata, consisting of year, ICD code and label. Optional arguments allow selection of entries by year, code or label. This is beneficial because the entire history is relatively large and rarely required in full.

### Value

`data.frame(year, icd3, icd_code, icd_normcode, icd_sub, label)`, see `icd_labels`

### Examples

```
get_icd_labels(year = 2019, icd3 = "I25")
get_icd_labels(year = 2019, search = "Asthma")
```

icd\_expand

*Expand list of ICD codes to include all possible subcodes***Description**

The function `icd_expand` takes a `data.frame` containing ICD codes and optional metadata as input. It returns a `data.frame` containing all ICD codes at or below the specified level of the hierarchy (e.g. the specification "E11" is expanded to include all three, four and five-digit codes beginning with E11).

**Usage**

```
icd_expand(
  icd_in,
  year,
  col_icd = "ICD",
  col_meta = NULL,
  type = "strict",
  ignore_icd_errors = FALSE
)
```

**Arguments**

<code>icd_in</code>	Data frame defining ICD codes of interest
<code>year</code>	ICD 10 version
<code>col_icd</code>	Column of <code>icd_in</code> containing ICD codes (Default: ICD)
<code>col_meta</code>	(Optional) Columns containing meta information to retain (e.g. Grouper, age or other criteria for later use). If left <code>NULL</code> , only <code>col_icd</code> is retained.
<code>type</code>	A character string determining how strictly matching should be performed, passed to <code>icd_parse</code> . This must be one of "strict" ( <code>str</code> contains a ICD code with no extraneous characters), bounded ( <code>str</code> contains an ICD code with a word boundary on both sides) or weak (ICD codes are extracted even if they are contained within a word, e.g. "E10Diabetes" would return "E10"). Default: <code>strict</code> .
<code>ignore_icd_errors</code>	logical. Whether to ignore incorrectly specified input (potentially leading to incomplete output) or stop if any ICD specification does not correspond to a valid ICD code. Default: <code>FALSE</code> , stop on error.

**Value**

`data.frame` with columns `YEAR`, `ICD_CODE`, `ICD_COMPRESSED`, `ICD_LABEL` and, if specified, columns specified by `col_meta`

**See Also**

[icd\\_history\(\)](#) to historize the output

**Examples**

```
# Incomplete or non-terminal codes expand to the right.
# This is useful to specified code blocks in a compact manner
icd_meta <- data.frame(ICD = "R1")
icd_expand(icd_meta, year = 2019)

# Optional metadata columns can be carried
# through with the specification
icd_meta <- data.frame(ICD = "M54", icd_label = "Back pain")
icd_expand(icd_meta, year = 2019, col_meta = "icd_label")
```

icd\_history

*Historize a list of ICD codes to cover the specified years***Description**

The function `icd_history` takes the result of `icd_expand`, specified for a particular year, and returns a `data.frame` containing all corresponding codes for the specified years (from 2003). To do this, it applies the ICD-10-GM transition tables to map codes between successive ICD-10-GM versions. Only automatic transitions are followed.

**Usage**

```
icd_history(icd_expand, years, custom_transitions = NULL)
```

**Arguments**

```
icd_expand      A data.frame (e.g. as generated by the function icd_expand)
years           Years to historize (e.g. 2005:2014)
custom_transitions
                (Optional) A data.frame containing custom transitions to complement the official transitions provided by icd_meta_transition.
```

**Value**

`data.frame` with columns `YEAR`, `ICD_CODE`, `ICD_COMPRESSED`, `ICD_LABEL` and, if specified, `DIAG_GROUP #` @seealso [icd\\_expand\(\)](#) to generate the necessary input

**Examples**

```
# Between 2018 and 2019, causalgia (G56.4) was reclassified
# under G90 as a complex regional pain syndrome
icd_meta <- data.frame(ICD = "G56.4", ICD_LABEL = "Causalgia")
icd_meta_expanded <- icd_expand(icd_meta, year = 2018, col_meta = "ICD_LABEL")
icd_history(icd_meta_expanded, years = 2018:2019)
```

