

Package ‘lcrisks’

July 10, 2023

Title Lung Cancer Death Risk Predictor

Version 4.1.1

Date 2023-07-05

Author Li C. Cheung [aut, cre],
Stephanie A. Kovalchik [aut],
Hormuzd A. Katki [aut]

Maintainer Bill Wheeler <wheelerb@imsweb.com>

Description In both the absence and presence of screening, the R package calculates individual risks of lung cancer and lung cancer death based on covariates: age, education, sex, race, smoking intensity/duration/quit-years, Body Mass Index, family history of lung-cancer, and self-reported emphysema. In the presence of CT screening akin to the NLST (3 yearly screens, 5 years of follow-up), it uses the covariates to estimate risk of false-positive CT screen as well as the reduction in risk of lung cancer death and increase in risk of lung cancer screening. This package also estimates the Life Years Gained From Screening-CT (LYFS-CT) as per Cheung et al., 2019. It requires the same variables as LCDRAT plus 12 additional comorbidities and the year of patient assessment.

Depends R (>= 4.0)

Imports survival, VGAM

License GPL-2

NeedsCompilation no

R topics documented:

cuts	1
lcrisk	2
Index	6

cuts	<i>Wrapper to the cut() function</i>
------	--------------------------------------

Description

A wrapper to the `cut` function, so that you can automatically break into quantiles as the default behavior, otherwise if the breakpoints are included, then just break on those. In all cases, `include.lowest` is set to `True`

Usage

```
cuts(data, npieces, simple.labels = TRUE, ...)
```

Arguments

data	A numeric vector
npieces	Number of cut points or numeric vector of cut points
simple.labels	TRUE or FALSE
...	Other arguments passed into <code>cut</code>

lcrisk

Lung Cancer Death Risk Predictor

Description

In both the absence and presence of screening, the R package calculates individual risks of lung cancer and lung cancer death based on covariates: age, education, sex, race, smoking intensity/duration/quit-years, Body Mass Index, family history of lung-cancer, and self-reported emphysema. In the presence of CT screening akin to the NLST (3 yearly screens, 5 years of follow-up), it uses the covariates to estimate risk of false-positive CT screen as well as the reduction in risk of lung cancer death and increase in risk of lung cancer screening. This package also estimates the Life Years Gained From Screening-CT (LYFS-CT) as per Cheung et al., 2019. A basic mortality model can be fitted using only age, gender, race, and smoking variables. To use the full mortality model, comorbidities and the year of patient assessment should be included.

Usage

```
lcrisk(x, nyears, impute.missing=TRUE, counterfactual.race=0,
       nyears.mortality=5)
```

Arguments

x	A numeric matrix containing individuals' covariates for the model. Columns 11-23 are needed only to estimate life gained from undergoing CT screening using the full mortality model. Covariates should be in the following column and format: <ul style="list-style-type: none"> • column 1 - current age (numeric); • column 2 - gender (1=Female, 0=Male); • column 3 - years smoked (numeric); • column 4 - years quit (numeric or NA); • column 5 - cigarettes per day (numeric); • column 6 - race (0=Non-hispanic white, 1=Non-hispanic Black/African American, 2=Hispanic, 3=Other Ethnicity); • column 7 - lung disease (1=COPD or Emphysema, 0=No COPD or Emphysema); • column 8 - number of parents with lung cancer (0,1,2); • column 9 - bmi;
---	---

- column 10 - highest education level (1=<12 grade, 2=HS graduate, 3=post hs, no college, 4=associate degree/some college, 5=bachelors degree, 6=graduate school);
- column 11 - prior history of cancer (1=Yes,0=No);
- column 12 - Hypertension (1=Yes,0=No);
- column 13 - Coronary Heart Disease (1=Yes,0=No);
- column 14 - Angina pectoris (1=Yes,0=No);
- column 15 - Heart Attack (1=Yes,0=No);
- column 16 - Other heart disease (1=Yes,0=No);
- column 17 - Stroke (1=Yes,0=No);
- column 18 - Diabetes (1=Yes,0=No);
- column 19 - Chronic bronchitis in past year (1=Yes,0=No);
- column 20 - Weak/failing kidneys in past year (1=Yes,0=No);
- column 21 - Liver condition in past year (1=Yes,0=No);
- column 22 - Health problem requiring special equipment (1=Yes,0=No);
- column 23 - Year of assessment.

`nyears` Number of years to calculate risks for (numeric, max of 10).

`impute.missing` TRUE or FALSE to impute missing variables using the NHIS 2015 survey (default=TRUE). The columns that can be imputed are 7-22. The imputation assumes there is complete data for columns 1-6.

`counterfactual.race`
0-3 to compute counterfactual estimates based on a different race, where (0=Non-hispanic white, 1=Non-hispanic Black/African American, 2=Hispanic, 3=Other Ethnicity). See details for more information. The default is 0.

