

# BayesMendel v2.1-4: An R package for cancer risk prediction

Amanda Blackford

March 20, 2018

## 1 Introduction

The BayesMendel working group is dedicated to the development of methodologies, models, and open source software for predicting who may carry a cancer susceptibility gene. We use statistical ideas that go back to Bayes and genetic models that go back to Mendel.

This vignette will show the user how to use BRCAPRO, MMRpro, PancPRO and MelaPRO to:

- Calculate probabilities of being a germline mutation carrier.
- Calculate future risk of cancer.
- Incorporate supplementary information (marker testing results, germline testing results, tumor information) into the models.

## 2 Using the models

### 2.1 BRCAPRO

#### 2.1.1 Family History

Before running your pedigree through **brcapro**, be sure it is structured as a numeric data frame with history of breast and ovarian cancers: *n* rows (where *n* is the number of family members, including the counselee) and 13 columns with column names:

| Column Name            | Content   |
|------------------------|---|
| ID                     | Member identifier   |
| Gender                 | Gender (0=female, 1=male)   |
| FatherID               | Father's identifier number  |
| MotherID               | Mother's identifier number  |
| AffectedBreast         | Breast cancer status (0=no cancer, 1=breast cancer, one breast involved; 2=bilateral breast cancer, NA=unknown status)  |
| AffectedOvary          | Ovarian cancer status (0=no cancer, 1=ovarian cancer, NA=unknown status)  |
| AgeBreast              | Age of onset of breast cancer if a breast cancer case.<br>Current age or age of death if not a breast cancer case.<br>NA if there is no age information.      |
| AgeOvary               | Age of onset of ovarian cancer if an ovarian cancer case.<br>Current age or age of death if not an ovarian cancer case.<br>NA if there is no age information. |
| AgeBreastContralateral | Age at onset of breast cancer, second breast.<br>Only for members with breast cancer status=2. For the rest enter a 0.  |
| Twins                  | Identifies siblings who are identical twins.<br>Each twin pair is identified by a unique number. For the rest enter a 0.                                      |
| ethnic                 | Identifies the ethnicity of each family member.<br>Enter "nonAJ", "AJ", "Italian", "Other" or NA (as recognized by is.na() function).                         |
| Death                  | Vital Status (0=Alive, 1=Dead)  |
| AgeDeath               | Family member's age at death or current age if alive.   |

If at least one family member is "AJ" the default is to use the prevalence associated with the "AJ" for family members with unknown ethnicity. Otherwise, the prevalence associated with "nonAJ" is used for family members with unknown ethnicity.

To begin using any BayesMendel models, load the package library:

```
> library(BayesMendel)
```

The parameters used by the model, including penetrance, allele frequency, and sensitivity/specificity of testing, are set using the function **brcaparams**. Any changes to the parameters can be made by calling this function.

```
> # Change future risk to be calculated in intervals of 2 y instead of the default of 5 y.
> # Leave all other parameters as set.
> myparams <- brcaparams(age.by=2)
> # Run BRCAPro with family history information for example family
> out = brcapro(family=brca.fam)
```

```
[1] "Warning: Unknown ages of some unaffected and affected family members have been imputed.
The probability of being a carrier is 0.3136387
  an BRCA1 carrier 0.1466462
  an BRCA2 carrier 0.1669257
The risks of developing cancers are
  By age Breast Ca Risk Ovarian Ca Risk
```

|   |    |            |            |
|---|----|------------|------------|
| 1 | 62 | 0.04628476 | 0.02418171 |
| 2 | 67 | 0.08870664 | 0.05238824 |
| 3 | 72 | 0.12664661 | 0.08257972 |
| 4 | 77 | 0.16414905 | 0.11320972 |
| 5 | 82 | 0.20058852 | 0.14269260 |

```
> slotNames(out)
```

```
[1] "family"      "posterior"    "probs"        "predictions"
[5] "counselee.id" "loglik"       "future.risk"
```

```
> out@probs
```

|   | Pr(Being a carrier) | Pr(BRCA1 mutation) | Pr(BRCA2 mutation) |
|---|---------------------|--------------------|--------------------|
| 1 | 0.3136387           | 0.1466462          | 0.1669257          |

|   | Pr(Both genes mutated) |
|---|------------------------|
| 1 | 6.683007e-05           |

```
> out@family
```

|    | ID | Gender | FatherID | MotherID | AffectedBreast | AffectedOvary |
|----|----|--------|----------|----------|----------------|---------------|
| 1  | 1  | 0      | 3        | 2        | 0              | 0             |
| 2  | 2  | 0      | 9        | 8        | 0              | 1             |
| 3  | 3  | 1      | 11       | 10       | 0              | 0             |
| 4  | 4  | 0      | 0        | 1        | 0              | 0             |
| 5  | 5  | 1      | 3        | 2        | 0              | 0             |
| 6  | 6  | 0      | 0        | 0        | 0              | 0             |
| 7  | 7  | 0      | 3        | 2        | 1              | 0             |
| 8  | 8  | 0      | 0        | 0        | 0              | 1             |
| 9  | 9  | 1      | 0        | 0        | 0              | 0             |
| 10 | 10 | 0      | 0        | 0        | 0              | 0             |
| 11 | 11 | 1      | 0        | 0        | 0              | 0             |
| 12 | 12 | 0      | 9        | 8        | 0              | 0             |
| 13 | 13 | 0      | 9        | 8        | 0              | 0             |
| 14 | 14 | 0      | 11       | 10       | 1              | 0             |
| 15 | 15 | 1      | 5        | 6        | 0              | 0             |
| 16 | 16 | 1      | 0        | 7        | 0              | 0             |
| 17 | 17 | 0      | 0        | 7        | 0              | 0             |
| 18 | 18 | 0      | 0        | 7        | 0              | 0             |
| 19 | 19 | 0      | 0        | 7        | 0              | 0             |
| 20 | 20 | 0      | 21       | 12       | 0              | 0             |
| 21 | 21 | 1      | 0        | 0        | 0              | 0             |

| 22 | 22        | 0        | 9                      | 8             | 2            | 0     |
|----|-----------|----------|------------------------|---------------|--------------|-------|
| 23 | 23        | 0        | 0                      | 22            | 0            | 0     |
| 24 | 24        | 1        | 5                      | 6             | 0            | 0     |
| 25 | 25        | 1        | 5                      | 6             | 0            | 0     |
|    | AgeBreast | AgeOvary | AgeBreastContralateral | Twins         | ethnic       | Death |
| 1  | 57        | 57       |                        | 0             | 0 nonAJ      | 0     |
| 2  | 70        | 69       |                        | 0             | 1 nonAJ      | 0     |
| 3  | 87        | 87       |                        | 0             | 0 nonAJ      | 0     |
| 4  | 32        | 32       |                        | 0             | 0 nonAJ      | 0     |
| 5  | 50        | 50       |                        | 0             | 0 nonAJ      | 0     |
| 6  | 57        | 57       |                        | 0             | 0 nonAJ      | 0     |
| 7  | 45        | 47       |                        | 0             | 0 nonAJ      | 0     |
| 8  | 65        | 65       |                        | 0             | 0 nonAJ      | 0     |
| 9  | 96        | 96       |                        | 0             | 0 nonAJ      | 0     |
| 10 | 75        | 75       |                        | 0             | 0 nonAJ      | 0     |
| 11 | 94        | 94       |                        | 0             | 0 nonAJ      | 0     |
| 12 | 85        | 85       |                        | 0             | 0 nonAJ      | 0     |
| 13 | 79        | 79       |                        | 0             | 0 nonAJ      | 0     |
| 14 | 1         | 1        |                        | 0             | 0 nonAJ      | 0     |
| 15 | 23        | 23       |                        | 0             | 0 nonAJ      | 0     |
| 16 | 12        | 12       |                        | 0             | 0 nonAJ      | 0     |
| 17 | 22        | 22       |                        | 0             | 0 nonAJ      | 0     |
| 18 | 19        | 19       |                        | 0             | 0 nonAJ      | 0     |
| 19 | 16        | 16       |                        | 0             | 0 nonAJ      | 0     |
| 20 | 54        | 54       |                        | 0             | 0 nonAJ      | 0     |
| 21 | 77        | 77       |                        | 0             | 0 nonAJ      | 0     |
| 22 | 40        | 70       |                        | 45            | 1 nonAJ      | 0     |
| 23 | 40        | 40       |                        | 0             | 0 nonAJ      | 0     |
| 24 | 17        | 17       |                        | 0             | 2 nonAJ      | 0     |
| 25 | 17        | 17       |                        | 0             | 2 nonAJ      | 0     |
|    | AgeDeath  | Relation | Mastectomy             | AgeMastectomy | Oophorectomy |       |
| 1  | 57        | 1        | 0                      | 1             | 0            |       |
| 2  | 70        | 4        | 0                      | 1             | 0            |       |
| 3  | 87        | 4        | 0                      | 1             | 0            |       |
| 4  | 32        | 3        | 0                      | 1             | 0            |       |
| 5  | 50        | 2        | 0                      | 1             | 0            |       |
| 6  | NA        | 15       | 0                      | 1             | 0            |       |
| 7  | 47        | 2        | 0                      | 1             | 0            |       |
| 8  | 65        | 7        | 0                      | 1             | 0            |       |
| 9  | 96        | 7        | 0                      | 1             | 0            |       |
| 10 | 75        | 5        | 0                      | 1             | 0            |       |
| 11 | 94        | 5        | 0                      | 1             | 0            |       |
| 12 | 85        | 8        | 0                      | 1             | 0            |       |
| 13 | 79        | 8        | 0                      | 1             | 0            |       |
| 14 | NA        | 6        | 0                      | 1             | 0            |       |

