BMJ Open

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

Journal:	*BMJ Open*
Manuscript ID	bmjopen-2020-045078
Article Type:	Original research
Date Submitted by the Author:	21-Sep-2020
Complete List of Authors:	Jantzen, Rodolphe; CHU Sainte-Justine, Centre de recherche Payette, Yves; CHU Sainte-Justine Centre de Recherche de Malliard, Thibault; CHU Sainte-Justine Centre de Recherche Labbé, Catherine; CHU Sainte-Justine Centre de Recherche Noisel, Nolwenn; CHU Sainte-Justine Centre de Recherche Broët, Philippe; CHU Sainte-Justine Centre de Recherche
Keywords:	Breast tumours < ONCOLOGY, EPIDEMIOLOGY, GENETICS, PREVENTIVE MEDICINE, PUBLIC HEALTH

I, the Submitting Author has the right to grant and does grant on behalf of all authors of the Work (as defined in the below author licence), an exclusive licence and/or a non-exclusive licence for contributions from authors who are: i) UK Crown employees; ii) where BMJ has agreed a CC-BY licence shall apply, and/or iii) in accordance with the terms applicable for US Federal Government officers or employees acting as part of their official duties; on a worldwide, perpetual, irrevocable, royalty-free basis to BMJ Publishing Group Ltd ("BMJ") its licensees and where the relevant Journal is co-owned by BMJ to the co-owners of the Journal, to publish the Work in this journal and any other BMJ products and to exploit all rights, as set out in our licence.

The Submitting Author accepts and understands that any supply made under these terms is made by BMJ to the Submitting Author unless you are acting as an employee on behalf of your employer or a postgraduate student of an affiliated institution which is paying any applicable article publishing charge ("APC") for Open Access articles. Where the Submitting Author wishes to make the Work available on an Open Access basis (and intends to pay the relevant APC), the terms of reuse of such Open Access shall be governed by a Creative Commons licence – details of these licences and which Creative Commons licence will apply to this Work are set out in our licence referred to above.

Other than as permitted in any relevant BMJ Author's Self Archiving Policies, I confirm this Work has not been accepted for publication elsewhere, is not being considered for publication elsewhere and does not duplicate material already published. I confirm all authors consent to publication of this Work and authorise the granting of this licence.

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,*]

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

[2]Université de Montréal, Montréal, Québec, Canada

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

[4]Assistance Publique-Hôpitaux de Paris (AP-HP), Hôpitaux Universitaires Paris-Sud, Hôpital Paul Brousse, 12 Avenue Paul Vaillant Couturier, 94807 Villejuif, France

*** Correspondence:**

Rodolphe Jantzen, rodolphe.jantzen@gmail.com

Philippe Broët, philippe.broet@umontreal.ca

Keywords: Breast cancer, Validation Study, Clinical Decision Rules, Polygenic risk score, BCRAT, IBIS

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) to predict the occurrence of invasive breast cancers at five 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 recruited during two phases (2009-2010 and 2013-2014).

Participants were randomly selected to be broadly representative of the population recorded on the Quebec administrative health insurance registries.

Participants: 10,200 women were included for validating BCRAT and IBIS and 4,555 with clinical and genetic information for validating the PRS and combined scores.

Outcome measures: We computed the absolute risks of breast cancer at five 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 five 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 expected-to-observed ratio of 1.01 [0.85-1.19] and 1.02 [0.86-1.21]. IBIS' c-index was significantly higher than BCRAT (63.42 [59.35-67.49] versus 58.63 [54.05-63.21], p=0.013). All the PRS scores had a global calibration around 0.82, with a confidence interval 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 (BCRAT+PRS) did not improve the results.

Conclusions: In this French-Canadian population-based cohort, BCRAT and IBIS are globally well calibrated but with modest discriminatory accuracy. Despite this modest discriminatory power, these tools can be of interest for primary care physicians for delivering a personalized message to their high risk patients, regarding screening and lifestyle counseling.

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 forty years old.

- Since the genotyping information was not available for all the cohort, the models had to be evaluated on two different sub-cohorts.

1 Introduction

Breast cancer is the most frequently diagnosed cancer and the second leading cause 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 BCRAT was developed by Dr. Mitchell Gail in 1989 using 5,998 American women from a case-control study [2]. 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 (e.g., 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 U.S. Food and Drug Administration recommended to consider chemo-prevention for women at high risk of breast cancer (i.e. a 5-year risk equal or higher than 1.66%), while the U.S. Preventive Services Task Force recommended chemo-prevention 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 U.S.-centered use led to an under-use 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 International Breast Cancer Intervention Study model (IBIS, also known as the Tyrer-Cuzick model) is also a widely used breast cancer risk prediction model, which takes into account multi-generational 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 the probability of carrying a *BRCA1* or *BRCA2* mutation, as well as 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 to 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 multi-generational pedigree information and *BRCA1/2* gene mutations. The IBIS model is a hybrid model combining a segregation model for familial risk together with a classical Cox model for non-genetic risk factors. The segregation model estimates the risk due to genetic factors conditional on woman's multi-generational family history of breast and ovarian cancer, and the results of tests for *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* gene mutation [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 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 ROC curve (AUC) increasing from 0.607 to 0.632 [15]. This kind of clinico-genetic 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 upon 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 (GWAS) [21]. The predictive accuracy of these PRSs compared to classical prediction models, such as the BCRAT and IBIS, should now be evaluated in various populations.

In Quebec, the Breast Cancer Screening Program consists of a mammogram every two years for women aged 50 to 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 organized 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 multi-countries cohorts. The fact that BCRAT and IBIS have not been evaluated in the French-Canadian population, which has specific genetic patterns, as compared to the general European population [25,26], with lifestyle risk factors (e.g., 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 polygenic risk scores to predict the occurrence of invasive breast cancers at five years in middle-aged and older French-Canadian women.

2 Materials and methods

2.1 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, we used two administrative health databases, the Quebec Health Insurance Board (RAMQ) and the Quebec Breast Cancer Registry (see Supplementary Methods), and an algorithm based on a previous report from the *Institut National de Santé Publique du Québec* (INSPQ) [28] and the Tonelli *et al.* algorithm [29]. Using the Breast Cancer Registry, we retrieved the incidence date of histologically confirmed breast cancers. Then, we selected all women having an abnormal mammography and retrieved, when available, the incidence date after the abnormal mammography from the RAMQ database for women with at least two claims in two years or one hospitalization with the appropriate International Classification of Diseases (ICD), Ninth or Tenth Revision codes (174 and C50). 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. Recruitment was unrelated to the last mammography screening. The validation of the BCRAT and IBIS models was done on the sub-cohort 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 sub-cohort of 4,555 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.

2.2 Genetic data

Only a fraction of the population cohort has been genotyped (n=12,062). These participants were selected to be genotyped through various scientific projects unrelated to breast cancer [30–32]. Single-nucleotide polymorphism (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 the Supplementary Methods.

2.3 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 five-years study period (administrative censoring) or at death.

2.4 Predictive scores

2.4.1 Absolute risk using the BCRAT and the IBIS models

The BCRAT and IBIS risk scores are calculated using baseline hazard rates calculated from the marginal hazard rates, and attributable hazard function estimates obtained from the United States population data (BCRAT) and the United Kingdom/Swedish population data (IBIS). In this article, the BCRAT and IBIS absolute risks of breast cancer at five years were calculated for each woman at the inclusion date using the National Institutes of Health R package "BCRA", version 2.1 [33] and the latest version of the "IBIS Breast Cancer Risk Evaluation Tool" (http://www.ems-trials.org/riskevaluator/ — version 8.0b, September 2017), respectively.

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

2.4.2 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 risk score distribution, obtained from the sampling at random of 10% of the individuals from the clinicogenetic-based cohort with small probability weights for the breast cancer cases. To avoid the optimism bias, 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 odds ratio can be found as Supplementary Table S1.

2.4.3 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".

As the hazard function obtained from the IBIS model is not an output of the software, we cannot combine the IBIS and PRS information in this work.

2.5 Statistical analysis

For comparing means between groups, we used a one-way ANOVA test. Relationships between categorical variables were tested using the χ^2 test. Statistical significance was considered as P-values less than 0.05. We plotted predictiveness curves (i.e., 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.

2.5.1 Calibration

We computed the expected-to-observed ratio (E/O), with the 95% confidence interval (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. For the few women with less than five years follow-up, their risk contributions were proportional to the follow-up time. 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), ≥1% and <1.66% (intermediate risk), ≥1.66% and <3% (average risk), ≥3% (high risk). The observed proportion at five-year 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 goodness of fit test statistic and compared this latter to the critical value from the chi-squared distribution with four degrees of freedom.

2.5.2 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]. Receiver operating curves (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].

2.5.3 Sensitivity and specificity

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

All statistical analyses were performed using R software, version 3.6 [40].

2.6 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.

3 Results

Overall, 10,200 women were included for validating the BCRAT and IBIS scores and 4,555 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 five-year follow-up, for the CC and CGC, respectively. The clinical characteristics of the two cohorts can be found in the Supplementary Table S2.

3.1 Breast cancer risk prediction models (BCRAT and IBIS) evaluated in the clinical-based validation cohort

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 [0.85-1.19]. 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.

3.2 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 sub-cohort, 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 expected/observed ratio 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 (Supplementary Figure S1). 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 zero and one, 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 to BCRAT and IBIS scores (down-left corner of the ROC curves, Supplementary Figure S2). Using a 1.66% threshold, all PRS scores increased both the sensitivity and the specificity as compared to the BCRAT and IBIS risk prediction score (Table 1).

All the combined models (BCRAT + PRS) 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=0.0478 and p=0.0471, 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 zero and one, respectively (Table 2).

The 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 1). The discrimination for women at higher risk was better for the Shieh and Mavaddat combined scores (down-left corner of the ROC curves, Supplementary Figure S2). Using a 1.66% threshold, only the Evans combined model increased both the sensitivity and the specificity as compared to the BCRAT and IBIS risk prediction score (Table 2).

4 Discussion

In this work, we reported the predictive performance of BCRAT, IBIS and four polygenic risk scores for predicting breast cancer occurrence within five years in a French-Canadian population. Results show that the BCRAT and IBIS models are globally well calibrated. 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 and intermediate risk groups (below 1.66% risk). In our study, IBIS produced slightly better discrimination than BCRAT. As compared to the clinical-based models, the genetic prediction models (PRS) did not provide a significant improvement of the discriminative capacity. Adding PRS to the BCRAT scores did not significantly increase the predictive power of BCRAT.

Despite an overall good calibration of the BCRAT model, the analysis of the four groups of risk shows a significant difference between expected and observed cases with an over-prediction 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, the prediction horizon was at 10 years, 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-standardized incidence rates since, based on information retrieved from national cancer databases [35,44,45], the incidence rates are comparable between the United States and Canada (250.4 [95%CI 209.0-298.3] cases per 100,000 per year for Canada and 236.8 [95%CI 235.5-238.1] for US). The IBIS model, the PRS models and the clinico-genetic model (BCRAT+PRS) had also an overall good calibration. However, the IBIS over and underestimated the risk in the low and intermediate groups, respectively. 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 to 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 emphasizes that the inclusion of new modifiable risk factors can increase discriminatory accuracy of predictive models. The PRS and the clinico-genetic model 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.

Some strengths of the present study should be highlighted. Firstly, 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. Secondly and to the best of our knowledge, this study is the first to evaluate the breast cancer risk assessment tools in a French-Canadian population for predicting breast cancer at five years.

This study has nevertheless some limitations. Firstly, our findings may not apply to younger women under forty years old. Secondly, 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 chemo-prophylaxis in the US [46,47] and its risk score 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. Thirdly, since the genotyping information was not available for all the cohort, the PRS, BCRAT and IBIS models had to be evaluated on different sub-cohorts. The ethnicity differences between the two sub-cohorts 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 sub-cohorts were similar, it would be useful to collect all genotype information for the entire cohort to validate the PRS results.

4.1 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 personalized message to their high risk patients, regarding screening and lifestyle counseling.

5 Figures

Figure 1 Flow-chart

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 to the critical value from the chi-squared distribution. E/O: expected-to-observed cases. (C) Discrimination power of the models according to sensitivity and specificity. C-index was calculated using the Inverse Probability of Censoring Weighting (IPCW) estimation of cumulative time-dependent ROC curve.

Figure 3 **Calibration according to BCRAT, IBIS and PRS scores' predictions groups.** PRS: polygenic risk score. E/O: expected-to-observed cases. Results from the clinicogenetic-based cohort. P values were computed using a goodness of fit test statistic compared to the critical value from the chi-squared distribution. Each PRS models name referred to the first author of the study from which the PRS were derived.

Figure 4 **Calibration according to BCRAT, IBIS and combined models' predictions groups.** PRS: polygenic risk score. E/O: expected-to-observed cases. Results from the clinicogenetic-based cohort. P values were computed using a goodness of fit test statistic compared to the critical value from the chi-squared distribution. Each PRS models name referred to the first author of the study from which the PRS were derived.

6 Tables

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

	BCRAT model / IBIS model	Mavaddat	Shieh	Evans	Wacholder
E/O	0.94 [0.73-1.22] 0.94 [0.73-1.22]	0.83 [0.64-1.08]	0.81 [0.63-1.05]	0.82 [0.63-1.06]	0.81 [0.62-1.04]
Goodness of fit	p=0.0415 p=268	p=0.0984	p=0.1009	p=0.1992	p=0.2770
Intercept	-2 [-4.4 - 0.2] -0.8 [-3.4 - 1.8]	- 0.3 [-2.4 - 1.8]	-1 [-2.5 - 0.5]	1 [-1.6 - 3.6]	0.9 [-1.8 - 3.5]
Slope	0.5 [0 - 1] 0.8 [0.2 - 1.4]	0.9 [0.4 - 1.4]	0.7 [0.4 - 1.1]	1.2 [0.6 - 1.8]	1.1 [0.5 - 1.7]
C-index	59.13 [52.96-65.29] 59.63 [53.26-66]	60.77 [53-68.53]	62.56 [54.54-70.59]	63.4 [56.65-70.16]	64.27 [58.09-70.44]
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% [11.2-33.4] 24.1% [13.9-37.2]	31% [19.5-44.5]	39.7% [27-53.4]	34.5% [22.5-48.1]	25.9% [15.3-39]
Specificity *	79% [77.7-80.3] 81.6% [80.4-82.8]	82.2% [81-83.4]	81.3% [80.1-82.5]	85.4% [84.2-86.4]	86.7% [85.6-87.7]

BCRAT: Breast Cancer Risk Assessment Tool; E/O: expected-to-observed ratio; IBIS: International Breast Cancer Intervention Study risk evaluation tool; NRI: Net Reclassification Index

Clinicogenetic-based cohort: validation of the PRS models and comparison with the BCRAT and IBIS models, genotyped women with all SNPs available. 10% of the cohort was used to obtain the risk score distribution while the remained 90% was used for computing the results.

95% confidence intervals in square brackets

* 1.66% threshold

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

	BCRAT model / IBIS model	**Combined Mavaddat**	**Combined Shieh**	**Combined Evans**	**Combined Wacholder**
E/O	0.94 [0.73-1.22] 0.94 [0.73-1.22]	0.86 [0.66-1.11]	0.83 [0.64-1.07]	0.83 [0.64-1.07]	0.82 [0.64-1.06]
Goodness of fit	p=0.0415 p=0.268	p=0.161	p=0.13	p=0.047	p=0.0475
Intercept	-2 [-4.4 - 0.2] -0.8 [-3.4 - 1.8]	- 1.5 [-3.3 - 0.1]	-1.6 [-3 - -0.3]	-1.2 [-3.1 - 0.6]	-1.3 [-3.2 - 0.5]
Slope	0.5 [0 - 1] 0.8 [0.2 - 1.4]	0.6 [0.2 - 1]	0.6 [0.3 - 0.9]	0.7 [0.3 - 1.1]	0.7 [0.2 - 1.1]
C-index	59.13 [52.96-65.29] 59.63 [53.26-66]	61.42 [54.05-68.78]	63.35 [55.58-71.12]	62.69 [55.88-69.50]	63.58 [57.46-69.69]
C-indexes comparison with:					
BCRAT model	-	p=0.50	p=0.28	p=0.12	p=0.059

IBIS model	-	p=0.66	p=0.42	p=0.38	p=0.22
Sensitivity *	20.7% [11.2-33.4]	36.2% [24-49.9]	37.9% [25.5-51.6]	25.9% [15.3-39]	22.4% [12.5-35.3]
	24.1% [13.9-37.2]				
Specificity *	79% [77.7-80.3]	80.5% [79.2-81.7]	81.5% [80.2-82.7]	82.1% [80.9-83.3]	83.8% [82.6-84.9]
	81.6% [80.4-82.8]				

BCRAT: Breast Cancer Risk Assessment Tool; E/O: expected-to-observed ratio; IBIS: International Breast Cancer Intervention Study risk evaluation tool; NRI: Net Reclassification Index

Combined scores: PRS scores combined with the BCRAT scores.

Clinicogenetic-based cohort: validation of the PRS models and comparison with the BCRAT and IBIS models, genotyped women with all SNPs available. 10% of the cohort was used to obtain the risk score distribution while the remained 90% was used for computing the results.

95% confidence intervals in square brackets

* 1.66% threshold

7 References

1 Canadian Cancer Statistics Advisory Committee. Canadian Cancer Statistics 2019. *Toronto, ON: Canadian Cancer Society* Published Online First: September 2019.cancer.ca/Canadian-Cancer-Statistics-2019-EN (accessed 13 Nov 2019).

2 Gail MH, Brinton LA, Byar DP, *et al.* Projecting Individualized Probabilities of Developing Breast Cancer for White Females Who Are Being Examined Annually. *JNCI Journal of the National Cancer Institute* 1989;**81**:1879–86. doi:10/djr43s

3 Tyrer J, Duffy SW, Cuzick J. A breast cancer prediction model incorporating familial and personal risk factors. *Statist Med* 2004;**23**:1111–30. doi:10.1002/sim.1668

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. doi:10/ckg9vc

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. doi:10/btbjf9

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. doi:10/dkmv6x

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**. doi:10/gc83rw

8 Farkas A, Vanderberg R, Merriam S, *et al.* Breast Cancer Chemoprevention: A Practical Guide for the Primary Care Provider. *Journal of Women's Health* 2020;**29**:46–56. doi:10.1089/jwh.2018.7643

9 Levine M, Moutquin J-M, 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. www.cancer.ca. http://www.cancer.ca/en/cancer-information/cancer-type/breast/risks/reducing-your-risk/?region=qc (accessed 28 Nov 2019).

11 Voruganti RT, 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. *BMJ Health & Care Informatics* 2015;**22**:333–9. doi:10/ggt2c8

12 Terry MB, Liao Y, Whittemore AS, *et al.* 10-year performance of four models of breast cancer risk: a validation study. *The Lancet Oncology* 2019;**20**:504–17. doi:10/ggt2dp

13 Gail MH. Performance of BCRAT in high-risk patients with breast cancer. *The Lancet Oncology* 2019;**20**:e285. doi:10.1016/S1470-2045(19)30301-8

14 Terry MB, Liao Y, Hopper JL, *et al.* Performance of BCRAT in high-risk patients with breast cancer – Authors' reply. *The Lancet Oncology* 2019;**20**:e286. doi:10.1016/S1470-2045(19)30311-0

15 Gail MH. Discriminatory Accuracy From Single-Nucleotide Polymorphisms in Models to Predict Breast Cancer Risk. *J Natl Cancer Inst* 2008;**100**:1037–41. doi:10/ff2x2r

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. doi:10.1158/1055-9965.EPI-15-0838

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. doi:10/fn7zxw

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/f3nv95

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. doi:10/ggt2dm

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. doi:10/f9nd2w

21 MacArthur J, Bowler E, Cerezo M, *et al.* The new NHGRI-EBI Catalog of published genome-wide association studies (GWAS Catalog). *Nucleic Acids Research* 2017;**45**:D896. doi:10/f9v7cp

22 Gouvernement du Québec. Québec Breast Cancer Screening Program. 2018.https://www.quebec.ca/en/health/advice-and-prevention/screening-and-carrier-testing-offer/quebec-breast-cancer-screening-program/ (accessed 9 Mar 2020).

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 (Oxford)* 2004;**43**:1178–81. doi:10/bq2cx6

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. doi:10.1371/journal.pgen.1003815

26 Scriver CR. Human Genetics: Lessons from Quebec Population. *Annu Rev Genom Hum Genet* 2001;**2**:69–101. doi:10.1146/annurev.genom.2.1.69

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. *International Journal of Epidemiology* 2013;**42**:1285–99. doi:10.1093/ije/dys160

28 Théberge I, Institut national de santé publique du Québec, Direction systèmes de soins et services. *Validation de stratégies pour obtenir le taux de détection du cancer, la valeur prédictive 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 données fournies par le programme québécois de dépistage du cancer du sein (PQDCS)*. Montréal: : Direction des systèmes de soins et services, Institut national de santé publique 2003.

29 Tonelli M, Wiebe N, Fortin M, *et al.* Methods for identifying 30 chronic conditions: application to administrative data. *BMC Medical Informatics and Decision Making* 2015;**15**:31. doi:10/f698hd

30 Akçimen F, Ross JP, Sarayloo F, *et al.* Genetic and epidemiological characterization of restless legs syndrome in Québec. *Sleep* 2019;**43**. doi:10.1093/sleep/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. doi:10/ggt2b8

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. doi:10/ggrr3s

33 Zhang F. *BCRA: breast cancer risk assessment*. 2018. https://CRAN.R-project.org/package=BCRA

34 Choudhury PP, Maas P, Wilcox A, *et al.* iCARE: An R package to build, validate and apply absolute risk models. *PLOS ONE* 2020;**15**:e0228198. doi:10/ggt2ch

35 Statistics Canada. Table 13-10-0111-01 Number and rates of new cases of primary cancer, by cancer type, age group and sex. doi.org/10.25318/1310011101-eng (accessed 10 Nov 2019).

36 Statistics Canada. Table 13-10-0392-01 Deaths and age-specific mortality rates, by selected grouped causes. https://doi.org/10.25318/1310039201-eng (accessed 10 Nov 2019).

37 Uno H, Cai T, Tian L, *et al.* Evaluating Prediction Rules for t-Year Survivors with Censored Regression Models. *Journal of the American Statistical Association* 2007;**102**:527–37. doi:10.1198/016214507000000149

38 Blanche P, Latouche A, Viallon V. Time-dependent AUC with right-censored data: a survey study. *arXiv:12106805 [statME]* Published Online First: 25 October 2012.http://arxiv.org/abs/1210.6805 (accessed 26 Sep 2017).

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. *Statist Med* 2013;**32**:5381–97. doi:10.1002/sim.5958

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. *JNCIJ* 2016;**108**:djv348. doi:10/ggt2db

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. doi:10/ggt2dc

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. doi:10/ggt2c4

44 Statistics Canada. Table 17-10-0005-01 Population estimates on July 1st, by age and sex. doi: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.www.cdc.gov/cancer/dataviz (accessed 9 Oct 2019).

46 Moyer VA. Medications for Risk Reduction of Primary Breast Cancer in Women: U.S. Preventive Services Task Force Recommendation Statement. *Ann Intern Med* Published Online First: 24 September 2013. doi:10/ggt2cx

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. doi:10/ggt2c3

8 Funding

This research received no specific grant from any funding agency in the public, commercial or not-for-profit sectors.

9 Competing Interests

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

10 Author Contributions

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

11 Data Availability Statement

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).

