Package ‘LFSPRO’

February 13, 2019

Type Package

Title TP53 germline mutation carrier estimation and cancer risk predictions

Version 2.0.0

Date 2018-09-10

Author Gang Peng, Seung Jun Shin, Jasper Chen, Wenyi Wang

Maintainer Wenyi Wang <wwang7@mdanderson.org>

Description TP53 germline mutations are the main cause of Li-Fraumeni syndrome. This package is designed to estimate probabilities that: 1) the counselee is a TP53 germline mutation carrier, 2) the counselee develops any cancer in future, 3) the counselee develops breast cancer, sarcoma or any other cancers in future, 4) the counselee develops a first or second primary cancer in future, on the basis of his/her family cancer history. The package also provides functions for using the LFS classic and Chompret criteria.

LazyData true

Lazyload true

Depends R (>= 2.10)

NeedsCompilation true

License GPL-3

Encoding UTF-8

RoxygenNote 6.1.1

R topics documented:

LFSPRO-package .......................................................... 2
calLK ................................................................. 4
calLK.cs ............................................................... 5
calLK.mpc ............................................................. 6
cancer.data ............................................................ 7
cancer.type.all ......................................................... 7
combined.risk.mpc ..................................................... 8
combinedata ........................................................... 9
convert.data .......................................................... 9
fam.cancer.data ....................................................... 10
fam.data ............................................................ 11
firstDegreeRelative .................................................. 12
TP53 germline mutations are the main cause of Li-Fraumeni syndrome. This package is designed to estimate probabilities that: 1) the counselee is a TP53 germline mutation carrier, 2) the counselee develops any cancer in future, 3) the counselee develops breast cancer, sarcoma or any other cancers in future, 4) the counselee develops a first or second primary cancer in future, on the basis of his/her family cancer history. The package also provides functions for using the LFS classic and Chompret criteria.

Details

<table>
<thead>
<tr>
<th>Package</th>
<th>LFSPRO</th>
</tr>
</thead>
<tbody>
<tr>
<td>Type</td>
<td>Package</td>
</tr>
<tr>
<td>Version</td>
<td>2.0.0</td>
</tr>
<tr>
<td>Date</td>
<td>2018-12-04</td>
</tr>
<tr>
<td>License</td>
<td>GPL-3</td>
</tr>
</tbody>
</table>

Note

New features in 2.0.0: lfspro.mode is added to replace lfspro to account for additional utilities in cancer-specific and multiple primary cancer risk predictions. Simply specify parameter "mode" to
call desired functions.

New features in 1.0.5: (1) add function lfsChompret2015, (2) update the default MAF and de novo mutation rate.

Author(s)

Gang Peng, SeungJun Shin, Jasper Chen, Wenyi Wang

Maintainer: Wenyi Wang <wwang7@mdanderson.org>

References


See Also

lfspro.mode, lfsClassic, lfsChompret2009

Examples

fam.id <- c("fam1","fam2","fam2","fam2","fam2")
id <- c(1,1,2,100,200)
counsellee.id <- data.frame(fam.id, id)

# LFS classic criteria
lfsClassic(fam.data, cancer.data, counsellee.id)

# Chompret criteria
lfsChompret2015(fam.data, cancer.data, counsellee.id)

# "1st.all" predict the probability of carrying TP53 mutations
lfspro.mode(fam.data, cancer.data, counsellee.id, "1st.all")

# "mpc" predict future risks of developing multiple primary cancers
lfspro.mode(fam.data, cancer.data, counsellee.id, "mpc")

# "1st.cs" predict future risks of having breast cancer, sarcoma, other cancers, and death
lfspro.mode(fam.data, cancer.data, counsellee.id, "1st.cs")
Calculate the likelihood \( Pr(D|G) \).

Description

Calculate the likelihood \( (Pr(D|G), \text{probability of disease status } D \text{ given the genotype } G \text{ for each individual in the family.}) \)

Usage

calLK(fam.cancer.data, penetrance.all)

Arguments

fam.cancer.data

Data including family information and cancer information. See fam.cancer.data for details.

penetrance.all

The penetrance of three genotype (wild type, one copy and two copies TP53 mutated alleles) for male and female in the population. See lfspenet.2010 for details.

Value

Return a n*3 matrix. The likelihood of observing the cancer outcome given genotype for each individual in the family. n denotes the total number of individuals in the family.

Author(s)

Gang Peng

References


See Also

lkNoneAffect, lfsproC, peelingRC

Examples

cancer.data$cancer.type <- c(2,4,3,7,2,2,6,50,50,2,4,50,50,4,50) #cancer types
fam.cancer.data <- combinedata(fam.data, cancer.data)
calLK(fam.cancer.data[[1]], lfspenet.2010) #likelihood in family 1
Calculate the likelihood $Pr(D|G)$ for cancer specific.

Description

Calculate the likelihood ($Pr(D|G)$, probability of disease status D given the genotype G for each individual in the family.