| 15 | 23              | 13    | 0     | 1         | 0                     |
|----|-----------------|-------|-------|-----------|-----------------------|
| 16 | 12              | 13    | 0     | 1         | 0                     |
| 17 | 22              | 13    | 0     | 1         | 0                     |
| 18 | 19              | 13    | 0     | 1         | 0                     |
| 19 | 16              | 13    | 0     | 1         | 0                     |
| 20 | 54              | 0     | 0     | 1         | 0                     |
| 21 | 77              | 0     | 0     | 1         | 0                     |
| 22 | 70              | 8     | 0     | 1         | 0                     |
| 23 | 40              | 0     | 0     | 1         | 0                     |
| 24 | 17              | 13    | 0     | 1         | 0                     |
| 25 | 17              | 13    | 0     | 1         | 0                     |
|    | AgeOophorectomy | BRCA1 | BRCA2 | TestOrder | ER PR CK14 CK5.6 HER2 |
| 1  | 1               | 0     | 0     | 0 0 0     | 0 0 0                 |
| 2  | 1               | 0     | 0     | 0 0 0     | 0 0 0                 |
| 3  | 1               | 0     | 0     | 0 0 0     | 0 0 0                 |
| 4  | 1               | 0     | 0     | 0 0 0     | 0 0 0                 |
| 5  | 1               | 0     | 0     | 0 0 0     | 0 0 0                 |
| 6  | 1               | 0     | 0     | 0 0 0     | 0 0 0                 |
| 7  | 1               | 0     | 0     | 0 0 0     | 0 0 0                 |
| 8  | 1               | 0     | 0     | 0 0 0     | 0 0 0                 |
| 9  | 1               | 0     | 0     | 0 0 0     | 0 0 0                 |
| 10 | 1               | 0     | 0     | 0 0 0     | 0 0 0                 |
| 11 | 1               | 0     | 0     | 0 0 0     | 0 0 0                 |
| 12 | 1               | 0     | 0     | 0 0 0     | 0 0 0                 |
| 13 | 1               | 0     | 0     | 0 0 0     | 0 0 0                 |
| 14 | 1               | 0     | 0     | 0 0 0     | 0 0 0                 |
| 15 | 1               | 0     | 0     | 0 0 0     | 0 0 0                 |
| 16 | 1               | 0     | 0     | 0 0 0     | 0 0 0                 |
| 17 | 1               | 0     | 0     | 0 0 0     | 0 0 0                 |
| 18 | 1               | 0     | 0     | 0 0 0     | 0 0 0                 |
| 19 | 1               | 0     | 0     | 0 0 0     | 0 0 0                 |
| 20 | 1               | 0     | 0     | 0 0 0     | 0 0 0                 |
| 21 | 1               | 0     | 0     | 0 0 0     | 0 0 0                 |
| 22 | 1               | 0     | 0     | 0 0 0     | 0 0 0                 |
| 23 | 1               | 0     | 0     | 0 0 0     | 0 0 0                 |
| 24 | 1               | 0     | 0     | 0 0 0     | 0 0 0                 |
| 25 | 1               | 0     | 0     | 0 0 0     | 0 0 0                 |

>

## 2.2 Age Imputation

By default, **brcapro** imputes the ages of family members with unknown current or affected ages, denoted either by the user with NA (new as of v2.1) or value 1 (used in previous versions).

Family members who are unaffected at an unknown age have their ages imputed using the approach taken in Lyte+ (see Biswas, S. Atienza, P., Chipman, J., Hughes, K., Gutierrez Barrera, A.M., Amos, C.I., Arun, B., Parmigiani, G. (2013) “Simplifying Clinical Use of the Genetic Risk Prediction Model BRCAPro”, Breast Cancer Research and Treatment, 139: 571-579.). Family members who are affected at an unknown age have their ages imputed using a multiple imputation approach that uses SEER incidence rates of breast and ovarian cancer to sample affection ages. The imputation can be turned off by using the option `imputeAges=F` in the `brcapro` function. Note that the imputation of relatives must also be turned off by using option `imputeRelatives=F` in `brcapro`, because by default ages are imputed for relatives who are imputed. These options apply to models `MMRpro`, `pancpro`, and `melapro`.

```
> # Turn off age imputation
> out <- brcapro(family=brca.fam, imputeAges=FALSE, imputeRelatives=FALSE)
> # Calculate risks with imputed ages
> out = brcapro(family=brca.fam, imputeAges=TRUE, imputeRelatives=TRUE)
```

```
[1] "Warning: Unknown ages of some unaffected and affected family members have been imputed.
The probability of being a carrier is 0.3136628
```

```
an BRCA1 carrier 0.1466613
```

```
an BRCA2 carrier 0.1669265
```

```
The risks of developing cancers are
```

|   | By age | Breast Ca Risk | Ovarian Ca Risk |
|---|--------|----------------|-----------------|
| 1 | 62     | 0.04628807     | 0.02418460      |
| 2 | 67     | 0.08871235     | 0.05239424      |
| 3 | 72     | 0.12665398     | 0.08258883      |
| 4 | 77     | 0.16415788     | 0.11322175      |
| 5 | 82     | 0.20059865     | 0.14270720      |

```
> # When age imputation is done, the original
> #family (with NA inputs re-coded to
> #unaffected, age = 1) is returned by brcapro
> out@family
```

|   | ID | Gender | FatherID | MotherID | AffectedBreast | AffectedOvary |
|---|----|--------|----------|----------|----------------|---------------|
| 1 | 1  | 0      | 3        | 2        | 0              | 0             |
| 2 | 2  | 0      | 9        | 8        | 0              | 1             |
| 3 | 3  | 1      | 11       | 10       | 0              | 0             |
| 4 | 4  | 0      | 0        | 1        | 0              | 0             |
| 5 | 5  | 1      | 3        | 2        | 0              | 0             |
| 6 | 6  | 0      | 0        | 0        | 0              | 0             |
| 7 | 7  | 0      | 3        | 2        | 1              | 0             |
| 8 | 8  | 0      | 0        | 0        | 0              | 1             |
| 9 | 9  | 1      | 0        | 0        | 0              | 0             |

|    |    |   |    |    |   |   |
|----|----|---|----|----|---|---|
| 10 | 10 | 0 | 0  | 0  | 0 | 0 |
| 11 | 11 | 1 | 0  | 0  | 0 | 0 |
| 12 | 12 | 0 | 9  | 8  | 0 | 0 |
| 13 | 13 | 0 | 9  | 8  | 0 | 0 |
| 14 | 14 | 0 | 11 | 10 | 1 | 0 |
| 15 | 15 | 1 | 5  | 6  | 0 | 0 |
| 16 | 16 | 1 | 0  | 7  | 0 | 0 |
| 17 | 17 | 0 | 0  | 7  | 0 | 0 |
| 18 | 18 | 0 | 0  | 7  | 0 | 0 |
| 19 | 19 | 0 | 0  | 7  | 0 | 0 |
| 20 | 20 | 0 | 21 | 12 | 0 | 0 |
| 21 | 21 | 1 | 0  | 0  | 0 | 0 |
| 22 | 22 | 0 | 9  | 8  | 2 | 0 |
| 23 | 23 | 0 | 0  | 22 | 0 | 0 |
| 24 | 24 | 1 | 5  | 6  | 0 | 0 |
| 25 | 25 | 1 | 5  | 6  | 0 | 0 |

|    | AgeBreast | AgeOvary | AgeBreastContralateral | Twins | ethnic  | Death |
|----|-----------|----------|------------------------|-------|---------|-------|
| 1  | 57        | 57       |                        | 0     | 0 nonAJ | 0     |
| 2  | 70        | 69       |                        | 0     | 1 nonAJ | 0     |
| 3  | 87        | 87       |                        | 0     | 0 nonAJ | 0     |
| 4  | 32        | 32       |                        | 0     | 0 nonAJ | 0     |
| 5  | 50        | 50       |                        | 0     | 0 nonAJ | 0     |
| 6  | 57        | 57       |                        | 0     | 0 nonAJ | 0     |
| 7  | 45        | 47       |                        | 0     | 0 nonAJ | 0     |
| 8  | 65        | 65       |                        | 0     | 0 nonAJ | 0     |
| 9  | 96        | 96       |                        | 0     | 0 nonAJ | 0     |
| 10 | 75        | 75       |                        | 0     | 0 nonAJ | 0     |
| 11 | 94        | 94       |                        | 0     | 0 nonAJ | 0     |
| 12 | 85        | 85       |                        | 0     | 0 nonAJ | 0     |
| 13 | 79        | 79       |                        | 0     | 0 nonAJ | 0     |
| 14 | 1         | 1        |                        | 0     | 0 nonAJ | 0     |
| 15 | 23        | 23       |                        | 0     | 0 nonAJ | 0     |
| 16 | 12        | 12       |                        | 0     | 0 nonAJ | 0     |
| 17 | 22        | 22       |                        | 0     | 0 nonAJ | 0     |
| 18 | 19        | 19       |                        | 0     | 0 nonAJ | 0     |
| 19 | 16        | 16       |                        | 0     | 0 nonAJ | 0     |
| 20 | 54        | 54       |                        | 0     | 0 nonAJ | 0     |
| 21 | 77        | 77       |                        | 0     | 0 nonAJ | 0     |
| 22 | 40        | 70       |                        | 45    | 1 nonAJ | 0     |
| 23 | 40        | 40       |                        | 0     | 0 nonAJ | 0     |
| 24 | 17        | 17       |                        | 0     | 2 nonAJ | 0     |
| 25 | 17        | 17       |                        | 0     | 2 nonAJ | 0     |