12 Ethics approval and consent to participate

This project has been approved by the Research Ethics Board of the CHU Sainte-Justine under the reference 2020-2427. In addition, CARTaGENE has obtained ethics approval by the CHU Sainte-Justine under the reference: MP-21-2011-345, 3297. The latest annual ethics renewal was granted on September 13, 2019. This latter approval implies that all participants have given their consent (cartagene.qc.ca/sites/default/files/documents/consent/brochure_en_0505_0.pdf). Consent was obtained from all the participants.

13 Acknowledgments

We would like to thank all the CARTaGENE participants for their generous investments in health research. We would also like to thank the RAMQ and the Commission d'accès à l'information (CAI) for their support in obtaining the data.

14 Supplementary Material

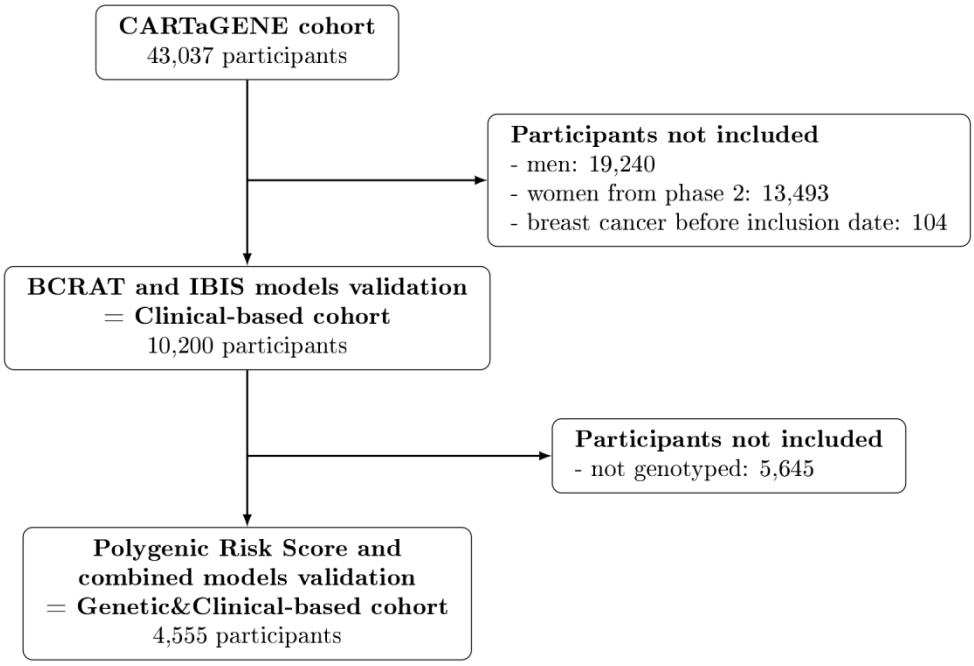
Supplementary Methods

Supplementary Table S1: SNPs used for each extended model and the associated gene and odds ratio.

Supplementary Table S2: Characteristics comparison of the women from the Clinical-based and the clinicogenetic-based cohorts.

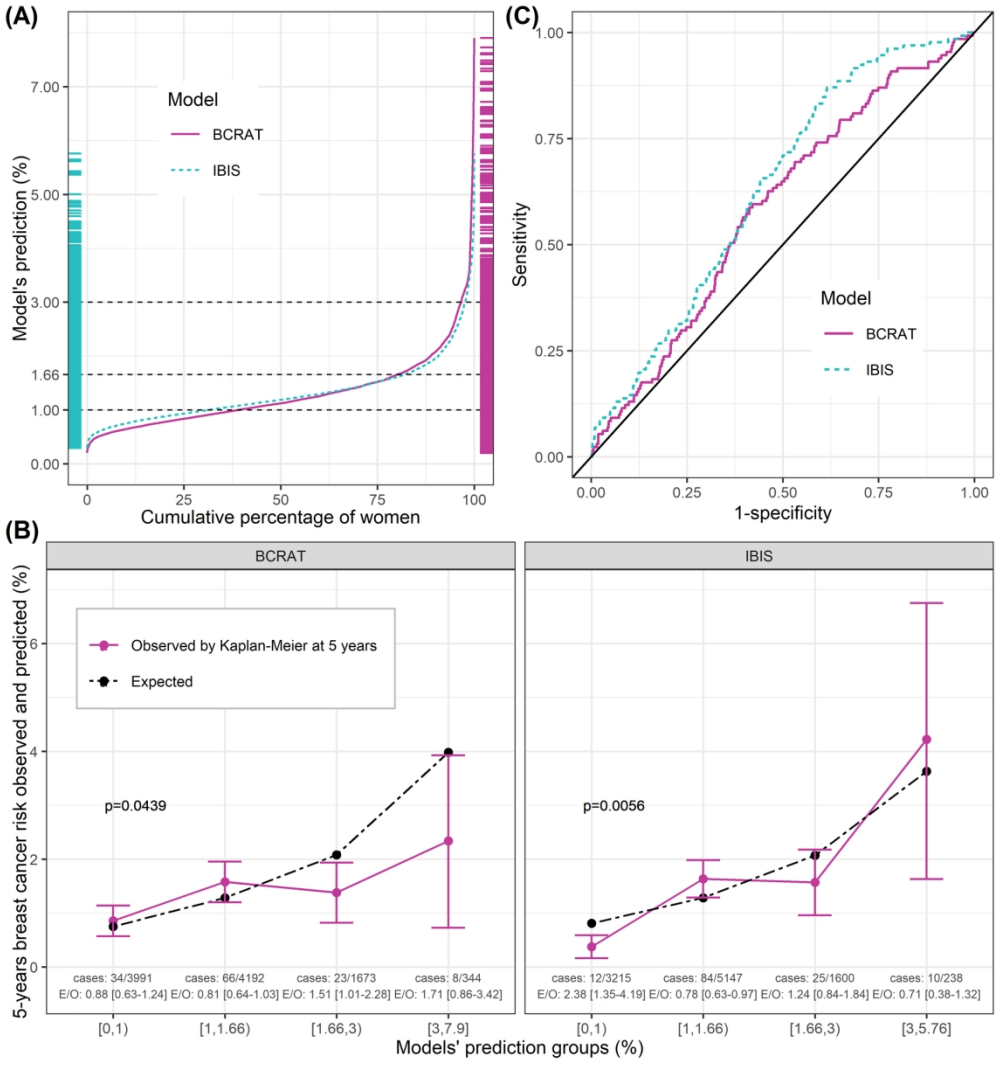
Supplementary Figure S1: Distribution of BCRAT, IBIS, PRS and combined scores predictions as a function of cumulative percentage of women. Results from the clinicogenetic-based cohort.

Supplementary Figure S2: Discrimination power of BCRAT, IBIS, PRS scores and combined models according to sensitivity and specificity. Results from the clinicogenetic-based cohort. C-indexes were calculated using the Inverse Probability of Censoring Weighting (IPCW) estimation of cumulative time-dependent ROC curve. Each PRS models name referred to the first author of the study from which the PRS were derived.



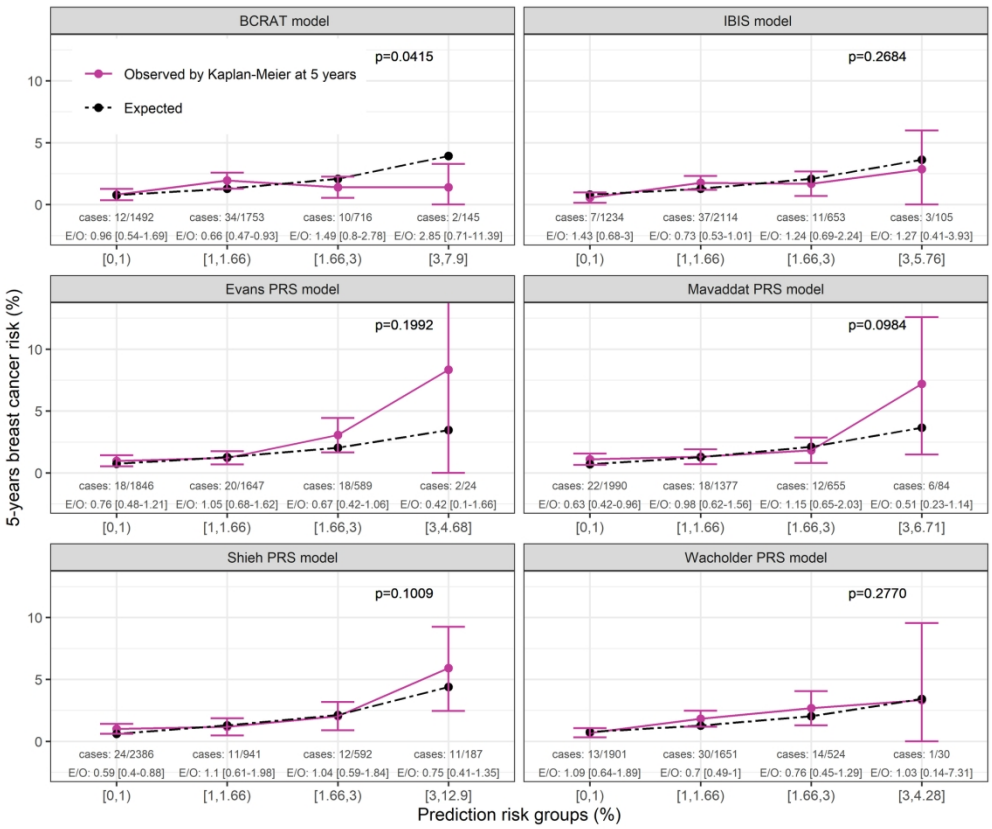
Flow-Chart

145x96mm (300 x 300 DPI)



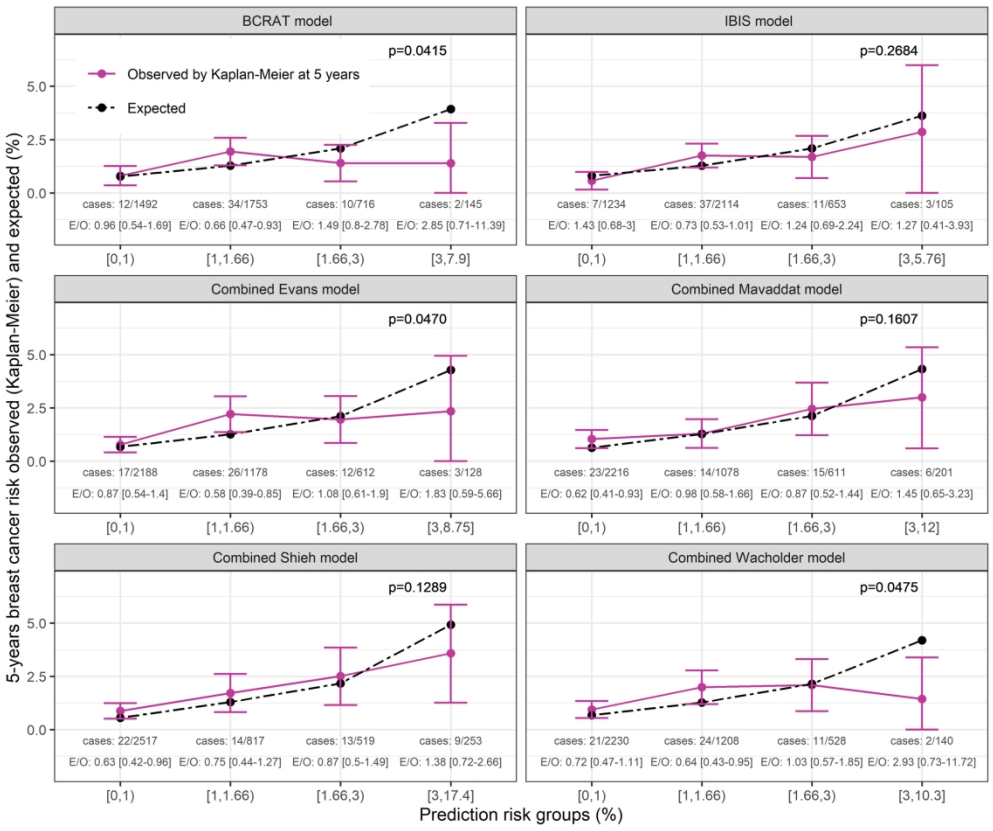
Absolute risk distribution and performance of the BCRAT and IBIS models in the Clinical-based cohort. (A) Distribution of models' predictions as a function of cumulative percentage of women. Rug plot on the y-axis. (B) Calibration according to the models' predictions groups. P values were computed using a goodness of fit test statistic compared to the critical value from the chi-squared distribution. E/O: expected-to-observed cases. (C) Discrimination power of the models according to sensitivity and specificity. C-index was calculated using the Inverse Probability of Censoring Weighting (IPCW) estimation of cumulative time-dependent ROC curve.

169x179mm (300 x 300 DPI)



Calibration according to BCRAT, IBIS and PRS scores' predictions groups. PRS: polygenic risk score. E/O: expected-to-observed cases. Results from the clinicogenetic-based cohort. P values were computed using a goodness of fit test statistic compared to the critical value from the chi-squared distribution. Each PRS models name referred to the first author of the study from which the PRS were derived.

179x149mm (300 x 300 DPI)



Calibration according to BCRAT, IBIS and combined models' predictions groups. PRS: polygenic risk score. E/O: expected-to-observed cases. Results from the clinicogenetic-based cohort. P values were computed using a goodness of fit test statistic compared to the critical value from the chi-squared distribution. Each PRS models name referred to the first author of the study from which the PRS were derived.

179x149mm (300 x 300 DPI)

Supplementary Methods

Health databases

For identifying participants who had breast cancer, we used two administrative health databases (AHD): 1) the MED-ÉCHO AHD: this database contains all the Quebec Health Insurance Board (RAMQ) diagnoses, hospitalizations and physician claims of insured patients (about 98% of Quebec residents [1]), excluding private healthcare; in the case of cancers, all patients are treated in the public sector. Data were available from January 1st, 1998 to March 31st, 2016. Dates of death were also retrieved from the RAMQ; 2) the Quebec Breast Cancer Registry: it contains information about the Quebec Breast Cancer Screening Program, such as mammograms' results and breast cancers histological confirmation. Data were available from May 15th, 1998 to December 31st, 2017.

References

1 RAMQ. Table PA.01 - Nombre de personnes inscrites et admissibles au régime d'assurance maladie du Québec selon le sexe, le groupe d'âge et la région sociosanitaire. 2017.https://www4.prod.ramq.gouv.qc.ca/IST/CD/CDF_DifsnInfoStats/CDF1_CnsulInfoStatsCNC_iut/DifsnInfoStats.aspx?ETAPE_COUR=3&IdPatronRapp=8&Annee=2017&Per=0&LANGUE=en-CA (accessed 25 Nov 2019).

Genetic data

Genotypes were included in the CaG database and were obtained from hybridation upon three different chips: Illumina Omni 2.5M (7.7% of the participants), Affymetrix Axiom UK biobank (8.2%) and Illumina Infinium Global Screening Array (84.1%). A quality control (QC) was made before the imputation (detailed pipeline can be found at www.cartagene.qc.ca/info-genetic-data): 1) QC sample: for replicated samples, samples with the lowest call rates were removed. Sample with a call rate below 95% were removed. Samples pairs with an identity by state (IBS) higher than 0.20 and similar to at least 50% of the whole set were removed. Then, for pair of samples with an IBS higher than 0.85, when the correct sample could not be identified with certainty, both samples of the pair were removed. Samples with discrepancy between sex chromosome genotypes and reported

gender were removed. 2) QC SNP: SNPs with a call rate lower than 95% or deviating from Hardy–Weinberg equilibrium (with a 10-6 threshold) were removed.

For the imputation, data were prepared using the Will Rayner toolbox (www.well.ox.ac.uk/~wrayner/tools/) with the Haplotype Reference Consortium (HRC) as reference panel [1]. To impute missing SNPs of our cohort, we used the Michigan Imputation Server with the Minimac4 algorithm [2], with separate chromosomes and chips. Imputation reference panel was the HRC r1.1 2016 European population, and the phasing was made with Eagle v2.4 [3]. A total of 39,131,578 SNPs were retrieved.

After imputation and after merging chromosomes, we used men and women to perform a sample QC based on the Anderson *et al.* protocole [4]: samples with a call rate lower than 95% and an heterozygosity higher than 3 standard deviation were removed. After LD pruning (window size: 50kb; step size: 5 variants; pairwise r^2 threshold: 0.2), for pair of participants with an IBS higher than 0.1875, the sample with the lowest call rate was removed. To remove samples with divergent ancestries, we used the two first principal components with the HapMap phase III reference panel. As we would like to have all SNPs available for calculating PRS, we did not perform an additional SNPs QC. QC process was performed using PLINK v1.90b6.2 and v2.00a2LM 64-bit ([5,6]; URL: pngu.mgh.harvard.edu/purcell/plink/).

References

1. the Haplotype Reference Consortium, McCarthy S, Das S, Kretzschmar W, Delaneau O, Wood AR, et al. A reference panel of 64,976 haplotypes for genotype imputation. Nat Genet. 2016 août;48:1279.
2. Das S, Forer L, Schönherr S, Sidore C, Locke AE, Kwong A, et al. Next-generation genotype imputation service and methods. Nat Genet. 2016;48(10):1284–7.
3. Loh P-R, Danecek P, Palamara PF, Fuchsberger C, A Reshef Y, K Finucane H, et al. Reference-based phasing using the Haplotype Reference Consortium panel. Nat Genet. 2016;48(11):1443–8.
4. Anderson CA, Pettersson FH, Clarke GM, Cardon LR, Morris AP, Zondervan KT. Data quality control in genetic case-control association studies. Nat Protoc. 2010 Sep;5(9):1564–73.

5. Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MAR, Bender D, et al. PLINK: a tool set for whole-genome association and population-based linkage analyses. Am J Hum Genet. 2007 Sep;81(3):559–75.
6. Chang CC, Chow CC, Tellier LC, Vattikuti S, Purcell SM, Lee JJ. Second-generation PLINK: rising to the challenge of larger and richer datasets. GigaScience. 2015 Dec;4(1):7.

Absolute risk of breast cancer

The absolute risk of breast cancer over an established period $[t_0,t_1]$ (five years in this study) is the probability that a woman who is free of a breast cancer at age t_0 and has a risk score S will be diagnosed with breast cancer over the period $[t_0,t_1]$.

Under the assumption of a multiplicative proportional hazard model (or Cox model), this latter conditional probability (denoted $AR(t_0,t_1;S)$) can be written such as:

$$AR(t_0,t_1;S) = \int_{t_0}^{t_1} \lambda_0(t)e^S exp\left[-\int_{t}^{t_0} \lambda_0(u)e^S + \gamma(u)du\right]dt$$

where $\lambda_0(t)$ and $\gamma(t)$ are the baseline age-specific hazard rate for breast cancer and the age-specific mortality hazard rate from other causes (competing risks), respectively. In practice, the absolute risk is computed using piece-wise constant hazard rates.

These baseline hazard rates are calculated using marginal (or composite) hazard rates obtained from registries, together with either the attributable hazard function or the risk factor distribution.

In this work, the timescale of the analyses was age of an individual so that t_0 was the age of a woman at entry into the cohort and t_1 was the age five years later.

For the IBIS model, the baseline age-specific hazard rate for breast cancer is replaced by a hazard rate estimate obtained from the segregation model conditionally on the woman's family history.

Variables extraction

Age at inclusion was calculated using the birthdate. We retrieved from the CARTaGENE questionnaire the first menstrual period, first live birth, number of first-degree relatives with breast cancer, ethnicity, menopause occurrence and age at menopause, height, weight, hormonal replacement therapy (HRT) use, length of HRT and last HRT use. If first menstrual period occurred after first live birth, both were considered as missing. We retrieved from the Quebec Breast Cancer Registry the previous breast biopsy and the number of biopsy with hyperplasia, atypical hyperplasia and lobular carcinoma *in situ*. We retrieved from the RAMQ the occurrence and age of ovary cancers.

snp;id*;genes;wacholder**;evans;mavaddat;shieh;shieh_asian***
rs1338704;88;0 88;0 88;1 06
rs1045485;96;-;-
rs999737;1 92;0 92;0 93
rs3817198;07;1 07;1 07
rs889312;5 12;1 12;1 12;1 05
rs7716600;5:44875005;AC093297.2 - AC114954.1;1.11/1.46;-;-;-;-
rs1328161; POU5F1B PCAT1;1.109;1 09;1 03
rs3803662;23;1 23;1 24;1 15
rs2981582;10:123352317;FGFR2;1.18/1.6;-;-;-;-
rs1124943;09;1 10;1 09;1 16
rs1099519(ZNF365;;0 86;0 86;0 86;0 94
rs1562430; CASC8 PCAT1;;0 90;-;1 16;1 16
rs909116;1 93;-;-;-
rs1156287;93;-;-;-
rs713588;1 01;-;-;-
rs8009944;04;-;-;-
rs1093193(04;-;-;-
rs1011970;05;1 05;1 06;1 06
rs704010;1 09;1 07;1 08;1 05
rs4973768;09;1 09;1 10;1 11
rs9790879;09;-;-;-
rs3757318;16;-;1 16;1 16
rs614367;1 21;-;1 21;1 29
rs2981579;27;1 25;1 27;1 27
rs1077139(86;0 86;1 15
rs865686;9 90;0 89;1 04
rs6828523;91;0 90;1 11
rs1735690;91;0 91;1 08
rs6472903;91;0 91;1 16
rs4849887;92;0 91;1 07
rs1353747; PDE4D;;-;(92;0 92;1 00
rs1292011;92;0 92;1 11
rs2236007;92;0 93;1 09
rs2823093;93;0 92;1 08
rs1781744(93;0 93;1 09
rs6504950;93;0 94;1 02
rs4808801;93;1 08;1 04
rs2736108;94;0 94;0 94
rs1124267(94;0 94;0 99
rs616488;1 94;0 94;1 06
rs1119991(94;0 95;1 03
rs3903072;94;0 95;1 05
rs1550623;94;1 06;1 21
rs720475;7 95;0 94;1 02
rs1436904; AQP4-AS1 95;0 96;1 02
rs2016394;95;-;-
rs527616;1 96;0 95;1 03
rs1182064(96;0 95;1 05
rs2380205;98;0 94;1 02
rs6678914;99;0 91;1 10
rs1006969(02;1 06;1 05
rs7591516(02;1 31;1 00
rs1242255(03;1 05;1 05
rs4245739;03;1 14;1 14
rs8170;19:(AC010463 BABAM1;; 03;1 15;1 00

rs2363956; 03;-;-
rs1047207(04;1 05;1 02
rs1271069(04;1 10;1 10
rs1107599! 04;1 11;1 11
rs7726159; 04;-;-
rs9790517; 05;1 05;1 02
rs204247;6 05;1 05;1 03
rs1075924; 05;1 06;1 05
rs1249360; 05;1 06;1 05
rs2046210; 05;1 15;1 27
rs1752911' 05;-;-
rs7904519; 06;1 06;1 02
rs3760982; 06;1 06;1 02
rs941764;1 06;1 06;1 05
rs7072776; 06;1 07;1 04
rs1178015(07;1 07;1 00
rs6762644; 07;1 07;1 03
rs9693444; 07;1 07;1 08
rs1432679; 07;1 07;1 09
rs2588809; 07;1 08;1 06
rs1685760! 07;1 08;1 07
rs1155244! 08;1 07;1 03
rs1332983! 08;1 08;1 02
rs132390;2 11;1 12;1 00
rs1094167! 12;1 13;1 08
rs554219;1 12;1 27;1 00
rs6001930; 13;1 12;1 03
rs2943559; 13;1 13;0 96
rs1266267(14;-;-
rs7854052(18;-;-
rs1181444! 22;1 26;1 08
rs1157183; 26;1 26;1 00
rs1787996' 36;1 36;1 00
rs1400681; 60;1 00
rs1082201; ZNF365;;-; 89;1 08
rs9485372; 90;1 11
rs1047435; 92;1 09
rs2290203; AC068831 PRC1-AS193;1 08
rs1753006! 05;1 05
rs9383938; 08;1 08
rs4951011; ZC3H11A; 09;1 09
rs2284378; 10;1 10
rs2392780; CASC8 PCAT1;;-;- 15;1 00
rs4415084; 17;1 00
rs3822625; 36;1 36
rs7726354; 37;1 37
* SNPs' position were based on build GRCh37/hg19;-;-;-;-;-;-;-;-
** OR for one allele/two alleles;-;-;-;-;-;-;-
*** OR from Shieh's study used for Asian women;-;-;-;-;-;-;-

