10-year performance of four models of breast cancer risk: a validation study

Mary Beth Terry, Yuyan Liao, Alice S Whittemore, Nicole Leoce, Richard Buchsbaum, Nur Zeinomar, Gillian S Dite, Wendy K Chung, Julia A Knight, Melissa C Southey, Roger L Milne, David Goldgar, Graham G Giles, Sue-Anne McLachlan, Michael L Friedlander, Prue C Weideman, Gord Glendon, Stephanie Nesci, Irene L Andrulis, Esther M John, Kelly-Anne Phillips, Mary B Daly, Saundra S Buys, John L Hopper, Robert J MacInnis

Lancet Oncol 2019; 20: 504–17

Published Online February 21, 2019 http://dx.doi.org/10.1016/S1470-2045(18)30902-1

See Comment page 463

Department of Epidemiology (Prof M B Terry PhD, Y Liao MS, N Leoce DrPH, N Zeinomar PhD) **and Department of Biostatistics** (R Buchsbaum), **Mailman School of Public Health, Department of Pediatrics** (Prof W K Chung MD), **and Department of Medicine** (Prof W K Chung), **Columbia University, New York, NY, USA; Herbert Irving Comprehensive Cancer Center, Columbia University Medical Center, New York, NY, USA** (Prof M B Terry, Prof W K Chung); **Department of Health Research and Policy** (Prof A S Whittemore PhD), **Department of Biomedical Data Science** (Prof A S Whittemore), **Department of Medicine** (Prof E M John PhD), **and Stanford Cancer Institute** (Prof E M John), **Stanford University School of Medicine, Stanford, CA, USA; Centre for Epidemiology and Biostatistics** (G S Dite PhD, R L Milne PhD, Prof G G Giles PhD, P C Weideman GDip, Prof K-A Phillips MD, Prof J L Hopper PhD, R J MacInnis PhD), **Genetic Epidemiology Laboratory, Department of Pathology** (Prof M C Southey PhD), **Department of Medicine** (S-A McLachlan MBBS), **and Sir Peter MacCallum Department of Oncology** (Prof K-A Phillips), **University of Melbourne, Parkville, VIC, Australia; Lunenfeld-Tanenbaum Research Institute, Sinai Health System, Toronto, ON, Canada** (Prof J A Knight PhD, G Glendon MSc, Prof I L Andrulis PhD); **Dalla Lana School of Public Health** (Prof J A Knight), **Department of Molecular Genetics**

Summary

Background Independent validation is essential to justify use of models of breast cancer risk prediction and inform decisions about prevention options and screening. Few independent validations had been done using cohorts for common breast cancer risk prediction models, and those that have been done had small sample sizes and short follow-up periods, and used earlier versions of the prediction tools. We aimed to validate the relative performance of four commonly used models of breast cancer risk and assess the effect of limited data input on each one's performance.

Methods In this validation study, we used the Breast Cancer Prospective Family Study Cohort (ProF-SC), which includes 18 856 women from Australia, Canada, and the USA who did not have breast cancer at recruitment, between March 17, 1992, and June 29, 2011. We selected women from the cohort who were 20–70 years old and had no previous history of bilateral prophylactic mastectomy or ovarian cancer, at least 2 months of follow-up data, and information available about family history of breast cancer. We used this selected cohort to calculate 10-year risk scores and compare four models of breast cancer risk prediction: the Breast and Ovarian Analysis of Disease Incidence and Carrier Estimation Algorithm model (BOADICEA), BRCAPRO, the Breast Cancer Risk Assessment Tool (BCRAT), and the International Breast Cancer Intervention Study model (IBIS). We compared model calibration based on the ratio of the expected number of breast cancer cases to the observed number of breast cancer cases in the cohort, and on the basis of their discriminatory ability to separate those who will and will not have breast cancer diagnosed within 10 years as measured with the concordance statistic (C-statistic). We did subgroup analyses to compare the performance of the models at 10 years in *BRCA1* or *BRCA2* mutation carriers (ie, *BRCA*-positive women), tested non-carriers and untested participants (ie, *BRCA*-negative women), and participants younger than 50 years at recruitment. We also assessed the effect that limited data input (eg, restriction of the amount of family history and non-genetic information included) had on the models' performance.

Findings After median follow-up of 11·1 years (IQR 6·0–14·4), 619 (4%) of 15 732 women selected from the ProF-SC cohort study were prospectively diagnosed with breast cancer after recruitment, of whom 519 (84%) had histologically confirmed disease. BOADICEA and IBIS were well calibrated in the overall validation cohort, whereas BRCAPRO and BCRAT underpredicted risk (ratio of expected cases to observed cases 1·05 [95% CI 0·97–1·14] for BOADICEA, 1·03 [0·96–1·12] for IBIS, 0·59 [0·55–0·64] for BRCAPRO, and 0·79 [0·73–0·85] for BRCAT). The estimated C-statistics for the complete validation cohort were 0·70 (95% CI 0·68–0·72) for BOADICEA, 0·71 (0·69–0·73) for IBIS, 0·68 (0·65–0·70) for BRCAPRO, and 0·60 (0·58–0·62) for BCRAT. In subgroup analyses by *BRCA* mutation status, the ratio of expected to observed cases for *BRCA*-negative women was 1·02 (95% CI 0·93–1·12) for BOADICEA, 1·00 (0·92–1·10) for IBIS, 0·53 (0·49–0·58) for BRCAPRO, and 0·97 (0·89–1·06) for BCRAT. For *BRCA*-positive participants, BOADICEA and IBIS were well calibrated, but BRCAPRO underpredicted risk (ratio of expected to observed cases 1·17 [95% CI 0·99–1·38] for BOADICEA, 1·14 [0·96–1·35] for IBIS, and 0·80 [0·68–0·95] for BRCAPRO). We noted similar patterns of calibration for women younger than 50 years at recruitment. Finally, BOADICEA and IBIS predictive scores were not appreciably affected by limiting input data to family history for first-degree and second-degree relatives.

Interpretation Our results suggest that models that include multigenerational family history, such as BOADICEA and IBIS, have better ability to predict breast cancer risk, even for women at average or below-average risk of breast cancer. Although BOADICEA and IBIS performed similarly, further improvements in the accuracy of predictions could be possible with hybrid models that incorporate the polygenic risk component of BOADICEA and the non-family-history risk factors included in IBIS.

Funding US National Institutes of Health, National Cancer Institute, Breast Cancer Research Foundation, Australian National Health and Medical Research Council, Victorian Health Promotion Foundation, Victorian Breast Cancer Research Consortium, Cancer Australia, National Breast Cancer Foundation, Queensland Cancer Fund, Cancer Councils of New South Wales, Victoria, Tasmania, and South Australia, and Cancer Foundation of Western Australia.

Research in context

Evidence before this study

We searched PubMed with the terms "BOADICEA", "IBIS", "BRCAPRO", "BCRAT", "cohort", and "validation" for articles published in English up to Oct 1, 2018, that had prospective follow-up to allow for calibration. Several independent validations of models of non-family-based breast cancer risk (such as the Breast Cancer Risk Assessment Tool [BCRAT]) have been done, but few have been done of three commonly used models based on family history: the Breast and Ovarian Analysis of Disease Incidence and Carrier Estimation Algorithm model (BOADICEA), the International Breast Cancer Intervention Study model (IBIS), and BRCAPRO. Validations with a cohort design of multiple models simultaneously that have been done were of previous versions of the models, and were further limited by small sample sizes and short follow-up.

Added value of this study

To our knowledge, this cohort study based on the Breast Cancer Prospective Family Study Cohort is the largest, independent validation of four commonly used models of breast cancer risk (BOADICEA, IBIS, BRCAPRO, and BCRAT). BOADICEA and IBIS were well calibrated overall for prediction of breast cancer risk in this cohort and were more accurate than the other models, even for women without *BRCA1* and *BRCA2* mutations and those with low familial risk of breast cancer. Furthermore, we found similar model performance when we limited family history to first-degree and second-degree relatives for all models. Assessment of risk models in cohorts enriched for familial risk of breast cancer like the Breast Cancer Prospective Family Study Cohort is important to assess performance across the spectrum of risk, especially in high-risk populations, in whom screening and risk predictions are crucial to patient–clinician shared decision-making.

Implications of all the available evidence

Family-history-based models—and particularly BOADICEA and IBIS—seem to more accurately predict breast cancer risk than those that do not include detailed family history, even for women who do not have a strong family history of breast cancer.

Introduction

Models of breast cancer risk based on established risk factors are used to inform decisions about primary prevention (eg, chemoprevention). Although age is still the main criterion used to inform screening guidelines, increasingly risk models are being used in risk-based screening programmes.[1] Independent prospective validation is essential to justify the clinical use of risk models.[2] However, unlike many cardiovascular models, there are few prospective, independent validations of models for estimation of cancer risk.[3] Model performance should be based on both calibration and discrimination. Calibration, which can be assessed only in prospective studies, compares the expected number of cases based on model estimates with the number reported.[4] Discrimination, which shows the model's ability to separate people who do and do not have the outcome, is often assessed by the receiver operator characteristic (ROC) curve and the concordance statistic (C-statistic).

Most models of cancer risk, including those for breast cancer, differ substantially in how they handle family history of cancer. Family history can be measured by a binary construct (ie, present or absent), by a number (such as number of affected relatives, usually first-degree only), or by pedigree information, which includes ages at diagnosis of affected relatives and ages or dates of death of unaffected relatives.[3] Family-history-based models based on pedigree information capture substantially more familial risk than binary or number-based exposure constructs. Models of breast cancer risk also differ in assumptions about the contribution of genes other than *BRCA1* and *BRCA2* to explain the remaining familial risk, as well as in whether risk factors not based on family history are included.