|   | AgeDeath | Relation | Mastectomy | AgeMastectomy | Oophorectomy |
|---|----------|----------|------------|---------------|--------------|
| 1 | 57       | 1        | 0          | 1             | 0            |
| 2 | 70       | 4        | 0          | 1             | 0            |

|    |    |    |   |   |   |
|----|----|----|---|---|---|
| 3  | 87 | 4  | 0 | 1 | 0 |
| 4  | 32 | 3  | 0 | 1 | 0 |
| 5  | 50 | 2  | 0 | 1 | 0 |
| 6  | NA | 15 | 0 | 1 | 0 |
| 7  | 47 | 2  | 0 | 1 | 0 |
| 8  | 65 | 7  | 0 | 1 | 0 |
| 9  | 96 | 7  | 0 | 1 | 0 |
| 10 | 75 | 5  | 0 | 1 | 0 |
| 11 | 94 | 5  | 0 | 1 | 0 |
| 12 | 85 | 8  | 0 | 1 | 0 |
| 13 | 79 | 8  | 0 | 1 | 0 |
| 14 | NA | 6  | 0 | 1 | 0 |
| 15 | 23 | 13 | 0 | 1 | 0 |
| 16 | 12 | 13 | 0 | 1 | 0 |
| 17 | 22 | 13 | 0 | 1 | 0 |
| 18 | 19 | 13 | 0 | 1 | 0 |
| 19 | 16 | 13 | 0 | 1 | 0 |
| 20 | 54 | 0  | 0 | 1 | 0 |
| 21 | 77 | 0  | 0 | 1 | 0 |
| 22 | 70 | 8  | 0 | 1 | 0 |
| 23 | 40 | 0  | 0 | 1 | 0 |
| 24 | 17 | 13 | 0 | 1 | 0 |
| 25 | 17 | 13 | 0 | 1 | 0 |

|    | Age | Oophorectomy | BRCA1 | BRCA2 | TestOrder | ER | PR | CK14 | CK5.6 | HER2 |
|----|-----|--------------|-------|-------|-----------|----|----|------|-------|------|
| 1  |     | 1            | 0     | 0     |           | 0  | 0  | 0    | 0     | 0    |
| 2  |     | 1            | 0     | 0     |           | 0  | 0  | 0    | 0     | 0    |
| 3  |     | 1            | 0     | 0     |           | 0  | 0  | 0    | 0     | 0    |
| 4  |     | 1            | 0     | 0     |           | 0  | 0  | 0    | 0     | 0    |
| 5  |     | 1            | 0     | 0     |           | 0  | 0  | 0    | 0     | 0    |
| 6  |     | 1            | 0     | 0     |           | 0  | 0  | 0    | 0     | 0    |
| 7  |     | 1            | 0     | 0     |           | 0  | 0  | 0    | 0     | 0    |
| 8  |     | 1            | 0     | 0     |           | 0  | 0  | 0    | 0     | 0    |
| 9  |     | 1            | 0     | 0     |           | 0  | 0  | 0    | 0     | 0    |
| 10 |     | 1            | 0     | 0     |           | 0  | 0  | 0    | 0     | 0    |
| 11 |     | 1            | 0     | 0     |           | 0  | 0  | 0    | 0     | 0    |
| 12 |     | 1            | 0     | 0     |           | 0  | 0  | 0    | 0     | 0    |
| 13 |     | 1            | 0     | 0     |           | 0  | 0  | 0    | 0     | 0    |
| 14 |     | 1            | 0     | 0     |           | 0  | 0  | 0    | 0     | 0    |
| 15 |     | 1            | 0     | 0     |           | 0  | 0  | 0    | 0     | 0    |
| 16 |     | 1            | 0     | 0     |           | 0  | 0  | 0    | 0     | 0    |
| 17 |     | 1            | 0     | 0     |           | 0  | 0  | 0    | 0     | 0    |
| 18 |     | 1            | 0     | 0     |           | 0  | 0  | 0    | 0     | 0    |
| 19 |     | 1            | 0     | 0     |           | 0  | 0  | 0    | 0     | 0    |
| 20 |     | 1            | 0     | 0     |           | 0  | 0  | 0    | 0     | 0    |
| 21 |     | 1            | 0     | 0     |           | 0  | 0  | 0    | 0     | 0    |



```

22          1      0      0          0 0 0      0      0      0
23          1      0      0          0 0 0      0      0      0
24          1      0      0          0 0 0      0      0      0
25          1      0      0          0 0 0      0      0      0

```

```

> # Can also impute ages, but not relatives.
> out = brcapro(family=brca.fam, imputeAges=TRUE, imputeRelatives=FALSE)

```

```

[1] "Warning: Unknown ages of some unaffected and affected family members have been imputed.
The probability of being a carrier is 0.3136413
an BRCA1 carrier 0.146648
an BRCA2 carrier 0.1669255
The risks of developing cancers are
  By age Breast Ca Risk Ovarian Ca Risk
1    62    0.04628512    0.02418204
2    67    0.08870726    0.05238892
3    72    0.12664741    0.08258075
4    77    0.16415001    0.11321107
5    82    0.20058963    0.14269423

>

```

### 2.2.1 Changing the penetrance or prevalence

Generally, the user can specify the prevalence of *BRCA1* and *BRCA2* directly in the pedigree through the “ethnic” column.

The user can input their own values for prevalence by specifying `ethnic = “Other”` and inputting the values using the `brcaparams` function.

The user can also specify the penetrance estimates to be used by `brcapro`. The default is the `BRCApenet.metaDSL.2008` object. To use the penetrance estimates for the Italian population:

```

> myparams <- brcaparams(penetrance = BRCApenet.Italian.2008)
> out <- brcapro(family=brca.fam, params=myparams)

```

```

[1] "Warning: Unknown ages of some unaffected and affected family members have been imputed.
The probability of being a carrier is 0.3202846
an BRCA1 carrier 0.1785461
an BRCA2 carrier 0.1416467
The risks of developing cancers are
  By age Breast Ca Risk Ovarian Ca Risk

```

|   |    |            |            |
|---|----|------------|------------|
| 1 | 62 | 0.03000620 | 0.02452284 |
| 2 | 67 | 0.05637127 | 0.05077518 |
| 3 | 72 | 0.08015165 | 0.07660710 |
| 4 | 77 | 0.10146797 | 0.09961138 |
| 5 | 82 | 0.11936882 | 0.11777586 |

### 2.2.2 Specifying race/ethnicity of the family

A set of race/ethnicity-specific baseline (non-carrier) penetrance values were recently added to `brcapro`. The current default assumes that the race/ethnicity of the input family is unknown, but the user can specify one of five different inputs: Asian, Black, Hispanic, NativeAmerican and White. Race/ethnicity categories and estimates were derived using the DevCan (<http://srab.cancer.gov/devcan/>) software provided by the National Cancer Institute (NCI). To specify a particular race, use the “race” input option in `brcapro`.

```
> out <- brcapro(family=brca.fam, race="Hispanic")
```

```
[1] "Warning: Unknown ages of some unaffected and affected family members have been imputed.
The probability of being a carrier is 0.3246615
  an BRCA1 carrier 0.1525616
  an BRCA2 carrier 0.1720064
The risks of developing cancers are
  By age Breast Ca Risk Ovarian Ca Risk
1    62    0.04448786    0.02483715
2    67    0.08418827    0.05375465
3    72    0.11875879    0.08480711
4    77    0.15255182    0.11637879
5    82    0.18461860    0.14657838
```

### 2.2.3 Germline Testing Results

If the results for *BRCA1* and *BRCA2* germline testings are available, the user can input the results in data frame `germline.testing` (0=no test, 1=positive test, 2=negative test) with column names “BRCA1”, “BRCA2” and “Test Order”.

```
> # Add the testing results for BRCA1 and BRCA2
> BRCA1 <- BRCA2 <- TestOrder <- rep(0,nrow(brca.fam))
> germline.testing <- data.frame(BRCA1,BRCA2,TestOrder)
> germline.testing[2,] <- c(2,0,1)
> out <- brcapro(family=brca.fam, germline.testing=germline.testing)
```

```
[1] "Warning: Unknown ages of some unaffected and affected family members have been imputed.
The probability of being a carrier is 0.3280651
```

```

an BRCA1 carrier 0.00312188
an BRCA2 carrier 0.324864
The risks of developing cancers are
  By age Breast Ca Risk Ovarian Ca Risk
1    62    0.04479404    0.01495994
2    67    0.08692135    0.03431950
3    72    0.12523838    0.05717367
4    77    0.16320553    0.08174144
5    82    0.20011388    0.10673470

```

## 2.2.4 Marker Testing Results

If the results for *BRCA1* prognostic markers are available, the user can input the results in data frame `marker.testing` with column names shown below. Note that even if not all the biomarker results listed below are available, all 4 columns must contain non-missing values, which should be set to zero for biomarkers that were not tested.

