

ORIGINAL ARTICLE

# KOHBRA *BRCA* risk calculator (KOHCal): a model for predicting *BRCA1* and *BRCA2* mutations in Korean breast cancer patients

Eunyoung Kang<sup>1</sup>, Sue K Park<sup>2,3,4</sup>, Jong Won Lee<sup>5</sup>, Zisun Kim<sup>6</sup>, Woo-Chul Noh<sup>7</sup>, Yongsik Jung<sup>8</sup>, Jung-Hyun Yang<sup>9</sup>, Sung Hoo Jung<sup>10</sup> and Sung-Won Kim<sup>11</sup> on behalf of Korean Hereditary Breast Cancer Study Group

The widely used Western *BRCA* mutation prediction models underestimated the risk of having a *BRCA* mutation in Korean breast cancer patients. This study aimed to identify predictive factors for *BRCA1/2* mutations and to develop a Korean *BRCA* risk calculator. The model was constructed by logistic regression model, and it was based on the Korean Hereditary Breast Cancer study, in which 1669 female patients were enrolled between May 2007 and December 2010. A separate data set of 402 patients, who were enrolled from Jan 2011 to August 2012, was used to test the performance of our model. In total, 264 (15.8%) and 67 (16.7%) *BRCA* mutation carriers were identified in the model and validation set, respectively. Multivariate analysis showed that age at breast cancer diagnosis, bilateral breast cancer, triple-negative breast cancer (TNBC) and the number of relatives with breast or ovarian cancer within third-degree relatives were independent predictors of the *BRCA* mutation among familial breast cancer patients. An age <35 years at diagnosis, bilateral breast cancer, both breast and ovarian cancer and TNBC remained significant predictors in non-familial breast cancer cases. Our model was developed based on logistic regression models. The validation results showed no differences between the observed and expected carrier probabilities. This model will be a useful tool for providing genetic risk assessments in Korean populations.

*Journal of Human Genetics* (2016) 61, 365–371; doi:10.1038/jhg.2015.164; published online 14 January 2016

## INTRODUCTION

Many genes that increase the risk for breast cancer have been identified, including *BRCA1*,<sup>1</sup> *BRCA2*,<sup>2</sup> *CHEK2*,<sup>3</sup> *ATM*,<sup>4</sup> *TP53*<sup>5</sup> and *PTEN*.<sup>6</sup> The genes most commonly associated with hereditary breast cancer are *BRCA1* and *BRCA2*. Therefore, many risk assessment models have been developed to predict the probability of carrying a *BRCA1* or *BRCA2* mutation. These models include empirical models (Myriad prevalence tables and Manchester scoring system),<sup>7,8</sup> genetic models (BRCAPRO and BOADICEA)<sup>9,10</sup> and logistic regression model (LAMBDA).<sup>11</sup> However, the performance of these models varies in different ethnic groups. For instance, the widely used BRCAPRO and Myriad II models underestimate the proportion of *BRCA1/2* mutation carriers in Asians.<sup>12</sup> We have shown that these models significantly underestimate the number of mutation carriers in Korean breast cancer patients; indeed, their performance is worse in Korean populations than in other ethnicities.<sup>13,14</sup> Therefore, a new risk assessment model based on data from Korean patients is needed.

In 2007, a large prospective nationwide database based on the Korean Hereditary Breast Cancer (KOHBRA) study was established to determine the prevalence of *BRCA1/2* mutations among breast cancer patients at risk for hereditary breast and ovarian cancer (HBOC).<sup>15</sup> According to the KOHBRA study, the overall prevalence of *BRCA* mutations among patients with breast cancer and with a family history of breast or ovarian cancer was 21.7% (*BRCA1* 9.3% and *BRCA2* 12.4%), and deleterious mutations were more frequently observed in patients with stronger family histories (higher number of relatives with breast or ovarian cancer).<sup>16</sup> Among the patients without a family history of breast or ovarian cancer, the mutation prevalence was 10.0% for patients with early onset (<35 years), 17.7% for patients with bilateral breast cancer, 5.9% for male breast cancer patients and 50.0% for patients with breast and ovarian cancer.<sup>17</sup> The mutation prevalence in most high-risk subgroups of Korean breast cancer patients was much higher than the 10% carrier probability threshold that is commonly used for selecting families for *BRCA*

<sup>1</sup>Department of Surgery, Seoul National University Bundang Hospital, Seoul National University College of Medicine, Seongnam, Korea; <sup>2</sup>Department of Preventive Medicine, Seoul National University College of Medicine, Seoul, Korea; <sup>3</sup>Department of Biomedical Science, Seoul National University Graduate School, Seoul, Korea; <sup>4</sup>Cancer Research Institute, Seoul National University, Seoul, Korea; <sup>5</sup>Department of Surgery, Asan Medical Center, College of Medicine, University of Ulsan, Seoul, Korea; <sup>6</sup>Department of Surgery, Soonchunhyang University Hospital, Bucheon, Korea; <sup>7</sup>Department of Surgery, Korea Institute of Radiological and Medical Science, Korea Cancer Center Hospital, Seoul, Korea; <sup>8</sup>Department of Surgery, Ajou University School of Medicine, Suwon, Korea; <sup>9</sup>Department of Surgery, Konkuk University School of Medicine, Seoul, Korea; <sup>10</sup>Department of Surgery, Chonbuk National University Hospital, Jeonju, Korea and <sup>11</sup>Department of Surgery, Daerim St Mary's Hospital, Seoul, Korea  
Correspondence: Dr S-W Kim, Department of Surgery, Daerim St Mary's Hospital, 657, Siheung-daero, Yongdungpo-gu, Seoul 150-822, Republic of Korea.  
E-mail: brcakorea@gmail.com

Received 26 August 2015; revised 26 November 2015; accepted 10 December 2015; published online 14 January 2016

genetic testing. However, the prediction strength of each risk factor for BRCA mutations is less definitive in Korea. The KOHBRA study is ongoing to establish a BRCA carrier cohort and provide accurate data on the prevalence and penetrance of the BRCA1/2 mutations. The large volumes of data from the KOHBRA study have made it possible to develop a BRCA risk prediction model for the Korean population. Our goal was to identify the pathogenic factors associated with BRCA mutations and to develop a Korean BRCA-based mutation risk calculator using a logistic regression model. The performance of the new model was evaluated against the KOHBRA data using a split sample validation method.