Four models that are commonly used to predict breast cancer risk are the International Breast Cancer Intervention Study[5] model (IBIS), the Breast and Ovarian Analysis of Disease Incidence and Carrier Estimation Algorithm model[6] (BOADICEA), BRCAPRO,[7] and the Breast Cancer Risk Assessment Tool (BCRAT; also referred to as the Gail model).[8] Although several large and independent validations of BCRAT have been reported,[9,10] few independent, prospective, validations of the other three models, which are all based on or include family pedigree information, have been done. Validations of these models that have been done had small sample sizes and short follow-up periods, and were based on previous versions of the models.[11–15] In addition to prospective, independent validation, another key unaddressed issue is how to maximise efficient use of these models to accurately assign risk, especially in clinical settings. Time and convenience of data collection are important considerations, because clinicians have reported challenges with the collection and entering of detailed family history data (including ages at diagnoses and ages of unaffected relatives).[16] We aimed to both prospectively and independently validate the performance of commonly used models of breast cancer risk and to assess the effect of limited data input on overall model performance.

Methods

Study design and participants

For this validation study, we used the Breast Cancer Prospective Family Study Cohort (ProF-SC), which comprised pooled data from two cohorts of women from Australia, Canada, and the USA (the Breast Cancer Family

(Prof I L Andrulis), **and Department of Laboratory Medicine and Pathobiology** (Prof I L Andrulis), **University of Toronto, Toronto, ON, Canada; Precision Medicine, School of Clinical Sciences at Monash Health** (Prof M C Southey, R L Milne) **and Department of Epidemiology and Preventive Medicine** (Prof G G Giles), **Monash University, Clayton, VIC, Australia; Cancer Epidemiology and Intelligence Division, Cancer Council Victoria, Melbourne, VIC, Australia** (R L Milne, Prof G G Giles, R J MacInnis); **Department of Dermatology** (Prof D Goldgar PhD), **Huntsman Cancer Institute** (Prof D Goldgar, Prof S S Buys MD), **and Department of Medicine** (Prof S S Buys), **University of Utah Health, Salt Lake City, UT, USA; Department of Oncology, St Vincent's Hospital, Parkville, VIC, Australia** (S-A McLachlan); **Prince of Wales Clinical School, University of New South Wales, Sydney, NSW, Australia** (Prof M L Friedlander MD); **Department of Medical Oncology, Prince of Wales Hospital, NSW, Australia** (Prof M L Friedlander); **Department of Medical Oncology, Peter MacCallum Cancer Centre, Melbourne, VIC, Australia** (S Nesci MPH, Prof K-A Phillips); **and Department of Clinical Genetics, Fox Chase Cancer Center, Philadelphia, PA, USA** (Prof M B Daly MD)

Correspondence to:
Prof Mary Beth Terry, Department of Epidemiology, Columbia University Mailman School of Public Health, New York, NY 10032, USA
mt146@cumc.columbia.edu

	BOADICEA	BRCAPRO	BCRAT	IBIS
Personal information				
Current age	Yes	Yes	Yes	Yes
Year of birth	Yes	NA	NA	NA
Race or ethnicity	NA	Yes	Yes	NA
Age at menarche	NA	NA	Yes	Yes
Parity	NA	NA	NA	Yes
Age at first birth	NA	NA	Yes	Yes
Menopausal status	NA	NA	NA	Yes
Menopausal hormone therapy use	NA	NA	NA	Yes
Body-mass index	NA	NA	NA	Yes
History of atypical hyperplasia	NA	NA	Yes	Yes
History of lobular carcinoma in situ	NA	NA	NA	Yes
Previous breast biopsy	NA	NA	Yes	Yes
Mammographic density	NA	NA	NA	Yes
Polygenic risk score	NA	NA	NA	Yes
Information about the individual and their family members				
First-degree relatives with breast cancer	Yes	Yes	Yes	Yes
Second-degree and third-degree relatives with breast cancer	Yes	Yes	NA	Yes
Identical twin with breast cancer	Yes	Yes	NA	NA
Age at cancer diagnosis	Yes	Yes	NA	Yes
Bilateral breast cancer*	Yes	Yes	NA	Yes
Ovarian cancer*	Yes	Yes	NA	Yes
Pancreatic cancer*	Yes	NA	NA	NA
Prostate cancer*	Yes	NA	NA	NA
Molecular subtype of breast tumour	Yes	Yes	NA	NA
Vital status of family members†	Yes	Yes	NA	Yes
BRCA1 and *BRCA2* mutation status	Yes	Yes	NA	Yes
Ashkenazi Jewish heritage	Yes	Yes	NA	Yes

These models do not necessarily require all data inputs for calculation of a risk estimate. BOADICEA=Breast and Ovarian Analysis of Disease Incidence and Carrier Estimation Algorithm model. BCRAT=Breast Cancer Risk Assessment Tool. IBIS=International Breast Cancer Intervention Study model. NA=not applicable. *In any family member. †Current age or age at death if deceased.

Table 1: **Summary of inputs needed for the four assessed models of breast cancer risk**

Registry and the Kathleen Cunningham Foundation Consortium for Research into Familial Breast Cancer [kConFab]) recruited between March 17, 1992, and June 29, 2011.[17–19] Full details of ProF-SC,[19] and baseline data collection and follow-up methods for the Breast Cancer Family Registry[17] and kConFab,[18] have been published previously.[19,20] At baseline, ProF-SC included 18 856 women who did not have breast cancer. We restricted analyses to women who were aged 20–70 years at baseline and had no previous history of bilateral prophylactic mastectomy or ovarian cancer, at least 2 months of follow-up data since recruitment, and information available about family history from family pedigrees (n=15 732). We limited eligibility to women aged between 20 and 70 years at baseline so that we could validate a 10-year predicted risk score during our average follow-up time of 10 years (some risk models do not predict risk beyond age 80 years). All participants in the Breast Cancer Family Registry and kConFab provided written informed consent before enrolment, and the study protocols at all sites were approved by institutional review boards at each participating institution.

For more on **IBIS** see http://www.ems-trials.org/riskevaluator/

For more on **BOADICEA** see https://pluto.srl.cam.ac.uk/cgi-bin/bd3/v3/bd.cgi

For more on **BCRAT** see https://dceg.cancer.gov/tools/risk-assessment/bcrasasmacro

For more on **BRCAPRO** see https://projects.iq.harvard.edu/bayesmendel/brcapro

Procedures

Participants completed the same baseline risk factor questionnaire, which included questions about participants' demographic characteristics, height, weight, history of benign and malignant breast disease, history of breast and ovarian surgeries, reproductive history, and lifestyle factors. The questionnaire also included questions about family history regarding breast and other cancers in the participants and their first-degree and second-degree relatives, including ages at cancer diagnoses and age last known to be alive. Family history and epidemiological information for the families participating in the ProF-SC cohort were collected at baseline and at each systematic follow-up interview; additional information was collected annually when families were contacted.

Screening for germline *BRCA1* and *BRCA2* mutations in the Breast Cancer Family Registry and kConFab[17,21] typically involved screening the youngest affected family member at baseline. If that person carried a mutation, other family members were also tested. *BRCA*-negative women were defined as women who were not known to be *BRCA1* or *BRCA2* carriers, which included both tested and untested women.

Outcome information about breast cancer diagnosis, including pathological report, biopsies, and treatment questionnaires, were collected at diagnosis, or after the diagnosis was identified by personal or family reports or by cancer registry confirmation. Because the cohorts are based on family information, cancer information for each participant was typically self-reported or reported by a first-degree relative and could come from more than one source.

In the validation cohort, data were generally available for all items required by these models, except for mammographic density and polygenic risk score used in IBIS. However, data were scarce for pathologically confirmed lobular carcinoma in situ (used in IBIS) and atypical hyperplasia (used in BCRAT and IBIS).

We used the latest version (at the time of analysis) of software packages for each risk prediction model, IBIS (version 8b), BOADICEA (version 3), BCRAT (version 2.0), and BRCAPRO (version 2.1-3) to calculate 10-year risk scores. The data inputs required for each model to calculate breast cancer risk are summarised in table 1. BOADICEA, BRCAPRO, and IBIS all incorporate familial risk via detailed family history information across multiple generations. The breast cancer hazard ratio calculated by the IBIS model depends on several lifestyle and personal risk factors in addition to familial risk.[5] BOADICEA[6] and BRCAPRO[7] do not include lifestyle and personal risk factors. BCRAT accounts for familial risk on the basis of the number of affected first-degree relatives diagnosed with breast cancer, and also includes other risk factors, such as age at menarche,

age at first birth, history of atypical hyperplasia, and previous breast biopsy.[8] BCRAT and IBIS allow for calculation of competing risk of mortality.

Statistical analysis

We counted follow-up time according to whichever came first, breast cancer diagnosis of invasive or ductal carcinoma in situ, bilateral mastectomy, death, and follow-up time reaching 10 years and 2 months (we did not count events within the first 2 months of recruitment). Because the four models were developed to calculate the risk of invasive cancer rather than of in situ disease, we censored women at the time of a diagnosis of in situ breast cancer because treatment for in situ cancers can affect the risk of future invasive diagnoses. Deaths from causes other than breast cancer were considered to be competing risks and were not censored in the main analysis.

We compared model performance at 10 years based on calibration and discrimination for all four models in all women and by mutation status (ie, *BRCA*-positive women and *BRCA*-negative women). For *BRCA*-positive women, we compared only the three models (BOADICEA, BRCAPRO, and IBIS) that included information specifically for carriers. We further compared performance for women younger than 50 years because this age is included in some countries' guidelines as the threshold for age-based screening.