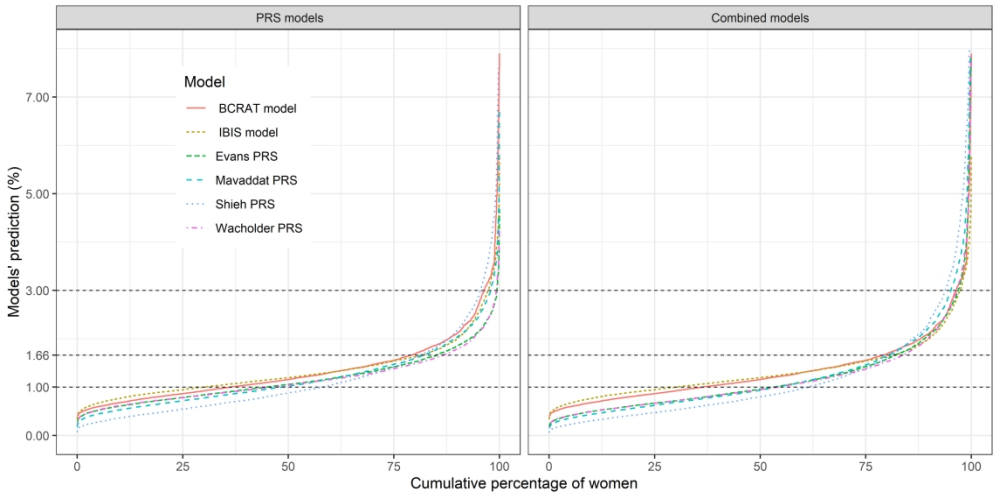
Table S2

	Clinical-based cohort	Clinicogenetic-based cohort
	N=10,200	**N=4,555**
Breast cancer within 5 years	131 (1.28%)	58 (1.27%)
BCRAT absolute risk (%)	1.3 (0.7)	1.3 (0.7)
Age at baseline (years)	54.1 (7.7)	54.1 (7.6)
Age categories:		
<=49	3,556 (34.9%)	1,557 (34.2%)
50-59	3,980 (39.0%)	1,839 (40.4%)
>=60	2,664 (26.1%)	1,159 (25.4%)
Birth province:		
In Canada outside Quebec	333 (3.3%)	74 (1.6%)
Outside Canada	1,490 (14.6%)	189 (4.1%)
Quebec	8,373 (82.1%)	4,292 (94.2%)
Missing	*4*	*0*
Ethnicity:		
Asian	188 (1.8%)	5 (0.1%)
Black African	182 (1.8%)	0 (0.0%)
Hispanic non-american	234 (2.3%)	1 (<0.1%)
Other	542 (5.3%)	86 (1.9%)
White/European	9,054 (88.8%)	4,463 (98.0%)
Age at menarche (years):		
<=11	2,305 (22.9%)	1,027 (22.7%)
12-13	4,754 (47.2%)	2,166 (47.9%)
>=14	3,021 (30.0%)	1,331 (29.4%)
Missing	*120*	*31*
Age at first live birth (years):		
<=19	1,124 (13.1%)	422 (11.1%)
20-24	2,955 (34.5%)	1,324 (34.8%)
25-29	2,814 (32.9%)	1,312 (34.5%)
>=30	1,621 (19.0%)	734 (19.3%)
Nulliparous	40 (0.5%)	14 (0.4%)
Missing	*1,646*	*749*
First-degree relatives with breast cancer:		
0	8,945 (87.7%)	3,949 (86.7%)
1	1,130 (11.1%)	556 (12.2%)
>=2	125 (1.23%)	50 (1.10%)
Previous breast biopsy:		
0	10,023 (98.3%)	4,463 (98.0%)
1	134 (1.31%)	71 (1.56%)

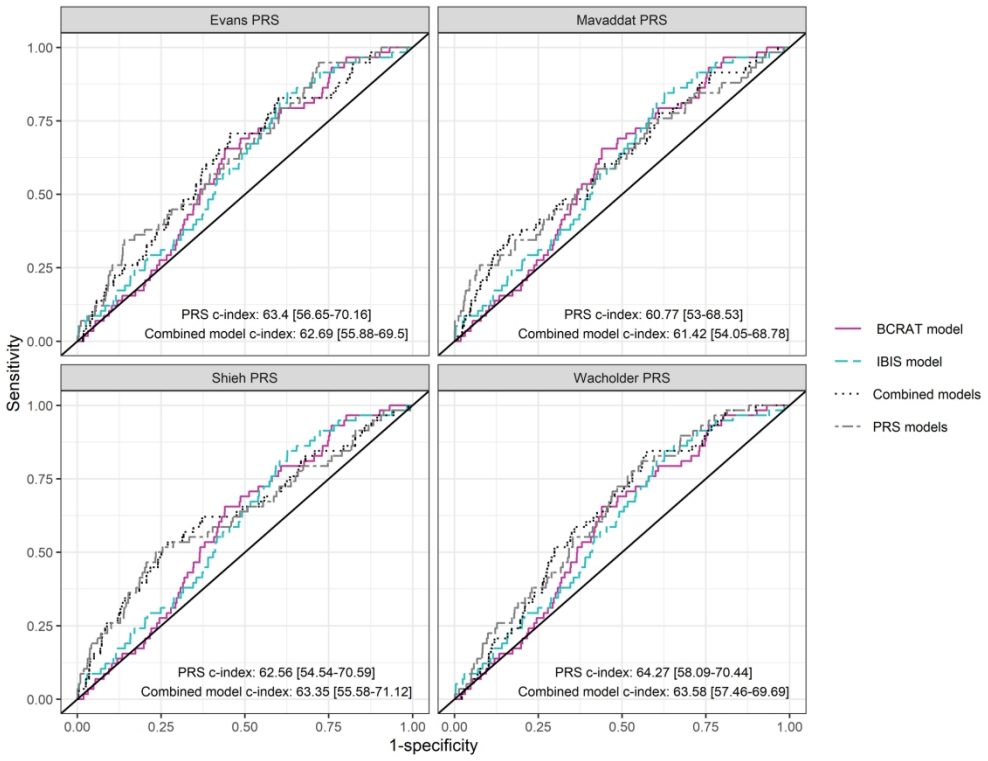
>=2	43 (0.42%)	21 (0.46%)
History of hyperplasia	6 (0.06%)	1 (0.02%)
History of atypical hyperplasia	1 (0.56%)	0 (0.00%)
Lobular Carcinoma In Situ	0	0
Weight (Kg)	67.4 [59.4;78.0]	66.8 [59.6;76.8]
Height (m)	1.61 [1.57;1.65]	1.61 [1.57;1.65]
History of ovary cancer	94 (0.92%)	46 (1.01%)
Menopause occurrence		
Pre-menopausal	4176 (40.9%)	1891 (41.5%)
Post-menopausal	5885 (57.7%)	2617 (57.5%)
Unknown	139 (1.36%)	47 (1.03%)
Use of HRT		
Never	7477 (73.3%)	3249 (71.3%)
Previous user (more than 5 years ago)	1126 (11.0%)	506 (11.1%)
Previous user (less than 5 years ago)	1285 (12.6%)	646 (14.2%)
Current user	312 (3.06%)	154 (3.38%)
HRT length of use (years)	0.00 [0.00;1.00]	0.00 [0.00;1.00]
Last HRT use (years)	0.00 [0.00;0.00]	0.00 [0.00;0.00]
Mother history of breast cancer	832 (8.16%)	412 (9.05%)
Mother history of ovary cancer	114 (1.12%)	60 (1.32%)
Father history of breast cancer	8 (0.08%)	2 (0.04%)

HRT: hormonal replacement therapy; PRS: polygenic risk score; clinical-based cohort: validation of the BCRAT and IBIS models, included women with a BCRAT and an IBIS score; clinicogenetic-based cohort: validation of the PRS models and comparison with the BCRAT and IBIS models, genotyped women with all SNPs available.

* Not available for the phase 2



254x127mm (300 x 300 DPI)



186x142mm (300 x 300 DPI)

As this is a validation study, the STROBE checklist is not fully adapted.

STROBE Statement—Checklist of items that should be included in reports of ***cohort studies***

	Item No	Recommendation	Page number
Title and abstract	1	(*a*) Indicate the study's design with a commonly used term in the title or the abstract	1
		(*b*) Provide in the abstract an informative and balanced summary of what was done and what was found	1-2
Introduction			
Background /rationale	2	Explain the scientific background and rationale for the investigation being reported	3-5
Objectives	3	State specific objectives, including any prespecified hypotheses	5
Methods			
Study design	4	Present key elements of study design early in the paper	5-6
Setting	5	Describe the setting, locations, and relevant dates, including periods of recruitment, exposure, follow-up, and data collection	5
Participants	6	(*a*) Give the eligibility criteria, and the sources and methods of selection of participants. Describe methods of follow-up	6
		(*b*) For matched studies, give matching criteria and number of exposed and unexposed	-
Variables	7	Clearly define all outcomes, exposures, predictors, potential confounders, and effect modifiers. Give diagnostic criteria, if applicable	7-8
Data sources/ measurement	8*	For each variable of interest, give sources of data and details of methods of assessment (measurement). Describe comparability of assessment methods if there is more than one group	7-8
Bias	9	Describe any efforts to address potential sources of bias	5
Study size	10	Explain how the study size was arrived at	6
Quantitative variables	11	Explain how quantitative variables were handled in the analyses. If applicable, describe which groupings were chosen and why	7-8
Statistical methods	12	(*a*) Describe all statistical methods, including those used to control for confounding	8-9
		(*b*) Describe any methods used to examine subgroups and interactions	-
		(*c*) Explain how missing data were addressed	7
		(*d*) If applicable, explain how loss to follow-up was addressed	8-9
		(*e*) Describe any sensitivity analyses	-
Results			
Participants	13*	(a) Report numbers of individuals at each stage of study—eg numbers potentially eligible, examined for eligibility, confirmed eligible, included in the study, completing follow-up, and analysed	9
		(b) Give reasons for non-participation at each stage	9
		(c) Consider use of a flow diagram	Figure 1
Descriptive data	14*	(a) Give characteristics of study participants (eg demographic, clinical, social) and information on exposures and potential confounders	Table S2
		(b) Indicate number of participants with missing data for each variable of	Table S2

		interest	
		(c) Summarise follow-up time (eg, average and total amount)	9
Outcome data	15*	Report numbers of outcome events or summary measures over time	9
Main results	16	(*a*) Give unadjusted estimates and, if applicable, confounder-adjusted estimates and their precision (eg, 95% confidence interval). Make clear which confounders were adjusted for and why they were included	10-12
		(*b*) Report category boundaries when continuous variables were categorized	Table S2
		(*c*) If relevant, consider translating estimates of relative risk into absolute risk for a meaningful time period	9
Other analyses	17	Report other analyses done—eg analyses of subgroups and interactions, and sensitivity analyses	10-12
Discussion			
Key results	18	Summarise key results with reference to study objectives	12
Limitations	19	Discuss limitations of the study, taking into account sources of potential bias or imprecision. Discuss both direction and magnitude of any potential bias	13
Interpretation	20	Give a cautious overall interpretation of results considering objectives, limitations, multiplicity of analyses, results from similar studies, and other relevant evidence	12-13
Generalisability	21	Discuss the generalisability (external validity) of the study results	13
Other information			
Funding	22	Give the source of funding and the role of the funders for the present study and, if applicable, for the original study on which the present article is based	21

*Give information separately for exposed and unexposed groups.

Note: An Explanation and Elaboration article discusses each checklist item and gives methodological background and published examples of transparent reporting. The STROBE checklist is best used in conjunction with this article (freely available on the Web sites of PLoS Medicine at http://www.plosmedicine.org/, Annals of Internal Medicine at http://www.annals.org/, and Epidemiology at http://www.epidem.com/). Information on the STROBE Initiative is available at http://www.strobe-statement.org.

BMJ Open

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

Journal:	*BMJ Open*
Manuscript ID	bmjopen-2020-045078.R1
Article Type:	Original research
Date Submitted by the Author:	22-Dec-2020
Complete List of Authors:	Jantzen, Rodolphe; CHU Sainte-Justine, Centre de recherche Payette, Yves; CHU Sainte-Justine Centre de Recherche de Malliard, Thibault; CHU Sainte-Justine Centre de Recherche Labbé, Catherine; CHU Sainte-Justine Centre de Recherche Noisel, Nolwenn; CHU Sainte-Justine Centre de Recherche Broët, Philippe; CHU Sainte-Justine Centre de Recherche
Primary Subject Heading:	Public health
Secondary Subject Heading:	Epidemiology, Genetics and genomics, Oncology
Keywords:	Breast tumours < ONCOLOGY, EPIDEMIOLOGY, GENETICS, PREVENTIVE MEDICINE, PUBLIC HEALTH

I, the Submitting Author has the right to grant and does grant on behalf of all authors of the Work (as defined in the below author licence), an exclusive licence and/or a non-exclusive licence for contributions from authors who are: i) UK Crown employees; ii) where BMJ has agreed a CC-BY licence shall apply, and/or iii) in accordance with the terms applicable for US Federal Government officers or employees acting as part of their official duties; on a worldwide, perpetual, irrevocable, royalty-free basis to BMJ Publishing Group Ltd ("BMJ") its licensees and where the relevant Journal is co-owned by BMJ to the co-owners of the Journal, to publish the Work in this journal and any other BMJ products and to exploit all rights, as set out in our licence.

The Submitting Author accepts and understands that any supply made under these terms is made by BMJ to the Submitting Author unless you are acting as an employee on behalf of your employer or a postgraduate student of an affiliated institution which is paying any applicable article publishing charge ("APC") for Open Access articles. Where the Submitting Author wishes to make the Work available on an Open Access basis (and intends to pay the relevant APC), the terms of reuse of such Open Access shall be governed by a Creative Commons licence – details of these licences and which Creative Commons licence will apply to this Work are set out in our licence referred to above.

Other than as permitted in any relevant BMJ Author's Self Archiving Policies, I confirm this Work has not been accepted for publication elsewhere, is not being considered for publication elsewhere and does not duplicate material already published. I confirm all authors consent to publication of this Work and authorise the granting of this licence.

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,*]

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

[2]Université de Montréal, Montréal, Québec, Canada

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

[4]Assistance Publique-Hôpitaux de Paris (AP-HP), Hôpitaux Universitaires Paris-Sud, Hôpital Paul Brousse, 12 Avenue Paul Vaillant Couturier, 94807 Villejuif, France

*** Correspondence:**

Rodolphe Jantzen, rodolphe.jantzen@gmail.com

Philippe Broët, philippe.broet@umontreal.ca

Keywords: Breast cancer, Validation Study, Clinical Decision Rules, Polygenic risk score, BCRAT, IBIS

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) to predict the occurrence of invasive breast cancers at five 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 4,555 with genetic information for validating the PRS and combined scores.

Outcome measures: We computed the absolute risks of breast cancer at five 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 five 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] versus 58.63 [54.05-63.21], $p=0.013$). PRS scores had a global calibration around 0.82, with a confidence interval 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 personalized message to their high risk patients, regarding screening and lifestyle counseling.

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 forty years old.

- Since the genotyping information was not available for all the cohort, the models had to be evaluated on two different sub-cohorts.

1 Introduction

Breast cancer is the most frequently diagnosed cancer and the second leading cause 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 BCRAT was developed by Dr. Mitchell Gail in 1989 using 5,998 American women from a case-control study [2]. 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 (e.g., 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 U.S. Food and Drug Administration recommended to consider chemo-prevention for women at high risk of breast cancer (i.e. a 5-year risk equal or higher than 1.66%), while the U.S. Preventive Services Task Force recommended chemo-prevention 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 U.S.-centered use led to an under-use 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 International Breast Cancer Intervention Study model (IBIS, also known as the Tyrer-Cuzick model) is also a widely used breast cancer risk prediction model, which takes into account multi-generational 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 to 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 multi-generational 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* gene mutation [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 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 ROC curve (AUC) increasing from 0.607 to 0.632 [15]. This kind of clinico-genetic 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 upon 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 (GWAS) [21], with weights that can be adapted to specific ethnicities [19]. The predictive accuracy of these PRSs compared to classical prediction models, such as the BCRAT and IBIS, should now be evaluated in various populations.

In Quebec, the Breast Cancer Screening Program consists of a mammogram every two years for women aged 50 to 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 organized 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 multi-countries cohorts.

The fact that BCRAT and IBIS have not been evaluated in the French-Canadian population, which has specific genetic patterns, as compared to the general European population [25,26], with lifestyle risk factors (e.g., 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 polygenic risk scores to predict the occurrence of invasive breast cancers at five years in middle-aged and older French-Canadian women.

2 Materials and methods

2.1 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 Supplementary Methods), and an algorithm based on a previous report from the *Institut National de Santé Publique du Québec* (INSPQ) [28] and the Tonelli *et al.* algorithm [29]. 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 (i.e., lesion suspected of malignancy) without histologically confirmed breast cancers and retrieved, when available, the incidence date after the abnormal mammography from the RAMQ database for women with at least two claims in two years or one hospitalization with the appropriate International Classification of Diseases (ICD), Ninth or Tenth Revision codes (174 and C50). 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 sub-cohort 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 sub-cohort of 4,555 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.

2.2 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]. Single-nucleotide polymorphism (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 the Supplementary Methods.

2.3 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 five-years study period (administrative censoring) or at death.

2.4 Predictive scores

2.4.1 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 (United States and United Kingdom incidence rates, respectively), and the attributable risk obtained from the United States population data (BCRAT) and the United Kingdom/Swedish population data (IBIS). In this article, the BCRAT and IBIS absolute risks of breast cancer at five years were calculated for each woman at the inclusion date using the National Institutes of Health R package "BCRA", version 2.1 [33] and the latest version of the "IBIS Breast Cancer Risk Evaluation Tool" (http://www.ems-trials.org/riskevaluator/ — version 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 Supplementary Methods for information about variables extraction and coding. Missing data can be handled in both BCRAT and IBIS models.

2.4.2 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 clinicogenetic-based cohort with small probability weights for the breast cancer cases. To avoid the optimism bias, 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 odds ratio can be found as Supplementary Table S1.

2.4.3 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".

As the hazard function obtained from the IBIS model is not an output of the software, we cannot combine the IBIS and PRS information in this work.

2.5 Statistical analysis

For comparing means between groups, we used a one-way ANOVA test. Relationships between categorical variables were tested using the χ^2 test. Statistical significance was considered as P-values less than 0.05. We plotted predictiveness curves (i.e., 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.

2.5.1 Calibration

We computed the expected-to-observed ratio (E/O), with the 95% confidence interval (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), ≥1% and <1.66% (intermediate risk), ≥1.66% and <3% (average risk), ≥3% (high risk). The observed proportion at five-year 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 ($G = \sum_{i=1}^{4} \sum (Oi - Ei)^2 / Ei$) where *Oi* and *Ei* are respectively the observed and expected number of events in group *i*, and compared this latter to the critical value from the chi-squared distribution with four degrees of freedom.

2.5.2 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]. Receiver operating characteristic (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].

2.5.3 Sensitivity and specificity

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

All statistical analyses were performed using R software, version 3.6 [40].

2.6 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.

3 Results

Overall, 10,200 women were included for validating the BCRAT and IBIS scores and 4,555 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 five-year follow-up, for the CC and CGC, respectively. The clinical characteristics of the two cohorts can be found in the Supplementary Table S2.

3.1 Breast cancer risk prediction models (BCRAT and IBIS) evaluated in the clinical-based cohort

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.

3.2 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 sub-cohort, 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 expected/observed ratio 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 (Supplementary Figure S1). 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 zero and one, 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 to BCRAT and IBIS scores (down-left corner of the ROC curves, Supplementary Figure S2). Using a 1.66% threshold, all PRS scores increased both the sensitivity and the specificity as compared to the BCRAT and IBIS risk prediction score (Table 1).

All the combined models (BCRAT + PRS) 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=0.0478 and p=0.0471, 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 zero and one, respectively (Table 2).

The 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 1). The discrimination for women at higher risk was better for the Shieh and Mavaddat combined scores (down-left corner of the ROC curves, Supplementary Figure S2). Using a 1.66% threshold, only the Evans combined model increased both the sensitivity and the specificity as compared to the BCRAT and IBIS risk prediction score (Table 2).

4 Discussion

In this work, we reported the predictive performance of BCRAT, IBIS and four polygenic risk scores for predicting breast cancer occurrence within five 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 and intermediate risk groups (below 1.66% risk). In our study, IBIS produced slightly better discrimination than BCRAT. As compared to the clinical-based models, the genetic prediction models (PRS) did not provide a significant improvement of the discriminative capacity. Adding PRS to the BCRAT scores did not significantly increase the predictive power of BCRAT.

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 over-prediction 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-standardized incidence rates since, based on information retrieved from national cancer databases [35,44,45], the incidence rates are comparable between the United States and Canada (250.4 [95%CI 209.0-298.3] cases per 100,000 per year for Canada and 236.8 [95%CI 235.5-238.1] for US). The IBIS model, the PRS models and the clinico-genetic model (BCRAT+PRS) had also an overall good mean calibration. However, when analyzing calibration across risk subgroups, the IBIS model had a significant goodness of fit test, with an over and underestimated the risk in the low and intermediate groups, respectively, probably explained by the United-Kingdom 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 to 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 emphasizes that the inclusion of new modifiable risk factors can increase discriminatory accuracy of predictive models. The PRS and the clinico-genetic model 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 and the PRSs models' relative risks. 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, even though it may lead to good predictive performance.

Some strengths of the present study should be highlighted. Firstly, 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. Secondly and to the best of our knowledge, this study is the first to evaluate the breast cancer risk assessment tools in a French-Canadian population for predicting breast cancer at five years.

This study has nevertheless some limitations. Firstly, our findings may not apply to younger women under forty years old. Secondly, 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 chemo-prophylaxis in the US [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. Thirdly, 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 to BCRAT between the two cohorts might be linked to the smaller size of the clinicogenetic-based cohort as compared to the clinic-based cohort. The ethnicity differences between the two sub-cohorts 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 sub-cohorts 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.

4.1 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 personalized message to their high risk patients, regarding screening and lifestyle counseling.