## MATERIALS AND METHODS

Between May 2007 and August 2012, 2071 female breast cancer patients (probands) at risk for HBOC underwent genetic counseling and testing for the BRCA1 and BRCA2 gene mutations through the KOHBRA study. The eligibility criteria for the KOHBRA study were as follows: (1) patients with breast cancer and a family history of breast or ovarian cancer (familial breast cancer patients); (2) patients with breast cancer without a family history of breast or ovarian cancer (non-familial breast cancer patients), including subjects aged  $\leq 40$  years at diagnosis, those with bilateral breast cancer, male patients or those diagnosed with another primary malignancy related to BRCA mutations; and (3) family members of BRCA1/2 mutation carriers.<sup>15</sup>

The prospectively collected KOHBRA database contains genetic, clinical and demographic data, including personal cancer history (age at diagnosis, bilateral breast cancer, coexistence of ovarian cancer and other BRCA1/2-related cancers); number of relatives with breast or ovarian cancer among first-, second- and third-degree relatives; and histological characteristics (histological type (invasive breast cancer vs ductal carcinoma *in situ*), estrogen receptor (ER), progesterone receptor (PR) and human epidermal growth factor receptor 2 (HER2) status). Personal and family histories of cancers were primarily obtained based on the epidemiological questionnaire and pedigree including at least three generations was obtained through counseling by the KOHBRA research nurses. The histological type, ER, PR and HER2 status determinations were based on pathology reports from each institution. In general, an immunohistochemical score of  $>3$  (at least 10% positive cells) was used to define the ER and PR positivity and an immunohistochemical score of  $3+$  was used to define the HER2 positivity. Triple-negative breast cancer (TNBC) was immunohistochemically defined as ER negative, PR negative and lacking overexpression of HER2. More information for data collection of the KOHBRA study can be found in previous protocols and interim report.<sup>15</sup>

For the analysis, 2071 female probands were included in the study subjects, and they all had a personal history of breast cancer. We defined the proband as the first family member recruited for BRCA1/2 genetic counseling and testing, and familial breast cancer patients as persons who has a family history of breast or ovarian cancer within third-degree relatives. In total, 1669 female breast cancer patients enrolled between May 2007 and December 2010 were used for the model construction (850 familial breast cancer patients and 819 non-familial breast cancer patients); 402 female breast cancer patients enrolled between January 2011 and August 2012 were used for validation. To build the model, we compared the demographics and clinical characteristics of the carriers and non-carriers using the chi-square or Fisher's exact test as categorical variables. Factors with  $P$ -values  $\leq 0.05$  were included in a multivariate logistic regression model constructed using the stepwise selection procedure to identify predictors of the BRCA mutation status. The standard logistic regression formula is as follows:

$$\text{Logit}(P) = \beta_0 + \beta_1\chi_1 + \beta_2\chi_2 + \beta_3\chi_3 + \dots + \beta_n\chi_n$$

$$P(\text{BRCAmutation} = 1) = \frac{1}{1 + e^{-(\beta_0 + \beta_1\chi_1 + \beta_2\chi_2 + \beta_3\chi_3 + \dots + \beta_n\chi_n)}}$$

' $P$ ' is the estimated probability of carrying the BRCA gene mutation; ' $\beta$ ' is the influence coefficient; ' $\beta_0$ ' is a constant; and ' $\chi$ ' is the influence factor. The model was developed for familial and non-familial breast cancer patient groups. To assess the performance of the model, we compared the observed and predicted numbers of BRCA1/2 mutation carriers using the chi-square or Fisher's exact test. We evaluated the ability of our model to discriminate between mutation

carriers and non-carriers by means of receiver operating characteristic (ROC) curves and assessing the area under the ROC curve. All data were analyzed using SPSS (version 21, IBM, NY, USA); a  $P$ -value  $\leq 0.05$  was considered to be statistically significant. This study was approved by the Seoul National University Bundang Hospital Institutional Review Board (IRB No. B-1007-1/05-009).

## RESULTS

### Development of the Korean BRCA risk calculator

In the model set, the mean ages of breast cancer diagnosis in the patients with familial and non-familial breast cancer were 43.5 and 36.9 years, respectively. Of the 1669 patients in the model set, 850 were familial breast cancer patients with a family history of breast or ovarian cancer; 819 were high-risk patients for HBOC without a family history. The overall prevalence of deleterious mutations was 15.8% (264/1669), with 106 (40.2%) in BRCA1 and 156 (59.1%) in BRCA2; 2 (0.8%) subjects carried deleterious mutations in both genes. In total, 199 (23.4%) and 65 (8.0%) mutation carriers were identified in 850 familial and 819 non-familial breast cancer patients, respectively.

Among the familial breast cancer patients, the univariate analysis revealed associations between the BRCA gene mutations and young age at first breast cancer diagnosis ( $P < 0.001$ ), bilateral breast cancer ( $P < 0.001$ ), invasiveness of histological type ( $P = 0.047$ ), TNBC ( $P < 0.001$ ) and a high number of relatives with breast cancer ( $P < 0.001$ ) and ovarian cancer ( $P < 0.001$ ). A personal history of ovarian cancer ( $P = 0.334$ ) and other BRCA-related cancers ( $P = 0.532$ ) were not significantly associated with the BRCA gene mutation (Table 1). The clinical and pathological factors significantly associated with BRCA1/2 mutations were used in the multivariate logistic regression model. With the exception of the histological type, five factors remained as significant predictors of a BRCA gene mutation (Table 2). The probability formula for predicting BRCA mutations in familial breast cancer patients using the logistic regression model is as follows:

$$P(\text{BRCAmutation} = 1) = \frac{1}{1 + e^{-(2.448 + 0.627\chi_1 + 1.016\chi_2 + 1.638\chi_3 + 1.252\chi_4 + 0.995\chi_5 + 0.744\chi_6 + 2.560\chi_7)}}$$

For familial breast cancer patients, the probability of carrying a BRCA mutation decreased with increasing age at diagnosis from  $< 36$  years to 36–45 years to  $\geq 46$  years. We selected this age grouping because it had a better  $C$ -statistic than the other age groupings ( $< 35$  and  $\geq 35$  years or  $< 50$  and  $\geq 50$  years).

