

Determining Carrier Probabilities for Breast Cancer Susceptibility Genes BRCA1 and BRCA2

USER'S MANUAL FOR C VERSION

G. Parmigiani, D.A. Berry, O. Aguilar
Institute of Statistics and Decision Sciences, Duke University

This document describes how to use the program BRCAPRO, which calculates the probability of carrying a mutation of the BRCA1 and BRCA2 breast cancer susceptibility genes. The calculation is based on an individual's family history of breast and ovarian cancer in first- and second-degree relatives and, if available, on the results of genetic testing for one or both of the genes. The ideas behind the calculations are presented in the ISDS Discussion Paper 96-06, available at www.isds.duke.edu. To request BRCAPRO, please contact Giovanni Parmigiani (gp@stat.duke.edu).

Compiling and running BRCAPRO

To begin you need a copy of the following two files:

`brcapro04.c` (the code)

`rates.mean.comb.aj.mar.dat` (the scientific data input file for the Ashkenazi population)

`rates.mean.comb.non.mar.dat` (the scientific data input file for the general population)

If the rates files are too cumbersome, you can rename them. Next you need to create an executable file. On a Unix system, you can do so using the command:

```
cc brcapro04.c -o brcapro
```

The code is self contained and only requires the files `stdio.h`, `math.h`, `stdlib.h`, `malloc.h`, `float.h`, which should be available locally.

If your compilation is successful, you are ready to use BRCAPRO. For every individual to be counseled, you need to create a pedigree input file containing the family history. Preparation is discussed in detail next. If the pedigree file is called, say, `family.dat`, you can then run the program by typing:

```
brcapro rates.mean.comb.aj.mar.dat family.dat probs.out
```

if the family is of Ashkenazi origin and

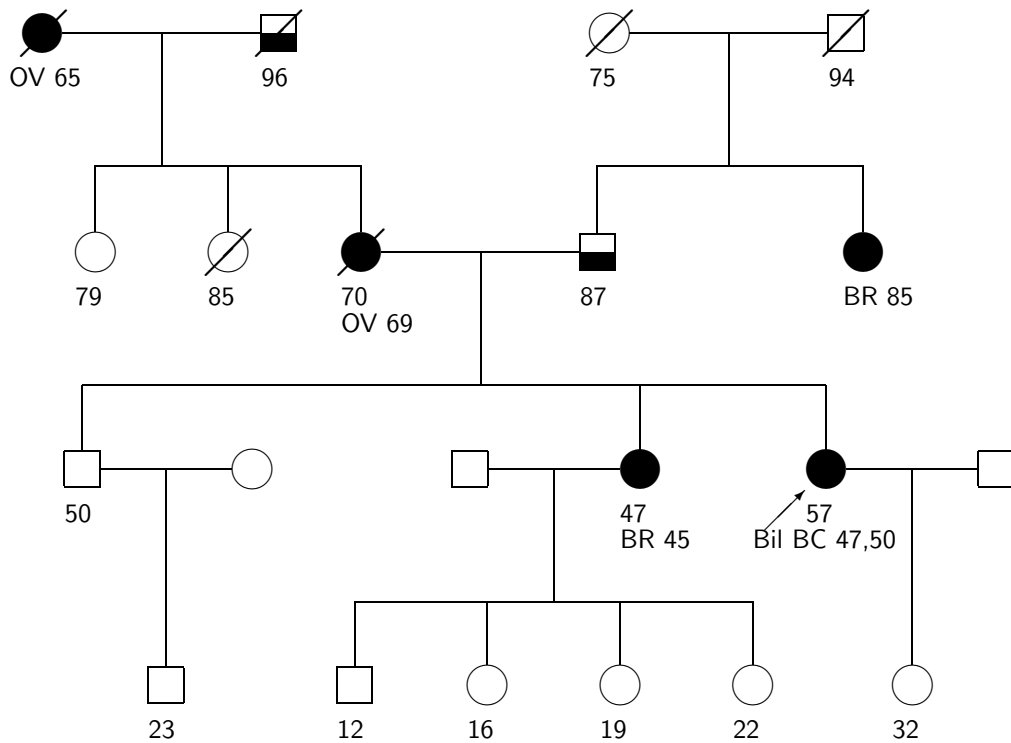


Figure 1: Example pedigree

```
brcapro rates.mean.comb.non.mar.dat family.dat probs.out
```

otherwise.

The file `probs.out` will contain the results of the calculations. The filenames `family.dat` and `probs.out` are just examples and can be changed to suit your needs.