Usage

calLK.cs(fam.cancer.data, penetrance.all)

Arguments

fam.cancer.data
Data including family information and cancer information. See fam.cancer.data for details.

penetrance.all
The penetrance of three genotype (wild type, one copy and two copies of TP53 mutated alleles) for male and female in the population. See lfspenet.cs.death for details.

Value

Return a n*3 matrix. The likelihood of observing the cancer outcome given genotype for each individual in the family. n denotes the total number of individuals in the family.

Author(s)

Seung Jun Shin

References


See Also

lkNoneAffect.cs, lfsproC.cs, peelingRC

Examples

cancer.data$cancer.type <- c(2,4,3,7,2,2,6,50,50,2,4,50,50,4,50) #cancer types
fam.cancer.data <- combinedata(fam.data, cancer.data)
calLK.cs(fam.cancer.data[[1]], lfspenet.cs.nodeath)
Calculate the likelihood \( \Pr(D|G) \) of multiple primary cancers.

### Description

Calculate the likelihood \( \Pr(D|G) \), the probability of disease status \( D \) given the genotype \( G \) for each individual in the family.

### Usage

```r
callK.mpc(data1, data2, parameter)
```

### Arguments

- **data1**: sex, father ID and mother ID
- **data2**: age of diagnosis, sex and multiple primary cancers index
- **parameter**: Parameter estimates for semiparametric recurrent event model of multiple primary cancers. See `parameter.mpc` for details.

### Value

Return a \( n \times 3 \) matrix. The likelihood of observing the cancer outcome given genotype for each individual in the family. \( n \) denotes the total number of individuals in the family.

### Author(s)

Seung Jun Shin

### References


### See Also

- `lfsproC.mpc`

### Examples

```r
fam.cancer.data <- combinedata(fam.data, cancer.data)
data.obj <- convert.data(fam.cancer.data)
data.obj1 <- data.obj[[1]]data.obj2 <- data.obj[[2]]callK.mpc(data.obj1[[2]], data.obj2[[2]], parameter.mpc)
```
**cancer.data**

**Built-in Cancer Information Data example**

**Description**
This built-in cancer information dataset is used to demonstrate the examples of the package. Input cancer data should maintain the same format, column names and data types of the data example to avoid warning or error messages.

**Usage**
cancer.data

**Format**
A data frame with 4 variables.

- **fam.id** Family id
- **id** Individual id
- **cancer.type** Cancer type. See lfspro.cancer.type for details
- **diag.age** The age when the individual was diagnosed with the corresponding cancer

**See Also**
lfspro.cancer.type, cancer.type.all, lfs.cut, invasive.cut

**Examples**
cancer.data

---

**cancer.type.all**

**Predefined cancer types in LFSPRO.**

**Description**
We classified the cancers into 11 groups according to NCCN Guidelines Version 1.2012 Li-Fraumeni Syndrome criteria. sts: soft tissue sarcoma; ost: osterosarcoma; brain: brain tumor; breast: breast cancer; acc: adrenocortical carcinoma; leukemia: leukemia; lung: lung bronchoalveolar cancer; choroid: choroid plexus carcinoma; other.lfs: other LFS spectrum cancers; non.lfs: non LFS spectrum invasive cancers; benign: benign tumors.

**Usage**
cancer.type.all

**Format**
"sts" "ost" "brain" "breast" "acc" "leukemia" "lung" "choroid" "other.lfs" "non.lfs" "benign"
combined.risk.mpc

Calculate the future cancer risk using the MPC model

Description

Calculate the future cancer risk of each individual who is either healthy or has only been diagnosed with one primary cancer using our MPC model (this model cannot predict the cancer risks for patients who have had multiple primary cancer).

Usage

combined.risk.mpc(posterior, risk, counselee.id)

Arguments

- **posterior**: Genotype likelihood, n*3 matrix.
- **risk**: Matrix, future cancer risks for each genotype, wild-type and mutated, in 5 years, 10 years and 15 years.
- **counselee.id**: Data frame including two variables: fam.id (family ids of counselees) and id (individual ids of counselees).

Value

Matrix, family ID, personal ID and the future cancer risks in 5, 10 and 15 years.

Author(s)

Jasper Chen

Examples

counselee.id <- data.frame(fam.id = "fam1", cid=1)
pp <- matrix(c(0.365762982, 0.634046770, 1.902481e-04),nrow=1, ncol=3)
risk <- matrix(c("fam1",1,35,0.008601,0.103538,0.020350,0.220971,0.035787,0.3410216),
nrow=1, ncol=9)
colnames(risk) <- c("fam.id", "ID", "age","5 years(wildtype)", "5 years(mutation)",
"10 years(wildtype)", "10 years(mutation)", "15 years(wildtype)",
"15 years(mutation)")
combined.risk.mpc(pp, risk, counselee.id)
combinedata

Combine the family information data and cancer information data into a single combined data object. The data object is organized family by family.