| Column Name | Content   |
|-------------|---|
| ER          | ER testing result. (0=no test, 1=positive test, 2=negative test)    |
| CK14        | CK14 testing result. (0=no test, 1=positive test, 2=negative test)  |
| CK5.6       | CK5/6 testing result. (0=no test, 1=positive test, 2=negative test) |
| PR          | PR testing result. (0=no test, 1=positive test, 2=negative test)    |
| HER2        | HER2 testing result. (0=no test, 1=positive test, 2=negative test)  |

When the testing result for ER is negative, and the results for CK14 and CK5/6 are both also available, these 3 markers are treated as a group, and the calculations of carrier probabilities will incorporate the joint conditional probabilities of them given genetic status. If the result for either CK14 or CK5/6 is not available, the calculations of carrier probabilities will involve either the marginal conditional probability of ER given genetic status, or if HER2 testing is available, the joint conditional probability of ER and HER2 given genetic status. Note that when ER is positive, the testing results for CK14 or CK5/6 are not considered. For any family member, if the testing result for ER is available, the testing result for PR will be ignored even if it is also available. That is, PR will not be included in carrier prediction when ER is available. PR will only be used when either PR only or PR and HER2 testing are available.

```

> # Add the testing results for breast cancer markers
> marker.testing <- data.frame(matrix(rep(0,nrow(brca.fam)*5),ncol=5))
> colnames(marker.testing) <- c("ER","CK14","CK5.6","PR","HER2")
> brca.fam[1,"AffectedBreast"] <- 1
> marker.testing[1,"ER"] <- 2
> out <- brcapro(family=brca.fam, germline.testing=germline.testing, marker.testing=marker.te

```

```

[1] "Warning: Unknown ages of some unaffected and affected family members have been imputed.
The probability of being a carrier is 0.7679244

```

an BRCA1 carrier 0.0274087

an BRCA2 carrier 0.7393757

The risks of developing cancers are

|   | By age | Contralateral Breast Ca Risk | Ovarian Ca Risk |
|---|--------|------------------------------|-----------------|
| 1 | 62     | 0.05170742                   | 0.03428485      |
| 2 | 67     | 0.12754901                   | 0.07837333      |
| 3 | 72     | 0.22130811                   | 0.13008675      |
| 4 | 77     | 0.32178954                   | 0.18521587      |
| 5 | 82     | 0.42065101                   | 0.24076770      |

### 2.2.5 Oophorectomy

If women in the pedigree have had an oophorectomy, this information can be included in the calculation by creating a data frame `oophorectomy`. Set up a data frame with two columns, one indicating if oophorectomy was done and the other with the age at oophorectomy. If no oophorectomy was done, an individual's current age should be used.

| Column Name     | Content  |
|-----------------|--|
| Oophorectomy    | Oophorectomy yes/no. (0=no oophorectomy, 1=oophorectomy) |
| AgeOophorectomy | Age at Oophorectomy.                                     |

```
> # Add the information for oophorectomy
> Oophorectomy <- c(1,rep(0,(nrow(brca.fam)-1)))
> AgeOophorectomy <- c(30,rep(1,(nrow(brca.fam)-1)))
> oophorectomy <- data.frame(Oophorectomy, AgeOophorectomy)
> out <- brcapro(family=brca.fam, germline.testing=germline.testing, marker.testing=marker.te
```

[1] "Warning: Unknown ages of some unaffected and affected family members have been imputed.

The probability of being a carrier is 0.7746747

an BRCA1 carrier 0.0430876

an BRCA2 carrier 0.7305298

The risks of developing cancers are

|   | By age | Contralateral Breast Ca Risk | Ovarian Ca Risk |
|---|--------|------------------------------|-----------------|
| 1 | 62     | 0.05202175                   | 0.009337688     |
| 2 | 67     | 0.12797834                   | 0.021279423     |
| 3 | 72     | 0.22163105                   | 0.035215751     |
| 4 | 77     | 0.32189038                   | 0.050023035     |
| 5 | 82     | 0.42050109                   | 0.064894221     |

### 2.2.6 Mastectomy

If women in the pedigree have had a bilateral mastectomy, this information can be included in the calculation by creating a data frame `mastectomy`. Set up a data frame with two columns,

one indicating if mastectomy was done and the other with the age at mastectomy. If no mastectomy was done, an individual's current age should be used. Only bilateral mastectomy should be included, and not mastectomy performed on only one breast.

| Column Name   | Content  |
|---------------|--|
| Mastectomy    | Mastectomy yes/no. (0=no mastectomy, 1=mastectomy) |
| AgeMastectomy | Age at Mastectomy.                                 |

```
> # Add the information for mastectomy
> Mastectomy <- c(1,rep(0,(nrow(brca.fam)-1)))
> AgeMastectomy <- c(57,rep(1,(nrow(brca.fam)-1)))
> mastectomy <- data.frame(Mastectomy, AgeMastectomy)
> out <- brcapro(family=brca.fam, mastectomy=mastectomy)
```

```
[1] "Warning: Unknown ages of some unaffected and affected family members have been imputed.
The probability of being a carrier is 0.7915955
  an BRCA1 carrier 0.4109466
  an BRCA2 carrier 0.3803036
The risks of developing cancers are
  By age Contralateral Breast Ca Risk Ovarian Ca Risk
1    62                0.005462367      0.06132724
2    67                0.012326646      0.13223144
3    72                0.020062308      0.20735528
4    77                0.028033310      0.28277503
5    82                0.035784893      0.35447415
```

## 2.3 MMRpro

### 2.3.1 Family History

Before running your pedigree through MMRpro, be sure it is structured as a numeric data frame with history of colon and endometrial cancers: n rows (where n is the number of family members, including the counselee) and 8 columns with required column names described below.

The family history includes the information on the counselee and his/her relatives. For each member, we need information on whether he or she has been diagnosed with colorectal cancer and either the age at diagnosis or, if cancer free, the current age or the age at death. We do the same for endometrial cancer, if the member is female.

The family cancer history must be entered in data frame form, with one row for each family member and columns containing the following information:

| Column              | Content   |
|---------------------|---|
| ID                  | Member identifier   |
| Gender              | Gender (0=female, 1=male)   |
| FatherID            | Father's identifier number  |
| MotherID            | Mother's identifier number  |
| AffectedColon       | Colorectal cancer status<br>(0=no cancer,1=colon/rectum cancer,NA=no information)   |
| AffectedEndometrium | Endometrial cancer status<br>(0=no cancer, 1=ovarian cancer, NA=no information)   |
| AgeColon            | Age of onset of colorectal cancer if a colorectal cancer case.<br>Current age or age of death if not a colorectal cancer case.<br>NA if there is no age information.      |
| AgeEndometrium      | Age of onset of endometrial cancer if an endometrial cancer case.<br>Current age or age of death if not an endometrial cancer case.<br>NA if there is no age information. |
| Twins               | Identifies siblings who are identical twins.<br>Each twin pair is identified by a unique number. For the rest enter a 0.  |

If it is known that a family member is affected, but age of diagnosis is unknown, either enter an estimate or evaluate the program at different plausible ages.

The parameters used by the model, including penetrance, allele frequency, and sensitivity/specificity of testing, are set using the function `MMRparams`. Any changes to the parameters can be made by calling this function.

```
> # Change future risk to be calculated up to age 95 instead of the default 85.
> # Leave all other parameters as set.
> myparams <- MMRparams(age.to=95)
> # Run MMRpro with family history information for example family
> out = MMRpro(family=MMR.fam, params=myparams)
```

```
[1] "Warning: individuals have an unknown age of diagnosis. The calculation would be improved"
```

```
The probability of being a carrier is 0.1186044
```

```
an MLH1 carrier 0.05474188
```

```
an MSH2 carrier 0.06224605
```

```
an MSH6 carrier 0.001649758
```

```
The risks of developing cancers are
```

|   | By age | Colorectal Ca Risk | Endometrial Ca Risk |
|---|--------|--------------------|---------------------|
| 1 | 60     | 0.009420773        | 0.02374388          |
| 2 | 65     | 0.017819581        | 0.04075237          |
| 3 | 70     | 0.026436806        | 0.04803817          |
| 4 | 75     | 0.035562328        | 0.05223954          |
| 5 | 80     | 0.045338604        | 0.05575708          |
| 6 | 85     | 0.055232548        | 0.05860274          |
| 7 | 90     | 0.064291008        | 0.06059296          |
| 8 | 95     | 0.069696469        | 0.06159716          |

>

### 2.3.2 Germline Testing

Information about germline testing results is included in the `germline.testing` object. If the results of germline testing are available, the user can input them into a data frame with `n` rows and 4 columns with column names “MLH1”, “MSH2”, “MSH6”, and Test Order which stores the mutation testing results for *MLH1*, *MSH2*, and *MSH6* (0=no test, 1=positive test, 2=negative test) and order in which family members were tested. If the testing order is unknown, we suggest evaluating the model multiple times, allowing each tested family member to be indicated as the first person tested.

```
> ## The counselee's father tested negative for MLH1 and MSH2.
> ## No testing for MSH6 was done.
> MLH1 <- MSH2 <- MSH6 <- TestOrder <- rep(0, nrow(MMR.fam))
> germline.testing = data.frame(MLH1, MSH2, MSH6, TestOrder)
> germline.testing[3,] <- c(2,2,0,1)
> out <- MMRpro(family=MMR.fam, germline.testing = germline.testing)
```

```
[1] "Warning: individuals have an unknown age of diagnosis. The calculation would be improved by
The probability of being a carrier is 0.04718041
  an MLH1 carrier 0.02110103
  an MSH2 carrier 0.02399553
  an MSH6 carrier 0.002091504
```