Preparing the pedigree input file

We explain how to prepare the input file using the example family shown in Figure 1. The arrow indicates the individual to be counseled (counseland). The counseland can be either a man or a woman. Computing carrier probabilities for a different counseland within the same family requires creating a different input file. The input file corresponding to the family of Figure 1 is shown in Figure 2.

The first figure to be entered is the number of individuals in the pedigree. The following four figures are the sensitivities and specificities of the genetic tests for BRCA1 and BRCA2 used on family members. These should be entered even if nobody has been tested (in which case the values entered will not affect the calculations). The final figure is the interval for computing the prediction of risk of cancer. Choosing 5, as in the example, will make the program calculate the risk of breast and ovarian cancer within the next 5, 10, 15... years, if the counseland is unaffected.

Num_members 19
Sensitivity_brca1 .9
Specificity_brca1 1
Sensitivity_brca2 .9
Specificity_brca2 1
Probabilities_every 5

Family

1	1	0	3	2	2	0	47	57	50	0	0
2	4	0	9	8	0	1	70	69	0	0	0
3	4	1	11	10	0	0	87	87	0	0	0
4	3	0	0	1	0	0	32	32	0	0	0
5	2	1	3	2	0	0	50	50	0	0	0
6	15	0	0	0	0	0	1	1	0	0	0
7	2	0	3	2	1	0	45	47	0	0	0
8	7	0	0	0	0	1	65	65	0	0	0
9	7	1	0	0	0	0	96	96	0	0	0
10	5	0	0	0	0	0	75	75	0	0	0
11	5	1	0	0	0	0	94	94	0	0	0
12	8	0	9	8	0	0	85	85	0	0	0
13	8	0	9	8	0	0	79	79	0	0	0
14	6	0	11	10	1	0	85	85	0	0	0
15	13	1	5	6	0	0	23	23	0	0	0
16	13	1	0	7	0	0	12	12	0	0	0
17	13	0	0	7	0	0	22	22	0	0	0
18	13	0	0	7	0	0	19	19	0	0	0
19	13	0	0	7	0	0	16	16	0	0	0

Figure 2: Pedigree input file corresponding to the family of Figure 1.

The format shown in Figure 2 should be followed closely. All of the typed text should be left in place. The only things that can be changed in the first 6 lines (that is the lines up to the one starting with the word Family included) are the numbers.

The information about the pedigree is entered next in matrix form, with one row for each family member and 12 columns, containing the following information:

- Column 1: Member identifier.
- Column 2: Relation to the counselland. (see Table 1)
- Column 3: Sex. (0=female, 1=male)
- Column 4: Father's identifier number.
- Column 5: Mother's identifier number.
- Column 6: Breast cancer status (0=no cancer or no information, 1=breast cancer, one breast involved; 2=bilateral breast cancer).
- Column 7: Ovarian cancer status. (0=no cancer, 1=ovarian cancer).
- Column 8: Age of onset of breast cancer if a breast cancer case. Current age or age of death if not a breast cancer case. 1 if unaffected and there is no age information.
- Column 9: Age of onset of ovarian cancer if an ovarian cancer case. Current age or age of death if not an ovarian cancer case. 1 if unaffected and there is no age information.
- Column 10: Age at onset of breast cancer, second breast. Only for members with breast cancer status =2. For the rest enter a 0.
- Column 11: BRCA1 testing result. (0=no test, 1=positive test, 2=negative test).
- Column 12: BRCA2 testing result. (0=no test, 1=positive test, 2=negative test).

Let's consider now family member 1 in detail. Family member 1 is the counselland. She was first diagnosed with breast cancer at age 47 and then had a new diagnosis of primary breast cancer at the other breast at age 50. She is alive and 57 years old. We enter 1 in the member identifier column; we enter 1 in the relation column, using Table 1; We enter 0 in the sex column; We enter 3 in the father's identifier number column —this will constrain us to input the father's information in the third row; We enter 2 in the mother's identifier number column —this will constrain us to input the mother's information in the second row; We enter 2 in the breast cancer status column; We enter 0 in the ovarian cancer status column; We enter 57 —the current age— in the age column for ovarian cancer; We enter 47 in the age column for the first breast cancer;

Number	Relation to the counseland
1	Counseland.
2	Brother or sister.
3	Son or daughter.
4	Parent.
5	Paternal grandparent.
6	Paternal aunt or uncle.
7	Maternal grandparent.
8	Maternal aunt or uncle.
13	Nephew or niece.
14	Husband.
15	Brother or sister in law.

Table 1: Relation codes

We enter 50 in the age column for the second breast cancer; We enter 0 in the BRCA1 test result column; We enter 0 in the BRCA2 test result column.

