

Performance of Single-Nucleotide Polymorphisms in Breast Cancer Risk Prediction Models: A Systematic Review and Meta-analysis

Si Ming Fung¹, Xin Yi Wong¹, Shi Xun Lee¹, Hui Miao², Mikael Hartman^{3,4,5}, and Hwee-Lin Wee^{1,3}



Abstract

Background: SNP risk information can potentially improve the accuracy of breast cancer risk prediction. We aim to review and assess the performance of SNP-enhanced risk prediction models.

Methods: Studies that reported area under the ROC curve (AUC) and/or net reclassification improvement (NRI) for both traditional and SNP-enhanced risk models were identified. Meta-analyses were conducted to compare across all models and within similar baseline risk models.

Results: Twenty-six of 406 studies were included. Pooled estimate of AUC improvement is 0.044 [95% confidence interval (CI), 0.038–0.049] for all 38 models, while estimates

by baseline models ranged from 0.033 (95% CI, 0.025–0.041) for BCRAT to 0.053 (95% CI, 0.018–0.087) for partial BCRAT. There was no observable trend between AUC improvement and number of SNPs. One study found that the NRI was significantly larger when only intermediate-risk women were included. Two other studies showed that majority of the risk reclassification occurred in intermediate-risk women.

Conclusions: Addition of SNP risk information may be more beneficial for women with intermediate risk.

Impact: Screening could be a two-step process where a questionnaire is first used to identify intermediate-risk individuals, followed by SNP testing for these women only.

Introduction

Breast cancer is the most common cancer among women and is rising in incidence worldwide (1, 2). In this era of precision medicine, there is interest in applying tailored breast cancer screening and prevention strategies based on a woman's specific risk (3). Many risk factors have been identified and risk prediction models developed to quantify the combined effect of these factors (4). These models can be used to estimate a woman's individual risk, advise patients, inform screening and direct breast cancer research (4, 5). The models include the Breast Cancer Risk Assessment Tool (BCRAT; ref. 6), International Breast Intervention Study (IBIS) Breast Cancer Risk Evaluation Tool (also known as Tyrer–Cuzick model; ref. 7), BRCAPRO developed by Parmigiani and colleagues (8), Breast and Ovarian Analysis of Disease Incidence and Carrier Estimation Algorithm (BOADICEA; refs. 9, 10) and the Breast Cancer Surveillance Consortium (BCSC) Breast Cancer Risk Calculator (11, 12).

Over the past few years, rapid expansion of next generation DNA sequencing has led to an increasing discovery of breast cancer predisposition genes beyond *BRCA1* and *BRCA2*. For example, Shimelis and colleagues identified six other germline mutations (*BARD1*, *PALB2*, *RAD51D*, *BRIP1*, *RAD51C*, and *TP53*) that were associated with moderate to high risk of triple-negative breast cancer, a notoriously aggressive subtype, in a case–control study (13). Still, these genes account for only a small fraction of familial (14) and sporadic breast cancer (15, 16), leaving most of it unexplained. Research suggests that much of the missing heritability could be polygenic (14). A single SNP is associated with only low to moderate breast cancer risk. When combined, their effects could be significant as they occur at higher frequencies than high-penetrance mutations (17).

Over the past decade, many common, low-penetrance risk alleles for breast cancer have been identified by genome-wide association studies (GWAS) (14, 18, 19). The GWAS catalog, a freely available database of published SNP-trait associations lists about 53 studies and 1,272 associations related to breast cancer (20). As many as 182 SNPs have been identified as being associated with breast cancer, increasing our understanding of its heritability (21). These SNPs have not been included in existing breast cancer risk prediction models except for BCRAT, but there is intention to include them in future versions (22), as researchers investigate the potential of adding these SNPs to traditional risk prediction models, which results in SNP-enhanced risk prediction models.

Breast cancer screening aims for early detection of the disease and reduction in the associated mortality (23). SNP-enhanced risk prediction models may be able to estimate breast cancer risk more accurately and can translate into a more efficient risk-based screening program. Potential risk-adapted screening practices include starting screening later in women with lower risk,

¹Department of Pharmacy, Faculty of Science, National University of Singapore, Singapore. ²Genome Institute of Singapore, Agency for Science, Technology and Research, Singapore. ³Saw Swee Hock School of Public Health, National University of Singapore and National University Health System, Singapore. ⁴Division of General Surgery (Breast Surgery), National University Hospital, Singapore. ⁵Department of Surgery, Yong Loo Lin School of Medicine, National University of Singapore, Singapore.

Note: Supplementary data for this article are available at Cancer Epidemiology, Biomarkers & Prevention Online (<http://cebp.aacrjournals.org/>).

Corresponding Author: Hwee-Lin Wee, National University of Singapore, Tahir Foundation Building, 12 Science Drive 2, #09-01U, Singapore 117549, Republic of Singapore. Phone: 656-516-4975; E-mail: weehweelin@nus.edu.sg

doi: 10.1158/1055-9965.EPI-18-0810

©2018 American Association for Cancer Research.

increasing the recommended screening interval for these women or possibly not inviting them for screening at all (24). High-risk women could be recommended to begin screening at an earlier age, attend annual screening (25) or even engage in supplemental screening beyond mammography (26).

A couple of reviews have studied the performance of breast cancer risk prediction models without SNPs. Anothaisintawee and colleagues (27) found that discriminatory accuracy was poor to fair in both internal validation [concordance statistic (*c*-statistic): 0.53–0.66] and external validation (*c*-statistic: 0.56–0.63). This could be due to insufficient knowledge about risk factors, heterogeneous nature of breast cancer, and varying distributions of risk factors across populations. Another review by Meads and colleagues (28) found that none of the models reliably discriminate between those who did and did not develop breast cancer. Models such as BRCAPRO, BOADICEA and IBIS, which are preferred for women who come from families known to have a *BRCA1* or *BRCA2* mutation were not included. One other review by Evans and colleagues (29) assessed discriminatory accuracy of models using data from 1,933 women who attended the Family History Evaluation and Screening Program in United Kingdom. AUC was 0.716 for Claus model, 0.735 for BCRAT, 0.737 for BRCAPRO and 0.762 for IBIS model.

In this systematic review and meta-analysis, we aim to identify existing SNP-enhanced breast cancer risk prediction models and assess their performance, measured by discriminatory accuracy and improvement in predictive ability (AUC and NRI). The extent of improvement in performance, from the addition of genetic information in the SNP-enhanced risk models, was also evaluated.

Materials and Methods

This review is written in accordance to the PRISMA guidelines (30, 31).

Literature search strategy

A literature search of the EMBASE, Scopus, and PubMed databases was completed on January 31, 2018, without a specified time frame for published articles. The search strategy utilized three keywords: "breast cancer," "single nucleotide polymorphism," "risk prediction," and their synonyms. The exact search terms used are shown in Supplementary Table S1 of the Online Supplementary Material.

Study selection

Only studies published in English were considered. Studies were included if they fulfilled the following criteria: (i) empirical studies assessing risk prediction models for breast cancer in women, which reported outcomes using AUC or its equivalent—*c*-statistic and/or NRI; (ii) compared between SNP-enhanced risk prediction models and traditional risk prediction models. Studies excluded were genome-wide association studies, reviews, narratives, prognostic or diagnostic studies, model development studies, nonrisk prediction studies and studies that included only *BRCA* mutation carriers. Full texts of relevant studies were independently screened by three reviewers (Shi Xun Lee, Xin Yi Wong, and Si Ming Fung) for inclusion. Any disagreement was resolved, and consensus was reached among the reviewers.

Discriminatory accuracy and improvement in predictive ability of model

The primary outcomes in this review are AUC and NRI. The ROC curve is a plot of true positive rate against false positive rate. AUC is interpreted as the probability of assigning a higher predicted risk to a randomly selected individual with the outcome of interest than another randomly selected individual without the outcome (32) and provides a measure of discriminatory accuracy of the risk prediction model. A guideline on interpreting AUC values by Swets (33) has been widely used. It classifies the discriminatory accuracy of models into the following categories: noninformative (AUC = 0.5), low accuracy (0.5 < AUC < 0.7), moderate accuracy (0.7 < AUC < 0.9), high accuracy (0.9 < AUC < 1.0) and perfect accuracy (AUC = 1.0). However, the increase in AUC for a new model is often of a very small magnitude (34) and has been criticized as an insensitive measure of model improvement (35–37). Hence, Pencina and colleagues (34) proposed a novel index, NRI, for evaluating the improvement in predictive ability of a new model over an old model by analyzing the reclassification of subjects. NRI is defined in the following equation,

$$\text{NRI} = P(\text{up}|\text{event}) - P(\text{down}|\text{event}) + P(\text{down}|\text{non-event}) - P(\text{up}|\text{non-event})$$

where *P*(*x*) represents predicted probabilities and upward movement (up) as a change into higher risk category based on the new model and downward movement (down) as a change in the opposite direction. Positive NRI values indicate an overall correct reclassification in the new model as cases were moved upwards to higher risk thresholds and controls moved downwards to lower risk thresholds (34).

NRI has been previously computed and analyzed in studies that predicted risk of cardiovascular disease (38, 39), atrial fibrillation (40), diabetes mellitus (41, 42), end-stage liver disease (43) and end-stage renal disease (44). In the study by Paynter and colleagues (39), addition of a genotype to a prediction model based on traditional risk factors, high-sensitivity C-reactive protein and family history of premature myocardial infarction had no effect on *c*-index, a generalization of AUC (0.807 to 0.809) nor showed improvement in net reclassification (NRI –0.2%, *P* = 0.59). Yet, in the study by Lubitz and colleagues (40), a significant improvement in *c*-statistic contributed by premature familial atrial fibrillation [0.842 (95% confidence interval (CI), 0.826–0.858) to 0.846 (95% CI, 0.831–0.862), *P* = 0.004] did not correspond to a significant NRI [0.011 (95% CI, –0.021–0.042), *P* = 0.51].

Data extraction

The following data were extracted from the studies: publication year, number of case and control participants, age and ethnicity of participants, breast cancer subtype or hormone receptor status of case participants, country of study, study design, risk factors considered in the models, number of SNPs and the loci of SNPs included, method of incorporating SNPs in SNP-enhanced models, measure for evaluating discriminatory accuracy, and improvement in predictive ability.

Standard of reporting and quality assessment

Studies were evaluated against the 25-item checklist provided by the "Strengthening the Reporting of Genetic Risk Prediction Studies" (GRIPS) statement (Supplementary Table S2A, Online Supplementary Material; ref. 45). The GRIPS statement suggested

a standard reporting guideline for genetic risk prediction studies, but it does not serve as a quality assessment for studies.