Among the non-familial breast cancer patients, an age at breast cancer diagnosis  $< 35$  ( $P = 0.021$ ), bilateral breast cancer ( $P < 0.001$ ), presence of both breast and ovarian cancer ( $P = 0.008$ ) and TNBC ( $P = 0.003$ ) were related to the BRCA gene mutations (Table 3). When these factors were included in the multivariate analysis, all were found to be independent predictors of BRCA gene mutations (Table 4). The mutation probability formula in non-familial breast cancer patients is as follows:

$$P(\text{BRCAmutation} = 1) = \frac{1}{1 + e^{-(3.578 + 1.044\chi_1 + 1.534\chi_2 + 2.964\chi_3 + 0.809\chi_4)}}$$

### Validation of the Korean BRCA risk calculator

The prevalence of BRCA1/2 mutations in the validation set was 16.7% (67/402), which is consistent with the model set ( $P = 0.677$ ). Twenty-

**Table 1 Univariate analysis of *BRCA1/2* mutations and clinical characteristics in familial breast cancer patients in the model set**

|                                    | Total |  | <i>BRCA</i> carriers |      | Non-carriers |      | P      |
|------------------------------------|-------|--|----------------------|------|--------------|------|--------|
|                                    | N     |  | N                    | %    | N            | %    |        |
| <i>Age at diagnosis, years</i>     |       |  |                      |      |              |      |        |
| <35                                | 175   |  | 59                   | 33.7 | 116          | 66.3 | <0.001 |
| 36–45                              | 322   |  | 85                   | 26.4 | 237          | 73.6 |        |
| ≥46                                | 353   |  | 55                   | 15.6 | 298          | 84.4 |        |
| <i>Bilateral BC</i>                |       |  |                      |      |              |      |        |
| Yes                                | 75    |  | 36                   | 48.0 | 39           | 52.0 | <0.001 |
| No                                 | 775   |  | 163                  | 21.0 | 612          | 79.0 |        |
| <i>BC and OC</i>                   |       |  |                      |      |              |      |        |
| Yes                                | 5     |  | 2                    | 60.0 | 3            | 40.0 | 0.334  |
| No                                 | 845   |  | 197                  | 23.3 | 648          | 76.7 |        |
| <i>BC and other cancer</i>         |       |  |                      |      |              |      |        |
| Yes                                | 40    |  | 11                   | 27.5 | 29           | 72.5 | 0.532  |
| No                                 | 810   |  | 188                  | 23.2 | 622          | 76.8 |        |
| <i>Histology</i>                   |       |  |                      |      |              |      |        |
| Invasive cancer                    | 737   |  | 181                  | 24.6 | 556          | 75.4 | 0.047  |
| DCIS                               | 97    |  | 15                   | 15.5 | 82           | 84.5 |        |
| Unknown                            | 16    |  | 3                    | 18.8 | 13           | 81.3 |        |
| <i>Triple-negative BC</i>          |       |  |                      |      |              |      |        |
| Yes                                | 152   |  | 67                   | 44.1 | 85           | 55.9 | <0.001 |
| No                                 | 634   |  | 118                  | 18.6 | 516          | 81.4 |        |
| Unknown                            | 64    |  | 14                   | 21.9 | 50           | 78.1 |        |
| <i>Family history</i>              |       |  |                      |      |              |      |        |
| BC only                            | 749   |  | 161                  | 21.5 | 588          | 78.5 | <0.001 |
| OC only                            | 72    |  | 22                   | 30.6 | 50           | 69.4 |        |
| BC and OC                          | 29    |  | 16                   | 55.2 | 13           | 44.8 |        |
| <i>Number of relatives with BC</i> |       |  |                      |      |              |      |        |
| 0                                  | 72    |  | 22                   | 30.6 | 50           | 69.4 | <0.001 |
| 1                                  | 640   |  | 124                  | 19.4 | 516          | 80.6 |        |
| ≥2                                 | 138   |  | 53                   | 38.4 | 85           | 61.6 |        |
| <i>Number of relatives with OC</i> |       |  |                      |      |              |      |        |
| 0                                  | 749   |  | 161                  | 21.5 | 588          | 78.5 | <0.001 |
| 1                                  | 97    |  | 35                   | 36.1 | 62           | 63.9 |        |
| ≥2                                 | 4     |  | 3                    | 75.0 | 1            | 25.0 |        |

Abbreviations: BC, breast cancer; OC, ovarian cancer; DCIS, ductal carcinoma *in situ*.

six (38.9%) patients carried a *BRCA1* mutation and 41 (61.2%) patients carried a *BRCA2* mutation. The observed proportions of *BRCA* mutations among the familial and non-familial breast cancer patients were 21.3% (39/183) and 12.8% (28/219); these observations were similar to the model predictions (25.4% and 8.5%, respectively). Table 5 shows the observed and the expected proportions of *BRCA1* or *BRCA2* mutations in the validation set according to the predicted carrier probabilities. There were no differences between the observed and expected carrier probabilities within each risk category. The area under the ROC curve was 0.756 (95% confidence interval (95% CI), 0.669–0.842) for the familial breast cancer model and 0.620 (95% CI, 0.498–0.741) for the non-familial breast cancer model (Figure 1). Both

values were significantly different from 0.5. Table 6 shows the performance of our models for familial and non-familial breast cancer at a carrier probability of 10%. Our model for familial breast cancer exhibited a high sensitivity (94.9%) at this cutoff value.

