



## **A new scoring system in cancer genetics: application to criteria for BRCA1 and BRCA2 mutation screening.**

Bernard Bonaïti, Flora Alarcon, Nadine Andrieu, Valérie Bonadona, Marie-Gabrielle Dondon, Sophie Penneç, Dominique Stoppa-Lyonnet, Catherine Bonaïti-Pellié, Hervé Perdry

### ► **To cite this version:**

Bernard Bonaïti, Flora Alarcon, Nadine Andrieu, Valérie Bonadona, Marie-Gabrielle Dondon, et al.. A new scoring system in cancer genetics: application to criteria for BRCA1 and BRCA2 mutation screening.. Journal of Medical Genetics, BMJ Publishing Group, 2013, 51 (2), epub ahead of print. 10.1136/jmedgenet-2013-101674 . hal-00923931

**HAL Id: hal-00923931**

**<https://hal.archives-ouvertes.fr/hal-00923931>**

Submitted on 6 Jan 2014

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

## ORIGINAL ARTICLE

A new scoring system in cancer genetics: application to criteria for *BRCA1* and *BRCA2* mutation screeningBernard Bonaïti,<sup>1,2</sup> Flora Alarcon,<sup>3,4,5</sup> Nadine Andrieu,<sup>5,6,7</sup> Valérie Bonadona,<sup>8,9,10</sup> Marie-Gabrielle Dondon,<sup>5,6,7</sup> Sophie Penneç,<sup>11</sup> Dominique Stoppa-Lyonnet,<sup>4,6,12</sup> Catherine Bonaïti-Pellié,<sup>2,13</sup> Hervé Perdry<sup>2,13</sup>

► Additional material is published online only. To view please visit the journal online (<http://dx.doi.org/10.1136/jmedgenet-2013-101674>).

For numbered affiliations see end of article.

**Correspondence to**

Dr Hervé Perdry, INSERM, UMR-S 669, Hôpital Paul Brousse, Bâtiment Inserm 15-16, Villejuif F-94807, France; [herve.perdry@inserm.fr](mailto:herve.perdry@inserm.fr)

Received 19 March 2013

Revised 8 November 2013

Accepted 14 November 2013

**ABSTRACT**

**Background** In hereditary forms of cancer due to mutations of genes such as *BRCA1* and *BRCA2*, methods have been proposed to predict the presence of a mutation in a family.

**Methods** Relying on carriage probability computation is the most predictive, but scores are a good proxy and avoid using computer software. An empirical method, the Manchester scoring system, has been elaborated for *BRCA1* and *BRCA2* mutation identification. We propose a general scoring system based on a transformation of the carriage probability. Up to an approximation, the transformed carriage probability becomes an additive score. We applied this new scoring system to the diagnosis of *BRCA1*-associated and *BRCA2*-associated breast-ovarian cancer predisposition. Using simulations, its performance was evaluated and compared with that of the Manchester scoring system and of the exact probability. Finally, the score system was used on a sample of 4563 families screened for *BRCA1* and *BRCA2* mutations.

**Results** The performance of the new scoring system was superior to the Manchester scoring system, but the probability computation remained the most predictive. The better performance of the new scoring system was attributed to accounting for unaffected family members and for the degree of kinship of relatives with the proband.

**Conclusions** The new scoring system has a theoretical basis and may be applied to any cancer family syndrome and, more generally, to any disease with monogenic subentities, in which the causal gene mutations have been identified. It will be easily modified when additional predictive factors are found.

**INTRODUCTION**

Hereditary forms of cancer have been described for many years, and some of them are due to known rare mutations of genes, such as *BRCA1* and *BRCA2* for breast-ovarian cancer and the mismatch repair genes for colorectal and gynaecological cancers. Individuals who have inherited a deleterious mutation have a highly increased risk of developing cancer.<sup>1 2</sup> It is essential to identify the families in which such mutations segregate to provide efficient counselling to individuals at risk and to ensure appropriate medical management of mutation carrier individuals.<sup>3-6</sup>

The low carriage frequency, the high incidence of these cancers and the high cost of testing render the exhaustive screening of mutations in the

population impractical. Hence, several models that use individual (age at onset, type of cancer, etc.) and family information have been proposed to more efficiently target the high-risk population, and criteria for recommendation of genetic testing to affected individuals have been defined.<sup>7</sup>

For *BRCA1* and *BRCA2* mutations, some of these criteria are based on carriage probability computed according to characteristics of individuals in a family through a variety of algorithms.<sup>8-10</sup> Another approach, the Manchester scoring system,<sup>11 12</sup> allocates a score to each affected individual in a family and computes the sum of the scores in the maternal and paternal lineage; if one of these two sums exceeds a given threshold, usually corresponding to a 10-20% mutation probability, a genetic test is recommended. Finally, some recommendations for genetic testing are based on a list of specific family configurations regarding affection status and age at onset of family members. They are the most commonly used in France.<sup>13 14</sup> Methods that compute the carriage probabilities are the most predictive ones, but they are not used by the majority of geneticists (genetic counsellors), particularly in France, because they require using a computer software and data entering for all family members, which they consider time consuming.<sup>14</sup> Scoring systems avoid data entry and are a good proxy to probability computations. Moreover, they may easily be modified to include potential new relevant predictive factors.<sup>15</sup>

The Manchester scoring system is empirical as the scores have been determined to fit the observations made in a group of selected families tested for *BRCA1* and *BRCA2* mutations. Moreover, it does not take into account information on unaffected family members, the presence of which is expected to decrease the probability of mutation, and gives the same weight to all affected family members whatever their degree of kinship.

In Bonaïti *et al*,<sup>14</sup> we advocated to the French physicians the use of a new scoring system for the diagnosis of *BRCA1*-associated or *BRCA2*-associated (referred as *BRCA1/2*) breast-ovarian cancer predisposition. In this scoring system, the scores can be calculated by the physician in a similar way to the Manchester scores. The presentation of the theoretical foundations on which it relies was deferred to a subsequent publication. The aim of the present study is to give a detailed description of this general scoring system, which is based on an approximation of the probability of an affected individual to be a

**To cite:** Bonaïti B, Alarcon F, Andrieu N, *et al*. *J Med Genet* Published Online First: [please include Day Month Year] doi:10.1136/jmedgenet-2013-101674

carrier conditional to his/her family history, and to compare the performance of this new scoring system with the Manchester scoring system and to the exact probability. We also used the score system on a sample of 4563 families from the Institut Curie in Paris tested for *BRCA1/2* mutations and present the results here.

