Package 'BioAge'

June 19, 2020

Type Package

Title Biological Age Calculations Using Several Biomarker Algorithms

Version 0.1.0

Author Dayoon Kwon

Maintainer Dayoon Kwon <dk1943@nyu.edu>

Description This package measures biological aging using data from the National Health and Nutrition Examination Survey (NHANES). The package uses published biomarker algorithms to calculate three biological age measures: Klemera-Doubal Method biological age, phenotypic age, and homeostatic dysregulation.

License GPL-3

Encoding UTF-8

LazyData true

RoxygenNote 7.1.0

Depends R (>= 2.10)

Imports dplyr,
ggplot2

Suggests broom,
flexsurv,
htmlTable,
survival,
testthat,
tidyr,
knitr,
rmarkdown

VignetteBuilder knitr

R topics documented:

`hd_calc` *hd_calc*

Description

Project HD algorithm onto new data.

Usage

```
hd_calc(data, reference, biomarkers)
```

Arguments

`data`	A projection dataset.
`reference`	A training dataset.
`biomarkers`	A character vector indicating the names of the biomarkers included in the HD algorithm.

Details

For HD algorithm., the constructed variable is based on a malhanobis distance statistic, which is theoretically the distance between observations and a hypothetically healthy, young cohort. You need to train separately for men and women who are between the ages of 20 and 30 and not pregnant, and have observe biomarker data within clinically accpetable distributions.

Value

An object of class "hd". This object is a list with two elements (data and fit). The dataset can be drawn by typing 'data'. The model can be drawn by typing 'fit'.

Examples

```
#HD using NHANES
hd = hd_calc(NHANES4, NHANES3,
             biomarkers=c("albumin_gL","lymph","mcv","glucose_mmol",
             "rdw","creat_umol","lncrp","alp","wbc"))

#Extract HD dataset
data = hd$data
```

hd_nhanes *hd_nhanes*

Description

Train HD algorithm in NHANES III and project into NHANES IV.

Usage

```
hd_nhanes(biomarkers)
```

Arguments

biomarkers	A character vector indicating the names of the biomarkers included in the HD algorithm.

Details

Training HD algorithm using the NHANES III (1988 - 1994) and projecting into NHANES IV (1999 - 2018) dataset. For this function, NHANES III included men and women who are between the ages of 20 and 30, and have observe biomarker data within clinically acceptable distributions.

Value

An object of class "hd". This object is a list with two elements (data and fit). The dataset can be drawn by typing 'data'. The model can be drawn by typing 'fit'.

Examples

```
#HD using NHANES
hd = hd_nhanes(biomarkers=c("albumin","lymph","mcv","glucose","rdw","creat","lncrp","alp","wbc"))

#Extract HD dataset
data = hd$data
```

kdm_calc *kdm_calc*

Description

Project KDM algorithm onto new data.

Usage

```
kdm_calc(data, biomarkers, fit = NULL, s_ba2 = NULL)
```

Arguments

data	A projection dataset.
biomarkers	A character vector indicating the names of the biomarkers included in the KDM Biological Age algorithm.
fit	An S3 object for model fit. If the value is NULL, then the parameters to use for training KDM Biological Age are calculated.
s_ba2	A particular fit parameter. Advanced users can modify this parameter to control the variance of KDM Biological Age. If left NULL, defaults are used.

Details

Projecting KDM algorithm onto new data.

Value

An object of class "kdm". This object is a list with two elements (data and fit). The dataset can be drawn by typing 'data'. The model can be drawn by typing 'fit'.

Examples

```
#Train using the NHANES III
train = kdm_calc(NHANES3,
                 biomarkers = c("fev","sbp","totchol","hba1c","albumin",
                 "creat","lncrp","alp","bun"))

#Project into the NHANES IV
kdm = kdm_calc(NHANES4,
               biomarkers = c("fev","sbp","totchol","hba1c","albumin",
               "creat","lncrp","alp","bun"),
               fit = train$fit,
               s_ba2 = train$fit$s_ba2)

#Extract KDM dataset
data = kdm$data
```

---

kdm_nhanes *kdm_nhanes*

Description

Train KDM algorithm in NHANES III and project into NHANES IV.

Usage

```
kdm_nhanes(biomarkers)
```

Arguments

biomarkers	A character vector indicating the names of the biomarkers included in the KDM Biological Age algorithm.

Details

Training KDM Biological Age algorithm using the NHANES III and projecting into NHANES IV dataset.

Value

An object of class "kdm". This object is a list with two elements (data and fit). The dataset can be drawn by typing 'data'. The model can be drawn by typing 'fit'.

Examples

```
#KDM using NHANES
kdm = kdm_nhanes(biomarkers=c("fev","sbp","totchol","hba1c","albumin","creat","lncrp","alp","bun"))

#Extract KDM dataset
data = kdm$data
```

---

phenoage_calc *phenoage_calc*

Description

Project Phenotypic Age algorithm onto new data.

