Received: 16 October 2020 | Revised: 30 April 2021 | Accepted: 6 May 2021

DOI: 10.1111/cas.14986

ORIGINAL ARTICLE

Cancer Science WILEY

Optimization of prediction methods for risk assessment of pathogenic germline variants in the Japanese population

Noriko Senda[1] | Nobuko Kawaguchi-Sakita[2] | Masahiro Kawashima[1] | Yukiko Inagaki-Kawata[1] | Kenichi Yoshida[3] | Masahiro Takada[1] | Masako Kataoka[4] | Masae Torii[5] | Tomomi Nishimura[1] | Kosuke Kawaguchi[1] | Eiji Suzuki[1] | Yuki Kataoka[6] | Yoshiaki Matsumoto[1] | Hiroshi Yoshibayashi[5] | Kazuhiko Yamagami[7] | Shigeru Tsuyuki[8] | Sachiko Takahara[9] | Akira Yamauchi[9] | Nobuhiko Shinkura[10] | Hironori Kato[11] | Yoshio Moriguchi[12] | Ryuji Okamura[13] | Norimichi Kan[14] | Hirofumi Suwa[15] | Shingo Sakata[16] | Susumu Mashima[17] | Fumiaki Yotsumoto[18] | Tsuyoshi Tachibana[19] | Mitsuru Tanaka[20] | Kaori Togashi[4] | Hironori Haga[21] | Takahiro Yamada[22] | Shinji Kosugi[22] | Takashi Inamoto[23] | Masahiro Sugimoto[24] | Seishi Ogawa[3] | Masakazu Toi[1]

[1]Department of Breast Surgery, Kyoto University, Kyoto, Japan

[2]Department of Clinical Oncology, Kyoto University Hospital, Kyoto, Japan

[3]Department of Pathology and Tumor Biology, Kyoto University, Kyoto, Japan

[4]Department of Diagnostic Imaging and Nuclear Medicine, Kyoto University, Kyoto, Japan

[5]Department of Breast Surgery, Japanese Red Cross Wakayama Medical Center, Wakayama, Japan

[6]Department of Healthcare Epidemiology, School of Public Health, in the Graduate School of Medicine, Kyoto University, Kyoto, Japan

[7]Department of Breast Surgery and Oncology, Shinko Hospital, Kobe, Japan

[8]Department of Breast Surgery, Osaka Red Cross Hospital, Osaka, Japan

[9]Department of Breast Surgery, Kitano Hospital, Osaka, Japan

[10]Department of Surgery, Ijinkai Takeda General Hospital, Kyoto, Japan

[11]Department of Breast Surgery, Kobe City Medical Center General Hospital, Kobe, Japan

[12]Department of Breast Surgery, Kyoto City Hospital, Kyoto, Japan

[13]Department of Breast Surgery, Yamatotakada Municipal Hospital, Yamatotakada, Japan

[14]Kan Norimichi Clinic, Kyoto, Japan

[15]Department of Breast Surgery, Hyogo Prefectural Amagasaki General Medical Center, Amagasaki, Japan

[16]Department of Breast Surgery, Rakuwakai Otowa Hospital, Kyoto, Japan

[17]Department of Surgery, Japan Community Health Care Organization, Yamato Koriyama Hospital, Yamato Koriyama, Japan

[18]Department of Breast Surgery, Shiga General Hospital, Moriyama, Japan

[19]Department of Breast Surgery, Otsu City Hospital, Otsu, Japan

[20]Department of Surgery, Hirakata Kohsai Hospital, Hirakata, Japan

[21]Department of Diagnostic Pathology, Kyoto University Hospital, Kyoto, Japan

[22]Department of Medical Ethics/Medical Genetics, Kyoto University, Kyoto, Japan

[23]Faculty of Health Care, Tenri Health Care University, Tenri, Japan

[24]Health Promotion and Preemptive Medicine, Research and Development Center for Minimally Invasive Therapies, Tokyo Medical University, Tokyo, Japan

Abbreviations: AUC, area under the curve; BMI, body mass index; CI, confidence interval; ER, estrogen receptor; HER2, human epidermal growth factor receptor 2; MGT, multigene testing; MMG, mammography; NCCN, National Comprehensive Cancer Network; PGV, pathogenic or likely pathogenic germline variant; ROC, receiver operating characteristic; TC model, Tyrer-Cuzick model; TNBC, triple-negative breast cancer.

Correspondence
Nobuko Kawaguchi-Sakita, Department of Clinical Oncology, Kyoto University Hospital, 54 Shogoin-Kawahara-cho, Sakyo-ku, Kyoto 606-8507, Japan.
Email: nobuko75@kuhp.kyoto-u.ac.jp

Funding information
AstraZeneca, Grant/Award Number: NCR-17-12919

Abstract

Predicting pathogenic germline variants (PGVs) in breast cancer patients is important for selecting optimal therapeutics and implementing risk reduction strategies. However, PGV risk factors and the performance of prediction methods in the Japanese population remain unclear. We investigated clinicopathological risk factors using the Tyrer-Cuzick (TC) breast cancer risk evaluation tool to predict *BRCA* PGVs in unselected Japanese breast cancer patients (n = 1,995). Eleven breast cancer susceptibility genes were analyzed using target-capture sequencing in a previous study; the PGV prevalence in *BRCA1*, *BRCA2*, and *PALB2* was 0.75%, 3.1%, and 0.45%, respectively. Significant associations were found between the presence of *BRCA* PGVs and early disease onset, number of familial cancer cases (up to third-degree relatives), triple-negative breast cancer patients under the age of 60, and ovarian cancer history (all $P < .0001$). In total, 816 patients (40.9%) satisfied the National Comprehensive Cancer Network (NCCN) guidelines for recommending multigene testing. The sensitivity and specificity of the NCCN criteria for discriminating PGV carriers from noncarriers were 71.3% and 60.7%, respectively. The TC model showed good discrimination for predicting *BRCA* PGVs (area under the curve, 0.75; 95% confidence interval, 0.69-0.81). Furthermore, use of the TC model with an optimized cutoff of TC score ≥0.16% in addition to the NCCN guidelines improved the predictive efficiency for high-risk groups (sensitivity, 77.2%; specificity, 54.8%; about 11 genes). Given the influence of ethnic differences on prediction, we consider that further studies are warranted to elucidate the role of environmental and genetic factors for realizing precise prediction.

KEYWORDS
BRCA, breast cancer, risk factor, pathogenic germline variant, Tyrer-Cuzick model

1 | INTRODUCTION

Breast cancer is the most prevalent malignancy in Japanese women, and its incidence continues to rise.[1] A high lifetime breast cancer risk has been reported in carriers of germline variants in genes associated with hereditary breast cancer (25%-90%).[2,3] In fact, the prevalence of hereditary breast cancer is reported to be 5%-10% in primary breast cancer patients.[2,4] Pathogenic or likely pathogenic germline variants (PGVs) of breast cancer susceptibility genes, including *BRCA1* and *BRCA2*, substantially impact breast cancer onset by increasing the risk of bilaterality, multiplicity, or early onset.[4-8] Recently, an increasing number of studies have been conducted to assess the clinical relevance of PGVs in cancers and to determine whether they are effective targets for novel cancer treatment and management strategies.[9,10]

The American Society of Clinical Oncology guidelines recommend that all patients diagnosed with malignancy undergo multigene testing (MGT).[11] Although MGT is an effective method for identifying PGV carriers, it remains too expensive to perform for all patients.[12,13] In contrast, models capable of predicting who might be carriers of PGVs would be convenient and essentially free to use. Accurately predicting which patients are PGV carriers might be useful for both clinicians and patients when considering whether to perform genetic testing. However, in the Japanese population in particular, precise testing criteria for PGVs have not yet been validated because the associated risk factors remain unclear. Moreover, the prediction models are based primarily on the incidence, factors, and parameter values in Western countries.