---

icd_meta_blocks	<i>data.frame containing metadata for the ICD-10-GM code blocks</i>
-----------------	---

---

### Description

The ICD blocks (German: "Gruppen") constitute a level in the hierarchy between the chapters and the three-digit codes. The three-digit codes are grouped in sequence to form 240 groups that represent similar aetiological diagnoses. Unlike some other grouper systems, the ICD blocks do not consider similar diagnoses from different chapters of the ICD classification, for example chronic pain coded as a nonspecific symptom (R52.1) and as a somatoform disorder (F45.4).

### Usage

```
icd_meta_blocks
```

### Format

An object of class `data.frame` with 4071 rows and 6 columns.

### Details

**year** Year of validity (from 2004)

**icd\_block\_first** First three-digit ICD code in the block

**icd\_block\_last** Last three-digit ICD code in the block

**chapter** ICD-10 chapter to which the block belongs

**block\_label** Label for the block

**block\_id** Short label for the block in format "A00-A09"

### Source

The source data was downloaded from the official download centre of the German Institute for Medical Documentation and Information (DIMDI). See also <https://www.dimdi.de/dynamic/en/classifications/icd/icd-10-gm/tabular-list/structure/>

### See Also

Other ICD-10-GM metadata: [icd\\_meta\\_chapters](#), [icd\\_meta\\_codes](#), [icd\\_meta\\_transition](#)

---

icd\_meta\_chapters      *data.frame containing metadata for the ICD-10-GM chapters*

---

**Description**

The ICD chapters group codes according to their aetiology.

**Usage**

```
icd_meta_chapters
```

**Format**

An object of class `data.frame` with 357 rows and 4 columns.

**Details**

**year** Year of validity (from 2004)

**chapter** Chapter number (arabic numerals)

**chapter\_roman** Chapter number (Roman numerals)

**chapter\_label** Label for the chapter

**Source**

The source data was downloaded from the official download centre of the German Institute for Medical Documentation and Information (DIMDI). See also <https://www.dimdi.de/dynamic/en/classifications/icd/icd-10-gm/tabular-list/structure/>

**See Also**

Other ICD-10-GM metadata: [icd\\_meta\\_blocks](#), [icd\\_meta\\_codes](#), [icd\\_meta\\_transition](#)

---

icd\_meta\_codes      *data.frame containing metadata for all ICD-10-GM codes*

---

**Description**

DIMDI provide a CSV file with metadata on all valid codes. This table is read in with only minor modifications to facilitate changes between versions.

**Usage**

```
icd_meta_codes
```

**Format**

A data.frame containing the following variables:

**year** Year of validity (from 2004)

**level** Level of the hierarchy (3, 4 or 5 digits)

**terminal** Whether the code is a terminal code (i.e. with no further subcodes) (T: yes; N: no)

**subcode\_type** Whether the subcode is pre- or postcombined (X: precombined; S: postcombined). Precombined codes are listed directly under the three-digit ICD code, whereas post-combined codes are lists of possible values for the fourth and fifth digits that are not specific to the particular code (e.g. the group E10-E14 shares a common list of postcombined fourth and fifth digits)

**chapter\_nr** Chapter number (arabic digits 1-22)

**icd\_block\_first** First code in the respective ICD block, can be used to join with the table ICD10gm::icd\_meta\_blocks

**icd\_code** Full icd code (up to 7 characters) with all symbols except the "dagger" (for aetiological codes that can be combined with an "asterisk" code to denote the manifestation)

**icd\_normcode** The ICD "normcode", consisting of up to 6 characters and without all symbols except the period (e.g. E11.30)