5 Tables

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 [0.73-1.22] 0.94 [0.73-1.22]	0.83 [0.65-1.08]	0.81 [0.63-1.05]	0.82 [0.63-1.06]	0.81 [0.62-1.04]
Goodness of fit	p=0.0415 p=268	p=0.0984	p=0.1009	p=0.1992	p=0.2770
Intercept	-2 [-4.4 - 0.2] -0.8 [-3.4 - 1.8]	- 0.3 [-2.4 - 1.8]	-1 [-2.5 - 0.5]	1 [-1.6 - 3.6]	0.9 [-1.8 - 3.5]
Slope	0.5 [0 - 1] 0.8 [0.2 - 1.4]	0.9 [0.4 - 1.4]	0.7 [0.4 - 1.1]	1.2 [0.6 - 1.8]	1.1 [0.5 - 1.7]
C-index	59.13 [52.96-65.29] 59.63 [53.26-66]	60.77 [53-68.53]	62.56 [54.54-70.59]	63.4 [56.65-70.16]	64.27 [58.09-70.44]
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% [11.2-33.4] 24.1% [13.9-37.2]	31% [19.5-44.5]	39.7% [27-53.4]	34.5% [22.5-48.1]	25.9% [15.3-39]
Specificity *	79% [77.7-80.3] 81.6% [80.4-82.8]	82.2% [81-83.4]	81.3% [80.1-82.5]	85.4% [84.2-86.4]	86.7% [85.6-87.7]

BCRAT: Breast Cancer Risk Assessment Tool; E/O: expected-to-observed ratio; IBIS: International Breast Cancer Intervention Study risk evaluation tool; NRI: Net Reclassification Index

Clinicogenetic-based validation cohort: validation of the PRS models and comparison with the BCRAT and IBIS models, genotyped women with all SNPs available. 10% of the cohort was used to obtain the relative risk distribution while the remained 90% was used for computing the results.

95% confidence intervals in square brackets

* 1.66% threshold

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

	BCRAT model / IBIS model	Combined Mavaddat	Combined Shieh	Combined Evans	Combined Wacholder
E/O	0.94 [0.73-1.22] 0.94 [0.73-1.22]	0.86 [0.66-1.11]	0.83 [0.64-1.07]	0.83 [0.64-1.08]	0.82 [0.64-1.06]
Goodness of fit	p=0.0415 p=0.268	p=0.161	p=0.13	p=0.047	p=0.0475
Intercept	-2 [-4.4 - 0.2] -0.8 [-3.4 - 1.8]	- 1.5 [-3.3 - 0.1]	-1.6 [-3 - -0.3]	-1.2 [-3.1 - 0.6]	-1.3 [-3.2 - 0.5]
Slope	0.5 [0 - 1] 0.8 [0.2 - 1.4]	0.6 [0.2 - 1]	0.6 [0.3 - 0.9]	0.7 [0.3 - 1.1]	0.7 [0.2 - 1.1]
C-index	59.13 [52.96-65.29] 59.63 [53.26-66]	61.42 [54.05-68.78]	63.35 [55.58-71.12]	62.69 [55.88-69.50]	63.58 [57.46-69.69]
C-indexes comparison with:					
BCRAT model	-	p=0.50	p=0.28	p=0.12	p=0.059
IBIS model	-	p=0.66	p=0.42	p=0.38	p=0.22
Sensitivity *	20.7% [11.2-33.4] 24.1% [13.9-37.2]	36.2% [24-49.9]	37.9% [25.5-51.6]	25.9% [15.3-39]	22.4% [12.5-35.3]
Specificity *	79% [77.7-80.3] 81.6% [80.4-82.8]	80.5% [79.2-81.7]	81.5% [80.2-82.7]	82.1% [80.9-83.3]	83.8% [82.6-84.9]

BCRAT: Breast Cancer Risk Assessment Tool; E/O: expected-to-observed ratio; IBIS: International Breast Cancer Intervention Study risk evaluation tool; NRI: Net Reclassification Index
Combined scores: PRS scores combined with the BCRAT scores.

Clinicogenetic-based validation cohort: validation of the PRS models and comparison with the BCRAT and IBIS models, genotyped women with all SNPs available. 10% of the cohort was used to obtain the relative risk distribution while the remained 90% was used for computing the results.

95% confidence intervals in square brackets

* 1.66% threshold

6 Figures

Figure 1 Flow-chart

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 to the critical value from the chi-squared distribution. E/O: expected-to-observed cases. (C) Discrimination power of the models according to sensitivity and specificity. C-index was calculated using the Inverse Probability of Censoring Weighting (IPCW) estimation of cumulative time-dependent ROC curve.

Figure 3 **Calibration according to BCRAT, IBIS and PRS scores' predictions groups.** PRS: polygenic risk score. E/O: expected-to-observed cases. Results from the clinicogenetic-based cohort. P values were computed using a goodness of fit test statistic compared to the critical value from the chi-squared distribution. Each PRS models name referred to the first author of the study from which the PRS were derived.

Figure 4 **Calibration according to BCRAT, IBIS and combined models' predictions groups.** PRS: polygenic risk score. E/O: expected-to-observed cases. Results from the clinicogenetic-based cohort. P values were computed using a goodness of fit test statistic compared to the critical value from the chi-squared distribution. Each PRS models name referred to the first author of the study from which the PRS were derived.

7 References

1 Canadian Cancer Statistics Advisory Committee. Canadian Cancer Statistics 2019. *Toronto, ON: Canadian Cancer Society* Published Online First: September 2019.cancer.ca/Canadian-Cancer-Statistics-2019-EN (accessed 13 Nov 2019).

2 Gail MH, Brinton LA, Byar DP, *et al.* Projecting Individualized Probabilities of Developing Breast Cancer for White Females Who Are Being Examined Annually. *JNCI Journal of the National Cancer Institute* 1989;**81**:1879–86. doi:10/djr43s

3 Tyrer J, Duffy SW, Cuzick J. A breast cancer prediction model incorporating familial and personal risk factors. *Statist Med* 2004;**23**:1111–30. doi:10.1002/sim.1668

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. doi:10/ckg9vc

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. doi:10/btbjf9

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. doi:10/dkmv6x

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**. doi:10/gc83rw

8 Farkas A, Vanderberg R, Merriam S, *et al.* Breast Cancer Chemoprevention: A Practical Guide for the Primary Care Provider. *Journal of Women's Health* 2020;**29**:46–56. doi:10.1089/jwh.2018.7643

9 Levine M, Moutquin J-M, 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. www.cancer.ca. http://www.cancer.ca/en/cancer-information/cancer-type/breast/risks/reducing-your-risk/?region=qc (accessed 28 Nov 2019).

11 Voruganti RT, 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. *BMJ Health & Care Informatics* 2015;**22**:333–9. doi:10/ggt2c8

12 Terry MB, Liao Y, Whittemore AS, *et al.* 10-year performance of four models of breast cancer risk: a validation study. *The Lancet Oncology* 2019;**20**:504–17. doi:10/ggt2dp

13 Gail MH. Performance of BCRAT in high-risk patients with breast cancer. *The Lancet Oncology* 2019;**20**:e285. doi:10.1016/S1470-2045(19)30301-8

14 Terry MB, Liao Y, Hopper JL, *et al.* Performance of BCRAT in high-risk patients with breast cancer – Authors' reply. *The Lancet Oncology* 2019;**20**:e286. doi:10.1016/S1470-2045(19)30311-0

15 Gail MH. Discriminatory Accuracy From Single-Nucleotide Polymorphisms in Models to Predict Breast Cancer Risk. *J Natl Cancer Inst* 2008;**100**:1037–41. doi:10/ff2x2r

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. doi:10.1158/1055-9965.EPI-15-0838

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. doi:10/fn7zxw

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/f3nv95

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. doi:10/ggt2dm

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. doi:10/f9nd2w

21 MacArthur J, Bowler E, Cerezo M, *et al.* The new NHGRI-EBI Catalog of published genome-wide association studies (GWAS Catalog). *Nucleic Acids Research* 2017;**45**:D896. doi:10/f9v7cp

22 Gouvernement du Québec. Québec Breast Cancer Screening Program. 2018.https://www.quebec.ca/en/health/advice-and-prevention/screening-and-carrier-testing-offer/quebec-breast-cancer-screening-program/ (accessed 9 Mar 2020).

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 (Oxford)* 2004;**43**:1178–81. doi:10/bq2cx6

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. doi:10.1371/journal.pgen.1003815

26 Scriver CR. Human Genetics: Lessons from Quebec Population. *Annu Rev Genom Hum Genet* 2001;**2**:69–101. doi:10.1146/annurev.genom.2.1.69

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. *International Journal of Epidemiology* 2013;**42**:1285–99. doi:10.1093/ije/dys160

28 Théberge I, Institut national de santé publique du Québec, Direction systèmes de soins et services. *Validation de stratégies pour obtenir le taux de détection du cancer, la valeur prédictive 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 données fournies par le programme québécois de dépistage du cancer du sein (PQDCS)*. Montréal: : Direction des systèmes de soins et services, Institut national de santé publique 2003.

29 Tonelli M, Wiebe N, Fortin M, *et al.* Methods for identifying 30 chronic conditions: application to administrative data. *BMC Medical Informatics and Decision Making* 2015;**15**:31. doi:10/f698hd

30 Akçimen F, Ross JP, Sarayloo F, *et al.* Genetic and epidemiological characterization of restless legs syndrome in Québec. *Sleep* 2019;**43**. doi:10.1093/sleep/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. doi:10/ggt2b8

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. doi:10/ggrr3s

33 Zhang F. *BCRA: breast cancer risk assessment*. 2018. https://CRAN.R-project.org/package=BCRA

34 Choudhury PP, Maas P, Wilcox A, *et al.* iCARE: An R package to build, validate and apply absolute risk models. *PLOS ONE* 2020;**15**:e0228198. doi:10/ggt2ch

35 Statistics Canada. Table 13-10-0111-01 Number and rates of new cases of primary cancer, by cancer type, age group and sex. doi.org/10.25318/1310011101-eng (accessed 10 Nov 2019).

36 Statistics Canada. Table 13-10-0392-01 Deaths and age-specific mortality rates, by selected grouped causes. https://doi.org/10.25318/1310039201-eng (accessed 10 Nov 2019).

37 Uno H, Cai T, Tian L, *et al.* Evaluating Prediction Rules for t-Year Survivors with Censored Regression Models. *Journal of the American Statistical Association* 2007;**102**:527–37. doi:10.1198/016214507000000149

38 Blanche P, Latouche A, Viallon V. Time-dependent AUC with right-censored data: a survey study. *arXiv:12106805 [statME]* Published Online First: 25 October 2012.http://arxiv.org/abs/1210.6805 (accessed 26 Sep 2017).

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. *Statist Med* 2013;**32**:5381–97. doi:10.1002/sim.5958

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. *JNCIJ* 2016;**108**:djv348. doi:10/ggt2db

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. doi:10/ggt2dc

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. doi:10/ggt2c4

44 Statistics Canada. Table 17-10-0005-01 Population estimates on July 1st, by age and sex. doi: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.www.cdc.gov/cancer/dataviz (accessed 9 Oct 2019).

46 Moyer VA. Medications for Risk Reduction of Primary Breast Cancer in Women: U.S. Preventive Services Task Force Recommendation Statement. *Ann Intern Med* Published Online First: 24 September 2013. doi:10/ggt2cx

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. doi:10/ggt2c3

8 Funding

This research received no specific grant from any funding agency in the public, commercial or not-for-profit sectors.

9 Competing Interests

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

10 Author Contributions

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

11 Data Availability Statement

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).

12 Ethics approval and consent to participate

This project has been approved by the Research Ethics Board of the CHU Sainte-Justine under the reference 2020-2427. In addition, CARTaGENE has obtained ethics approval by the CHU Sainte-Justine under the reference: MP-21-2011-345, 3297. The latest annual ethics renewal was granted on September 13, 2019. This latter approval implies that all participants have given their consent (cartagene.qc.ca/sites/default/files/documents/consent/brochure_en_0505_0.pdf). Consent was obtained from all the participants.

13 Acknowledgments

We would like to thank all the CARTaGENE participants for their generous investments in health research. We would also like to thank the RAMQ and the Commission d'accès à l'information (CAI) for their support in obtaining the data.

14 Supplementary Material

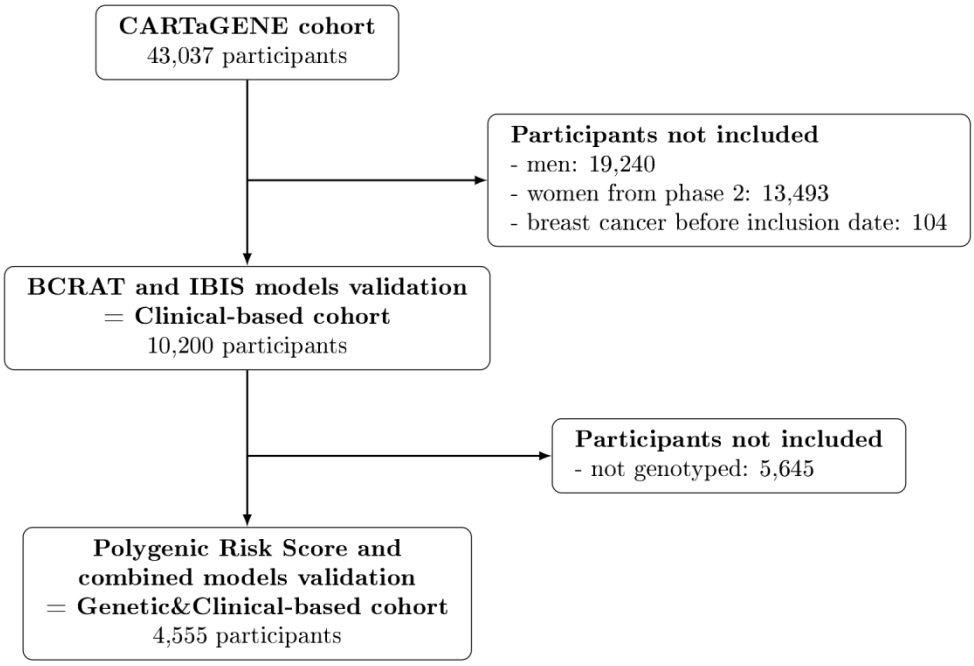
Supplementary Methods

Supplementary Table S1: SNPs used for each extended model and the associated gene and odds ratio.

Supplementary Table S2: Characteristics comparison of the women from the Clinical-based and the clinicogenetic-based cohorts.

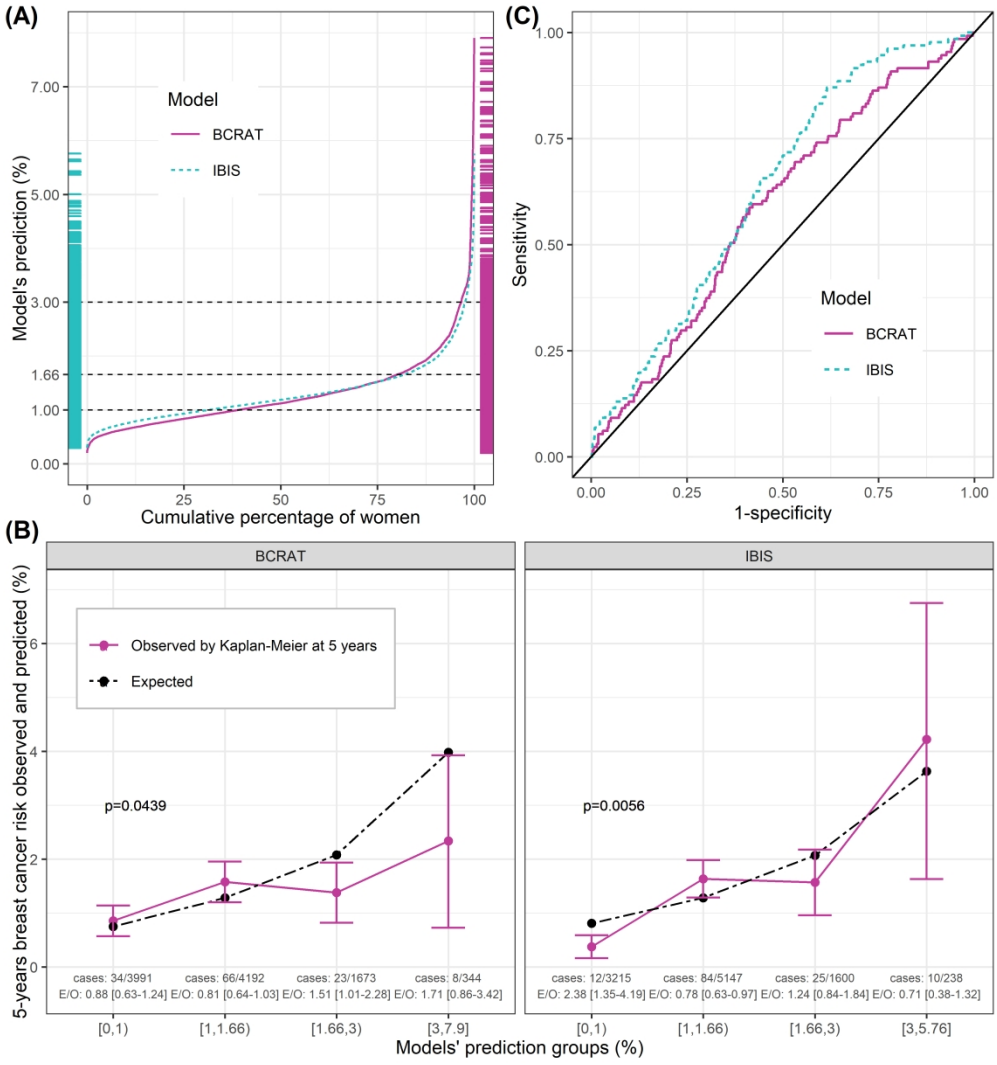
Supplementary Figure S1: Distribution of BCRAT, IBIS, PRS and combined scores predictions as a function of cumulative percentage of women. Results from the clinicogenetic-based cohort.

Supplementary Figure S2: Discrimination power of BCRAT, IBIS, PRS scores and combined models according to sensitivity and specificity. Results from the clinicogenetic-based cohort. C-indexes were calculated using the Inverse Probability of Censoring Weighting (IPCW) estimation of cumulative time-dependent ROC curve. Each PRS models name referred to the first author of the study from which the PRS were derived.



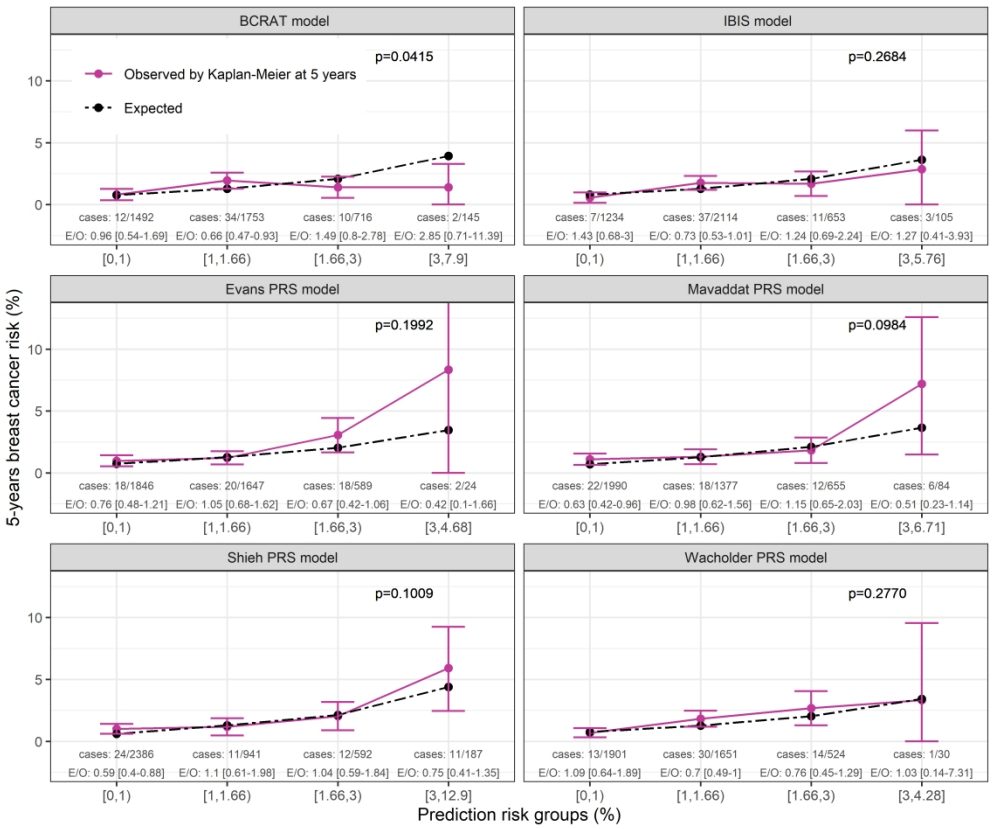
Flow-Chart

145x96mm (300 x 300 DPI)



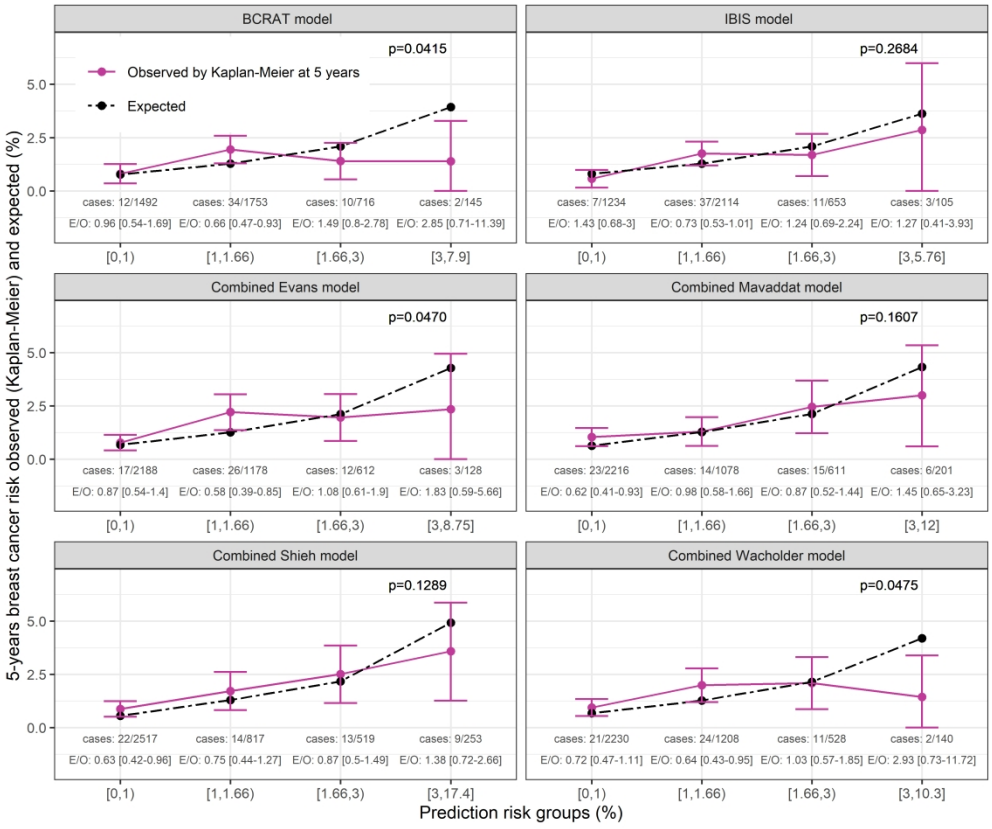
Absolute risk distribution and performance of the BCRAT and IBIS models in the Clinical-based cohort. (A) Distribution of models' predictions as a function of cumulative percentage of women. Rug plot on the y-axis. (B) Calibration according to the models' predictions groups. P values were computed using a goodness of fit test statistic compared to the critical value from the chi-squared distribution. E/O: expected-to-observed cases. (C) Discrimination power of the models according to sensitivity and specificity. C-index was calculated using the Inverse Probability of Censoring Weighting (IPCW) estimation of cumulative time-dependent ROC curve.

