


BMJ Open Validation of breast cancer risk assessment tools on a French-Canadian population-based cohort

Rodolphe Jantzen ^{1,2}, Yves Payette,¹ Thibault de Malliard,¹ Catherine Labbé,¹ Nolwenn Noisel,^{1,2} Philippe Broët^{1,2,3,4}

To cite: Jantzen R, Payette Y, de Malliard T, *et al.* Validation of breast cancer risk assessment tools on a French-Canadian population-based cohort. *BMJ Open* 2021;**11**:e045078. doi:10.1136/bmjopen-2020-045078

► Prepublication history and additional material for this paper are available online. To view these files, please visit the journal online (<http://dx.doi.org/10.1136/bmjopen-2020-045078>).

Received 21 September 2020
Revised 17 February 2021
Accepted 12 March 2021



© Author(s) (or their employer(s)) 2021. Re-use permitted under CC BY-NC. No commercial re-use. See rights and permissions. Published by BMJ.

¹CARTaGENE, Research Center, CHU Sainte-Justine, Montreal, Quebec, Canada

²Université de Montréal, Montréal, Québec, Canada

³CESP, INSERM, University Paris-Saclay, Villejuif, France

⁴Assistance Publique-Hôpitaux de Paris (AP-HP), Hôpitaux Universitaires Paris-Sud, Hôpital Paul Brousse, Villejuif, France

Correspondence to

Dr Rodolphe Jantzen;
rodolphe.jantzen@gmail.com

ABSTRACT

Objectives Evaluate the accuracy of the Breast Cancer Risk Assessment Tool (BCRAT), International Breast Cancer Intervention Study risk evaluation tool (IBIS), Polygenic Risk Scores (PRS) and combined scores (BCRAT+PRS and IBIS+PRS) to predict the occurrence of invasive breast cancers at 5 years in a French-Canadian population.

Design Population-based cohort study.

Setting We used the population-based cohort CARTaGENE, composed of 43 037 Quebec residents aged between 40 and 69 years and broadly representative of the population recorded on the Quebec administrative health insurance registries.

Participants 10 200 women recruited in 2009–2010 were included for validating BCRAT and IBIS and 4555 with genetic information for validating the PRS and combined scores.

Outcome measures We computed the absolute risks of breast cancer at 5 years using BCRAT, IBIS, four published PRS and combined models. We reported the overall calibration performance, goodness-of-fit test and discriminatory accuracy.

Results 131 (1.28%) women developed a breast cancer at 5 years for validating BCRAT and IBIS and 58 (1.27%) for validating PRS and combined scores. Median follow-up was 5 years. BCRAT and IBIS had an overall expected-to-observed ratio of 1.01 (0.85–1.19) and 1.02 (0.86–1.21) but with significant differences when partitioning by risk groups ($p < 0.05$). IBIS' c-index was significantly higher than BCRAT (63.42 (59.35–67.49) vs 58.63 (54.05–63.21), $p = 0.013$). PRS scores had a global calibration around 0.82, with a CI including one, and non-significant goodness-of-fit tests. PRS' c-indexes were non-significantly higher than BCRAT and IBIS, the highest being 64.43 (58.23–70.63). Combined models did not improve the results.

Conclusions In this French-Canadian population-based cohort, BCRAT and IBIS have good mean calibration that could be improved for risk subgroups, and modest discriminatory accuracy. Despite this modest discriminatory power, these tools can be of interest for primary care physicians for delivering a personalised message to their high-risk patients, regarding screening and lifestyle counselling.

INTRODUCTION

Breast cancer is the most frequently diagnosed cancer and the second leading cause

Strengths and limitations of this study

- First study to evaluate risk assessment tools in a French-Canadian population for predicting breast cancer.
- Population based-cohort representative of the French-Canadian urban population of middle-aged and older adults.
- Linkage with administrative health databases and the Quebec Breast Cancer Registry, which improved the outcome quality and accuracy, and made possible to use variables usually difficult to obtain.
- May not apply to younger women under 40 years old.
- Since the genotyping information was not available for all the cohort, the models had to be evaluated on two different subcohorts.

of death by cancer among the Canadian women.¹ However, assessing the individual risk of breast cancer remains a challenge. In this context, risk prediction models have been developed and implemented. The two most widely used are the Breast Cancer Risk Assessment Tool (BCRAT) and the International Breast Cancer Intervention Study risk evaluation tool (IBIS).^{2,3}

The National Cancer Institute's (NCI) BCRAT was developed by Gail *et al*² using 5998 American women from a case-control study. It provides an estimate of a woman's risk of developing invasive breast cancer over a specific period, knowing her personal risk factors. After its first release, this model has been validated in an American cohort,⁴ mainly composed of white women, and was later calibrated for African American, Hispanic, Asian and Pacific Islander women.^{5,6} The most recent version uses six clinical risk factors: current age, age at first menstrual period, age at first live birth, number of first-degree relatives with breast cancer, history of previous breast biopsy and ethnicity. Several studies have assessed or updated the BCRAT model to specific populations (eg, Asian,

Oceanian).⁷ It is worth noting that this model, designed for use in the general population, is not intended to be used for women carrying inherited *BRCA1/2* mutations. The BCRAT model is used to guide physicians on breast cancer prevention strategies. As an example, the US Food and Drug Administration recommended to consider chemoprevention for women at high risk of breast cancer (ie, a 5-year risk equal or higher than 1.66%), while the US Preventive Services Task Force recommended chemoprevention for a risk equal or higher than 3%.⁸ The Canadian Task Force, as well as the Canadian Cancer Society, used a threshold of 1.66%.^{9,10} Despite its implementation on the NCI's website (bcrisktool.cancer.gov/), the lack of recent Canadian guidelines combined with its US-centred use led to an underuse of the BCRAT model by Canadian primary care physicians. Indeed, a recent qualitative study showed that two-third of primary care physicians from two Canadian provinces (Ontario and Alberta) were unaware of the BCRAT tool.¹¹

The IBIS model, also known as the Tyrer-Cuzick model, is also a widely used breast cancer risk prediction model, which takes into account multigenerational family history data and *BRCA1/2* mutation information. It has been developed with data from the International Breast Cancer Intervention Study including a cohort of daughters of patients diagnosed with the disease and has focused on the estimation of breast cancer lifetime risks through the analysis of family history, reproductive and hormonal factors, and individual characteristics.³ The IBIS model takes into account non-genetic risk factors (current age, age at menarche, number of live births, age at first live birth, age at menopause, height, weight, history of hyperplasia, breast density, history and age of ovarian cancer, hormone replacement therapy) together with multigenerational pedigree information and *BRCA1/2* gene mutations. IBIS can be used even for women without a family history of breast cancer and without *BRCA1/2* gene mutations information. A recent study suggested that IBIS has better ability to assess breast cancer risk than BCRAT but with close performance in women not known to have mutations in *BRCA1* or *BRCA2* genes.^{12–14}

With the increasing availability and affordability of genetic information, there is a growing interest to incorporate individual-level genotype data into risk prediction models for increasing their discriminatory accuracy. The integration of such information into the BCRAT model has already been performed with the addition of seven single-nucleotide polymorphisms (SNPs) associated with breast cancer. Results showed that the performance of the predicted breast cancer's risk was slightly improved, with an area under the receiver operating characteristic (ROC) curve (AUC) increasing from 0.607 to 0.632.¹⁵ This kind of clinicogenetic model has also been done with IBIS leading to an improvement in the discriminative ability.¹⁶ Alongside these works, many genetic based or 'Polygenic Risk Scores' (PRS) have been published for breast cancer prediction. Most of them rely on linear combinations of the risk-conferring variant alleles weighted by their

effect sizes.^{17–20} The list of these risk alleles with their corresponding weights is usually obtained from large case-control genome-wide association studies,²¹ with weights that can be adapted to specific ethnicities.¹⁹ The predictive accuracy of these PRSs compared with classical prediction models, such as the BCRAT and IBIS, should now be evaluated in various populations.