**Usage**

```r
combinedata(fam.data, cancer.data)
```

**Arguments**

- `fam.data`: A data frame storing family information. See `fam.data` for details.
- `cancer.data`: A data frame storing cancer information. See `cancer.data` for details.

**Value**

A list. Each component in the list stores the family and cancer information for a family. See `fam.cancer.data` for details.

**Author(s)**

Gang Peng

**See Also**

`lfspro.mode`, `lfsproC`

**Examples**

```r
cancer.data$cancer.type <- c(2,4,3,7,2,2,6,50,50,2,4,50,50,4,50) # cancer types
combinedata(fam.data, cancer.data)
```

---

convert.data

Convert the combined family information data and cancer information data into separate data objects.

**Description**

Separate the family information data and cancer information data into two data objects used in multiple primary cancer predictions.

**Usage**

```r
convert.data(fam.cancer.data)
```
fam.cancer.data

Arguments

fam.cancer.data

Data frame storing family (ID, sex, father ID and mother ID) and cancer information (ID, time sex, test, death status (Yes/No) and counts of multiple primary cancers).

Value

A list. Each component in the list stores the family and cancer information for a family. See fam.cancer.data for details.

Author(s)

Seung Jun Shin

Examples

fam.cancer.data <- combinedata(fam.data, cancer.data)
data.obj <- convert.data(fam.cancer.data)

fam.cancer.data

Data object with Family and Cancer Information Example

Description

Data object example with integrated family and cancer information. After combining family and cancer information with combinedata, the data object should be similar to this example.

Usage

fam.cancer.data

Format

List. Each component of the list stores the family and cancer information of one family. Each component is also a list with the following variables:

- fam.id family id
- id individual id
- fid father id
- mid mother id
- sex sex, 0: female, 1: male
- age If the individual is dead, it is age at death, otherwise it is the age at last contact.
- cancer.info Cancer information. List. Each component of the list stores cancer information for each individual in the family. For each component, if there is no cancer for the individual, it is a 0 by 0 data frame. Otherwise, it is a data frame with two columns, cancer.type (coded number) and diag.age (the age at which the cancer was diagnosed).

Examples

fam.cancer.data
Description

This built-in family information dataset is used to demonstrate the examples of the package. Input family data should maintain the same format, column names and data types of the data example to avoid warning or error messages.

Usage

fam.data

Format

A data frame containing 9 variables.

- fam.id  family id
- id  individual id
- fid  father id
- mid  mother id
- sex  sex. 0: female, 1: male
- age  If the individual was alive at the last contact date, use the age at that time. otherwise, use the age at death. Note: do not record age at diagnosis, which should be recorded in cancer.data
- vital  Vital sign. D: death, A: alive

Details

Family id could include characters, but id, father id and mother id must be integers larger than 0.

See Also

cancer.data

Examples

fam.data
**firstDegreeRelative**    *First degree relatives*

**Description**

Find out the first degree relatives of a given individuals in a family.

**Usage**

`firstDegreeRelative(pedigree, ii)`

**Arguments**

- `pedigree`: A data frame including ID, fID and mID. These IDs should be a positive integer.
- `ii`: Individual index (not ID here). If you want to get the first degree relatives of the second individual in the family, set ii as 2.

**Value**

The index of individuals who are the first degree relatives to the given individual.

**Author(s)**

Gang Peng

**See Also**

`secondDegreeRelative`, `lfsClassic`, `lfsChompret2015`

**Examples**

```r
define cancer types using numbers. A cancer is considered malignant if its number is less than invasive.cut, and benign otherwise.

**Usage**

`invasive.cut`

```
*lfs.cut*

**Format**

The format is: num 100

**Examples**

invasive.cut

---

**lfs.cut**

*Cutoff for Li-Fraumeni Syndrome spectrum cancer*

**Description**

A cancer is considered within LFS spectrum if its number is less than lfs.cut, and non-LFS spectrum otherwise.

**Usage**

lfs.cut

---

**lfsChompret2009**

*The Chompret criteria for Li-Fraumeni syndrome*

**Description**

Use the Chompret criteria to identify TP53 germline mutation carriers in the family.

**Usage**

lfsChompret2009(fam.data, cancer.data, counselee.id)

**Arguments**

- **fam.data**
  - Family information data. See *fam.data* for details.

- **cancer.data**
  - Cancer information data. See *cancer.data* for details.

- **counselee.id**
  - Data frame including two variables: fam.id (family id of counselees) and id (individual id of counselees).

**Value**

A data frame of three columns: fam.id, id and result. The result column is a vector of TRUE/FALSE for all counselees indicating whether they are predicted to be TP53 mutation carriers. TRUE: carrier. FALSE: non-carrier.
Author(s)
Gang Peng

References


See Also
- `lfsClassic`, `lfspro.mode`

Examples
```r
fam.id <- c("fam1","fam2","fam2","fam2","fam2")
id <- c(0,0,2,100,200)
counselee.id <- data.frame(fam.id, id)
lfsChompret2009(fam.data, cancer.data, counselee.id)
```

Description
Use the updated Chompret criteria 2015 version to identify TP53 germline mutation carriers in the family.