169x179mm (600 x 600 DPI)



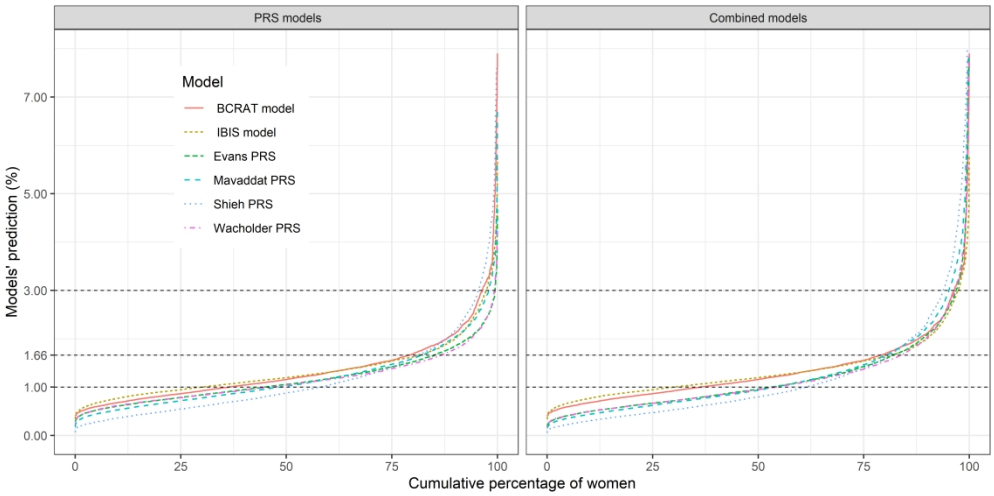
Calibration according to BCRAT, IBIS and PRS scores' predictions groups. PRS: polygenic risk score. E/O: expected-to-observed cases. Results from the clinicogenetic-based cohort. P values were computed using a goodness of fit test statistic compared to the critical value from the chi-squared distribution. Each PRS models name referred to the first author of the study from which the PRS were derived.

179x149mm (600 x 600 DPI)

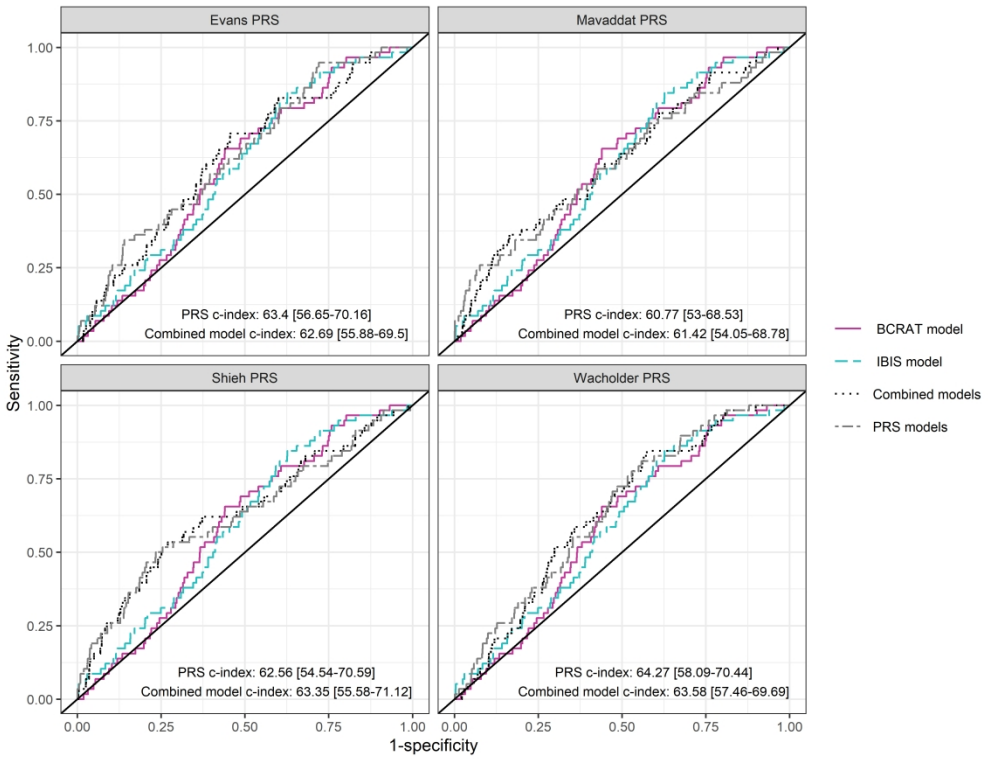


Calibration according to BCRAT, IBIS and combined models' predictions groups. PRS: polygenic risk score. E/O: expected-to-observed cases. Results from the clinicogenetic-based cohort. P values were computed using a goodness of fit test statistic compared to the critical value from the chi-squared distribution. Each PRS models name referred to the first author of the study from which the PRS were derived.

179x149mm (600 x 600 DPI)



254x127mm (600 x 600 DPI)



186x142mm (600 x 600 DPI)

Supplementary Methods

Health databases

For identifying participants who had breast cancer, we used two administrative health databases (AHD): 1) the MED-ÉCHO AHD: this database contains all the Quebec Health Insurance Board (RAMQ) diagnoses, hospitalizations and physician claims of insured patients (about 98% of Quebec residents [1]), excluding private healthcare; in the case of cancers, all patients are treated in the public sector. Data were available from January 1st, 1998 to March 31st, 2016. Dates of death were also retrieved from the RAMQ; 2) the Quebec Breast Cancer Registry: it contains information about the Quebec Breast Cancer Screening Program, such as mammograms' results and breast cancers histological confirmation. Data were available from May 15th, 1998 to December 31st, 2017.

References

1 RAMQ. Table PA.01 - Nombre de personnes inscrites et admissibles au régime d'assurance maladie du Québec selon le sexe, le groupe d'âge et la région sociosanitaire. 2017.https://www4.prod.ramq.gouv.qc.ca/IST/CD/CDF_DifsnInfoStats/CDF1_CnsulInfoStatsCNC_iut/DifsnInfoStats.aspx?ETAPE_COUR=3&IdPatronRapp=8&Annee=2017&Per=0&LANGUE=en-CA (accessed 25 Nov 2019).

Genetic data

Genotypes were included in the CaG database and were obtained from hybridation upon three different chips: Illumina Omni 2.5M (7.7% of the participants), Affymetrix Axiom UK biobank (8.2%) and Illumina Infinium Global Screening Array (84.1%). A quality control (QC) was made before the imputation (detailed pipeline can be found at www.cartagene.qc.ca/info-genetic-data): 1) QC sample: for replicated samples, samples with the lowest call rates were removed. Sample with a call rate below 95% were removed. Samples pairs with an identity by state (IBS) higher than 0.20 and similar to at least 50% of the whole set were removed. Then, for pair of samples with an IBS higher than 0.85, when the correct sample could not be identified with certainty, both samples of the pair were removed. Samples with discrepancy between sex chromosome genotypes and reported

gender were removed. 2) QC SNP: SNPs with a call rate lower than 95% or deviating from Hardy–Weinberg equilibrium (with a 10-6 threshold) were removed.

For the imputation, data were prepared using the Will Rayner toolbox (www.well.ox.ac.uk/~wrayner/tools/) with the Haplotype Reference Consortium (HRC) as reference panel [1]. To impute missing SNPs of our cohort, we used the Michigan Imputation Server with the Minimac4 algorithm [2], with separate chromosomes and chips. Imputation reference panel was the HRC r1.1 2016 European population, and the phasing was made with Eagle v2.4 [3]. A total of 39,131,578 SNPs were retrieved.

After imputation and after merging chromosomes, we used men and women to perform a sample QC based on the Anderson *et al.* protocole [4]: samples with a call rate lower than 95% and an heterozygosity higher than 3 standard deviation were removed. After LD pruning (window size: 50kb; step size: 5 variants; pairwise r^2 threshold: 0.2), for pair of participants with an IBS higher than 0.1875, the sample with the lowest call rate was removed. To remove samples with divergent ancestries, we used the two first principal components with the HapMap phase III reference panel. As we would like to have all SNPs available for calculating PRS, we did not perform an additional SNPs QC. QC process was performed using PLINK v1.90b6.2 and v2.00a2LM 64-bit ([5,6]; URL: pngu.mgh.harvard.edu/purcell/plink/).

References

1. the Haplotype Reference Consortium, McCarthy S, Das S, Kretzschmar W, Delaneau O, Wood AR, et al. A reference panel of 64,976 haplotypes for genotype imputation. Nat Genet. 2016 août;48:1279.
2. Das S, Forer L, Schönherr S, Sidore C, Locke AE, Kwong A, et al. Next-generation genotype imputation service and methods. Nat Genet. 2016;48(10):1284–7.
3. Loh P-R, Danecek P, Palamara PF, Fuchsberger C, A Reshef Y, K Finucane H, et al. Reference-based phasing using the Haplotype Reference Consortium panel. Nat Genet. 2016;48(11):1443–8.
4. Anderson CA, Pettersson FH, Clarke GM, Cardon LR, Morris AP, Zondervan KT. Data quality control in genetic case-control association studies. Nat Protoc. 2010 Sep;5(9):1564–73.

5. Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MAR, Bender D, et al. PLINK: a tool set for whole-genome association and population-based linkage analyses. Am J Hum Genet. 2007 Sep;81(3):559–75.
6. Chang CC, Chow CC, Tellier LC, Vattikuti S, Purcell SM, Lee JJ. Second-generation PLINK: rising to the challenge of larger and richer datasets. GigaScience. 2015 Dec;4(1):7.

Absolute risk of breast cancer

The absolute risk of breast cancer over an established period $[t_0,t_1]$ (five years in this study) is the probability that a woman who is free of a breast cancer at age t_0 and has a risk score S will be diagnosed with breast cancer over the period $[t_0,t_1]$.

Under the assumption of a multiplicative proportional hazard model (or Cox model), this latter conditional probability (denoted $AR(t_0,t_1;S)$) can be written such as:

$$AR(t_0,t_1;S)=\int_{t_0}^{t_1}\lambda_0(t)e^S\exp\left[-\int_{t_\square}^{t_0}\lambda_0(u)e^S+\gamma(u)du\right]dt$$

where $\lambda_0(t)$ and $\gamma(t)$ are the baseline age-specific hazard rate for breast cancer and the age-specific mortality hazard rate from other causes (competing risks), respectively. In practice, the absolute risk is computed using piece-wise constant hazard rates.

These baseline hazard rates are calculated using marginal (or composite) hazard rates obtained from registries, together with either the attributable hazard function or the risk factor distribution.

In this work, the timescale of the analyses was age of an individual so that t_0 was the age of a woman at entry into the cohort and t_1 was the age five years later.

For the IBIS model, the baseline age-specific hazard rate for breast cancer is replaced by a hazard rate estimate obtained from the segregation model conditionally on the woman's family history.

Variables extraction and coding

Age at inclusion was calculated using the birthdate. We retrieved from the CARTaGENE questionnaire the first menstrual period, first live birth, number of first-degree relatives with breast cancer, ethnicity, menopause occurrence and age at menopause, height, weight, hormonal replacement therapy (HRT) use, length of HRT and last HRT use. If first menstrual period occurred after first live birth, both were considered as missing. We retrieved from the Quebec Breast Cancer Registry the previous breast biopsy and the number of biopsy with hyperplasia, atypical hyperplasia and lobular carcinoma *in situ*. We retrieved from the RAMQ the occurrence and age of ovary cancers.

How the variables were coded for the IBIS model can be found online (https://ems-trials.org/riskevaluator/), in the Documentation section, file “Risk program input file format (v6-8)”

Supplementary Table S1

snp	id*	genes	wacholder**	evans	mavaddat	shieh	shieh_asian***
rs13387042	2:217905832	AC007749.1 - RN7SKP43	0.8/0.7	0,88	0,88	0,88	1,06
rs1045485	2:202149589	-	0.89/0.69	-	0,96	-	-
rs999737	14:69034682	RAD51B	0.91/0.67	-	0,92	0,92	0,93
rs3817198	11:1909006	LSP1	1.04/1.18	-	1,07	1,07	1,07
rs889312	5:56031884	C5orf67 - AC008940.1	1.05/1.1	1,12	1,12	1,12	1,05
rs7716600	5:44875005	AC093297.2 - AC114954.1	1.11/1.46	-	-	-	-
rs13281615	8:128355618	CASC8, POU5F1B, PCAT1	1.14/1.36	-	1,09	1,09	1,03
rs3803662	16:52586341	CASC16	1.16/1.44	1,23	1,23	1,24	1,15
rs2981582	10:123352317	FGFR2	1.18/1.6	-	-	-	-
rs11249433	1:121280613	EMBP1	1.23/1.3	1,09	1,10	1,09	1,16
rs10995190	10:64278682	AC024598.1, ZNF365		0,86	0,86	0,86	0,94
rs1562430	8:128387852	POU5F1B, CASC8, PCAT1		0,90	-	1,16	1,16
rs909116	11:1941946	TNNT3		0,93	-	-	-
rs1156287	17:53076799	-		0,93	-	-	-
rs713588	10:5886962	-		1,01	-	-	-
rs8009944	14:69039588	-		1,04	-	-	-
rs10931936	2:202143928	-		1,04	-	-	-
rs1011970	9:22062134	CDKN2B-AS1		1,05	1,05	1,06	1,06
rs704010	10:80841148	ZMIZ1		1,09	1,07	1,08	1,05
rs4973768	3:27416013	SLC4A7		1,09	1,09	1,10	1,11
rs9790879	5:44899885	-		1,09	-	-	-
rs3757318	6:151914113	CCDC170		1,16	-	1,16	1,16
rs614367	11:69328764	LINC01488 - CCND1		1,21	-	1,21	1,29
rs2981579	10:123337335	FGFR2		1,27	1,25	1,27	1,27
rs10771399	12:28155080	PTHLH - CCDC91		-	0,86	0,86	1,15
rs865686	9:110888478	CHCHD4P2 - AL353742.1		-	0,90	0,89	1,04
rs6828523	4:175846426	ADAM29		-	0,91	0,90	1,11
rs17356907	12:96027759	PGAM1P5		-	0,91	0,91	1,08
rs6472903	8:76230301	CASC9		-	0,91	0,91	1,16
rs4849887	2:121245122	LINC01101 - AC073257.2		-	0,92	0,91	1,07
rs1353747	5:58337481	AC092343.1, PDE4D		-	0,92	0,92	1,00
rs1292011	12:115836522	AC078880.2 - AC009803.2		-	0,92	0,92	1,11
rs2236007	14:37132769	PAX9		-	0,92	0,93	1,09
rs2823093	21:16520832	AF127577.5 - AF246928.1		-	0,93	0,92	1,08
rs17817449	16:53813367	FTO		-	0,93	0,93	1,09
rs6504950	17:53056471	STXBP4		-	0,93	0,94	1,02
rs4808801	19:18571141	ELL		-	0,93	1,08	1,04
rs2736108	5:1297488	TERT - MIR4457		-	0,94	0,94	0,94
rs11242675	6:1318878	FOXQ1 - LINC01394		-	0,94	0,94	0,99
rs616488	1:10566215	PEX14		-	0,94	0,94	1,06
rs11199914	10:123093901	LINC01153 - RN7SKP167		-	0,94	0,95	1,03
rs3903072	11:65583066	AP001266.1 - CFL1		-	0,94	0,95	1,05
rs1550623	2:174212894	AC092573.2		-	0,94	1,06	1,21
rs720475	7:144074929	ARHGEF5		-	0,95	0,94	1,02
rs1436904	18:24570667	CHST9, AQP4-AS1		-	0,95	0,96	1,02
rs2016394	2:172972971	DLX2-DT		-	0,95	-	-
rs527616	18:24337424	AQP4-AS1		-	0,96	0,95	1,03
rs11820646	11:129461171	AP003500.2		-	0,96	0,95	1,05

Supplementary Table S1

rs2380205	10:5886734	GDI2 - ANKRD16	-	0,98	0,94	1,02
rs6678914	1:202187176	LGR6	-	0,99	0,91	1,10
rs10069690	5:1279790	TERT	-	1,02	1,06	1,05
rs75915166	11:69379161	LINC01488 - CCND1	-	1,02	1,31	1,00
rs12422552	12:14413931	GNAI2P1 - RPL30P11	-	1,03	1,05	1,05
rs4245739	1:204518842	MDM4	-	1,03	1,14	1,14
rs8170	19:17389704	USHBP1, AC010463.1, BABAM1	-	1,03	1,15	1,00
rs2363956	19:17394124	ANKLE1	-	1,03	-	-
rs10472076	5:58184061	AC008852.1 - PDE4D	-	1,04	1,05	1,02
rs12710696	2:19320803	LINC01376	-	1,04	1,10	1,10
rs11075995	16:53855291	FTO	-	1,04	1,11	1,11
rs7726159	5:1282319	TERT	-	1,04	-	-
rs9790517	4:106084778	TET2	-	1,05	1,05	1,02
rs204247	6:13722523	RANBP9 - MCUR1	-	1,05	1,05	1,03
rs10759243	9:110306115	PPIAP88 - RNU6-996P	-	1,05	1,06	1,05
rs12493607	3:30682939	TGFBR2	-	1,05	1,06	1,05
rs2046210	6:151948366	CCDC170 - ESR1	-	1,05	1,15	1,27
rs17529111	6:82128386	AL590824.1 - TENT5A	-	1,05	-	-
rs7904519	10:114773927	TCF7L2	-	1,06	1,06	1,02
rs3760982	19:44286513	KCNN4 - LYPD5	-	1,06	1,06	1,02
rs941764	14:91841069	CCDC88C	-	1,06	1,06	1,05
rs7072776	10:22032942	MLLT10 - DNAJC1	-	1,06	1,07	1,04
rs11780156	8:129194641	PVT1	-	1,07	1,07	1,00
rs6762644	3:4742276	ITPR1	-	1,07	1,07	1,03
rs9693444	8:29509616	RPL17P33 - LINC00589	-	1,07	1,07	1,08
rs1432679	5:158244083	EBF1	-	1,07	1,07	1,09
rs2588809	14:68660428	RAD51B	-	1,07	1,08	1,06
rs16857609	2:218296508	DIRC3	-	1,07	1,08	1,07
rs11552449	1:114448389	DCLRE1B	-	1,08	1,07	1,03
rs13329835	16:80650805	CDYL2	-	1,08	1,08	1,02
rs132390	22:29621477	EMID1	-	1,11	1,12	1,00
rs10941679	5:44706498	AC093292.1 - RN7SL383P	-	1,12	1,13	1,08
rs554219	11:69331642	LINC01488 - CCND1	-	1,12	1,27	1,00
rs6001930	22:40876234	MRTFA	-	1,13	1,12	1,03
rs2943559	8:76417937	HNF4G	-	1,13	1,13	0,96
rs12662670	6:151918856	CCDC170	-	1,14	-	-
rs78540526	11:69331418	LINC01488 - CCND1	-	1,18	-	-
rs11814448	10:22315843	DNAJC1 - ADIPOR1P1	-	1,22	1,26	1,08
rs11571833	13:32972626	BRCA2	-	1,26	1,26	1,00
rs17879961	22:29121087	CHEK2	-	1,36	1,36	1,00
rs140068132	6:151954834	CCDC170 - ESR1	-	-	0,60	1,00
rs10822013	10:64251977	AC024598.1, ZNF365	-	-	0,89	1,08
rs9485372	6:149608874	TAB2	-	-	0,90	1,11
rs10474352	5:90732225	ARRDC3-AS1	-	-	0,92	1,09
rs2290203	15:91512067	PRC1, AC068831.7, PRC1-AS1	-	-	0,93	1,08
rs17530068	6:82193109	AL590824.1 - TENT5A	-	-	1,05	1,05
rs9383938	6:151987357	ESR1	-	-	1,08	1,08
rs4951011	1:203766331	ZBED6, ZC3H11A	-	-	1,09	1,09
rs2284378	20:32588095	RALY	-	-	1,10	1,10

Supplementary Table S1

rs2392780	8:128388025	POU5F1B, CASC8, PCAT1	-	-	1,15	1,00
rs4415084	5:44662515	LINC02224 - AC093292.1	-	-	1,17	1,00
rs3822625	5:56178111	MAP3K1	-	-	1,36	1,36
rs7726354	5:56256483	MIER3	-	-	1,37	1,37

* SNPs' position were based on build GRCh37/hg19

** OR for one allele/two alleles

*** OR from Shieh's study used for Asian women

Supplementary Table S1

Table S2

	Clinical-based cohort	Clinicogenetic-based cohort
	N=10,200	**N=4,555**
Breast cancer within 5 years	131 (1.28%)	58 (1.27%)
BCRAT absolute risk (%)	1.30 (0.74)	1.33 (0.73)
IBIS absolute risk (%)	1.31 (0.59)	1.33 (0.60)
Age at baseline (years)	54.1 (7.7)	54.1 (7.6)
Age categories:		
<=49	3,556 (34.9%)	1,557 (34.2%)
50-59	3,980 (39.0%)	1,839 (40.4%)
>=60	2,664 (26.1%)	1,159 (25.4%)
Birth province:		
In Canada outside Quebec	333 (3.3%)	74 (1.6%)
Outside Canada	1,490 (14.6%)	189 (4.1%)
Quebec	8,373 (82.1%)	4,292 (94.2%)
Missing	*4*	*0*
Ethnicity:		
Asian	188 (1.8%)	5 (0.1%)
Black African	182 (1.8%)	0 (0.0%)
Hispanic non-american	234 (2.3%)	1 (<0.1%)
Other	542 (5.3%)	86 (1.9%)
White/European	9,054 (88.8%)	4,463 (98.0%)
Age at menarche (years):		
<=11	2,305 (22.9%)	1,027 (22.7%)
12-13	4,754 (47.2%)	2,166 (47.9%)
>=14	3,021 (30.0%)	1,331 (29.4%)
Missing	*120*	*31*
Age at first live birth (years):		
<=19	1,124 (13.1%)	422 (11.1%)
20-24	2,955 (34.5%)	1,324 (34.8%)
25-29	2,814 (32.9%)	1,312 (34.5%)
>=30	1,621 (19.0%)	734 (19.3%)
Nulliparous	40 (0.5%)	14 (0.4%)
Missing	*1,646*	*749*
First-degree relatives with breast cancer:		
0	8,945 (87.7%)	3,949 (86.7%)
1	1,130 (11.1%)	556 (12.2%)
>=2	125 (1.23%)	50 (1.10%)
Previous breast biopsy:		
0	10,023 (98.3%)	4,463 (98.0%)

1	134 (1.31%)	71 (1.56%)
>=2	43 (0.42%)	21 (0.46%)
History of hyperplasia	6 (0.06%)	1 (0.02%)
History of atypical hyperplasia	1 (0.56%)	0 (0.00%)
Lobular Carcinoma In Situ	0	0
Weight (Kg)	67.4 [59.4;78.0]	66.8 [59.6;76.8]
Height (m)	1.61 [1.57;1.65]	1.61 [1.57;1.65]
History of ovary cancer	94 (0.92%)	46 (1.01%)
Menopause occurrence		
Pre-menopausal	4176 (40.9%)	1891 (41.5%)
Post-menopausal	5885 (57.7%)	2617 (57.5%)
Unknown	139 (1.36%)	47 (1.03%)
Use of HRT		
Never	7477 (73.3%)	3249 (71.3%)
Previous user (more than 5 years ago)	1126 (11.0%)	506 (11.1%)
Previous user (less than 5 years ago)	1285 (12.6%)	646 (14.2%)
Current user	312 (3.06%)	154 (3.38%)
HRT length of use (years)	0.00 [0.00;1.00]	0.00 [0.00;1.00]
Last HRT use (years)	0.00 [0.00;0.00]	0.00 [0.00;0.00]
Mother history of breast cancer	832 (8.16%)	412 (9.05%)
Mother history of ovary cancer	114 (1.12%)	60 (1.32%)
Father history of breast cancer	8 (0.08%)	2 (0.04%)

HRT: hormonal replacement therapy; PRS: polygenic risk score; clinical-based cohort: validation of the BCRAT and IBIS models, included women with a BCRAT and an IBIS score; clinicogenetic-based cohort: validation of the PRS models and comparison with the BCRAT and IBIS models, genotyped women with all SNPs available.
* Not available for the phase 2

As this is a validation study, the STROBE checklist is not fully adapted.

