

# Package ‘comorbidity’

June 27, 2019

**Type** Package

**Version** 0.4.1

**Title** Computing Comorbidity Scores

**Description** Computing comorbidity scores such as the weighted Charlson score (Charlson, 1987 <doi:10.1016/0021-9681(87)90171-8>) and the Elixhauser comorbidity score (Elixhauser, 1998 <doi:10.1097/00005650-199801000-00004>) using ICD-9-CM or ICD-10 codes (Quan, 2005 <doi:10.1097/01.mlr.0000182534.19832.83>).

**Maintainer** Alessandro Gasparini <ag475@leicester.ac.uk>

**URL** <https://github.com/ellessenne/comorbidity>

**BugReports** <https://github.com/ellessenne/comorbidity/issues>

**License** GPL (>= 3)

**Depends** R (>= 2.10)

**Imports** checkmate, parallel, stats, stringi, utils

**Suggests** covr, knitr, rmarkdown, testthat

**RoxygenNote** 6.1.1

**Encoding** UTF-8

**LazyData** true

**ByteCompile** true

**VignetteBuilder** knitr

**Language** en\_GB

**NeedsCompilation** no

**Author** Alessandro Gasparini [aut, cre]  
(<<https://orcid.org/0000-0002-8319-7624>>),  
Hojjat Salmasian [ctb] (ICD-9-CM scores)

**Repository** CRAN

**Date/Publication** 2019-06-27 13:00:03 UTC

## R topics documented:

|                               |           |
|-------------------------------|-----------|
| comorbidity-package . . . . . | 2         |
| australia10 . . . . .         | 2         |
| comorbidity . . . . .         | 3         |
| icd10cm_2017 . . . . .        | 7         |
| icd10cm_2018 . . . . .        | 7         |
| icd10_2009 . . . . .          | 8         |
| icd10_2011 . . . . .          | 9         |
| icd9_2015 . . . . .           | 9         |
| nhds2010 . . . . .            | 10        |
| sample_diag . . . . .         | 11        |
| <b>Index</b>                  | <b>12</b> |

---

|                     |                                     |
|---------------------|-------------------------------------|
| comorbidity-package | <i>Computing Comorbidity Scores</i> |
|---------------------|-------------------------------------|

---

### Description

Computing comorbidity scores such as the weighted Charlson score (Charlson, 1987 doi: [10.1016/00219681\(87\)901718](https://doi.org/10.1016/00219681(87)901718)) and the Elixhauser comorbidity score (Elixhauser, 1998 doi: [10.1097/00005650-19980100000004](https://doi.org/10.1097/00005650-19980100000004)) using ICD-9-CM or ICD-10 codes (Quan, 2005 doi: [10.1097/01.mlr.0000182534.19832.83](https://doi.org/10.1097/01.mlr.0000182534.19832.83)).

### Author(s)

Alessandro Gasparini (ag475@leicester.ac.uk)

---

|             |  |
|-------------|--|
| australia10 | <i>Australian mortality data, 2010</i> |
|-------------|--|

---

### Description

A dataset containing australian mortality data, obtained from Stata version 15.

### Usage

```
australia10
```

### Format

A data frame with 3,322 rows and 3 variables:

**cause** ICD-10 code representing cause of death

**sex** Gender

**deaths** Number of deaths

**Note**

The R code used to download and process the dataset from Stata is available [here](#).

---

 comorbidity

*Compute comorbidity scores.*


---

**Description**

Computes comorbidity scores such as the weighted Charlson score and the Elixhauser comorbidity score.

**Usage**

```
comorbidity(x, id, code, score, icd = "icd10", assign0,
  factorise = FALSE, labelled = TRUE, tidy.codes = TRUE,
  parallel = FALSE, mc.cores = parallel::detectCores())
```

