Package ‘ICD10gm’

August 22, 2019

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

**Version**  1.0.4

**Date**  2019-08-21

**Description**  Provides convenient access to the German modification of the International Classification of Diseases, 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://github.com/edonnachie/ICD10gm,
https://doi.org/10.5281/zenodo.2542833

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

**LazyData**  true

**RoxygenNote**  6.1.1

**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**  2019-08-21 22:00:02 UTC
get_icd_history

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.

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)

Value

data.frame, see icd_hist

Examples

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

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.

Description

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.

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 agrep when searching with icd_label

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)
icd_history

Historize a list of ICD codes to cover the specified years

```
# 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")
```

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

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")
```

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

Examples

```r
# 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

**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 categories. Sequential codes are grouped to form 240 groups that represent similar aetiological diagnoses. Unlike other grouper systems, the ICD blocks do not consider similar diagnoses from different chapters of the ICD classification, for example chronic pain as a unspecific symptom (R52.1) and as a somatoform disorder (F45.4).

**Usage**

```r
icd_meta_blocks
```

**Format**

An object of class `data.frame` with 3831 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/

<table>
<thead>
<tr>
<th>icd_meta_chapters</th>
<th>Metadata for the ICD-10-GM chapters</th>
</tr>
</thead>
</table>

Description

The ICD chapters group codes according to their aetiology.

Usage

icd_meta_chapters

Format

An object of class data.frame with 336 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/

<table>
<thead>
<tr>
<th>icd_meta_codes</th>
<th>Metadata for all ICD-10-GM codes</th>
</tr>
</thead>
</table>

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
icd_meta_codes 7

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**  The ICD "normcode", consisting of up to 5 characters and without any symbols (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)
**icd_meta-transition**  
Transition between ICD-10-GM 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 201295 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)
ICD codes can be extracted from a character vector using the `icd_parse` function. The function supports different matching modes: strict, bounded, or weak.

**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](https://www.dimdi.de/dynamic/en/classifications/icd/icd-10-gm/tabular-list/#crosswalks).

**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.

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

Examples

```r
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 icd_showchanges
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.
```

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

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

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

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

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, 5
   icd_meta_chapters, 6
   icd_meta_codes, 6
   icd_meta_transition, 8

get_icd_history, 2
get_icd_labels, 3

icd_expand, 3
icd_history, 4
icd_meta_blocks, 5
icd_meta_chapters, 6
icd_meta_codes, 6
icd_meta_transition, 8
icd_parse, 9
icd_showchanges, 10
icd_showchanges_icd3, 11
is_icd_code, 11