## METHODS

### Disease model

We consider a predisposing cancer syndrome, due to a dominant mutated gene, involved in various tumour localisations  $j$  and various tumour characteristics  $c$ , such as the 'triple-negative' (TN) phenotype in breast cancer<sup>16</sup> or the MSI phenotype in colorectal cancer.<sup>17</sup> Let  $F(t; g, j)$  be the cumulative risk of developing a tumour at localisation  $j$  by age  $t$  (in years) for carrier status  $g$  ( $g=1$  for a carrier and  $g=0$  for a non-carrier). For affected individuals, the probability of tumour characteristic  $c$  is designed by  $\rho(c; g, j)$ . An observation is defined by  $(s, c, t)$ , where  $s$  equals 1 or 0 according to the affected status of the individual, and  $t$  is age at onset for affected persons and censored age for unaffected ones. The probability  $\Phi(s, c, t; g, j)$  of this observation depends on carrier status  $g$  and on localisation  $j$ :

$$\begin{aligned} \Phi(s, c, t; g, j) &= (F(t+1; g, j) - F(t; g, j))\rho(c; g, j) & \text{if } s = 1 \\ &= 1 - F(t+1; g, j) & \text{if } s = 0 \end{aligned}$$

Sex is directly taken into account by considering different values of  $j$  for men and women for a same localisation.

### Elaboration of the scoring system

The scoring system is proposed as a surrogate for the probability of a proband, in a given family, to be a mutation carrier. This scoring system is derived from a conditional probability  $P$  which is evaluated by taking into account the proband's personal and family history of the disease and may be written according to Bayes' theorem as

$$P = \Pr(g_1 = 1|H) = \frac{p}{p + (1-p) \frac{\Pr(H|g_1=0)}{\Pr(H|g_1=1)}}$$

where  $H$  represents the personal and family history, that is, all the observations  $(s_{ij}, c_{ij}, t_{ij})$  for each localisation  $j$  and each family member  $i$ ,  $g_1$  is the carrier status of the proband and  $p$  is the carrier frequency in the population.

A first approximation consists in considering that two family members have no other common relative than the proband and that  $\Pr(g_i = 1|g_1 = 1) = 0.5^{d_i}$ , where  $d_i$  is the degree of kinship of family member  $i$  with the proband (0 for the proband, 1 for first-degree relatives, 2 for second-degree relatives, etc.). We obtain

$$\Pr(H|g_1 = 0) \simeq \prod_i \prod_j \Phi(s_{ij}, c_{ij}, t_{ij}; 0, j)$$

$$\Pr(H|g_1 = 1) \simeq$$

$$\prod_i \left( 0.5^{d_i} \prod_j \Phi(s_{ij}, c_{ij}, t_{ij}; 1, j) + (1 - 0.5^{d_i}) \prod_j \Phi(s_{ij}, c_{ij}, t_{ij}; 0, j) \right)$$

Hence, we can write

$$\frac{\Pr(H|g_1 = 1)}{\Pr(H|g_1 = 0)} \simeq \prod_i \left( 1 + 0.5^{d_i} \left( \prod_j \frac{\Phi(s_{ij}, c_{ij}, t_{ij}; 1, j)}{\Phi(s_{ij}, c_{ij}, t_{ij}; 0, j)} - 1 \right) \right),$$

and

$$\begin{aligned} S &:= \log \left( \frac{\Pr(H|g_1=1)}{\Pr(H|g_1=0)} \right) \simeq \sum_i \log \left( 1 + 0.5^{d_i} \left( \prod_j \frac{\Phi(s_{ij}, c_{ij}, t_{ij}; 1, j)}{\Phi(s_{ij}, c_{ij}, t_{ij}; 0, j)} - 1 \right) \right), \\ P &\simeq \frac{p}{p + (1-p)\exp(-S)}. \end{aligned}$$

The sum over all family members in  $S$ , which we call the score of the family, is then approximated as

$$\begin{aligned} S &\simeq \sum_i \left[ \Gamma_i \sum_j I_{ij} \log \left( 1 + 0.5^{d_i} (\Lambda(s_{ij}, c_{ij}, t_{ij}; j) - 1) \right) \right. \\ &\quad \left. + (1 - \Gamma_i) \log \left( 1 + 0.5^{d_i} (\Lambda(s_{ij}, c_{ij}, t_{ij}; j) - 1) \right) \right] \end{aligned}$$

where  $\Gamma_i$  and  $I_{ij}$  are two dummy variables:  $\Gamma_i=1$  if the individual is affected by at least one localisation and 0 if not,  $I_{ij}=1$  if the individual is affected by localisation  $j$  and 0 if unaffected, and

$$\Lambda(s_{ij}, c_{ij}, t_{ij}; j) = \frac{\Phi(s_{ij}, c_{ij}, t_{ij}; 1, j)}{\Phi(s_{ij}, c_{ij}, t_{ij}; 0, j)}.$$

If  $s=1$ , this ratio is the relative risk for a carrier of being affected by a tumour of localisation  $j$ .

Using this approximation (S), one neglects the fact that a person affected with a given cancer is unaffected by the other cancer localisations and considers that cancers for a given individual occur independently.

The computation of the global score  $S$  for a family is then obtained by summing the elementary scores:

- ▶  $s(t, j, c) = \log(1 + 0.5^{d_i} (\Lambda(1, c, t; j) - 1))$  for each occurrence of a cancer in the family
- ▶  $sn(t, x) = \log(1 + 0.5^{d_i} (\prod_{j \in J_x} \Lambda(0, 0, t; j) - 1))$  for each unaffected individual of sex  $x$ , where  $J_x$  represents the localisations related to sex  $x$

If there are mutations on several genes, we propose to use for  $\Lambda(s_{ij}, c_{ij}, t_{ij}; j)$  the average value in carriers of the different gene mutations weighted by their frequency in the population.