To assess model calibration, we compared the expected with the observed number of breast cancer cases. If follow-up time was less than 10 years, we used the method described by Amir and colleagues to adjust the predicted risk to the available follow-up.[11] The 95% CI for the ratio of expected to observed cases was calculated assuming a Poisson distribution:[10]

$$(\text{expected}/\text{observed} \times e\,(\pm 1{\cdot}96 \times \sqrt{(1/\text{observed})}))$$

We also assessed model calibration by comparing the mean model-assigned 10-year risk with the observed breast cancer incidence for each of the four assigned risk quantiles. We used the same cutoffs for all four models, but the cutoffs were separate for *BRCA*-positive and *BRCA*-negative women. For *BRCA*-negative women, the cutoffs used (<1·7%, ≥1·7% to <3·4%, ≥3·4% to <5%, and ≥5%) corresponded approximately to the quartiles of risk distribution. Furthermore, the 3·4% cutoff corresponds to the 10-year risk of an average 60-year-old woman and is roughly double the 5-year risk of 1·67%, which has been used in some chemoprevention trials to define high risk for the purposes of eligibility. The cutoffs used for *BRCA*-positive women (<10%, ≥10% to <20%, ≥20% to <25%, and ≥25%) corresponded approximately to the quartiles of the risk distribution for carriers. To compare the calibration across models, we calculated the estimated calibration index, which we defined as the mean squared difference between observed and predicted risk, with values closer to 0 indicating better calibration.[22,23]

	Unaffected after 10 years (N=8620)	Follow-up <10 years (N=6078)	Died within 10 years (N=415)	Breast cancer diagnosis within 10 years (N=619)
Age at baseline questionnaire, years				
20–29	1197 (14%)	1053 (17%)	5 (1%)	34 (6%)
30–39	2004 (23%)	1383 (23%)	24 (6%)	119 (19%)
40–49	2239 (26%)	1508 (25%)	55 (13%)	177 (29%)
50–59	1877 (22%)	1292 (21%)	117 (28%)	162 (26%)
60–70	1303 (15%)	842 (14%)	214 (52%)	127 (21%)
Race or ethnicity				
Non-Hispanic white	7390 (86%)	4281 (70%)	318 (77%)	522 (84%)
Non-Hispanic black	236 (3%)	448 (7%)	38 (9%)	16 (3%)
Hispanic	309 (4%)	905 (15%)	34 (8%)	38 (6%)
Asian	336 (4%)	223 (4%)	9 (2%)	26 (4%)
Other	251 (3%)	163 (3%)	11 (3%)	11 (2%)
Unknown	98 (1%)	58 (1%)	5 (1%)	6 (1%)
Age at menarche, years				
≤11	1466 (17%)	1091 (18%)	83 (20%)	91 (15%)
12–13	4543 (53%)	3069 (50%)	188 (45%)	331 (53%)
≥14	2530 (29%)	1823 (30%)	140 (34%)	190 (31%)
Unknown	81 (1%)	95 (2%)	4 (1%)	7 (1%)
Body-mass index, kg/m^2				
<25	4830 (56%)	2924 (48%)	174 (42%)	309 (50%)
25 to <30	2213 (26%)	1617 (27%)	125 (30%)	177 (29%)
≥30	1402 (16%)	1413 (23%)	103 (25%)	128 (21%)
Unknown	175 (2%)	124 (2%)	13 (3%)	5 (1%)
Age at first livebirth, years				
<20	993 (12%)	913 (15%)	96 (23%)	71 (11%)
20–24	2520 (29%)	1675 (28%)	167 (40%)	206 (33%)
25–29	1953 (23%)	1188 (20%)	64 (15%)	121 (20%)
≥30	922 (11%)	642 (11%)	29 (7%)	94 (15%)
Nulliparous	2232 (26%)	1660 (27%)	59 (14%)	127 (21%)
Menopausal hormone therapy use				
Ever	1930 (22%)	1137 (19%)	169 (41%)	161 (26%)
Never	6438 (75%)	4827 (79%)	235 (57%)	436 (70%)
Unknown	252 (3%)	114 (2%)	11 (3%)	22 (4%)
Menopausal status				
Premenopausal	4946 (57%)	3730 (61%)	66 (16%)	298 (48%)
Post-menopausal	2855 (33%)	1998 (33%)	326 (79%)	268 (43%)
Unknown	819 (10%)	350 (6%)	23 (6%)	53 (9%)
Personal history of benign breast disease				
Yes	2353 (27%)	1623 (27%)	105 (25%)	219 (35%)
No	5972 (69%)	4330 (71%)	302 (73%)	377 (61%)
Unknown	295 (3%)	125 (2%)	8 (2%)	23 (4%)
***BRCA1* and *BRCA2* mutation status**				
BRCA positive				
BRCA1 mutation carrier	184 (2%)	301 (5%)	17 (4%)	82 (13%)
BRCA2 mutation carrier	129 (1%)	294 (5%)	13 (3%)	55 (9%)
BRCA negative*	8307 (96%)	5483 (90%)	385 (93%)	482 (78%)
Individual tested negative	2716/8307 (33%)	1344/5483 (25%)	111/385 (29%)	184/482 (38%)
Affected family member tested negative	2551/8307 (31%)	2121/5483 (39%)	134/385 (35%)	101/482 (21%)

(Table 2 continues on next page)

	Unaffected after 10 years (N=8620)	Follow-up <10 years (N=6078)	Died within 10 years (N=415)	Breast cancer diagnosis within 10 years (N=619)
(Continued from previous page)				
No information about individual or family members	3040/8307 (37%)	2018/5483 (37%)	140/385 (36%)	197/482 (41%)
First-degree relatives with breast cancer				
0	1477 (17%)	1157 (19%)	88 (21%)	83 (13%)
1	5633 (65%)	3878 (64%)	232 (56%)	347 (56%)
≥2	1510 (18%)	1043 (17%)	95 (23%)	189 (31%)
Second-degree relatives with breast cancer				
0	4173 (48%)	2678 (44%)	203 (49%)	240 (39%)
1	2858 (33%)	2042 (34%)	139 (33%)	227 (37%)
≥2	1589 (18%)	1358 (22%)	73 (18%)	152 (25%)

Data are n (%) or n/N (%). All eligible participants in the Breast Cancer Prospective Family Study Cohort are shown (n=15 732). *Women not known to be *BRCA1* or *BRCA2* carriers (includes both tested and untested women).

Table 2: **Risk factors by breast cancer outcome**

	Expected number of breast cancer cases	Observed number of breast cancer cases	Expected/observed ratio (95% CI)	Concordance statistic (95% CI)
Overall population (n=15 732)				
BOADICEA	652·5	619	1·05 (0·97–1·14)	0·70 (0·68–0·72)
BRCAPRO	367·7	619	0·59 (0·55–0·64)	0·68 (0·65–0·70)
BCRAT	488·6	619	0·79 (0·73–0·85)	0·60 (0·58–0·62)
IBIS	640·3	619	1·03 (0·96–1·12)	0·71 (0·69–0·73)
***BRCA* negative (n=14 657)**				
BOADICEA	492·3	482	1·02 (0·93–1·12)	0·65 (0·62–0·67)
BRCAPRO	257·6	482	0·53 (0·49–0·58)	0·62 (0·59–0·64)
BCRAT	466·7	482	0·97 (0·89–1·06)	0·64 (0·61–0·66)
IBIS	484·1	482	1·00 (0·92–1·10)	0·66 (0·64–0·68)
***BRCA* positive (n=1075)**				
BOADICEA	160·2	137	1·17 (0·99–1·38)	0·59 (0·53–0·64)
BRCAPRO	110·1	137	0·80 (0·68–0·95)	0·57 (0·52–0·63)
BCRAT	..	..	..	..
IBIS	156·2	137	1·14 (0·96–1·35)	0·60 (0·55–0·66)
Age <50 years at baseline (n=9798)				
BOADICEA	331·5	330	1·00 (0·90–1·12)	0·75 (0·73–0·78)
BRCAPRO	181·0	330	0·55 (0·49–0·61)	0·73 (0·70–0·76)
BCRAT	181·7	330	0·55 (0·49–0·61)	0·60 (0·57–0·63)
IBIS	328·7	330	1·00 (0·89–1·11)	0·75 (0·72–0·78)
Age <50 years at baseline and *BRCA* negative*(n=8397)				
BOADICEA	221·0	218	1·01 (0·89–1·16)	0·67 (0·64–0·70)
BRCAPRO	97·5	218	0·45 (0·39–0·51)	0·63 (0·59–0·66)
BCRAT	169·3	218	0·78 (0·68–0·89)	0·64 (0·61–0·67)
IBIS	209·6	218	0·96 (0·84–1·10)	0·67 (0·63–0·70)

All eligible participants in the Breast Cancer Prospective Family Study Cohort are shown for each category. BOADICEA=Breast and Ovarian Analysis of Disease Incidence and Carrier Estimation Algorithm model. BCRAT=Breast Cancer Risk Assessment Tool. IBIS=International Breast Cancer Intervention Study model. *Women not known to be *BRCA1* or *BRCA2* carriers (includes both tested and untested women).

Table 3: **Calibration and concordance statistics of the four assessed models of breast cancer risk in the Breast Cancer Prospective Family Study Cohort**

We assessed each model's ability to discriminate between women who did and did not develop breast cancer within 10 years by estimating and testing for significant differences in the C-statistic and plotting ROC curves to account for incomplete follow-up.[24] The C-statistic ranges from 0·5 (no discriminative ability) to 1·0 (perfect discrimination). To estimate the sensitivity of the risk scores, we used the 10-year breast cancer risk cutoff 3·4% for *BRCA*-negative women and 20% for *BRCA*-positive women. We derived the corresponding specificity values from the ROC curve output.

We did sensitivity analyses in which we did not censor women with diagnoses of in situ breast cancer. We compared model performance with and without accounting for competing mortality, and when selected model inputs (eg, detailed information about second-degree and third-degree relatives and non-genetic risk factors) were excluded. We created separate datasets for BOADICEA, BRCAPRO, and IBIS on the basis of the exclusion of family history data according to each data-completeness scenario. We also created separate datasets for BCRAT and IBIS in which we excluded all non-family-history data. We did additional sensitivity analyses in which we compared country-specific model performance for Australia, Canada, and the USA, examined model performance at shorter risk projection (ie, at 5 years) based on the calibration for all four models, and examined whether the risk scores differed between those with and without pathologically confirmed cancer. Based on 619 observed breast cancer diagnoses, the 95% CI of the ratio of expected cases to observed cases, if perfectly calibrated, is 0·92–1·10. Overall, assuming a type I error of 0·05, we could detect differences between two C-statistics of 0·02 or higher when the correlation between the two risk measures was 0·60 or higher.

The calibration and discrimination analyses were done in Stata (version 14.2) and with the Risk Model Assessment Package (versions 0.03-01) and Calibration Curves in R. All other analyses were done in SAS (version 9.4).