Usage
```r
lfsChompret2015(fam.data, cancer.data, counselee.id)
```

Arguments
- **fam.data**: Family information data. See `fam.data` for details.
- **cancer.data**: Cancer information data. See `cancer.data` for details.
- **counselee.id**: Data frame including two variables: fam.id (family id of counselees) and id (individual id of counselees).

Value
A data frame of three columns: fam.id, id and result. The result column is a vector of TRUE/FALSE for all counselees indicating whether they are predicted to be TP53 mutation carriers. TRUE: carrier. FALSE: non-carrier.

Author(s)
Gang Peng
References


See Also

lfsClassic, lfspro.mode, lfsChompret2009

Examples

```r
options(stringsAsFactors = FALSE)
fam.id <- c("fam1","fam2","fam2","fam2","fam2")
id <- c(0,0,2,100,200)
counsellee.id <- data.frame(fam.id, id)
lfsChompret2015(fam.data, cancer.data, counsellee.id)
```

Description

Use the Classic criteria to identify TP53 germline mutation carriers in the family.

Usage

```r
lfsClassic(fam.data, cancer.data, counsellee.id)
```

Arguments

- **fam.data**: Family information data. See `fam.data` for details.
- **cancer.data**: Cancer information data. See `cancer.data` for details.
- **counsellee.id**: Data frame including two variables: fam.id (family id of counselees) and id (individual id of counsellee).

Value

A data frame of three columns: fam.id, id and result. The result column is a vector of TRUE/FALSE for all counsellees indicating whether they are predicted to be TP53 mutation carriers. TRUE: carrier. FALSE: non-carrier.

Author(s)

Gang Peng

References

lfspenet.2010  
*Penetrance for TP53 mutations*

**Description**

Penetrance table of three kinds of genotype (wild type, one allele mutation, two allele mutation) of TP53 gene at age from 1 to 110 for both male and female.

**Usage**

lfspenet.2010

**Format**

List. Two component of list indicates the penetrance of male and female (fMX: male, fFX: female). For each component, it is a 110*3 matrix containing the penetrance from age 1 to age 110 for three kinds of genotype: P530 (wild type), P531 (one allele mutation) and P532 (two allele mutation).

**References**


**Examples**

lfspenet.2010

---

lfspenet.cs.death

*Penetrance table for cancer-specific mutation with four competing risks (breast cancer, sarcoma, other cancers and death).*

**Description**

Penetrance table

**Usage**

lfspenet.cs.death
Format

List. Two lists indicate the cancer-specific penetrances of male and female (fMX: male, fFX: female), respectively. For each list, it contains 4 of 110x3 penetrance matrices, row: from age 1 to age 110, column: three kinds of genotype: P530 (wild type, neither of allele mutated), P531 (one allele mutated in P53) and P532 (two allele mutated in P53). Each matrix represents one penetrance matrix of breast cancer, sarcoma, other cancers, and death, respectively.

Examples

1fspenet.cs.death

1fspenet.cs.nodeath

Penetrance table for cancer-specific mutation with three competing risks (breast cancer, sarcoma and other cancers). Default penetrance table in predicting cancer-specific risks.

Description

Penetrance table

Usage

1fspenet.cs.nodeath

Format

List. Two lists indicate the cancer-specific penetrances of male and female (fMX: male, fFX: female), respectively. For each list, it contains 3 of 110x3 penetrance matrices, row: from age 1 to age 110, column: three kinds of genotype: P530 (wild type, neither of allele mutated), P531 (one allele mutated in P53) and P532 (two allele mutated in P53). Each matrix represents one penetrance matrix of breast cancer, sarcoma and other cancers, respectively.

Examples

1fspenet.cs.nodeath

1fspro.cancer.type

Predefined cancer types and the corresponding number in lfspro

Description

We classified the cancers into 11 groups according to NCCN Guidelines Version 1.2012 Li-Fraumeni Syndrome criteria. And then we numbered these different groups of cancer as below.

sts (soft tissue sarcoma): 1
ost (osterosarcoma): 2
brain (brain tumor): 3
breast (breast cancer): 4
acc (adrenocortical carcinoma): 5
Usage

1fspro.cancer.type

Format

A vector with names of different cancer types.

sts ost brain breast acc leukemia lung choroid other.lfs non.lfs benign

Examples

1fspro.cancer.type

1fspro.mode

Estimate TP53 mutation probability and predict cancer risk for families with Li-Fraumeni Syndrome with mode choice

Description

We use a Mendelian risk prediction model to estimate the probability of the counselee being a TP53 germline mutation carrier on the basis of his/her family cancer history.