In Quebec, the Breast Cancer Screening Programme consists of a mammogram every 2 years for women aged 50–69.²² Although this screening decreased the number of deaths from breast cancer,²³ it could be stressful with non-negligible costs for the public health system. In this context, risk assessment tools could be helpful for primary care physicians to enhance screening uptake among high risk patients who are less likely to participate in organised screening. Some previous studies have assessed the accuracy of the BCRAT risk predictions in Canadian women,^{12,24} but they were limited to specific ethnic populations or were part of multicountries cohorts. The fact that BCRAT and IBIS have not been evaluated in the French-Canadian population, which has specific genetic patterns, as compared with the general European population,^{25,26} with lifestyle risk factors (eg, nutrition) that are at the intersection between North America and Europe, prompted us to evaluate their predictive abilities in the population-based cohort CARTaGENE from Quebec.

In this study, we report the predictive accuracies of the BCRAT model, the IBIS model and PRS to predict the occurrence of invasive breast cancers at 5 years in middle-aged and older French-Canadian women.

MATERIALS AND METHODS

Design and participants selection

The CARTaGENE population-based cohort is composed of 43 037 Quebec residents aged between 40 and 69 years, recruited during two phases (2009–2010 and 2013–2014). With a rich collection of data including phenotyping and genotyping data, CARTaGENE is the largest ongoing prospective population cohort and biobank in Québec, Canada.²⁷ Details on recruitment and sample selection have been described previously.²⁷

To comprehensively identify participants with an invasive breast cancer and the incidence date, we used two administrative health databases, the Quebec Health Insurance Board (RAMQ) and the Quebec Breast Cancer Registry (see online supplemental methods 1), and an algorithm based on a previous report from the Institut National de Santé Publique du Québec²⁸ and the Tonelli *et al.*²⁹ algorithm. Using the Breast Cancer Registry, we retrieved the incidence date of histologically confirmed breast cancers. Then, as some women with a breast cancer might not have a histologically confirmed cancers in the Breast Cancer Registry, we selected in this registry all women having an abnormal mammography (ie, lesion suspected of malignancy) without histologically confirmed breast cancers and retrieved, when available, the incidence date after the abnormal mammography

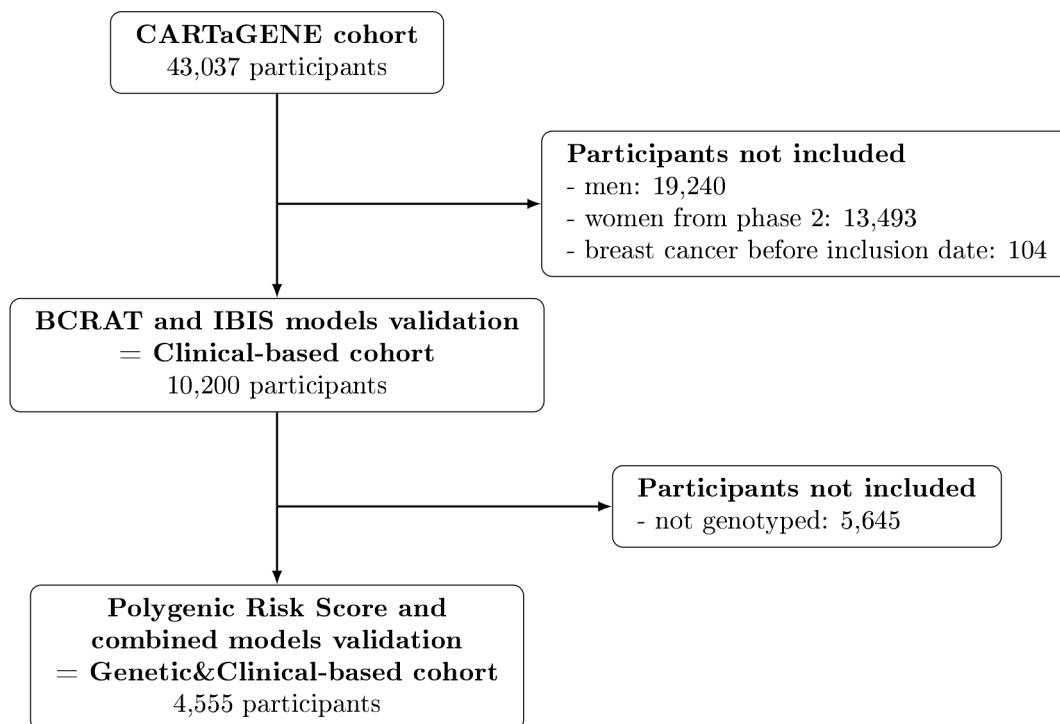


Figure 1 Flow chart. BCRAT, Breast Cancer Risk Assessment Tool; IBIS, International Breast Cancer Intervention Study.

from the RAMQ database for women with at least two claims in 2 years or one hospitalisation with the appropriate International Classification of Diseases, 9th or 10th Revision codes (174 and C50). The Quebec breast cancer registry's data were available from 15 May 1998 to 31 December 2017, while the RAMQ's data were available from 1 January 1998 to 31 March 2016. Adherence to mammography was not available.

For this study, we have considered the women without a breast cancer before the inclusion date from the CARTaGENE first phase of recruitment as the family history of breast cancer was not available for the participants of the phase 2. Recruitment was unrelated to the last mammography screening. The validation of the BCRAT and IBIS models was done on the subcohort of 10 200 women with available information for computing the BCRAT and IBIS models (hereinafter referred as clinical-based cohort (CC)). The validation of the PRS was done on the subcohort of 4555 women with available genotyping information (hereinafter referred as clinicogenetic-based cohort (CGC)) (figure 1). We also compared PRS to the BCRAT and IBIS models on the CGC cohort.

Genetic data

Only a fraction of the CARTaGENE population cohort has been genotyped. These participants were selected to be genotyped through various scientific projects unrelated to breast cancer.^{30–32} SNP positions were based on build GRCh37/hg19. The detailed pipeline about quality control and imputation can be found at www.cartagene.qc.ca/info-genetic-data and in online supplemental methods 1.

Outcome

The outcome of interest was the time of occurrence of the breast cancer from the enrollment in the cohort. Patients without breast cancer occurrence were censored at the end of the 5 years study period (administrative censoring) or at death.

Predictive scores

Absolute risk using the BCRAT and the IBIS models

The absolute risk of breast cancer estimated by BCRAT and IBIS is calculated using baseline hazard functions calculated from the marginal hazard functions (USA and UK incidence rates, respectively), and the attributable risk obtained from the United States population data (BCRAT) and the UK/Swedish population data (IBIS). In this article, the BCRAT and IBIS absolute risks of breast cancer at 5 years were calculated for each woman at the inclusion date using the National Institutes of Health R package 'BCRA', V.2.1³³ and the latest version of the 'IBIS Breast Cancer Risk Evaluation Tool' (<http://www.ems-trials.org/riskevaluator/>—V.8.0b, September 2017), respectively. Death as a competing risk was taken into account for both models.

All variables of the BCRAT model could be retrieved, while some variables of the IBIS model were not available and were considered missing: breast density, Ashkenazi Jewish heritage, HRT type, length of time woman intends to use HRT in the future, *BRCA1/2* genetic testing (participant and relatives), mother bilateral mastectomy, relatives' age of breast and ovary cancers, variables related to each sister, brother, grandmother, aunt, uncle and daughter. See online supplemental methods

1 for information about variables extraction and coding. Missing data can be handled in both BCRAT and IBIS models.

Absolute risk using PRS

For estimating the absolute risk of breast cancer using PRS, we have considered the procedure implemented in the iCARE package.³⁴ It requires the marginal (composite) rates for breast cancer and death, obtained here from Canada Health,^{35 36} and the relative risk distribution, obtained from the sampling at random of 10% of the individuals from the CGC with small probability weights for the breast cancer cases. We reported the results obtained using the 90% remaining (hereinafter referred as ‘validation CGC’).