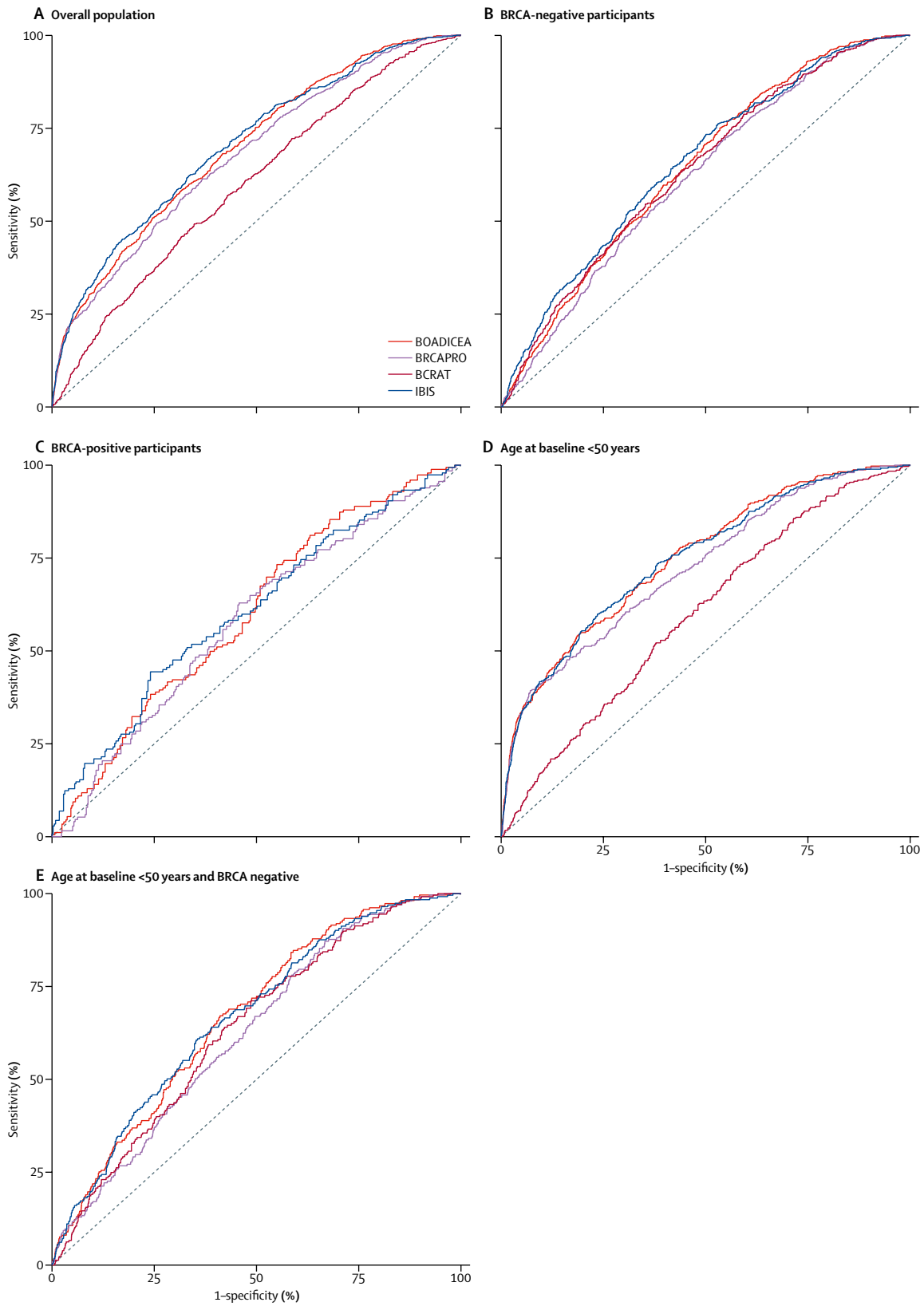
Role of the funding source

The study funders had no role in the study design; data collection, analysis, or interpretation; or writing of the report. The corresponding author and RJM had full access to all the data in the study and take responsibility for the integrity of the data and the accuracy of the data analysis. The corresponding author had final responsibility for the decision to submit for publication.

Results

Of 15732 eligible women in the validation cohort, 619 (4%) women were diagnosed with breast cancer during median follow-up of 11·1 years (IQR 6·0–14·4). 519 (84%) cases of breast cancer were histologically confirmed. Table 2 summarises participants' baseline characteristics.

BOADICEA and IBIS were well calibrated, closely predicting the observed number of women who developed



For more on **Risk Model Assessment Package** see https://gailg.github.io/rmap/

For more on **Calibration Curves** see https://github.com/BavoDC/CalibrationCurves

Figure 1: **Performance in terms of the sensitivity and specificity for BOADICEA, BRCAPRO, BCRAT, and IBIS**
Receiver operator characteristic curves of each model in the overall population (A), *BRCA*-negative participants (B), *BRCA*-positive participants (C), women younger than 50 years at baseline (D), and *BRCA*-negative participants younger than 50 years at baseline (E) are shown. *BRCA*-negative participants include those who tested negative for *BRCA*1 and *BRCA*2 mutations, and untested participants. The dotted line represents the line of no discrimination. BOADICEA=Breast and Ovarian Analysis of Disease Incidence and Carrier Estimation Algorithm model. BCRAT=Breast Cancer Risk Assessment Tool. IBIS=International Breast Cancer Intervention Study model.

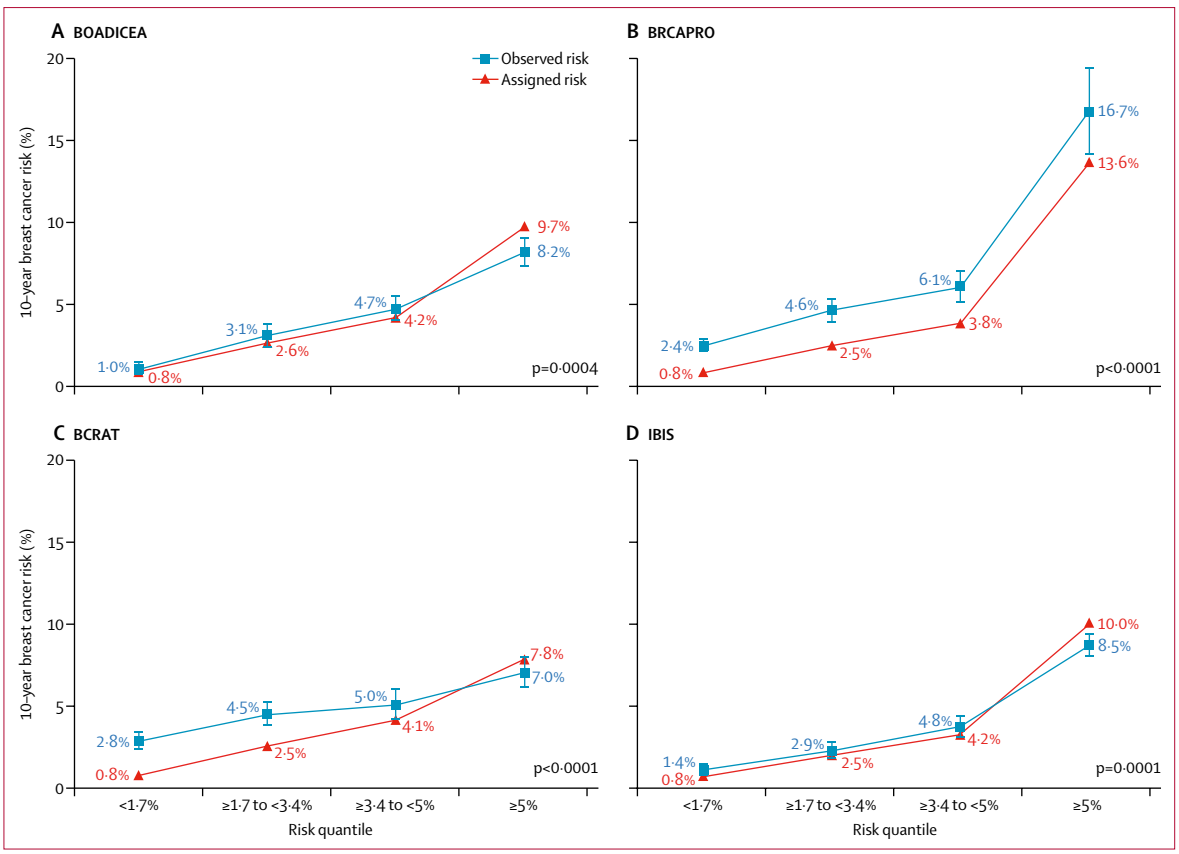


Figure 2: **Calibration of 10-year breast cancer risk prediction scores for BOADICEA (A), BRCAPRO (B), BCRAT (C), and IBIS (D) in overall cohort, by risk quantile**
Triangles represent the mean risk predicted by the models, whereas squares represent the mean observed risk for each quantile. Error bars represent 95% CIs. p values represent the test of goodness of fit across all four risk quantiles. All eligible participants included in the analysis (n=15 732). BOADICEA=Breast and Ovarian Analysis of Disease Incidence and Carrier Estimation Algorithm model. BCRAT=Breast Cancer Risk Assessment Tool. IBIS=International Breast Cancer Intervention Study model.

breast cancer in the complete validation cohort. The ratio of expected to observed cases was 1·05 (95% CI 0·97–1·14) for BOADICEA, 0·59 (0·55–0·64) for BRCAPRO, 0·79 (0·73–0·85) for BRCAT, and 1·03 (0·96–1·12) for IBIS (table 3). Figure 1 shows discriminatory power of the models in terms of relative sensitivity and specificity in the complete validation cohort, and subgroups by *BRCA* mutation status and age. The estimated sensitivity and specificity in the complete validation cohort were 81·9% (95% CI 78·6–84·9) and 42·8% (42·1–43·6) for BOADICEA, 45·9% (41·9–49·9) and 76·1% (75·4–76·8) for BRCAPRO, 57·8% (53·8–61·8) and 55·0% (54·2–55·8) for BCRAT, and 79·5% (76·1–82·6) and 46·7% (45·9–47·5) for IBIS.