Finally, we note the two reciprocal relationships between  $S$  and  $P$ :

$$P \simeq \frac{p}{p + (1-p)\exp(-S)}$$

and

$$S \simeq -\log \left( \frac{(1-P)/P}{(1-p)/p} \right).$$

### Application to the diagnosis of *BRCA1/2*-associated breast-ovarian cancer predisposition

Carriers of *BRCA1/2* gene mutations have an increased risk not only of breast and ovarian cancer but also of pancreas and prostate cancer.<sup>18 19</sup> In addition, it has been shown that the type of breast pathology, particularly TN tumours, that is, negative for oestrogen receptor (ER), progesterone receptor (PR) and *HER2* expression, influences the probability of detecting a *BRCA1* mutation, but not a *BRCA2* mutation.<sup>16</sup>

For assessing the ratios of probabilities  $\Lambda(s, c, t; j)$  necessary to the determination of the scores, penetrance values were taken from the meta-analysis of Chen and Parmigiani<sup>1</sup> for female breast cancer and ovarian cancer, from Tai *et al*<sup>20</sup> for male breast cancer and from van Asperen *et al*<sup>19</sup> and Risch *et al*<sup>18</sup> for

pancreas and prostate cancers. In cases where the risks at greater ages were not provided by these sources, we inferred the missing data from available data by fitting a mixture model with a subgroup of non-susceptible carrier individuals and a subgroup of susceptible carriers for which the risk is modelled by a Weibull function. This model is very similar to the 'cured model' of de Angelis,<sup>21</sup> also used for penetrance estimation in the genotype-restricted likelihood method.<sup>22</sup> Cumulative risks in the general population used were those published for France based on data of cancer registries.<sup>23</sup>

Probabilities for the TN status of a breast tumour were obtained from the proportion of TN tumours among women tested for *BRCA1/2* data in studies without any selection on morphological characteristics.<sup>15 24 25</sup> Weighting these proportions by the number of women, we obtained an average proportion of TN of 68% among women with *BRCA1* mutation and 13% among women with no *BRCA1* mutation (including those with a *BRCA2* mutation).<sup>14</sup>

Mutations in *BRCA1* and *BRCA2* are rare and involved in breast cancer in almost equal proportions<sup>26</sup> for the two genes. Neglecting the very infrequent occurrence where individuals are carriers for both genes, one on *BRCA1* and one on *BRCA2*, we computed  $\Lambda(s_{ij}, c_{ij}, t_{ij}; j)$  as the average of risks associated with the mutations of the two genes weighted by the frequency of mutation carriers on each gene. For allele frequencies, estimates as different as 0.001<sup>26</sup> and 0.005<sup>18</sup> for both genes were found. We used an intermediate value of 0.001 for each gene, which was shown to best fit published data<sup>27-31</sup> for frequency of carriers among affected individuals.<sup>14</sup>

### Evaluation of the performance of the new scoring system

The performance of the new scoring system was evaluated using simulations of families with breast and/or ovarian cancer. The main advantage of simulations compared with observed data is to provide an estimation of absolute sensitivity, that is, the proportion of carrier individuals that fulfil the criteria recommended for genetic testing among all affected ones, whereas observed data only provide an estimation of relative sensitivity of criteria among the only families that underwent genetic testing.

The simulation of pedigrees was performed in order to be representative of families of individuals affected by breast-ovarian cancer in the French general population. We considered an affected case, the proband, his/her age (not exceeding 80 years); sex and phenotype (breast, including male breast cancer, or ovarian cancer) were drawn at random according to French cancer registries.<sup>23</sup> His/her genotype (carrier/non-carrier) was drawn at random according to the conditional probability of genotypes, given the disease status obtained from the frequencies of the mutated alleles and the penetrance values mentioned above, using Bayes' theorem. For breast cancer, the TN status of the tumour was simulated according to its probability conditional on genotype, as mentioned above.

The pedigree structure was built using demographic data for France.<sup>32 33</sup> In order to take into account the impact of mortality of mother on sibship structure, the size and structure of sibships were defined as follows:

Women are fertile between 15 and 49 years.

Maternal age at first birth was taken as  $x_1 = 16 + X$ , where  $X$  is a random variable from an exponential distribution with parameter  $\lambda = 0.1$ . This parameter value was chosen to fit the age at first birth in France according to the birth cohort of the mother.

For  $i \geq 1$ , age of women at the  $(i+1)$ th birth was  $x_{i+1} = x_i + 1 + X$ , where  $X$  follows an exponential distribution with parameter  $\lambda$  depending on the birth cohort of the mother. The value of  $\lambda$  was chosen to fit female fertility in France. Note that we use  $x_{i+1}$  instead of  $x_i$  to account for the minimum interval between two births at 1 year.

Fertility could be interrupted either by death or by reaching the end of reproductive period (age exceeding 49).

The birth ranks of the proband and of his/her parents within their own sibships were set at random from a uniform distribution.

The resulting pedigrees include at least three generations, from the proband's grandparents to the proband and his/her first cousins, and, if the proband and his/her cousins are old enough to have children, the pedigrees include a fourth or even a fifth generation.

The genotype of each member of the family conditional on the proband's genotype was simulated following Mendel laws and using Bayes' theorem. Their phenotype (disease status and tumour characteristics) was simulated using the parameters mentioned above.

We simulated a total of 10 000 000 families, and we computed for each family its conditional probability  $P$  that the proband is a mutation carrier given the personal and family history of the disease (H), without knowledge of the simulated genotype, the Manchester combined score (MCS)<sup>15</sup> and the score obtained by the new system (NSS). The probability  $P$  was computed using the Elston-Stewart algorithm<sup>34</sup> and is referred as  $P_{\text{Probability}}$  computed using the Elston-Stewart algorithm (PREST). For each of the three methods (PREST, MCS, NSS), the sensitivity was estimated as the proportion of families with a score (or a probability  $P$ ) above a given threshold, among families with a mutation carrier proband. Similarly, the specificity at this threshold was estimated as the proportion of families with a score (or a probability  $P$ ) below the threshold, among families with a non-carrier proband. The positive predictive value (PPV) was estimated as the proportion of carriers among the families exceeding this threshold. The performance of the three systems was compared by receiver operating characteristic (ROC) curves, which represent sensitivity as a function of one minus specificity for various threshold values, and also by the curve of PPV as a function of sensitivity.

### Family sample

The scoring system was used on a sample 4563 families screened for *BRCA1* and *BRCA2* mutations between 1995 and 2012 in the Institut Curie in Paris. These families have been tested following the current recommended criteria in France.<sup>3 13 35</sup> Proband's had given their informed consent for *BRCA1/2* genetic testing in the setting of a family cancer clinic and their consent to the use of their anonymised data. Proband's and relatives who gave authorisations to their medical record access were informed of the computer recording of their family history. The family database had been declared to the French national ad hoc authority (CNIL). As the study does not involve intervention on patients, there was no need to consult an ethics committee. The point and small-size mutation analyses of the *BRCA1* and *BRCA2* genes have been performed from 1995 to the present by different screening methods, followed by Sanger sequencing of amplicons showing abnormal screening patterns: DGGE, DHPLC and EMMA. Regarding large gene rearrangements of both genes, they have been searched using DNA combing, and later QMPSF and EMMA.

## RESULTS

## Simulated families

The values of the new *BRCA1/2* mutation scores for the different members of a family, according to cancer localisation and type (TN status for breast cancer), age class at diagnosis and degree of kinship with the proband are given in table 1 for affected members and in table 2 for unaffected members. The score values displayed were obtained by multiplying the original scores by 2 and by rounding them to integer values in order to be easily summed by hand. As expected, for cancer at a given age and localisation, the proband always has the highest score compared with his/her relatives, and the scores of relatives decrease with increasing degree of kinship with the proband.