In this study, woman’s genotyping information were used for computing four different published PRS: Wacholder *et al*¹⁷ (10 SNPs), Mavaddat *et al*¹⁸ (77 SNPs), Shieh *et al*¹⁹ (86 SNPs) and Evans *et al*²⁰ (18 SNPs). In the following, each PRS is referred to the name of the first author of the study. The SNPs and associated OR can be found as online supplemental table 1.

Absolute risk using a combination of BCRAT and PRS

For estimating the absolute risk of breast cancer with a combination of BCRAT and PRS (hereinafter referred as ‘combined scores’), we summed the PRS and BCRAT scores (relative hazard regression scores), and used the same procedure as described in the section ‘Absolute risk using PRS’.

Absolute risk using a combination of IBIS and PRS

As the clinical risk score obtained from the IBIS model is not an output of the software, we cannot estimate the absolute risk associated with a combination of the IBIS clinical risk score and PRS using the iCARE package in the same way we did for BCRAT (see above). In practice, the V.8.0b of the IBIS risk evaluation tool allows to compute the absolute risk by incorporating the PRS scores, but these absolute risks are different from the ones that would be obtained with the iCARE package. Keeping in mind this issue, we have used the IBIS breast cancer risk evaluation tool and incorporate the PRS scores. More precisely, and for taking into account the distribution of the PRS, we incorporated a shifted PRS that corresponds to the PRS minus the logarithm of the expected value of the relative risk associated to the PRS in our population. This latter transformation is due to the fact that the baseline hazard rate can be approximated by the composite hazard divided by the expected value of the relative risk score in the underlying population.³⁴

Statistical analysis

For comparing means between groups, we used a one-way analysis of variance test. Relationships between categorical variables were tested using the χ^2 test. Statistical significance was considered as $p < 0.05$. We plotted predictiveness curves (ie, the risk quantile against the

corresponding cumulative proportion of the population with risks below this quantile) with rug plots.

To assess the performance of the BCRAT, IBIS and PRS procedures for predicting invasive breast cancer risk, we reported calibration performance and discriminatory accuracy (see hereafter). We also reported the results obtained with the BCRAT and IBIS procedures in the validation CGC.

Calibration

We computed the expected-to-observed ratio (E/O), with the 95% CI, from the sum of the estimated risk divided by the number of observed cases. An E/O of 1 corresponds to perfect global calibration. We reported the intercept and slope estimates from logistic regression models (observed outcomes with the logit of the predicted probabilities as the independent variable).

We also compared the predicted and observed proportion of breast cancers in four absolute risk groups: $<1\%$ (low risk), $\geq 1\%$ and $<1.66\%$ (intermediate risk), $\geq 1.66\%$ and $<3\%$ (average risk), $\geq 3\%$ (high risk). The observed proportion at 5years in each risk group was calculated using a Kaplan-Meier estimator. To test the null hypothesis of a global agreement between the observed and expected values across these groups, we computed a global test statistic ($G = \sum (O_i - E_i)^2 / E_i$) where O_i and E_i are respectively the observed and expected number of events in group i , and compared this latter to the critical value from the χ^2 distribution with four df.

Discrimination

The global discrimination was assessed by the c-statistic with an inverse probability of censoring weighting (IPCW) estimation of cumulative time-dependent ROC curve with their 95% CI.^{37–39} ROC curves were plotted.

In the validation CGC, the c-indexes calculated with the BCRAT and IBIS scores were compared with those calculated with each PRS scores by using the independent and identically distributed-representation of the c-index estimators.³⁹

Sensitivity and specificity

Since the Canadian recommendation for chemoprophylaxis is a BCRAT absolute risk of breast cancer of 1.66% or higher at 5years, we calculated sensitivity and specificity using this threshold.

All statistical analyses were performed using R software, V.3.6.⁴⁰

Patient and public involvement

Patients or the public were not involved in the design, conduct, reporting or dissemination plans of this study. However, the CARTaGENE cohort received an ethical approval from thirteen ethics committees before its development and implementation. Each ethics committee includes participants and public representatives, which had the opportunity to ask questions and make recommendations.

RESULTS

Overall, 10 200 women were included for validating the BCRAT and IBIS scores and 4555 women with available genotype data were selected for the validation of the PRS scores and combined scores (figure 1). The median age was 53.1 years (quartile: 47.8–60.4) and 53.1 years (quartile 48–60.1) for the participants of CC and CGC, respectively. The median follow-up time was of 5 years in both cohorts. We observed 131 (1.28%) and 58 (1.27%) women developing a breast cancer for the CC and CGC, respectively. In total, there was 42 (0.41%) and 11 (0.24%) deaths during the 5 years follow-up, for the CC and CGC, respectively. The clinical characteristics of the two cohorts can be found in online supplemental table 2.

Breast cancer risk prediction models (BCRAT and IBIS) evaluated in the CC

Using the BCRAT model, 19.8% of women were classified into the group with an absolute risk equal or higher than 1.66% (figure 2A). There was a global agreement between the predicted and observed number of breast cancer incident cases, with an E/O of 1.01 (0.85–1.20). However, the goodness-of-fit test for the four risk groups showed a significant difference between observed and expected values ($p=0.0439$). Among the four risk groups, the E/O was significantly different from one for the average risk group (E/O: 1.51% (1.01–2.28)). There was also a slight overestimation in the high-risk group (figure 2B). This finding was in agreement with the estimate values obtained from the calibration plot with an intercept lower than zero (intercept: -1.9 (-3.4 – -0.4)) and a slope smaller than 1 (slope: 0.6 (0.2 – 0.9)). The BCRAT model had a modest discriminatory accuracy, with a c-index of 58.63 (54.05–63.21) (figure 2C). The sensitivity and specificity for the 1.66% threshold were 23.7% (16.7–31.9) and 80.3% (79.5–81), respectively.

Using the IBIS model, 18.0% of women were classified into the group with an absolute risk higher or equal to 1.66% (figure 2A). There was also a global agreement between the predicted and observed number of breast cancer incident cases, with an E/O of 1.02 (0.86–1.21). However, the goodness-of-fit test for the four risk groups showed a significant difference between observed and expected values ($p=0.0056$). The IBIS risk prediction score overestimated the number of cases in the low risk group (E/O: 2.38 (1.35–4.19)) and underestimated the number of cases in the intermediate risk group (E/O: 0.78 (0.63–0.97)), while the E/O were non-significant in the two higher risk groups (figure 2B). The intercept and slope were not significantly different from zero and one, respectively (0.4 (-1.3 – 2) and 1.1 (0.7–1.5), respectively). The IBIS model produced a slightly better discriminatory accuracy than BCRAT, with a c-index of 63.42 (59.35–67.49) ($p=0.013$) (figure 2C). The sensitivity and specificity for the 1.66% threshold were 26.7% (19.4–35.2) and 82.1% (81.3–82.8), respectively.

Breast cancer risk prediction models (BCRAT, IBIS, PRS and combined scores) evaluated in the clinicogenetic-based validation cohort

Results obtained in the validation CGC cohort that included participants with all the genetic and clinical information are reported in tables 1 and 2.

In this subcohort, BCRAT and IBIS models classified 21% and 18.5% of women into the two higher risk groups, respectively. There was a global agreement between the predicted and observed number of breast cancer cases, with an E/O of 0.94 (0.73–1.22) and 0.94 (0.73–1.22), respectively. The discriminatory accuracy of the BCRAT and IBIS models were of 59.13 (52.96–65.29) and 59.63 (53.26–66), respectively.

Using the Mavaddat, Shieh, Evans and Wacholder PRS scores, 18%, 19%, 15% and 13.5% of women were classified into the group with an absolute risk equal or higher than 1.66%, respectively (online supplemental figure 1). All the PRS scores had an E/O around 0.82, with a 95% CI including one (table 1). None of the goodness-of-fit test showed a significant departure from the null hypothesis (figure 3). The intercepts and slopes for the calibration plot were not significantly different from 0 and 1, respectively (table 1).