Usage

1fspro.mode(fam.data, cancer.data, penetrance.all, counselee.id, allef, nloci, mRate, mode)

Arguments

fam.data Family information data. See fam.data for details.
cancer.data Cancer information data. See cancer.data for details.
penetrance.all Penetrance data by default for each mode. lfspenet.2010 is for predicting the probability of carrying TP53 germline mutation, lfspenet.cs.death is used in cancer-specific risk prediction model and parameter.mpc is multiple primary cancer model parameter.
counselee.id Data frame including two variables: fam.id (family id of counselees) and id (individual id of counselees).
AllelFrequency for each locus/gene. If there is only one gene and two alleles in the gene (allele frequency is 0.1 and 0.9), allef = list(c(0.1, 0.9)). If there are two genes, two alleles (allele frequency is 0.1 and 0.9) for gene 1 and three alleles (allele frequency is 0.2, 0.2, and 0.6) for gene 2, allef = list(c(0.1, 0.9, c(0.2, 0.2, 0.6)). We suggest to use list(c(1-maf, maf)). We set it as allef = list(c(0.9994, 0.0006)) by default.

Number of loci/genes in the model. It’s set to be 1 here.

Mutation rate. We set it as 0.00012 by default.

Mode, "1st.all", "1st.cs" or "mpc". "1st.all" for only predicting the probability of carrying TP53 germline mutation, "1st.cs" for predicting cancer-specific risks, and "mpc" for predicting risks of developing multiple primary cancers.

When mode is "1st.all", it outputs the TP53 mutation carrier probability for each counselee. It is a data frame with 3 variables: fam.id (family id), id (individual id) and pp (posterior probability that the counselee is a TP53 mutation carrier). When mode is "1st.cs", it outputs a list, contains a data frame of TP53 mutation carrier probability and a list of future risks of having breast cancer, sarcoma, other cancers and death for each counselee. When mode is "mpc", it outputs a data frame of TP53 mutation carrier probability and a data frame of future cancer risk prediction of developing multiple primary cancer for each counselee.

Author(s)
Gang Peng, Seung Jun Shin, Jasper Chen, Wenyi Wang

References

Examples
```r
fam.id <- c("fam1", "fam2", "fam2", "fam2", "fam2")
id <- c(0, 0, 2, 100, 200)
counselee.id <- data.frame(fam.id, id)
lfspro.mode(fam.data, cancer.data, counselee.id, "1st.all")
lfspro.mode(fam.data, cancer.data, counselee.id, "mpc")
lfspro.mode(fam.data, cancer.data, counselee.id, "1st.cs")
```
Calculate the posterior probability of p53 mutations on the basis of family history

**Description**

LfsproC is used to calculate posterior probability of carrying p53 mutation on the basis of counselee's family cancer history.

**Usage**

```r
lfsproC(fam.cancer.data, penetrance.all, counselee.id, allef, nloci, mRate)
```

**Arguments**

- `fam.cancer.data`: Combined family and cancer information data for ONE FAMILY ONLY. See `fam.cancer.data` for details.
- `penetrance.all`: Penetrance data. See `lfspenet.2010` for details.
- `counselee.id`: Individual id for the counselee. If you want to estimate multiple samples at the same time, just set counselee.id as a vector of IDs for all.
- `allef`: List. Allele frequency for each locus/gene. If there is only one gene and two alleles in the gene (allele frequency is 0.1 and 0.9), allef = list(c(0.1,0.9)). If there are two genes, two alleles (allele frequency is 0.1 and 0.9) for gene 1 and three alleles (allele frequency is 0.2, 0.2 and 0.6) for gene 2, allef = list(c(0.1,0.9),c(0.2,0.2,0.6)). We set it as `allef=list(c(0.9994,0.0006))` by default.
- `nloci`: Number of loci/genes in the model. It’s set to be 1 here.
- `mRate`: Mutation rate. We set it as 0.00012 by default.

**Value**

A data frame of posterior probabilities of having wild type, one allele mutated and two alleles mutated in TP53 for each counselee.

**Author(s)**

Gang Peng, Wenyi Wang

**References**


**See Also**

`lfsClassic`, `lfsChompret2009`, `lfsChompret2015`, `lfspro.mode`, `peelingRC`
Examples

```r
# convert cancer type to specific number and check the cancer type
cancer.data$cancer.type <- c(2,4,3,7,2,2,6,50,50,2,4,50,50,4,50) # cancer types
allef <- allef.g <- list(c(0.9997,0.0003))
mRate.g <- 6e-05
fam.cancer.data <- combinedata(fam.data, cancer.data)
lfsproC(fam.cancer.data[1], lfspenet.2010, 0, allef.g, 1, mRate.g)
```

The function `lfsproC.cs` is used to calculate the posterior probability of p53 mutations for cancer-specific risk prediction model on the basis of family history.