The overall results did not change substantially when we did sensitivity analyses in which women diagnosed with in situ disease were not censored and were followed up until invasive diagnosis or last follow-up (appendix p 2), nor did results for IBIS when we considered different assumptions regarding competing mortality events (appendix p 3).

See Online for appendix

In *BRCA*-negative participants, the models' discriminatory ability based on C-statistic estimates ranged from 0·62 (95% CI 0·59–0·64) to 0·66 (0·64–0·68; table 3). The assigned risks from the different models were moderately correlated with each other (appendix p 5), particularly for *BRCA*-negative participants, and model discrimination based on the C-statistic did not differ significantly different between BOADICEA, IBIS, and BCRAT for *BRCA*-negative women (appendix p 5). BOADICEA, BCRAT, and IBIS were well calibrated for *BRCA*-negative participants, but BRCAPRO substantially underpredicted breast cancer risk (table 3). Although BOADICEA and IBIS were well calibrated in both *BRCA*-positive and *BRCA*-negative participants, they overpredicted risk for women in the highest risk quantile (figures 2, 3). BCRAT underpredicted risk for women in the two lowest risk quantiles, whereas BRCAPRO underpredicted risk for all except for those in the highest risk quantile (all p<0·0001 for difference between predicted risk and observed risk; figures 2, 3). The country-specific calibration analysis showed that BOADICEA, but not the other three models, was well calibrated for participants from all three countries (appendix pp 6–7).

Discrimination was lower for *BRCA*-positive than for *BRCA*-negative participants (table 3). Model discrimination

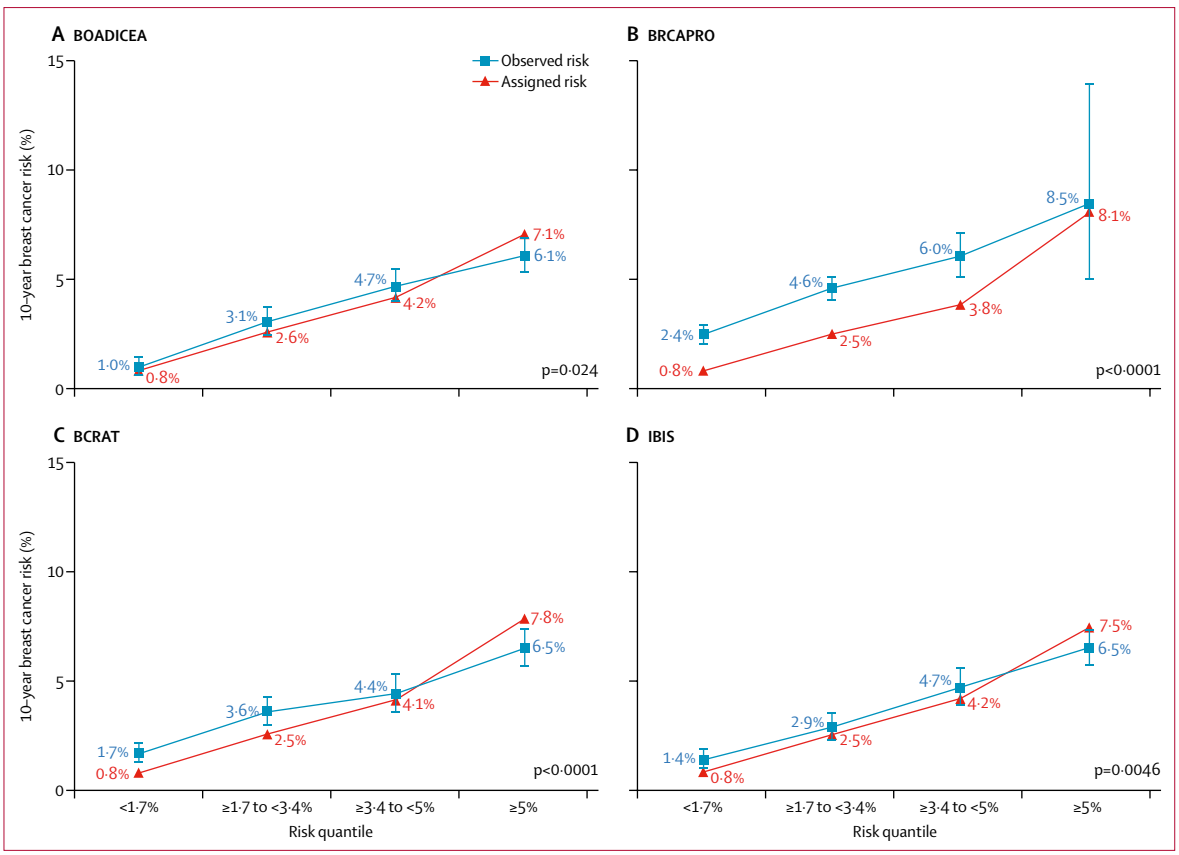


Figure 3: **Calibration of 10-year breast cancer risk prediction scores for BOADICEA (A), BRCAPRO (B), BCRAT (C), and IBIS (D) in *BRCA*-negative participants by risk quantile**

Triangles represent the mean risk predicted by the models, whereas squares represent the mean observed risk for each quantile. Error bars represent 95% CIs. p values represent the test of goodness of fit across all four risk quantiles. All *BRCA*-negative participants (those who tested negative for *BRCA1* and *BRCA2* mutations and untested participants) were included in the analysis (n=14 657). BOADICEA=Breast and Ovarian Analysis of Disease Incidence and Carrier Estimation Algorithm model. BCRAT=Breast Cancer Risk Assessment Tool. IBIS=International Breast Cancer Intervention Study model.

did not differ significantly between IBIS, BOADICEA, or BRCAPRO in women with *BRCA1* and *BRCA2* mutations (appendix p 5). Both BOADICEA and IBIS models were well calibrated for *BRCA*-positive women, but BRCAPRO underpredicted the breast cancer risk of these participants (table 3). BOADICEA overpredicted risk for women in the highest quantile, and BRCAPRO and IBIS both underpredicted risk for women in the lowest quantile (figure 4).

Regarding age, in participants younger than 50 years, independently of their *BRCA* mutation status, discrimination estimates were similar for BOADICEA and IBIS (p=0·64); both models had higher discrimination estimates than BRCAPRO and BCRAT (table 3; appendix p 5). BCRAT had low discrimination because it does not include mutation data. Similarly, BOADICEA and IBIS had higher discrimination than BRCAPRO and BCRAT in *BRCA*-negative participants younger than 50 years (table 3; appendix p 5). BOADICEA and IBIS models were well calibrated for all risk quantiles in all women younger than 50 years and also specifically in non-carriers younger than 50 years, whereas both BRCAPRO and BCRAT underpredicted risk for most quantiles (figure 5; table 3).

BOADICEA and IBIS had higher (and similar) sensitivities than the other two models for each analysed subgroup on the basis of 10-year risk scores (table 4). The sensitivities of BOADICEA, BRCAPRO, and IBIS were higher in *BRCA*-negative than in *BRCA*-positive participants, but lowest in *BRCA*-negative participants younger than 50 years (table 4). For all women, the estimated calibration index was 0·124 for BOADICEA and 0·035 for IBIS, suggesting slightly better calibration for IBIS than BOADICEA; this pattern held for all subgroups except for *BRCA*-negative participants (appendix p 8).

Results for the 5-year analysis of breast cancer risk were mostly similar to those for the 10-year risk overall and in all subgroups (appendix pp 8–13). However, at 5 years, IBIS underpredicted the number of breast cancers in *BRCA*-negative women younger than 50 years (ratio of expected to observed cases 0·83 [95% CI 0·69–0·99]).

We also assessed whether model performance was affected if some inputs (baseline information) were

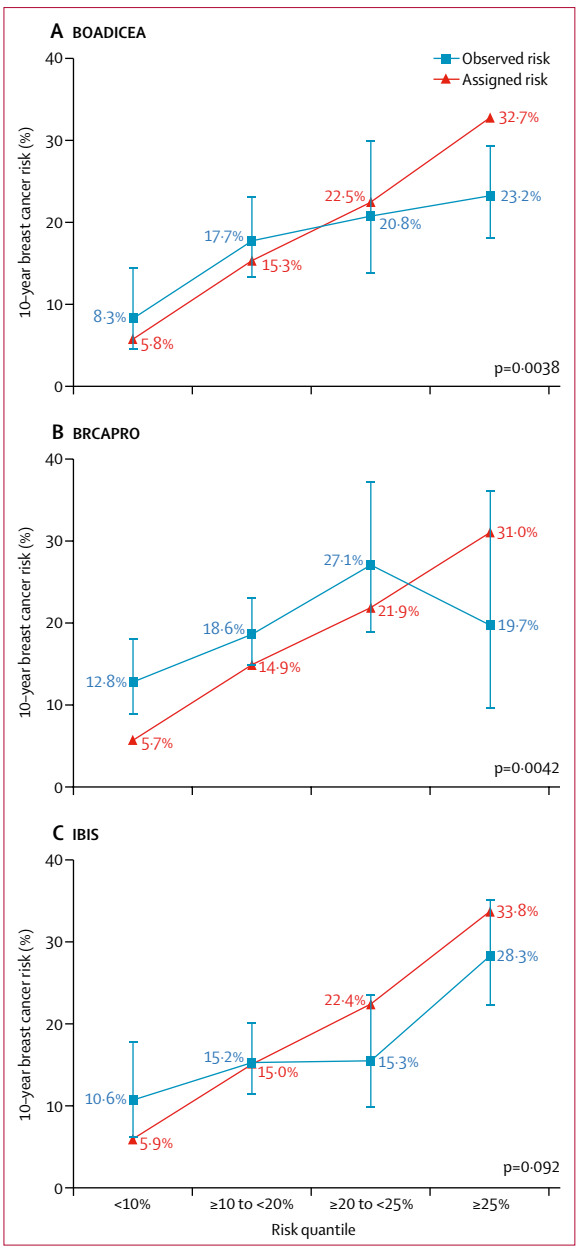


Figure 4: **Calibration of 10-year breast cancer risk prediction scores for BOADICEA (A), BRCAPRO (B), BCRAT (C), and IBIS (D) in *BRCA*-positive participants by risk quantile**