## DISCUSSION

We built our model from the largest database of prospective, multi-institutional *BRCA1/2* mutation data in Korea. The prevalence of *BRCA1/2* mutations was 15.8% of 1669 patients, with 23.4% of the familial and 8.0% of the non-familial breast cancer patients carrying deleterious *BRCA1/2* mutations. Younger age at breast cancer diagnosis, bilateral breast cancer, TNBC and a high number of relatives with breast or ovarian cancer were predictors of a *BRCA* gene mutation in the familial breast cancer. An age at diagnosis <35 years, bilateral breast cancer, presence of both breast and ovarian cancer and TNBC were independent predictors of a *BRCA* gene mutation in the non-familial breast cancer. These factors were applied to a logistic regression model to develop a Korean *BRCA* risk calculator, which accurately predicted the *BRCA* mutation rates for each risk category in the validation set.

Several related factors increase the probability of *BRCA* mutations, including early-onset breast cancer, bilateral breast cancer, history of both breast and ovarian cancer, male breast cancer and familial clustering of breast or ovarian cancer.<sup>7,18,19</sup> Various prediction models have been developed based on these factors to identify patients with a high risk of HBOC. The Myriad and BRCAPRO models are the most commonly used models for the convenient selection of *BRCA1/2* candidates worldwide. The Myriad model initially involved logistic regression analyses, which identified the following predictive factors of *BRCA1/2* mutations: breast cancer before 50 years of age and ovarian cancer in the family.<sup>7</sup> BRCAPRO uses Bayes' theory to calculate the probability of carrying a *BRCA* mutation. The model is based on published *BRCA1/2* mutation frequencies, information about the first- and second-degree relatives, individual cancer status, cancer penetrance, bilateral breast cancer and male breast cancer.<sup>9</sup> The approach in our model was oriented to the Myriad I model, which used logistic regression in predicting *BRCA1*. Our model incorporated personal and familial breast or ovarian cancer histories, as well as the histological characteristics of each breast tumor and TNBC. Previous studies have found that the ER, PR and HER2 negativities of breast cancer are significantly linked to *BRCA1* mutations.<sup>20–23</sup> The incorporation of breast tumor markers, such as ER, PR and HER2, has been shown to improve the performance of BRCAPRO and BOADICEA.<sup>24–26</sup> We confirmed that most *BRCA1* breast cancers are TNBC (data not shown), although we did not derive a formula to calculate the separate risk for *BRCA1* and *BRCA2*, because our goal was to develop a simple and easy model to give a composite probability for selecting candidates for *BRCA1* or *BRCA2* genetic testing in Korean populations.

Familial clustering of breast and ovarian cancer is a strong predictor of a *BRCA* mutation.<sup>27</sup> Family studies have documented the relationship of *BRCA* gene mutation risk and the age of onset of cancer and the number of these cancers among first- and second-degree relatives, and the majority of the clinical guidelines for HBOC recommend genetic risk assessment in cases of two or more individuals with breast and/or ovarian cancer within second-degree relatives. Previously, the result of the KOHBRA study for the prevalence of *BRCA* mutations among familial breast cancer showed that the prevalence satisfied the 10% cutoff among breast cancer patients with a weak family history (only one relative with breast cancer (17.5%) and third-degree relatives as the closest relatives with breast cancer (14.0%)).<sup>16</sup>

**Table 2 Predictors for *BRCA1/2* mutations in familial breast cancer patients in the model set**

| Characteristics                     | Variable | B      | s.e.  | P      | OR     | 95%CI |         |
|-------------------------------------|----------|--------|-------|--------|--------|-------|---------|
|                                     |          |        |       |        |        | Upper | Lower   |
| <i>Age at diagnosis, years</i>      |          |        |       |        |        |       |         |
| ≥46                                 |          |        |       |        |        |       |         |
| 36–45                               | X1       | 0.627  | 0.215 | 0.004  | 1.873  | 1.229 | 2.854   |
| ≤35                                 | X2       | 1.016  | 0.245 | <0.001 | 2.763  | 1.710 | 4.464   |
| Bilateral BC (yes vs no)            | X3       | 1.638  | 0.282 | <0.001 | 5.147  | 2.959 | 8.953   |
| TNBC (yes vs no)                    | X4       | 1.252  | 0.208 | <0.001 | 3.498  | 2.327 | 5.259   |
| No. of relatives with BC (≥2 vs <2) | X5       | 0.995  | 0.226 | <0.001 | 2.705  | 1.737 | 4.212   |
| <i>No. of relatives with OC</i>     |          |        |       |        |        |       |         |
| 0                                   |          |        |       |        |        |       |         |
| 1                                   | X6       | 0.744  | 0.262 | 0.005  | 2.104  | 1.259 | 3.516   |
| ≥2                                  | X7       | 2.560  | 1.249 | 0.040  | 12.934 | 1.126 | 148.538 |
| Constant                            |          | -2.448 | 0.193 | <0.001 | 0.086  |       |         |

Abbreviations: BC, breast cancer; CI, confidence interval; OC, ovarian cancer; OR, odds ratio; TNBC, triple-negative breast cancer.

**Table 3 Univariate analysis of *BRCA1/2* mutations and clinical characteristics in non-familial breast cancer patients in the model set**