The risks of developing cancers are

|   | By age | Colorectal Ca Risk | Endometrial Ca Risk |
|---|--------|--------------------|---------------------|
| 1 | 60     | 0.005417425        | 0.01094528          |
| 2 | 65     | 0.011313262        | 0.01990627          |
| 3 | 70     | 0.018501100        | 0.02535474          |
| 4 | 75     | 0.026942196        | 0.02952298          |
| 5 | 80     | 0.036548404        | 0.03314471          |
| 6 | 85     | 0.046501591        | 0.03607283          |

>

### 2.3.3 Marker Testing

Information about the colorectal tumor is included in the `marker.testing` object. This object is a data frame with `n` rows and 2 columns with information about MSI testing and location of the colorectal tumor. For more information on determining MSI, please refer to Boland (1998). If immunohistochemistry (IHC) was performed, enter 1 if any protein expression was shown to be abnormal or 2 if all were normal.

| Column Name | Content   |
|-------------|---|
| MSI         | Microsatellite instability result<br>enter 1 if high instability is present<br>2 if low instability or stability is present, or<br>0 if no MSI test has been performed. |
| location    | Location of the colorectal tumor:<br>enter 1 if found in the proximal colon<br>2 if found in the distal colon, or<br>0 if the location of the tumor is unknown.         |

```

> ## Now let's say the counselee's sister has a colorectal tumor
>
> MMR.fam[7, "AffectedColon"] <- 1
> ## The counselee's sister's tumor was found to be MSI high.
> ## Add in this MSI result.
>
> MSI <- location <- rep(0, nrow(MMR.fam))
> marker.testing <- data.frame(MSI, location)
> marker.testing[7, "MSI"] <- 1
> out <- MMRpro(family = MMR.fam, marker.testing = marker.testing)

```

```

[1] "Warning: individuals have an unknown age of diagnosis. The calculation would be improved
The probability of being a carrier is 0.3698298
  an MLH1 carrier 0.1705205
  an MSH2 carrier 0.1936391
  an MSH6 carrier 0.006845196
The risks of developing cancers are
  By age Colorectal Ca Risk Endometrial Ca Risk
1    60          0.02353879          0.06822635
2    65          0.04071173          0.11357192
3    70          0.05433141          0.12794211
4    75          0.06584419          0.13288967
5    80          0.07620384          0.13652084
6    85          0.08587415          0.13943420

```

```

>

```



## 2.4 PancPRO

### 2.4.1 Family History

Before running your pedigree through **pancpro**, be sure it is structured as a numeric data frame with history of pancreas cancer: *n* rows (where *n* is the number of family members, including the counselee) and 6 columns with required column names described below.

The family history includes the information on the counselee and his/her relatives. For each member, we need information on whether he or she has been diagnosed with colorectal cancer and either the age at diagnosis or, if cancer free, the current age or the age at death. We do the same for endometrial cancer, if the member is female.

The family cancer history must be entered in data frame form, with one row for each family member and columns containing the following information:

| Column           | Content  |
|------------------|--|
| ID               | Member identifier  |
| Gender           | Gender (0=female, 1=male)  |
| FatherID         | Father's identifier number   |
| MotherID         | Mother's identifier number   |
| AffectedPancreas | Pancreatic cancer status<br>(0=no cancer, 1=pancreatic cancer, NA=no information)  |
| AgePancreas      | Age of onset of pancreatic cancer if a pancreas cancer case.<br>Current age or age of death if not a pancreas cancer case.<br>NA if there is no age information. |
| Twins            | Identifies siblings who are identical twins.<br>Each twin pair is identified by a unique number. For the rest enter a 0.   |

If it is known that a family member is affected, but age of diagnosis is unknown, either enter an estimate or evaluate the program at different plausible ages.

The parameters used by the model, including penetrance, allele frequency, and sensitivity/specificity of testing, are set using the function **pancparams**. Any changes to the parameters can be made by calling this function.

```
> # Change the output for future risk to be calculated
> # in age intervals of 1 year up to
> # age 65 instead of the default 5 years.
> # Leave all other parameters as set.
> myparams <- pancparams(age.by=1, age.to=65)
> # Run PancPRO with family history information for example family
> pancpro(family=panc.fam, params=myparams)
```

```
[1] "Warning: individuals have an unknown age of diagnosis. The calculation would be improved
The probability of being a carrier is 0.8949114"
```

The risks of developing cancers are

By age Pancreatic Ca Risk

|   |    |             |
|---|----|-------------|
| 1 | 58 | 0.006774714 |
| 2 | 59 | 0.014135801 |
| 3 | 60 | 0.022095263 |
| 4 | 61 | 0.030656679 |
| 5 | 62 | 0.039813749 |
| 6 | 63 | 0.049549002 |
| 7 | 64 | 0.059832813 |
| 8 | 65 | 0.070622830 |

An object of class "BayesMendel"

Slot "family":

|    | ID | Relation | Gender | FatherID | MotherID | AffectedPancreas |
|----|----|----------|--------|----------|----------|------------------|
| 1  | 1  | 1        | 0      | 3        | 2        | 0                |
| 2  | 2  | 4        | 0      | 9        | 8        | 0                |
| 3  | 3  | 4        | 1      | 11       | 10       | 0                |
| 4  | 4  | 3        | 0      | 0        | 1        | 0                |
| 5  | 5  | 2        | 1      | 3        | 2        | 0                |
| 6  | 6  | 15       | 0      | 0        | 0        | 0                |
| 7  | 7  | 2        | 0      | 3        | 2        | 1                |
| 8  | 8  | 7        | 0      | 0        | 0        | 0                |
| 9  | 9  | 7        | 1      | 0        | 0        | 0                |
| 10 | 10 | 5        | 0      | 0        | 0        | 0                |
| 11 | 11 | 5        | 1      | 0        | 0        | 0                |
| 12 | 12 | 8        | 0      | 9        | 8        | 0                |
| 13 | 13 | 8        | 0      | 9        | 8        | 0                |
| 14 | 14 | 6        | 0      | 11       | 10       | 1                |
| 15 | 15 | 13       | 1      | 5        | 6        | 0                |
| 16 | 16 | 13       | 1      | 0        | 7        | 0                |
| 17 | 17 | 13       | 0      | 0        | 7        | 0                |
| 18 | 18 | 13       | 0      | 0        | 7        | 0                |
| 19 | 19 | 13       | 0      | 0        | 7        | 0                |

|    | AgePancreas | Twins | Death | AgeDeath | ethnic |
|----|-------------|-------|-------|----------|--------|
| 1  | 57          | 1     | 0     | 57       | Panc   |
| 2  | 70          | 0     | 0     | 70       | Panc   |
| 3  | 87          | 0     | 0     | 87       | Panc   |
| 4  | 32          | 0     | 0     | 32       | Panc   |
| 5  | 50          | 0     | 0     | 50       | Panc   |
| 6  | 57          | 0     | 0     | NA       | Panc   |
| 7  | 45          | 1     | 0     | 45       | Panc   |
| 8  | 65          | 0     | 0     | 65       | Panc   |
| 9  | 96          | 0     | 0     | 96       | Panc   |
| 10 | 75          | 0     | 0     | 75       | Panc   |
| 11 | 94          | 0     | 0     | 94       | Panc   |
| 12 | 85          | 0     | 0     | 85       | Panc   |

|    |    |   |   |    |      |
|----|----|---|---|----|------|
| 13 | 79 | 0 | 0 | 79 | Panc |
| 14 | 1  | 0 | 0 | NA | Panc |
| 15 | 23 | 0 | 0 | 23 | Panc |
| 16 | 12 | 0 | 0 | 12 | Panc |
| 17 | 22 | 0 | 0 | 22 | Panc |
| 18 | 19 | 0 | 0 | 19 | Panc |
| 19 | 16 | 0 | 0 | 16 | Panc |

Slot "posterior":

|      | PANCO     | PANC1     | PANC2       |
|------|-----------|-----------|-------------|
| [1,] | 0.1050886 | 0.8935459 | 0.001365491 |

Slot "probs":

| Pr(Being a carrier) |
|---------------------|
| 0.8949114           |

Slot "predictions":

| By | age | Pancreatic Ca Risk |
|----|-----|--------------------|
| 1  | 58  | 0.006774714        |
| 2  | 59  | 0.014135801        |
| 3  | 60  | 0.022095263        |
| 4  | 61  | 0.030656679        |
| 5  | 62  | 0.039813749        |
| 6  | 63  | 0.049549002        |
| 7  | 64  | 0.059832813        |
| 8  | 65  | 0.070622830        |

Slot "counselee.id":

[1] 1

Slot "loglik":