Figure 1 shows the ROC curves for the new scoring system, for the PREST probability and for the Manchester scoring system. Figure 2 shows PPV as a function of sensitivity for these three systems. In figures 1 and 2, the PREST curves show that the PREST method performs the best. The NSS curves are consistently closer to the PREST curves than the MCS curves. The area under the ROC curve is 0.92 for PREST, 0.88 for the Manchester scoring system and 0.90 for the new system.

To investigate the performance gained in accounting for unaffected family members and for the degree of kinship of relatives with the proband, we compared the PPV curves of the NSS with that of a NSS modified to ignore either unaffected family members or kinship coefficients, and finally both. Each information type alone slightly improves the prediction and

**Table 1** New scores for the proband and affected relatives in the family (reprinted by permission from *Bulletin du Cancer*)

	Proband	First-degree relative	Second-degree relative	Third-degree relative
Women with breast cancer according to age at onset (years)				
<20	11	9	8	7
20–29	7	6	5	4
30–39	5	4	3	2
40–49	4	3	2	1
50–59	4	3	2	1
60–69	2	2	1	1
≥70	0	0	0	0
Women with ovarian cancer according to age at onset (years)				
<20	5	4	3	2
20–29	7	6	5	3
30–39	8	7	6	4
40–59	7	6	5	4
≥60	6	5	4	2
Men with breast cancer according to age at onset (years)				
<30	23	22	21	19
30–39	16	14	13	11
40–49	11	10	8	7
50–59	10	9	8	6
60–69	8	7	6	4
≥70	7	6	5	3
Triple-negative tumours (if information available)				
Yes	3	2	1	1
No	-2	-1	0	0
Prostate cancer	2	1	0	0
Pancreas cancer	3	2	1	0

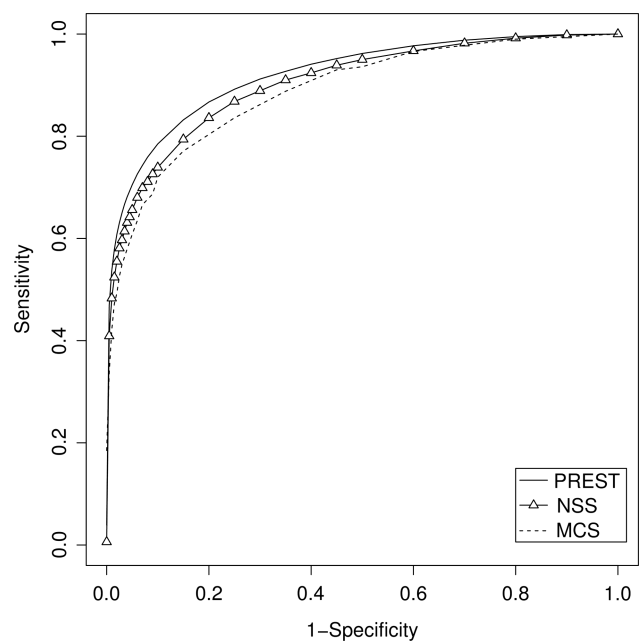
**Table 2** New scores for unaffected relatives (reprinted by permission from *Bulletin du Cancer*)

Age class (years)	Women			Men		
	First degree	Second degree	Third degree	First degree	Second degree	Third degree
<30	0	0	0	0	0	0
30–49	-0.3	-0.1	0	0	0	0
50–69	-0.7	-0.4	-0.2	0	0	0
≥70	-1	-0.4	-0.2	-0.2	0	0

taken into account all together substantially increases the performance of the system, as shown by the PPV curves (figure 3).

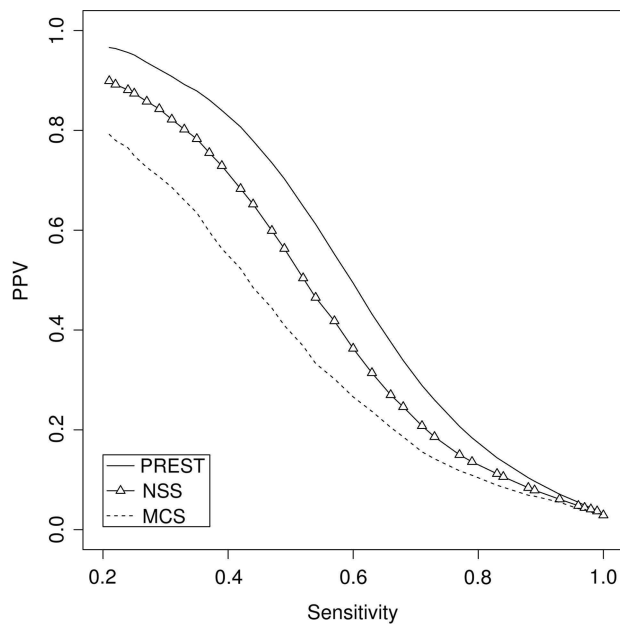
Table 3 gives the average probability  $Pr_0$  obtained by PREST for various score values of the simulated families, the probability  $Pr_1$  estimated from the proportion of simulated families with a mutation among those with a score within each score class, and the probability  $Pr_2$  expected from the approximations made for elaborating the score (see the relationship between P and S in the 'Methods' section). The observed proportion  $Pr_1$  is very close to the  $Pr_0$  probability calculated by PREST for all values of scores. The probability  $Pr_2$  is very close to  $Pr_0$  and  $Pr_1$  for moderate values of scores but slightly deviates from these probabilities as the score increases.

Table 4 shows the sensitivity, PPV and specificity of the NSS for various values of the score threshold recommended for genetic testing. For comparison, we calculated these parameters for the Manchester scoring system with a threshold of 16 assumed to correspond to a threshold mutation probability of 10%. The sensitivity was 72%, the PPV was 16% and the specificity was 88%. We obtained similar PPV and specificity with the new system for a score threshold of 5, but the sensitivity was higher at 77% (see table 4). We recommend this score threshold of 5 to ensure a good sensitivity.



**Figure 1** Receiver operating characteristic curves for the probability of mutation (PREST), the new scoring system (NSS) and the Manchester combined score system (MCS).