**Arguments**

|         |  |
|---------|--|
| x       | A tidy data frame with one column containing an individual ID and a column containing all diagnostic codes.  |
| id      | Column of x containing the individual ID.  |
| code    | Column of x containing diagnostic codes. Codes must be in upper case with no punctuation in order to be properly recognised.   |
| score   | The comorbidity score to compute. Possible choices are the weighted Charlson score ( <code>charlson</code> ) and the weighted Elixhauser score ( <code>elixhauser</code> ). Values are case-insensitive.   |
| icd     | The version of ICD coding to use. Possible choices are ICD-9-CM ( <code>icd9</code> ) or ICD-10 ( <code>icd10</code> ). Defaults to <code>icd10</code> , and values are case-insensitive.  |
| assign0 | Apply a hierarchy of comorbidities. If TRUE, should a comorbidity be present in a patient with different degrees of severity, then the milder form will be assigned to 0 and therefore not counted. By doing this, a type of comorbidity is not counted more than once in each patient. In particular, the comorbidities that are affected by this argument are: <ul style="list-style-type: none"> <li>• "Mild liver disease" (<code>mld</code>) and "Moderate/severe liver disease" (<code>msld</code>) for the Charlson score;</li> <li>• "Diabetes" (<code>diab</code>) and "Diabetes with complications" (<code>diabwc</code>) for the Charlson score;</li> <li>• "Cancer" (<code>canc</code>) and "Metastatic solid tumour" (<code>metacanc</code>) for the Charlson score;</li> <li>• "Hypertension, uncomplicated" (<code>hypunc</code>) and "Hypertension, complicated" (<code>hypc</code>) for the Elixhauser score;</li> <li>• "Diabetes, uncomplicated" (<code>diabunc</code>) and "Diabetes, complicated" (<code>diabc</code>) for the Elixhauser score;</li> </ul> |

|            |  |
|------------|--|
|            | <ul style="list-style-type: none"> <li>• "Solid tumour" (solidtum) and "Metastatic cancer" (metacanc) for the Elixhauser score.</li> </ul>   |
| factorise  | Return comorbidities as factors rather than numeric, where (1 = presence of comorbidity, 0 = otherwise). Defaults to FALSE.  |
| labelled   | Attach labels to each comorbidity, compatible with the RStudio viewer via the <code>utils::View()</code> function. Defaults to TRUE.   |
| tidy.codes | Tidy diagnostic codes? If TRUE, all codes are converted to upper case and all non-alphanumeric characters are removed using the regular expression <code>[^[:alnum:]]</code> . Defaults to TRUE. |
| parallel   | Run the computation in parallel? See the Notes section for more information. Defaults to FALSE.  |
| mc.cores   | The number of cores to use when running the computations in parallel. Defaults to all available cores.   |

### Details

The ICD-10 and ICD-9-CM coding for the Charlson and Elixhauser scores is based on work by Quan *et al.* (2005). Weights for the Charlson score are based on the original formulation by Charlson *et al.* in 1987, while weights for the Elixhauser score are based on work by Moore *et al.* and van Walraven *et al.* Finally, the categorisation of scores and weighted scores is based on work by Menendez *et al.* See `vignette("comorbiditiescores", package = "comorbidity")` for further details on the comorbidity scores and the weighting algorithm. ICD-10 and ICD-9 codes must be in upper case and with alphanumeric characters only in order to be properly recognised; set `tidy.codes = TRUE` to properly tidy the codes automatically. As a convenience, a message is printed to the R console when non-alphanumeric characters are found. To run the calculations in parallel set `parallel = TRUE`. This is based on `parallel::parSapply()`, and it is possible to set the number of cores to use via the `mc.cores` argument, which defaults to using all the cores available.

### Value

A data frame with `id`, columns relative to each comorbidity domain, comorbidity score, weighted comorbidity score, and categorisations of such scores, with one row per individual.

For the Charlson score, the following variables are included in the dataset:

- The `id` variable as defined by the user;
- `ami`, for acute myocardial infarction;
- `chf`, for congestive heart failure;
- `pvd`, for peripheral vascular disease;
- `cevd`, for cerebrovascular disease;
- `dementia`, for dementia;
- `copd`, chronic obstructive pulmonary disease;
- `rheumd`, for rheumatoid disease;
- `pud`, for peptic ulcer disease;
- `mld`, for mild liver disease;

- diab, for diabetes without complications;
- diabwc, for diabetes with complications;
- hp, for hemiplegia or paraplegia;
- rend, for renal disease;
- canc, for cancer (any malignancy);
- msld, for moderate or severe liver disease;
- metacanc, for metastatic solid tumour;
- aids, for AIDS/HIV;
- score, for the non-weighted version of the Charlson score;
- index, for the non-weighted version of the grouped Charlson index;
- wscore, for the weighted version of the Charlson score;
- windex, for the weighted version of the grouped Charlson index.