Description

`lfsproC.cs` is used to calculate posterior probability of p53 mutation based on cancer-specific risk prediction model specifically integrated counselee’s family cancer history.

Usage

```r
lfsproC.cs(fam.cancer.data, penetrance.all, counselee.id, allef, nloci, mRate)
```

Arguments

- **fam.cancer.data**: Combined family and cancer information data for **ONE FAMILY ONLY**. See `fam.cancer.data` for details.
- **penetrance.all**: Penetrance data. See `lfspenet.2010` for details.
- **counselee.id**: Individual id for the counselee. If you want to estimate multiple samples at the same time, just set counselee.id as a vector of IDs for all.
- **allef**: List. Allele frequency for each locus/gene. If there is only one gene and two alleles in the gene (allele frequency is 0.1 and 0.9), `allef = list(c(0.1,0.9))`. If there are two genes, two alleles (allele frequency is 0.1 and 0.9) for gene 1 and three alleles (allele frequency is 0.2, 0.2 and 0.6) for gene 2, `allef = list(c(0.1,0.9),c(0.2,0.2,0.6))`. We set it as `allef=list(c(0.9994,0.0006))` by default.
- **nloci**: Number of loci/genes in the model. It’s set to be 1 here.
- **mRate**: Mutation rate. We set it as 0.00012 by default.

Value

A data frame of posterior probabilities of having wild type, one allele mutated and two alleles mutated in TP53 for each counselee.

Author(s)

Seung Jun Shin, Wenyi Wang

References

**lfsproC.mpc**

Calculate the posterior probability of p53 mutations for multiple primary cancer on the basis of family history

### Description

lfsproC.mpc is used to calculate posterior probability of p53 mutation based on predicting multiple primary cancer model specifically on the basis of counselee's family cancer history.

### Usage

```r
lfsproC.mpc(fam.cancer.data, parameter, data1, data2, counselee.id, allef, nloci, mRate)
```

### Arguments

- **fam.cancer.data**
  Combined family and cancer information data for *ONE FAMILY ONLY*. See `fam.cancer.data` for details.

- **parameter**
  See `parameter.mpc` for details.

- **data1**
  sex, father ID and mother ID

- **data2**
  age of diagnosis, sex and multiple primary cancers index

- **counselee.id**
  Individual id for the counselee. If you want to estimate multiple samples at the same time, just set counselee.id as a vector of IDs for all.

- **allef**
  List. Allele frequency for each locus/gene. If there is only one gene and two alleles in the gene (allele frequency is 0.1 and 0.9), allef = list(c(0.1,0.9)). If there are two genes,two alleles (allele frequency is 0.1 and 0.9) for gene 1 and three alleles (allele frequency is 0.2, 0.2 and 0.6) for gene 2, allef = list(c(0.1,0.9),c(0.2,0.2,0.6)). We set it as allef=list(c(0.9994,0.0006)) by default.

- **nloci**
  Number of loci/genes in the model. It's set to be 1 here.

- **mRate**
  Mutation rate. We set it as 0.00012 by default.

### Value

A data frame of posterior probabilities of having wild type, one allele mutated and two alleles mutated in TP53 for each counselee.
Author(s)