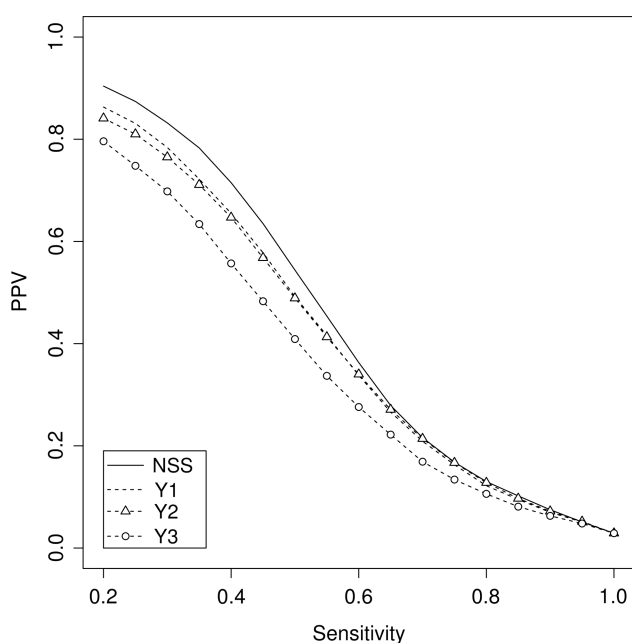


**Figure 2** Positive predictive value as a function of sensitivity for the probability of mutation (PREST), the new scoring system (NSS) and the Manchester combined score system (MCS).

The online supplementary appendix gives an example of computation of score in a family. The maternal branch reaches the score of 9 and would lead to *BRCA1/2* screening as it largely exceeds the score threshold of 5 that we recommend. The Manchester score would be 16, which is exactly the threshold recommended for mutation screening.

#### Families screened at the Institut Curie

Both NSS and PREST were computed in the 4563 families. The scores range from  $-4$  to 31 (to the exception of a single family



**Figure 3** Positive predictive value as a function of sensitivity for the probability of mutation (PREST), the new scoring system (NSS) and the Manchester combined score system (MCS).

**Table 3** Probability of mutation according to score values in the new system

Score class	Pr0	Pr1	Pr2
[0, 1[	0.01	0.01	0.00
[1, 2[	0.01	0.01	0.00
[2, 3[	0.01	0.01	0.01
[3, 4[	0.02	0.02	0.01
[4, 5[	0.03	0.03	0.02
[5, 6[	0.03	0.04	0.03
[6, 7[	0.05	0.05	0.05
[7, 8[	0.08	0.08	0.08
[8, 9[	0.12	0.13	0.12
[9, 10[	0.18	0.18	0.18
[10, 11[	0.25	0.26	0.27
[11, 12[	0.35	0.37	0.38
[12, 13[	0.47	0.49	0.50
[13, 14[	0.57	0.58	0.62
[14, 15[	0.66	0.67	0.73
[15–16[	0.74	0.75	0.82

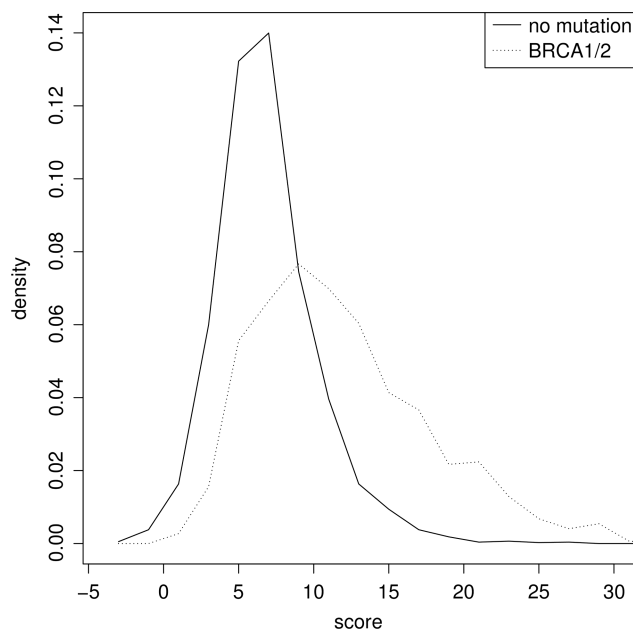
Pr0, probability of mutation obtained by PREST; Pr1, probability estimated from the proportion of simulated families with a mutation; Pr2, probability expected from the approximation made for elaborating the score.

with a score of 48, which was discarded from the plots). Figure 4 shows the frequency polygons of the score values, separately in families where no mutation was found and in families where a *BRCA1/2* mutation was found. Figure 5 illustrates the good concordance of the score and PREST values by showing the score plotted against the transformation of PREST (P) given at the very end of the 'Methods' section.

Table 5 shows the proportion of families with a mutation according to score value and score threshold. The proportion of mutations among families according to score value is close to the results obtained from simulated families, except for low-score values. On the other hand, the proportion of mutations according to score threshold is greater than the VPP obtained from simulated families. This is most likely due to the selection of families in which the prescription of a genetic test was based on stringent criteria that favoured very high scores.

**Table 4** Performance of the new scoring system according to threshold for recommending genetic testing

Score threshold	Sensitivity	Positive predictive value	specificity
1	0.93	0.06	0.57
2	0.89	0.08	0.69
3	0.86	0.10	0.76
4	0.82	0.12	0.82
5	0.77	0.15	0.87
6	0.71	0.21	0.92
7	0.65	0.29	0.95
8	0.59	0.38	0.97
9	0.54	0.48	0.98
10	0.49	0.57	0.99
11	0.44	0.66	0.99
12	0.39	0.74	0.99
13	0.34	0.79	0.99
14	0.30	0.84	0.99
15	0.26	0.87	0.99

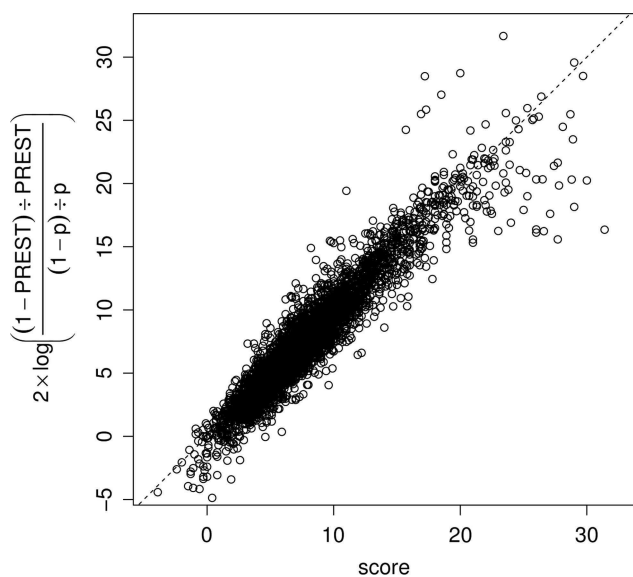


**Figure 4** Frequency polygon of the score values in families where no mutation was detected and in families where a *BRCA1/2* mutation was detected.