There are some additional rules and restrictions that you should be aware of when preparing a pedigree input file:

Order. The only restrictions in the order of the family members is that the counseland's husband, if applicable, must be entered immediately after the counseland. The same applies for the brothers' and the sisters' husbands. If you are entering pedigrees by hand, we suggest that you begin by creating the first three columns for all the individuals, and then create the father's and mother's identifiers columns.

Missing Information. In general, if information about a family member, other than the counseland, is missing entirely, the member can be omitted without affecting the calculations. Unaffected members are very important in the calculation, but they provide information only if their current age or age at death is known. In our example, we have no information about family member 6, the counseland's brother's wife, so we could have omitted that row. If information about the breast cancer status is missing, enter 0 in the status column 1 in the age at onset column. When the breast cancer is not 0, the age at onset of breast cancer columns must be different from one. The same applies to ovarian cancer.

Nieces. There is one exception to the rule above. If there is information about a niece of the counseland, it is necessary to include a record (that is a row in the matrix) for the counseland's sibling that is a parent of the niece in question. This will come natural in most cases, but it must be done even in the case where there is no information about that counseland's sibling.

Interpreting the output of BRCAPRO

The output file of BRCAPRO for pedigree 1 is shown in Figure 3. In this case, computations took about 40 seconds on an Alpha station 500. Larger pedigrees require more computations.

FRACTION OF BRCA1 CARRIERS, GENERAL POPULATION: 0.001200
 FRACTION OF BRCA2 CARRIERS, GENERAL POPULATION: 0.000440

JOINT PROBABILITY OF CARRYING BRCA1 AND BRCA2

		BRCA1			
		0	1	2	Marginal
B		-----			-----
R	0	0.034219	0.588630	0.000175	0.623025
C	1	0.376628	0.000308	0.000000	0.376936
A	2	0.000039	0.000000	0.000000	0.000039
2		-----			-----
	Marginal	0.410886	0.588939	0.000176	1
		-----			-----

PROBABILITY OF BEING A BRCA1 CARRIER: 0.589114
 PROBABILITY OF BEING A BRCA2 CARRIER: 0.376975

Figure 3: Output file corresponding to the family of Figure 1.

Also, expect a longer running time on older unix machines and most PC's.

The bottom line of Figure 3 is what you are probably most interested in: the probabilities of carrying BRCA1 and BRCA2, given the family history. These are marginal probabilities. We report 6 significant digits, but we suggest that counseling should be based on at most 2, to avoid conveying a false sense of accuracy. The precision of our estimates is discussed in the technical report.

If you need to have a closer look at the problem, we provide the joint probabilities of carrying the two genes. A counseland can carry 0, 1 or 2 mutations at of each the two genes (carrying two mutations is rare, but in high risk families it cannot be ruled out without biasing the calculations). BRCAPRO calculated the probability that the counseland carries x mutations of BRCA1 and y mutations of BRCA2 (x and y can be 0,1 or 2). These probabilities, a total of 9, are arranged in a 3X3 table at the top of the output. For example, the probability that the counseland has no mutation of either genes is on the 0,0 cell of the table, that is the one at the top left, and is 0.034219. This number is usually a little bit different from what you get if you subtract the two figures in the bottom line from 1, because some people may carry mutations at both genes. The marginal probabilities for this 3X3 table are also provided. For example, the probability of carrying no mutations of BRCA1, irrespective of BRCA2 is 0.410886.

```

Num_members 19
Sensitivity_brca1 .9
Specificity_brca1 1
Sensitivity_brca2 .9
Specificity_brca2 1
Probabilities_every 5

Family
1 1 0 3 2 0 0 57 57 0 0 0
2 4 0 9 8 0 1 70 69 0 0 0
3 4 1 11 10 0 0 87 87 0 0 0
4 3 0 0 1 0 0 32 32 0 0 0
5 2 1 3 2 0 0 50 50 0 0 0
6 15 0 0 0 0 0 1 1 0 0 0
7 2 0 3 2 1 0 45 47 0 0 0
8 7 0 0 0 0 1 65 65 0 0 0
9 7 1 0 0 0 0 96 96 0 0 0
10 5 0 0 0 0 0 75 75 0 0 0
11 5 1 0 0 0 0 94 94 0 0 0
12 8 0 9 8 0 0 85 85 0 0 0
13 8 0 9 8 0 0 79 79 0 0 0
14 6 0 11 10 1 0 85 85 0 0 0
15 13 1 5 6 0 0 23 23 0 0 0
16 13 1 0 7 0 0 12 12 0 0 0
17 13 0 0 7 0 0 22 22 0 0 0
18 13 0 0 7 0 0 19 19 0 0 0
19 13 0 0 7 0 0 16 16 0 0 0

```

Figure 4: Pedigree input file corresponding to the family of Figure 1.