Seung Jun Shin, Wenyi Wang

References


Examples

```r
fam.cancer.data <- combinedata(fam.data, cancer.data)
data.obj <- convert.data(fam.cancer.data)
data.obj1 <- data.obj[[1]]
data.obj2 <- data.obj[[2]]
allef.g <- list(c(0.9997,0.0003))
nloci.g <-1
mRate.g <- 6e-05
cid <- c(0,2,100,200)
lfsproC.mpc(fam.cancer.data[[1]], parameter.mpc,
            data.obj[[1]], data.obj2[[1]], cid, allef.g, nloci.g, mRate.g)

lkNoneAffect

Likelihood for un-affected individuals

Description

The probability of observing cancer outcome given genotype for each un-affected individual in the family.

Usage

`lkNoneAffect(penetrance, age)`

Arguments

penetrance List. Two component of list indicates the penetrance of male and female (fMX: male, fFX: female). For each component, it is a 110*3 matrix containing the penetrance from age 1 to age 110 for three kinds of genotype: P530 (wild type), P531 (one allele mutation) and P532 (two alleles mutation). See lfspenet.2010 for details.

age The age when the individual is still healthy.

Value

The likelihood (Pr(DIG)) for the un-affected individual at 'age' years old. G, genotype, can be TP530 (wild type), TP531 (one allele of TP53 mutated) and PT532 (both alleles of TP53 mutated).

Author(s)

Gang Peng
See Also
calLK

Examples
lkNoneAffect.cs(lfspenet.2010$fFX, 50)

Description
The probability of observing cancer outcome given genotype for each un-affected individual in the family in the cancer-specific model.

Usage
lkNoneAffect.cs(penetrance, age)

Arguments
penetrance List. Two lists indicate the cancer-specific penetrances of male and female (fMX: male, fFX: female), respectively. For each list, it contains 4 of 110x3 penetrance matrices, row: from age 1 to age 110, column: three kinds of genotype: P530 (wild type, neither of allele mutated), P531 (one allele mutated in P53) and P532 (two alleles mutated in P53). Each matrix represents one penetrance matrix of breast cancer, sarcoma, other cancers, and death, respectively. See lfspenet.cs.death for details.
age The age when the individual is still healthy.

Value
The likelihood (Pr(D|G)) for the un-affected individual at ‘age’ years old. G, genotype, can be TP530 (wild type), TP531 (one allele of TP53 mutated) and TP532 (both alleles of TP53 mutated).

Author(s)
Seung Jun Shin

See Also
calLK.cs

Examples
lkNoneAffect.cs(lfspenet.cs.death$fFX, 50)
Estimated parameter for semiparametric recurrent event model of multiple primary cancers

Description
Parameters used in the multiple primary cancer model.

Usage
parameter.mpc

Format
List
beta Genotype, sex, cancer status, interaction of genotype and sex, interaction of genotype and cancer status.
gamma Non-negative flat priors for the baseline intensity.

References

Examples
parameter.mpc

Peeling interface in R.

Description
peelingRC is based on Elston-Stewart algorithm and it is a low level function has been integrated in lfsproC, lfsproC.cs and lfsproC.mpc. We implemented the function in C++ to improve computation so peelingRC linked the peeling algorithm in C++ version to R.

Usage
peelingRC(allel, LIK, ped, counslee.id, nloci = 1, mRate = 0)
Arguments

allef  List. Allele frequency for each locus/gene. If there is only one gene and two alleles in the gene (allele frequency is 0.1 and 0.9), allef = list(c(0.1, 0.9)). If there are two genes, two alleles (allele frequency is 0.1 and 0.9) for gene 1 and three alleles (allele frequency is 0.2, 0.2 and 0.6) for gene 2. allef = list(c(0.1, 0.9), c(0.2, 0.2, 0.6)).

LIK  Matrix, likelihood, Pr(D|G), for three genotypes for all individuals in the family. D: healthy status. G: genotype.

ped  Pedigree structure. A data frame with variables: ID (individual id), Gender (sex, 0: female, 1: male), FatherID (father id) and MotherID (mother id).

counsel.id  Individual id for the counselee. If you want to estimate multiple samples at the same time, just set counsel.id as a vector of IDs for all the samples.

nloci  Number of loci/genes in the model.

mRate  Mutation rate.

Details

One family a time.

Value

The posterior probability (Pr(G|D)) for each counselee.

Author(s)

Gang Peng

References


See Also

lfsproC

---

reformatForClassicChompret  
Reformat input data of LFSPRO for evaluation using the Classic or the Chompret criteria

Description

Reformat input data of LFSPRO for evaluation using the Classic or the Chompret criteria

Usage

reformatForClassicChompret(fam.cancer.data)
risk.cs

Arguments

fam.cancer.data
Combined family and cancer information data for ONE FAMILY ONLY. See fam.cancer.data for details.

Value

Data with format used in lfsClassic and lfsChompret. See lfsClassic, lfsChompret2015 and lfsChompret2009 for details.

Note

One family a time.

Author(s)

Gang Peng

See Also

lfsClassic, lfsChompret2009, lfsChompret2015, lfsproC

Examples

fam.cancer.data <- combinedata(fam.data, cancer.data)
reformatForClassicChompret(fam.cancer.data[[1]])

Description

Predict the future risks of having breast cancer, sarcoma, other cancers, and death in a few years for each counselee, who never had cancers. This function has been integrated in lfspro.mode to predict cancer risks when mode is "1st.cs".

Usage

risk.cs(fam.cancer.data, penetrance.all, counselee.id, posterior)

Arguments

fam.cancer.data
Data frame storing family and cancer information. See fam.cancer.data for details.

penetrance.all List. Two lists indicate the cancer-specific penetrances of male and female (fMX: male, fFX: female), respectively. For each list, it contains 4 of 110x3 penetrance matrices, row: from age 1 to age 110, column: three kinds of genotype: P530 (wild type, neither of allele mutated), P531 (one allele mutated in P53) and P532 (two allele mutated in P53). Each matrix represents one penetrance matrix of breast cancer, sarcoma, other cancers, and death, respectively. See lfspenet.cs.death for details.
counselee.id  Individual id for the counselee. If you want to estimate multiple samples at the same time, just set counselee.id as a vector of IDs for all the samples.

posterior  Posterior probability of carrying TP53 mutation

Value

List. Each matrix of the list indicates the risks of counselees in 5, 10, 15, 20 years of having breast cancer, sarcoma, other cancers and death, respectively.

Author(s)

Jasper Chen

Examples

nloci <- 1
allef <- list(c(1 - 1.0e-4, 1.0e-4))
counselee.id <- data.frame(fam.data[1:5,1:2])
mRate <- 1.0e-6
cancer.data <- data.frame(fam.id = c("fam1", "fam1", "fam1"),
id = c(0,2,4),
cancer.type = c(2,4,1),
diag.age = c(12,50,14))
fam.cancer.data <- combinedata(fam.data, cancer.data)
cid <- c(0,5,10)
pp <- lfsproC.cs(fam.cancer.data[[1]], lfspenet.cs.death, cid, allef, nloci, mRate)
risk.cs(fam.cancer.data[[1]], lfspenet.cs.death, cid, pp)

Description

Calculate the future cancer risk of each counselee who is either healthy or has had one primary cancer before using the MPC model (this model is not built to predict the cancer risks for patients who have had multiple primary cancer). This function has been integrated in lfspro.mode to predict cancer risks when mode is "mpc".

Usage

risk.mpc(fam.cancer.data, cancer.data, cid, data2, parameter)

Arguments

fam.cancer.data  Data frame storing family and cancer information. See fam.cancer.data for details.
cancer.data  Data frame storing cancer information. See cancer.data for details.
cid  Individual id for the counselee. If you want to estimate multiple samples at the same time, just set counselee.id as a vector of IDs for all the samples.
data2  Converted family and cancer data in a good format
parameter  See parameter.mpc for details.
secondDegreeRelative

Value

Matrix, family ID, personal ID and the cancer risks in 5, 10 and 15 years.

Author(s)

Jasper Chen

References

Shin SJ, Li J, Ning J, Bodjadjieva J, Strong LC and Wang W. A Bayesian estimation of semiparametric recurrent event model with applications to the penetrance estimation of multiple primary cancers in Li-Fraumeni syndrome. *Biostatistics*

Examples