Conversely, for the Elixhauser score the dataset contains the following variables:

- The id variable as defined by the user;
- chf, for congestive heart failure;
- carit, for cardiac arrhythmias;
- valv, for valvular disease;
- pcd, for pulmonary circulation disorders;
- pvd, for peripheral vascular disorders;
- hypunc, for hypertension, uncomplicated;
- hypc, for hypertension, complicated;
- para, for paralysis;
- ond, for other neurological disorders;
- cpd, for chronic pulmonary disease;
- diabunc, for diabetes, uncomplicated;
- diabc, for diabetes, complicated;
- hypothy, for hypothyroidism;
- rf, for renal failure;
- ld, for liver disease;
- pud, for peptic ulcer disease, excluding bleeding;
- aids, for AIDS/HIV;
- lymph, for lymphoma;
- metacanc, for metastatic cancer;
- solidtum, for solid tumour, without metastasis;
- rheumd, for rheumatoid arthritis/collaged vascular disease;
- coag, for coagulopathy;
- obes, for obesity;

- `wloss`, for weight loss;
- `fed`, for fluid and electrolyte disorders;
- `blane`, for blood loss anaemia;
- `dane`, for deficiency anaemia;
- `alcohol`, for alcohol abuse;
- `drug`, for drug abuse;
- `psycho`, for psychoses;
- `depre`, for depression;
- `score`, for the non-weighted version of the Elixhauser score;
- `index`, for the non-weighted version of the grouped Elixhauser index;
- `wscore_ahrq`, for the weighted version of the Elixhauser score using the AHRQ algorithm (Moore *et al.*, 2017);
- `wscore_vw`, for the weighted version of the Elixhauser score using the algorithm in van Walraven *et al.* (2009);
- `windex_ahrq`, for the weighted version of the grouped Elixhauser index using the AHRQ algorithm (Moore *et al.*, 2017);
- `windex_vw`, for the weighted version of the grouped Elixhauser index using the algorithm in van Walraven *et al.* (2009).

Labels are presented to the user when using the RStudio viewer (e.g. via the `utils::View()` function) for convenience.

## References

- Quan H, Sundararajan V, Halfon P, Fong A, Burnand B, Luthi JC, et al. *Coding algorithms for defining comorbidities in ICD-9-CM and ICD-10 administrative data*. Medical Care 2005; 43(11):1130-1139.
- Charlson ME, Pompei P, Ales KL, et al. *A new method of classifying prognostic comorbidity in longitudinal studies: development and validation*. Journal of Chronic Diseases 1987; 40:373-383.
- Moore BJ, White S, Washington R, Coenen N, and Elixhauser A. *Identifying increased risk of readmission and in-hospital mortality using hospital administrative data: the AHRQ Elixhauser comorbidity index*. Medical Care 2017; 55(7):698-705.
- van Walraven C, Austin PC, Jennings A, Quan H and Forster AJ. *A modification of the Elixhauser comorbidity measures into a point system for hospital death using administrative data*. Medical Care 2009; 47(6):626-633.
- Menendez ME, Neuhaus V, van Dijk CN, Ring D. *The Elixhauser comorbidity method outperforms the Charlson index in predicting inpatient death after orthopaedic surgery*. Clinical Orthopaedics and Related Research 2014; 472(9):2878-2886.

## Examples

```
set.seed(1)
x <- data.frame(
  id = sample(1:15, size = 200, replace = TRUE),
```

```
code = sample_diag(200),
stringsAsFactors = FALSE
)

# Charlson score based on ICD-10 diagnostic codes:
comorbidity(x = x, id = "id", code = "code", score = "charlson", assign0 = FALSE)

# Elixhauser score based on ICD-10 diagnostic codes:
comorbidity(x = x, id = "id", code = "code", score = "elixhauser", assign0 = FALSE)
```

---

icd10cm\_2017

*ICD-10-CM Diagnostic Codes, 2017 Version*

---

### **Description**

A dataset containing the 2017 version of the ICD10-CM coding system.

### **Usage**

```
icd10cm_2017
```

### **Format**

A data frame with 71,486 rows and 2 variables:

**Code** ICD-10-CM diagnostic code

**Description** Description of each code

### **Note**

The R code used to download and process the dataset from the CDC website is available [here](#).

---

icd10cm\_2018

*ICD-10-CM Diagnostic Codes, 2018 Version*

---

### **Description**

A dataset containing the 2018 version of the ICD10-CM coding system.

### **Usage**

```
icd10cm_2018
```

**Format**

A data frame with 71,704 rows and 2 variables:

**Code** ICD-10-CM diagnostic code

**Description** Description of each code

**Note**

The R code used to download and process the dataset from the CDC website is available [here](#).

---

icd10\_2009

*ICD-10 Diagnostic Codes, 2009 Version*

---

**Description**

A dataset containing the 2009 version of the ICD-10 codes.