Triangles represent the mean risk predicted by the models, whereas squares represent the mean observed risk for each quantile. Error bars represent 95% CIs. p values represent the test of goodness of fit across all four risk quantiles. All *BRCA*-positive participants were included in the analysis (n=1075). BOADICEA=Breast and Ovarian Analysis of Disease Incidence and Carrier Estimation Algorithm model. IBIS=International Breast Cancer Intervention Study model.

omitted (appendix p 4). For *BRCA*-negative participants, BOADICEA remained well calibrated after exclusion of information about third-degree relatives, and the C-statistic estimate was virtually unchanged (p=0·63; appendix p 4). IBIS remained well calibrated after exclusion of information about third-degree relatives and unaffected second-degree relatives, and the C-statistic estimate was largely unaffected (p=0·24; appendix p 4). However, when we limited analyses to women who had at least one affected second-degree relative, exclusion of data for all second-degree relatives (including for those known not to have breast cancer) adversely affected calibration for both BOADICEA and IBIS, resulting in prediction of fewer events than were observed (appendix p 4). For *BRCA*-negative women, IBIS predicted a similar number of cases after exclusion of all non-family-history risk factors and information about third-degree relatives and unaffected second-degree relatives (appendix p 4), and the discriminatory power was similar to that for the overall model (p=0·54). For *BRCA*-negative women aged 50 years or older, BCRAT remained well calibrated and had similar discrimination after exclusion of non-family-history risk factors (appendix p 4).

Finally, predicted risk scores did not differ between women diagnosed with breast cancer with or without confirmed pathology report (figure 6).

Discussion

To our knowledge, our study is the largest independent validation of four commonly used models of breast cancer risk. We examined model performance for several subgroups, including women younger than 50 years and *BRCA1* and *BRCA2* mutation carriers. IBIS and BOADICEA generally gave better discrimination than either BCRAT or BRCAPRO, and were well calibrated in all subgroups considered. BCRAT, which was developed in a population not selected on the basis of mutation status, was well calibrated in *BRCA*-negative women. BRCAPRO was not well calibrated in any subgroup. According to our results, the four models tend to overpredict risk in the highest quantiles of familial risk and underpredict risk in the lower quantiles, even if, on average they seem to predict risk accurately overall. This misclassification means that some women at increased risk of breast cancer might make decisions about risk-reducing surgeries and chemoprevention that are based on inaccurate information. Similarly, some women in the lower risk quantiles might be under-screened or under-targeted for prevention. Because breast cancer incidence is rising worldwide,[25] risk models that incorporate available knowledge about family history along with polygenic risk scores, mammography-based risk measures, and new phenotypic markers (eg, DNA repair phenotype) could improve model performance for specific quantiles.

We selected BOADICEA, BRCAPRO, IBIS, and BCRAT because they vary substantially in terms of model inputs, ranging from those that use counts of family members with cancer along with non-family-history information to those including multigenerational pedigree and genetic information, and those that include non-family-history, genetic, and pedigree information. We also compared

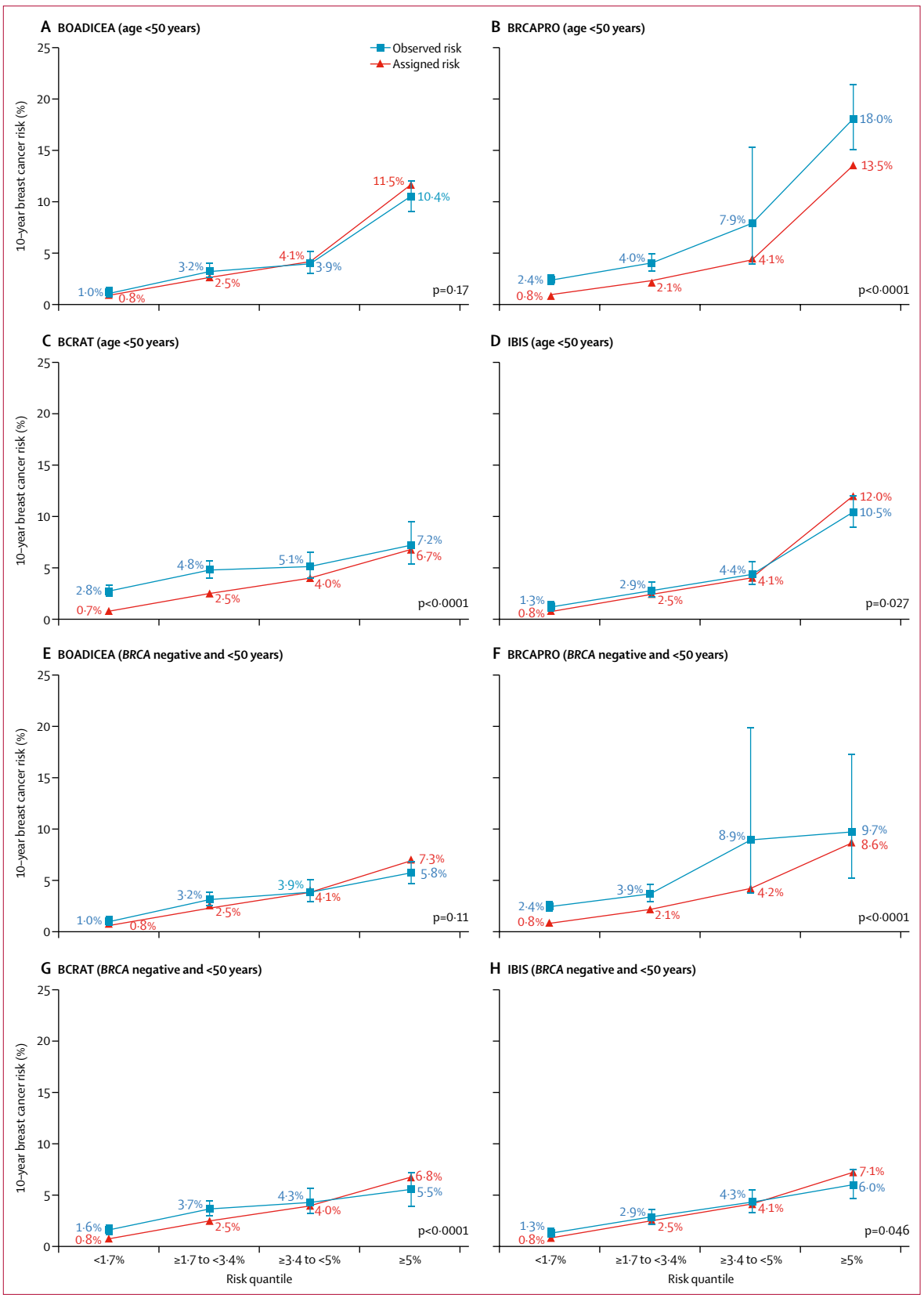


Figure 5: **Calibration of 10-year breast cancer risk prediction scores by age and risk quantile**
BOADICEA (A), BRCAPRO (B), BCRAT (C), and IBIS (D) in women younger than 50 years at baseline (n=9798) and BOADICEA (E), BRCAPRO (F), BCRAT (G), and IBIS (H) in *BRCA*-negative participants younger than 50 years at baseline (n=8937). Triangles represent the mean risk predicted by the models, whereas squares represent the mean observed risk for each quantile. Error bars denote 95% CIs. p values represent the test of goodness of fit across all four risk quantiles. *BRCA*-negative participants include those who tested negative for *BRCA1* and *BRCA2* mutations and untested participants. BOADICEA=Breast and Ovarian Analysis of Disease Incidence and Carrier Estimation Algorithm model. BCRAT=Breast Cancer Risk Assessment Tool. IBIS=International Breast Cancer Intervention Study model.

	Total breast cancer cases	Sensitivity (95% CI)	Specificity (95% CI)
Overall population			
N	619	..	..
BOADICEA	..	81·9% (78·6–84·9)	42·8% (42·1–43·6)
≥3·4%	507	..	..
<3·4%	112	..	..
BRCAPRO	..	45·9% (41·9–49·9)	76·1% (75·4–76·8)
≥3·4%	284	..	..
<3·4%	335	..	..
BCRAT	..	57·8% (53·8–61·8)	55·0% (54·2–55·8)
≥3·4%	358	..	..
<3·4%	261	..	..
IBIS	..	79·5% (76·1–82·6)	46·7% (45·9–47·5)
≥3·4%	492	..	..
<3·4%	127	..	..
***BRCA* negative**			
N	482	..	..
BOADICEA	..	77·0% (73·0–80·7)	44·4% (43·6–45·2)
≥3·4%	371	..	..
<3·4%	111	..	..
BRCAPRO	..	31·5% (27·4–35·9)	79·3% (78·7–80·0)
≥3·4%	152	..	..
<3·4%	330	..	..
BCRAT	..	64·3% (59·9–68·6)	54·4% (53·5–55·2)
≥3·4%	310	..	..
<3·4%	172	..	..
IBIS	..	73·9% (69·7–77·7)	48·4% (47·6–49·2)
≥3·4%	356	..	..
<3·4%	126	..	..
***BRCA* positive**			
N	137	..	..
BOADICEA	..	58·4% (49·7–66·8)	51·5% (48·3–54·7)
≥20%	80	..	..
<20%	57	..	..
BRCAPRO	..	25·5% (18·5–33·7)	80·7% (78·2–83·1)
≥20%	35	..	..
<20%	102	..	..
IBIS	..	57·7% (48·9–66·1)	56·5% (53·3–59·7)
≥20%	79	..	..
<20%	58	..	..
Age <50 years at baseline			
N	330	..	..
BOADICEA	..	68·8% (63·5–73·8)	63·4% (62·4–64·4)
≥3·4%	227	..	..
<3·4%	103	..	..
BRCAPRO	..	37·3% (32·0–42·7)	93·4% (92·8–94·0)