We used the Newcastle-Ottawa Quality Assessment Scale (NOS; Supplementary Table S3A, Online Supplementary Material; ref. 46) to assess quality. NOS was identified by Cochrane (47) as a tool to assess methodological quality or risk of bias in non-randomized studies. It was based on three categories: selection of cases and controls, comparability of cases and controls and ascertainment of the exposure of interest (46). NOS scores range from 0 to 9. To the best of our knowledge, there are no established cut-offs for low, moderate and high quality. Hence, we have relied on previous literature (48) to define low quality as a score ≤ 5 , moderate quality as a score between 6 and 7 and high quality as a score between 8 and 9.

Statistical analyses

We extracted the number of case and control participants, the improvement in AUC (AUC of SNP-enhanced model minus AUC of baseline risk model without SNPs), overall AUC of SNP-enhanced model and the corresponding 95% confidence interval (CI) for our meta-analyses. These were conducted to pool studies by similar baseline risk models. We grouped the baseline risk prediction models into six groups: (i) BCRAT (5–6 risk factors of BCRAT), (ii) partial BCRAT (2–4 risk factors of BCRAT), (iii) partial BCRAT with additional risk factors, (iv) BCSC, (v) IBIS and (vi) other models. Forest plots were simultaneously obtained to visualize the trends, within each baseline model group.

A preformatted datasheet developed by Neyeloff and colleagues (49) was used to perform the meta-analyses and forest plots. The method suggested by Hanley and McNeil (32, 50) was used to calculate the standard error for the AUC of the studies in the datasheet. Heterogeneity was assessed using I^2 -statistic to assess the extent of variation between study population estimates. If significant heterogeneity was present (I^2 -statistic $\geq 50\%$), the random-effects model will be applied (51, 52).

Results

Figure 1 presents the literature search process. A total of 259 unique studies were identified from the databases. Reviewing of titles and abstracts yielded 84 potentially eligible articles of which 20 were included after a full-text review. Six more studies were identified from the references of relevant studies and 26 studies were included in the quality assessment and systematic review. One study did not evaluate AUC (53), so the remaining 25 studies were included in the meta-analysis.

Study population characteristics

Table 1 summarizes the study population characteristics, risk prediction model characteristics and model performance of the 26 studies which evaluated the discriminatory accuracy and/or predictive ability of baseline risk models and SNP-enhanced models. Supplementary Table S4 (Online Supplementary Material) provides more detailed information about the method of incorporating SNPs in model. The studies were published between 2010 to 2018. All were case-control in nature and no sample overlap was found in any study. The largest number of studies ($n = 14$) came from the United States (54–67). 13 studies evaluated risk prediction models among Caucasian or European populations (23, 24, 54–57, 59, 61–64, 68, 69), eight in Asian populations (53, 67, 70–75), three within mixed populations

(58, 60, 65), one among Australians (76), another among African Americans and Hispanics (66). Sample sizes ranged from 324 to 37,033 women. Reported measures of central tendency of age ranged from 44.2 (mean) to 64.6 (mean) years. Most studies evaluated the risk prediction models in the general population (24, 53, 55–58, 60, 61, 63, 64, 66–68, 70, 71, 73–76) while others specifically studied non-BRCA mutation carriers (54, 62, 69), post-menopausal (23, 59, 65, 72) or pre-menopausal women (72). Most studies included only patients with invasive breast cancer as case participants (23, 53, 56–61, 63, 64, 66–70, 73, 74, 76), while six included invasive or *in situ* breast cancer patients (24, 54, 55, 62, 71, 75) and two focused on estrogen receptor (ER)—positive patients with breast cancer (65, 72).

Risk prediction model characteristics

The number of SNPs incorporated in the risk prediction models ranged from 2 to 92. From the 25 studies, 38 risk prediction models were selected for our meta-analysis (Supplementary Table S5). Some studies evaluated more than one risk prediction model by including different risk factors in the baseline model (23, 55, 69) or different number of SNPs in the SNP-enhanced model (56, 66, 72). Seven models used BCRAT as baseline model (55, 59, 66, 69, 75, 76) and another four used partial BCRAT (57, 63, 68, 74). 11 models used partial BCRAT with additional risk factors (23, 55, 62, 64, 67, 70–73), four used BCSC (60, 61, 65), four used IBIS (24, 66, 69) and eight used other types of baseline models (54, 56, 58, 69). Risk factors included in the various traditional baseline models are summarized in Supplementary Table S6 (Online Supplementary Material). Most were designed for use in the general population (69). IBIS was developed using data from postmenopausal women and intended for use in high-risk populations (66, 69) while BRCAPRO was developed based on studies among individuals of Ashkenazi Jewish and European ancestry for pretest BRCA mutation (77). BOADICEA was developed using complex segregation analysis of breast and ovarian cancer based on a combination of families identified through population-based studies of breast cancer, and families with multiple affected individuals who had been screened for *BRCA1* and *BRCA2* mutations (78).

A variety of methods for incorporating SNPs into the models were observed. Among the 38 models, seven models (54, 56, 62, 63, 68) used the number of risk alleles present (0, 1, or 2) in each SNP, to incorporate genetic information in the model. Ten models (59, 67, 69, 76) used the Mealiffe method (59) to incorporate the SNPs in the model. In the Mealiffe method, the polygenic risk score (PRS) was calculated as the product of genotype relative risk values. Risk allele frequencies and published estimates of odds ratio per allele were based on a log-additive model. Independence between non-genetic risk factors and SNP risk for breast cancer was assumed. In five models (61, 64, 67, 71, 75), the PRS was calculated as the sum of the product of the number of risk allele copies of the selected SNPs and corresponding log odds ratio. Another three models used a Bayesian approach to calculate the PRS (60, 65). In four models (24, 70, 73, 74), a genetic risk score was calculated by other ways to include genetic information in the model (Supplementary Table S4, Online Supplementary Material). Two models (23) used a multiplicative penetrance model and one model (56) used a multiple log-additive model for incorporating the SNPs. Six models (55, 57, 58, 72) used other methods to include genetic information in the model.

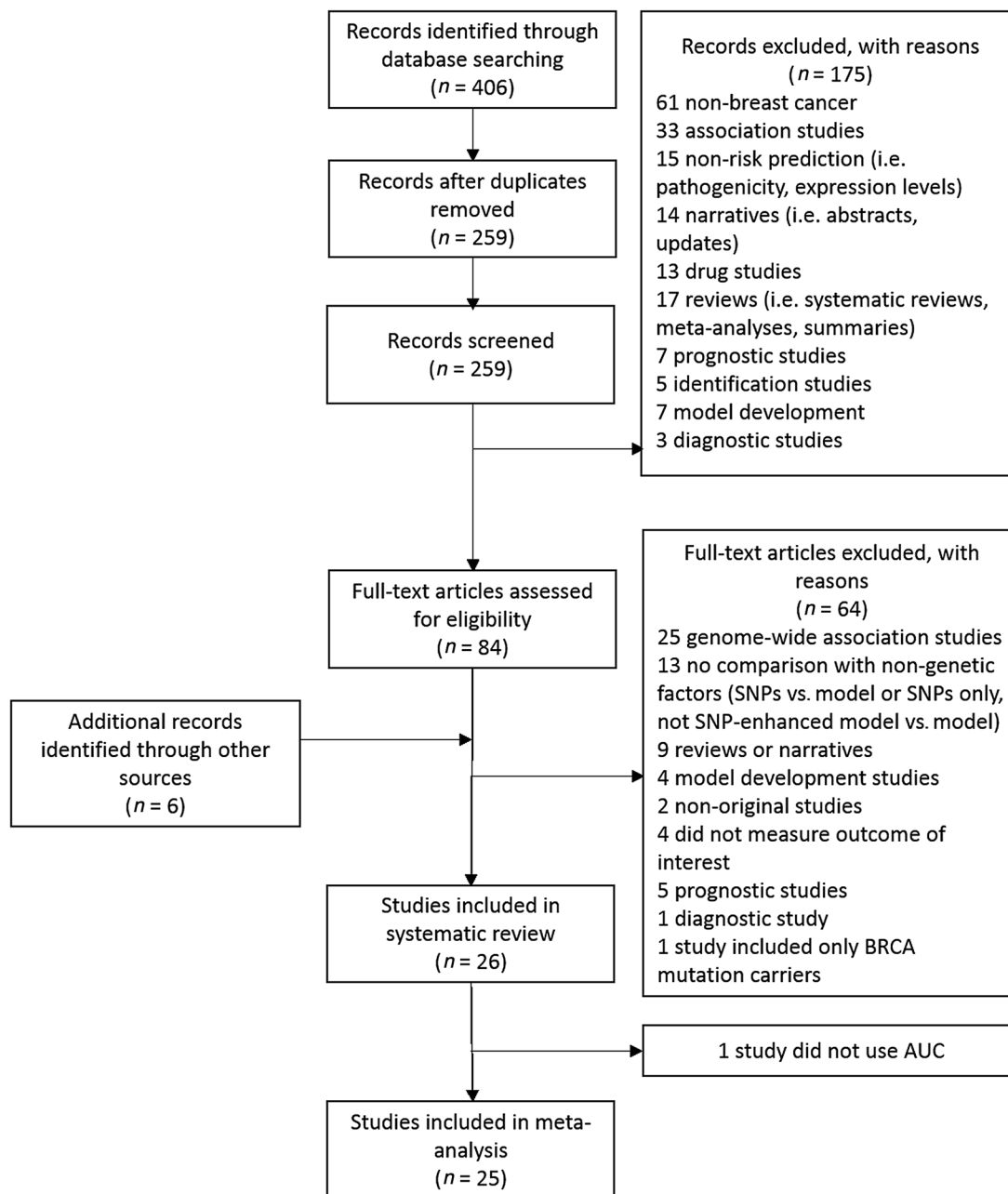


Figure 1.

Flowchart of selection process to include studies in systematic review and meta-analysis. The flowchart depicts the flow of information through the different phases of the systematic review and meta-analysis. It maps out the number of records identified, included and excluded, and the reasons for exclusions.