**icd\_sub** Complete ICD code without any symbols or punctuation, consisting of up to 5 characters (e.g. E1130)

**label** ICD label for the complete code.

**label\_icd3** ICD label for the three-digit ICD code.

**label\_icd4** ICD label for fourth digit of the ICD code.

**label\_icd5** ICD label for the fifth digit of the ICD code.

**usage\_295** Usage of the code in the ambulatory sector (Paragraph 295 SGB V) (P: primary code; O: only as a "star" code in conjunction with a "dagger" code for aetiology; Z: only an optional "!" code in conjunction with a primary code; V: not to be used for coding)

**usage\_301** Usage of the code in the stationary (hospital) sector (Paragraph 301 SGB V) (P: primary code; O: only as a "star" code in conjunction with a "dagger" code for aetiology; Z: only an optional "!" code in conjunction with a primary code; V: not to be used for coding)

**mort\_list1** Key to join with the WHO mortality list 1

**mort\_list2** Key to join with the WHO mortality list 2

**mort\_list3** Key to join with the WHO mortality list 3

**mort\_list4** Key to join with the WHO mortality list 4

**morb\_list** Key to join with the WHO morbidity list

**gender\_specific** Whether the diagnosis is gender specific (M: male; W: female; 9: Not gender specific)

**gender\_error\_type** Type of error implied by the field gender\_specific (9: irrelevant; K: possible error)

**age\_min** Minimum age for which the diagnosis is plausible (T001: from one day; Y005: from five years)

**age\_max** Maximum age for which the diagnosis is plausible (T010: up to 10 days; Y005: up to five years)



**age\_error\_type** Type of error resulting from implausible age (9: irrelevant; M: always an error ("Muss-Fehler"); K: possible error ("Kann-Fehler"))

**rare\_in\_central\_europe** Indicates whether the diagnosis is rare in Central Europe (J: yes; N: no)

**code\_with\_content** Indicates whether the code has content associated with it (J: yes; N: no, leads to an error)

**notifiable** Indicates whether the diagnosis is notifiable in Germany (J: yes; N: no)

**notifiable\_lab** Indicates whether the diagnosis is notifiable for laboratories in Germany (J: yes; N: no)

## Details

This metadata is not suitable for operative coding and does not include all relevant information concerning the codes. For example, the file contains neither the inclusion and exclusion notes nor the detailed definitions (where present, mainly in Chapter V). DIMDI provide additional reference material for operative coding and detailed research.

The block U00-U49 contains reserved codes that can be allocated quickly for the documentation of new diseases or epidemiological phenomena. Such usage is allowed only when mandated by DIMDI. In particular, the codes may not be utilised on the initiative of other parties, for example, for clinical trials or contractual purposes. Notable uses of the reserved codes are for the Zika and COVID-19 viruses. These are included in the DIMDI online documentation, but not in the download files. They are therefore added manually to this data set as documented in the [package source](#).

Die Schlüsselnummern U05.0-U05.9 dieser Kategorie sollen ein schnelles Reagieren auf aktuelle epidemiologische Phänomene ermöglichen. Sie dürfen nur zusätzlich benutzt werden, um einen anderenorts klassifizierten Zustand besonders zu kennzeichnen. Die Schlüsselnummern dieser Kategorie dürfen nur über das Deutsche Institut für Medizinische Dokumentation und Information (DIMDI) mit Inhalten belegt werden; eine Anwendung für andere Zwecke ist nicht erlaubt. DIMDI wird den Anwendungszeitraum solcher Schlüsselnummern bei Bedarf bekannt geben.

## Source

The source data was downloaded from the official download centre of the German Institute for Medical Documentation and Information (DIMDI). See also <https://www.dimdi.de/dynamic/en/classifications/icd/icd-10-gm/tabular-list/#metadata>

## See Also

Other ICD-10-GM metadata: [icd\\_meta\\_blocks](#), [icd\\_meta\\_chapters](#), [icd\\_meta\\_transition](#)

---

icd_meta_transition	<i>data.frame</i> detailing the changes in ICD-10-GM codes between versions
---------------------	---

---

## Description

A data.frame providing old and new ICD codes (identical if no changes) and information as to whether the transition is automatic when transitioning forwards or backwards

**Usage**

```
icd_meta_transition
```

**Format**

An object of class `data.frame` with 215033 rows and 12 columns.