The PRS' c-indexes were all slightly higher than those obtained from the BCRAT and IBIS scores, Wacholder score leading to the highest c-index (64.27 (58.09–70.44)). However, none of the c-indexes was statistically different from the ones computed with the BCRAT and IBIS models (table 1). The discrimination for women at higher risk was better for the Shieh, Evans and Mavaddat PRS scores compared with BCRAT and IBIS scores (down-left corner of the ROC curves, online supplemental figure 2). Using a 1.66% threshold, all PRS scores increased both the sensitivity and the specificity as compared with the BCRAT and IBIS risk prediction score (table 1).

The distribution of the combined models' absolute risks can be found in online supplemental figure 1. All the BCRAT +PRS combined models had an E/O around 0.84, with all 95% CI including one (table 2). The goodness-of-fit test using the four risk groups showed a significant departure from the null hypothesis for the Wacholder and Evans combined models ($p=0.0475$ and $p=0.0470$, respectively) (figure 4). While the Mavaddat and Shieh combined models underestimated the number of cases in the low risk group (E/O: 0.62 (0.41–0.93) and 0.63 (0.42–0.96), respectively), the Evans and Wacholder combined models underestimated the number of cases in the intermediate risk group (E/O: 0.58 (0.39–0.85) and 0.64 (0.43–0.95), respectively). Other groups' E/O were not different from one. The Shieh combined model had an intercept and slope significantly different from 0 and 1, respectively (table 2).

The BCRAT+PRS combined models' c-indexes were all slightly higher than the BCRAT and IBIS scores, but none of them were statistically different from the ones computed with the BCRAT and IBIS models (table 2). The discrimination for women at higher risk was better

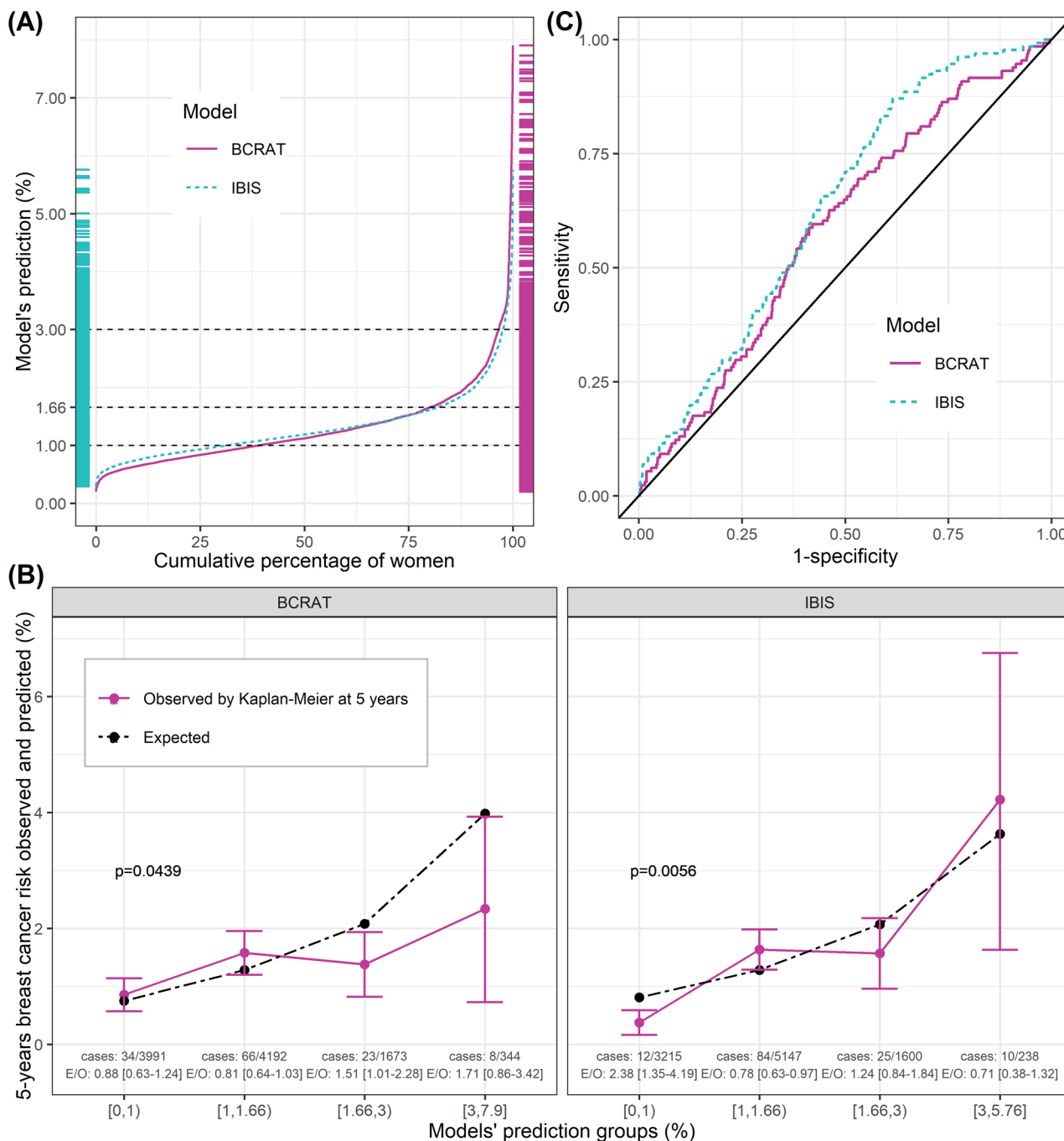


Figure 2 Absolute risk distribution and performance of the BCRAT and IBIS models in the clinical-based cohort. (A) Distribution of models' predictions as a function of cumulative percentage of women. Rug plot on the y-axis. (B) Calibration according to the models' predictions groups. P values were computed using a goodness-of-fit test statistic compared with the critical value from the χ^2 distribution. (C) Discrimination power of the models according to sensitivity and specificity. C-index was calculated using the inverse probability of censoring weighting estimation of cumulative time-dependent ROC curve. BCRAT, Breast Cancer Risk Assessment Tool; E/O, expected-to-observed cases; IBIS, International Breast Cancer Intervention Study; ROC, receiver operating characteristic.

for the Shieh and Mavaddat combined scores (down-left corner of the ROC curves, (online supplemental figure 2)). Using a 1.66% threshold, only the Evans combined model increased both the sensitivity and the specificity as compared with the BCRAT and IBIS risk prediction score (table 2).

Regarding the IBIS+PRS combined models, the E/O were the same as the BCRAT and IBIS models (0.94

(0.73–1.22)) with non-significant goodness-of-fit tests (table 2). All the combined models had an E/O that included one in each four risk groups (figure 4). Intercepts and slopes were not different from 0 and 1, respectively (table 2). The c-indexes were all slightly higher than those obtained from the BCRAT and IBIS scores, but none of them were statistically different. The discrimination for women at higher risk was also better for the Shieh

Table 1 Comparison of BCRAT, IBIS and PRS scores using the clinicogenetic-based validation cohort

	BCRAT model/IBIS model	Mavaddat	Shieh	Evans	Wacholder
E/O	0.94 (0.73 to 1.22) 0.94 (0.73 to 1.22)	0.83 (0.65 to 1.08)	0.81 (0.63 to 1.05)	0.82 (0.63 to 1.06)	0.81 (0.62 to 1.04)
Goodness of fit	P=0.0415 P=0.268	P=0.0984	P=0.1009	P=0.1992	P=0.2770
Intercept	-2 (-4.4 to 0.2) -0.8 (-3.4 to 1.8)	-0.3 (-2.4 to 1.8)	-1 (-2.5 to 0.5)	1 (-1.6 to 3.6)	0.9 (-1.8 to 3.5)
Slope	0.5 (0 to 1) 0.8 (0.2 to 1.4)	0.9 (0.4 to 1.4)	0.7 (0.4 to 1.1)	1.2 (0.6 to 1.8)	1.1 (0.5 to 1.7)
C-index	59.13 (52.96 to 65.29) 59.63 (53.26 to 66)	60.77 (53 to 68.53)	62.56 (54.54 to 70.59)	63.4 (56.65 to 70.16)	64.27 (58.09 to 70.44)
C-indexes comparison with					
BCRAT model	-	P=0.72	P=0.46	P=0.23	P=0.18
IBIS model	-	P=0.81	P=0.57	P=0.34	P=0.26
Sensitivity*	20.7% (11.2 to 33.4) 24.1% (13.9 to 37.2)	31% (19.5 to 44.5)	39.7% (27 to 53.4)	34.5% (22.5 to 48.1)	25.9% (15.3 to 39)
Specificity*	79% (77.7 to 80.3) 81.6% (80.4 to 82.8)	82.2% (81 to 83.4)	81.3% (80.1 to 82.5)	85.4% (84.2 to 86.4)	86.7% (85.6 to 87.7)

Clinicogenetic-based validation cohort: validation of the PRS models and comparison with the BCRAT and IBIS models, genotyped women with all SNPs available. Ten per cent of the cohort was used to obtain the relative risk distribution while the remained 90% was used for computing the results.