NULL

Slot "future.risk":

|    | hFX0         | hFX1        |
|----|--------------|-------------|
| 1  | 0.000000e+00 | 0.000000000 |
| 2  | 0.000000e+00 | 0.000000000 |
| 3  | 0.000000e+00 | 0.000000000 |
| 4  | 0.000000e+00 | 0.000000000 |
| 5  | 0.000000e+00 | 0.000000000 |
| 6  | 0.000000e+00 | 0.000000000 |
| 7  | 0.000000e+00 | 0.000000000 |
| 8  | 0.000000e+00 | 0.000000000 |
| 9  | 0.000000e+00 | 0.000000000 |
| 10 | 0.000000e+00 | 0.000000000 |

|    |              |             |
|----|--------------|-------------|
| 11 | 0.000000e+00 | 0.000000000 |
| 12 | 0.000000e+00 | 0.000000000 |
| 13 | 0.000000e+00 | 0.000000000 |
| 14 | 0.000000e+00 | 0.000000000 |
| 15 | 0.000000e+00 | 0.000000000 |
| 16 | 0.000000e+00 | 0.000000000 |
| 17 | 0.000000e+00 | 0.000000000 |
| 18 | 0.000000e+00 | 0.000000000 |
| 19 | 0.000000e+00 | 0.000000000 |
| 20 | 0.000000e+00 | 0.000000000 |
| 21 | 0.000000e+00 | 0.000000000 |
| 22 | 0.000000e+00 | 0.000000000 |
| 23 | 0.000000e+00 | 0.000000000 |
| 24 | 0.000000e+00 | 0.000000000 |
| 25 | 0.000000e+00 | 0.000000000 |
| 26 | 0.000000e+00 | 0.000000000 |
| 27 | 0.000000e+00 | 0.000000000 |
| 28 | 0.000000e+00 | 0.000000000 |
| 29 | 0.000000e+00 | 0.000000000 |
| 30 | 0.000000e+00 | 0.000000000 |
| 31 | 0.000000e+00 | 0.000000000 |
| 32 | 0.000000e+00 | 0.000000000 |
| 33 | 0.000000e+00 | 0.000000000 |
| 34 | 0.000000e+00 | 0.000000000 |
| 35 | 0.000000e+00 | 0.000000000 |
| 36 | 0.000000e+00 | 0.000000000 |
| 37 | 0.000000e+00 | 0.000000000 |
| 38 | 0.000000e+00 | 0.000000000 |
| 39 | 0.000000e+00 | 0.000000000 |
| 40 | 0.000000e+00 | 0.000000000 |
| 41 | 0.000000e+00 | 0.000000000 |
| 42 | 0.000000e+00 | 0.000000000 |
| 43 | 0.000000e+00 | 0.000000000 |
| 44 | 0.000000e+00 | 0.000000000 |
| 45 | 0.000000e+00 | 0.000000000 |
| 46 | 0.000000e+00 | 0.000000000 |
| 47 | 0.000000e+00 | 0.000000000 |
| 48 | 0.000000e+00 | 0.000000000 |
| 49 | 0.000000e+00 | 0.000000000 |
| 50 | 0.000000e+00 | 0.000000000 |
| 51 | 0.000000e+00 | 0.000000000 |
| 52 | 0.000000e+00 | 0.000000000 |
| 53 | 0.000000e+00 | 0.000000000 |
| 54 | 0.000000e+00 | 0.000000000 |
| 55 | 0.000000e+00 | 0.000000000 |

```

56 0.000000e+00 0.000000000
57 0.000000e+00 0.000000000
58 7.987868e-05 0.007560882
59 1.705722e-04 0.015775725
60 2.735389e-04 0.024657767
61 3.904320e-04 0.034210816
62 5.231253e-04 0.044427609
63 6.737418e-04 0.055288377
64 8.446862e-04 0.066759732
65 1.038681e-03 0.078794031

```

```
>
```

## 2.4.2 Germline and Marker Testing

Because the PANC gene is a hypothetical gene, there are no germline or marker testing results to add to the calculation.

## 2.5 MelaPRO

### 2.5.1 Family History

Before running your pedigree through **melapro**, be sure it is structured as a numeric data frame with history of melanomas: *n* rows (where *n* is the number of family members, including the counselee) and 6 columns with required column names described below.

The family history includes the information on the counselee and his/her relatives. For each member, we need information on whether he or she has been diagnosed with colorectal cancer and either the age at diagnosis or, if cancer free, the current age or the age at death. We do the same for endometrial cancer, if the member is female.

The family cancer history must be entered in data frame form, with one row for each family member and columns containing the following information:

| Column       | Content  |
|--------------|--|
| ID           | Member identifier  |
| Gender       | Gender (0=female, 1=male)  |
| FatherID     | Father's identifier number   |
| MotherID     | Mother's identifier number   |
| AffectedSkin | Number of diagnosed melanomas<br>0=no cancer, 1=single melanoma, 2=multiple melanomas, NA=no information                               |
| AgeSkin      | Age of onset of melanomas if a cancer case.<br>Current age or age of death if not a cancer case.<br>NA if there is no age information. |
| Twins        | Identifies siblings who are identical twins.<br>Each twin pair is identified by a unique number. For the rest enter a 0.               |

If it is known that a family member is affected, but age of diagnosis is unknown, either enter an estimate or evaluate the program at different plausible ages.

The parameters used by the model, including penetrance, allele frequency, and sensitivity/specificity of testing, are set using the function `melaparams`. Any changes to the parameters can be made by calling this function.

```
> # Change likelihood ratio for single melanomas
> # among noncarriers from default 0.702 to 0.80
> myparams <- melaparams(spm.lr.noncarrier=0.80)
> # Run PancPRO with family history information for example family
> melapro(family=mela.fam, params=myparams)
```

```
[1] "Warning: individuals have an unknown age of diagnosis. The calculation would be improved
The probability of being a carrier is 0.2398589
The risk of developing cancer is
```

```
By age Melanoma Risk
1      35      0.02230654
2      40      0.04576840
3      45      0.06953031
4      50      0.09294148
5      55      0.11551843
6      60      0.13691796
7      65      0.15691443
8      70      0.17537927
9      75      0.19226226
10     80      0.20757453
11     85      0.22137343
```

```
An object of class "BayesMendel"
```

```
Slot "family":
```

|   | ID | Gender | FatherID | MotherID | AffectedSkin | AgeSkin | Twins | Death |
|---|----|--------|----------|----------|--------------|---------|-------|-------|
| 1 | 1  | 0      | 2        | 3        | 0            | 30.0    | 0     | 0     |
| 2 | 2  | 1      | 20       | 21       | 1            | 40.0    | 0     | 0     |
| 3 | 3  | 0      | 11       | 12       | 0            | 36.0    | 0     | 0     |

|    |    |   |    |    |   |       |   |   |
|----|----|---|----|----|---|-------|---|---|
| 4  | 4  | 0 | 2  | 3  | 1 | 29.0  | 0 | 0 |
| 5  | 5  | 0 | 2  | 3  | 0 | 50.0  | 0 | 0 |
| 6  | 6  | 1 | 0  | 1  | 0 | 24.0  | 0 | 0 |
| 7  | 7  | 0 | 0  | 1  | 0 | 23.0  | 0 | 0 |
| 8  | 8  | 1 | 0  | 1  | 0 | 20.0  | 0 | 0 |
| 9  | 9  | 0 | 0  | 5  | 0 | 26.0  | 0 | 0 |
| 10 | 10 | 0 | 0  | 5  | 0 | 22.0  | 0 | 0 |
| 11 | 11 | 1 | 0  | 0  | 0 | 63.0  | 0 | 0 |
| 12 | 12 | 0 | 0  | 0  | 0 | 92.0  | 0 | 0 |
| 13 | 13 | 1 | 11 | 12 | 0 | 64.0  | 0 | 0 |
| 14 | 14 | 1 | 11 | 12 | 0 | 74.0  | 0 | 0 |
| 15 | 15 | 0 | 14 | 0  | 1 | 1.0   | 0 | 0 |
| 16 | 16 | 0 | 14 | 0  | 0 | 30.0  | 0 | 0 |
| 17 | 17 | 1 | 14 | 0  | 0 | 30.0  | 1 | 0 |
| 18 | 18 | 1 | 14 | 0  | 0 | 30.0  | 1 | 0 |
| 19 | 19 | 1 | 14 | 0  | 0 | 30.0  | 0 | 0 |
| 20 | 20 | 1 | 0  | 0  | 0 | 99.0  | 0 | 0 |
| 21 | 21 | 0 | 0  | 0  | 0 | 100.0 | 0 | 0 |
| 22 | 22 | 0 | 20 | 21 | 0 | 68.5  | 0 | 0 |
| 23 | 23 | 0 | 20 | 21 | 1 | 1.0   | 0 | 0 |
| 24 | 24 | 0 | 20 | 21 | 1 | 1.0   | 0 | 0 |
| 25 | 25 | 1 | 20 | 21 | 0 | 16.0  | 0 | 0 |
| 26 | 26 | 0 | 0  | 24 | 0 | 30.0  | 0 | 0 |
| 27 | 27 | 1 | 0  | 24 | 0 | 30.0  | 0 | 0 |
| 28 | 28 | 1 | 0  | 23 | 0 | 30.0  | 0 | 0 |

|    | AgeDeath | ethnic | Relation | P16 | TestOrder |
|----|----------|--------|----------|-----|-----------|
| 1  | 30       | HBI    | 1        | 0   | 0         |
| 2  | 40       | HBI    | 4        | 0   | 0         |
| 3  | 36       | HBI    | 4        | 0   | 0         |
| 4  | 29       | HBI    | 2        | 0   | 0         |
| 5  | 50       | HBI    | 2        | 0   | 0         |
| 6  | 24       | HBI    | 3        | 0   | 0         |
| 7  | 23       | HBI    | 3        | 0   | 0         |
| 8  | 20       | HBI    | 3        | 0   | 0         |
| 9  | 26       | HBI    | 13       | 0   | 0         |
| 10 | 22       | HBI    | 13       | 0   | 0         |
| 11 | 63       | HBI    | 7        | 0   | 0         |
| 12 | 92       | HBI    | 7        | 0   | 0         |
| 13 | 64       | HBI    | 8        | 0   | 0         |
| 14 | 74       | HBI    | 8        | 0   | 0         |
| 15 | NA       | HBI    | 0        | 0   | 0         |
| 16 | NA       | HBI    | 0        | 0   | 0         |
| 17 | NA       | HBI    | 0        | 0   | 0         |
| 18 | NA       | HBI    | 0        | 0   | 0         |
| 19 | NA       | HBI    | 0        | 0   | 0         |

|    |     |     |   |   |   |
|----|-----|-----|---|---|---|
| 20 | 99  | HBI | 5 | 0 | 0 |
| 21 | 100 | HBI | 5 | 0 | 0 |
| 22 | NA  | HBI | 6 | 0 | 0 |
| 23 | NA  | HBI | 6 | 0 | 0 |
| 24 | NA  | HBI | 6 | 0 | 0 |
| 25 | 16  | HBI | 6 | 0 | 0 |
| 26 | NA  | HBI | 0 | 0 | 0 |
| 27 | NA  | HBI | 0 | 0 | 0 |
| 28 | NA  | HBI | 0 | 0 | 0 |

