Title  Metadata Processing for the German Modification of the ICD-10 Coding System
Version 1.2.2
Date 2021-04-09
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
BugReports https://github.com/edonnachie/ICD10gm/issues/
LazyData true
LazyDataCompression xz
RoxygenNote 7.1.1
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@donnachie.net>
Repository CRAN
Date/Publication 2021-04-09 22:00:02 UTC
R topics documented:

charlson_rcs ................................................................. 2
charlson_sundararajan .................................................. 3
get_icd_history ............................................................ 4
get_icd_labels ............................................................. 4
icd_expand ................................................................. 5
icd_history ............................................................... 6
icd_meta_blocks ........................................................... 7
icd_meta_chapters ....................................................... 8
icd_meta_codes ............................................................ 9
icd_meta_transition .................................................... 11
icd_parse ................................................................. 12
icd_showchanges .......................................................... 13
icd_showchanges_icd3 .................................................. 14
is_icd_code .............................................................. 14

Index 16

| charlson_rcs | Charlson Comorbidities (Royal College of Surgeons) |

Description

Specification of the Charlson comorbidity index in the version of the Royal College of Surgeons (2010).

Usage

charlson_rcs

Format

An object of class spec_tbl_df (inherits from tbl_df, tbl, data.frame) with 121 rows and 2 columns.

Details

The specification can be expanded using the icd_expand function to return all corresponding ICD-10-GM codes.

This table was created on the basis of the publication referenced below. It is provided as is with no guarantee of accuracy. Furthermore, the applicability of the codes in the context of the German ICD-10-GM is unclear.

Condition Disease entity

ICD_SPEC Secification of the corresponding ICD-10 codes, suitable for input to icd_expand
Source
doi: 10.1002/bjs.6930

See Also
Other Charlson: charlson_sundararajan

---

charlson_sundararajan  

Charlson Comorbidities (Sundararajan)

Description
Specification of the Charlson comorbidity index in the version of Sundararahan et al. (2004).

Usage
charlson_sundararajan

Format
An object of class spec_tbl_df (inherits from tbl_df, tbl, data.frame) with 17 rows and 3 columns.

Details
The specification can be expanded using the icd_expand function to return all corresponding ICD-10-GM codes.

This table was created on the basis of the publication referenced below. It is provided as is with no guarantee of accuracy. Furthermore, the applicability of the codes in the context of the German ICD-10-GM is unclear.

Condition  Disease entity
Weight  Controbition of the disease entity towards the combined comorbidity index
ICD_10_AM  Specification of the corresponding ICD-10 codes, suitable for input to icd_expand

Source
doi: 10.1016/j.jclinepi.2004.03.012

See Also
Other Charlson: charlson_rcs
get_icd_history

Get ICD history metadata

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 grep)

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

Get or query ICD-10 labels

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, ...)

**icd_expand**

### 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 agrep 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

```r
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

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

Arguments

icd_in      Data frame defining ICD codes of interest
year        ICD 10 version
col_icd     Column of icd_in containing ICD codes (Default: ICD)
col_meta    (Optional) Columns containing meta information to retain (e.g. Grouper, age or other criteria for later use). If left NULL, only col_icd is retained.
type        A character string determining how strictly matching should be performed, passed to icd_parse. This must be one of "strict" (str contains a ICD code with no extraneous characters), bounded (str 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: strict.
ignore_icd_errors 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: FALSE, 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**

data.frame containing metadata for the ICD-10-GM code blocks

**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 unspecific symptom (R52.1) and as a somatoform disorder (F45.4).

**Usage**

```
icd_meta_blocks
```

**Format**

An object of class data.frame with 4329 rows and 6 columns.
`icd_meta_chapters`

**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/](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 396 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/](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 postcombinated (X: precombinated; S: postcombinated). Precombinated codes are listed directly under the three-digit ICD code, whereas postcombinated 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 postcombinated 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 | data.frame 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 228848 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 forward 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

str Character vector from which to extract all ICD codes

type A character string determining how strictly matching should be performed. This must be one of "strict" (str contains a ICD code with no extraneous characters), bounded (str 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.

bind_rows logical. Whether to convert the matrix output of stringi::stri_match_all to a data.frame, with additional icd_sub 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.
icd_showchanges

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  Test whether a string is a valid ICD code

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)
**is_icd_code**

**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**

```r
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

* Charlson
  charlson_rcs, 2
  charlson_sundararajan, 3

* ICD-10-GM metadata
  icd_meta_blocks, 7
  icd_meta_chapters, 8
  icd_meta_codes, 9
  icd_meta_transition, 11

* datasets
  charlson_rcs, 2
  charlson_sundararajan, 3
  icd_meta_blocks, 7
  icd_meta_chapters, 8
  icd_meta_codes, 9
  icd_meta_transition, 11

charlson_rcs, 2, 3
charlson_sundararajan, 3, 3

get_icd_history, 4
get_icd_labels, 4

icd_expand, 2, 3, 5
icd_expand(), 7, 14
icd_history, 6
icd_history(), 6
icd_meta_blocks, 7, 9, 11, 12
icd_meta_chapters, 8, 8, 11, 12
icd_meta_codes, 4, 8, 9, 9, 12
icd_meta_transition, 4, 8, 9, 11, 11
icd_parse, 12
icd_parse(), 15
icd_showchanges, 13
icd_showchanges(), 14
icd_showchanges_icd3, 14
icd_showchanges_icd3(), 13
is_icd_code, 14
is_icd_code(), 13