**Details**

**year\_from** Year of validity of the old code (from 2004)

**year\_to** Year of validity of the new code (from 2005)

**icd\_from** Old ICD code

**icd\_to** New ICD code

**automatic\_forward** Whether the transition is automatic in the forward direction (i.e. the old code can always be converted to the new code). (A: automatic, otherwise NA)

**automatic\_backward** Whether the transition is automatic in the backward direction (i.e. the new code can always be converted to the old code) (A: automatic, otherwise NA)

**change\_5** Whether the change relates to the fifth digit of the ICD-10 code (TRUE/FALSE).

**change\_4** Whether the change relates to the fourth digit of the ICD-10 code (TRUE/FALSE).

**change\_3** Whether the change relates to the three-digit ICD-10 code (TRUE/FALSE).

**icd3** The first three digits of `icd_from`.

**icd\_chapter** The first character of `icd_from` (i.e. the letter denoting the chapter).

**Source**

The source data was downloaded from the official download centre of the German Institute for Medical Documentation and Information (DIMDI). See also <https://www.dimdi.de/dynamic/en/classifications/icd/icd-10-gm/tabular-list/#crosswalks>

**See Also**

Other ICD-10-GM metadata: [icd\\_meta\\_blocks](#), [icd\\_meta\\_chapters](#), [icd\\_meta\\_codes](#)

---

```
icd_parse
```

*Extract all ICD codes from a character vector*

---

**Description**

An ICD code consists of, at a minimum, a three digit ICD-10 code (i.e. one upper-case letter followed by two digits). This may optionally be followed by a two digit subcode, selected punctuation symbols (cross "\*", dagger "U2020" or exclamation mark "!"). Both the period separating the three-digit code from the subcode, and the hyphen indicating an "incomplete" subcode, are optional. Finally, in the ambulatory system, an additional letter G, V, Z or A may be appended to signify the status ("security") of the diagnosis.

## Usage

```
icd_parse(str, type = "bounded", bind_rows = TRUE)
```

## Arguments

<code>str</code>	Character vector from which to extract all ICD codes
<code>type</code>	A character string determining how strictly matching should be performed. This must be one of "strict" ( <code>str</code> contains a ICD code with no extraneous characters), bounded ( <code>str</code> contains an ICD code with a word boundary on both sides) or weak (ICD codes are extracted even if they are contained within a word, e.g. "E10Diabetes" would return "E10"). Default: bounded.
<code>bind_rows</code>	logical. Whether to convert the matrix output of <code>stringi::stri_match_all</code> to a <code>data.frame</code> , with additional <code>icd_sub</code> to uniquely represent the code and allow lookup of the code

## Details

By default, the function returns a `data.frame` containing the matched codes and the standardised three digit code (`icd3`), subcode (`icd_subcode`), normcode (`icd_norm`) and code without period (`icd_sub`).

If `bind_rows = FALSE`, the list output of `stringi::stri_match_all_regex` is returned. This is particularly useful to retrieve the matches from each element of the `str` vector separately.

## Value

`data.frame` (if `bind_rows = TRUE`) or `matrix`

## See Also

[is\\_icd\\_code\(\)](#)

## Examples

```
icd_parse("E11.7")
icd_parse("Depression: F32")
icd_parse(c(
  "Backpain (M54.9) is one of the most common diagnoses in primary care",
  "Codes for chronic pain include R52.1 and F45.4"
))
```

---

icd\_showchanges      *Show historical changes to selected ICD-10-GM codes*

---

### Description

Show all changes in ICD history relating to the 3-digit codes contained in the data.frame icd\_in. The output of icd\_expand can be passed directly to this function to display relevant changes.