Ethnic-specific differences are a critical aspect that must be addressed to allow for personalized approaches and may inform development of more effective breast cancer prevention strategies. Earlier studies have reported ethnic-specific differences in breast cancer.[14] There are two primary differences: environmental factors (including epigenetic factors) and genetic factors. Environmental factors such as climate, diet, alcohol consumption, smoking, estradiol exposure, and infection are considered to be involved in modifying the promotion and suppression of onset.[15] Genetic factors differ between ethnicities in terms of single-nucleotide polymorphisms and variants.[16-19]

In the present study, we investigated risk factors for being a carrier of breast cancer susceptibility genes, especially PGVs in *BRCA1* and *BRCA2*,[4] in order to assess the application of specific testing criteria and a *BRCA* PGV carrier prediction model, namely the

Tyrer-Cuzick (TC) model,[20] in a cohort of 1995 unselected Japanese women with primary breast cancer.

2 | MATERIALS AND METHODS

2.1 | Patient cohort and clinical data collection

We analyzed a cohort of 1995 unselected Japanese women with primary breast cancer registered at Kyoto Breast Cancer Research Network institutions, including Kyoto University Hospital and 15 affiliated hospitals from September 2011 to October 2016. Genomic DNA samples from the peripheral blood of these patients were analyzed previously.[21] The dataset used in this study was updated with relevant clinical information until December 31, 2019. Written informed consent was obtained from all participants. This study was reviewed and approved by the Ethics Committee of Kyoto University Graduate School and Faculty of Medicine and Kyoto University Hospital (G424) and was performed in accordance with the Declaration of Helsinki and its later amendments.

We obtained the following data from the registration database for each patient: age; menopause history; reproductive history; biopsy history of benign and precancerous lesions, such as atypical ductal hyperplasia, lobular carcinoma in situ, and hyperplasia; hormone replacement therapy history; personal history of cancer; family history of cancer (up to third-degree relatives); imaging data, including mammary gland density in mammography (MMG); epidemiological information; clinicopathological data of breast cancer; BMI; and other factors.

Age was defined as the age at first breast cancer diagnosis. MMG density was classified by two certified physicians into four categories according to the American College of Radiology's BI-RADS® (Breast Imaging Reporting and Data System) Atlas, 5th edition. MMG assessment was limited to the untreated mammary gland during the first visit. Synchronous and metachronous bilateral breast cancer cases were excluded from MMG evaluation. These data were cross-checked with the patient's records at each institution. Missing values were identified by cross-checking the records during the respective institution visits and were updated before December 31, 2019.

2.2 | Profiling the PGVs of the cohort

This study used previously analyzed data.21 Briefly, targeted sequencing was used to analyze genomic DNA samples from the patients' peripheral blood. In this study, variants classified as "pathogenic" or "likely pathogenic" were defined as PGVs.

2.3 | Family history scoring

To represent family history of cancer as a score, we determined the sum of family history of cancer (up to third-degree relatives) as a "cancer family history score," which indicates the number of breast, ovarian, pancreatic, and prostate cancers in the family up to third-degree relatives. These specific cancers were included because they are described as related cancers that should be regarded as family history of cancer in the National Comprehensive Cancer Network (NCCN) guideline criteria for predicting PGV probability of breast cancer susceptibility genes.[22] Bilateral breast cancer was counted as two cancers. For example, if a patient has a sister with bilateral breast cancer, and her mother has ovarian cancer, the cancer family history score would be 3 (ie, 2 breast cancers and 1 ovarian cancer.)

2.4 | Validation of the NCCN guideline criteria

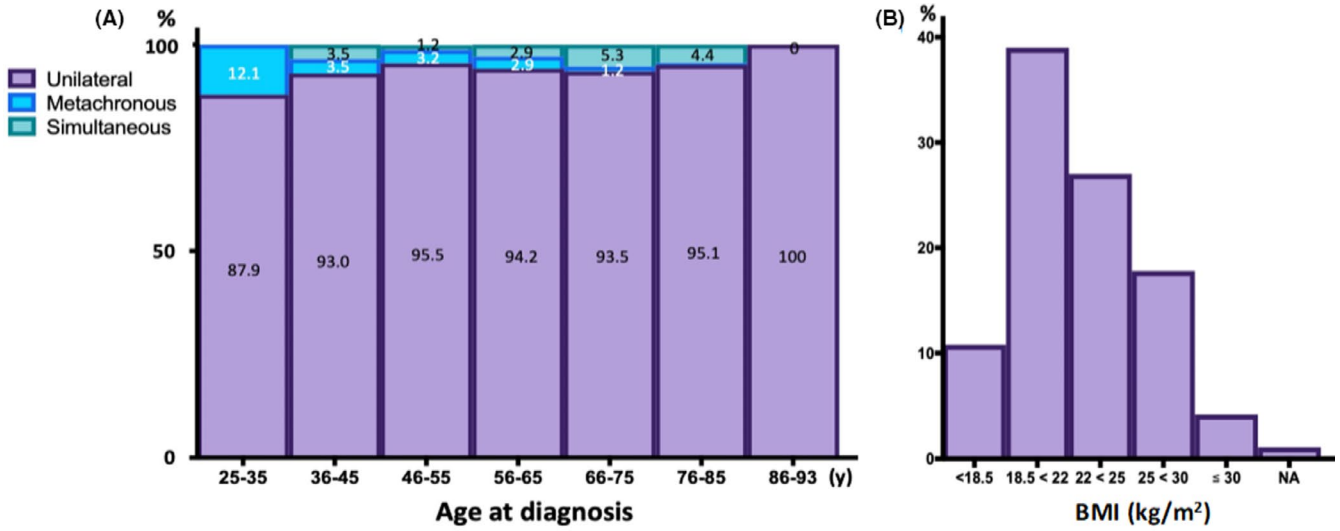
We assessed high-risk patients according to personal history of cancer (breast, ovarian, and pancreatic) and family history of cancer (breast, ovarian, pancreatic, and prostate) based on the NCCN Guidelines, Genetic/Familial High-Risk Assessment: Breast, Ovarian, and Pancreatic (Version 1.2020, December 4, 2019),[22] and then we validated the results.