95% CIs in parenthesis

*1.66% threshold.

BCRAT, breast cancer risk assessment tool; E/O, expected-to-observed ratio; IBIS, International Breast Cancer Intervention Study; PRS, Polygenic Risk Scores; SNP, single-nucleotide polymorphism.

and Mavaddat combined scores (down-left corner of the ROC curves, online supplemental figure 2). Compared with the BCRAT and IBIS models, sensitivities values were higher while specificities values were lower (table 2).

DISCUSSION

In this work, we reported the predictive performance of BCRAT, IBIS and four PRS for predicting breast cancer occurrence within 5 years in a French-Canadian population. Results show that the BCRAT and IBIS models are globally well calibrated, with an E/O close to one. However, when focusing on predicted risk subgroups, the BCRAT model overestimates the number of cases in the average risk group (1.66%–3% risk) while the IBIS model was miscalibrated in the low-risk and intermediate-risk groups (below 1.66% risk). In our study, IBIS produced slightly better discrimination than BCRAT. As compared with the clinical-based models, the genetic prediction models (PRS) did not provide a significant improvement of the discriminative capacity. Adding PRS to the BCRAT or IBIS scores did not significantly increase the predictive power of both models.

Despite an overall good mean calibration of the BCRAT model, the calibration across risk subgroups could be

improved. The analysis of the four groups of risk shows a significant difference between expected and observed cases with an overprediction in women with a risk equal or higher than 1.66%. This finding is in accordance with previous studies.^{41–43} Opposite results have also been reported in a recent large study with pooled data from two cohorts of women where the BCRAT model underestimated the risk for values between 1.7% and 3.4%.¹² However, in this latter study, eligible women were aged between 20 and 70 years at the enrollment and recruited since 1991, while our population was aged between 40 and 70 years and enrolled since 2009. The overestimation of the BCRAT risk prediction model for women with a risk higher than 1.66% cannot be explained by differences in age-standardised incidence rates since, based on information retrieved from national cancer databases,^{35 44 45} the incidence rates are comparable between the USA and Canada (250.4 (95% CI 209.0 to 298.3) cases per 100 000 per year for Canada and 236.8 (95% CI 235.5 to 238.1) for US). The IBIS model, the PRS models and the clinicogenetic model (BCRAT+PRS) had also an overall good mean calibration. However, when analysing calibration across risk subgroups, the IBIS model had a significant goodness-of-fit test, with an over and underestimated risk

Table 2 Comparison of BCRAT, IBIS and combined scores using the clinicogenetic-based validation cohort

	Combined scores				
	BCRAT model/IBIS model with IBIS	Mavaddat with BCRAT/IBIS	Shieh with BCRAT/with IBIS	Evans with BCRAT/with IBIS	Wacholder with BCRAT/with IBIS
E/O	0.94 (0.73 to 1.22)	0.86 (0.66 to 1.11)	0.83 (0.64 to 1.07)	0.83 (0.64 to 1.08)	0.82 (0.64 to 1.06)
	0.94 (0.73 to 1.22)	0.95 (0.73 to 1.22)	0.94 (0.73 to 1.22)	0.94 (0.73 to 1.22)	0.94 (0.73 to 1.22)
Goodness of fit	P=0.0415	P=0.161	P=0.130	P=0.047	P=0.048
	P=0.268	P=0.470	P=0.519	P=0.993	P=0.627
Intercept	-2 (-4.4 to 0.2)	-1.5 (-3.3 to 0.1)	-1.6 (-3 to -0.3)	-1.2 (-3.1 to 0.6)	-1.3 (-3.2 to 0.5)
	-0.8 (-3.4 to 1.8)	-0.9 (-2.7 to 0.8)	-1.3 (-2.7 to 0)	-0.3 (-2.3 to 1.7)	-0.5 (-2.5 to 1.5)
Slope	0.5 (0 to 1)	0.6 (0.2 to 1)	0.6 (0.3 to 0.9)	0.7 (0.3 to 1.1)	0.7 (0.2 to 1.1)
	0.8 (0.2 to 1.4)	0.8 (0.4 to 1.2)	0.7 (0.3 to 1)	0.9 (0.4 to 1.4)	0.9 (0.4 to 1.3)
C-index	59.13 (52.96 to 65.29)	61.42 (54.05 to 68.78)	63.35 (55.58 to 71.12)	62.69 (55.88 to 69.50)	63.58 (57.46 to 69.69)
	59.63 (53.26 to 66)	62.73 (55.34 to 70.12)	63.83 (56.27 to 71.39)	63.35 (56.44 to 70.26)	64.21 (57.88 to 70.54)
C-indexes comparison with BCRAT model	-	P=0.50	P=0.28	P=0.12	P=0.059
	-	P=0.369	P=0.265	P=0.214	P=0.135
C-indexes comparison with IBIS model	-	P=0.66	P=0.42	P=0.38	P=0.22
	-	P=0.393	P=0.316	P=0.169	P=0.080
Sensitivity*	20.7% (11.2 to 33.4)	36.2% (24 to 49.9)	37.9% (25.5 to 51.6)	25.9% (15.3 to 39)	22.4% (12.5 to 35.3)
	24.1% (13.9 to 37.2)	36.2% (24 to 49.9)	44.8% (31.7 to 58.5)	41.4% (28.6 to 55.1)	37.9% (25.5 to 51.6)
Specificity*	79% (77.7 to 80.3)	80.5% (79.2 to 81.7)	81.5% (80.2 to 82.7)	82.1% (80.9 to 83.3)	83.8% (82.6 to 84.9)
	81.6% (80.4 to 82.8)	76.6% (75.2 to 77.9)	76.9% (75.5 to 78.2)	76.9% (75.6 to 78.2)	77.9% (76.6 to 79.2)

Combined scores: PRS scores combined with the BCRAT or IBIS scores.

Clinicogenetic-based validation cohort: validation of the PRS models and comparison with the BCRAT and IBIS models, genotyped women with all SNPs available. Ten per cent of the cohort was used to obtain the relative risk distribution while the remained 90% was used for computing the results.

95% CIs in parenthesis.

*1.66% threshold.

BCRAT, breast cancer risk assessment tool; E/O, expected-to-observed ratio; IBIS, International Breast Cancer Intervention Study; PRS, Polygenic Risk Scores; SNP, single-nucleotide polymorphism.