Standards of reporting and quality assessment

Details on the evaluation of studies against the GRIPS checklist can be found in Supplementary Table S2B (Online Supplementary Material). Many studies did not carry out internal validation or cross-validation (23, 24, 53, 55, 58, 61, 62, 64, 66–69, 72, 73, 76; Item 10). However, some used risk factors from established risk models (e.g., BCRA1) and SNPs identified from previously published GWAS and hence, the study itself might be considered a validation. Some studies did not carry out additional analyses or

had stated the results without discussion (23, 53, 54, 57–60, 62, 63, 66, 67, 69, 73, 74, 76; Item 20). Still, the main results were not significantly affected, as these analyses were not considered in the primary outcome. Some studies did not state the numbers or reasons for non-participation (23, 53–56, 62, 64–67, 69–74, 76; Item 14). Most studies did not report the measures of association between risk factors and outcome except that for SNPs (23, 55, 56, 58, 61, 65–67, 73–76; Item 16). In general, majority of the studies sufficiently reported their objectives, methods, results and

Table 1. Characteristics of studies that evaluated the predictive ability of SNP-enhanced breast cancer risk prediction models

Author	Mean (SD) or [Median (IQR)]	Age range	Sample size	Ethnic group	Type of participants	Patients' BC stage/ HRS	Country of study	Baseline model	No. of SNPs	Method of incorporating SNPs in model ^a	AUC of model without SNPs (95% CI)	AUC of SNP-enhanced model (95% CI)	Improv in AUC (95% CI) ^b	NRI (95% CI) ^c
Mealiffe, 2010 (59)	-	50-79	3,300	White	Post-menopausal	Invasive	US	BCRAT: Age, ethnicity, AGEMEN, AGEFLB, NUMREL, NBIOPS	7	Mealiffe method	0.557 (0.537-0.575)	0.594 (0.575-0.612)	0.037 (0.025-0.051)*	0.085
	-	50-79	11,588	European	General	Invasive	US, Poland	BCRAT: NUMREL, AGEMEN, AGEFLB, NBIOPS	10	Allele count	0.580	0.618	0.038	-
Zheng, 2010 (70)	Case: 49.7 (8.4) Control: 49.9 (8.8)	-	6,121	Chinese	General	Invasive	China	AGEMEN, AGEFLB; waist-to-hip ratio, benign breast disease (yes/no), family history of BC (yes/no), BMI, parity	8	Risk score	c-statistic: 0.618	c-statistic: 0.630	c-statistic: 0.012	-
	[Case: 50.0] [Control: 47.0]	-	718	Mixed	General	Invasive	US	Race, age, BMI	4	Other method	0.530	Non-epi: 0.600 Epi: 0.680	Non-epi: 0.070 Epi: 0.150*	-
Dai, 2012 (74)	Case: 50.1 (11.4) Control: 48.6 (12.3)	-	1,881	Chinese	General	-	China	AGEFLB, AGEMEN	5	Allele count	0.633	0.637 (0.619-0.655)	0.004	-
	-	-	3,093	Swedish	Post-menopausal	-	Sweden	Swe-Gail model: AGEMEN, AGEFLB, family history of BC (yes/no), benign breast disease (yes/no) Full model: mammographic percentage density, BMI, components of Swe-Gail model	18	Multiplicative penetrance model	0.548 (0.527-0.568)	0.615 (0.595-0.634)	0.067*	-
Darabi, 2012 (23)	-	-	1,873	-	-	-	-	-	-	Modified Gail model	0.638	0.658 (0.640-0.676)	0.020	-
	-	-	2,737	Caucasian (CBCS)	General	Invasive or <i>in situ</i> BC	US	BCRAT Model: age, AGEMEN, AGEFLB, biopsy histopathology, family history of BC (yes/no), benign breast surgery (yes/no)	9	Other method	0.588	0.615 (0.594-0.644)	0.027	-
Higginbotham, 2012 (55)	Case: 20-74 Control: 16-96	-	1,760	Caucasian (NBC)	Women with certain conditions ^d	-	-	BCRAT Model: age, AGEMEN, AGEFLB, NUMREL, NBIOPS, biopsy histopathology	-	-	0.573	0.601 (0.571-0.597)	0.048*	-

(Continued on the following page)

Table 1. Characteristics of studies that evaluated the predictive ability of SNP-enhanced breast cancer risk prediction models (Cont'd)

Author	Mean (SD) or [Median (IQR)] age	Age range	Sample size	Ethnic group	Type of participants	Patients' BC stage/ HRS	Country of study	Baseline model	No. of SNPs	Method of incorporating SNPs in model ^a	AUC of model without SNPs (95% CI)	AUC of SNP-enhanced model (95% CI)	Improv in AUC (95% CI) ^b	NRI (95% CI) ^c
Hüsing, 2012 (56)	Case: 57.9 (8.7)	-	13,836	European	General	Invasive	US	BMI, BMI interaction with menopausal status, age at first parity, age at menopause, no. of full term pregnancies, use of hormone replacement therapy, AGEMEN, alcohol consumption, smoking status	7	Allele count	0.564 (0.547–0.581)	0.591 (0.574–0.608)	0.027	-
	9								0.595 (0.579–0.612)		0.031	-		
	18								Log-additive model		0.605 (0.589–0.622)	0.041	0.083*	
	32								Allele count		0.604 (0.588–0.621)	0.040	-	
Sueta, 2012 (73)	Case: 51.6 (11.0) Control: 51.4 (11.0)	-	2,091	Japanese	General	-	Japan	Age, AGEMEN, AGEFLB, menopausal status, family history of breast cancer, BMI, regular exercise, referral pattern to hospital	7	Risk score	c-statistic: 0.665	c-statistic: 0.693	c-statistic: 0.028	-
Dite, 2013 (76)	Case: 45.9 (7.5) Control: 44.2 (7.4)	-	1,425	Australian	General	Invasive	Australia	BCRAT: age, ethnicity, AGEMEN, AGEFLB, NUMREL	7	Mealiffe method	Overall: 0.580 (0.550–0.610) 35–39 years: 0.600 (0.550–0.650) 40–49 years: 0.610 (0.550–0.660) 50–59 years: 0.540 (0.480–0.600)	0.610 (0.580–0.640)	0.030*	0.028
											0.650 (0.600–0.700)	0.050	0.021	
											0.630 (0.570–0.690)	0.020	0.074	
											0.560 (0.510–0.620)	0.020	-0.029	

(Continued on the following page)

Table 1. Characteristics of studies that evaluated the predictive ability of SNP-enhanced breast cancer risk prediction models (Cont'd.)

Author	Mean (SD) or [Median (IQR)] age	Age range	Sample size	Ethnic group	Type of participants	Patients' BC stage/ HRS	Country of study	Baseline model	No. of SNPs	Method of incorporating SNPs in model ^a	AUC of model without SNPs (95% CI)	AUC of SNP-enhanced model (95% CI)	Improv in AUC (95% CI) ^b	NRI (95% CI) ^c	
Xu, 2013 (63)	Case: 57 Control: 55	-	910	White Women	Women who have had a biological sister with BC	-	US	NUMREL, AGEMEN, AGEFLB, NBIOPS	9	Allele count	0.560	0.612	0.052	-	
Jupe, 2014 (57)	Case: 49.7 (9.2) Control: 49.7 (9.3)	-	5,022	Caucasian	General	-	US	BCRAT: age, AGEMEN, AGEFLB, NUMREL	22	Other method	35-39 years: 0.580 ^e	0.690	0.110*	-	
Lee, 2014 ^f (53)	Case: 54.9 (7.6) Control: 54.9 (7.5)	-	1,623	Chinese	General	-	Singapore	Level of education, AGEMEN, history of past breast biopsy, family history, BMI	51	Sum of product method	-	-	-	0.062	
Lee, 2015 (75)	-	50-64	24,161	Asian	General	Invasive or <i>in situ</i> BC	Singapore	BCRAT: age, ethnicity, AGEMEN, AGEFLB, NUMREL, breast biopsy history, BMI, mammographic density	75	Sum of product method	0.660 (0.640-0.680)	0.680 (0.660-0.690)	0.020	-	
Allman, 2015 (66)	-	> 49	7,421	African American	General	-	US	BCRAT, IBIS (no information on presence of atypical hyperplasia and family history variables)	75	Mealiffe method	BCRAT: 0.56 (0.53-0.59), IBIS: 0.51 (0.48-0.54)	BCRAT: 0.59 (0.56-0.61), IBIS: 0.55 (0.52-0.58)	BCRAT: 0.03, IBIS: 0.04	BCRAT: 0.033 (0.025-0.089), IBIS: 0.060 (0.005-0.113)	-
Vachon, 2015 (61)	Case: 60.8 (10.8) Control: 57.5 (11.5)	-	3,348	Hispanic	General	Invasive	US	BCSC v1.0 (Age, race/ethnicity, first degree family history, NBIOPS, BI-RADS density)	71	Sum of product method	BCRAT: 0.55 (0.51-0.60), IBIS: 0.53 (0.48-0.57)	BCRAT: 0.61 (0.56-0.66), IBIS: 0.59 (0.54-0.64)	BCRAT: 0.06, IBIS: 0.06	BCRAT: 0.082 (0.003-0.162), IBIS: 0.181 (0.085-0.273)	Case: 0.030 (0.007-0.049)* Control: 0.02 (-0.01-0.05)
Wu, 2015 (62)	Case: 62.0 (12.8)	29-90	768	Western	Non-BRCA mutation	Invasive or <i>in situ</i>	US	BCRAT (Age at biopsy, AGEMEN, NBIOPS,	10	Allele count	0.597	0.638	0.041*	-	

(Continued on the following page)

Table 1. Characteristics of studies that evaluated the predictive ability of SNP-enhanced breast cancer risk prediction models (Cont'd.)

Author	Mean (SD) or Median [IQR] age	Age range	Sample size	Ethnic group	Type of participants	Patients' BC stage/ HRS	Country of study	Baseline model	No. of SNPs	Method of incorporating SNPs in model ^a	AUC of model without SNPs (95% CI)	AUC of SNP-enhanced model (95% CI)	Improv in AUC (95% CI) ^b	NRI (95% CI) ^c
Wen, 2016 (67)	-	-	23,567	European	carriers	BC	US	NUMREL, parity), 5 BI-RADS mammographic features	44	Sum of product method	0.563	0.602	0.039 (0.026–0.051)	-
Burnside, 2016 (54)	62.0 (12.8)	395	768	East Asian	General	-	US	AGEMEN, AGEFLB, waist-to-hip ratio, prior benign breast disease (yes/no), family history of BC (yes/no)	10	Allele count	0.547 (0.50–0.587)	0.601 (0.562–0.641)	0.054	-
Dite, 2016 (69)	-	<50	1,155	White Women	Non-BRCA mutation carriers	Invasive or <i>in situ</i> BC	Australia	Age; Demographic model: age, AGEMEN, NUMREL, NBIOPS, parity	77	Mealiffe method	BOADICEA: 0.660 (0.630–0.700)	BOADICEA: 0.700 (0.670–0.730)	0.0400	BOADICEA: 0.040 (0.007–0.073)*
Shieh, 2016 (60)	56	36–86	931	Mixed	General	Invasive	US	BOADICEA and BRCAPRO, BCRAT (biopsies or atypical hyperplasia information not available), IBIS (hyperplasia or lobular carcinoma <i>in situ</i> information not available)	83	Bayesian method	BRCAPRO: 0.650 (0.620–0.680)	BRCAPRO: 0.690 (0.660–0.720)	0.040	BRCAPRO: 0.063 (0.030–0.094)*
Maas, 2016 (64)	-	-	37,033	White Women	General	Invasive	US	Fitted-BCSCv2.0 (First degree relative with BC (yes/no), BI-RADS breast density)	92	Sum of product method	BCRAT: 0.640 (0.600–0.680)	BCRAT: 0.67 (0.63–0.700)	0.030	BCRAT: 0.066 (0.019–0.110)*
Hsieh, 2017 (71)	Case: 51.3 (9.8) Control: 51.0 (10.1)	-	960	Asian	General	Invasive or <i>in situ</i> BC	Taiwan	Family history, AGEFLB, parity, AGEMEN, height, menopausal status, age at menopause, BMI, menopausal hormone therapy use, level of alcohol consumption, smoking status	13	Sum of product method	IBIS: 0.630 (0.570–0.600)	IBIS: 0.650 (0.610–0.680)	0.060	IBIS: 0.052 (0.015–0.088)*

(Continued on the following page)

Table 1. Characteristics of studies that evaluated the predictive ability of SNP-enhanced breast cancer risk prediction models (Cont'd.)