STROBE Statement—Checklist of items that should be included in reports of ***cohort studies***

	Item No	Recommendation	Page number
Title and abstract	1	(*a*) Indicate the study's design with a commonly used term in the title or the abstract	1
		(*b*) Provide in the abstract an informative and balanced summary of what was done and what was found	1-2
Introduction			
Background /rationale	2	Explain the scientific background and rationale for the investigation being reported	3-5
Objectives	3	State specific objectives, including any prespecified hypotheses	5
Methods			
Study design	4	Present key elements of study design early in the paper	5-6
Setting	5	Describe the setting, locations, and relevant dates, including periods of recruitment, exposure, follow-up, and data collection	5
Participants	6	(*a*) Give the eligibility criteria, and the sources and methods of selection of participants. Describe methods of follow-up	6
		(*b*) For matched studies, give matching criteria and number of exposed and unexposed	-
Variables	7	Clearly define all outcomes, exposures, predictors, potential confounders, and effect modifiers. Give diagnostic criteria, if applicable	7-8
Data sources/ measurement	8*	For each variable of interest, give sources of data and details of methods of assessment (measurement). Describe comparability of assessment methods if there is more than one group	7-8
Bias	9	Describe any efforts to address potential sources of bias	5
Study size	10	Explain how the study size was arrived at	6
Quantitative variables	11	Explain how quantitative variables were handled in the analyses. If applicable, describe which groupings were chosen and why	7-8
Statistical methods	12	(*a*) Describe all statistical methods, including those used to control for confounding	8-9
		(*b*) Describe any methods used to examine subgroups and interactions	-
		(*c*) Explain how missing data were addressed	7
		(*d*) If applicable, explain how loss to follow-up was addressed	8-9
		(*e*) Describe any sensitivity analyses	-
Results			
Participants	13*	(a) Report numbers of individuals at each stage of study—eg numbers potentially eligible, examined for eligibility, confirmed eligible, included in the study, completing follow-up, and analysed	9
		(b) Give reasons for non-participation at each stage	9
		(c) Consider use of a flow diagram	Figure 1
Descriptive data	14*	(a) Give characteristics of study participants (eg demographic, clinical, social) and information on exposures and potential confounders	Table S2
		(b) Indicate number of participants with missing data for each variable of	Table S2

		interest	
		(c) Summarise follow-up time (eg, average and total amount)	9
Outcome data	15*	Report numbers of outcome events or summary measures over time	9
Main results	16	(*a*) Give unadjusted estimates and, if applicable, confounder-adjusted estimates and their precision (eg, 95% confidence interval). Make clear which confounders were adjusted for and why they were included	10-12
		(*b*) Report category boundaries when continuous variables were categorized	Table S2
		(*c*) If relevant, consider translating estimates of relative risk into absolute risk for a meaningful time period	9
Other analyses	17	Report other analyses done—eg analyses of subgroups and interactions, and sensitivity analyses	10-12
Discussion			
Key results	18	Summarise key results with reference to study objectives	12
Limitations	19	Discuss limitations of the study, taking into account sources of potential bias or imprecision. Discuss both direction and magnitude of any potential bias	13
Interpretation	20	Give a cautious overall interpretation of results considering objectives, limitations, multiplicity of analyses, results from similar studies, and other relevant evidence	12-13
Generalisability	21	Discuss the generalisability (external validity) of the study results	13
Other information			
Funding	22	Give the source of funding and the role of the funders for the present study and, if applicable, for the original study on which the present article is based	21

*Give information separately for exposed and unexposed groups.

Note: An Explanation and Elaboration article discusses each checklist item and gives methodological background and published examples of transparent reporting. The STROBE checklist is best used in conjunction with this article (freely available on the Web sites of PLoS Medicine at http://www.plosmedicine.org/, Annals of Internal Medicine at http://www.annals.org/, and Epidemiology at http://www.epidem.com/). Information on the STROBE Initiative is available at http://www.strobe-statement.org.

BMJ Open

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

Journal:	*BMJ Open*
Manuscript ID	bmjopen-2020-045078.R2
Article Type:	Original research
Date Submitted by the Author:	17-Feb-2021
Complete List of Authors:	Jantzen, Rodolphe; CHU Sainte-Justine, Centre de recherche Payette, Yves; CHU Sainte-Justine Centre de Recherche de Malliard, Thibault; CHU Sainte-Justine Centre de Recherche Labbé, Catherine; CHU Sainte-Justine Centre de Recherche Noisel, Nolwenn; CHU Sainte-Justine Centre de Recherche Broët, Philippe; CHU Sainte-Justine Centre de Recherche
Primary Subject Heading:	Public health
Secondary Subject Heading:	Epidemiology, Genetics and genomics, Oncology
Keywords:	Breast tumours < ONCOLOGY, EPIDEMIOLOGY, GENETICS, PREVENTIVE MEDICINE, PUBLIC HEALTH

I, the Submitting Author has the right to grant and does grant on behalf of all authors of the Work (as defined in the below author licence), an exclusive licence and/or a non-exclusive licence for contributions from authors who are: i) UK Crown employees; ii) where BMJ has agreed a CC-BY licence shall apply, and/or iii) in accordance with the terms applicable for US Federal Government officers or employees acting as part of their official duties; on a worldwide, perpetual, irrevocable, royalty-free basis to BMJ Publishing Group Ltd ("BMJ") its licensees and where the relevant Journal is co-owned by BMJ to the co-owners of the Journal, to publish the Work in this journal and any other BMJ products and to exploit all rights, as set out in our licence.

The Submitting Author accepts and understands that any supply made under these terms is made by BMJ to the Submitting Author unless you are acting as an employee on behalf of your employer or a postgraduate student of an affiliated institution which is paying any applicable article publishing charge ("APC") for Open Access articles. Where the Submitting Author wishes to make the Work available on an Open Access basis (and intends to pay the relevant APC), the terms of reuse of such Open Access shall be governed by a Creative Commons licence – details of these licences and which Creative Commons licence will apply to this Work are set out in our licence referred to above.

Other than as permitted in any relevant BMJ Author's Self Archiving Policies, I confirm this Work has not been accepted for publication elsewhere, is not being considered for publication elsewhere and does not duplicate material already published. I confirm all authors consent to publication of this Work and authorise the granting of this licence.

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,*]

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

[2]Université de Montréal, Montréal, Québec, Canada

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

[4]Assistance Publique-Hôpitaux de Paris (AP-HP), Hôpitaux Universitaires Paris-Sud, Hôpital Paul Brousse, 12 Avenue Paul Vaillant Couturier, 94807 Villejuif, France

*** Correspondence:**

Rodolphe Jantzen, rodolphe.jantzen@gmail.com

Philippe Broët, philippe.broet@umontreal.ca

Keywords: Breast cancer, Validation Study, Clinical Decision Rules, Polygenic risk score, BCRAT, IBIS

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 five 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 4,555 with genetic information for validating the PRS and combined scores.

Outcome measures: We computed the absolute risks of breast cancer at five 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 five 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] versus 58.63 [54.05-63.21], $p=0.013$). PRS scores had a global calibration around 0.82, with a confidence interval 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 personalized message to their high risk patients, regarding screening and lifestyle counseling.

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 forty years old.

- Since the genotyping information was not available for all the cohort, the models had to be evaluated on two different sub-cohorts.

1 Introduction

Breast cancer is the most frequently diagnosed cancer and the second leading cause 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 BCRAT was developed by Dr. Mitchell Gail in 1989 using 5,998 American women from a case-control study [2]. 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 (e.g., 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 U.S. Food and Drug Administration recommended to consider chemo-prevention for women at high risk of breast cancer (i.e. a 5-year risk equal or higher than 1.66%), while the U.S. Preventive Services Task Force recommended chemo-prevention 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 U.S.-centered use led to an under-use 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 International Breast Cancer Intervention Study model (IBIS, also known as the Tyrer-Cuzick model) is also a widely used breast cancer risk prediction model, which takes into account multi-generational 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 to 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 multi-generational 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* gene mutation [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 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 ROC curve (AUC) increasing from 0.607 to 0.632 [15]. This kind of clinico-genetic 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 upon 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 (GWAS) [21], with weights that can be adapted to specific ethnicities [19]. The predictive accuracy of these PRSs compared to classical prediction models, such as the BCRAT and IBIS, should now be evaluated in various populations.

In Quebec, the Breast Cancer Screening Program consists of a mammogram every two years for women aged 50 to 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 organized 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 multi-countries cohorts. The fact that BCRAT and IBIS have not been evaluated in the French-Canadian population, which has specific genetic patterns, as compared to the general European population [25,26], with lifestyle risk factors (e.g., 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 polygenic risk scores to predict the occurrence of invasive breast cancers at five years in middle-aged and older French-Canadian women.

2 Materials and methods

2.1 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 Supplementary Methods), and an algorithm based on a previous report from the *Institut National de Santé Publique du Québec* (INSPQ) [28] and the Tonelli *et al.* algorithm [29]. 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 (i.e., lesion suspected of malignancy) without histologically confirmed breast cancers and retrieved, when available, the incidence date after the abnormal mammography from the RAMQ database for women with at least two claims in two years or one hospitalization with the appropriate International Classification of Diseases (ICD), Ninth or Tenth Revision codes

(174 and C50). The Quebec breast cancer registry's data were available from May 15th, 1998 to December 31st, 2017, while the RAMQ's data were available from January 1st, 1998 to March 31st, 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 sub-cohort 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 sub-cohort of 4,555 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.

2.2 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]. Single-nucleotide polymorphism (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 the Supplementary Methods.

2.3 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 five-years study period (administrative censoring) or at death.

2.4 Predictive scores

2.4.1 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 (United States and United Kingdom incidence rates, respectively), and the attributable risk obtained from the United States population data (BCRAT) and the United Kingdom/Swedish population data (IBIS). In this article, the BCRAT and IBIS absolute risks of breast cancer at five years were calculated for each woman at the inclusion date using the National Institutes of Health R package "BCRA", version 2.1 [33] and the latest

version of the "IBIS Breast Cancer Risk Evaluation Tool" (http://www.ems-trials.org/riskevaluator/ — version 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 Supplementary Methods for information about variables extraction and coding. Missing data can be handled in both BCRAT and IBIS models.

2.4.2 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 clinicogenetic-based cohort 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 odds ratio can be found as Supplementary Table S1.

2.4.3 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".

2.4.4 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 version 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]).

2.5 Statistical analysis

For comparing means between groups, we used a one-way ANOVA test. Relationships between categorical variables were tested using the χ^2 test. Statistical significance was considered as P-values less than 0.05. We plotted predictiveness curves (i.e., 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.

2.5.1 Calibration

We computed the expected-to-observed ratio (E/O), with the 95% confidence interval (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), ≥1% and <1.66% (intermediate risk), ≥1.66% and <3% (average risk), ≥3% (high risk). The observed proportion at five-year 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 ($G = \sum(Oi-Ei)^2/Ei$) where Oi and Ei are respectively the observed and expected number of events in group i, and compared this latter to the critical value from the chi-squared distribution with four degrees of freedom.

2.5.2 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]. Receiver operating characteristic (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].

2.5.3 Sensitivity and specificity

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

All statistical analyses were performed using R software, version 3.6 [40].

2.6 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.

3 Results

Overall, 10,200 women were included for validating the BCRAT and IBIS scores and 4,555 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 five-year follow-up, for the CC and CGC, respectively. The clinical characteristics of the two cohorts can be found in the Supplementary Table S2.

3.1 Breast cancer risk prediction models (BCRAT and IBIS) evaluated in the clinical-based cohort

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.

3.2 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 sub-cohort, 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 expected/observed ratio 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 (Supplementary Figure S1). 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 zero and one, 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 to BCRAT and IBIS scores (down-left corner of the ROC curves, Supplementary Figure S2). Using a 1.66% threshold, all PRS scores increased both the sensitivity and the specificity as compared to the BCRAT and IBIS risk prediction score (Table 1).

The distribution of the combined models' absolute risks can be found in the Supplementary Figure S1. 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 zero and one, 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 for the Shieh and Mavaddat combined scores (down-left corner of the ROC curves, Supplementary Figure S2). Using a

1.66% threshold, only the Evans combined model increased both the sensitivity and the specificity as compared to 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 zero and one, 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 and Mavaddat combined scores (down-left corner of the ROC curves, Supplementary Figure S2). Compared to the BCRAT and IBIS models, sensitivities values were higher while specificities values were lower (Table 2).

4 Discussion

In this work, we reported the predictive performance of BCRAT, IBIS and four polygenic risk scores for predicting breast cancer occurrence within five 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 and intermediate risk groups (below 1.66% risk). In our study, IBIS produced slightly better discrimination than BCRAT. As compared to 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 over-prediction 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-standardized incidence rates since, based on information retrieved from national cancer databases [35,44,45], the incidence rates are comparable between the United States and

Canada (250.4 [95%CI 209.0-298.3] cases per 100,000 per year for Canada and 236.8 [95%CI 235.5-238.1] for US). The IBIS model, the PRS models and the clinico-genetic model (BCRAT+PRS) had also an overall good mean calibration. However, when analyzing calibration across risk subgroups, the IBIS model had a significant goodness of fit test, with an over and underestimated the risk in the low and intermediate groups, respectively, probably explained by the United-Kingdom 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 to 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 emphasizes 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 clinico-genetic 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 to the BCRAT+PRS models. However, the absolute risk of IBIS combined models were not obtained 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. Firstly, 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. Secondly and to the best of our knowledge, this study is the first to evaluate the breast cancer risk assessment tools in a French-Canadian population for predicting breast cancer at five years.

This study has nevertheless some limitations. Firstly, our findings may not apply to younger women under forty years old. Secondly, 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 chemo-prophylaxis in the US [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. Thirdly, 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 to BCRAT between the two cohorts might be linked to the smaller size of the clinicogenetic-based cohort as compared to the clinic-based cohort. The ethnicity differences between the two sub-cohorts 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 sub-cohorts 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.

4.1 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 personalized message to their high risk patients, regarding screening and lifestyle counseling.

5 Tables

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 [0.73-1.22] 0.94 [0.73-1.22]	0.83 [0.65-1.08]	0.81 [0.63-1.05]	0.82 [0.63-1.06]	0.81 [0.62-1.04]
Goodness of fit	p=0.0415 p=268	p=0.0984	p=0.1009	p=0.1992	p=0.2770
Intercept	-2 [-4.4 - 0.2] -0.8 [-3.4 - 1.8]	- 0.3 [-2.4 - 1.8]	-1 [-2.5 - 0.5]	1 [-1.6 - 3.6]	0.9 [-1.8 - 3.5]
Slope	0.5 [0 - 1] 0.8 [0.2 - 1.4]	0.9 [0.4 - 1.4]	0.7 [0.4 - 1.1]	1.2 [0.6 - 1.8]	1.1 [0.5 - 1.7]
C-index	59.13 [52.96-65.29] 59.63 [53.26-66]	60.77 [53-68.53]	62.56 [54.54-70.59]	63.4 [56.65-70.16]	64.27 [58.09-70.44]
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% [11.2-33.4] 24.1% [13.9-37.2]	31% [19.5-44.5]	39.7% [27-53.4]	34.5% [22.5-48.1]	25.9% [15.3-39]
Specificity *	79% [77.7-80.3] 81.6% [80.4-82.8]	82.2% [81-83.4]	81.3% [80.1-82.5]	85.4% [84.2-86.4]	86.7% [85.6-87.7]

BCRAT: Breast Cancer Risk Assessment Tool; E/O: expected-to-observed ratio; IBIS: International Breast Cancer Intervention Study risk evaluation tool

Clinicogenetic-based validation cohort: validation of the PRS models and comparison with the BCRAT and IBIS models, genotyped women with all SNPs available. 10% of the cohort was used to obtain the relative risk distribution while the remained 90% was used for computing the results.

95% confidence intervals in square brackets

* 1.66% threshold

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

	BCRAT model / IBIS model	Combined scores			
		Mavaddat with BCRAT / with IBIS	Shieh with BCRAT / with IBIS	Evans with BCRAT / with IBIS	Wacholder with BCRAT / with IBIS
E/O	0.94 [0.73-1.22]	0.86 [0.66-1.11]	0.83 [0.64-1.07]	0.83 [0.64-1.08]	0.82 [0.64-1.06]
	0.94 [0.73-1.22]	0.95 [0.73-1.22]	0.94 [0.73-1.22]	0.94 [0.73-1.22]	0.94 [0.73-1.22]
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 - 0.2]	- 1.5 [-3.3 - 0.1]	-1.6 [-3 - -0.3]	-1.2 [-3.1 - 0.6]	-1.3 [-3.2 - 0.5]
	-0.8 [-3.4 - 1.8]	-0.9 [-2.7 - 0.8]	-1.3 [-2.7 - 0]	-0.3 [-2.3 - 1.7]	-0.5 [-2.5 - 1.5]
Slope	0.5 [0 - 1]	0.6 [0.2 - 1]	0.6 [0.3 - 0.9]	0.7 [0.3 - 1.1]	0.7 [0.2 - 1.1]
	0.8 [0.2 - 1.4]	0.8 [0.4 - 1.2]	0.7 [0.3 - 1]	0.9 [0.4 - 1.4]	0.9 [0.4 - 1.3]
C-index	59.13 [52.96-65.29]	61.42 [54.05-68.78]	63.35 [55.58-71.12]	62.69 [55.88-69.50]	63.58 [57.46-69.69]
	59.63 [53.26-66]	62.73 [55.34-70.12]	63.83 [56.27-71.39]	63.35 [56.44-70.26]	64.21 [57.88-70.54]
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	-	p=0.66	p=0.42	p=0.38	p=0.22

comparison with IBIS model	-	p=0.393	p=0.316	p=0.169	p=0.080
Sensitivity *	20.7% [11.2-33.4]	36.2% [24-49.9]	37.9% [25.5-51.6]	25.9% [15.3-39]	22.4% [12.5-35.3]
	24.1% [13.9-37.2]	36.2% [24-49.9]	44.8% [31.7-58.5]	41.4% [28.6-55.1]	37.9% [25.5-51.6]
Specificity *	79% [77.7-80.3]	80.5% [79.2-81.7]	81.5% [80.2-82.7]	82.1% [80.9-83.3]	83.8% [82.6-84.9]
	81.6% [80.4-82.8]	76.6% [75.2-77.9]	76.9% [75.5-78.2]	76.9% [75.6-78.2]	77.9% [76.6-79.2]

BCRAT: Breast Cancer Risk Assessment Tool; E/O: expected-to-observed ratio; IBIS: International Breast Cancer Intervention Study risk evaluation tool

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. 10% of the cohort was used to obtain the relative risk distribution while the remained 90% was used for computing the results.

95% confidence intervals in square brackets

* 1.66% threshold

6 Contributorship statement

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

7 Competing Interests

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

8 Funding

This research received no specific grant from any funding agency in the public, commercial or not-for-profit sectors.

9 Data Sharing Statement

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).

10 Figures caption

Figure 1 Flow-chart

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 to the critical value from the chi-squared distribution. E/O: expected-to-observed cases. (C) Discrimination power of the models according to sensitivity and specificity. C-index was calculated using the Inverse Probability of Censoring Weighting (IPCW) estimation of cumulative time-dependent ROC curve.

Figure 3 **Calibration according to BCRAT, IBIS and PRS scores' predictions groups.** PRS: polygenic risk score. E/O: expected-to-observed cases. Results from the clinicogenetic-based cohort. P values were computed using a goodness of fit test statistic compared to the critical value from the chi-squared distribution. Each PRS models name referred to the first author of the study from which the PRS were derived.

Figure 4 **Calibration according to BCRAT, IBIS and combined models' predictions groups.** PRS: polygenic risk score. E/O: expected-to-observed cases. Results from the clinicogenetic-based cohort. P values were computed using a goodness of fit test statistic compared to the critical value from the chi-squared distribution. Each PRS models name referred to the first author of the study from which the PRS were derived.

11 References

1 Canadian Cancer Statistics Advisory Committee. Canadian Cancer Statistics 2019. *Toronto, ON: Canadian Cancer Society* Published Online First: September 2019.cancer.ca/Canadian-Cancer-Statistics-2019-EN (accessed 13 Nov 2019).

2 Gail MH, Brinton LA, Byar DP, *et al.* Projecting Individualized Probabilities of Developing Breast Cancer for White Females Who Are Being Examined Annually. *JNCI Journal of the National Cancer Institute* 1989;**81**:1879–86. doi:10/djr43s

3 Tyrer J, Duffy SW, Cuzick J. A breast cancer prediction model incorporating familial and personal risk factors. *Statist Med* 2004;**23**:1111–30. doi:10.1002/sim.1668

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. doi:10/ckg9vc

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. doi:10/btbjf9

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. doi:10/dkmv6x

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**. doi:10/gc83rw

8 Farkas A, Vanderberg R, Merriam S, *et al.* Breast Cancer Chemoprevention: A Practical Guide for the Primary Care Provider. *Journal of Women's Health* 2020;**29**:46–56. doi:10.1089/jwh.2018.7643

9 Levine M, Moutquin J-M, 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. www.cancer.ca. http://www.cancer.ca/en/cancer-information/cancer-type/breast/risks/reducing-your-risk/?region=qc (accessed 28 Nov 2019).

11 Voruganti RT, 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. *BMJ Health & Care Informatics* 2015;**22**:333–9. doi:10/ggt2c8

12 Terry MB, Liao Y, Whittemore AS, *et al.* 10-year performance of four models of breast cancer risk: a validation study. *The Lancet Oncology* 2019;**20**:504–17. doi:10/ggt2dp

13 Gail MH. Performance of BCRAT in high-risk patients with breast cancer. *The Lancet Oncology* 2019;**20**:e285. doi:10.1016/S1470-2045(19)30301-8

14 Terry MB, Liao Y, Hopper JL, *et al.* Performance of BCRAT in high-risk patients with breast cancer – Authors' reply. *The Lancet Oncology* 2019;**20**:e286. doi:10.1016/S1470-2045(19)30311-0

15 Gail MH. Discriminatory Accuracy From Single-Nucleotide Polymorphisms in Models to Predict Breast Cancer Risk. *J Natl Cancer Inst* 2008;**100**:1037–41. doi:10/ff2x2r

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. doi:10.1158/1055-9965.EPI-15-0838

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. doi:10/fn7zxw

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/f3nv95

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. doi:10/ggt2dm

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. doi:10/f9nd2w

21 MacArthur J, Bowler E, Cerezo M, *et al.* The new NHGRI-EBI Catalog of published genome-wide association studies (GWAS Catalog). *Nucleic Acids Research* 2017;**45**:D896. doi:10/f9v7cp

22 Gouvernement du Québec. Québec Breast Cancer Screening Program. 2018.https://www.quebec.ca/en/health/advice-and-prevention/screening-and-carrier-testing-offer/quebec-breast-cancer-screening-program/ (accessed 9 Mar 2020).

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 (Oxford)* 2004;**43**:1178–81. doi:10/bq2cx6

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. doi:10.1371/journal.pgen.1003815

26 Scriver CR. Human Genetics: Lessons from Quebec Population. *Annu Rev Genom Hum Genet* 2001;**2**:69–101. doi:10.1146/annurev.genom.2.1.69

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. *International Journal of Epidemiology* 2013;**42**:1285–99. doi:10.1093/ije/dys160

28 Théberge I, Institut national de santé publique du Québec, Direction systèmes de soins et services. *Validation de stratégies pour obtenir le taux de détection du cancer, la valeur prédictive 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 données fournies par le programme québécois de dépistage du cancer du sein (PQDCS)*. Montréal: : Direction des systèmes de soins et services, Institut national de santé publique 2003.