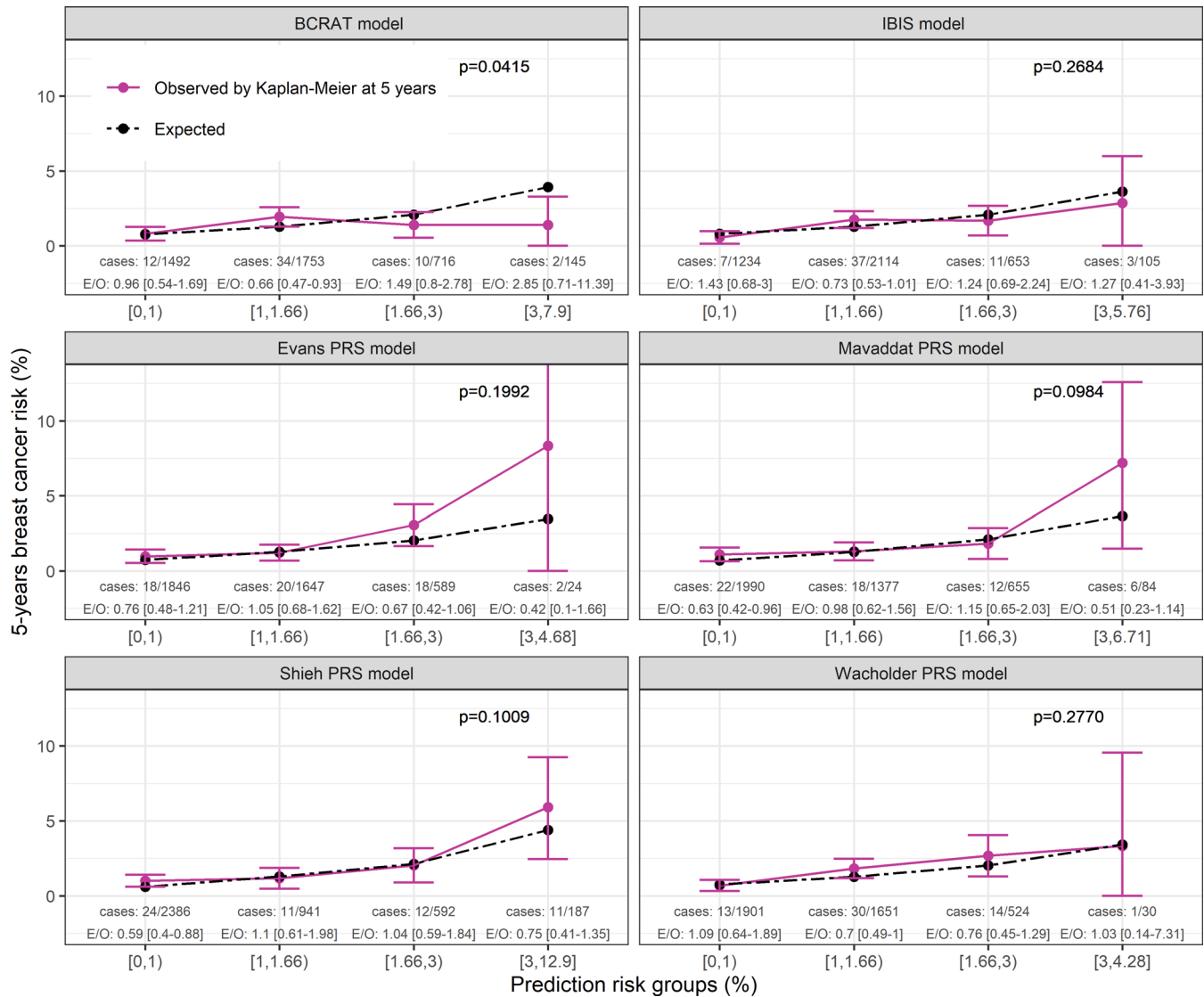


Figure 3 Calibration according to BCRAT, IBIS and PRS scores' predictions groups. Results from the clinicogenetic-based cohort. P values were computed using a goodness-of-fit test statistic compared with the critical value from the χ^2 distribution. Each PRS models name referred to the first author of the study from which the PRS were derived. BCRAT, Breast Cancer Risk Assessment Tool; E/O, expected-to-observed cases; IBIS, International Breast Cancer Intervention Study; PRS, Polygenic Risk Scores.

in the low and intermediate groups, respectively, probably explained by the UK incidence rates used by the IBIS model. This is not the case for the PRS models but this result should be cautiously interpreted in light of the reduced number of breast cancers in the genetic cohort.

The discriminatory accuracy of the BCRAT risk prediction model is modest in our population (58.6%) but is in accordance to the meta-analysis of Wang *et al*⁷ that reported a pooled AUC close to our c-index (0.60 (0.58–0.62)). The IBIS model produced a better discrimination estimate (63.4%) than BCRAT. Since we did not collect multi-generational pedigree or *BRCA1/2* gene mutations data in our cohort, the gain in discrimination for the IBIS model as compared with BCRAT model may be linked to the non-genetic risk factors. HRT use and the menopausal status, that are risk factors for the IBIS model, are significantly associated in our series with the outcome

($p < 0.05$, results not shown) and may explain the gain in discriminative accuracy. It emphasises that the inclusion of new modifiable risk factors can increase discriminatory accuracy of predictive models.

Although the calibration and discriminative power of the PRS and the clinicogenetic models were satisfactory, they did not provide a significantly better discrimination. This is not surprising since when combining SNPs the gains in prediction are usually small.¹⁵ Moreover, these non-significant results should also be interpreted in light of the modest size of our cohort having genetic information and the different baseline populations used for calculating the BCRAT, IBIS and PRSs models' relative risks. It should be noted that the combined IBIS+PRS models had a better calibration regarding the four risk groups compared with the BCRAT+PRS models. However, the absolute risk of IBIS combined models were not obtained