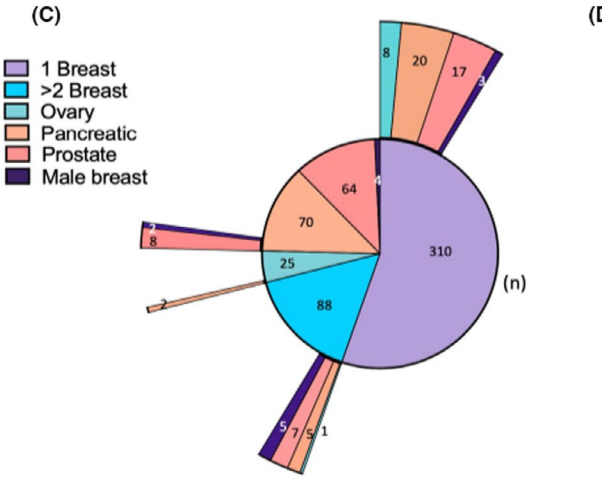
2.5 | Evaluation of BRCA carrier probability

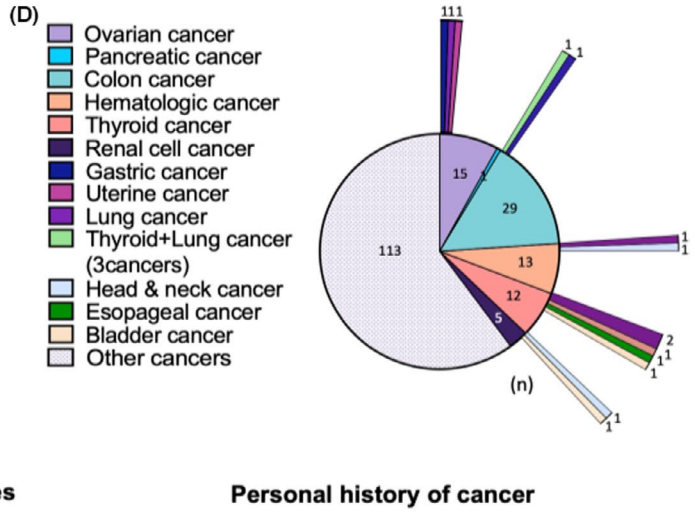
To predict the risk of carrying PGV, we applied the TC model (software package IBIS v8). We used the TC model because it was more useful compared with the risk models BOADICEA[23] and BRCAPro,[24] given that the TC model can tolerate missing values, and the information required for input could be obtained through interviews and from MMG in clinical practice. In contrast, our data were insufficient to use the risk model Penn II.[25] The TC model is a logistic regression breast cancer risk assessment tool based on the International Breast Intervention Study and UK National Cancer Statistics. It calculates a risk prediction score based on the following information: personal history of cancer, family history of cancer, BMI, age, MMG breast density, reproductive history, number of past breast biopsies, menopausal status, hormone replacement therapy history. We input these data to the TC model, and the *BRCA* personal % was calculated. The probability was independently assessed and confirmed by more than two researchers. Disagreement between two researchers was resolved through discussion. This model was able to tolerate missing data. We used the sum of *BRCA1* and *BRCA2* personal % as *BRCA* personal %.

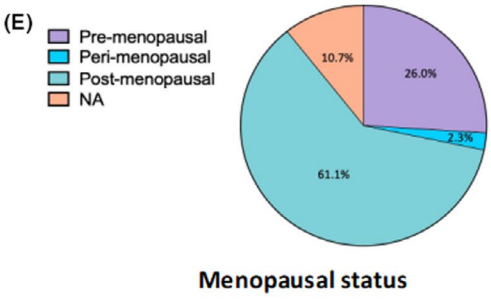
2.6 | Statistical analysis

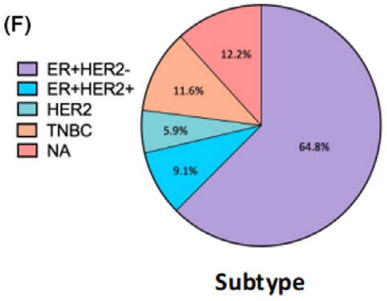
To assess the risk factors, multivariate logistic regression analysis was performed using patients' clinical information, including age, BMI, subtype, family history of cancer, personal history of cancer, and MMG breast density according to the BI-RADS classification in terms of *P*-values and odds ratios (ORs). All statistical analyses were performed using JMP Pro (ver., 14.1.0; SAS Institute Inc).


(A)
%
100
50
0
Unilateral
Metachronous
Simultaneous
12.1
3.5
3.5
1.2
3.2
2.9
2.9
5.3
4.4
87.9
93.0
95.5
94.2
93.5
95.1
100
25-35
36-45
46-55
56-65
66-75
76-85
86-93 (y)
Age at diagnosis
(B)
%
40
30
20
10
0
<18.5
18.5 < 22
22 < 25
25 < 30
≤ 30
NA
BMI (kg/m²)


(C)
1 Breast
>2 Breast
Ovary
Pancreatic
Prostate
Male breast
8
20
17
64
70
25
88
310
(n)
Family history of cancer up to third-degree relatives


(D)
Ovarian cancer
Pancreatic cancer
Colon cancer
Hematologic cancer
Thyroid cancer
Renal cell cancer
Gastric cancer
Uterine cancer
Lung cancer
Thyroid+Lung cancer (3cancers)
Head & neck cancer
Esopageal cancer
Bladder cancer
Other cancers
111
15
29
113
13
12
5
(n)
Personal history of cancer


(E)
Pre-menopausal
Peri-menopausal
Post-menopausal
NA
10.7%
26.0%
2.3%
61.1%
Menopausal status


(F)
ER+HER2-
ER+HER2+
HER2
TNBC
NA
12.2%
11.6%
5.9%
9.1%
64.8%
Subtype

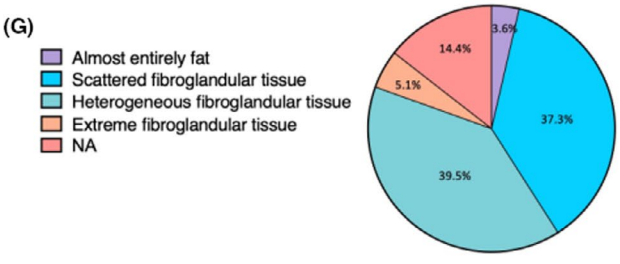

(G)
Almost entirely fat
Scattered fibroglandular tissue
Heterogeneous fibroglandular tissue
Extreme fibroglandular tissue
NA
3.6%
14.4%
5.1%
37.3%
39.5%
Mammography breast density

FIGURE 1 Patient and tumor characteristics (n = 1995 cases). A, The ratio of metachronous and simultaneous bilateral breast cancer cases by the age at first diagnosis. The metachronous bilateral breast cancer is defined the contralateral breast cancer onset from their enrollment to December 2019. B, The prevalence of body mass index. C, The prevalence of cases with family history of cancer: third-degree relatives. There were 561 cases with family history of cancer relating to hereditary breast ovarian cancer syndrome (ie, breast, ovarian, pancreatic, prostate, and male breast cancer). 483 cases had family history of 1 cancer, 74 cases had 2 cancers, and 4 cases had 3 cancers. Inner pie chart shows the distribution of cases by cancer type; outer shows the cases with more than two cancers. D, The personal history of cancer. There are 188 cases with personal history of cancer. 172 cases had personal history of 1 cancer, 13 cases had 2 cancers, and 1 case had 3 cancers. Inner pie chart shows the distribution of cases by cancer type; outer shows cases with more than two cancers. E, Percent of menopausal status in the first diagnosis. F, Percent of subtypes of the tumor. G, Percent of mammography breast density. The assessment was limited to the untreated mammary gland during the first visit. Synchronous and metachronous bilateral breast cancer cases were excluded from evaluation