|                                | <i>BRCA</i> |    |      |              |      |        |
|--------------------------------|-------------|----|------|--------------|------|--------|
|                                | Total       |    |      | Non-carriers |      | P      |
|                                | N           | N  | %    | N            | %    |        |
| <i>Age at diagnosis, years</i> |             |    |      |              |      |        |
| <35                            | 343         | 36 | 10.5 | 307          | 89.5 | 0.021  |
| ≥35                            | 476         | 29 | 6.1  | 447          | 93.9 |        |
| <i>Bilateral BC</i>            |             |    |      |              |      |        |
| Yes                            | 148         | 23 | 15.5 | 125          | 84.5 | <0.001 |
| No                             | 671         | 42 | 6.3  | 629          | 93.7 |        |
| <i>BC and OC</i>               |             |    |      |              |      |        |
| Yes                            | 6           | 3  | 50.0 | 3            | 50.0 | 0.008  |
| No                             | 813         | 62 | 7.6  | 751          | 92.4 |        |
| <i>BC and other cancer</i>     |             |    |      |              |      |        |
| Yes                            | 69          | 4  | 5.8  | 65           | 94.2 | 0.664  |
| No                             | 750         | 61 | 8.1  | 689          | 91.9 |        |
| <i>Histology</i>               |             |    |      |              |      |        |
| Invasive cancer                | 747         | 62 | 8.3  | 685          | 91.7 | 0.134  |
| DCIS                           | 65          | 2  | 3.1  | 63           | 96.9 |        |
| Unknown                        | 7           | 1  | 14.3 | 6            | 85.7 |        |
| <i>Triple-negative BC</i>      |             |    |      |              |      |        |
| Yes                            | 150         | 21 | 14.0 | 129          | 86.0 | 0.003  |
| No                             | 621         | 41 | 6.6  | 580          | 93.4 |        |
| Unknown                        | 48          | 3  | 6.3  | 45           | 93.8 |        |

Abbreviations: BC, breast cancer; OC, ovarian cancer; DCIS, ductal carcinoma *in situ*.

Therefore, our model for familial breast cancer was developed based on extended family data, including information on third-degree relatives. The National Comprehensive Cancer Network Clinical Practice Guidelines in Oncology for Genetic/Familial High-Risk

Assessment: Breast and Ovarian also suggest that close relatives for risk assessment include first-, second- and third-degree relatives. For breast and ovarian analysis of disease incidence and carrier estimation algorithm (BOADICEA), which is a model of genetic susceptibility for breast and ovarian cancer developed by Antoniou *et al.*,<sup>10</sup> the inclusion of extended family data slightly improved the accuracy of the risk prediction of a *BRCA* mutation compared with the risk prediction obtained using limited information on second-degree relatives.<sup>28</sup>

Our findings also showed that a higher number of relatives with breast cancer is related to a higher risk of carrying a *BRCA* mutation, and the age of breast cancer of the proband is an important predictor of a *BRCA1/2* mutation when the patient has a weak family history of breast cancer (presence of one relative with breast cancer in the family). In our study, when patients > 45 years with breast cancer had only one relative with breast cancer, the predicted prevalence of a *BRCA1/2* mutation did not satisfy the 10% probability requirement. In the other age groups (<36 and 36–45 years), patients with one relative with breast cancer had a *BRCA1/2* mutation probability of 19.3% and 13.9%, respectively.

Most prediction models include a family history of breast or ovarian cancer, making it difficult to score the risk of a *BRCA* mutation for a single case of breast cancer in a family. Therefore, we developed a separate risk calculation for isolated breast cancer patients without a family history of breast or ovarian cancer. Our study suggests *BRCA1/2* genetic testing is not justified in patients with early-onset breast cancer between the ages of 35 and 40 years who have no other risk factors (the predicted mutation rate is 2.7%). The predicted prevalence in patients with breast cancer aged <35 years without other risk factors (7.4%) and patients with TNBC aged ≥35 years (5.9%) did not satisfy the 10% probability requirement. However, patients with TNBC aged <35 years had a *BRCA1/2* mutation probability of 15.1%, which satisfied the 10% probability requirement. Previous studies suggest the test is valid in isolated breast cancer if it is triple-negative and the patient is <35 years of age at diagnosis.<sup>29,30</sup> Therefore, *BRCA1/2* genetic testing should be recommended for young women with an isolated case of breast cancer based on the age-of-onset and the pathological features (ER, PR and HER2 negativities) of the tumor.

**Table 4 Predictors for *BRCA1/2* mutations in non-familial breast cancer patients of model set**

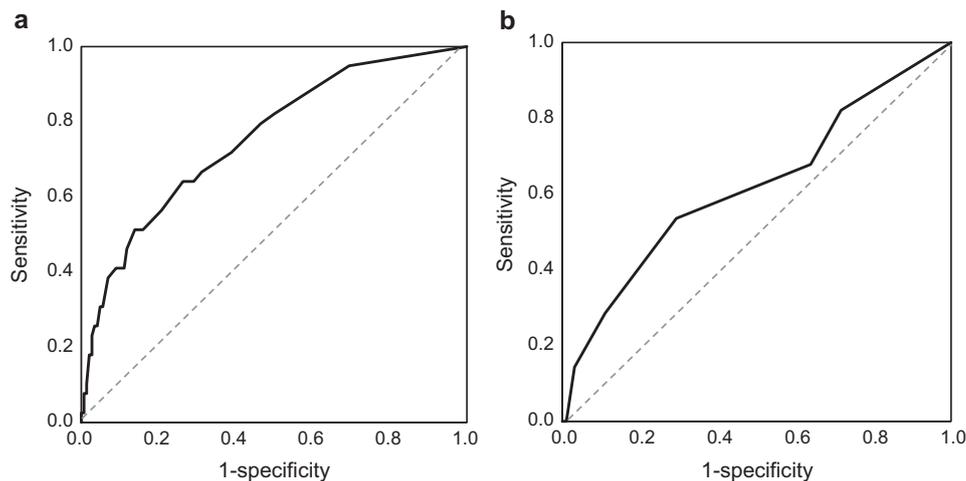
| Characteristics                    | Variable | B      | S.E   | P      | OR     | 95%CI |         |
|------------------------------------|----------|--------|-------|--------|--------|-------|---------|
|                                    |          |        |       |        |        | upper | lower   |
| Age at diagnosis (years)           |          |        |       |        |        |       |         |
| <35 vs ≥35                         | X'1      | 1.044  | 0.314 | 0.001  | 2.842  | 1.534 | 5.264   |
| Bilateral BC (yes vs no)           | X'2      | 1.534  | 0.334 | <0.001 | 4.639  | 2.410 | 8.927   |
| Personal history of OC (yes vs no) | X'3      | 2.964  | 0.914 | 0.001  | 19.368 | 3.228 | 116.221 |
| TNBC (yes vs no)                   | X'4      | 0.809  | 0.308 | 0.009  | 2.246  | 1.228 | 4.107   |
| Constant                           |          | -3.578 | 0.288 | <0.001 | 0.028  |       |         |

Abbreviations: BC, breast cancer; OC, ovarian cancer; OR, odds ratio; TNBC, triple-negative breast cancer.