```r
fam.id <- unique(fam.data$fam.id)
famdata <- fam.data[fam.data$fam.id == fam.id[1], c(1:6)]
cancerdata <- cancer.data[cancer.data$fam.id == fam.id[1],]
cid <- c(0,1,5,10)
fam.cancer.data <- combinedata(fam.data, cancer.data)
data.obj <- convert.data(fam.cancer.data)
data.obj1 <- data.obj[[1]]
data.obj2 <- data.obj[[2]]
risk.mpc(fam.cancer.data[[1]], cancerdata, cid, data.obj2[[1]], parameter.mpc)
```

Description

Find out the second degree relatives of a given individuals in a family.

Usage

`secondDegreeRelative(pedigree, ii)`

Arguments

- `pedigree`: A data frame including ID, fID and mID. These IDs should be positive integers.
- `ii`: Individual index (not ID here). If you want to get the second degree relatives of the second individual in the family, set ii as 2.

Value

The indexes of individuals who are the second degree relatives to the given individual.

Author(s)

Gang Peng

See Also

`firstDegreeRelative`, `lfsClassic`, `lfsChompret2009`, `lfsChompret2015`
Examples

```r
fam.cancer.data <- combinedata(fam.data, cancer.data)
lfsData <- reformatForClassicChompret(fam.cancer.data[[1]])
pedigree <- data.frame(lfsData$ID, lfsData$fID, lfsData$mID, stringsAsFactors=FALSE)
names(pedigree) <- c("ID","fID","mID")
secondDegreeRelative(pedigree,2)
```
Index

callK, 4, 24
 callK.cs, 5, 24
 callK.mpc, 6
 cancer.data, 7, 8, 9, 11, 13–15, 18, 28
 cancer.type.all, 7, 7
 combined.risk.mpc, 8
 combinedata, 9, 10
 convert.data, 9

fam.cancer.data, 4, 5, 9, 10, 10, 20–22, 27, 28
fam.data, 9, 11, 13–15, 18
firstDegreeRelative, 12, 29

invasive.cut, 7, 8, 12
lfs.cut, 7, 8, 13
lfsChompret2009, 3, 13, 15, 20, 22, 27, 29
lfsChompret2015, 12, 14, 16, 20, 22, 27, 29
lfsClassic, 3, 12, 14, 15, 15, 20, 22, 27, 29
lfspenet.2010, 4, 16, 18, 20, 21, 23
lfspenet.cs.death, 5, 16, 18, 24, 27
lfspenet.cs.nodeath, 17
LFSPRO-package, 2
lfspro.cancer.type, 7, 8, 17
lfspro.mode, 3, 9, 14–16, 18, 20, 22, 27, 28
lfsproC, 4, 9, 20, 25–27
lfsproC.cs, 5, 21, 25
lfsproC.mpc, 6, 22, 25
lkNoneAffect, 4, 23
lkNoneAffect.cs, 5, 24

parameter.mpc, 6, 18, 22, 25, 28
peelingRC, 4, 5, 20, 25

reformatForClassicChompret, 26
risk.cs, 27
risk.mpc, 28

secondDegreeRelative, 12, 29