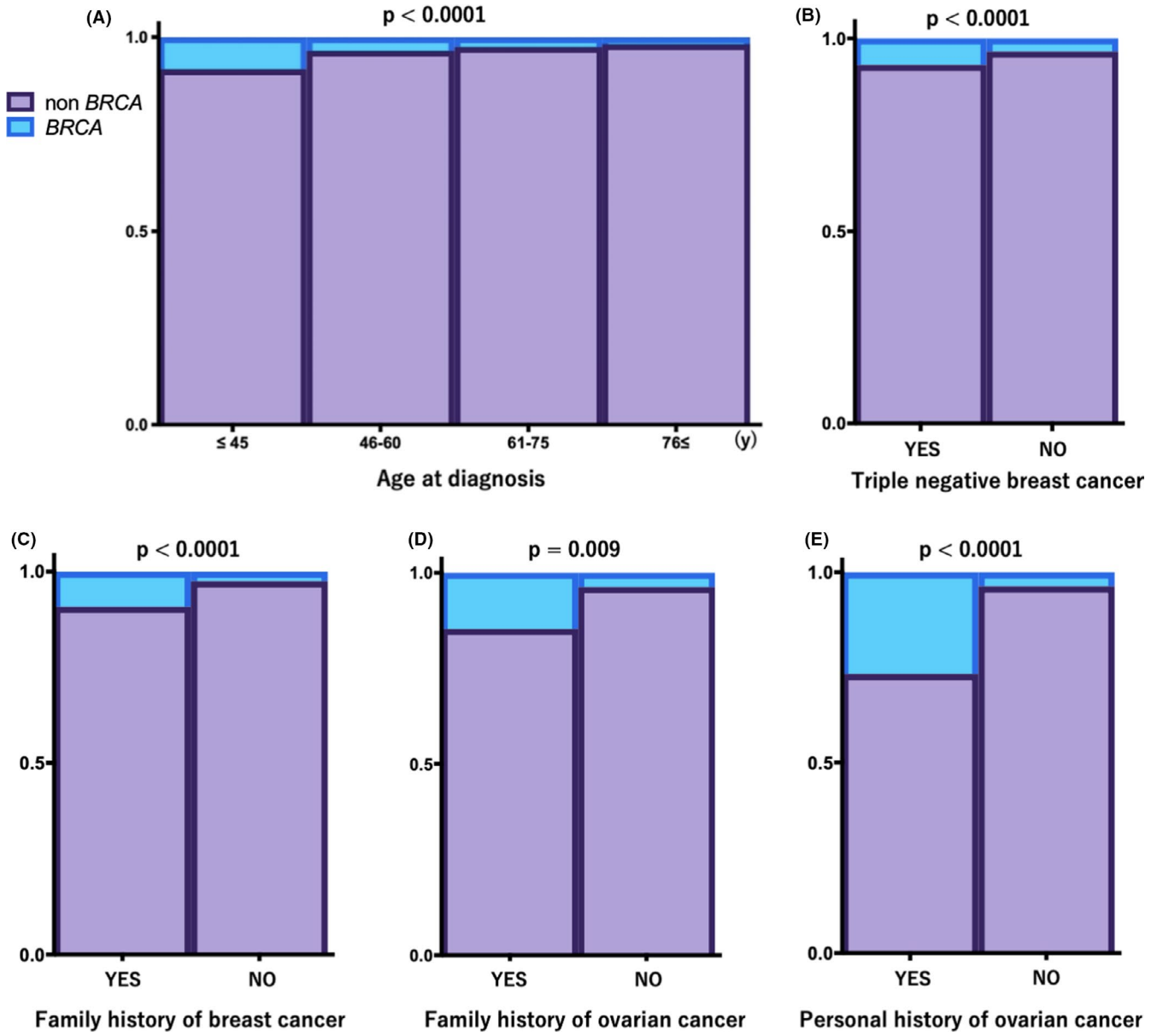


FIGURE 2 The correlation between major clinicopathological factors and the presence of *BRCA* pathogenic or likely pathogenic germline variant (PGV). Univariate analysis was performed on *BRCA* PGV and following clinicopathological factors: age at diagnosis A, triple-negative breast cancer B, family history of breast cancer C, family history of ovarian cancer D, personal history of ovarian cancer E. These factors showed significant correlation with *BRCA* PGV

To evaluate the NCCN guidelines, our cohort was classified into three groups, "indicated," "considered," and "not recommended" for MGT according to the NCCN guideline criteria. Here, "indicated" was defined as meeting these criteria, and "considered" and "not recommended" were defined as not meeting them. Correlation analysis between two groups ("indicated" vs. the combined "considered" and "not recommended" group) and PGV carrier status was performed using the chi-square test on GraphPad Prism (ver 8.4.3; GraphPad Software).

To validate the ability of the TC model to discriminate PGV carriers from noncarriers, we performed receiver operating characteristic (ROC) curve analysis and calculated the areas under the curve (AUCs) with 95% confidence intervals (CI) to determine sensitivity and specificity using GraphPad Prism (ver. 8.4.3, GraphPad Software). The PGV carrier *P*-value, probability, and specificity at the optimal cutoff point were recorded. The cutoff point was calculated using JMP Pro (ver. 14.1.0, SAS Institute Inc).

All other statistical analyses were performed using GraphPad Prism (ver. 8.4.3, GraphPad Software). Statistical significance was represented as 95% CIs and *P*-values <0.05.

3 | RESULTS

3.1 | Patient characteristics

Patient characteristics are shown in Figure 1 and Table S1. In this cohort of 1995 women with primary breast cancer, the PGV frequencies were as follows: *BRCA1* (0.75%), *BRCA2* (3.1%), *PALB2* (0.45%),

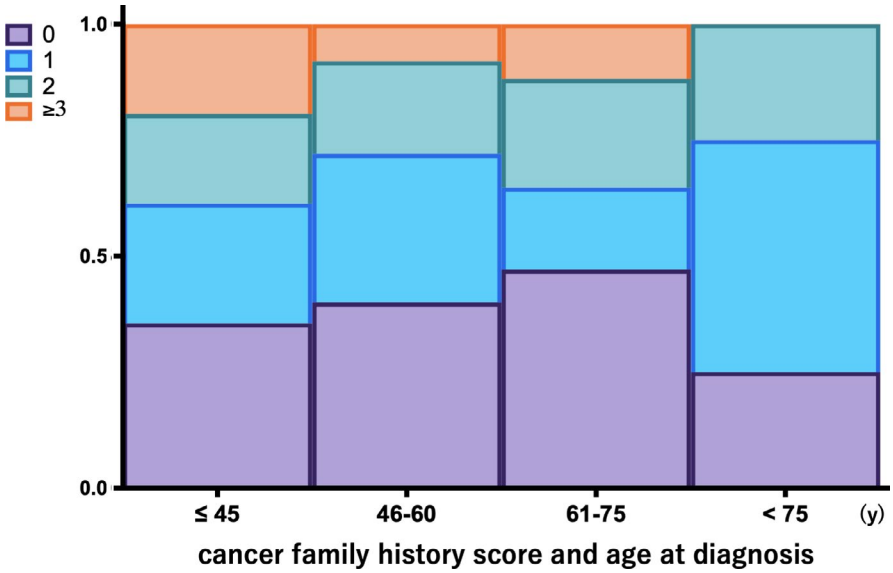


FIGURE 3 Stacked bar chart of the ratio of cancer family history score by age at diagnosis. Cancer family history score: number of cancers developed in one's family (up to third-degree relatives). Breast cancer, ovarian cancer, pancreatic cancer, and prostate cancer were included in this scoring. Bilateral breast cancer was counted as two cancers

TABLE 1 Multivariate logistic regression analysis for *BRCA* PGV

		Multivariate analysis				Univariate analysis			
Factors		OR	95% CI		*P*-value	OR	95% CI		*P*-value
Age, y	≤45	3.7	1.27	10.81	0.0059	4.96	1.93	16.8	0.0004
	46-60	1.6	0.54	4.71	0.371	2.07	0.79	7.09	0.1458
	61-75	1.31	0.43	3.98	0.6259	1.4	0.53	0.51	0.532
	≥75								
Cancer family history score	0								
	1	2.4	1.35	4.25	0.0038	2.43	1.28	4.5	0.0074
	2	7.33	3.82	14.07	<0.0001	6.01	2.77	12.29	<0.0001
	≥3	13.75	6.04	31.33	<0.0001	16.72	6.83	38.44	<0.0001
TNBC ≤60 y or bilateral breast cancer or personal history of ovarian cancer	Yes	3.04	1.77	5.2	<0.0001	1.69	0.98	2.8	0.0594
	No								

Note: Breast cancer, ovarian cancer, pancreas cancer, and prostate cancer were included in this scoring. Bilateral breast cancer was counted as two cancers. Cancer family history score indicates the number of cancers developed up to third-degree relatives.