**Table 5 Observed and predicted mutation frequencies in the KOHCal validation set, stratified by predicted carrier probability**

|  | No. of probands | Expected mutations (%) | Observed mutations (%) | E/O ratio | 95% CI      | P-value |
|--|-----------------|------------------------|------------------------|-----------|-------------|---------|
| <i>Familial breast cancer patients</i>     |                 |                        |                        |           |             |         |
| <10%                                       | 46              | 8.0                    | 4.3                    | 1.850     | 0.348–9.837 | 0.677   |
| 10–20%                                     | 66              | 16.2                   | 16.7                   | 0.973     | 0.451–2.099 | 1.000   |
| ≥20%                                       | 71              | 45.3                   | 36.6                   | 1.239     | 0.831–1.845 | 0.306   |
| Total                                      | 183             | 25.4                   | 21.3                   | 1.192     | 0.821–1.731 | 0.339   |
| <i>Non-familial breast cancer patients</i> |                 |                        |                        |           |             |         |
| <10%                                       | 148             | 5.3                    | 8.8                    | 0.608     | 0.259–1.428 | 0.258   |
| 10–20%                                     | 61              | 12.6                   | 18.0                   | 0.700     | 0.299–1.637 | 0.454   |
| ≥20%                                       | 10              | 31.2                   | 40.0                   | 0.775     | 0.234–2.564 | 1.000   |
| Total                                      | 219             | 8.5                    | 12.8                   | 0.668     | 0.384–1.163 | 0.165   |

Abbreviations: CI, confidence interval; KOHCal, KOHBRA *BRCA* risk calculator.



**Figure 1** Receiver operating characteristic (ROC) curves for familial (a) and non-familial (b) breast cancer in the validation set (the area under ROC curve (AUC) for familial breast cancer patients was 0.756 ( $P < 0.001$ ) and the AUC for non-familial breast cancer patients was 0.620 ( $P = 0.041$ )).

Personal history of ovarian cancer is one of the major indicators for *BRCA* genetic screening. Approximately 10% of unselected incident cases of ovarian cancer have mutations in the *BRCA1* and *BRCA2* genes.<sup>31</sup> In our study, a personal history of ovarian cancer was a predictor of a *BRCA1/2* mutation in only the non-familial breast cancer patients. The univariate analysis revealed that a personal history of ovarian cancer in familial breast cancer patients tended to increase the mutation risk but was not statistically significant (odds ratio 2.193, 95% CI 0.364–13.217). This lack of significance was likely observed

because there were only five subjects with a personal history of ovarian cancer, which is too small to observe statistical significance in patients who had a strong predictor, such as a family history of breast cancer and/or ovarian cancer ( $n = 850$ ).

Prediction models generally account for the prevalence of *BRCA1/2* mutation in specific ethnic populations; therefore, their validity must be verified before they can be applied to different populations. Previously, we evaluated the performance of the Myriad II and BRCAPRO models for Korean breast cancer patients who received

**Table 6 KOHCal performance with carrier probability of 10% in familial and non-familial breast cancer**

|                           | Familial | Non-familial |
|---------------------------|----------|--------------|
| Sensitivity               | 94.9%    | 53.6%        |
| Specificity               | 30.6%    | 70.7%        |
| False negative rate       | 5.1%     | 46.4%        |
| False positive rate       | 69.4%    | 29.3%        |
| Positive predictive value | 27.0%    | 21.2%        |
| Negative predictive value | 95.7%    | 91.2%        |
| Positive likelihood ratio | 1.4      | 1.8          |

Abbreviation: KOHCal, KOHBRA BRCA risk calculator.

BRCA genetic testing.<sup>13</sup> The overall prevalence of a BRCA1/2 mutation was significantly higher than that predicted by these models. We also found the models performed poorly for patients with only one relative with breast cancer (probands' age at breast cancer diagnosis > 50 years) and patients with non-familial early-onset breast cancer or bilateral breast cancer. At a 10% threshold, the sensitivities of BRCAPRO and Myriad II among Korean women were only 47.8% and 50.0%, respectively, which are poorer than their sensitivities in other races (70.7–88.2%).<sup>32,33</sup> This finding is due to the penetrance differences, and multiple factors, such as modifying genes, environmental factors, birth cohorts and race, influence the penetrance of BRCA1 and BRCA2 mutations. Our model for familial breast cancer patients successfully predicted the number of mutation carriers and had a high sensitivity (94.9%). For the non-familial model, although it was able to predict the observed mutation proportion according to the predicted probabilities, it had a relatively low sensitivity (53.6%) at a carrier probability of 10%. In evaluating the performance of a prediction model, a sufficient number of test sets in each level of risk categories are required to increase the predictive power of the model. Therefore, it appears to be due to the small and uneven distribution of the number of subjects according to the level of probabilities, as the number of subjects with at least a 10% predicted probability was half of the number of subjects predicted with a probability of <10%. Because the existence of non-familial breast cancer patients with high mutation probabilities is in fact likely to be low, we assume that the non-familial model would be difficult to have a high sensitivity. Although our non-familial model did not show a high sensitivity (at a 10% cutoff value) due to the limited number of test sets in high-risk groups, the proportion of mutation carriers in the different ranges of probabilities corresponded well to the expected mutation rates in both the familial and non-familial models. Therefore, we consider that our models are appropriate to screen candidates for BRCA1/2 gene testing in high-risk patients for HBOC. However, our model may not be representative of women with breast cancer in the general population or from other ethnicities because our subjects were derived from a Korean population of patients at high risk for HBOC. Therefore, the performance of our model should be evaluated before being applied in other populations.

