

Working with World Health Organization ICD codes and national editions, and different languages

Jack O. Wasey

`icd` 4.0+ offers some international support.

- World Health Organization (WHO) codes for 2016 in English, the latest release.
- World Health Organization (WHO) codes for 2008 in French, the latest available online data.
- ICD-10-FR in French, with modifications as used in France.
- Translations of the US ICD-10-CM codes in French and Dutch, which are used in Belgium.

All of this should be considered beta quality, and I would much appreciate testing and bug reports. I would especially be interested if someone has experience internationalizing an R package like this.

Function and variable names are likely to change in the future. The current versions are more consistent with naming of data files.

WHO

Surprisingly, there are potential copyright restrictions on redistributing the WHO ICD-9 and ICD-10 definitions, which form the basis of ICD code systems around the world. These definitions are themselves also used as the internationally standard way to codify morbidity and mortality for public health.

`icd` asks the user whether they wish to download the data from the WHO web site, whereupon it is made available as functions, named similarly to existing data, e.g., `get_icd10who2016()`, which is the latest ICD-10 release from the WHO at the time of writing. For French codes, `get_icd10fr2016()`, with the localized synonym `get_cim10fr2016()` having internationalized column names, but this would duplicate the data making package storage and loading bigger, which is increasingly a problem as more data is included.

Unfortunately, the availability and URL of the WHO resources sometimes change. If you see errors, check that the WHO resource is available at that time, and, if you think the URL has been changed, then please file an issue.

```
# if not already done, set up the downloaded data cache:
setup_icd_data()
# or setup and optionally download everything at once (currently ~350 MB)
# and a few minutes to generate the data structures.
download_all_icd_data()

get_icd10who2016()[1:5, ]
#>   code leaf          desc three_digit
#> 1  A00 FALSE          Cholera          A00
#> 2 A000 TRUE Cholera due to Vibrio cholerae 01, biovar cholerae A00
#> 3 A001 TRUE Cholera due to Vibrio cholerae 01, biovar eltor   A00
#> 4 A009 TRUE          Cholera, unspecified A00
#> 5  A01 FALSE Typhoid and paratyphoid fevers A01
#>      major sub_sub_chapter      sub_chapter
#> 1 Cholera <NA> Intestinal infectious diseases
#> 2 Cholera <NA> Intestinal infectious diseases
```

```

#> 3 Cholera <NA> Intestinal infectious diseases
#> 4 Cholera <NA> Intestinal infectious diseases
#> 5 Typhoid and paratyphoid fevers <NA> Intestinal infectious diseases
#> chapter
#> 1 Certain infectious and parasitic diseases
#> 2 Certain infectious and parasitic diseases
#> 3 Certain infectious and parasitic diseases
#> 4 Certain infectious and parasitic diseases
#> 5 Certain infectious and parasitic diseases
get_icd10fr2019()[1:5, ]
#> code short_desc
#> 1 A00 CHOLERA
#> 2 A000 CHOLERA A VIBRIO CHOLERAE 01, BIOVAR CHOLERAE
#> 3 A001 CHOLERA A VIBRIO CHOLERAE 01, BIOVAR EL TOR
#> 4 A009 CHOLERA, SAI
#> 5 A01 FIEVRES TYPHOIDE ET PARATYPHOIDE
#> long_desc
#> 1 Choléra
#> 2 Choléra à Vibrio cholerae 01, biovar cholerae
#> 3 Choléra à Vibrio cholerae 01, biovar El Tor
#> 4 Choléra, sans précision
#> 5 Fièvres typhoïde et paratyphoïde
#> major three_digit
#> 1 Choléra A00
#> 2 Choléra A00
#> 3 Choléra A00
#> 4 Choléra A00
#> 5 Fièvres typhoïde et paratyphoïde A01
summary(get_icd10who2016())
#> code leaf desc three_digit
#> Length:12187 Mode :logical Length:12187 A04 : 11
#> Class1:icd10who FALSE:1580 Class :character A06 : 11
#> Class2:icd10 TRUE :10607 Mode :character A15 : 11
#> Class3:character A66 : 11
#> Mode :character B20 : 11
#> B37 : 11
#> (Other):12121
#> major sub_sub_chapter sub_chapter chapter
#> Length:12187 Length:12187 Length:12187 Length:12187
#> Class :character Class :character Class :character Class :character
#> Mode :character Mode :character Mode :character Mode :character
#>
#>
#>
#>
#> code leaf desc three_digit
#> 1 A00 FALSE Cholera A00
#> 2 A000 TRUE Cholera due to Vibrio cholerae 01, biovar cholerae A00
#> 3 A001 TRUE Cholera due to Vibrio cholerae 01, biovar eltor A00
#> 4 A009 TRUE Cholera, unspecified A00
#> 5 A01 FALSE Typhoid and paratyphoid fevers A01
#> major sub_sub_chapter sub_chapter
#> 1 Cholera <NA> Intestinal infectious diseases