Usage

```
phenoage_calc(data, biomarkers, fit = NULL, orig = FALSE)
```

Arguments

data	A projection dataset.
biomarkers	A character vector indicating the names of the biomarkers included in the Phenotypic Age algorithm.
fit	An S3 object for model fit. If the value is NULL, then the parameters to use for training Phenotypic Age are calculated.
orig	TRUE to compute the origianl Phenotypic Age.

Details

Projecting Phenotypic Age algorithm onto new data.

Value

An object of class "phenoage". This object is a list with two elements (data and fit). The dataset can be drawn by typing 'data'. The model can be drawn by typing 'fit'.

Examples

```
#Train using the NHANES III
train = phenoage_calc(NHANES3,
                      biomarkers = c("albumin_gL","lymph","mcv","glucose_mmol",
                      "rdw","creat_umol","lncrp","alp","wbc"))

#Project into the NHANES IV
phenoage = phenoage_calc(NHANES4,
                         biomarkers = c("albumin_gL","lymph","mcv","glucose_mmol",
                         "rdw","creat_umol","lncrp","alp","wbc"),
                         fit = train$fit)

#Extract phenoage dataset
data = phenoage$data
```

---

phenoage_nhanes *phenoage_nhanes*

Description

Train Phenotypic Age algorithm in NHANES III and project into NHANES IV.

Usage

```
phenoage_nhanes(biomarkers)
```

Arguments

biomarkers	A character vector indicating the names of the biomarkers included in the Phenotypic Age algorithm.

Details

Training Phenotypic Age algorithm using the NHANES III and projecting into NHANES IV dataset.

Value

An object of class "phenoage". This object is a list with two elements (data and fit). The dataset can be drawn by typing 'data'. The model can be drawn by typing 'fit'.

Examples

```
#Phenoage using NHANES
phenoage = phenoage_nhanes(biomarkers=c("albumin_gL","lymph","mcv","glucose_mmol",
                           "rdw","creat_umol","lncrp","alp","wbc"))

#Extract phenoage dataset
data = phenoage$data
```

plot_ba *plot_ba*

Description

Plot association of biological aging measures with chronological age.

Usage

```
plot_ba(data, agevar, label)
```

Arguments

data	A dataset with projected biological aging measures for analysis.
agevar	A character vector indicating the names of the biological aging measures.
label	A character vector indicating the labels of the biological aging measures.

Details

Plot association of biological aging measures with chronological age.

Note

Chronological age and gender variables need to be named "age" and "gender".

Examples

```
#Plot age vs bioage
f1 = plot_ba(data = data, agevar = c("kdm", "phenoage", "hd"),
             label = c("Modified-KDM\nBiological\nAge",
                       "Modified-Levine\nPhenotypic\nAge",
                       "Homeostatic\nDysregulation"))

f1
```

plot_baa *plot_baa*

Description

Plot correlations among biological aging measures.

Usage

```
plot_baa(data, agevar, label, axis_type)
```

Arguments

data	A dataset with projected biological aging measures for analysis.
agevar	A character vector indicating the names of the biological aging measures.
label	A character vector indicating the labels of the biological aging measures. Values should be formatted for displaying along diagonal of the plot. Names should be used to match variables and order is preserved.
axis_type	A character vector indicating the axis type (int or float). Use variable name to define the axis type.

Details

The figure plots associations among the different biological aging measures. Cells below the diagonal show scatter plots of the measures listed above the cell (x-axis) and to the right (y-axis). Cells above the diagonal show the Pearson correlations for the measures listed below the cell and to the left. For this analysis, KDM Biological Age and Levine Phenotypic Age measures are differenced from chronological age (i.e. plotted values = BA-CA).

Examples

```
#Create corplot of BAA with chronologicl age
agevar = c("kdm_advance0",
           "phenoage_advance0",
           "kdm_advance",
           "phenoage_advance",
           "hd",
           "hd_log")

label = c("KDM\nBiological Age",
          "Levine\nPhenotypic Age",
          "Modified-KDM\nBiological Age",
          "Modified-Levine\nPhenotypic Age",
          "Homeostatic\nDysregulation",
          "Log\nHomeostatic\nDysregulation")

axis_type = c("kdm_advance0"="float",
              "phenoage_advance0"="float",
              "kdm_advance"="float",
              "phenoage_advance"="flot",
              "hd"="flot",
              "hd_log"="float")

f2 = plot_baa(data, agevar, labels, axis_type)

f2
```

---

table_health *table_health*

Description

Associations of biological aging measures with healthspan-related characteristics.