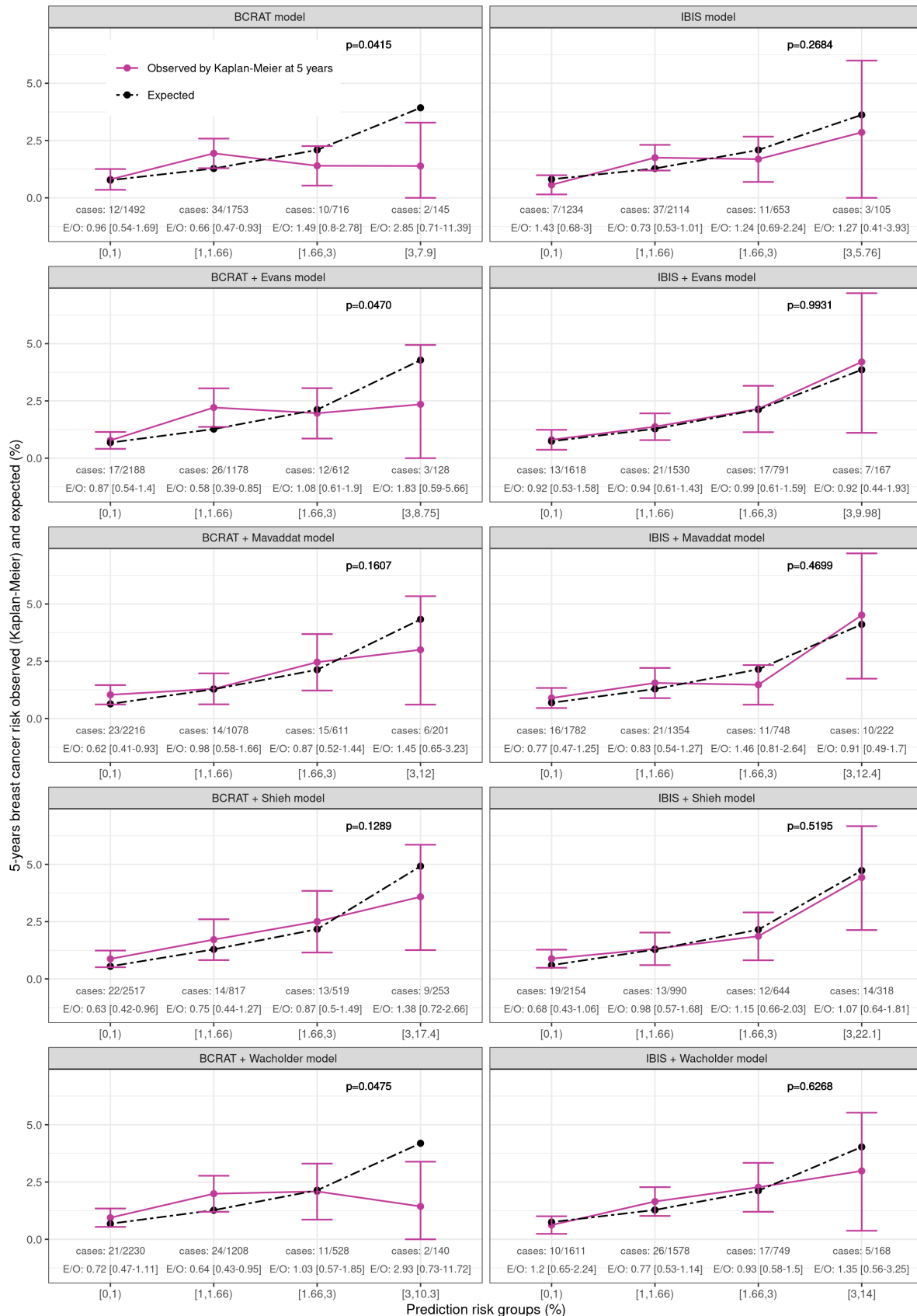


Figure 4 Calibration according to BCRAT, IBIS and combined models' predictions groups. Results from the clinicogenetic-based cohort. P values were computed using a goodness-of-fit test statistic compared with the critical value from the χ^2 distribution. Each PRS models name referred to the first author of the study from which the PRS were derived. BCRAT, Breast Cancer Risk Assessment Tool; E/O, expected-to-observed cases; IBIS, International Breast Cancer Intervention Study; PRS, Polygenic Risk Scores.

with the same procedures as for BCRAT, which makes the results not straightforward to compare. Moreover, it is worth noting that combining both clinical and genetic information in an oversimplified additive way has nevertheless some limitations from an explanatory point of view.

Some strengths of the present study should be highlighted. First, this validation study relies on the CARTaGENE cohort, which is representative of the French-Canadian urban population of middle-aged and older adults. Moreover, the linkage with administrative health databases and the Quebec Breast Cancer Registry improved the outcome quality and accuracy, and made possible to use variables usually difficult to obtain such as the history of breast biopsy or atypical hyperplasia. Second and to the best of our knowledge, this study is the first to evaluate the BCRAT in a French-Canadian population for predicting breast cancer at 5 years.

This study has nevertheless some limitations. First, our findings may not apply to younger women under forty years old. Second, we have limited our study to BCRAT and IBIS risk prediction models. The main reason was that both models were well documented and implemented. The BCRAT model is used for prevention purpose with chemoprophylaxis in the USA^{46 47} and is composed of clinical variables, easy to obtain in real clinical practice. The IBIS model is also implemented and can be used even with missing data such as multi-generational pedigree and *BRCA1/2* gene mutations data. Third, since the genotyping information was not available for all the cohort, the number of incident cases for validating the combined scores was lower than for validating BCRAT and IBIS. Moreover, the PRS, BCRAT and IBIS models had to be evaluated on different sub-cohorts. The larger decrease of IBIS's c-index compared with BCRAT between the two cohorts might be linked to the smaller size of the CGC as compared with the clinic-based cohort. The ethnicity differences between the two subcohorts could be explained by the divergent ancestry step of the quality control of genotype data. The highest breast cancer risk among genotyped women (higher age at first live birth and more relatives with breast cancer) could not be explained by the women preferentially genotyped, as they were selected for studies unrelated with breast cancers.^{30–32} Even though these two subcohorts were similar, it would be useful to collect all genotype information for the entire cohort to validate the PRS results. Finally, regarding family history included in the IBIS model, we only had maternal and paternal history of breast cancer and maternal history of ovary cancer. However, the IBIS model can handle missing data and the performance of the model remained good without this information. Therefore, the IBIS model should be more accurate with more family history variables.

CONCLUSION

BCRAT and IBIS produced overall good calibration in our French-Canadian cohort but with moderate performance in terms of discriminative ability. These results are in accordance to previous validation studies. IBIS had the better discriminatory accuracy. PRS models did not significantly improve the discrimination. Despite the modest discriminatory power of BCRAT and IBIS, these tools can be of interest for primary care physicians for delivering a personalised message to their high risk patients, regarding screening and lifestyle counselling.

Contributors RJ: conceptualisation, data curation, formal analysis, investigation, methodology, visualisation, writing-original draft, writing-review and editing. YP: data curation, software, writing-review and editing. TdM: data curation, software. CL: resources. NN: conceptualisation, resources, writing-review and editing. PB: conceptualisation, formal analysis, methodology, project administration, supervision, validation, writing-review and editing. All authors read and approved the final manuscript.

Funding The authors have not declared a specific grant for this research from any funding agency in the public, commercial or not-for-profit sectors.

Competing interests None declared.

Patient consent for publication Not required.

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement Data may be obtained from a third party and are not publicly available. The data that support the findings of this study are available from CARTaGENE but restrictions apply to the availability of these data. Data are, however, available directly from CARTaGENE (<http://cartagene.qc.ca>; access@cartagene.qc.ca; +1 514-345-2156).

Supplemental material This content has been supplied by the author(s). It has not been vetted by BMJ Publishing Group Limited (BMJ) and may not have been peer-reviewed. Any opinions or recommendations discussed are solely those of the author(s) and are not endorsed by BMJ. BMJ disclaims all liability and responsibility arising from any reliance placed on the content. Where the content includes any translated material, BMJ does not warrant the accuracy and reliability of the translations (including but not limited to local regulations, clinical guidelines, terminology, drug names and drug dosages), and is not responsible for any error and/or omissions arising from translation and adaptation or otherwise.

Open access This is an open access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited, appropriate credit is given, any changes made indicated, and the use is non-commercial. See: <http://creativecommons.org/licenses/by-nc/4.0/>.

ORCID iD

Rodolphe Jantzen <http://orcid.org/0000-0003-1356-238X>

REFERENCES

- 1 Canadian Cancer Statistics Advisory Committee. *Canadian cancer statistics 2019*. Toronto, ON: Canadian Cancer Society, 2019. cancer.ca/Canadian-Cancer-Statistics-2019-EN
- 2 Gail MH, Brinton LA, Byar DP, *et al*. Projecting individualized probabilities of developing breast cancer for white females who are being examined annually. *J Natl Cancer Inst* 1989;81:1879–86.
- 3 Tyrer J, Duffy SW, Cuzick J. A breast cancer prediction model incorporating familial and personal risk factors. *Stat Med* 2004;23:1111–30.
- 4 Costantino JP, Gail MH, Pee D, *et al*. Validation studies for models projecting the risk of invasive and total breast cancer incidence. *J Natl Cancer Inst* 1999;91:1541–8.
- 5 Gail MH, Costantino JP, Pee D, *et al*. Projecting individualized absolute invasive breast cancer risk in African American women. *J Natl Cancer Inst* 2007;99:1782–92.
- 6 Matsuno RK, Costantino JP, Ziegler RG, *et al*. Projecting individualized absolute invasive breast cancer risk in Asian