29 Tonelli M, Wiebe N, Fortin M, *et al.* Methods for identifying 30 chronic conditions: application to administrative data. *BMC Medical Informatics and Decision Making* 2015;**15**:31. doi:10/f698hd

30 Akçimen F, Ross JP, Sarayloo F, *et al.* Genetic and epidemiological characterization of restless legs syndrome in Québec. *Sleep* 2019;**43**. doi:10.1093/sleep/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. doi:10/ggt2b8

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. doi:10/ggrr3s

33 Zhang F. *BCRA: breast cancer risk assessment*. 2018. https://CRAN.R-project.org/package=BCRA

34 Choudhury PP, Maas P, Wilcox A, *et al.* iCARE: An R package to build, validate and apply absolute risk models. *PLOS ONE* 2020;**15**:e0228198. doi:10/ggt2ch

35 Statistics Canada. Table 13-10-0111-01 Number and rates of new cases of primary cancer, by cancer type, age group and sex. doi.org/10.25318/1310011101-eng (accessed 10 Nov 2019).

36 Statistics Canada. Table 13-10-0392-01 Deaths and age-specific mortality rates, by selected grouped causes. https://doi.org/10.25318/1310039201-eng (accessed 10 Nov 2019).

37 Uno H, Cai T, Tian L, *et al.* Evaluating Prediction Rules for t-Year Survivors with Censored Regression Models. *Journal of the American Statistical Association* 2007;**102**:527–37. doi:10.1198/016214507000000149

38 Blanche P, Latouche A, Viallon V. Time-dependent AUC with right-censored data: a survey study. *arXiv:12106805 [statME]* Published Online First: 25 October 2012.http://arxiv.org/abs/1210.6805 (accessed 26 Sep 2017).

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. *Statist Med* 2013;**32**:5381–97. doi:10.1002/sim.5958

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. *JNCIJ* 2016;**108**:djv348. doi:10/ggt2db

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. doi:10/ggt2dc

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. doi:10/ggt2c4

44 Statistics Canada. Table 17-10-0005-01 Population estimates on July 1st, by age and sex. doi: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.www.cdc.gov/cancer/dataviz (accessed 9 Oct 2019).

46 Moyer VA. Medications for Risk Reduction of Primary Breast Cancer in Women: U.S. Preventive Services Task Force Recommendation Statement. *Ann Intern Med* Published Online First: 24 September 2013. doi:10/ggt2cx

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. doi:10/ggt2c3

12 Ethics approval and consent to participate

This project has been approved by the Research Ethics Board of the CHU Sainte-Justine under the reference 2020-2427. In addition, CARTaGENE has obtained ethics approval by the CHU Sainte-Justine under the reference: MP-21-2011-345, 3297. The latest annual ethics renewal was granted on September 13, 2019. This latter approval implies that all participants have given their consent (cartagene.qc.ca/sites/default/files/documents/consent/brochure_en_0505_0.pdf). Consent was obtained from all the participants.

13 Acknowledgments

We would like to thank all the CARTaGENE participants for their generous investments in health research. We would also like to thank the RAMQ and the Commission d'accès à l'information (CAI) for their support in obtaining the data.

14 Supplementary Material

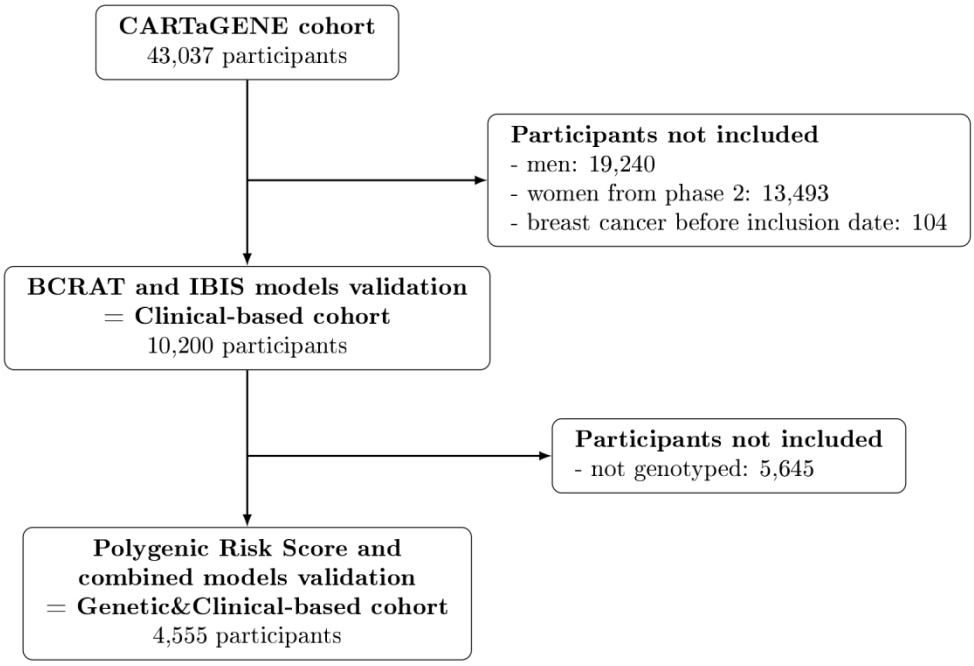
Supplementary Methods

Supplementary Table S1: SNPs used for each extended model and the associated gene and odds ratio.

Supplementary Table S2: Characteristics comparison of the women from the Clinical-based and the clinicogenetic-based cohorts.

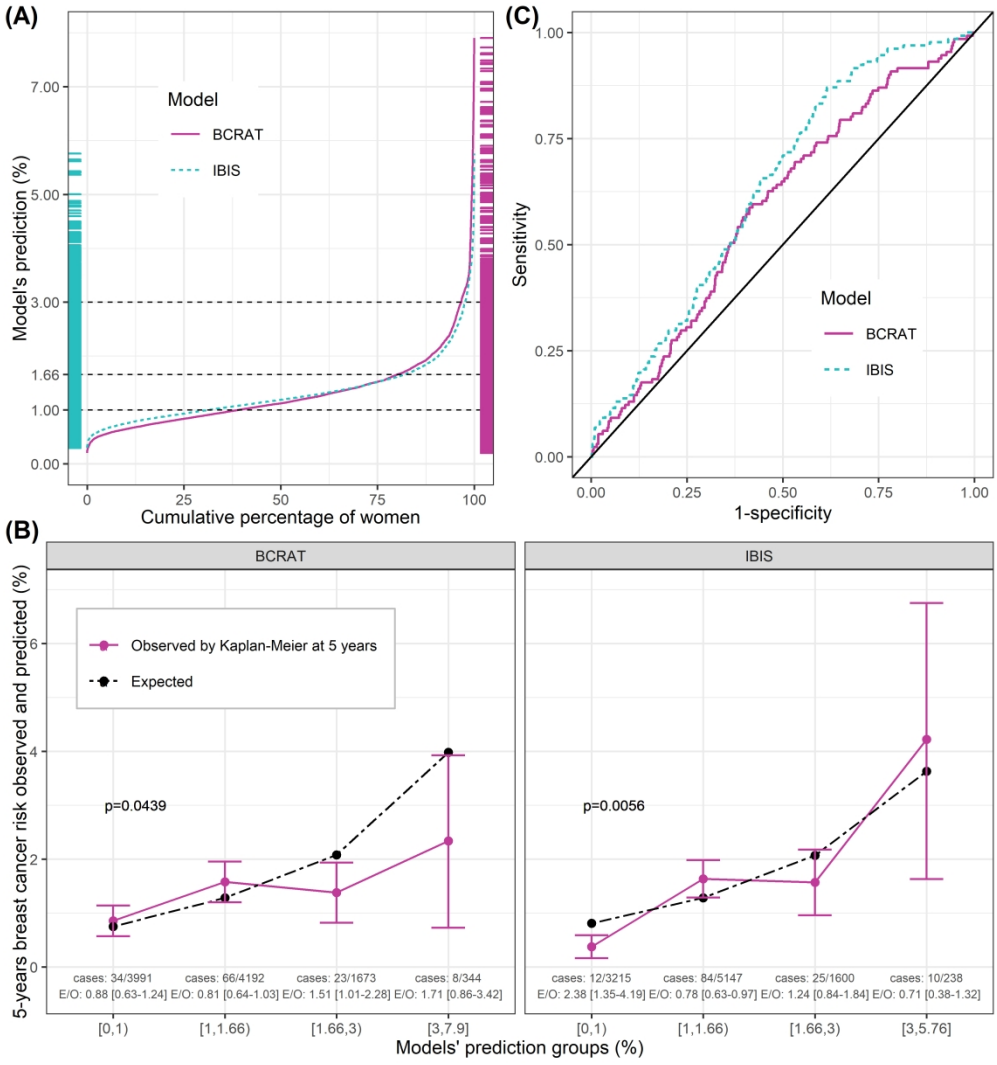
Supplementary Figure S1: Distribution of BCRAT, IBIS, PRS and combined scores predictions as a function of cumulative percentage of women. Results from the clinicogenetic-based cohort.

Supplementary Figure S2: Discrimination power of BCRAT, IBIS, PRS scores and combined models according to sensitivity and specificity. Results from the clinicogenetic-based cohort. C-indexes were calculated using the Inverse Probability of Censoring Weighting (IPCW) estimation of cumulative time-dependent ROC curve. Each PRS models name referred to the first author of the study from which the PRS were derived.



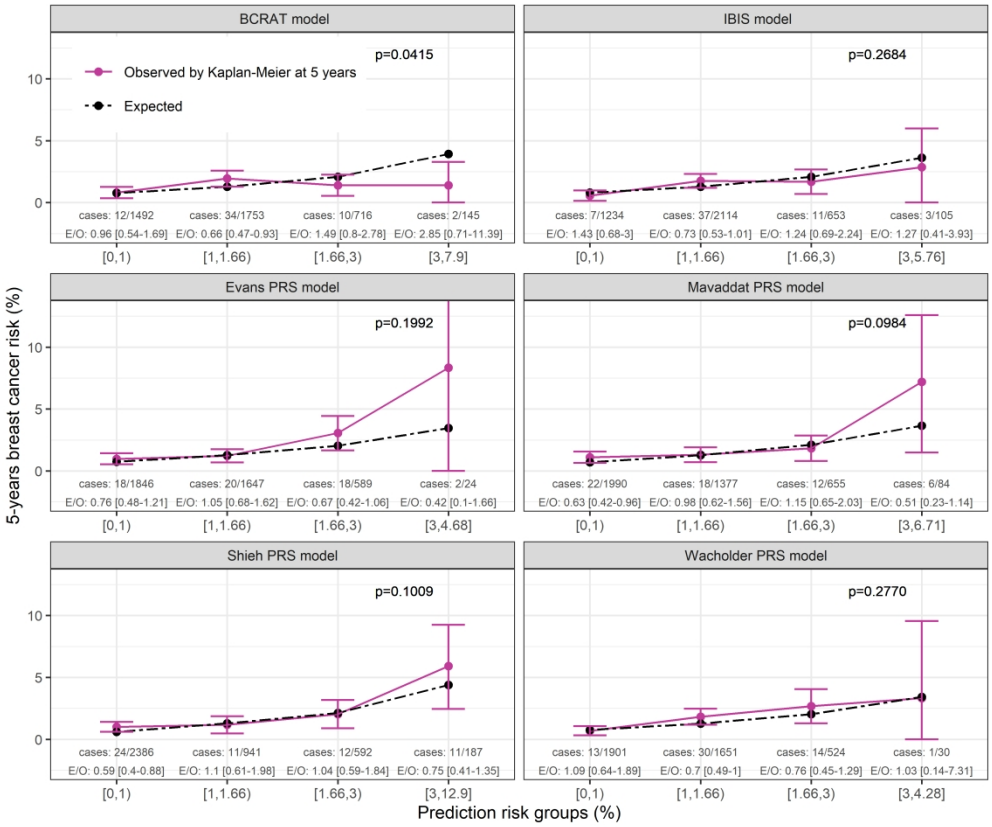
Flow-Chart

145x96mm (300 x 300 DPI)



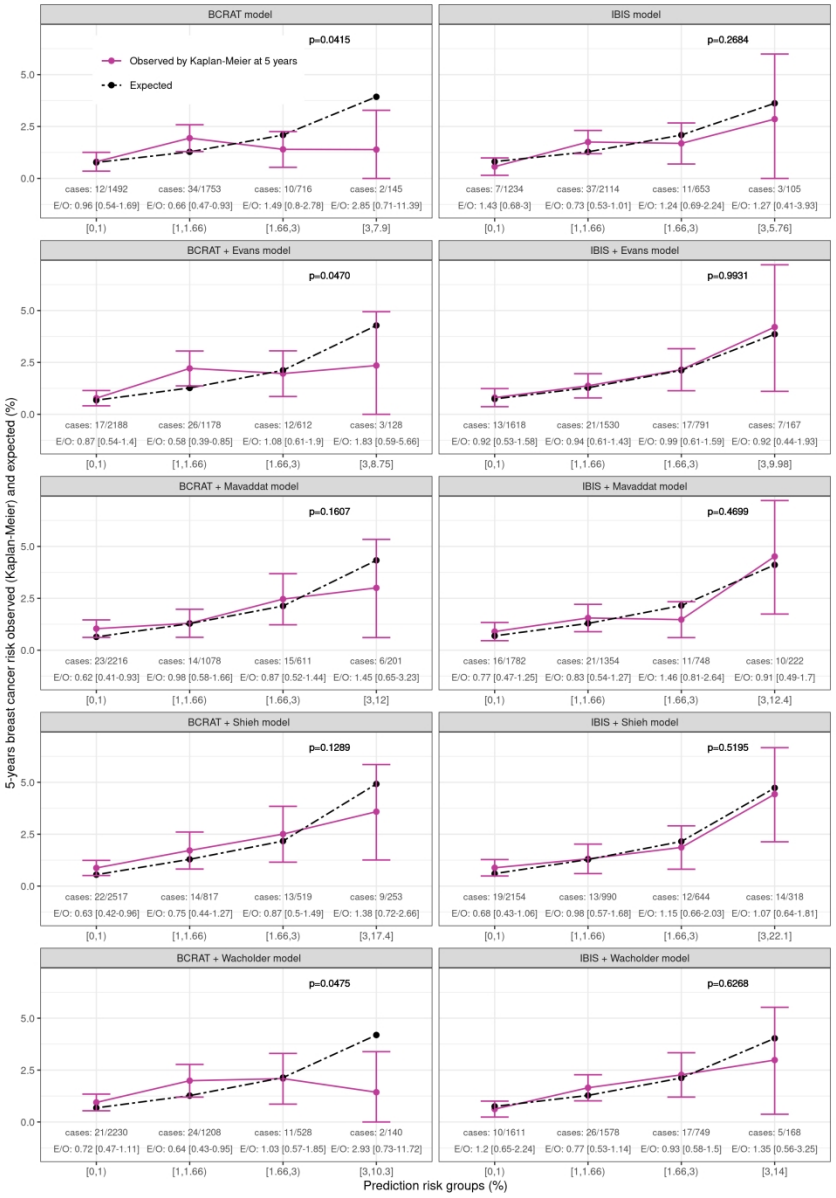
Absolute risk distribution and performance of the BCRAT and IBIS models in the Clinical-based cohort. (A) Distribution of models' predictions as a function of cumulative percentage of women. Rug plot on the y-axis. (B) Calibration according to the models' predictions groups. P values were computed using a goodness of fit test statistic compared to the critical value from the chi-squared distribution. E/O: expected-to-observed cases. (C) Discrimination power of the models according to sensitivity and specificity. C-index was calculated using the Inverse Probability of Censoring Weighting (IPCW) estimation of cumulative time-dependent ROC curve.

169x179mm (600 x 600 DPI)



Calibration according to BCRAT, IBIS and PRS scores' predictions groups. PRS: polygenic risk score. E/O: expected-to-observed cases. Results from the clinicogenetic-based cohort. P values were computed using a goodness of fit test statistic compared to the critical value from the chi-squared distribution. Each PRS models name referred to the first author of the study from which the PRS were derived.

179x149mm (600 x 600 DPI)



Calibration according to BCRAT, IBIS and combined models' predictions groups.

PRS: polygenic risk score. E/O: expected-to-observed cases. Results from the clinicogenetic-based cohort. P values were computed using a goodness of fit test statistic compared to the critical value from the chi-squared distribution. Each PRS models name referred to the first author of the study from which the PRS were derived.

209x299mm (300 x 300 DPI)

Supplementary Methods

Health databases

For identifying participants who had breast cancer, we used two administrative health databases (AHD): 1) the MED-ÉCHO AHD: this database contains all the Quebec Health Insurance Board (RAMQ) diagnoses, hospitalizations and physician claims of insured patients (about 98% of Quebec residents [1]), excluding private healthcare; in the case of cancers, all patients are treated in the public sector. Data were available from January 1st, 1998 to March 31st, 2016. Dates of death were also retrieved from the RAMQ; 2) the Quebec Breast Cancer Registry: it contains information about the Quebec Breast Cancer Screening Program, such as mammograms' results and breast cancers histological confirmation. Data were available from May 15th, 1998 to December 31st, 2017.

References

1 RAMQ. Table PA.01 - Nombre de personnes inscrites et admissibles au régime d'assurance maladie du Québec selon le sexe, le groupe d'âge et la région sociosanitaire. 2017.https://www4.prod.ramq.gouv.qc.ca/IST/CD/CDF_DifsnInfoStats/CDF1_CnsulInfoStatsCNC_iut/DifsnInfoStats.aspx?ETAPE_COUR=3&IdPatronRapp=8&Annee=2017&Per=0&LANGUE=en-CA (accessed 25 Nov 2019).

Genetic data

Genotypes were included in the CaG database and were obtained from hybridation upon three different chips: Illumina Omni 2.5M (7.7% of the participants), Affymetrix Axiom UK biobank (8.2%) and Illumina Infinium Global Screening Array (84.1%). A quality control (QC) was made before the imputation (detailed pipeline can be found at www.cartagene.qc.ca/info-genetic-data): 1) QC sample: for replicated samples, samples with the lowest call rates were removed. Sample with a call rate below 95% were removed. Samples pairs with an identity by state (IBS) higher than 0.20 and similar to at least 50% of the whole set were removed. Then, for pair of samples with an IBS higher than 0.85, when the correct sample could not be identified with certainty, both samples of the pair were removed. Samples with discrepancy between sex chromosome genotypes and reported

gender were removed. 2) QC SNP: SNPs with a call rate lower than 95% or deviating from Hardy–Weinberg equilibrium (with a 10-6 threshold) were removed.

For the imputation, data were prepared using the Will Rayner toolbox (www.well.ox.ac.uk/~wrayner/tools/) with the Haplotype Reference Consortium (HRC) as reference panel [1]. To impute missing SNPs of our cohort, we used the Michigan Imputation Server with the Minimac4 algorithm [2], with separate chromosomes and chips. Imputation reference panel was the HRC r1.1 2016 European population, and the phasing was made with Eagle v2.4 [3]. A total of 39,131,578 SNPs were retrieved.

After imputation and after merging chromosomes, we used men and women to perform a sample QC based on the Anderson *et al.* protocole [4]: samples with a call rate lower than 95% and an heterozygosity higher than 3 standard deviation were removed. After LD pruning (window size: 50kb; step size: 5 variants; pairwise r^2 threshold: 0.2), for pair of participants with an IBS higher than 0.1875, the sample with the lowest call rate was removed. To remove samples with divergent ancestries, we used the two first principal components with the HapMap phase III reference panel. As we would like to have all SNPs available for calculating PRS, we did not perform an additional SNPs QC. QC process was performed using PLINK v1.90b6.2 and v2.00a2LM 64-bit ([5,6]; URL: pngu.mgh.harvard.edu/purcell/plink/).

References

1. the Haplotype Reference Consortium, McCarthy S, Das S, Kretzschmar W, Delaneau O, Wood AR, et al. A reference panel of 64,976 haplotypes for genotype imputation. Nat Genet. 2016 août;48:1279.
2. Das S, Forer L, Schönherr S, Sidore C, Locke AE, Kwong A, et al. Next-generation genotype imputation service and methods. Nat Genet. 2016;48(10):1284–7.
3. Loh P-R, Danecek P, Palamara PF, Fuchsberger C, A Reshef Y, K Finucane H, et al. Reference-based phasing using the Haplotype Reference Consortium panel. Nat Genet. 2016;48(11):1443–8.
4. Anderson CA, Pettersson FH, Clarke GM, Cardon LR, Morris AP, Zondervan KT. Data quality control in genetic case-control association studies. Nat Protoc. 2010 Sep;5(9):1564–73.

5. Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MAR, Bender D, et al. PLINK: a tool set for whole-genome association and population-based linkage analyses. Am J Hum Genet. 2007 Sep;81(3):559–75.
6. Chang CC, Chow CC, Tellier LC, Vattikuti S, Purcell SM, Lee JJ. Second-generation PLINK: rising to the challenge of larger and richer datasets. GigaScience. 2015 Dec;4(1):7.

Absolute risk of breast cancer

The absolute risk of breast cancer over an established period $[t_0,t_1]$ (five years in this study) is the probability that a woman who is free of a breast cancer at age t_0 and has a risk score S will be diagnosed with breast cancer over the period $[t_0,t_1]$.

Under the assumption of a multiplicative proportional hazard model (or Cox model), this latter conditional probability (denoted $AR(t_0,t_1;S)$) can be written such as:

$$AR(t_0,t_1;S)=\int_{t_0}^{t_1}\lambda_0(t)e^S\exp\left[-\int_{t_\square}^{t_0}\lambda_0(u)e^S+\gamma(u)du\right]dt$$

where $\lambda_0(t)$ and $\gamma(t)$ are the baseline age-specific hazard rate for breast cancer and the age-specific mortality hazard rate from other causes (competing risks), respectively. In practice, the absolute risk is computed using piece-wise constant hazard rates.

These baseline hazard rates are calculated using marginal (or composite) hazard rates obtained from registries, together with either the attributable hazard function or the risk factor distribution.

In this work, the timescale of the analyses was age of an individual so that t_0 was the age of a woman at entry into the cohort and t_1 was the age five years later.

For the IBIS model, the baseline age-specific hazard rate for breast cancer is replaced by a hazard rate estimate obtained from the segregation model conditionally on the woman's family history.

Variables extraction and coding

Age at inclusion was calculated using the birthdate. We retrieved from the CARTaGENE questionnaire the first menstrual period, first live birth, number of first-degree relatives with breast cancer, ethnicity, menopause occurrence and age at menopause, height, weight, hormonal replacement therapy (HRT) use, length of HRT and last HRT use. If first menstrual period occurred after first live birth, both were considered as missing. We retrieved from the Quebec Breast Cancer Registry the previous breast biopsy and the number of biopsy with hyperplasia, atypical hyperplasia and lobular carcinoma *in situ*. We retrieved from the RAMQ the occurrence and age of ovary cancers.