This Korean BRCA mutation prediction model, known as KOHCal (KOHBRA BRCA risk calculator) is available on the KOHBRA study website (www.kohbra.kr). KOHCal is simple to use and data input takes little time. The model provides quick guidance for clinicians to identify candidates for BRCA1/2 mutation testing in Korea. The model will also add to our understanding of BRCA gene mutation risk in breast cancer by quantifying the probability of a BRCA gene mutation.

## CONFLICT OF INTEREST

The authors declare no conflict of interest.

## ACKNOWLEDGEMENTS

This study was supported by a grant from the National R&D Program for Cancer Control, Ministry for Health, Welfare, and Family Affairs, Republic of Korea (No. 1020350). We thank the Medical Research Collaborating Center at Seoul National University Bundang Hospital for statistical analyses and all investigators of the KOHBRA Study: Beom Seok Kwak, Byeong-Woo Park, Byung Ho Son, Byung-In Moon, Cha Kyong Yom, Chan Heun Park, Chan Seok Yoon, Chang Hyun Lee, Dae Sung Yoon, Dong-Young Noh, Doo Ho Choi, Eundeok Chang, Eun-Kyu Kim, Eunyoung Kang, Hae Kyung Lee, Hai-Lin Park, Hyde Lee, Hyeong-Gon Moon, Hyun-Ah Kim, Il-Kyun Lee, Jeong Eon Lee, Jihyoun Lee, Jong Won Lee, Jong-Han Yu, Joon Jeong, Jung Han Yoon, Jung-Hyun Yang, Keumhee Kwak, Ki-Tae Hwang, Ku Sang Kim, Lee Su Kim, Min Hee Hur, Min Ho Park, Min Hyuk Lee, Myung Chul Chang, Nam Sun Paik, Sang Ah Han, Sang Seol Jung, Sang Uk Woo, Se Jeong Oh, Sehwan Han, Sei Joong Kim, Sei-Hyun Ahn, Seok-Jin Nam, Seung Sang Ko, Sung Hoo Jung, Sung Soo Kang, Sung Yong Kim, Sung-Won Kim, Tae Hyun Kim, Tae Wan Won, Tae Woo Kang, Wonshik Han, Woo-Chul Noh, Yong Lai Park, Yongsik Jung, Young Jin Suh, Young Tae Bae, Young Up Cho, Young-Ik Hong, Sue K. Park, Yoon Joo Jung, Su Yun Choi, Young Bum Yoo, and Soo-Jung Lee.

- Miki, Y., Swensen, J., Shattuck-Eidens, D., Futreal, P. A., Harshman, K., Tavtigian, S. et al. A strong candidate for the breast and ovarian cancer susceptibility gene BRCA1. *Science* **266**, 66–71 (1994).
- Wooster, R., Bignell, G., Lancaster, J., Swift, S., Seal, S., Mangion, J. et al. Identification of the breast cancer susceptibility gene BRCA2. *Nature* **378**, 789–792 (1995).
- Meijers-Heijboer, H., van den Ouweland, A., Klijn, J., Wasielewski, M., de Snoo, A., Oldenburg, R. et al. Low-penetrance susceptibility to breast cancer due to CHEK2(\*) 1100delC in noncarriers of BRCA1 or BRCA2 mutations. *Nat. Genet.* **31**, 55–59 (2002).
- Swift, M., Morrell, D., Massey, R. B. & Chase, C. L. Incidence of cancer in 161 families affected by ataxia-telangiectasia. *N. Engl. J. Med.* **325**, 1831–1836 (1991).
- Malkin, D., Li, F. P., Strong, L. C., Fraumeni, J. F. Jr., Nelson, C. E., Kim, D. H. et al. Germ line p53 mutations in a familial syndrome of breast cancer, sarcomas, and other neoplasms. *Science* **250**, 1233–1238 (1990).
- Eng, C. Genetics of Cowden syndrome: through the looking glass of oncology. *Int. J. Oncol.* **12**, 701–710 (1998).
- Frank, T. S., Deffenbaugh, A. M., Reid, J. E., Hulick, M., Ward, B. E., Lingenfelter, B. et al. Clinical characteristics of individuals with germline mutations in BRCA1 and BRCA2: analysis of 10,000 individuals. *J. Clin. Oncol.* **20**, 1480–1490 (2002).
- Evans, D. G., Eccles, D. M., Rahman, N., Young, K., Bulman, M., Amir, E. et al. A new scoring system for the chances of identifying a BRCA1/2 mutation outperforms existing models including BRCAPRO. *J. Med. Genet.* **41**, 474–480 (2004).
- Parmigiani, G., Berry, D. & Aguilar, O. Determining carrier probabilities for breast cancer-susceptibility genes BRCA1 and BRCA2. *Am. J. Hum. Genet.* **62**, 145–158 (1998).
- Antoniou, A. C., Pharoah, P. P., Smith, P. & Easton, D. F. The BOADICEA model of genetic susceptibility to breast and ovarian cancer. *Br. J. Cancer* **91**, 1580–1590 (2004).
- Apicella, C., Andrews, L., Hodgson, S. V., Fisher, S. A., Lewis, C. M., Solomon, E. et al. Log odds of carrying an Ancestral Mutation in BRCA1 or BRCA2 for a Defined personal and family history in an Ashkenazi Jewish woman (LAMBDA). *Breast Cancer Res.* **5**, R206–R216 (2003).
- Kurian, A. W., Gong, G. D., Chun, N. M., Mills, M. A., Staton, A. D., Kingham, K. E. et al. Performance of BRCA1/2 mutation prediction models in Asian Americans. *J. Clin. Oncol.* **26**, 4752–4758 (2008).
- Kang, E., Park, S. K., Yang, J. J., Park, B., Lee, M. H., Lee, J. W. et al. Accuracy of BRCA1/2 mutation prediction models in Korean breast cancer patients. *Breast Cancer Res. Treat.* **134**, 1189–1197 (2012).
- Kurian, A. W., Gong, G. D., John, E. M., Miron, A., Felberg, A., Phipps, A. I. et al. Performance of prediction models for BRCA mutation carriage in three racial/ethnic groups: findings from the Northern California Breast Cancer Family Registry. *Cancer Epidemiol. Biomarkers Prev.* **18**, 1084–1091 (2009).
- Han, S. A., Park, S. K., Hyun Ahn, S., Hyuk Lee, M., Noh, D. Y., Kim, L. S. et al. The Korean Hereditary Breast Cancer (KOHBRA) Study: Protocols and Interim Report. *Clin. Oncol.* **23**, 434–441 (2011).
- Han, S. A., Kim, S. W., Kang, E., Park, S. K., Ahn, S. H., Lee, M. H. et al. The prevalence of BRCA mutations among familial breast cancer patients in Korea: results of the Korean Hereditary Breast Cancer study. *Fam. Cancer* **12**, 75–81 (2013).