`nyears.mortality`
Number of years to calculate risks for all-cause mortality. A maximum age of 98 is used for this calculation. If age is greater than 98, then the calculation is performed at age 98 and a warning is produced.

Details

The age range for estimation is from 40-98; results may not be accurate for ages outside of this range. The `counterfactual.race` option is used to compute additional life expectancy and life gained estimates. These estimates will appear as columns with the prefix `Counterfactual` in the returned object.

Value

A numeric matrix containing individuals' predictions:

- column 1 - An indicator variable for whether the individual is eligible for CT lung screening according to US Preventive Services Task Force (USPSTF) recommendations.
- column 2 - Number of years predictions are for.
- column 3 - Among 1000 people in the US with this risk-factor profile, this is the number who will die from lung cancer if they do not attend screening.
- column 4 - In the NLST, those who underwent 3 rounds of annual CT screening had their risk reduced by 20 percent. Therefore, among those who would have died from lung cancer, this is the number who will not die from lung cancer death,

- column 5 - Among 1000 people in the US with this risk-factor profile, this is the number who will be diagnosed with lung cancer if they do not attend screening (LCRAT).
- column 6 - In the NLST, those who underwent CT screening had 12.4 percent more lung cancer diagnosed, all of which require treatment. Therefore, among 1000 people with this risk-factor profile, this is the number of extra lung cancer that would be diagnosed, if they undergo 3 yearly CT lung screens as in the NLST.
- column 7 - Out of 1000 NLST participants with this risk profile, this is the number who had at least one false-positive CT screen out of 3 screens.
- column 8 - Days of life expectancy gained from undergoing 3 rounds of CT screening.
- column 9 - Life expectancy without CT screening.
- column 10 - Life expectancy with CT screening.
- column 11 - Years of life gained if lung cancer is found early due to screening.
- column 12 - Years of life gained if lung cancer death is averted due to screening
- column 13 - All-cause mortality risk
- column 14 - Counterfactual days of life expectancy gained from undergoing 3 rounds of CT screening.
- column 15 - Counterfactual life expectancy without CT screening.
- column 16 - Counterfactual life expectancy with CT screening.
- column 17 - Counterfactual years of life gained if lung cancer is found early due to screening.
- column 18 - Counterfactual years of life gained if lung cancer death is averted due to screening
- column 19 - Counterfactual all-cause mortality risk

Model Objects in Package

- LCDRAT - model for lung cancer death in absence of screening;
- LCRAT - model for lung cancer incidence in absence of screening;
- cox.death - model for deaths from causes other than lung cancer;
- morat - model for overall mortality;
- polytmod - polytomous model for false positive CT lung screens.

Author(s)

Li C. Cheung, <li.cheung@nih.gov>, Stephanie A. Kovalchik, Hormuzd A. Katki

References

- Katki HA, Kovalchik SA, Berg CD, Cheung LC, Chaturvedi AK. Development and validation of risk models to select ever-smokers for CT lung cancer screening. *JAMA*. 2016; 315(21):2300-2311. doi: 10.1001/jama.2016.6255.
- Cheung LC, Berg CD, Castle PE, Katki HA, Chaturvedi AK. Life-gained-based versus risk-based selection of smokers for lung cancer screening. *Ann Intern Med*.;171:623-632. doi:107326/M19-1263.

Examples

```
age <- c(66,58,75,72,56)
bmi <- c(23,28,26,27,24)
cpd <- c(36,36,40,24,40)
emp <- c(0,1,1,0,1)
fam.lung.trend <- c(0,2,0,2,0)
female <- c(0,1,0,1,0)
smkyears <- c(43,37,45,42,29)
qyears <- c(NA,NA,9,6,6)
race <- c(0,1,2,2,3)
edu6 <- c(3,5,4,5,5)
prior.cancer <- c(0,0,0,0,0)
hypertension <- c(0,0,1,0,1)
chd <- c(0,0,0,0,0)
angina <- c(0,0,0,0,0)
heartattack <- c(0,0,0,0,1)
heartdisease <- c(0,0,0,0,0)
stroke <- c(0,0,0,0,0)
diab <- c(1,0,0,0,0)
bron <- c(0,1,0,0,1)
kidney <- c(0,0,0,0,0)
liver <- c(0,0,0,0,0)
speceq <- c(0,1,0,0,0)
yearassessed <- rep(2019,5)
years <- 5

persons <- cbind(age,
                 female,
                 smkyears,
                 qyears,
                 cpd,
                 race,
                 emp,
                 fam.lung.trend,
                 bmi,
                 edu6,
                 prior.cancer,
                 hypertension,
                 chd,
                 angina,
                 heartattack,
                 heartdisease,
                 stroke,
                 diab,
                 bron,
                 kidney,
                 liver,
                 speceq,
                 yearassessed)

persons_predictions <- lcrisk(persons,years)
persons_predictions
```

Index

cut, [1](#), [2](#)

cuts, [1](#)

lcrisk, [2](#)