(Table 4 continues on next page)

two pedigree models, BRCAPRO and IBIS, which do not include a polygenic risk factor to account for the familial risk not explained by *BRCA1* and *BRCA2*, with BOADICEA, which does.[6] A key advantage of validating the models in the enriched ProF-SC is that we have more statistical power than cohorts not enriched on the basis of family history to examine how well the four models work across a wide spectrum of familial risk, including women with the highest risk of cancer, in whom accurate prediction is especially important. Most epidemiological cohorts include only a small number of participants at very high familial risk and, therefore, are useful only for validation of models in women with low or average familial risk of breast cancer. Both the Breast Cancer Family Registry and kConFab have been deliberately enriched for familial risk of breast cancer, which increases the statistical power for comparisons of risk models across a spectrum of absolute familial risk.[19]

Previous independent validations of models of breast cancer risk include two single-site prospective studies[13,15] that were included in ProF-SC, but our study includes data for a further five sites and has much longer follow-up data. A large validation study[26] of 567 prospective invasive cases in a screening cohort of 50 061 women compared IBIS (version 6) with BCRAT and showed improved performance with the pedigree information (ie, with IBIS), but the cohort was only followed up for a median of 3·2 years. A study[27] published in 2018 assessed 10-year risk with IBIS (version 8), which incorporated mammographic density data. IBIS was well calibrated, irrespective of the inclusion of mammographic density data, although their inclusion shifted some women towards higher risk quantiles. The independent, prospective validations of BOADICEA that have been done were smaller, with 15–115 cases in cohorts of 358–4176 women.[12,14,15] Only one study[11] has been done in which several pedigree models were compared with BCRAT; it included 64 prospective cases in a cohort of 3050 followed up for a mean of 5·3 years (range 0·1–15·0). The longer follow-up in ProF-SC compared with previous studies allowed us to assess the accuracy of the four included models in estimation of 10-year risks. Furthermore, the models that we used had been updated since most of the previous validation studies. We previously reported that discordances across models of breast cancer risk are greater for prediction of lifetime risks than for short fixed-time periods (eg, 5-year or 10-year risks) because the models were not developed on the basis of cohorts that have been followed up for a lifetime.[2] We are not aware of any models that have been validated for lifetime risk estimates, and the performance statistics that we report are based on, to our knowledge, the longest prospective follow-up data available to date.

Our results provide support for the use of pedigree models, such as BOADICEA or IBIS, to predict breast cancer risk in women younger than 50 years and showed that pedigree models work better than non-pedigree

models for all women aged 20–70 years in terms of calibration in women in lower familial risk quantiles. The key point—that the familial risk depends strongly on the age at diagnosis of the affected relative or relatives—is lost when binary constructs are used. Sensitivity analyses showed that the models did not differ in terms of whether they computed absolute risk after consideration of competing causes. We did not observe a difference in inference when we compared IBIS with and without this assumption. Competing mortality is likely, however, to have a greater effect when computing risks for women older than 70 years.

This study has some limitations. Even though our sample size was larger than that included in previous validations of pedigree-based models, it was still too small to assess model performance by race or ethnicity, or in more strictly defined age groups. Another limitation was that complete data were not available for all model inputs. There was a particular lack of mammographic density data, which are used in the latest version of IBIS. Data for lobular carcinoma in situ, which are included in some models, were also incomplete. We also did not have complete data for *BRCA1* and *BRCA2* mutations, but tried to minimise misclassification based on our mutation testing strategy. Although our study sample enriched with data for breast cancer family history is, on average, different from the general population, the value of ProF-SC is the wide range of absolute risk across women with average, intermediate, and high familial risk. This design ensures sufficient power to examine modifiable factors across the spectrum of risk[28] and to validate model performance in individuals at varying absolute risk. Although our study captures the full range of absolute risk, validations in other settings and countries with different absolute risks and exposures are needed.

Models that include multigenerational family history and comprehensive non-family-history risk factors produced more accurate risk estimates and had greater discrimination than those that did not include multigenerational family history data, but the amount of data needed for the models could potentially limit widespread clinical use, particularly in primary care settings. Obtaining the needed information directly from electronic medical records or having women enter their information directly into an online calculator before medical consultations[29] could help to alleviate the time pressure involved in collection of family history data. We found that adequate estimates of risk can be obtained using only information about unaffected and affected first-degree and second-degree relatives, including ages at diagnosis. Inclusion of additional data for non-family-history risk factors contributed little to risk discrimination as shown by sensitivity analyses in which this information was included. IBIS accurately predicted breast cancer risk on the basis of information about first-degree relatives and affected second-degree relatives only.

	Total breast cancer cases	Sensitivity (95% CI)	Specificity (95% CI)
(Continued from previous page)			
≥3·4%	123	..	..
<3·4%	207	..	..
BCRAT	..	31·8% (26·8–37·1)	77·0% (76·2–77·8)
≥3·4%	105	..	..
<3·4%	225	..	..
IBIS	..	70·3% (65·1–75·2)	63·2% (62·2–64·2)
≥3·4%	232	..	..
<3·4%	98	..	..
Age <50 years at baseline and *BRCA* negative			
N	218	..	..
BOADICEA	..	53·2% (46·4–60·0)	66·5% (65·5–67·5)
≥3·4%	116	..	..
<3·4%	102	..	..
BRCAPRO	..	6·9% (3·9–11·1)	98·0% (97·6–98·3)
≥3·4%	15	..	..
<3·4%	203	..	..
BCRAT	..	36·2% (29·9–43·0)	76·7% (75·9–77·6)
≥3·4%	79	..	..
<3·4%	139	..	..
IBIS	..	55·5% (48·6–62·2)	66·2% (65·2–67·2)
≥3·4%	121	..	..
<3·4%	97	..	..

Only participants in the Breast Cancer Prospective Family Study Cohort who had a breast cancer diagnosis during follow-up are shown (n=619); specificity values are based on those who were not diagnosed with breast cancer. BOADICEA=Breast and Ovarian Analysis of Disease Incidence and Carrier Estimation Algorithm model. BCRAT=Breast Cancer Risk Assessment Tool. IBIS=International Breast Cancer Intervention Study model.

Table 4: **Sensitivity and specificity of 10-year breast cancer risk prediction by model in participants diagnosed with breast cancer by *BRCA* mutation status**

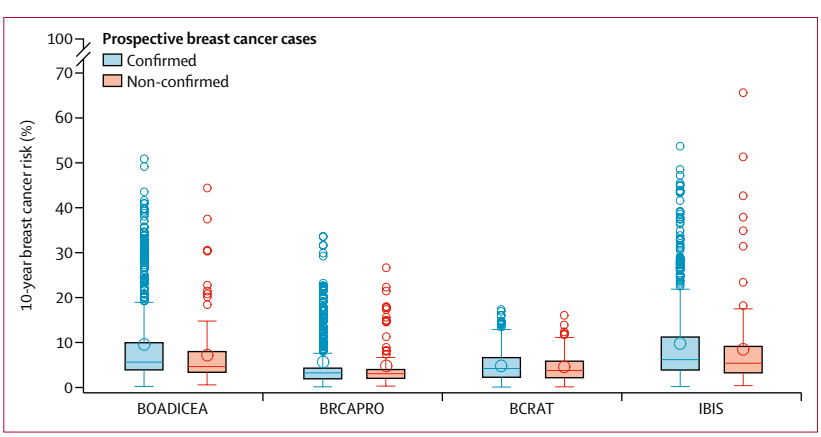


Figure 6: **10-year breast cancer risk prediction scores by model and diagnosis status**
Boxes represent the IQR of the distribution. The horizontal line within the box represents the median, and the vertical lines represent ±1·5 times the IQR. The large circles represent the mean, whereas the smaller circles represent outliers. Confirmed breast cancer cases include participants whose diagnosis was histologically confirmed (n=519). Non-confirmed breast cancer cases include those diagnoses not histologically confirmed (n=100). BOADICEA=Breast and Ovarian Analysis of Disease Incidence and Carrier Estimation Algorithm model. BCRAT=Breast Cancer Risk Assessment Tool. IBIS=International Breast Cancer Intervention Study model.

BCRAT accurately predicted risk in women aged 50 years or older on the basis of information about affected first-degree relatives only. Based on our results, we do not recommend use of BCRAT in women younger than 50 years. Accurate prediction of breast cancer risk can be achieved with BOADICEA only when information is available for all first-degree and second-degree relatives. We found that pedigree-based models were better irrespective of age, and thus the assumption that family history is an important predictor only in young women is incorrect. Improvements could be made to IBIS and BOADICEA to calibrate risk estimates when information is available only for first-degree relatives. The accuracy of risk estimates could be substantially improved by ensuring that family history data (particularly for all first-degree and second-degree relatives) are collected from all women, if possible.

Although both BOADICEA and IBIS performed similarly, further improvements in the accuracy of predictions could be possible with hybrid models that incorporate the polygenetic risk component of BOADICEA and the non-family-history risk factors included in IBIS. Reports suggest that the addition of polygenic risk scores based on breast-cancer-associated single-nucleotide polymorphisms will further increase breast cancer risk discrimination.[30,31] Most models are being extended to include one or more of mammographic density, polygenic risk factors, or other major genetic risk factors. Furthermore, new definitions of mammographic density seem to result in stronger risk discrimination at younger ages.[32] There is also a strong association between breast cancer risk and markers of DNA repair[33] and adducts,[34] suggesting that incorporation of these variables into risk models might lead to gains in discrimination. In addition to realising the full potential of polygenic risk and new phenotypic biomarkers, inclusion of modifiable lifestyle factors and environmental exposures could improve risk assessment in pedigree models. In particular, quantification of the effect on risk estimates of modifiable exposures—including physical activity, alcohol consumption, and breastfeeding—could help to improve prevention and risk assessment.[35]

In summary, our study shows that even though some pedigree models are well calibrated, they overpredict risk for women in the upper quantiles of familial risk and underpredict risk for those in the lower quantiles. Our findings suggest that all women would benefit from risk assessment that involves collection of detailed family histories, and that all risk models would be improved by inclusion of pedigree information.