Author	Mean (SD) or Median [IQR] age	Age range	Sample size	Ethnic group	Type of participants	Patients' BC stage/ HRS	Country of study	Baseline model	No. of SNPs	Method of incorporating SNPs in model ^a	AUC of model without SNPs (95% CI)	AUC of SNP-enhanced model (95% CI)	Improv in AUC (95% CI) ^b	NRI (95% CI) ^c
Guo, 2017 (72)	Case: 47.0 (3.8) Control: 48.7 (2.5)	-	406	Japanese	Pre-menopausal	ER-positive, HER2-negative, invasive and non-invasive cancers	Japan	Age, BMI, AGEMEN, Age at menopause (postmenopausal), no. of pregnancy, parity (yes/no), age at primiparity, breastfeeding, benign breast disease (yes/no), family history of breast cancer (yes/no), hyperlipidemia (yes/no), diabetes mellitus (yes/no), testosterone serum levels, 25-hydroxyvitamin D serum levels	4	Other method	0.708	0.785	0.077	-
	Case: 64.6 (7.2) Control: 61.5 (6.3)	-	752		Post-menopausal				2		0.693	0.764	0.071	-
Shieh, 2017 (65)	Case: 61.5 (57-67) Control: 62 (57-67)]	-	324	Mixed	Postmenopausal	ER-positive	US	BCSC v1.0 (Age, race, presence of first degree relative with BC, history of breast biopsy, BMI, BI-RADS breast density) BCSC v1.0; estradiol	83	Bayesian method	0.580 (0.500-0.650)	0.640 (0.570-0.710)	0.060	-
	Case: 60 Control: 59	53-65	9,363	European	General	Invasive or <i>in situ</i> BC	England	IBIS, VAS mammographic density	18	Risk score	0.640 (0.600-0.680)	0.670 (0.620-0.710)	0.030	-

NOTE: The risk factors of the BCRAT have been indicated as NUMREL: NUMBER of first-degree Relatives with breast cancer; AGEMEN: AGE at MENarche; AGEFLB: AGE at First Live Birth; NBIOPS: Number of previous breast BIOPSies according to a code by Gail (6).

Abbreviations: AUC, area under the ROC curve; BC, breast cancer; BCRAT, Breast Cancer Risk Assessment Tool (6); BCSC, Breast Cancer Surveillance Consortium (11, 12); BI-RADS, Breast Imaging Reporting and Data System; BMI, Body Mass Index; BOADICEA, Breast and Ovarian Analysis of Disease Incidence and Carrier Estimation Algorithm (9, 10); BRCAPRO (8); CBCS, Collaborative Breast Cancer Study; CI, confidence interval; ER, estrogen receptor; HER2, human epidermal growth factor receptor 2; HRS, hormone receptor status; IBIS, International Breast Intervention Study (7); Improv in AUC, improvement in AUC; IQR, interquartile range; NBC, Nashville Breast Cohort; NRI, net reclassification improvement; SD, standard deviation; VAS, visual analogue scale.

^aThe SNP-enhanced model showed a statistically significant improvement ($P < 0.05$) over the model without SNPs.

^bRefer to Supplementary Table S4 for detailed methods.

^cImprovement in AUC of SNP-enhanced model calculated by AUC of SNP-enhanced model – AUC of baseline model.

^dNRI of adding genetic information to baseline model.

^eWomen with benign parenchyma or fibroadenoma.

^fJupe et al study (57) only provided clear AUC values for this age group.

^gLee et al study (53) evaluated NRI but not AUC hence the study was not included in the meta-analysis.

presented appropriate discussion and conclusion. As for the study quality, the average NOS score obtained by the studies was 6.1 (range: 2–8), out of a maximum of 9. Out of 26 studies, nine were of low quality, 12 were of moderate quality and 5 of high quality (Supplementary Table S3B). Only two studies reported that the non-response rates were the same for both cases and controls.

Discriminatory accuracy and improvement in predictive ability of model

Two models reported the AUC-equivalent *c*-statistic values (70, 73) while the rest provided AUC values. Using the AUC classification provided by Swets (33), four SNP-enhanced models (65, 69, 72) had moderate accuracy ($0.7 < \text{AUC} < 0.9$) and the other 34 SNP-enhanced models (23, 24, 54–64, 66–68, 70, 71, 73–76) showed low accuracy ($0.5 < \text{AUC} < 0.7$). Nonetheless, all studies showed an improvement in AUC/*c*-statistic from baseline risk models to SNP-enhanced models, and the improvement ranged from 0.011 to 0.15. Statistically significant improvement ($P < 0.05$) in either AUC or NRI was observed in 15 models (see asterisks in Table 1, last two columns).

There were 13 models (53, 56, 59, 61, 66, 69, 76) that reported the NRI of the effect of including genetic information in the SNP-enhanced models (Table 2). The reported NRI values for all the overall populations were positive, which indicates an improvement in classification when genetic information was included. However, one study found that the addition of SNPs to the baseline model did not improve classification among a subgroup of women aged 50 to 59 years (76).

Meta-analysis of AUC

We found no observable trend between improvement in AUC of the SNP-enhanced risk prediction model and number of SNPs in the model. In the test for trend in AUC improvement across ordered groups according to panel size (2 to 9 SNPs: small; 10 to 44 SNPs: medium and 71 to 92 SNPs: large), we found no significant association ($P = 0.244$). We also estimated the Pearson's correlation coefficient between AUC improvement and number of SNPs, which is -0.0305 ($P = 0.856$). Hence, we conclude that there is no linear correlation between AUC improvement and number of SNPs.

In Fig. 2, the pooled estimate of AUC improvement is 0.044 (95% CI, 0.038–0.049) for all 38 models, while the pooled estimates of AUC improvement by baseline models ranged from 0.033 (95% CI, 0.025–0.041; Fig. 3A) for BCRAT to 0.053 (95% CI: 0.018 to 0.087, Fig. 3B) for partial BCRAT. The overall AUC of the SNP-enhanced models generally fall short of the cut-off of 0.7 for moderate accuracy (Supplementary Fig. S7). Pooled estimates of overall AUC was highest at 0.671 (95% CI, 0.649–0.694) for BCSC models, followed by 0.653 (95% CI, 0.632–0.674) when partial BCRAT and additional risk factors model was used as the baseline, 0.643 (95% CI, 0.602–0.685) for partial BCRAT, 0.627 (95% CI, 0.608–0.645) for other models, 0.622 (95% CI, 0.591–0.654) for BCRAT model and finally 0.612 (95% CI, 0.584–0.640) for IBIS model (Supplementary Fig. S8).

Prediction of risk by breast cancer subtype

It appears that the addition of genetic information offers greater benefit when the risk models were used to predict ER-positive diseases. Among studies that used partial BCRAT and additional

Table 2. Reported NRI values from included studies

Model	No. of risk factors	No. of SNPs	NRI Value (95% CI) ^a	Classification categories	Author, year
BCRAT (5–6 risk factors of BCRAT)	6	7	Overall: 0.085	<1.5%,	Mealiffe, 2010
	5	7	Overall: 0.028	1.5%–2.0%,	Dite, 2013
			35–39 years: 0.021	>2.0%	
			40–49 years: 0.074		
			50–59 years: –0.029		
		5	77	0.066 (0.019–0.110)	
Partial BCRAT (2–4 risk factors of BCRAT)	6	75	0.033 (0.025–0.089)	<1.5%,	Allman, 2015
	6	71	0.082 (0.003–0.162)	1.5%–<2.0%,	
	2 + 4	51	0.062	≥2.0%	Lee, 2014
				<1.0%,	
			1.0%–<1.5%,		
			1.5%–<2.0%,		
			2.0%–<2.5%		
			≥2.5%		
BOADICEA	8	77	0.040 (0.007–0.073)	<1.5%,	Dite, 2016
BRCAPRO	8	77	0.063 (0.030–0.094)	1.5%–2.0%,	
IBIS	12	77	0.052 (0.015–0.088)	>2.0%	
	6	75	0.060 (0.005–0.113)	<1.5%,	Allman, 2015
	6	71	0.181 (0.085–0.273)	1.5%–<2.0%,	
				≥2.0%	
BCSC	5	76	Case: 0.110 (0.070–0.150), Control 0.020 (–0.010–0.050)	<3%, ≥3%	Vachon, 2015
Other models	10	18	0.083	<1%, 1%–<1.66%, 1.66%–<3.5%, >3.5%	Hüsing, 2012

Abbreviations: BCRAT, Breast Cancer Risk Assessment Tool (6); BCSC, Breast Cancer Surveillance Consortium (11, 12); BOADICEA, Breast and Ovarian Analysis of Disease Incidence and Carrier Estimation Algorithm (9, 10); BRCAPRO (8); CI, confidence interval; IBIS, International Breast Intervention Study (7); NRI, net reclassification improvement.

^aPositive NRI values indicate an overall correct reclassification.