## DISCUSSION

Our new scoring system is based on the conditional probability  $P$  that a proband is a carrier, given all relevant predictive information in the family. It is therefore possible to apply it to other cancer family syndromes such as Lynch syndrome or Li-Fraumeni syndrome, and even any disease with monogenic subtentities due to specific gene mutations, such as Alzheimer or Parkinson disease. The only requirement is to know the relative risks of disease(s) in mutation carriers associated with predictive factors and the frequency of mutation carriers in the general population.

In breast-ovarian cancer, the most commonly used system for identifying *BRCA1/2* mutation carriers in France is based on



**Figure 5** The score values for the 4563 families plotted against the transformation of PREST, which is approximated by the score.

**Table 5** Proportion of families of the Institut Curie with a mutation according to score value and score threshold

Score class	Number of families	% with mutation	Score threshold	Number of families	% with mutation
<1	68	1.5	1	4495	16.4
[1, 2[	81	2.5	2	4414	16.6
[2, 3[	137	4.4	3	4277	17.0
[3, 4[	287	5.2	4	3990	17.9
[4, 5[	580	6.4	5	3410	19.8
[5, 6[	503	7.8	6	2907	21.9
[6, 7[	645	8.1	7	2262	25.9
[7, 8[	543	9.8	8	1719	30.9
[8, 9[	339	13.0	9	1380	35.4
[9, 10[	342	17.5	10	1038	41.2
[10, 11[	266	24.8	11	772	46.9
[11, 12[	173	24.9	12	599	53.3
[12, 13[	135	31.1	13	464	59.7
[13, 14[	94	41.5	14	270	64.3
[14, 15[	76	61.8	15	294	65.0
≥15	294	65.0	16	233	73.0

family configurations.<sup>13 14</sup> This system was shown to perform rather well, but its performance is difficult to evaluate when new relevant criteria are added. Our new scoring system was shown to be an efficient alternative.<sup>14</sup> Moreover, as any scoring system, it may be easily modified if new predictive factors are found and relevant information is available. In addition, the score threshold of 5 that we recommended may be easily modified if mutation screening cost decreases as expected in the near future.

An advantage of our method is to rely on theoretical foundations and not on an empirical calculation of scores from a sample of families selected on familial aggregation of breast or ovarian cancer cases. Indeed, a method based on observed proportion of mutations in such families should correct for the criteria recommended for genetic testing. This would probably be not even sufficient since not all families fulfilling these criteria are expected to undergo genetic testing. Such a bias likely explains the discrepancy in the proportion of mutations in families with low-score values between the families screened at the Institut Curie and those from the simulated sample.

The improvement of our new scoring system compared with the Manchester scoring system is mainly the result of incorporating degrees of kinship of relatives with the proband and unaffected relatives. This valuable improvement comes at the cost of a moderately more complicated computation of scores of families. In the pedigree given in the online supplementary appendix, the computation of scores may be achieved in a few seconds using the Manchester scoring system and in about 2 min with the new one. To overcome this complication, we proposed to use a two-step procedure, with first a rough screening on simple criteria<sup>3</sup> by physicians for recommending genetic counselling followed, if judged relevant, by a fine evaluation by genetic counsellors using the new scoring system for recommending genetic testing.<sup>14</sup>

The approximation of the probability  $P$  that was made for elaborating the score system was found to be valid for moderate values of scores, but slightly differed from the true value as the score increased. This is only a minor drawback as any system would recommend genetic testing to the families that have such high-score values.

Another limit is that the system relies on uncertain values for disease risks. However, these risks should be known with increasing precision in the future and it will be possible to modify the score values and threshold accordingly. Note that this limit also holds for methods based on probability computations. This uncertainty on risks estimates might also affect the simulations but not the comparison between PREST and the score systems, given that the same parameters were used.

Finally, the model on which the scoring system is elaborated does not account for familial factors other than *BRCA1* and *BRCA2* that could be another source of familial aggregation. This is done, for instance, in the BOADICEA model, which includes, in addition to *BRCA1/2* genes, a polygenic component<sup>9 26</sup> and was shown to perform better for carrier prediction than other algorithms.<sup>36</sup> Such an addition would have considerably complicated the elaboration of the score system and was not expected to have any impact on sensitivity, but might have slightly affected PPV and specificity.

In the present study, we used an original method that generates families as close as possible to real families in the general population, particularly by taking into account mortality at childbearing ages. The interest of using simulations is to provide estimations of absolute sensitivity, whereas observed data only provide relative sensitivity. For instance, using simulations permitted to demonstrate that criteria based on positive family history for *BRCA1/2* genetic testing have low sensitivity.<sup>37</sup> Estimation of absolute sensitivity is important as recommendations for genetic testing more and more focus on this parameter, and not only on PPV.<sup>35</sup> In this study, we have shown that the usual threshold probability of mutation of 10%, corresponding to a score threshold of about 8 in the new system (table 3), would lead to a low sensitivity of less than 60% (table 4). In addition, we must keep in mind that the observed proportion of families in which a mutation was found may be a poor estimation of PPV as they are most likely selected towards high values of scores. Indeed, families for which a test is recommended may not all undergo genetic testing, particularly when only a few family members are affected. Therefore, families with a score greater than a given threshold, say 5, have an overrepresentation of families with higher scores and are not representative of all families exceeding a score threshold of 5. Therefore, the value of 15% for PPV found from the simulations is probably a better estimation of this parameter than the proportion of nearly 20% found in the sample.

#### Author affiliations

<sup>1</sup>INRA-GABI, Jouy-en-Josas, France

<sup>2</sup>INSERM, UMR-S 669, Villejuif, France

<sup>3</sup>MAP5, CNRS UMR 8145, Paris, France

<sup>4</sup>Université Paris-Descartes, Paris, France

<sup>5</sup>INSERM, U900, Paris, France

<sup>6</sup>Institut Curie, Paris, France

<sup>7</sup>École des Mines de Paris, ParisTech, Fontainebleau, France

<sup>8</sup>Université Lyon 1, Lyon, France

<sup>9</sup>CNRS UMR 5558, Lyon, France

<sup>10</sup>Centre Léon Bérard, Lyon, France

<sup>11</sup>INED, Paris, France

<sup>12</sup>INSERM, U830, Paris, France

<sup>13</sup>Université Paris-Sud, Villejuif, France

**Contributors** BB was responsible for design of the scoring system, design of the study, computer simulation study and writing of the paper. FA and HP were responsible for design of the study, computer simulation study and writing of the paper. NA and DS-L were responsible for design of the study, data collection and writing of the paper. VB was responsible for design of the study and writing of the paper. M-GD was responsible for data management and writing of the paper. SP was responsible for design of the study and computer simulation study. CB-P was responsible for design of the scoring system, design of the study and writing of the paper.