Usage

```
table_health(data, agevar, outcome, label)
```

Arguments

data	A dataset with projected biological aging measures for analysis.
agevar	A character vector indicating the names of the biological aging measures.
outcome	A character vector indicating the name of the healthspan-related characteristics.
label	A character vector indicating the labels of the biological aging measures.

Details

Coefficients are from linear regressions of healthspan-related characteristics on biological aging measures. Outcome variables were standardized to have M=0, SD=1 for analysis. KDM Biological Age and Levine Phenotypic Age measures were differenced from chronological age for analysis (i.e. values = BA-CA). These differenced values were then standardized to have M=0, SD=1 separately for men and women within the analysis sample so that effect-sizes are denominated in terms of a sex-specific 1 SD unit increase in biological age advancement. Models included covariates for chronological age and sex.

Value

The result is a list with two elements (table and n). The regression table can be drawn by typing 'table'. The sample size table can be drawn by typing 'n'.

Note

Chronological age, gender, and race/ethnicity variables need to be named "age", "gender", and "race".

Examples

```
table2 = table_health(data,
                      agevar = c("kdm_advance0","phenoage_advance0",
                                 "kdm_advance","phenoage_advance",
                                 "hd","hd_log"),
                      outcome = c("health","adl","lnwalk","grip_scaled"),
                      label = c("KDM\nBiological Age\nAdvancement",
                                "Levine\nPhenotypic Age\nAdvancement",
                                "Modified-KDM\nBiological Age\nAdvancement",
                                "Modified-Levine\nPhenotypic Age\nAdvancement",
                                "Homeostatic\nDysregulation",
                                "Log\nHomeostatic\nDysregulation"))

table2$table
table2$n
```

table_ses *table_ses*

Description

Associations of socioeconomic circumstances measures with measures of biological aging.

Usage

```
table_ses(data, agevar, exposure, label)
```

Arguments

data	A dataset with projected biological aging measures for analysis.
agevar	A character vector indicating the names of the biological aging measures.
exposure	A character vector indicating the name of the socioeconomic circumstances.
label	A character vector indicating the labels of the biological aging measures.

Details

Coefficients are from linear regressions of biological aging measures on measures of socioeconomic circumstances. KDM Biological Age and Levine Phenotypic Age measures were differenced from chronological age for analysis (i.e. values = BA-CA). These differenced values were then standardized to have M=0, SD=1 separately for men and women within the analysis sample. Socioeconomic circumstances measures were standardized to M=0, SD=1 for analysis so that effect-sizes are denominated in terms of a 1 SD unit improvement in socioeconomic circumstances.

Value

The result is a list with two elements (table and n). The regression table can be drawn by typing 'table'. The sample size table can be drawn by typing 'n'.

Note

Chronological age, gender, and race/ethnicity variables need to be named "age", "gender", and "race".

Examples

```
table3 = table_ses(data,
                   agevar = c("kdm_advance0","phenoage_advance0",
                              "kdm_advance","phenoage_advance",
                              "hd","hd_log"),
                   exposure = c("edu","annual_income","poverty_ratio"),
                   label = c("KDM\nBiological Age\nAdvancement",
                             "Levine\nPhenotypic Age\nAdvancement",
                             "Modified-KDM\nBiological Age\nAdvancement",
                             "Modified-Levine\nPhenotypic Age\nAdvancement",
                             "Homeostatic\nDysregulation",
                             "Log\nHomeostatic\nDysregulation"))

table3$table
```

```
table3$n
```

---

table_surv *table_surv*

Description

Associations of biological aging measures with mortality.

Usage

```
table_surv(data, agevar, label)
```

Arguments

data	A dataset with projected biological aging measures for analysis.
agevar	A character vector indicating the names of the biological aging measures.
label	A character vector indicating the labels of the biological aging measures.

Details

BioAge coefficients in the table are hazard ratios estimated from Cox proportional hazard regressions. KDM Biological Age and Levine Phenotypic Age measures were differenced from chronological age for analysis (i.e. values = BA-CA). These differenced values were then standardized to have M=0, SD=1 separately for men and women within the analysis sample so that effect-sizes are denominated in terms of a sex-specific 1 SD unit increase in biological age advancement. Models included covariates for chronological age and sex.

Note

Chronological age, gender, and race/ethnicity variables need to be named "age", "gender", and "race".

Examples

```
table1 = table_surv(data,
                    agevar = c("kdm_advance0","phenoage_advance0",
                               "kdm_advance","phenoage_advance",
                               "hd","hd_log"),
                    label = c("KDM\nBiological Age\nAdvancement",
                    "Levine\nPhenotypic Age\nAdvancement",
                    "Modified-KDM\nBiological Age\nAdvancement",
                    "Modified-Levine\nPhenotypic Age\nAdvancement",
                    "Homeostatic\nDysregulation",
                    "Log\nHomeostatic\nDysregulation"))

table1
```

Index