Slot "posterior":

|      | P160      | P161      | P162         |
|------|-----------|-----------|--------------|
| [1,] | 0.7601411 | 0.2398483 | 1.056473e-05 |

Slot "probs":

Pr(Being a carrier)  
0.2398589

Slot "predictions":

|    | By age | Melanoma Risk |
|----|--------|---------------|
| 1  | 35     | 0.02230654    |
| 2  | 40     | 0.04576840    |
| 3  | 45     | 0.06953031    |
| 4  | 50     | 0.09294148    |
| 5  | 55     | 0.11551843    |
| 6  | 60     | 0.13691796    |
| 7  | 65     | 0.15691443    |
| 8  | 70     | 0.17537927    |
| 9  | 75     | 0.19226226    |
| 10 | 80     | 0.20757453    |
| 11 | 85     | 0.22137343    |

Slot "counselee.id":

[1] 1

Slot "loglik":

NULL

Slot "future.risk":

|   | hFX0         | hFX1       |
|---|--------------|------------|
| 1 | 0.0000000000 | 0.00000000 |
| 2 | 0.0000000000 | 0.00000000 |
| 3 | 0.0000000000 | 0.00000000 |
| 4 | 0.0000000000 | 0.00000000 |
| 5 | 0.0000000000 | 0.00000000 |



|    |              |            |
|----|--------------|------------|
| 6  | 0.0000000000 | 0.00000000 |
| 7  | 0.0000000000 | 0.00000000 |
| 8  | 0.0000000000 | 0.00000000 |
| 9  | 0.0000000000 | 0.00000000 |
| 10 | 0.0000000000 | 0.00000000 |
| 11 | 0.0000000000 | 0.00000000 |
| 12 | 0.0000000000 | 0.00000000 |
| 13 | 0.0000000000 | 0.00000000 |
| 14 | 0.0000000000 | 0.00000000 |
| 15 | 0.0000000000 | 0.00000000 |
| 16 | 0.0000000000 | 0.00000000 |
| 17 | 0.0000000000 | 0.00000000 |
| 18 | 0.0000000000 | 0.00000000 |
| 19 | 0.0000000000 | 0.00000000 |
| 20 | 0.0000000000 | 0.00000000 |
| 21 | 0.0000000000 | 0.00000000 |
| 22 | 0.0000000000 | 0.00000000 |
| 23 | 0.0000000000 | 0.00000000 |
| 24 | 0.0000000000 | 0.00000000 |
| 25 | 0.0000000000 | 0.00000000 |
| 26 | 0.0000000000 | 0.00000000 |
| 27 | 0.0000000000 | 0.00000000 |
| 28 | 0.0000000000 | 0.00000000 |
| 29 | 0.0000000000 | 0.00000000 |
| 30 | 0.0000000000 | 0.00000000 |
| 31 | 0.0002759750 | 0.01713412 |
| 32 | 0.0005614004 | 0.03457191 |
| 33 | 0.0008560248 | 0.05227499 |
| 34 | 0.0011596159 | 0.07020707 |
| 35 | 0.0014719584 | 0.08833378 |
| 36 | 0.0017928516 | 0.10662258 |
| 37 | 0.0021221081 | 0.12504260 |
| 38 | 0.0024595520 | 0.14356456 |
| 39 | 0.0028050179 | 0.16216070 |
| 40 | 0.0031583498 | 0.18080468 |
| 41 | 0.0035194001 | 0.19947156 |
| 42 | 0.0038880289 | 0.21813767 |
| 43 | 0.0042641032 | 0.23678066 |
| 44 | 0.0046474965 | 0.25537935 |
| 45 | 0.0050380881 | 0.27391375 |
| 46 | 0.0054357626 | 0.29236500 |
| 47 | 0.0058404097 | 0.31071532 |
| 48 | 0.0062519235 | 0.32894800 |
| 49 | 0.0066702025 | 0.34704733 |
| 50 | 0.0070951491 | 0.36499861 |

```

51 0.0075266691 0.38278807
52 0.0079646722 0.40040288
53 0.0084090709 0.41783108
54 0.0088597808 0.43506158
55 0.0093167204 0.45208413
56 0.0097798105 0.46888928
57 0.0102489749 0.48546833
58 0.0107241393 0.50181336
59 0.0112052317 0.51791715
60 0.0116921823 0.53377317
61 0.0121849231 0.54937558
62 0.0126833882 0.56471914
63 0.0131875130 0.57979926
64 0.0136972352 0.59461191
65 0.0142124936 0.60915366
66 0.0147332287 0.62342159
67 0.0152593824 0.63741330
68 0.0157908981 0.65112690
69 0.0163277204 0.66456095
70 0.0168697950 0.67771447
71 0.0174170691 0.69058690
72 0.0179694909 0.70317808
73 0.0185270097 0.71548824
74 0.0190895758 0.72751798
75 0.0196571407 0.73926821
76 0.0202296565 0.75074020
77 0.0208070767 0.76193551
78 0.0213893552 0.77285596
79 0.0219764472 0.78350368
80 0.0225683085 0.79388102
81 0.0231648957 0.80399057
82 0.0237661661 0.81383513
83 0.0243720779 0.82341772
84 0.0249825900 0.83274151
85 0.0255976619 0.84180987

```

```
>
```

## 2.5.2 Germline and Marker Testing

Information about germline testing results is included in the `germline.testing` object. If the results of germline testing are available, the user can input them into a data frame with `n` rows and 2 columns with column names “P16”, and “Test Order” which stores the mutation testing results for *P16* (0=no test, 1=positive test, 2=negative test) and order in which family

members were tested. If the testing order is unknown, we suggest evaluating the model multiple times, allowing each tested family member to be indicated as the first person tested.

```
> # The counselee's sister was tested for
> # germline mutations in P16, and one was found.
> # Proband was also tested, but no mutation was found.
> P16 <- TestOrder <- rep(0, nrow(mela.fam))
> germline.testing = data.frame(P16, TestOrder)
> germline.testing[4,] <- c(1,1)
> germline.testing[1,] <- c(2,2)
> out <- melapro(family=mela.fam, germline.testing = germline.testing)
```

```
[1] "Warning: individuals have an unknown age of diagnosis. The calculation would be improved
The probability of being a carrier is 0.0002101976
```

```
The risk of developing cancer is
```

|    | By age | Melanoma Risk |
|----|--------|---------------|
| 1  | 35     | 0.001490217   |
| 2  | 40     | 0.003195691   |
| 3  | 45     | 0.005094605   |
| 4  | 50     | 0.007170380   |
| 5  | 55     | 0.009409789   |
| 6  | 60     | 0.011801922   |
| 7  | 65     | 0.014337549   |
| 8  | 70     | 0.017008703   |
| 9  | 75     | 0.019808401   |
| 10 | 80     | 0.022730437   |
| 11 | 85     | 0.025769228   |

```
>
```