**Competing interests** None.

**Ethics approval** CNIL.

**Provenance and peer review** Not commissioned; externally peer reviewed.

**Data sharing statement** The cancer familial data will not be shared.

#### REFERENCES

- Chen S, Parmigiani G. Meta-analysis of *BRCA1* and *BRCA2* penetrance. *J Clin Oncol* 2007;25:1329–33.
- Bonadona V, Bonaïti B, Olschwang S, Grandjouan S, Huiart L, Longy M, Guimbaud R, Buecher B, Bignon Y-J, Caron O, Colas C, Noguès C, Lejeune-Dumoulin S, Olivier-Faivre L. Cancer risks associated with germline mutations in *MLH1*, *MSH2*, and *MSH6* genes in Lynch syndrome. *JAMA* 2011;305:2304–10.
- Eisinger F, Bressac B, Castaigne D, Cottu PH, Lansac J, Lefranc JP, Lesur A, Noguès C, Pierret J, Puy-Pernias S, Sobol H, Tardivon A, Tristant H, Villet R. Identification et prise en charge des prédispositions héréditaires aux cancers du sein et de l'ovaire (mise à jour 2004). *Bull Cancer* 2004;91:219–37.
- Robson M, Offit K. Management of an inherited predisposition to breast cancer. *N Engl J Med* 2007;357:154–62.
- Olschwang S, Bonaïti C, Feingold J, Frebourg T, Grandjouan S, Lasset C, Laurent-Puig P, Lecuru F, Millat B, Sobol H, Thomas G, Eisinger F. Identification et prise en charge du syndrome HNPCC (Hereditary NonPolyposis Colon Cancer), prédisposition héréditaire aux cancers du côlon, du rectum et de l'utérus. *Bull Cancer* 2004;91:303–15.
- Lindor NM, Petersen GM, Hadley DW, Kinney AY, Miesfeldt S, Lu KH, Lynch P, Burke W, Press N. Recommendations for the care of individuals with an inherited predisposition to Lynch syndrome: a systematic review. *JAMA* 2006;296:1507–17.
- Antoniou AC, Easton DF. Risk prediction models for familial breast cancer. *Future Oncol* 2006;2:257–74.
- Parmigiani G, Berry D, Aguilar O. Determining carrier probabilities for breast cancer-susceptibility genes *BRCA1* and *BRCA2*. *Am J Hum Genet* 1998;62:145–58.
- Antoniou AC, Pharoah PP, Smith P, Easton DF. The BOADICEA model of genetic susceptibility to breast and ovarian cancer. *Br J Cancer* 2004;91:1580–90.
- Tyrer J, Duffy SW, Cuzick J. A breast cancer prediction model incorporating familial and personal risk factors. *Stat Med* 2004;23:1111–1130.
- Evans DG, Eccles DM, Rahman N, Young K, Bulman M, Amir E, Shenton A, Howell A, Lalloo F. A new scoring system for the chances of identifying a *BRCA1/2* mutation outperforms existing models including BRCAPRO. *J Med Genet* 2004;41:474–80.
- Evans DG, Lalloo F, Wallace A, Rahman N. Update on the Manchester Scoring System for *BRCA1* and *BRCA2* testing. *J Med Genet* 2005;42:e39.
- Stoppa-Lyonnet D, Noguès C, Sobol H, Eisinger F. Consultations de génétique et prédisposition aux cancers du sein. In *Risques héréditaires de cancers du sein et de l'ovaire: Quelle prise en charge?* Paris: Inserm, 1998.
- Bonaïti B, Alarcon F, Bonadona V, et al. Nouveau système de score pour le diagnostic des prédispositions aux cancers du sein et de l'ovaire associées à *BRCA1/2*. *Bull Cancer* 2011;98:779–95.
- Evans DG, Lalloo F, Cramer A, Jones EA, Knox F, Amir E, Howell A. Addition of pathology and biomarker information significantly improves the performance of the Manchester scoring system for *BRCA1* and *BRCA2* testing. *J Med Genet* 2009;46:811–17.
- Lakhani SR, Van de Vijver MJ, Jacquemier J, Anderson TJ, Osin PP, McGuffog L, Easton DF. The pathology of familial breast cancer: predictive value of immunohistochemical markers estrogen receptor, progesterone receptor, and p53 in patients with mutations in *BRCA1* and *BRCA2*. *J Clin Oncol* 2002;20:2310–18.
- Olschwang S. Germline mutation and genome instability. *Eur J Cancer Prev* 1999;8 (Suppl 1):S33–7.
- Risch HA, McLaughlin JR, Cole DE, Rosen B, Bradley L, Fan I, Tang J, Li S, Zhang S, Shaw PA, Narod SA. Population *BRCA1* and *BRCA2* mutation frequencies and cancer penetrances: a kin-cohort study in Ontario, Canada. *J Natl Cancer Inst* 2006;98:1694–706.
- van Asperen CJ, Brohet RM, Meijers-Heijboer EJ, Hoogerbrugge N, Verhoef S, Vasen HF, Ausems MG, Menko FH, Gomez Garcia EB, Klijn JG, Hogervorst FB, van Hooftelingen JC, van't Veer LJ, Rookus MA, van Leeuwen FE. Cancer risks in *BRCA2* families: estimates for sites other than breast and ovary. *J Med Genet* 2005;42:711–19.
- Tai YC, Domchek S, Parmigiani G, Chen S. Breast cancer risk among male *BRCA1* and *BRCA2* mutation carriers. *J Natl Cancer Inst* 2007;99:1811–14.
- De Angelis R, Capocaccia R, Hakulinen T, Soderman B, Verdecchia A. Mixture models for cancer survival analysis: application to population-based data with covariates. *Stat Med* 1999;18:441–54.
- Bonaïti B, Bonadona V, Perdry H, Andrieu N, Bonaïti-Pellie C. Estimating penetrance from multiple case families with predisposing mutations: extension of the 'genotype-restricted likelihood' (GRL) method. *Eur J Hum Genet* 2011;19:173–9.