- 17 Son, B. H., Ahn, S. H., Kim, S. W., Kang, E., Park, S. K., Lee, M. H. *et al*. Prevalence of BRCA1 and BRCA2 mutations in non-familial breast cancer patients with high risks in Korea: the Korean Hereditary Breast Cancer (KOHBRA) Study. *Breast Cancer Res. Treat.* **133**, 1143–1152 (2012).
- 18 Srivastava, A., McKinnon, W. & Wood, M. E. Risk of breast and ovarian cancer in women with strong family histories. *Oncology (Williston Park)* **15**, 889–902; discussion 902, 905–887, 911–813 (2001).
- 19 Couch, F. J., DeShano, M. L., Blackwood, M. A., Calzone, K., Stopfer, J., Campeau, L. *et al*. BRCA1 mutations in women attending clinics that evaluate the risk of breast cancer. *N. Engl. J. Med.* **336**, 1409–1415 (1997).
- 20 Lakhani, S. R., Van De Vijver, M. J., Jacquemier, J., Anderson, T. J., Osin, P. P., McGuffog, L. *et al*. The pathology of familial breast cancer: predictive value of immunohistochemical markers estrogen receptor, progesterone receptor, HER-2, and p53 in patients with mutations in BRCA1 and BRCA2. *J. Clin. Oncol.* **20**, 2310–2318 (2002).
- 21 Armes, J. E., Trute, L., White, D., Southey, M. C., Hammet, F., Tesoriero, A. *et al*. Distinct molecular pathogenesis of early-onset breast cancers in BRCA1 and BRCA2 mutation carriers: a population-based study. *Cancer Res.* **59**, 2011–2017 (1999).
- 22 Palacios, J., Honrado, E., Osorio, A., Cazorla, A., Sarrio, D., Barroso, A. *et al*. Phenotypic characterization of BRCA1 and BRCA2 tumors based in a tissue microarray study with 37 immunohistochemical markers. *Breast Cancer Res. Treat.* **90**, 5–14 (2005).
- 23 Mavaddat, N., Barrowdale, D., Andrulis, I. L., Domchek, S. M., Eccles, D., Nevanlinna, H. *et al*. Pathology of breast and ovarian cancers among BRCA1 and BRCA2 mutation carriers: results from the Consortium of Investigators of Modifiers of BRCA1/2 (CIMBA). *Cancer Epidemiol. Biomarkers Prev.* **21**, 134–147 (2012).
- 24 Biswas, S., Tankhiwale, N., Blackford, A., Barrera, A. M., Ready, K., Lu, K. *et al*. Assessing the added value of breast tumor markers in genetic risk prediction model BRCAPRO. *Breast Cancer Res. Treat.* **133**, 347–355 (2012).
- 25 Mavaddat, N., Rebbeck, T. R., Lakhani, S. R., Easton, D. F. & Antoniou, A. C. Incorporating tumour pathology information into breast cancer risk prediction algorithms. *Breast Cancer Res.* **12**, R28 (2010).
- 26 Tai, Y. C., Chen, S., Parmigiani, G. & Klein, A. P. Incorporating tumor immunohistochemical markers in BRCA1 and BRCA2 carrier prediction. *Breast Cancer Res.* **10**, 401 (2008).
- 27 Ford, D., Easton, D. F., Stratton, M., Narod, S., Goldgar, D., Devilee, P. *et al*. Genetic heterogeneity and penetrance analysis of the BRCA1 and BRCA2 genes in breast cancer families. The Breast Cancer Linkage Consortium. *Am. J. Hum. Genet.* **62**, 676–689 (1998).
- 28 Barcnas, C. H., Hosain, G. M., Arun, B., Zong, J., Zhou, X., Chen, J. *et al*. Assessing BRCA carrier probabilities in extended families. *J. Clin. Oncol.* **24**, 354–360 (2006).
- 29 Young, S. R., Pilarski, R. T., Donenberg, T., Shapiro, C., Hammond, L. S., Miller, J. *et al*. The prevalence of BRCA1 mutations among young women with triple-negative breast cancer. *BMC Cancer* **9**, 86 (2009).
- 30 Lidereau, R., Eisinger, F., Champeme, M. H., Nogues, C., Bieche, I., Birnbaum, D. *et al*. Major improvement in the efficacy of BRCA1 mutation screening using morphoclinical features of breast cancer. *Cancer Res.* **60**, 1206–1210 (2000).
- 31 Risch, H. A., McLaughlin, J. R., Cole, D. E., Rosen, B., Bradley, L., Kwan, E. *et al*. Prevalence and penetrance of germline BRCA1 and BRCA2 mutations in a population series of 649 women with ovarian cancer. *Am. J. Hum. Genet.* **68**, 700–710 (2001).
- 32 Huo, D., Senie, R. T., Daly, M., Buys, S. S., Cummings, S., Ogutha, J. *et al*. Prediction of BRCA mutations using the BRCAPRO model in clinic-based African American, Hispanic, and other minority families in the United States. *J. Clin. Oncol.* **27**, 1184–1190 (2009).
- 33 Antoniou, A. C., Hardy, R., Walker, L., Evans, D. G., Shenton, A., Eeles, R. *et al*. 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.* **45**, 425–431 (2008).