- and Pacific Islander American women. *J Natl Cancer Inst* 2011;103:951–61.
- 7 Wang X, Huang Y, Li L, *et al.* Assessment of performance of the Gail model for predicting breast cancer risk: a systematic review and meta-analysis with trial sequential analysis. *Breast Cancer Res* 2018;20:18.
 - 8 Farkas A, Vanderberg R, Merriam S, *et al.* Breast cancer chemoprevention: a practical guide for the primary care provider. *J Womens Health* 2020;29:46–56.
 - 9 Levine M, Moutquin JM, Walton R, *et al.* Chemoprevention of breast cancer. A joint guideline from the Canadian Task force on preventive health care and the Canadian breast cancer initiative's Steering Committee on clinical practice guidelines for the care and treatment of breast cancer. *CMAJ* 2001;164:1681–90.
 - 10 Canadian Cancer Society. Reducing your risk for breast cancer, 2019. Available: <http://www.cancer.ca/en/cancer-information/cancer-type/breast/risks/reducing-your-risk/?region=qc>
 - 11 Voruganti T, O'Brien MA, Straus SE, *et al.* Primary care physicians' perspectives on computer-based health risk assessment tools for chronic diseases: a mixed methods study. *J Innov Health Inform* 2015;22:333–9.
 - 12 Terry MB, Liao Y, Whittmore AS, *et al.* 10-Year performance of four models of breast cancer risk: a validation study. *Lancet Oncol* 2019;20:504–17.
 - 13 Gail MH. Performance of BCRAT in high-risk patients with breast cancer. *Lancet Oncol* 2019;20:e285.
 - 14 Terry MB, Liao Y, Hopper JL, *et al.* Performance of BCRAT in high-risk patients with breast cancer - Authors' reply. *Lancet Oncol* 2019;20:e286.
 - 15 Gail MH. Discriminatory accuracy from single-nucleotide polymorphisms in models to predict breast cancer risk. *J Natl Cancer Inst* 2008;100:1037–41.
 - 16 Dite GS, MacInnis RJ, Bickerstaffe A, *et al.* Breast cancer risk prediction using clinical models and 77 independent risk-associated SNPs for women aged under 50 years: Australian breast cancer family registry. *Cancer Epidemiol Biomarkers Prev* 2016;25:359–65.
 - 17 Wacholder S, Hartge P, Prentice R, *et al.* Performance of common genetic variants in breast-cancer risk models. *N Engl J Med* 2010;362:986–93.
 - 18 Mavaddat N, Pharoah PDP, Michailidou K, *et al.* Prediction of breast cancer risk based on profiling with common genetic variants. *J Natl Cancer Inst* 2015;107. doi:10.1093/jnci/djv036. [Epub ahead of print: 08 04 2015].
 - 19 Shieh Y, Hu D, Ma L, *et al.* Breast cancer risk prediction using a clinical risk model and polygenic risk score. *Breast Cancer Res Treat* 2016;159:513–25.
 - 20 Evans DG, Brentnall A, Byers H, *et al.* The impact of a panel of 18 SNPs on breast cancer risk in women attending a UK familial screening clinic: a case-control study. *J Med Genet* 2017;54:111–3.
 - 21 MacArthur J, Bowler E, Cerezo M, *et al.* The new NHGRI-EBI catalog of published genome-wide association studies (GWAS catalog). *Nucleic Acids Res* 2017;45:D896.
 - 22 Gouvernement du Québec. Québec breast cancer screening program, 2018. Available: <https://www.quebec.ca/en/health/advice-and-prevention/screening-and-carrier-testing-offer/quebec-breast-cancer-screening-program/>
 - 23 Shields M, Wilkins K. An update on mammography use in Canada. *Health Rep* 2009;20:7–19.
 - 24 Bernatsky S, Clarke A, Ramsey-Goldman R, *et al.* Hormonal exposures and breast cancer in a sample of women with systemic lupus erythematosus. *Rheumatology* 2004;43:1178–81.
 - 25 Casals F, Hodgkinson A, Hussin J, *et al.* Whole-Exome sequencing reveals a rapid change in the frequency of rare functional variants in a founding population of humans. *PLoS Genet* 2013;9:e1003815.
 - 26 Scriver CR. Human genetics: lessons from Quebec population. *Annu Rev Genom Hum Genet* 2001;2:69–101.
 - 27 Awadalla P, Boileau C, Payette Y, *et al.* Cohort profile of the CARTaGENE study: Quebec's population-based Biobank for public health and personalized genomics. *Int J Epidemiol* 2013;42:1285–99.
 - 28 Th  berge I. Institut national de sant   publique Du Qu  bec, direction syst  mes de soins et services. validation de strat  gies pour obtenir Le taux de d  tection Du cancer, La valeur pr  dictive positive, La proportion des cancers in situ, La proportion des cancers infiltrants de petite taille et La proportion des cancers infiltrants SANS envahissement ganglionnaire dans Le cadre des donn  es fournies PAR Le programme qu  b  cois de d  pistage Du cancer Du sein (PQDCS). Montr  al: Direction des syst  mes de soins et services, Institut national de sant   publique, 2003.
 - 29 Tonelli M, Wiebe N, Fortin M, *et al.* Methods for identifying 30 chronic conditions: application to administrative data. *BMC Med Inform Decis Mak* 2015;15:31.
 - 30 Akcimen F, Ross JP, Sarayloo F, *et al.* Genetic and epidemiological characterization of restless legs syndrome in Qu  bec. *Sleep* 2020;43:zsz265.
 - 31 Hodgkinson A, Idaghdour Y, Gbeha E, *et al.* High-Resolution genomic analysis of human mitochondrial RNA sequence variation. *Science* 2014;344:413–5.
 - 32 Hussin JG, Hodgkinson A, Idaghdour Y, *et al.* Recombination affects accumulation of damaging and disease-associated mutations in human populations. *Nat Genet* 2015;47:400–4.
 - 33 Zhang F. BCRA: breast cancer risk assessment, 2018. Available: <https://CRAN.R-project.org/package=BCRA>
 - 34 Pal Choudhury P, Maas P, Wilcox A, *et al.* iCARE: an R package to build, validate and apply absolute risk models. *PLoS One* 2020;15:e0228198.
 - 35 Statistics Canada. Table 13-10-0111-01 number and rates of new cases of primary cancer, by cancer type, age group and sex, 2019. Available: <https://doi.org/10.25318/1310011101-eng>
 - 36 Statistics Canada. Table 13-10-0392-01 deaths and age-specific mortality rates by selected grouped causes, 2019. Available: <https://doi.org/10.25318/1310039201-eng>
 - 37 Uno H, Cai T, Tian L, *et al.* Evaluating prediction rules for *t*-Year survivors with censored regression models. *J Am Stat Assoc* 2007;102:527–37.
 - 38 Blanche P, Latouche A, Viallon V. *Time-Dependent AUC with right-censored data: a survey study*. New York, NY: Springer, 2012: 239–51. <http://arxiv.org/abs/1210.6805>
 - 39 Blanche P, Dartigues J-F, Jacqmin-Gadda H. Estimating and comparing time-dependent areas under receiver operating characteristic curves for censored event times with competing risks. *Stat Med* 2013;32:5381–97.
 - 40 R Core Team. *R: a language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing, 2020. <https://www.R-project.org/>
 - 41 Schonberg MA, Li VW, Eliassen AH, *et al.* Performance of the breast cancer risk assessment tool among women age 75 years and older. *J Natl Cancer Inst* 2016;108. doi:10.1093/jnci/djv348. [Epub ahead of print: 30 11 2015].
 - 42 Nickson C, Procopio P, Velentzis LS, *et al.* Prospective validation of the NCI breast cancer risk assessment tool (Gail model) on 40,000 Australian women. *Breast Cancer Res* 2018;20:155.
 - 43 Glynn RJ, Colditz GA, Tamimi RM, *et al.* Comparison of questionnaire-based breast cancer prediction models in the nurses' health study. *Cancer Epidemiol Biomarkers Prev* 2019;28:1187–94.
 - 44 Statistics Canada. Table 17-10-0005-01 population estimates on July 1st, by age and sex, 2021. Available: <http://doi.org/10.25318/1710000501-eng>
 - 45 U.S. Cancer Statistics Working Group. U.S. cancer statistics data Visualizations tool, based on November 2018 submission data (1999–2016): U.S. department of health and human services, centers for disease control and prevention and National cancer Institute, 2019. Available: www.cdc.gov/cancer/dataviz
 - 46 Moyer VA. Medications for risk reduction of primary breast cancer in women: U.S. preventive services Task force recommendation statement. *Ann Intern Med* 2013;159:698–708.
 - 47 Louro J, Posso M, Hilton Boon M, *et al.* A systematic review and quality assessment of individualised breast cancer risk prediction models. *Br J Cancer* 2019;121:76–85.