```

```

#> 2 Cholera <NA> Intestinal infectious diseases
#> 3 Cholera <NA> Intestinal infectious diseases
#> 4 Cholera <NA> Intestinal infectious diseases
#> 5 Typhoid and paratyphoid fevers <NA> Intestinal infectious diseases
#> chapter
#> 1 Certain infectious and parasitic diseases
#> 2 Certain infectious and parasitic diseases
#> 3 Certain infectious and parasitic diseases
#> 4 Certain infectious and parasitic diseases
#> 5 Certain infectious and parasitic diseases
#> code short_desc
#> 1 A00 CHOLERA
#> 2 A000 CHOLERA A VIBRIO CHOLERAEE 01, BIOVAR CHOLERAEE
#> 3 A001 CHOLERA A VIBRIO CHOLERAEE 01, BIOVAR EL TOR
#> 4 A009 CHOLERA, SAI
#> 5 A01 FIEVRES TYPHOIDE ET PARATYPHOIDE
#> long_desc
#> 1 Choléra
#> 2 Choléra à Vibrio cholerae 01, biovar cholerae
#> 3 Choléra à Vibrio cholerae 01, biovar El Tor
#> 4 Choléra, sans précision
#> 5 Fièvres typhoïde et paratyphoïde
#> major three_digit
#> 1 Choléra A00
#> 2 Choléra A00
#> 3 Choléra A00
#> 4 Choléra A00
#> 5 Fièvres typhoïde et paratyphoïde A01
#> code leaf desc three_digit
#> Length:12187 Mode :logical Length:12187 A04 : 11
#> Class1:icd10who FALSE:1580 Class :character A06 : 11
#> Class2:icd10 TRUE :10607 Mode :character A15 : 11
#> Class3:character A66 : 11
#> Mode :character B20 : 11
#> B37 : 11
#> (Other):12121
#> major sub_sub_chapter sub_chapter chapter
#> Length:12187 Length:12187 Length:12187 Length:12187
#> Class :character Class :character Class :character Class :character
#> Mode :character Mode :character Mode :character Mode :character
#>
#>
#>
#>

```

French edition of ICD-10

ICD-10-FR is significantly different from the WHO version, and is used by official bodies and health systems in France. As with all the features in this document, this is an early attempt at broadening the international capabilities of this package.

```

fr <- get_icd10fr2019()
fr[fr$code == "C43", ]
#> code short_desc long_desc

```

```
#> 1149 C43 MELANOME MALIN DE LA PEAU Mélanome malin de la peau
#>                                     major three_digit
#> 1149 Mélanome malin de la peau C43
```

Working with different ICD-10-CM annual versions

There is an initial beta-quality mechanism for selecting a particular version of ICD-10-CM to use. This is a US coding system, but Belgian authorities provide translations into French and Dutch, and appear to have made no modifications.

```
# get the active version string
get_icd10cm_active_year()
#> [1] "2019"
# get the data itself:
devnull <- get_icd10cm_active()
# set the active version, and check it is the one we expected
set_icd10cm_active_year("2015")
identical(get_icd10cm_active(), get_icd10cm2015())
#> [1] TRUE
```

The function `with_icd10cm_version` allows temporarily using a particular data set for a computation, analogous to the functions in the `withr` package:

```
# The code "C4311" goes from being a leaf in 2018, to a parent in 2019
with_icd10cm_version(
  ver = "2018",
  is_leaf(as.icd10cm("C4311"))
)
#> [1] TRUE
with_icd10cm_version(
  ver = "2019",
  is_leaf(as.icd10cm("C4311"))
)
#> [1] FALSE
# In 2018 the following code was not even defined, but in 2019 was a new child
# code of "C4311"s
with_icd10cm_version(ver = "2018", {
  print(is_defined(as.icd10cm("C4311")))
  print(is_leaf(as.icd10cm("C4311")))
})
#> [1] TRUE
#> [1] TRUE
with_icd10cm_version(
  ver = "2019",
  is_leaf("C43111")
)
#> [1] TRUE
```