Abbreviations: CI, confidence interval; OR, odds ratio; PGV, pathogenic or likely pathogenic germline variant; TNBC, triple-negative breast cancer.

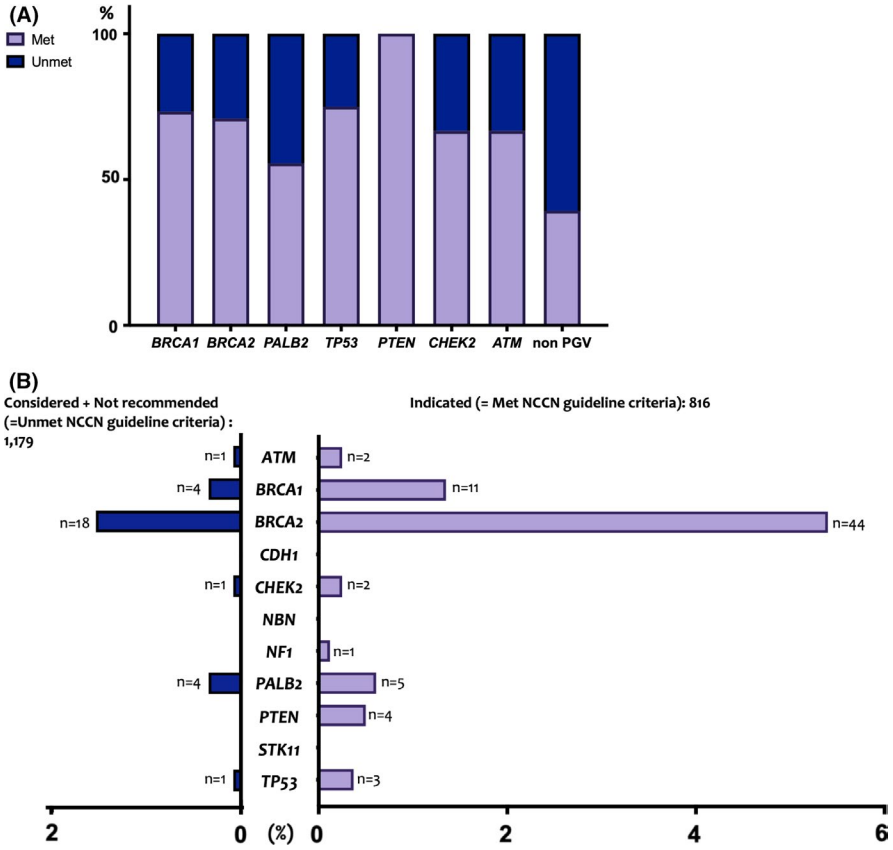


FIGURE 4 Prevalence of pathogenic or likely pathogenic germline variants (PGVs) in the groups assigned according to the National Comprehensive Cancer Network (NCCN) criteria. A, The prevalence of cases by each gene PGV. B, Comparison of prevalence of PGVs between met and unmet NCCN criteria. The 1995 patients in the cohort were classified into three groups according to the NCCN guideline criteria. A, The prevalence of cases with *BRCA* PGV was 71.4% in the "met" group, and 28.6% in the "unmet" group. The "met" group had significantly more PGVs compared with the "unmet" groups (P <.0001). B, Comparison of prevalence of PGVs between met and unmet NCCN guideline criteria. Based on the NCCN guidelines, more likely to carry PGVs than to not carry PGVs (3.7% vs 1.4%; P <.0001). Similar results were observed for the six high-risk genes (2.0% vs 0.6%; P =.005) and for *BRCA1/2* (6.7% vs 1.7%; P <.0001)

PTEN (0.20%), *TP53* (0.20%), *CHEK2* (0.15%), *ATM* (0.15%), *NF1* (0.05%), *CDH1* (0%), *STK11* (0%), and *NBN* (0%).

3.2 | Risk factors and prediction

We assessed the prevalence of *BRCA* and other genes PGV carriers in subgroups of risk factors. The risk factor assessment results are shown in Figure 2. Presence of the *BRCA* PGV was significantly correlated with age at disease onset (P <.0001). Additionally, early onset of the disease (age ≤45 years; P <.0001), personal history of ovarian cancer (P <.0001), family history of breast cancer (P <.0001), family history of ovarian cancer (P <.0009), triple-negative breast cancer (TNBC; P =.0085), and bilateral breast cancer onset (P =.0425) were also risk factors.

Figure 3 shows composites of age of onset, TNBC, and family cancer history score for four types of cancer. The cancer family history score, particularly ≥2, was significantly correlated with increased *BRCA* PGV rates (P <.05).

Table 1 shows the multivariate and univariate analysis results. Significant associations were identified between the presence of *BRCA* PGV and early onset, cancer family history score ≥1, and TNBC development in patients aged ≤60 years with bilateral breast cancer, or a personal history of ovarian cancer (P =.0059, P <.0001, and P <.0001, respectively).

BMI (P =.85) and MMG density (P =.45) were not significantly associated with the presence of *BRCA* variants. In 14.5% of cases, the MMG density was not applicable due to the presence of bilateral breast cancer, T4b (ulceration, bleeding), large lumps, referrals from other hospitals, pregnancy, lactation period, pacemaker implantation, pain, or a lack of imaging information for other reasons. Risk factors of other genes with PGVs were not detected in this study (data not shown).

3.3 | NCCN guideline criteria

The 1995 patients in our cohort were classified into three groups according to the NCCN criteria (Table S2). Of these cases, 40.9% met the NCCN criteria for MGT and 59.1% did not (sensitivity 71.3% and specificity 60.7%, P <.0001). The prevalence of cases with *BRCA* PGV and PGVs in other genes was respectively 6.7% and 2.1% in the "indicated" group, 6.1% and 0% in the "considered" group, and 1.7% and 0.6% in the "not recommended" group (Table S2). Figure 4 shows the number of cases and the ratio of those meeting to those not meeting the NCCN criteria for each of the 11 genes analyzed in this study.

3.4 | TC model

The TC model demonstrated a significant ability to predict the probability of carrying *BRCA1* (AUC, 0.77; 95% CI, 0.64-0.91), *BRCA2* (AUC, 0.74; 95% CI, 0.67-0.81), and *BRCA1/2* (AUC, 0.75; 95% CI, 0.69-0.81; Figure 5). However, the probability was lower than

expected from the results of the previous study, and the best cutoffs from the ROC were determined to be 0.15% in *BRCA1*, 0.12% in *BRCA2*, and 0.16% in *BRCA1/2*. Based on the TC model used in this study, the sensitivity for predicting PGVs was 53.3% in *BRCA1*, 59.0% in *BRCA2*, and 63.2% in *BRCA1/2*, and the specificity was found to be 92.2% in *BRCA1*, 79.2% in *BRCA2*, and 76.8% in *BRCA1/2*. Given that the NCCN guidelines use a cutoff of >5% for recommendation and of 2.5%-5% for consideration for further screening, we also assessed the TC performance with these cutoffs in each gene (Figure 6, Table S3). For cases meeting the NCCN criteria and/or TC ≥0.16%, the sensitivity for predicting BRCA1/2 PGVs was 79.2% and the specificity was 54.5%. And for the same population, the sensitivity for predicting 11 genes PGVs was 77.2% and the specificity was 54.8%.