If the counselland is unaffected by cancer, BRCAPRO provides, in addition to the results shown in Figure 3, the probability of developing cancer in the future. To illustrate this feature, consider again the pedigree of Figure 1 and imagine that the counselland is 57 and unaffected. The appropriate input file is shown in Figure 4. It differs from that of Figure 2 only in the first line of the family matrix.

The results of BRCAPRO are shown in Figure 5. The ages are chosen by adding intervals of 5 years, beginning with the counselland's current age. The interval size is controlled by the parameter Probabilities_every in the input file. The probability that the counselland will develop breast cancer within the next ten years, that is by age 67, is .049, or 4.9%.

FRACTION OF BRCA1 CARRIERS, GENERAL POPULATION: 0.001200
 FRACTION OF BRCA2 CARRIERS, GENERAL POPULATION: 0.000440

JOINT PROBABILITY OF CARRYING BRCA1 AND BRCA2

		BRCA1			
		0	1	2	Marginal
B	0	0.939375	0.035909	0.000011	0.975294
C	1	0.024698	0.000005	0.000000	0.024703
A	2	0.000003	0.000000	0.000000	0.000003
2	Marginal	0.964075	0.035914	0.000011	1

PROBABILITY OF BEING A BRCA1 CARRIER: 0.035925
 PROBABILITY OF BEING A BRCA2 CARRIER: 0.024706

PROBABILITY OF DEVELOPING CANCER BEFORE THE AGE INDICATED

AGE	BREAST	OVARIAN
62	0.025616	0.009344
67	0.049138	0.018019
72	0.070339	0.024906
77	0.088817	0.029690
82	0.103863	0.032599
87	0.114244	0.034267

Figure 5: Output file corresponding to the family of Figure 1 assuming mother of the counseland is cancer free, but her father was diagnosed with breast cancer at age 70, and lived until 87.

FRACTION OF BRCA1 CARRIERS, GENERAL POPULATION: 0.001200
 FRACTION OF BRCA2 CARRIERS, GENERAL POPULATION: 0.000440

JOINT PROBABILITY OF CARRYING BRCA1 AND BRCA2

		BRCA1			
		0	1	2	Marginal
B	0	0.046073	0.446525	0.000024	0.492621
C	1	0.507092	0.000234	0.000000	0.507326
A	2	0.000053	0.000000	0.000000	0.000053
2	Marginal	0.553218	0.446759	0.000024	1

PROBABILITY OF BEING A BRCA1 CARRIER: 0.446782
 PROBABILITY OF BEING A BRCA2 CARRIER: 0.507379

Figure 6: Output file corresponding to the family of Figure 1 after the 32 year old daughter of the counseland tested negative for BRCA1.

Two more variants on the example

We are now going to look at two more variants of our example, to illustrate genetic testing and male breast cancer. In the first example, we know in addition that the 32 year old daughter of the counseland tested negative for BRCA1. She hasn't been tested for BRCA2. The fourth line of the matrix in the family input file should be changed to:

4 3 0 0 1 0 0 32 32 0 2 0

The new results are shown in Figure 6.

Next we consider the case in which the mother of the counseland is cancer free, but her father was diagnosed with breast cancer at age 70, and lived until 87. Lines 2 and 3 of the matrix in the family input file should be changed to:

2 4 0 9 8 0 0 70 70 0 0 0
 3 4 1 11 10 1 0 70 87 0 0 0

The new results are shown in Figure 7.

FRACTION OF BRCA1 CARRIERS, GENERAL POPULATION: 0.001200
 FRACTION OF BRCA2 CARRIERS, GENERAL POPULATION: 0.000440

JOINT PROBABILITY OF CARRYING BRCA1 AND BRCA2

		BRCA1			
		0	1	2	Marginal
B					
R	0	0.038890	0.014096	0.000002	0.052988
C	1	0.946185	0.000558	0.000000	0.946743
A	2	0.000269	0.000000	0.000000	0.000269
2					
	Marginal	0.985345	0.014653	0.000002	1

PROBABILITY OF BEING A BRCA1 CARRIER: 0.014655
 PROBABILITY OF BEING A BRCA2 CARRIER: 0.947012

Figure 7: Output file corresponding to the family of Figure 1 assuming mother of the counselland is cancer free, but her father was diagnosed with breast cancer at age 70, and lived until 87.