### Usage

```
icd_showchanges(icd_in, col_icd = "icd_sub")
```

### Arguments

icd\_in              Data frame defining ICD codes of interest  
col\_icd             Column of icd\_in containing ICD codes (Default: ICD)

### Value

data.frame with columns YEAR, ICD\_CODE, ICD\_LABEL and, if specified, DIAG\_GROUP

### See Also

[icd\\_showchanges\\_icd3\(\)](#) to provide one or more three-digit codes as input

### Examples

```
dat_icd <- icd_expand(
  data.frame(ICD_SPEC = c("K52.9")),
  col_icd = "ICD_SPEC",
  year = 2019)
icd_showchanges(dat_icd)
```

---

icd\_showchanges\_icd3      *Show historical changes to selected three-digit ICD-10-GM codes*

---

### Description

Show all changes in ICD history relating to the 3-digit codes contained in a given vector icd

### Usage

```
icd_showchanges_icd3(icd3)
```

**Arguments**

icd3                    Vector of three-digit ICD codes

**Value**

data.frame with columns YEAR, ICD\_CODE, ICD\_LABEL and, if specified, DIAG\_GROUP

**See Also**

[icd\\_showchanges\(\)](#) if the input has been generated by [icd\\_expand\(\)](#)

**Examples**

```
icd_showchanges_icd3("A09")
```

---

is_icd_code	<i>Test whether a string is a valid ICD code</i>
-------------	--

---

**Description**

An ICD code consists of, at a minimum, a three digit ICD-10 code (i.e. one upper-case letter followed by two digits). This may optionally be followed by a two digit subcode, selected punctuation symbols (cross "\*", dagger "U2020" or exclamation mark "!"). Both the period separating the three-digit code from the subcode, and the hyphen indicating an "incomplete" subcode, are optional. Finally, in the ambulatory system, an additional letter G, V, Z or A may be appended to signify the status ("security") of the diagnosis.

**Usage**

```
is_icd_code(str, year = NULL, parse = TRUE)
```

**Arguments**

str                    Character vector to be tested

year                   Year for which to test whether the specification is a valid code. Default: NULL (test whether str matches a code from any year since 2003)

parse                  logical. Whether to first parse the input str using [icd\\_parse](#) (Default: TRUE). If FALSE, assumes that str is already formatted as [icd\\_sub](#) (i.e. without separating period or other punctuation)

**Value**

Logical vector the same length as the character input

**See Also**

[icd\\_parse\(\)](#)

**Examples**

```
is_icd_code("A09.9")
is_icd_code("A099")
is_icd_code("A09.9-")
```

```
is_icd_code("AA9")
```

```
# The following code is syntactically correct but
# has never been in use
is_icd_code("E15.9")
```

# Index

## \*Topic **datasets**

- icd\_meta\_blocks, [6](#)
- icd\_meta\_chapters, [7](#)
- icd\_meta\_codes, [7](#)
- icd\_meta\_transition, [9](#)

- get\_icd\_history, [2](#)
- get\_icd\_labels, [3](#)

- icd\_expand, [4](#)
- icd\_expand(), [5](#), [13](#)
- icd\_history, [5](#)
- icd\_history(), [4](#)
- icd\_meta\_blocks, [6](#), [7](#), [9](#), [10](#)
- icd\_meta\_chapters, [6](#), [7](#), [9](#), [10](#)
- icd\_meta\_codes, [3](#), [6](#), [7](#), [7](#), [10](#)
- icd\_meta\_transition, [2](#), [6](#), [7](#), [9](#), [9](#)
- icd\_parse, [10](#)
- icd\_parse(), [13](#)
- icd\_showchanges, [12](#)
- icd\_showchanges(), [13](#)
- icd\_showchanges\_icd3, [12](#)
- icd\_showchanges\_icd3(), [12](#)
- is\_icd\_code, [13](#)
- is\_icd\_code(), [11](#)