4 | DISCUSSION

Predicting PGVs is important for managing breast cancer. However, PGV differences based on ethnicity remain unclear. We evaluated the risk factors for breast cancer–susceptible PGVs, particularly *BRCA1/2*, as well as the performance of the TC model in predicting pathogenic or likely pathogenic *BRCA* carriers in a Japanese cohort of 1995 unselected women with primary breast cancer.

The prevalence of *BRCA* PGV was 3.9%, while 2.1-5.3% have been reported.[6,7,16,17,26] The risk factors of *BRCA* PGV were similar to those reported previously. Combining these risk factors (TNBC, age, cancer family history score) might lead to the establishment of criteria for genetic testing (Figure 3). Moreover, the criteria for the Japanese Breast Cancer Society composite of these risk factors proved informative (Table S4).[27,28] The NCCN criteria also showed sufficient discriminatory ability, as reported previously.[26,29] A study conducted in the USA to validate the NCCN guidelines also reported the ratio for the "indicated" population as 47.9%, which is comparable to our results.[26] However, as shown in Figure 1, the prevalence of the related gene appears to differ between the two studies. Hence, our data can be used to guide clinical practice, particularly in Japan.

The TC model performed well in this Japanese breast cancer cohort. ROC curve analysis revealed significant predictive values for carrying *BRCA1/2* PGVs. Interestingly, although ROC curve analysis demonstrated the high discriminatory power of the TC model, our data showed that the best cutoff was 0.16% for *BRCA1/2*, which is nearly the same as the prevalence of *BRCA1/2* in the healthy population.[6,16,17] A possible reason for this is that the TC model is based on the incidence, factors, and parameter values from Western countries. The incidence of breast and ovarian cancers by age group, which also affects family or personal history of cancer, differs between Western countries and Japan.[30] This difference in incidence might be explained by genetic and environmental factors.

As for genetic factors, the prevalence of PGVs in *BRCA* or other genes may differ according to country or region.[6,7,14,31,32] In Asian countries, reports indicate that the prevalence of *BRCA2* PGVs appears to be higher than that of *BRCA1* PGVs.[6-8,33] The prevalence of

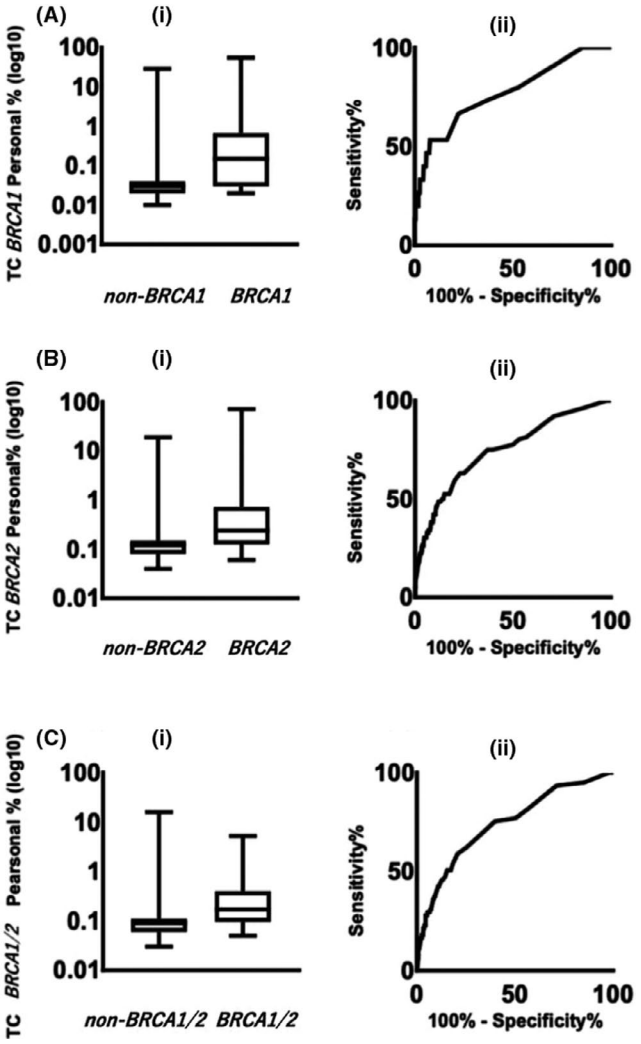


FIGURE 5 Distribution of probability from Tyrer-Cuzick (TC) model and receiver operating characteristic (ROC) curves. A, (i) & (ii): Distribution of patients carrying *BRCA1* PGVs. B, (i) & (ii): Distribution of patients carrying *BRCA2* pathogenic or likely pathogenic germline variants (PGVs). C, (i) & (ii): Distribution of patients carrying *BRCA1/2* PGVs. Boxes represent the interquartile range of the distribution. A(i), B(i), C(i): TC *BRCA1*, *BRCA2*, *BRCA1/2* personal % was significantly higher in PGV carriers than in non-PGV carriers. The horizontal line within the box represents the median, and the vertical lines represent the 95% confidence interval (CI). A(ii), B(ii), C(ii): ROC prediction curves for patients carrying *BRCA1*, *BRCA2*, and *BRCA1/2* PGVs. The TC model demonstrates discrimination in predicting the probability of carrying *BRCA1* PGV (area under the curve [AUC], 0.77; 95% CI, 0.64-0.91), *BRCA2* PGV (AUC, 0.74; 95% CI, 0.67-0.81), and *BRCA1/2* PGV (AUC, 0.75; 95% CI, 0.69-0.81)

each PGV in the same gene differs by country because some PGVs are not reported in the archive ClinVar (Figure S1) in our cohort or in previous studies.[6,34] Moreover, the risk of cancer associated with these PGVs remains unknown because it has been reported that some PGVs in *BRCA* are related to ovarian cancer, whereas others are related to breast cancer, indicating that phenotypes might differ depending on the PGVs.[35] These differences in the type and nature of PGVs might affect the incidence of some cancers. As for