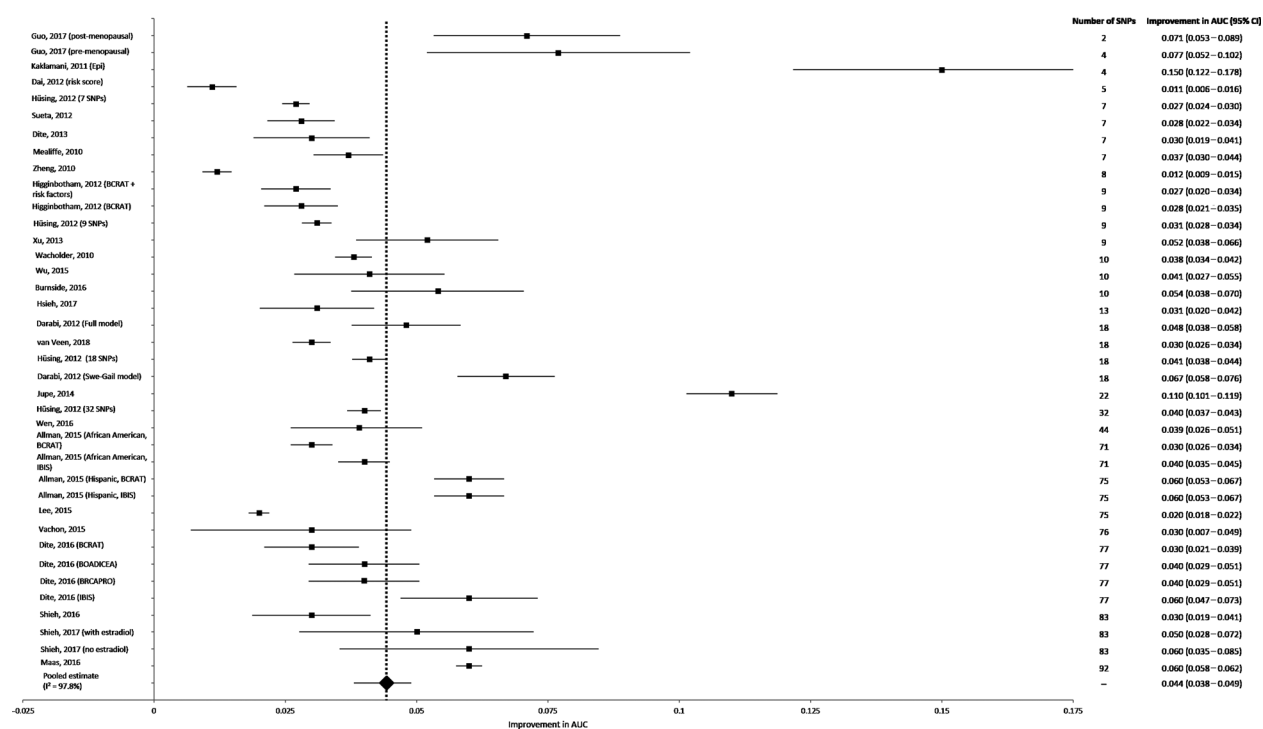


Figure 2.

Meta-analysis of the improvement in AUC, from the addition of SNPs into breast cancer risk prediction models by increasing number of SNPs (38 models). A forest plot for the improvement in AUC and increasing number of SNPs is shown. The 95% CIs are denoted by the black tails originating from the black square, which is the point estimate for improvement in AUC in each model. The combined improvement estimate across all models is represented by the black diamond, where the diamond width corresponds to the 95% CI bounds. Heterogeneity was assessed using I^2 -statistic to determine the extent of variation between study population estimates. As the I^2 -statistic is more than 50%, the random-effects model is applied.

risk factors, Guo and colleagues (72) had the largest AUC improvement and overall AUC (Fig. 3C; Supplementary Fig. S7C). It only involved patients with ER-positive, HER-2 negative breast cancers, while other studies in the same group included invasive or *in situ* breast cancer cases regardless of hormone receptor status (23, 55, 62, 64, 70, 71, 73). Another study by Shieh and colleagues (65), which included only ER-positive cases, also had larger improvement in AUC compared to the other studies within the group (Fig. 3D). Hüsing and colleagues investigated breast cancer risk by subtype and the predictive quality of the models were markedly better for ER-positive cancers rather than ER-negative (56).

However, it should be noted that majority of breast cancer cases used in the discovery and validation of SNPs from GWAS thus far have been ER-positive. Most of the known breast cancer loci show differences in relative risk by subtype. In particular, 6 of the 14 loci associated with ER-negative disease at genome-wide significance show no evidence of association with ER-positive disease (79). A recent meta-analysis (80) identified four new susceptibility loci for ER-negative disease, and that brings the total count to only 23. Studies included in our review did not focus on ER-negative specific SNPs or ER-negative cancers. This is likely why the risk prediction models appear to perform better in ER-positive cancers.

Discussion

The findings of this study suggest that the addition of genetic information into traditional risk prediction models for breast cancer improved model performance, although only slightly.

Four out of 38 models showed moderate discriminatory accuracy ($0.7 < \text{AUC} < 0.9$) (65, 69, 72). Of these, three were focused on ER-positive breast cancers and included endogenous hormones as risk factors in the baseline models (65, 72). All other studies displayed low discriminatory accuracy. All studies showed an overall improvement in discriminatory accuracy or reclassification when genetic information was added to the baseline risk prediction model, apart from a subgroup analysis among women aged 50 to 59 years (76). The greatest gain in AUC was in the study by Kaklamani and colleagues (58) using race, age and body mass index in the baseline model. They genotyped breast cancer cases and controls for four fat mass and obesity associated gene SNPs: *rs7206790*, *rs8047395*, *rs9939609*, and *rs1477196* and found that these genotypes provided powerful classifiers to predict breast cancer risk. A model containing epistatic interactions further improved the prediction accuracy to an AUC of 0.68 (improvement of 0.15, from 0.53).

The greatest gain in NRI was in the study by Allman and colleagues (66) using the IBIS as baseline model (Table 2, Fig. 3E). They studied the extent to which clinical breast cancer risk prediction models are improved by including information on susceptibility SNPs in African American or Hispanic women. The addition of 71 SNPs resulted in an NRI of 0.181 in Hispanic women. The IBIS model includes an extensive set of risk factors and is one of the most sensitive models for detecting risk for breast cancer (81, 82). For instance, it includes extended family history, *BRCA1/2* genetic status with non-genetic risk factors such as age, age at menarche, age at first live birth, age at menopause, parity,

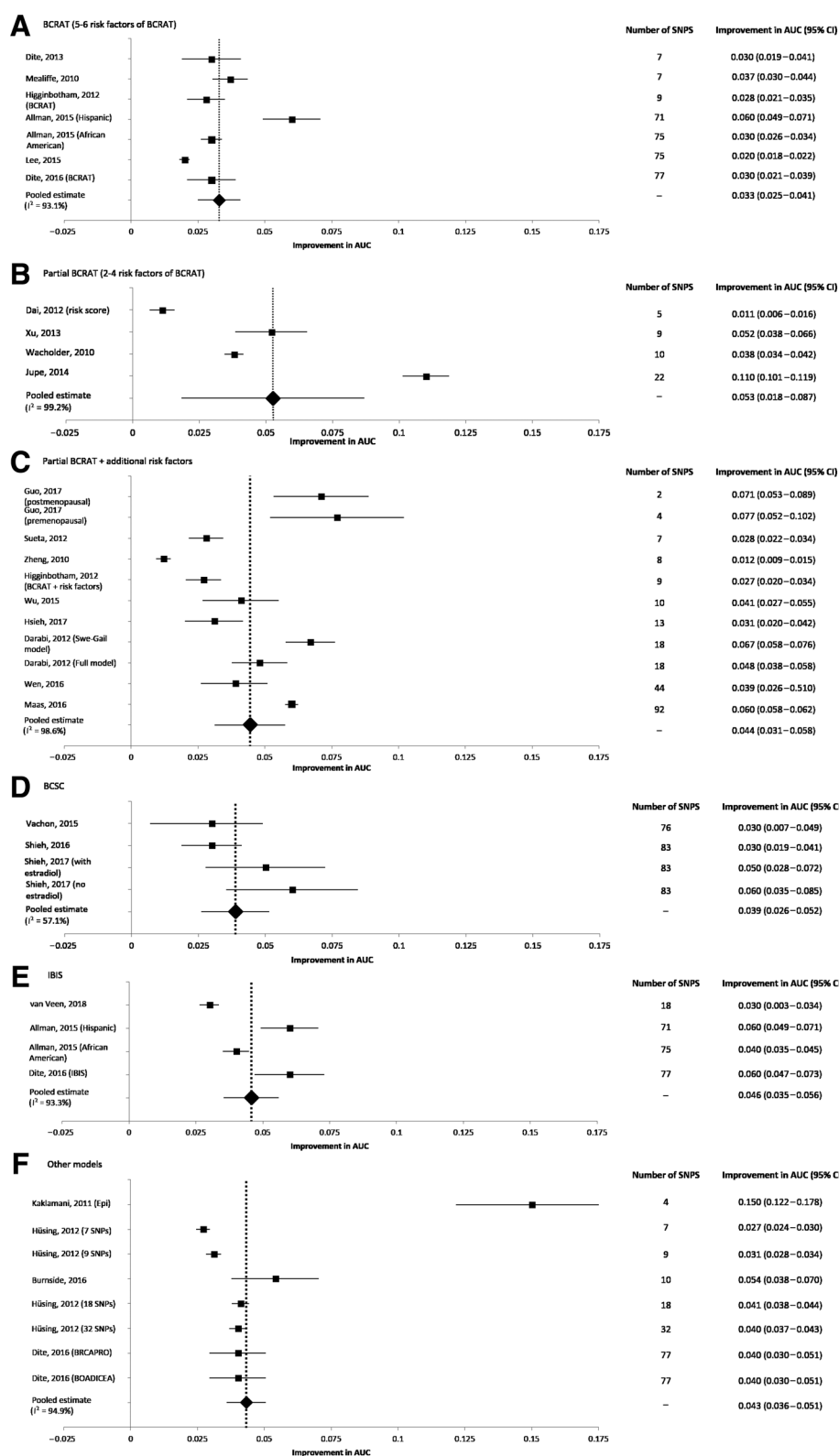


Figure 3. Meta-analysis of the improvement in AUC, from the addition of SNPs into breast cancer risk prediction models by increasing number of SNPs and baseline model group. **A–F**, BCRAT (5–6 risk factors of BCRAT; **A**); partial BCRAT (2–4 risk factors of BCRAT; **B**); partial BCRAT + additional risk factors (**C**), BCSC (**D**), IBIS (**E**), and other models (**F**). A forest plot for the improvement in AUC and increasing number of SNPs by baseline model group is shown. The 95% CIs are denoted by the black tails originating from the black square, which is the point estimate for improvement in AUC in each model. The combined improvement estimate within each baseline model group is represented by the black diamond, where the diamond width corresponds to the 95% CI bounds. Heterogeneity was assessed using I²-statistic to determine the extent of variation between study population estimates. As the I²-statistic is more than 50% for all model groups, the random-effects model is applied.

history of hormone replacement therapy use, history of atypical hyperplasia, history of lobular carcinoma *in situ*, height, and body mass index (7). Clinicians typically use models like BCRAT for

women deemed at average risk and models like IBIS for women whose family history and genetic information indicate above average risk (83).