Contributors

MBT and JLH conceived the study, obtained funding, collected and analysed data, and interpreted the findings. MBT and RJM analysed data, interpreted the findings, and co-led writing of the Article. ILA, EMJ, K-AP, MBD, and SSB conceived the study, obtained funding, collected data, and interpreted the findings. YL, ASW, NL, NZ, and GSD analysed data and interpreted the findings. WKC, JAK, MCS, RLM, DG, GGG, S-AM, MLF, and GG helped to conceive the study and obtain funding, collected data, and interpreted the findings. RB, PCW, and SN, coordinated the data collection and interpreted the findings. All authors contributed to the writing of the Article.

Declaration of interests

GSD reports grants from Genetic Technologies. MLF reports personal fees and grants from AstraZeneca, and personal fees from MSD. K-AP holds a patent for the System and Process of Cancer Risk Estimation (Australian Innovation Patent). All other authors declare competing interests.

Acknowledgments

This work was primarily supported by a grant from the US National Institute of Health (RO1CA159868). Additionally, the cohorts were supported through additional resources. The Australian Breast Cancer Family Registry was supported by the Australian National Health and Medical Research Council, the New South Wales Cancer Council, the Victorian Health Promotion Foundation, the Victorian Breast Cancer Research Consortium, Cancer Australia, and the National Breast Cancer Foundation. The six sites of the Breast Cancer Family Registry were supported by grants from the US National Cancer Institute (UM1 CA164920 and U01CA164920). This work was also supported by grants to kConFab and the kConFab follow-up study from Cancer Australia (809195 and 1100868), the Australian National Breast Cancer Foundation (IF 17 kConFab), the National Health and Medical Research Council (454508, 288704, and 145684), the Queensland Cancer Fund, the cancer councils of New South Wales, Victoria, Tasmania, and South Australia, and the Cancer Foundation of Western Australia. K-AP is a Practitioner Fellow of the Australian National Breast Cancer Foundation (PRAC-17-004). MCS is a Senior Research Fellow of the National Health and Medical Research Council, Australia. We thank the investigators and entire teams (past and current) of the Breast Cancer Family Registry and kConFab; Heather Thorne, Eveline Niedermayr, Lucy Stanhope, Yasmin Lisson, and Sandra Picken; the heads and staff of the Family Cancer Clinics (the genetic testing clinics where we identified probands); and the many families who contribute to the Breast Cancer Family Registry and kConFab. The content of this Article does not necessarily reflect the views or policies of the US National Cancer Institute or any of the collaborating centres in the Breast Cancer Family Registry. Mention of trade names, commercial products, or organisations does not imply endorsement by the US Government or the Breast Cancer Family Registry.

References

1 Shieh Y, Eklund M, Madlensky L, et al. Breast cancer screening in the precision medicine era: risk-based screening in a population-based trial. *J Natl Cancer Inst* 2017; **109:** djw290.

2 Quante AS, Whittemore AS, Shriver T, Hopper JL, Strauch K, Terry MB. Practical problems with clinical guidelines for breast cancer prevention based on remaining lifetime risk. *J Natl Cancer Inst* 2015; **107:** djv124.

3 Cintolo-Gonzalez JA, Braun D, Blackford AL, et al. Breast cancer risk models: a comprehensive overview of existing models, validation, and clinical applications. *Breast Cancer Res Treat* 2017; **164:** 263–84.

4 Whittemore AS. Evaluating health risk models. *Stat Med* 2010; **29:** 2438–52.

5 Tyrer J, Duffy SW, Cuzick J. A breast cancer prediction model incorporating familial and personal risk factors. *Stat Med* 2004; **23:** 1111–30.

6 Antoniou AC, Cunningham AP, Peto J, et al. The BOADICEA model of genetic susceptibility to breast and ovarian cancers: updates and extensions. *Br J Cancer* 2008; **98:** 1457–66.

7 Parmigiani G, Berry D, Aguilar O. Determining carrier probabilities for breast cancer-susceptibility genes *BRCA1* and *BRCA2*. *Am J Hum Genet* 1998; **62:** 145–58.

8 Gail MH, Brinton LA, Byar DP, et al. Projecting individualized probabilities of developing breast cancer for white females who are being examined annually. *J Natl Cancer Inst* 1989; **81:** 1879–86.

9 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:** 1541–48.

10 Rockhill B, Spiegelman D, Byrne C, Hunter DJ, Colditz GA. Validation of the Gail et al. model of breast cancer risk prediction and implications for chemoprevention. *J Natl Cancer Inst* 2001; **93:** 358–66.

11 Amir E, Evans DG, Shenton A, et al. Evaluation of breast cancer risk assessment packages in the family history evaluation and screening programme. *J Med Genet* 2003; **40:** 807–14.

12 Stahlbom AK, Johansson H, Liljegren A, von Wachenfeldt A, Arver B. Evaluation of the BOADICEA risk assessment model in women with a family history of breast cancer. *Fam Cancer* 2012; **11:** 33–40.

13 Quante AS, Whittemore AS, Shriver T, Strauch K, Terry MB. Breast cancer risk assessment across the risk continuum: genetic and nongenetic risk factors contributing to differential model performance. *Breast Cancer Res* 2012; **14:** R144.

14 Laitman Y, Simeonov M, Keinan-Boker L, Liphshitz I, Friedman E. Breast cancer risk prediction accuracy in Jewish Israeli high-risk women using the BOADICEA and IBIS risk models. *Genet Res (Camb)* 2013; **95:** 174–77.

15 MacInnis RJ, Bickerstaffe A, Apicella C, et al. Prospective validation of the breast cancer risk prediction model BOADICEA and a batch-mode version BOADICEACentre. *Br J Cancer* 2013; **109:** 1296–301.

16 Bredart A, Kop JL, Antoniou AC, et al. Use of the BOADICEA web application in clinical practice: appraisals by clinicians from various countries. *Fam Cancer* 2018; **17:** 31–41.

17 John EM, Hopper JL, Beck JC, et al. The Breast Cancer Family Registry: an infrastructure for cooperative multinational, interdisciplinary and translational studies of the genetic epidemiology of breast cancer. *Breast Cancer Res* 2004; **6:** R375–89.

18 Mann GJ, Thorne H, Balleine RL, et al. Analysis of cancer risk and *BRCA1* and *BRCA2* mutation prevalence in the kConFab familial breast cancer resource. *Breast Cancer Res* 2006; **8:** R12.

19 Terry MB, Phillips KA, Daly MB, et al. Cohort profile: The Breast Cancer Prospective Family Study Cohort (ProF-SC). *Int J Epidemiol* 2016; **45:** 683–92.

20 Phillips KA, Butow PN, Stewart AE, et al. Predictors of participation in clinical and psychosocial follow-up of the kConFab breast cancer family cohort. *Fam Cancer* 2005; **4:** 105–13.

21 Neuhausen SL, Ozcelik H, Southey MC, et al. *BRCA1* and *BRCA2* mutation carriers in the Breast Cancer Family Registry: an open resource for collaborative research. *Breast Cancer Res Treat* 2009; **116:** 379–86.

22 Van Calster B, Nieboer D, Vergouwe Y, De Cock B, Pencina MJ, Steyerberg EW. A calibration hierarchy for risk models was defined: from utopia to empirical data. *J Clin Epidemiol* 2016; **74:** 167–76.

23 Van Hoorde K, Van Huffel S, Timmerman D, Bourne T, Van Calster B. A spline-based tool to assess and visualize the calibration of multiclass risk predictions. *J Biomed Inform* 2015; **54:** 283–93.

24 Newson RB. Comparing the predictive powers of survival models using Harrell's C or Somers' D. *Stata J* 2010; **10:** 1–19.

25 Torre LA, Islami F, Siegel RL, Ward EM, Jemal A. Global cancer in women: burden and trends. *Cancer Epidemiol Biomarkers Prev* 2017; **26:** 444–57.

26 Brentnall AR, Harkness EF, Astley SM, et al. Mammographic density adds accuracy to both the Tyrer-Cuzick and Gail breast cancer risk models in a prospective UK screening cohort. *Breast Cancer Res* 2015; **17:** 147.

27 Brentnall AR, Cuzick J, Buist DSM, Bowles EJA. Long-term accuracy of breast cancer risk assessment combining classic risk factors and breast density. *JAMA Oncol* 2018; **4:** e180174.

28 Hopper JL, Dite GS, MacInnis RJ, et al. Age-specific breast cancer risk by BMI and familial risk: prospective family study cohort (ProF-SC). *Breast Cancer Res* 2018; **20:** 132.

29 Collins IM, Bickerstaffe A, Ranaweera T, et al. iPrevent(R): a tailored, web-based, decision support tool for breast cancer risk assessment and management. *Breast Cancer Res Treat* 2016; **156:** 171–82.

30 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.

31 van Veen EM, Brentnall AR, Byers H, et al. Use of single-nucleotide polymorphisms and mammographic density plus classic risk factors for breast cancer risk prediction. *JAMA Oncol* 2018; **4:** 476–82.

32 Nguyen TL, Choi YH, Aung YK, et al. Breast cancer risk associations with digital mammographic density by pixel brightness threshold and mammographic system. *Radiology* 2018; **286:** 433–42.

33 Machella N, Terry MB, Zipprich J, et al. Double-strand breaks repair in lymphoblastoid cell lines from sisters discordant for breast cancer from the New York site of the BCFR. *Carcinogenesis* 2008; **29:** 1367–72.

34 Shen J, Liao Y, Hopper JL, Goldberg M, Santella RM, Terry MB. Dependence of cancer risk from environmental exposures on underlying genetic susceptibility: an illustration with polycyclic aromatic hydrocarbons and breast cancer. *Br J Cancer* 2017; **116:** 1229–33.

35 Colditz GA, Bohlke K, Berkey CS. Breast cancer risk accumulation starts early: prevention must also. *Breast Cancer Res Treat* 2014; **145:** 567–79.