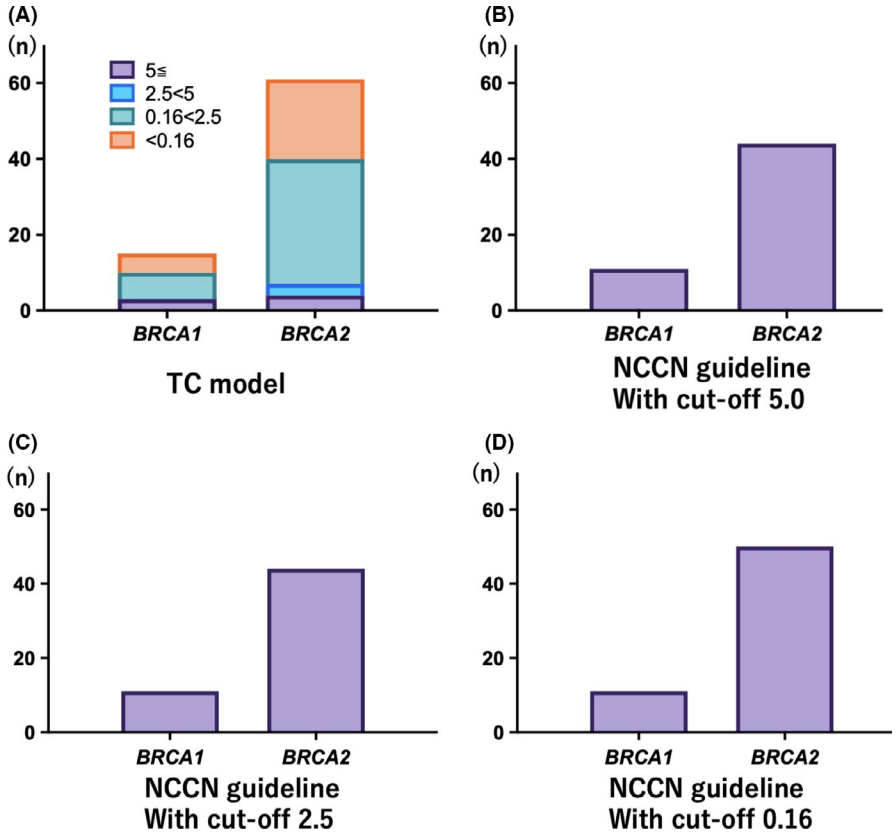


FIGURE 6 Distribution of *BRCA1/2* pathogenic or likely pathogenic germline variant (PGV) carriers (n = 77) in groups classified according to the National Comprehensive Cancer Network (NCCN) guidelines and Tyrer-Cuzick (TC) score. (A) TC score (B) the NCCN guidelines with TC score cutoff of >5% (C) the NCCN guidelines with TC score cutoff of 2.5% and (D) the NCCN guidelines with TC score cutoff of TC score ≥0.16% (details are shown in Table S3).

single-nucleotide polymorphisms, germline variants and polygenetic single-nucleotide combinations have recently been used to improve the prediction probability of *BRCA* carriers and breast cancer risk assumptions, which is an area of research that warrants further investigation.[36]

Environmental risk factors for breast cancer such as alcohol intake, obesity, and MMG density may differ according to ethnicity.[30,37] The rate of BMI ≥30% was 4.2% in our cohort and 4.4% in the general Japanese female population, compared with 37.3% in the USA and 29.5% in the United Kingdom.[30] The distributions of age-specific patterns in MMG density also differ by country or region.[38-40] These factors may be related to the lower breast cancer incidence in Japan. Further nationwide and global studies are needed to elucidate the impact on cancer incidence of both genetic and environmental factors as well as their interaction because we were unable to validate these in the present study.

In the future, it is expected that the cost of MGT will decrease. However, considering the high price currently, it is difficult in practice to recommend MGT for all breast cancer patients worldwide. Furthermore, our data showed that some PGV carriers had late onset of breast cancer, suggesting a need to identify the population of PGV carriers who would benefit from risk-reduction surgery. The PGV carrier prediction model is an essentially free, convenient, and useful tool that can be used to support the decision-making process for high-risk individuals and might also be useful for cancer risk assessment. The model is based mainly on family and personal history, similar to the TC model and can be used for the healthy population. Because the genetic and environmental factors remain unclear, as discussed above, it is practical and beneficial to use these models to screen high-risk patients or high-risk healthy individuals and to recommend genetic testing and regular checkups based on the results.

This study had some limitations. First, most patients in this cohort were from western Japan. Because there are few reports of regional PGV differences in Japan, our results are expected to be almost the same as those in studies using other unselected cohorts in Japan.[6] Second, the patient histories might be underestimated given that they were obtained through a questionnaire survey distributed at hospital outpatient clinics. Third, PGV carriers of breast cancer susceptibility genes except *BRCA* were rare in our study, and thus we were unable to precisely validate prediction methods for these genes. Compared with a previous study of a high-risk cohort, there might be differences in the influence of family history on PGV prediction for each gene, suggesting that it is necessary to study such prediction to improve the model in the future.[41]

In conclusion, we assessed the risk factors of PGVs and established a composite of clinicopathological parameters that might help us predict the probability that a person is carrying *BRCA1/2* PGV. In addition, we demonstrated that the NCCN guidelines and the TC model can be readily applied for Japanese women with breast cancer to identify *BRCA* PGV carriers. However, other prediction models should also be examined, given that precise prediction of germline variants is becoming increasingly crucial for future breast cancer management and prevention. Further studies are also needed to detect ethnic-specific PGV differences, particularly environmental and genetic factors, as well as other genes containing PGVs.

ACKNOWLEDGMENTS

We thank all the patients who participated in this study of the Kyoto Breast Cancer Research Network and Breast Oncology Research Network Bio-Bank, and to all the members of these organizations. We would also like to thank the members of Seishi Ogawa's laboratory, where the previous target sequencing study (submitted) was conducted. We also appreciate the Hereditary Breast and Ovarian Cancer subunit group members, especially, Takahiro Yamada, Hiromi Murakami, and Sayaka Honda of Kyoto University Hospital for providing valuable information. This research was partially supported by an AstraZeneca Externally Sponsored Research grant (NCR-17-12919).

CONFLICT OF INTEREST

Masakazu Toi discloses research grants from AstraZeneca, Pfizer, Eisai, Chugai Pharma, Astellas Pharma, Taiho Pharmaceutical, Shimadzu, the Japan Breast Cancer Research Group, Nippon Kayaku, GL Sciences, Luxonus, and Tokyo University; and honoraria from AstraZeneca, Eli Lilly, Kansai Medical Net, and Terumo Corporation. Eiji Suzuki discloses research grants from Daiichi Sankyo, Astellas Pharma, Kyowa Kirin, Chugai Pharma, and MSD. Masahiro Takada discloses research grants from AstraZeneca, Daiichi Sankyo, the Kyoto Breast Cancer Research Network, and the Austrian Breast & Colorectal Study Group; and honoraria from Chugai Pharma. Masahiro Kawashima discloses a research grant from Nippon Kayaku. Kosuke Kawaguchi discloses a research grant from Terumo Corporation. The remaining authors declare no conflict of interest.

ORCID

Noriko Senda https://orcid.org/0000-0002-5836-7556
Yoshiaki Matsumoto https://orcid.org/0000-0002-4256-2349
Masakazu Toi https://orcid.org/0000-0003-1488-9958