Downloaded from <http://aacrjournals.org/cebp/article-pdf/28/3/506/2285701/506.pdf> by guest on 22 April 2025

There was no observable trend between improvement in AUC of the SNP-enhanced risk prediction model and number of SNPs included in the model. Other factors, such as age, ethnicity, loci of SNPs, and method of incorporating SNPs may also affect risk prediction. While we have stratified the studies by baseline risk models, variation in abovementioned study characteristics still exist within each group and could have confounded the findings and limited the ability to conclude on the association between number of SNPs and improvement in AUC. Therefore, we looked at studies that compared incremental number of SNPs internally. Hüsing and colleagues (56) showed that adding SNPs using allele count method resulted in a small but steady increase in discriminatory accuracy, when the number of genetic variants increase (7, 9, and 32 SNPs; Table 1). However, when 18 SNPs were added using a multiple log-additive model, the AUC was slightly higher than that of 32 SNPs added using allele count.

In another study that included only SNPs in the risk prediction model (84), the AUC increased from 0.591 to 0.622 (10 versus 22 SNPs) and from 0.622 to 0.684 (22 versus 77 SNPs). This suggests that the more SNPs the prediction model includes, the more discriminative it becomes. However, this was not observed in the model with 153 SNPs, where AUC was 0.650, lower than that achieved by model with 77 SNPs (84). This raises the question of whether the upper limit of SNPs predictive power has been reached. Also, GWAS have been primarily designed to capture common variation, and are thus underpowered to detect the effects of rare variants (85). We may need to relook at the strategies for genetic variants identification, such as employing whole genome sequencing as the technology advances (86).

Addition of genetic information in SNP-enhanced models may not offer benefit when the risk prediction models were used among older women. Jupe and colleagues (57) reported that the improvement in AUC between SNP-enhanced model and baseline model among women aged 50 to 54 years did not reach a statistically significant increase. Another study reported that the addition of genetic information did not improve classification among 50 to 59 years, as the NRI for this subgroup was negative (76). Under the partial BCRAT baseline model group, both the improvement and overall AUC of the study by Jupe and colleagues were much higher than the others (Fig. 3B; Supplementary Fig. S7B, Online Supplementary Material). This could be due to lower age (35–39 years) as compared with the other three studies which involved older women (63, 68, 74). While this may suggest that the addition of genetic information results in better prediction among younger women, it is worth noting that these women came from Marin County, California, a region with very high incidence of breast cancer (57). Also, in a study that investigated the value of using 77 SNPs as a PRS for risk stratification, Mavaddat and colleagues (87) found that the degree of attenuation of the family history odds ratio when adjusted by PRS was lower among younger women (below 40 years), due to the higher familial relative risk in this subgroup. This suggests that rarer genetic variants may be more important at young ages.

Currently, there is no clear evidence to guide the inclusion of SNP-SNP interactions in the SNP-enhanced model. Hüsing and colleagues (56) performed individual pair-wise SNP-interaction tests and found no evidence to include genetic interaction terms into the risk models. A large-scale study assessed all 2.5 billion possible two-way interactions between 70,917 breast cancer associated-SNPs and found no significant associations with breast cancer risk (88). Still, the authors cautioned that despite the large

sample size, the study might have been underpowered to detect very small interaction effects. Kaklamani and colleagues (58) reported a larger than expected improvement and overall AUC (Fig. 3F; Supplementary Fig. S7F). This study found that the inclusion of epistatic interactions that were significantly associated with breast cancer risk improved the model fit and reduced out-of-sample prediction error.

Few studies reported NRI as it was introduced recently in 2008 (34). NRI should also be used in conjunction with complementary statistical measures, such as AUC. An increase in AUC was observed for all the SNP-enhanced models and positive NRI values observed for all the overall models, which indicates overall correct reclassification by the addition of genetic information in the SNP-enhanced models. If we consider the incorporation of SNP information as a refinement to existing risk models, the benefit may be the greatest among those who are at borderline high or borderline low risk for breast cancer. This was reflected in the study by Mealiffe and colleagues (59), where the NRI value was significantly larger only when women with intermediate risk was included compared to when women with all risk categories were included. Two other studies in our review have shown that majority of the risk reclassification occurred in the group of women with intermediate risk. Shieh and colleagues (60) found that both cases and controls within the BCSC 5-year average- and intermediate-risk strata (1.00% to 1.66% and 1.67% to 2.49% respectively) were reclassified into the low-, high- and very high-risk strata, when PRS from 83 SNPs were added to the original BCSC model. Notably, the BCSCv2-PRS model classified nearly three times as many cases into the high-risk ($\geq 3\%$) strata compared with the BCSCv2 model. This points towards the possibility of administering the SNPs test in women with average/intermediate risk.

Another study by Darabi and colleagues (23) found that by adding mammographic density, body mass index and genetic information from 18 SNPs, 58% of those in the intermediate risk category under the original Swe-Gail model were reclassified into low or high risk. This is higher than that in the low- and high-risk groups (24% and 41%).

A study that investigated the cost-effectiveness of a 7SNP test for breast cancer risk (89) found it most cost-effective when given to patients with an intermediate lifetime risk of breast cancer. When they limited the test to patients who are most likely to have their risk category changed, they found that testing those with an intermediate Gail risk near 20% was relatively efficient. Two issues arise when we consider the role of SNP testing. The first issue is whether the screening should be a two-step process where a questionnaire is first used to identify individuals at intermediate risk and followed by administration of SNP testing for these women only. The second issue is whether this marginal gain in discriminatory accuracy offers value for money, that is, whether SNP testing is cost-effective. There are several advantages for the two-step process. Women at higher risk of breast cancer may be more likely to take up SNP testing as it can guide prevention and surveillance strategies (90, 91). SNP testing a smaller group of women may be more cost-effective than testing the whole population. Furthermore, fewer participants would be subjected to the potential loss of privacy of their genetic information. Still, we acknowledge that the evidence for targeted SNPs testing is in the early stage and more research is warranted.

Comprehensive risk assessment in a clinic setting is limited by time constraints, leading to incomplete and variable risk

evaluation (92). While a genetic test may be more efficient, it does not incorporate well-established risk factors such as hormonal and reproductive factors, family history of breast and/or ovarian cancer and mammographic density and cannot accurately predict a woman's breast cancer risk. Studies in our review shown that the AUC for SNPs-only models are typically in the 0.5 to 0.6 range, indicating low accuracy (56, 59, 60, 73, 74). An AUC of 0.5 indicates a random classification model not capable of predicting cancer.

A more opportune moment to assess breast cancer risk and inform women about risk reduction measures could be during the first mammography screening appointment, as highlighted by Evans and colleagues (22). The Predicting Risk Of Cancer At Screening study (93) has shown that it is feasible to collect questionnaire data from women when they attend the screening episode. In future, this may also allow the introduction of risk-stratified screening.

In the pooled estimates grouped by baseline models, overall AUC was the highest for BCSC followed by partial BCRAT and additional risk factors, partial BCRAT, other models, BCRAT and IBIS model (Supplementary Fig. S8). Researchers may keep this in mind when selecting the baseline model for their studies. BCSC was developed and validated in a multiracial and multiethnic population of over 1 million women undergoing mammography in the United States (11).

Some limitations include those associated with study design. The studies in this review were all case-control in nature, of which nine were nested case-control studies (24, 53, 55, 56, 59–61, 64, 65). The calibration of absolute event rates could not be evaluated in such a study design (59). In many studies (23, 53, 56, 58–61, 64, 65, 70, 71, 73, 74), the cases and controls were matched by age, and hence the discriminatory effect of age on breast cancer risk prediction could not be evaluated. The discrimination of the risk prediction score is likely to be higher in large, unmatched studies (65). There is a need for evaluation of model calibration in population-based cohorts so that the clinical validity of the models can be assessed further.

Another common limitation is that the SNP-enhanced risk prediction model in the studies might not be generalizable to other populations, especially those of different ethnicities. Different populations have different SNP profiles. Thus, for higher accuracy, the associations between SNPs and breast cancer should be specifically validated in ethnicities used in the study (94). For instance, in one study, 7 out of 8 breast cancer-associated SNPs were initially identified in previously published studies conducted in European women but were applied in the risk prediction model among Chinese women (70). The SNP-enhanced model showed low discriminatory accuracy (AUC = 0.630) and lower improvement in AUC (0.012) compared with the other included studies (70). Most studies also had missing information on some risk factors, such as history of atypical hyperplasia for BCRAT (57, 59, 63, 68, 69, 74, 76) and IBIS (69) and number of breast biopsies for BCRAT (57, 69, 76) and BCSC (65).

References

- Ginsburg O, Bray F, Coleman MP, Vanderpuye V, Eniu A, Kotha SR, et al. The global burden of women's cancers: a grand challenge in global health. *Lancet* 2017;389:847–60.
- Global Burden of Disease Cancer Collaboration, Fitzmaurice C, Allen C, Barber RM, Barregard L, Bhutta ZA, et al. Global, regional, and national cancer incidence, mortality, years of life lost, years lived with disability, and disability-adjusted life-years for 32 cancer groups, 1990 to 2015: a systematic analysis for the Global Burden of Disease Study. *JAMA Oncol* 2017;3:524–48.
- Esserman LJ, Wisdom Study and Athena Investigators. The WISDOM Study: breaking the deadlock in the breast cancer screening debate. *NPJ Breast Cancer* 2017;3:34.

In this review, we obtained information from published literature which did not always provide access to primary data. Given that all included studies showed an improvement in overall discriminatory accuracy and reclassification, publication bias is likely.

Conclusion

Genetic information improved the discrimination accuracy when added to traditional risk prediction models for breast cancer, with overall AUC being the highest in SNP-enhanced BCSC model. SNP-enhanced models have also demonstrated an improvement in overall reclassification for risk groups. We did not observe any association ($P = 0.244$ for non-parametric test for trend across ordered panel sizes) nor linear correlation ($P = 0.856$ for Pearson correlation coefficient) between AUC improvement and number of SNPs added (range: 2–92). In addition, we observed significant heterogeneity in the choice of baseline model, method of incorporating SNP information and population studied. To further advance knowledge in this field, guidance on appropriate study design and/or standardization of methodology may be required. To overcome the limitation of GWAS in identifying rare variants, newer technologies such as whole genome sequencing may be employed. Our findings suggest that the addition of genetic information in SNP-enhanced models may offer greater benefit when the models are used for risk prediction among subgroups, particularly women with intermediate risk. This implies that perhaps screening could be a two-step process where a questionnaire is first used to identify individuals at intermediate risk and followed by administration of SNP testing for these women only. Further research on targeted SNPs testing is warranted.