How the variables were coded for the IBIS model can be found online (https://ems-trials.org/riskevaluator/), in the Documentation section, file “Risk program input file format (v6-8)”

Supplementary Table S1

snp	id*	genes	wacholder**	evans	mavaddat	shieh	shieh_asian***
rs13387042	2:217905832	AC007749.1 - RN7SKP43	0.8/0.7	0,88	0,88	0,88	1,06
rs1045485	2:202149589	-	0.89/0.69	-	0,96	-	-
rs999737	14:69034682	RAD51B	0.91/0.67	-	0,92	0,92	0,93
rs3817198	11:1909006	LSP1	1.04/1.18	-	1,07	1,07	1,07
rs889312	5:56031884	C5orf67 - AC008940.1	1.05/1.1	1,12	1,12	1,12	1,05
rs7716600	5:44875005	AC093297.2 - AC114954.1	1.11/1.46	-	-	-	-
rs13281615	8:128355618	CASC8, POU5F1B, PCAT1	1.14/1.36	-	1,09	1,09	1,03
rs3803662	16:52586341	CASC16	1.16/1.44	1,23	1,23	1,24	1,15
rs2981582	10:123352317	FGFR2	1.18/1.6	-	-	-	-
rs11249433	1:121280613	EMBP1	1.23/1.3	1,09	1,10	1,09	1,16
rs10995190	10:64278682	AC024598.1, ZNF365		0,86	0,86	0,86	0,94
rs1562430	8:128387852	POU5F1B, CASC8, PCAT1		0,90	-	1,16	1,16
rs909116	11:1941946	TNNT3		0,93	-	-	-
rs1156287	17:53076799	-		0,93	-	-	-
rs713588	10:5886962	-		1,01	-	-	-
rs8009944	14:69039588	-		1,04	-	-	-
rs10931936	2:202143928	-		1,04	-	-	-
rs1011970	9:22062134	CDKN2B-AS1		1,05	1,05	1,06	1,06
rs704010	10:80841148	ZMIZ1		1,09	1,07	1,08	1,05
rs4973768	3:27416013	SLC4A7		1,09	1,09	1,10	1,11
rs9790879	5:44899885	-		1,09	-	-	-
rs3757318	6:151914113	CCDC170		1,16	-	1,16	1,16
rs614367	11:69328764	LINC01488 - CCND1		1,21	-	1,21	1,29
rs2981579	10:123337335	FGFR2		1,27	1,25	1,27	1,27
rs10771399	12:28155080	PTHLH - CCDC91		-	0,86	0,86	1,15
rs865686	9:110888478	CHCHD4P2 - AL353742.1		-	0,90	0,89	1,04
rs6828523	4:175846426	ADAM29		-	0,91	0,90	1,11
rs17356907	12:96027759	PGAM1P5		-	0,91	0,91	1,08
rs6472903	8:76230301	CASC9		-	0,91	0,91	1,16
rs4849887	2:121245122	LINC01101 - AC073257.2		-	0,92	0,91	1,07
rs1353747	5:58337481	AC092343.1, PDE4D		-	0,92	0,92	1,00
rs1292011	12:115836522	AC078880.2 - AC009803.2		-	0,92	0,92	1,11
rs2236007	14:37132769	PAX9		-	0,92	0,93	1,09
rs2823093	21:16520832	AF127577.5 - AF246928.1		-	0,93	0,92	1,08
rs17817449	16:53813367	FTO		-	0,93	0,93	1,09
rs6504950	17:53056471	STXBP4		-	0,93	0,94	1,02
rs4808801	19:18571141	ELL		-	0,93	1,08	1,04
rs2736108	5:1297488	TERT - MIR4457		-	0,94	0,94	0,94
rs11242675	6:1318878	FOXQ1 - LINC01394		-	0,94	0,94	0,99
rs616488	1:10566215	PEX14		-	0,94	0,94	1,06
rs11199914	10:123093901	LINC01153 - RN7SKP167		-	0,94	0,95	1,03
rs3903072	11:65583066	AP001266.1 - CFL1		-	0,94	0,95	1,05
rs1550623	2:174212894	AC092573.2		-	0,94	1,06	1,21
rs720475	7:144074929	ARHGEF5		-	0,95	0,94	1,02
rs1436904	18:24570667	CHST9, AQP4-AS1		-	0,95	0,96	1,02
rs2016394	2:172972971	DLX2-DT		-	0,95	-	-
rs527616	18:24337424	AQP4-AS1		-	0,96	0,95	1,03
rs11820646	11:129461171	AP003500.2		-	0,96	0,95	1,05

Supplementary Table S1

rs2380205	10:5886734	GDI2 - ANKRD16	-	0,98	0,94	1,02
rs6678914	1:202187176	LGR6	-	0,99	0,91	1,10
rs10069690	5:1279790	TERT	-	1,02	1,06	1,05
rs75915166	11:69379161	LINC01488 - CCND1	-	1,02	1,31	1,00
rs12422552	12:14413931	GNAI2P1 - RPL30P11	-	1,03	1,05	1,05
rs4245739	1:204518842	MDM4	-	1,03	1,14	1,14
rs8170	19:17389704	USHBP1, AC010463.1, BABAM1	-	1,03	1,15	1,00
rs2363956	19:17394124	ANKLE1	-	1,03	-	-
rs10472076	5:58184061	AC008852.1 - PDE4D	-	1,04	1,05	1,02
rs12710696	2:19320803	LINC01376	-	1,04	1,10	1,10
rs11075995	16:53855291	FTO	-	1,04	1,11	1,11
rs7726159	5:1282319	TERT	-	1,04	-	-
rs9790517	4:106084778	TET2	-	1,05	1,05	1,02
rs204247	6:13722523	RANBP9 - MCUR1	-	1,05	1,05	1,03
rs10759243	9:110306115	PPIAP88 - RNU6-996P	-	1,05	1,06	1,05
rs12493607	3:30682939	TGFBR2	-	1,05	1,06	1,05
rs2046210	6:151948366	CCDC170 - ESR1	-	1,05	1,15	1,27
rs17529111	6:82128386	AL590824.1 - TENT5A	-	1,05	-	-
rs7904519	10:114773927	TCF7L2	-	1,06	1,06	1,02
rs3760982	19:44286513	KCNN4 - LYPD5	-	1,06	1,06	1,02
rs941764	14:91841069	CCDC88C	-	1,06	1,06	1,05
rs7072776	10:22032942	MLLT10 - DNAJC1	-	1,06	1,07	1,04
rs11780156	8:129194641	PVT1	-	1,07	1,07	1,00
rs6762644	3:4742276	ITPR1	-	1,07	1,07	1,03
rs9693444	8:29509616	RPL17P33 - LINC00589	-	1,07	1,07	1,08
rs1432679	5:158244083	EBF1	-	1,07	1,07	1,09
rs2588809	14:68660428	RAD51B	-	1,07	1,08	1,06
rs16857609	2:218296508	DIRC3	-	1,07	1,08	1,07
rs11552449	1:114448389	DCLRE1B	-	1,08	1,07	1,03
rs13329835	16:80650805	CDYL2	-	1,08	1,08	1,02
rs132390	22:29621477	EMID1	-	1,11	1,12	1,00
rs10941679	5:44706498	AC093292.1 - RN7SL383P	-	1,12	1,13	1,08
rs554219	11:69331642	LINC01488 - CCND1	-	1,12	1,27	1,00
rs6001930	22:40876234	MRTFA	-	1,13	1,12	1,03
rs2943559	8:76417937	HNF4G	-	1,13	1,13	0,96
rs12662670	6:151918856	CCDC170	-	1,14	-	-
rs78540526	11:69331418	LINC01488 - CCND1	-	1,18	-	-
rs11814448	10:22315843	DNAJC1 - ADIPOR1P1	-	1,22	1,26	1,08
rs11571833	13:32972626	BRCA2	-	1,26	1,26	1,00
rs17879961	22:29121087	CHEK2	-	1,36	1,36	1,00
rs140068132	6:151954834	CCDC170 - ESR1	-	-	0,60	1,00
rs10822013	10:64251977	AC024598.1, ZNF365	-	-	0,89	1,08
rs9485372	6:149608874	TAB2	-	-	0,90	1,11
rs10474352	5:90732225	ARRDC3-AS1	-	-	0,92	1,09
rs2290203	15:91512067	PRC1, AC068831.7, PRC1-AS1	-	-	0,93	1,08
rs17530068	6:82193109	AL590824.1 - TENT5A	-	-	1,05	1,05
rs9383938	6:151987357	ESR1	-	-	1,08	1,08
rs4951011	1:203766331	ZBED6, ZC3H11A	-	-	1,09	1,09
rs2284378	20:32588095	RALY	-	-	1,10	1,10

Supplementary Table S1

rs2392780	8:128388025	POU5F1B, CASC8, PCAT1	-	-	1,15	1,00
rs4415084	5:44662515	LINC02224 - AC093292.1	-	-	1,17	1,00
rs3822625	5:56178111	MAP3K1	-	-	1,36	1,36
rs7726354	5:56256483	MIER3	-	-	1,37	1,37

* SNPs' position were based on build GRCh37/hg19

** OR for one allele/two alleles

*** OR from Shieh's study used for Asian women

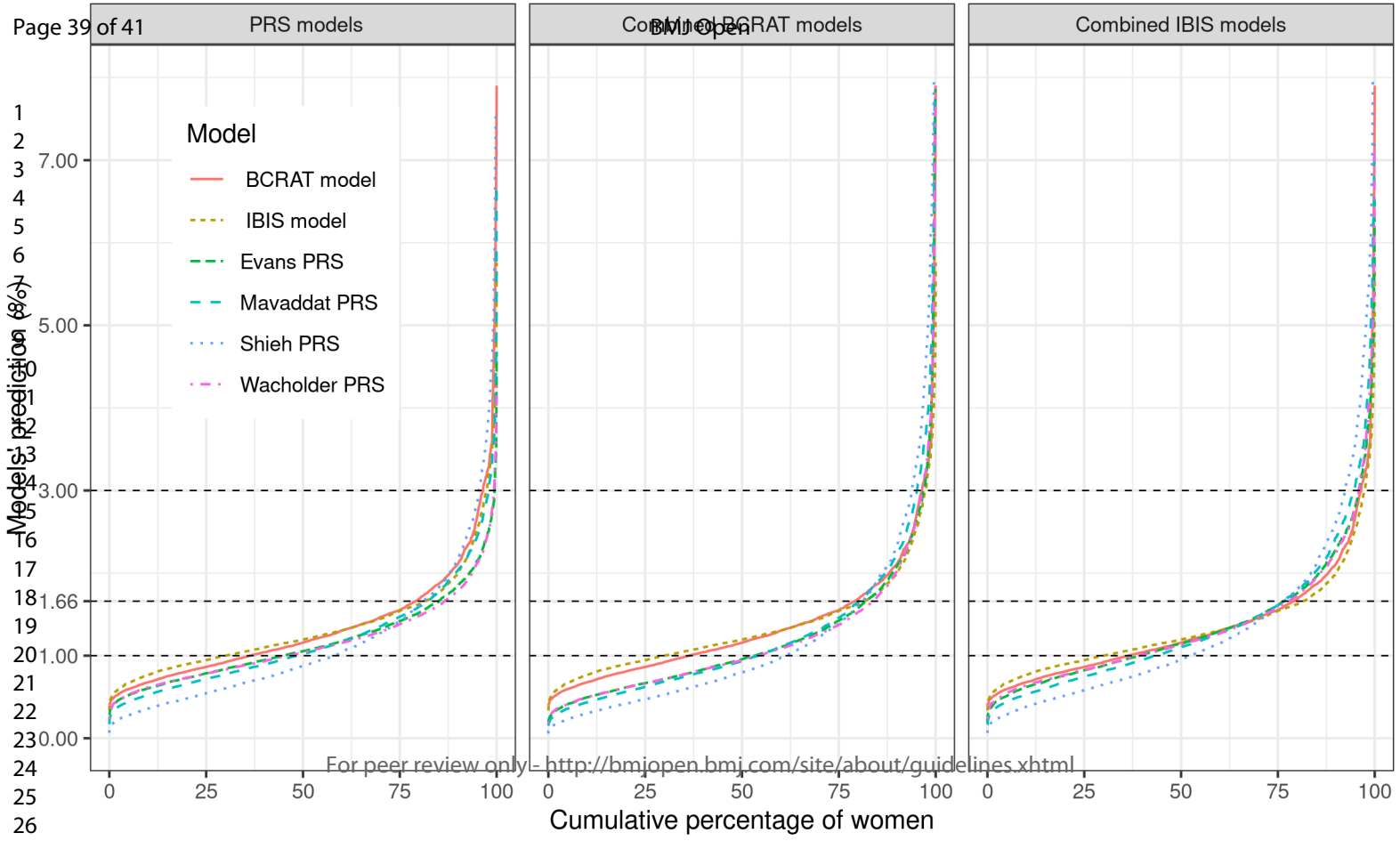
Supplementary Table S1

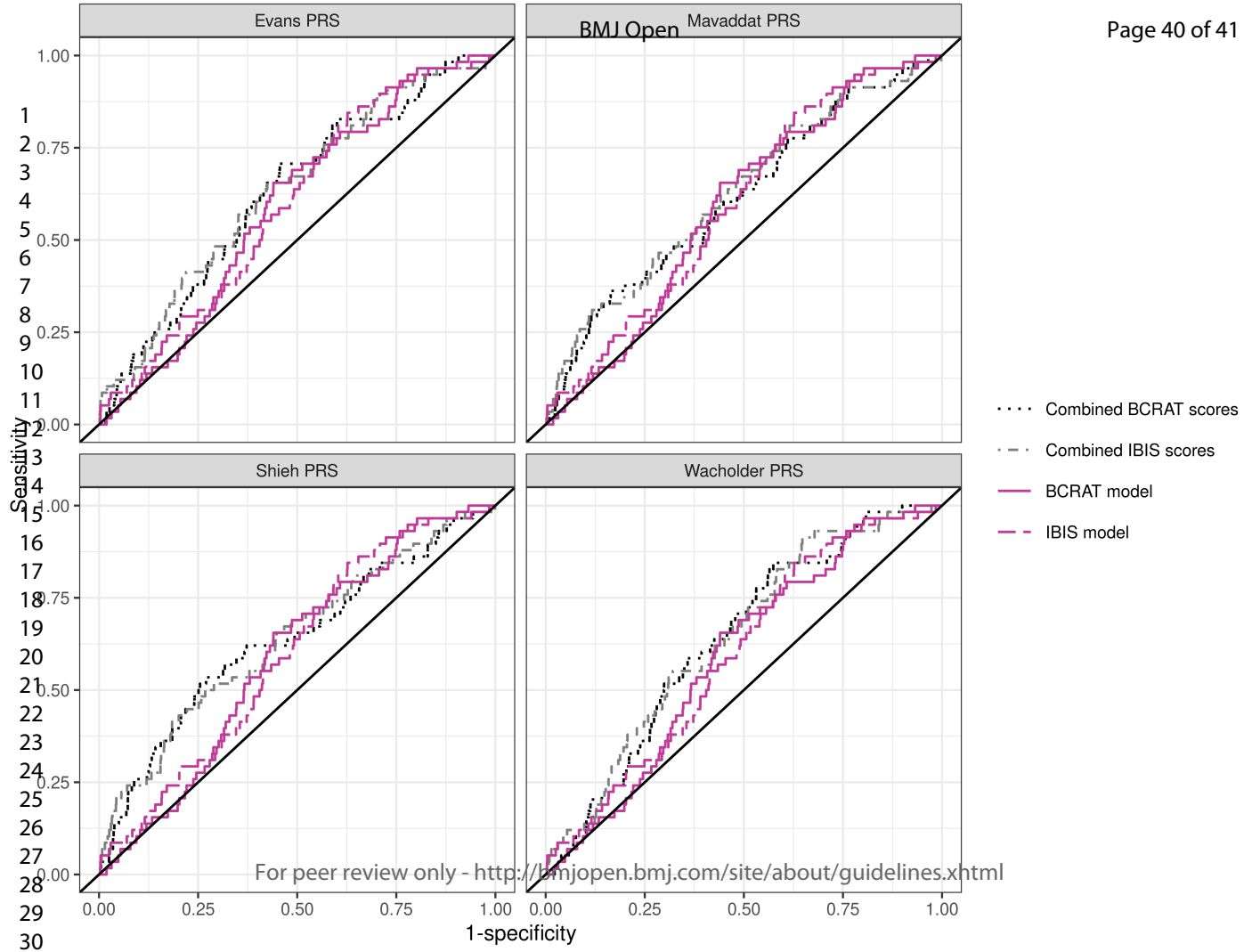
Table S2

	Clinical-based cohort	Clinicogenetic-based cohort
	N=10,200	**N=4,555**
Breast cancer within 5 years	131 (1.28%)	58 (1.27%)
BCRAT absolute risk (%)	1.30 (0.74)	1.33 (0.73)
IBIS absolute risk (%)	1.31 (0.59)	1.33 (0.60)
Age at baseline (years)	54.1 (7.7)	54.1 (7.6)
Age categories:		
<=49	3,556 (34.9%)	1,557 (34.2%)
50-59	3,980 (39.0%)	1,839 (40.4%)
>=60	2,664 (26.1%)	1,159 (25.4%)
Birth province:		
In Canada outside Quebec	333 (3.3%)	74 (1.6%)
Outside Canada	1,490 (14.6%)	189 (4.1%)
Quebec	8,373 (82.1%)	4,292 (94.2%)
Missing	*4*	*0*
Ethnicity:		
Asian	188 (1.8%)	5 (0.1%)
Black African	182 (1.8%)	0 (0.0%)
Hispanic non-american	234 (2.3%)	1 (<0.1%)
Other	542 (5.3%)	86 (1.9%)
White/European	9,054 (88.8%)	4,463 (98.0%)
Age at menarche (years):		
<=11	2,305 (22.9%)	1,027 (22.7%)
12-13	4,754 (47.2%)	2,166 (47.9%)
>=14	3,021 (30.0%)	1,331 (29.4%)
Missing	*120*	*31*
Age at first live birth (years):		
<=19	1,124 (13.1%)	422 (11.1%)
20-24	2,955 (34.5%)	1,324 (34.8%)
25-29	2,814 (32.9%)	1,312 (34.5%)
>=30	1,621 (19.0%)	734 (19.3%)
Nulliparous	40 (0.5%)	14 (0.4%)
Missing	*1,646*	*749*
First-degree relatives with breast cancer:		
0	8,945 (87.7%)	3,949 (86.7%)
1	1,130 (11.1%)	556 (12.2%)
>=2	125 (1.23%)	50 (1.10%)
Previous breast biopsy:		
0	10,023 (98.3%)	4,463 (98.0%)

1	134 (1.31%)	71 (1.56%)
>=2	43 (0.42%)	21 (0.46%)
History of hyperplasia	6 (0.06%)	1 (0.02%)
History of atypical hyperplasia	1 (0.56%)	0 (0.00%)
Lobular Carcinoma In Situ	0	0
Weight (Kg)	67.4 [59.4;78.0]	66.8 [59.6;76.8]
Height (m)	1.61 [1.57;1.65]	1.61 [1.57;1.65]
History of ovary cancer	94 (0.92%)	46 (1.01%)
Menopause occurrence		
Pre-menopausal	4176 (40.9%)	1891 (41.5%)
Post-menopausal	5885 (57.7%)	2617 (57.5%)
Unknown	139 (1.36%)	47 (1.03%)
Use of HRT		
Never	7477 (73.3%)	3249 (71.3%)
Previous user (more than 5 years ago)	1126 (11.0%)	506 (11.1%)
Previous user (less than 5 years ago)	1285 (12.6%)	646 (14.2%)
Current user	312 (3.06%)	154 (3.38%)
HRT length of use (years)	0.00 [0.00;1.00]	0.00 [0.00;1.00]
Last HRT use (years)	0.00 [0.00;0.00]	0.00 [0.00;0.00]
Mother history of breast cancer	832 (8.16%)	412 (9.05%)
Mother history of ovary cancer	114 (1.12%)	60 (1.32%)
Father history of breast cancer	8 (0.08%)	2 (0.04%)

HRT: hormonal replacement therapy; PRS: polygenic risk score; clinical-based cohort: validation of the BCRAT and IBIS models, included women with a BCRAT and an IBIS score; clinicogenetic-based cohort: validation of the PRS models and comparison with the BCRAT and IBIS models, genotyped women with all SNPs available.
* Not available for the phase 2





STROBE Statement—Checklist of items that should be included in reports of ***cohort studies***

	Item No	Recommendation	Page number
Title and abstract	1	(*a*) Indicate the study's design with a commonly used term in the title or the abstract	1
		(*b*) Provide in the abstract an informative and balanced summary of what was done and what was found	1-2
Introduction			
Background /rationale	2	Explain the scientific background and rationale for the investigation being reported	3-5
Objectives	3	State specific objectives, including any prespecified hypotheses	5
Methods			
Study design	4	Present key elements of study design early in the paper	5-6
Setting	5	Describe the setting, locations, and relevant dates, including periods of recruitment, exposure, follow-up, and data collection	5
Participants	6	(*a*) Give the eligibility criteria, and the sources and methods of selection of participants. Describe methods of follow-up	6
		(*b*) For matched studies, give matching criteria and number of exposed and unexposed	-
Variables	7	Clearly define all outcomes, exposures, predictors, potential confounders, and effect modifiers. Give diagnostic criteria, if applicable	7-8
Data sources/ measurement	8*	For each variable of interest, give sources of data and details of methods of assessment (measurement). Describe comparability of assessment methods if there is more than one group	7-8
Bias	9	Describe any efforts to address potential sources of bias	5
Study size	10	Explain how the study size was arrived at	6
Quantitative variables	11	Explain how quantitative variables were handled in the analyses. If applicable, describe which groupings were chosen and why	7-8
Statistical methods	12	(*a*) Describe all statistical methods, including those used to control for confounding	8-9
		(*b*) Describe any methods used to examine subgroups and interactions	-
		(*c*) Explain how missing data were addressed	7
		(*d*) If applicable, explain how loss to follow-up was addressed	8-9
		(*e*) Describe any sensitivity analyses	-
Results			
Participants	13*	(a) Report numbers of individuals at each stage of study—eg numbers potentially eligible, examined for eligibility, confirmed eligible, included in the study, completing follow-up, and analysed	9
		(b) Give reasons for non-participation at each stage	9
		(c) Consider use of a flow diagram	Figure 1
Descriptive data	14*	(a) Give characteristics of study participants (eg demographic, clinical, social) and information on exposures and potential confounders	Table S2
		(b) Indicate number of participants with missing data for each variable of interest	Table S2
		(c) Summarise follow-up time (eg, average and total amount)	9
Outcome	15*	Report numbers of outcome events or summary measures over time	9

data			
Main results	16	(*a*) Give unadjusted estimates and, if applicable, confounder-adjusted estimates and their precision (eg, 95% confidence interval). Make clear which confounders were adjusted for and why they were included	10-12
		(*b*) Report category boundaries when continuous variables were categorized	Table S2
		(*c*) If relevant, consider translating estimates of relative risk into absolute risk for a meaningful time period	9
Other analyses	17	Report other analyses done—eg analyses of subgroups and interactions, and sensitivity analyses	10-12
Discussion			
Key results	18	Summarise key results with reference to study objectives	12
Limitations	19	Discuss limitations of the study, taking into account sources of potential bias or imprecision. Discuss both direction and magnitude of any potential bias	13
Interpretation	20	Give a cautious overall interpretation of results considering objectives, limitations, multiplicity of analyses, results from similar studies, and other relevant evidence	12-13
Generalisability	21	Discuss the generalisability (external validity) of the study results	13
Other information			
Funding	22	Give the source of funding and the role of the funders for the present study and, if applicable, for the original study on which the present article is based	21

*Give information separately for exposed and unexposed groups.

Note: An Explanation and Elaboration article discusses each checklist item and gives methodological background and published examples of transparent reporting. The STROBE checklist is best used in conjunction with this article (freely available on the Web sites of PLoS Medicine at http://www.plosmedicine.org/, Annals of Internal Medicine at http://www.annals.org/, and Epidemiology at http://www.epidem.com/). Information on the STROBE Initiative is available at http://www.strobe-statement.org.