**Usage**

```
icd10_2009
```

**Format**

A data frame with 10,817 rows and 4 variables:

**Code** ICD-10 diagnostic code

**Code.clean** ICD-10 diagnostic code, removing all punctuation

**ICD.title** Code description, in plain English.

**Status** Additional information, if available.

**Note**

The R code used to download and process the dataset from the CDC website is available [here](#).

**Source**

CDC Website: <https://goo.gl/6e2mvp>



---

`icd10_2011`*ICD-10 Diagnostic Codes, 2011 Version*

---

**Description**

A dataset containing the 2011 version of the ICD-10 codes.

**Usage**`icd10_2011`**Format**

A data frame with 10,856 rows and 4 variables:

**Code** ICD-10 diagnostic code

**Code.clean** ICD-10 diagnostic code, removing all punctuation

**ICD.title** Code description, in plain English.

**Status** Additional information, if available.

**Note**

The R code used to download and process the dataset from the CDC website is available [here](#).

**Source**

CDC Website: <https://goo.gl/rcTJJ2>

---

`icd9_2015`*ICD-9 Diagnostic Codes, 2015 Version (v32)*

---

**Description**

A dataset containing the version of the ICD-9 codes effective October 1, 2014.

**Usage**`icd9_2015`**Format**

A data frame with 14,567 rows and 3 variables:

**Code** ICD-9 diagnostic code

**Long\_description** Long description of each code

**Short\_description** Short description of each code

**Note**

The R code used to download and process the dataset from the CMS.gov website is available [here](#).

**Source**

CMS.gov Website: <https://www.cms.gov/Medicare/Coding/ICD9ProviderDiagnosticCodes/codes.html>

---

nhds2010

*Adult same-day discharges, 2010*

---

**Description**

A dataset containing adult same-day discharges from 2010.

**Usage**

nhds2010

**Format**

A data frame with 2,210 rows and 15 variables:

**ageu** Units for age

**age** Age

**sex** Sex

**race** Race

**month** Discharge month

**status** Discharge status

**region** Region

**atype** Type of admission

**dx1** Diagnosis 1, ICD9-CM

**dx2** Diagnosis 2, ICD9-CM

**dx3** Diagnosis 3, ICD9-CM, imported incorrectly

**dx3corr** Diagnosis 3, ICD9-CM, corrected

**pr1** Procedure 1

**wgt** Frequency weight

**recid** Order of record (raw data)

**Note**

The R code used to download and process the dataset from Stata is available [here](#).

---

|             |   |
|-------------|---|
| sample_diag | <i>Simulate ICD-10 and ICD-9 diagnostic codes</i> |
|-------------|---|

---

**Description**

A simple function to simulate ICD-10 and ICD-9 diagnostic codes at random.

**Usage**

```
sample_diag(n = 1, version = "ICD10_2011")
```

**Arguments**

|         |   |
|---------|---|
| n       | Number of ICD codes to simulate.  |
| version | The version of the ICD coding scheme to use. Possible choices are ICD10_2009, ICD10_2011, and ICD9_2015; defaults to ICD10_2011. See <a href="#">comorbidity::icd10_2009</a> , <a href="#">comorbidity::icd10_2011</a> , and <a href="#">comorbidity::icd9_2015</a> for further information on the different schemes. |

**Value**

A vector of n ICD diagnostic codes.

**Examples**

```
# Simulate 10 ICD-10 codes
sample_diag(10)

# Simulate a tidy dataset with 15 individuals and 200 rows
set.seed(1)
x <- data.frame(
  id = sample(1:15, size = 200, replace = TRUE),
  code = sample_diag(n = 200),
  stringsAsFactors = FALSE
)
head(x)
```

# Index

## \*Topic **datasets**

- australia10, [2](#)
- icd10\_2009, [8](#)
- icd10\_2011, [9](#)
- icd10cm\_2017, [7](#)
- icd10cm\_2018, [7](#)
- icd9\_2015, [9](#)
- nhds2010, [10](#)

australia10, [2](#)

comorbidity, [3](#)

comorbidity-package, [2](#)

comorbidity::icd10\_2009, [11](#)

comorbidity::icd10\_2011, [11](#)

comorbidity::icd9\_2015, [11](#)

icd10\_2009, [8](#)

icd10\_2011, [9](#)

icd10cm\_2017, [7](#)

icd10cm\_2018, [7](#)

icd9\_2015, [9](#)

nhds2010, [10](#)

parallel::parSapply(), [4](#)

sample\_diag, [11](#)

utils::View(), [4, 6](#)