Disclosure of Potential Conflicts of Interest

No potential conflicts of interest were disclosed.

Acknowledgments

S.M. Fung was directly supported by the Singapore Ministry of Health Health Services Research Competitive Research Grant, administered by the National Medical Research Council (grant number: HSRG/13MAY006). The study sponsor had no role in the study design, collection, analysis and interpretation of data; in the writing of the manuscript; and in the decision to submit the manuscript for publication. We would like to thank Drs. Joanne Ngeow and Kelvin Bryan Tan, and members of the thesis advisory committee for X.Y. Wong for their comments on this manuscript, which is part of a thesis being prepared in fulfillment of the requirements for a Doctor of Philosophy at the National University of Singapore. We would also like to thank three anonymous reviewers for suggesting substantial improvements.

The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked *advertisement* in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

Received July 23, 2018; revised October 30, 2018; accepted December 3, 2018; published first December 18, 2018.

4. Cintolo-Gonzalez JA, Braun D, Blackford AL, Mazzola E, Acar A, Plichta JK, et al. Breast cancer risk models: a comprehensive overview of existing models, validation, and clinical applications. *Breast Cancer Res Treat* 2017;164:263–84.
5. Bevers TB, Anderson BO, Bonaccio E, Buys S, Daly MB, Dempsey PJ, et al. NCCN clinical practice guidelines in oncology: breast cancer screening and diagnosis. *J Natl Compr Canc Netw* 2009;7:1060–96.
6. Gail MH, Brinton LA, Byar DP, Corle DK, Green SB, Schairer C, 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.
7. Tyrer J, Duffy SW, Cuzick J. A breast cancer prediction model incorporating familial and personal risk factors. *Stat Med* 2004;23:1111–30.
8. Parmigiani G, Berry DA, Aguilar O. Determining carrier probabilities for breast cancer-susceptibility genes BRCA1 and BRCA2. *Am J Hum Genet* 1998;62:145–58.
9. Antoniou AC, Pharoah PP, Smith P, Easton DF. The BOADICEA model of genetic susceptibility to breast and ovarian cancer. *Br J Cancer* 2004;91:1580–90.
10. Antoniou AC, Cunningham AP, Peto J, Evans DG, Lalloo F, Narod SA, et al. The BOADICEA model of genetic susceptibility to breast and ovarian cancers: updates and extensions. *Br J Cancer* 2008;98:1457–66.
11. Tice JA, Cummings SR, Smith-Bindman R, Ichikawa L, Barlow WE, Kerlikowske K. Using clinical factors and mammographic breast density to estimate breast cancer risk: development and validation of a new predictive model. *Ann Intern Med* 2008;148:337–47.
12. Tice JA, Miglioretti DL, Li CS, Vachon CM, Gard CC, Kerlikowske K. Breast density and benign breast disease: risk assessment to identify women at high risk of breast cancer. *J Clin Oncol* 2015;33:3137–43.
13. Shimelis H, LaDuca H, Hu C, Hart SN, Na J, Thomas A, et al. Triple-negative breast cancer risk genes identified by multigene hereditary cancer panel testing. *J Natl Cancer Inst* 2018;110:855–62.
14. Sud A, Kinnersley B, Houlston RS. Genome-wide association studies of cancer: current insights and future perspectives. *Nat Rev Cancer* 2017;17:692–704.
15. Ricks-Santi LJ, Nie J, Marian C, Ochs-Balcom HM, Trevisan M, Edge SB, et al. BRCA1 polymorphisms and breast cancer epidemiology in the Western New York exposures and breast cancer (WEB) study. *Genet Epidemiol* 2013;37:504–11.
16. Kenemans P, Verstraeten RA, Verheijen RH. Oncogenic pathways in hereditary and sporadic breast cancer. *Maturitas* 2008;61:141–50.
17. Zheng W, Zhang B, Cai Q, Sung H, Michailidou K, Shi J, et al. Common genetic determinants of breast-cancer risk in East Asian women: a collaborative study of 23 637 breast cancer cases and 25 579 controls. *Hum Mol Genet* 2013;22:2539–50.
18. Easton DF, Eeles RA. Genome-wide association studies in cancer. *Hum Mol Genet* 2008;17:R109–15.
19. Gail MH. Value of adding single-nucleotide polymorphism genotypes to a breast cancer risk model. *J Natl Cancer Inst* 2009;101:959–63.
20. MacArthur J, Bowler E, Cerezo M, Gil L, Hall P, Hastings E, et al. The new NHGRI-EBI Catalog of published genome-wide association studies (GWAS Catalog). *Nucleic Acids Res*, 2017;45:D896–901.
21. Lilyquist J, Ruddy KJ, Vachon CM, Couch FJ. Common genetic variation and breast cancer risk - past, present, and future. *Cancer Epidemiol Biomarkers Prev* 2018;27.
22. Evans DG, Howell A. Can the breast screening appointment be used to provide risk assessment and prevention advice? *Breast Cancer Res* 2015;17:84.
23. Darabi H, Czene K, Zhao W, Liu J, Hall P, Humphreys K. Breast cancer risk prediction and individualised screening based on common genetic variation and breast density measurement. *Breast Cancer Res* 2012;14:R25.
24. Van Veen EM, Brentnall AR, Byers H, Harkness EF, Astley SM, Sampson S, 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.
25. Peres J. Studies support risk-based mammography screening. *J Natl Cancer Inst* 2012;104:1428–30.
26. Berg WA. Tailored supplemental screening for breast cancer: what now and what next? *AJR Am J Roentgenol* 2009;192:390–9.
27. Anothaisintawee T, Teerawattananon Y, Wiratkapun C, Kasamesup V, Thakkinstian A. Risk prediction models of breast cancer: a systematic review of model performances. *Breast Cancer Res Treat* 2011;133:1–10.
28. Meads C, Ahmed I, Riley RD. A systematic review of breast cancer incidence risk prediction models with meta-analysis of their performance. *Breast Cancer Res Treat* 2012;132:365–77.
29. Evans DG, Howell A. Breast cancer risk-assessment models. *Breast Cancer Res* 2007;9:213.
30. Moher D, Liberati A, Tetzlaff J, Altman DG; PRISMA Group. Preferred reporting items for systematic reviews and meta-analyses: The PRISMA Statement. *Int J Surg* 2010;8:336–41.
31. Moher D, Shamseer L, Clarke M, Ghersi D, Liberati A, Petticrew M, et al. Preferred reporting items for systematic reviews and meta-analysis protocols (PRISMA-P) 2015 statement. *Syst Rev* 2015;4:1.
32. Hanley JA, McNeil BJ. The meaning and use of the area under a receiver operating characteristic (ROC) curve. *Radiology* 1982;143:29–36.
33. Swets JA. Measuring the accuracy of diagnostic systems. *Science* 1988;240:1285–93.
34. Pencina MJ, D'Agostino RB Sr, D'Agostino RB Jr, Vasan RS. Evaluating the added predictive ability of a new marker: From area under the ROC curve to reclassification and beyond. *Stat Med* 2008;27:157–72.
35. Baker SG, Schuit E, Steyerberg EW, Pencina MJ, Vickers A, Moons KG, et al. How to interpret a small increase in AUC with an additional risk prediction marker: decision analysis comes through. *Stat Med* 2014;33:3946–59.
36. Cook NR. Assessing the incremental role of novel and emerging risk factors. *Curr Cardiovasc Risk Rep* 2010;4:112–9.
37. Cook NR. Use and misuse of the receiver operating characteristic curve in risk prediction. *Circulation* 2007;115:928–35.
38. Cook NR, Ridker PM. Advances in measuring the effect of individual predictors of cardiovascular risk: the role of reclassification measures. *Ann Intern Med* 2009;150:795–802.
39. Paynter NP, Chasman DI, Buring JE, Shiffman D, Cook NR, Ridker PM. Cardiovascular disease risk prediction with and without knowledge of genetic variation at chromosome 9p21.3. *Ann Intern Med* 2009;150:65–72.
40. Lubitz SA, Yin X, Fontes JD, Magnani JW, Rienstra M, Pai M, et al. Association between familial atrial fibrillation and risk of new-onset atrial fibrillation. *JAMA* 2010;304:2263–9.
41. Lyssenko V, Jonsson A, Almgren P, Pulizzi N, Isomaa B, Tuomi T, et al. Clinical risk factors, DNA variants, and the development of type 2 diabetes. *N Engl J Med* 2008;359:2220–32.
42. Meigs JB, Shrader P, Sullivan LM, McAteer JB, Fox CS, Dupuis J, et al. Genotype score in addition to common risk factors for prediction of type 2 diabetes. *N Engl J Med* 2008;359:2208–19.
43. Kim WR, Biggins SW, Kremers WK, Wiesner RH, Kamath PS, Benson JT, et al. Hyponatremia and mortality among patients on the liver-transplant waiting list. *N Engl J Med* 2008;359:1018–26.
44. Peralta CA. Detection of chronic kidney disease with creatinine, cystatin C, and urine albumin-to-creatinine ratio and association with progression to end-stage renal disease and mortality. *JAMA* 2011;305:1545.
45. Janssens AC, Ioannidis JP, Bedrosian S, Boffetta P, Dolan SM, Dowling N, et al. Strengthening the reporting of genetic risk prediction studies (GRIPS): explanation and elaboration. *Eur J Clin Invest* 2011;41:1010–35.
46. Wells GA, Shea B, O'Connell D, Peterson J, Welch V, Losos M, et al. The Newcastle-Ottawa Scale (NOS) for assessing the quality of nonrandomised studies in meta-analyses. Available from: http://www.ohri.ca/programs/clinical_epidemiology/oxford.asp.
47. Reeves BC, Deeks JJ, Higgins JPT, Wells GA. Chapter 13: Including non-randomized studies. In: Higgins JPT, Green S (editors), *Cochrane Handbook for Systematic Reviews of Interventions* Version 5.1.0. The Cochrane Collaboration; 2011.
48. Ahmed S, Shahid RK, Leis A, Haider K, Kanthan S, Reeder B, et al. Should noncurative resection of the primary tumour be performed in patients with stage IV colorectal cancer? A systematic review and meta-analysis. *Curr Oncol* 2013;20:e420–41.
49. Neyeloff JL, Fuchs SC, Moreira LB. Meta-analyses and Forest plots using a Microsoft excel spreadsheet: step-by-step guide focusing on descriptive data analysis. *BMC Res Notes* 2012;5:52.
50. Hajian-Tilaki KO, Hanley JA. Comparison of three methods for estimating the standard error of the area under the curve in ROC analysis of quantitative data. *Acad Radiol* 2002;9:1278–85.