## Cancer genetics

- 23 Belot A, Grosclaude P, Bossard N, Jouglu E, Benhamou E, Delafosse P, Guizard AV, Molinie F, Danzon A, Bara S, Bouvier AM, Tretarre B, Binder-Foucard F, Colonna M, Daubisse L, Hedelin G, Launoy G, Le Stang N, Maynadie M, Monnerieu A, Troussard X, Faivre J, Collignon A, Janoray I, Arveux P, Buemi A, Raverdy N, Schvartz C, Bovet M, Cherie-Challine L, Esteve J, Remontet L, Velten M. Cancer incidence and mortality in France over the period 1980–2005. *Rev Epidemiol Sante Publique* 2008;56:159–75.
- 24 Atchley DP, Albarracin CT, Lopez A, Valero V, Amos CI, Gonzalez-Angulo AM, Hortobagyi GN, Arun BK. Clinical and pathologic characteristics of patients with BRCA-positive and BRCA-negative breast cancer. *J Clin Oncol* 2008;26:4282–8.
- 25 Haffty BG, Yang Q, Reiss M, Kearney T, Higgins SA, Weidhaas J, Harris L, Hait W, Toppmeyer D. Locoregional relapse and distant metastasis in conservatively managed triple negative early-stage breast cancer. *J Clin Oncol* 2006;24:5652–7.
- 26 Antoniou AC, Pharoah PD, McMullan G, Day NE, Stratton MR, Peto J, Ponder BJ, Easton DF. A comprehensive model for familial breast cancer incorporating BRCA1, BRCA2 and other genes. *Br J Cancer* 2002;86:76–83.
- 27 Brozek I, Ochman K, Debniak J, Morzuch L, Ratajska M, Stepnowska M, Stukan M, Emerich J, Limon J. High frequency of BRCA1/2 germline mutations in consecutive ovarian cancer patients in Poland. *Gynecol Oncol* 2008;108:433–7.
- 28 Jacobi CE, van Ierland Y, van Asperen CJ, Hallensleben E, Devilee P, Jan Fleuren G, Kenter GG. Prediction of BRCA1/2 mutation status in patients with ovarian cancer from a hospital-based cohort. *Genet Med* 2007;9:173–9.
- 29 Moslehi R, Chu W, Karlan B, Fishman D, Risch H, Fields A, Smotkin D, Ben-David Y, Rosenblatt J, Russo D, Schwartz P, Tung N, Warner E, Rosen B, Friedman J, Brunet JS, Narod SA. BRCA1 and BRCA2 mutation analysis of 208 Ashkenazi Jewish women with ovarian cancer. *Am J Hum Genet* 2000;66:1259–72.
- 30 Risch HA, McLaughlin JR, Cole DE, Rosen B, Bradley L, Kwan E, Jack E, Vesprini DJ, Kuperstein G, Abrahamson JL, Fan I, Wong B, Narod SA. Prevalence and penetrance of germline BRCA1 and BRCA2 mutations in a population series of 649 women with ovarian cancer. *Am J Hum Genet* 2001;68:700–10.
- 31 Alarcon F, Bourgain C, Gauthier-Villars M, Plante-Bordeneuve V, Stoppa-Lyonnet D, Bonaiti-Pellie C. PEL: an unbiased method for estimating age-dependent genetic disease risk from pedigree data unselected for family history. *Genet Epidemiol* 2009;33:379–85.
- 32 Pennec S. Four-Generation families in France. *Popul Eng Sel* 1997;9:75–101.
- 33 Toulemon L. How many children and how many siblings in France in the last century? *Popul Societies* 2001;373:1–4.
- 34 Elston RC, Stewart J. A general model for the genetic analysis of pedigree data. *Hum Hered* 1971;21:523–42.
- 35 Bonaiti-Pellie C, Andrieu N, Arveux P, Bonadona V, Buecher B, Delpech M, Jolly D, Julian-Reynier C, Luporsi E, Nogues C, Nowak F, Olschwang S, Orsi F, Pujol P, Saurin JC, Sinilnikova O, Stoppa-Lyonnet D, Thepot F. Oncogénétique: estimation des besoins de la populations en France pour les 10 ans à venir. *Bull Cancer* 2009;96:875–900.
- 36 Antoniou AC, Hardy R, Walker L, Evans DG, Shenton A, Eeles R, Shanley S, Pichert G, Izatt L, Rose S, Douglas F, Eccles D, Morrison PJ, Scott J, Zimmern RL, Easton DF, Pharoah PD. Predicting the likelihood of carrying a BRCA1 or BRCA2 mutation: validation of BOADICEA, BRCAPRO, IBIS, Myriad and the Manchester scoring system using data from UK genetics clinics. *J Med Genet* 2008;45:425–31.
- 37 Cui J, Hopper JL. Why are the majority of hereditary cases of early-onset breast cancer sporadic? A simulation study. *Cancer Epidemiol Biomarkers Prev* 2000;9:805–12.



## A new scoring system in cancer genetics: application to criteria for *BRCA1* and *BRCA2* mutation screening

Bernard Bonaïti, Flora Alarcon, Nadine Andrieu, et al.

*J Med Genet* published online December 16, 2013  
doi: 10.1136/jmedgenet-2013-101674

---

Updated information and services can be found at:  
<http://jmg.bmj.com/content/early/2013/12/16/jmedgenet-2013-101674.full.html>

---

*These include:*

**Data Supplement**

"Supplementary Data"

<http://jmg.bmj.com/content/suppl/2013/12/17/jmedgenet-2013-101674.DC1.html>

**References**

This article cites 36 articles, 12 of which can be accessed free at:

<http://jmg.bmj.com/content/early/2013/12/16/jmedgenet-2013-101674.full.html#ref-list-1>

**P<P**

Published online December 16, 2013 in advance of the print journal.

**Email alerting service**

Receive free email alerts when new articles cite this article. Sign up in the box at the top right corner of the online article.

---

**Topic Collections**

Articles on similar topics can be found in the following collections

[Breast cancer](#) (217 articles)  
[Screening \(oncology\)](#) (210 articles)  
[Epidemiology](#) (565 articles)  
[Molecular genetics](#) (1140 articles)

---

Advance online articles have been peer reviewed, accepted for publication, edited and typeset, but have not yet appeared in the paper journal. Advance online articles are citable and establish publication priority; they are indexed by PubMed from initial publication. Citations to Advance online articles must include the digital object identifier (DOIs) and date of initial publication.

---

To request permissions go to:

<http://group.bmj.com/group/rights-licensing/permissions>

To order reprints go to:

<http://journals.bmj.com/cgi/reprintform>

To subscribe to BMJ go to:

<http://group.bmj.com/subscribe/>

## Notes

---

Advance online articles have been peer reviewed, accepted for publication, edited and typeset, but have not yet appeared in the paper journal. Advance online articles are citable and establish publication priority; they are indexed by PubMed from initial publication. Citations to Advance online articles must include the digital object identifier (DOIs) and date of initial publication.

---

To request permissions go to:

<http://group.bmj.com/group/rights-licensing/permissions>

To order reprints go to:

<http://journals.bmj.com/cgi/reprintform>

To subscribe to BMJ go to:

<http://group.bmj.com/subscribe/>