REFERENCES

1. Hori M, Matsuda T, Shibata A, Katanoda K, Sobue T, Nishimoto H. Cancer incidence and incidence rates in Japan in 2009: a study of 32 population-based cancer registries for the Monitoring of Cancer Incidence in Japan (MCIJ) project. *Jpn J Clin Oncol*. 2015;45(9):884-891.
2. Foulkes WD. Inherited susceptibility to common cancers. *N Engl J Med*. 2008;359:2143-2153.
3. Whittemore AS. Risk of breast cancer in carriers of *BRCA* gene mutations. *N Engl J Med*. 1997;337:788-789.
4. Easton DF, Pharoah PD, Antoniou AC, et al. Gene-panel sequencing and the prediction of breast-cancer risk. *N Engl J Med*. 2015;372:2243-2257.
5. Miki Y, Swensen J, Shattuck-Eidens D, et al. A strong candidate for the breast and ovarian cancer susceptibility gene *BRCA1*. *Science*. 1994;266:66-71.
6. Momozawa Y, Iwasaki Y, Parsons MT, et al. Germline pathogenic variants of 11 breast cancer genes in 7,051 Japanese patients and 11,241 controls. *Nat Commun*. 2018;9:4083.
7. Sun J, Meng H, Yao L, et al. Germline Mutations in cancer susceptibility genes in a large series of unselected breast cancer patients. *Clin Cancer Res*. 2017;23:6113-6119.
8. Nakamura S, Takahashi M, Tozaki M, et al. Prevalence and differentiation of hereditary breast and ovarian cancers in Japan. *Breast Cancer*. 2015;22:462-468.
9. Finch AP, Lubinski J, Møller P, et al. Impact of oophorectomy on cancer incidence and mortality in women with a *BRCA1* or *BRCA2* mutation. *J Clin Oncol*. 2014;32:1547-1553.
10. Li X, You R, Wang X, et al. Effectiveness of prophylactic surgeries in *BRCA1* or *BRCA2* mutation carriers: A meta-analysis and systematic review. *Clin Cancer Res*. 2016;22:3971-3981.
11. Tung NM, Boughey JC, Pierce LJ, et al. Management of hereditary breast cancer: American Society of Clinical Oncology, American Society for Radiation Oncology, and Society of Surgical Oncology Guideline. *J Clin Oncol*. 2020;38:2080-2106.
12. Valencia OM, Samuel SE, Viscusi RK, Riall TS, Neumayer LA, Aziz H. The role of genetic testing in patients with breast cancer: a review. *JAMA Surg*. 2017;152:589-594.
13. Owens DK, Davidson KW, Krist AH, et al. Risk assessment, genetic counseling, and genetic testing for *BRCA*-related cancer: US Preventive services task force recommendation statement. *JAMA*. 2019;322:652-665.
14. Armstrong N, Ryder S, Forbes C, Ross J, Quek RG. A systematic review of the international prevalence of *BRCA* mutation in breast cancer. *Clin Epidemiol*. 2019;11:543-561.
15. Lichtenstein P, Holm NV, Verkasalo PK, et al. Environmental and heritable factors in the causation of cancer–analyses of cohorts of twins from Sweden, Denmark, and Finland. *N Engl J Med*. 2000;343:78-85.
16. Dorling L, Carvalho S, Allen J, et al. Breast cancer risk genes - association analysis in more than 113,000 women. *N Engl J Med*. 2021;384(5):428–439.
17. Hu C, Hart SN, Gnanaolivu R, et al. A population-based study of genes previously implicated in breast cancer. *N Engl J Med*. 2021;384(5):440-451.
18. Apostolou P, Fostira F. Hereditary breast cancer: the era of new susceptibility genes. *Biomed Res Int*. 2013;2013:1-11.
19. Meindl A, Ditsch N, Kast K, Rhiem K, Schmutzler RK. Hereditary breast and ovarian cancer: new genes, new treatments, new concepts. *Dtsch Arztebl Int*. 2011;108(19):323-330.
20. Tyrer J, Duffy SW, Cuzick J. A breast cancer prediction model incorporating familial and personal risk factors. *Stat Med*. 2004;23:1111-1130.
21. Inagaki-Kawata Y, Yoshida K, Kawaguchi-Sakita N, et al. Genetic and clinical landscape of breast cancers with germline *BRCA1/2* variants. *Commun Biol*. 2020;3(1):578.
22. Daly MB, Pilarski R, Yurgelun MB, et al. NCCN guidelines insights: Genetic/familial high-risk assessment: breast, ovarian, and pancreatic, version 1.2020. *J Natl Compr Canc Netw*. 2020;18:380-391.
23. Lee A, Mavaddat N, Wilcox AN, et al. BOADICEA: a comprehensive breast cancer risk prediction model incorporating genetic and non-genetic risk factors. *Genet Med*. 2019;21(8):1708-1718.
24. Mazzola E, Blackford A, Parmigiani G, Biswas S. Recent enhancements to the genetic risk prediction model BRCAPRO. *Cancer Inform*. 2015;14(Suppl 2):147-157.
25. Lindor NM, Johnson KJ, Harvey H, et al. Predicting *BRCA1* and *BRCA2* gene mutation carriers: comparison of PENN II model to previous study. *Fam Cancer*. 2010;9(4):495-502.
26. Yadav S, Hu C, Hart SN, et al. Evaluation of germline genetic testing criteria in a hospital-based series of women with breast cancer. *J Clin Oncol*. 2020;38:1409-1418.
27. Shimoi T, Nagai SE, Yoshinami T, et al. The Japanese Breast Cancer Society Clinical Practice Guidelines for systemic treatment of breast cancer. *Breast Cancer*. 2020;27(3):322-331.
28. Japanese Organization of Hereditary Breast and Ovarian (JOHBOC). Guidebook for diagnosis and treatment of hereditary

breast and ovarian cancer syndrome. 2017. http://johboc.jp/guidebook2017/
29. Beck AC, Yuan H, Liao J, et al. Rate of *BRCA* mutation in patients tested under NCCN genetic testing criteria. *Am J Surg*. 2020;219:145-149.
30. WHO. *Cancer Key Facts*. Geneva: WHO; 2018. https://www.who.int/news-room/fact-sheets/detail/cancer
31. Tung N, Battelli C, Allen B, et al. Frequency of mutations in individuals with breast cancer referred for *BRCA1* and *BRCA2* testing using next-generation sequencing with a 25-gene panel. *Cancer*. 2015;121:25-33.
32. Tung N, Lin NU, Kidd J, et al. Frequency of germline mutations in 25 cancer susceptibility genes in a sequential series of patients with breast cancer. *J Clin Oncol*. 2016;34:1460-1468.
33. Lang GT, Shi JX, Hu X, et al. The spectrum of *BRCA* mutations and characteristics of *BRCA*-associated breast cancers in China: Screening of 2,991 patients and 1,043 controls by next-generation sequencing. *Int J Cancer*. 2017;141:129-142.
34. Bhaskaran SP, Chandratre K, Gupta H, et al. Germline variation in *BRCA1/2* is highly ethnic-specific: Evidence from over 30,000 Chinese hereditary breast and ovarian cancer patients. *Int J Cancer*. 2019;145(4):962-973.
35. Rebbeck TR, Mitra N, Wan F, et al. Association of type and location of *BRCA1* and *BRCA2* mutations with risk of breast and ovarian cancer. *JAMA*. 2015;313:1347-1361.
36. Li Q, Seo JH, Stranger B, et al. Integrative eQTL-based analyses reveal the biology of breast cancer risk loci. *Cell*. 2013;152(3):633-634.
37. Renehan AG, Tyson M, Egger M, Heller RF, Zwahlen M. Body-mass index and incidence of cancer: a systematic review and meta-analysis of prospective observational studies. *Lancet*. 2008;371:569-578.
38. Arora N, King TA, Jacks LM, et al. Impact of breast density on the presenting features of malignancy. *Ann Surg Oncol*. 2010;17(Suppl 3):211-218.
39. Burton A, Maskarinec G, Perez-Gomez B, et al. Mammographic density and ageing: A collaborative pooled analysis of cross-sectional data from 22 countries worldwide. *PLOS Med*. 2017;14(6):e1002335.
40. Boyd NF, Martin LJ, Yaffe MJ, Minkin S. Mammographic density and breast cancer risk: current understanding and future prospects. *Breast Cancer Res*. 2011;13:223.
41. Kaneyasu T, Mori S, Yamauchi H, et al. Prevalence of disease-causing genes in Japanese patients with *BRCA*1/2-wildtype hereditary breast and ovarian cancer syndrome. *NPJ Breast Cancer*. 2020;6:25.

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section.

How to cite this article: Senda N, Kawaguchi-Sakita N, Kawashima M, et al. Optimization of prediction methods for risk assessment of pathogenic germline variants in the Japanese population. *Cancer Sci*. 2021;112:3338–3348. https://doi.org/10.1111/cas.14986