51. Dinnes J, Deeks J, Kirby J, Roderick P. A methodological review of how heterogeneity has been examined in systematic reviews of diagnostic test accuracy. *Health Technol Assess* 2005;9:1–113.
52. Higgins JP, Thompson SG, Deeks JJ, Altman DG. Measuring inconsistency in meta-analyses. *BMJ* 2003;327:557–60.
53. Lee CP, Irwanto A, Salim A, Yuan JM, Liu J, Koh WP, et al. Breast cancer risk assessment using genetic variants and risk factors in a Singapore Chinese population. *Breast Cancer Res* 2014;16:R64.
54. Burnside ES, Liu J, Wu Y, Onitilo AA, McCarty CA, Page CD, et al. Comparing mammography abnormality features to genetic variants in the prediction of breast cancer in women recommended for breast biopsy. *Acad Radiol* 2016;23:62–9.
55. Higginbotham KS, Breyer JP, McReynolds KM, Bradley KM, Schuyler PA, Plummer WD, et al. A multistage genetic association study identifies breast cancer risk loci at 10q25 and 16q24. *Cancer Epidemiol Biomarkers Prev* 2012;21:1565–73.
56. Hüsing A, Canzian F, Beckmann L, Garcia-Closas M, Diver WR, Thun MJ, et al. Prediction of breast cancer risk by genetic risk factors, overall and by hormone receptor status. *J Med Genet* 2012;49:601–8.
57. Jupe ER, Dalessandri KM, Mulvihill JJ, Milke R, Knowlton NS, Pugh TW, et al. A steroid metabolizing gene variant in a polyfactorial model improves risk prediction in a high incidence breast cancer population. *BBA Clin* 2014;2:94–102.
58. Kaklamani V, Yi N, Sadim M, Siziopikou K, Zhang K, Xu Y, et al. The role of the fat mass and obesity associated gene (FTO) in breast cancer risk. *BMC Med Genet* 2011;12:52.
59. Mealiffe ME, Stokowski RP, Rhees BK, Prentice RL, Pettinger M, Hinds DA. Assessment of clinical validity of a breast cancer risk model combining genetic and clinical information. *J Natl Cancer Inst* 2010;102:1618–27.
60. Shieh Y, Hu D, Ma L, Gard CC, Leung J, Tice JA, et al. Breast cancer risk prediction using a combined clinical risk calculator and polygenic risk score. *J Gen Intern Med* 2016;31:S141.
61. Vachon CM, Pankratz VS, Scott CG, Haeberle L, Ziv E, Jensen MR, et al. The contributions of breast density and common genetic variation to breast cancer risk. *J Natl Cancer Inst* 2015;107:pii:dju937.
62. Wu Y, Liu J, Del Rio AM, Page DC, Alagoz O, Peissig P, et al. Developing a clinical utility framework to evaluate prediction models in radiogenomics. *Proc SPIE Int Soc Opt Eng* 2015;9416:pii:941617.
63. Xu Z, Bolick SC, Deroo LA, Weinberg CR, Sandler DP, Taylor JA. Epigenome-wide association study of breast cancer using prospectively collected sister study samples. *J Natl Cancer Inst* 2013;105:694–700.
64. Maas P, Bairdahl M, Joshi AD, Auer PL, Gaudet MM, Milne RL, et al. Breast cancer risk from modifiable and nonmodifiable risk factors among white women in the United States. *JAMA Oncol* 2016;2:1295–302.
65. Shieh Y, Hu D, Ma L, Huntsman S, Gard CC, Leung JWT, et al. Joint relative risks for estrogen receptor-positive breast cancer from a clinical model, polygenic risk score, and sex hormones. *Breast Cancer Res Treat* 2017;166:603–12.
66. Allman R, Dite GS, Hopper JL, Gordon O, Starlard-Davenport A, Chlebowski R, et al. SNPs and breast cancer risk prediction for African American and Hispanic women. *Breast Cancer Res Treat* 2015;154:583–9.
67. Wen W, Shu XO, Guo X, Cai Q, Long J, Bolla MK, et al. Prediction of breast cancer risk based on common genetic variants in women of East Asian ancestry. *Breast Cancer Res* 2016;18:124.
68. Wacholder S, Hartge P, Prentice R, Garcia-Closas M, Feigelson HS, Diver WR, et al. Performance of common genetic variants in breast-cancer risk models. *N Engl J Med* 2010;362:986–93.
69. Dite GS, Macinnis RJ, Bickerstaffe A, Dowty JG, Allman R, Apicella C, 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.
70. Zheng W, Wen W, Gao YT, Shyr Y, Zheng Y, Long J, et al. Genetic and clinical predictors for breast cancer risk assessment and stratification among Chinese women. *J Natl Cancer Inst* 2010;102:972–81.
71. Hsieh YC, Tu SH, Su CT, Cho EC, Wu CH, Hsieh MC, et al. A polygenic risk score for breast cancer risk in a Taiwanese population. *Breast Cancer Res Treat* 2017;163:131–8.
72. Guo J, Sueta A, Nakamura K, Yoshimoto N, Baba M, Ishida N, et al. Genetic and environmental factors and serum hormones, and risk of estrogen receptor-positive breast cancer in pre- and postmenopausal Japanese women. *Oncotarget* 2017;8:65759–69.
73. Sueta A, Ito H, Kawase T, Hirose K, Hosono S, Yatabe Y, et al. A genetic risk predictor for breast cancer using a combination of low-penetrance polymorphisms in a Japanese population. *Breast Cancer Res Treat* 2012;132:711–21.
74. Dai J, Hu Z, Jiang Y, Shen H, Dong J, Ma H, et al. Breast cancer risk assessment with five independent genetic variants and two risk factors in Chinese women. *Breast Cancer Res* 2012;14:R17.
75. Lee CP, Choi H, Soo KC, Tan MH, Chay WY, Chia KS, et al. Mammographic breast density and common genetic variants in breast cancer risk prediction. *PLoS One* 2015;10:e0136650.
76. Dite GS, Mahmoodi M, Bickerstaffe A, Hammet F, Macinnis RJ, Tsimiklis H, et al. Using SNP genotypes to improve the discrimination of a simple breast cancer risk prediction model. *Breast Cancer Res Treat* 2013;139:887–96.
77. Huo D, Senie RT, Daly M, Buys SS, Cummings S, Ogutha J, et al. Prediction of BRCA mutations using the BRCAPRO model in clinic-based African American, hispanic, and other minority families in the United States. *J Clin Oncol* 2009;27:1184–90.
78. Lee AJ, Cunningham AP, Kuchenbaecker KB, Mavaddat N, Easton DF, Antoniou AC. BOADICEA breast cancer risk prediction model: updates to cancer incidences, tumour pathology and web interface. *Br J Cancer* 2014;110:535–45.
79. Michailidou K, Beesley J, Lindstrom S, Canisius S, Dennis J, Lush MJ, et al. Genome-wide association analysis of more than 120,000 individuals identifies 15 new susceptibility loci for breast cancer. *Nat Genet* 2015;47:373–80.
80. Couch FJ, Kuchenbaecker KB, Michailidou K, Mendoza-Fandino GA, Nord S, Lilyquist J, et al. Identification of four novel susceptibility loci for oestrogen receptor negative breast cancer. *Nat Commun* 2016;7:11375.
81. Amir E, Freedman OC, Seruga B, Evans DG. Assessing women at high risk of breast cancer: a review of risk assessment models. *J Natl Cancer Inst* 2010;102:680–91.
82. Amir E, Shenton A, Laloo F, Moran A, Boggis C, Wilson M, Howell A. Evaluation of breast cancer risk assessment packages in the family history evaluation and screening programme. *J Met Genet* 2003;40:807–14.
83. Quante AS, Whitemore 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.
84. Wu Y, Abbey CK, Liu J, Ong I, Peissig P, Onitilo AA, et al. Discriminatory power of common genetic variants in personalized breast cancer diagnosis. *Proc SPIE Int Soc Opt Eng* 2016;9787:978706.
85. Mägi R, Asimit JL, Day-Williams AG, Zeggini E, Morris AP. Genome-wide association analysis of imputed rare variants: application to seven common complex diseases. *Genet Epidemiol* 2012;36:785–96.
86. Guo X, Shi J, Cai Q, Shu X, He J, Wen W, et al. Use of deep whole-genome sequencing data to identify structure risk variants in breast cancer susceptibility genes. *Hum Mol Genet* 2018;27:853–9.
87. Mavaddat N, Pharoah PD, Michailidou K, Tyrer J, Brook MN, Bolla MK, et al. Prediction of breast cancer risk based on profiling with common genetic variants. *J Natl Cancer Inst* 2015;107:pii:djv036.
88. Milne RL, Herranz J, Michailidou K, Dennis J, Tyrer JP, Zamora MP, et al. A large-scale assessment of two-way SNP interactions in breast cancer susceptibility using 46,450 cases and 42,461 controls from the breast cancer association consortium. *Hum Mol Genet* 2014;23:1934–46.
89. Folse HJ, Green LE, Kress A, Allman R, Dinh TA. Cost-effectiveness of a genetic test for breast cancer risk. *Cancer Prev Res* 2013;6:1328–36.
90. Howe R, Miron-Shatz T, Hanoch Y, Omer ZB, O'Donoghue C, Ozanne EM. Personalized medicine through SNP testing for breast cancer risk: clinical implementation. *J Genet Couns* 2015;24:744–51.
91. Hall TO, Renz AD, Snapinn KW, Bowen DJ, Edwards KL. Awareness and uptake of direct-to-consumer genetic testing among cancer cases, their relatives, and controls: the Northwest Cancer Genetics Network. *Genet Test Mol Biomarkers* 2012;16:744–8.
92. Sabatino SA, McCarthy EP, Phillips RS, Burns RB. Breast cancer risk assessment and management in primary care: provider attitudes, practices, and barriers. *Cancer Detect Prev* 2007;31:375–83.
93. Evans DG, Astley S, Stavrinou P, Harkness E, Donnelly LS, Dawe S, et al. Improvement in risk prediction, early detection and prevention of breast cancer in the NHS Breast Screening Programme and family history clinics: a dual cohort study. *Programme Grants Appl Res* 2016:4.
94. Maxwell KN, Nathanson KL. Common breast cancer risk variants in the post-COGS era: a comprehensive review. *Breast Cancer Res* 2013;15:212.