### 3 Other Features

#### 3.1 Plotting a pedigree

The family history data frame can be displayed graphically in a traditional pedigree plot. There are two options for plotting your pedigree. If you want to plot your pedigree without running it through any of the models, the family history data frame `family` must be set to be part of the BayesMendel class and then plotted by simply using the generic function `plot`. If the vital status of family members is known, it can be included by adding a column labeled “status” can be added to the family data frame. Enter 0 if the individual is alive, or 1 if not alive.

```
> pdf("brcafamplot.pdf")
> brca.fam$Death <- rbinom(nrow(brca.fam), 1, 0.2)
```

```

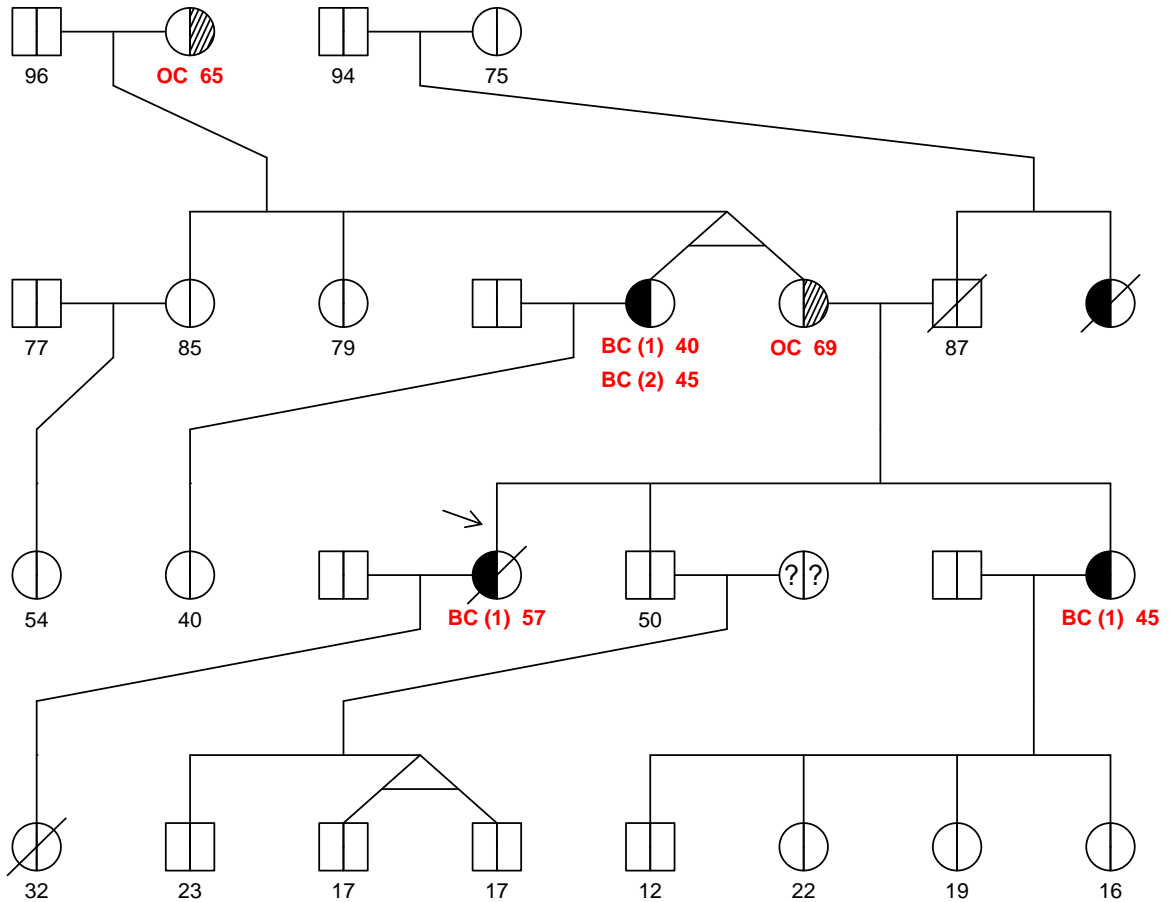
> myfamily <- new("BayesMendel", family=brca.fam, counselee.id=1)
> plot.BayesMendel(myfamily, cex=0.2)
> dev.off()

```

```

null device
      1

```



The pedigree can also be run through any of the models and plotted with the carrier probabilities displayed on the graph.

```

> pdf("mmrfamplot.pdf")

```

```
> MMR.fam$Death <- rbinom(nrow(MMR.fam), 1, 0.2)
> mmrpro.out <- MMRpro(family=MMR.fam, counselee.id=1)
```

```
[1] "Warning: individuals have an unknown age of diagnosis. The calculation would be improved"
```

```
The probability of being a carrier is 0.3537069
```

```
an MLH1 carrier 0.1629603
```

```
an MSH2 carrier 0.1852711
```

```
an MSH6 carrier 0.005677268
```

```
The risks of developing cancers are
```

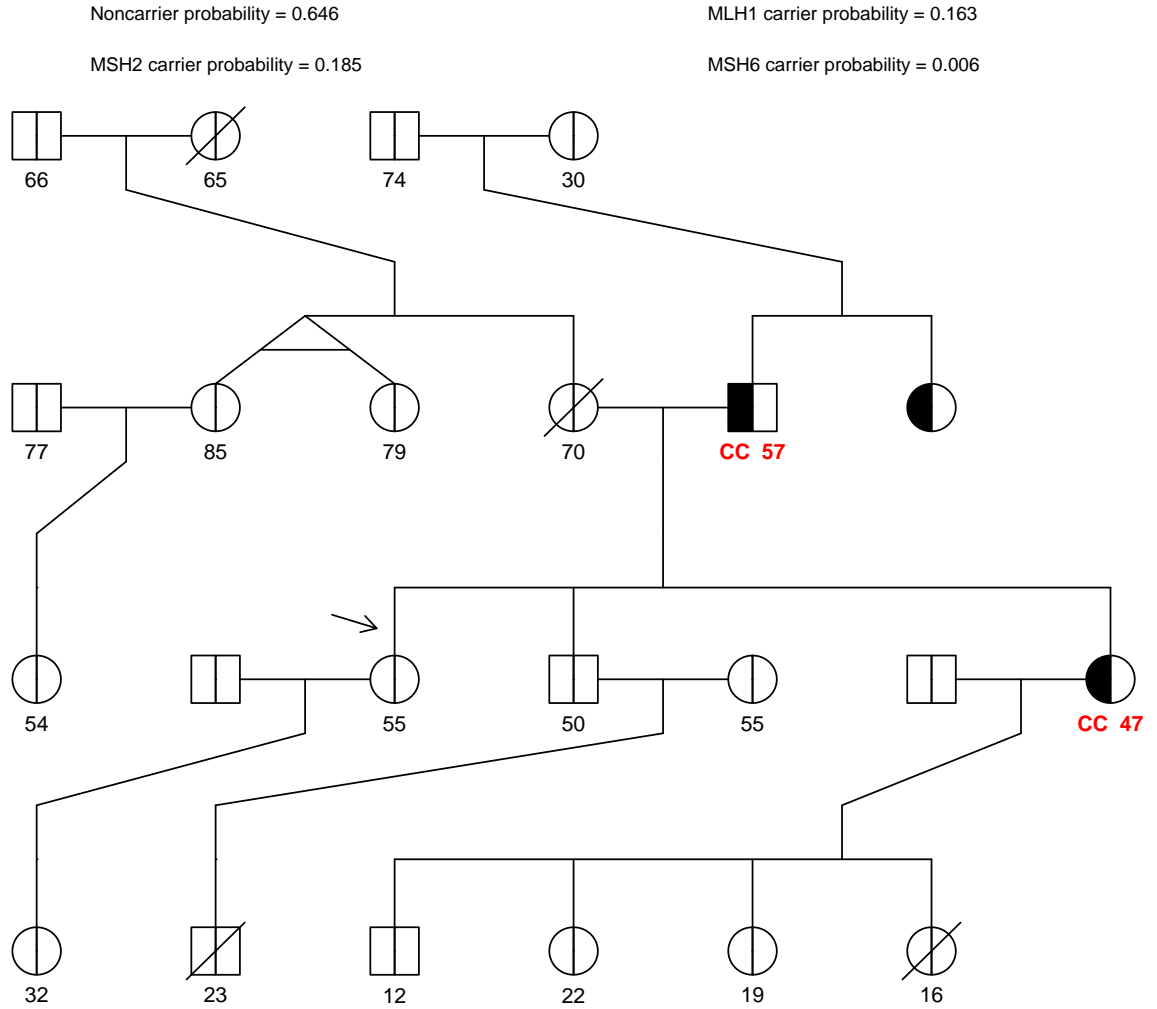
|   | By age | Colorectal Ca Risk | Endometrial Ca Risk |
|---|--------|--------------------|---------------------|
| 1 | 60     | 0.02258293         | 0.06529823          |
| 2 | 65     | 0.03916625         | 0.10873190          |
| 3 | 70     | 0.05244782         | 0.12254724          |
| 4 | 75     | 0.06379638         | 0.12736524          |
| 5 | 80     | 0.07411129         | 0.13092782          |
| 6 | 85     | 0.08379079         | 0.13379105          |

```
> plot(mmrpro.out, cex=0.2)
```

```
> dev.off()
```

```
null device
```

```
1
```



## 4 Interpreting the Risk Predictions

The  $t$ -year risk predictions in these models can be interpreted as the probability of developing the disease within  $t$  years, conditional on surviving and being disease-free at the current age. Implicit in this interpretation is the assumption that the counselee has not died before developing the disease. To understand this more formally, we introduce some notation:

Let  $T_C$  be the theoretical (discrete, in years) age of the specific cancer of interest of the counselee. Thus, in the hypothetical scenario where the counselee does not die before this age, the counselee would develop the cancer of interest at this age. It is important to note that the

counselee may or may not actually observe this outcome. Now let  $T_D$  be the age of death for the counselee, and let  $T = \min(T_C, T_D)$  be the age of the first outcome, either the cancer of interest or death. Let  $J = C$  if  $T = T_C$ ; i.e., if the counselee actually develops the cancer of interest, and let  $J = D$  if  $T = T_D$ .

Using this formulation, we can define the penetrance functions in BayesMendel (at a time  $t$ ) as the probability of developing the cancer of interest at time  $t$  and not having died up to time  $t$ ; i.e.,  $P(T = t, J = C)$ . Here we ignore the dependency on the gender and genotype in the penetrance functions.

Putting this together, we can formally interpret the outputted  $t$ -year risk for the cancer of interest for a  $t_0$ -year-old counselee as

$$P(T \leq t_0 + t, J = C | T > t_0).$$

## 5 Further Information

More information about our methods and software can be found at our website <http://http://bcb.dfci.harvard.edu/bayesmendel>. We can also be reached by email at [BayesMendel@jimmy.harvard.edu](mailto:BayesMendel@jimmy.harvard.edu).