Research Article

Cancer Epidemiology, Biomarkers & Prevention

Genome-Based Prediction of Breast Cancer Risk in the General Population: A Modeling Study Based on Meta-Analyses of Genetic Associations

Moniek van Zitteren[1], Jeroen B. van der Net[2,3], Suman Kundu[1], Andrew N. Freedman[4], Cornelia M. van Duijn[1], and A. Cecile J.W. Janssens[1,2]

Abstract

Background: Genome-wide association studies identified novel breast cancer susceptibility variants that could be used to predict breast cancer in asymptomatic women. This review and modeling study aimed to investigate the current and potential predictive performance of genetic risk models.

Methods: Genotypes and disease status were simulated for a population of 10,000 women. Genetic risk models were constructed from polymorphisms from meta-analysis including, in separate scenarios, all polymorphisms or statistically significant polymorphisms only. We additionally investigated the magnitude of the odds ratios (OR) for 1 to 100 hypothetical polymorphisms that would be needed to achieve similar discriminative accuracy as available prediction models [modeled range of area under the receiver operating characteristic curve (AUC) 0.70–0.80].

Results: Of the 96 polymorphisms that had been investigated in meta-analyses, 41 showed significant associations. AUC was 0.68 for the genetic risk model based on all 96 polymorphisms and 0.67 for the 41 significant polymorphisms. Addition of 50 additional variants, each with risk allele frequencies of 0.30, requires per-allele ORs of 1.2 to increase this AUC to 0.70, 1.3 to increase AUC to 0.75, and 1.5 to increase AUC to 0.80. To achieve AUC of 0.80, even 100 additional variants would need per-allele ORs of 1.3 to 1.7, depending on risk allele frequencies.

Conclusion: The predictive ability of genetic risk models in breast cancer has the potential to become comparable to that of current breast cancer risk models.

Impact: Risk prediction based on low susceptibility variants becomes a realistic tool in prevention of nonfamilial breast cancer. *Cancer Epidemiol Biomarkers Prev; 20(1); 9–22. ©2011 AACR.*

Introduction

Recent genome-wide association studies (GWAS) have identified various novel breast cancer susceptibility variants (1–4). In contrast to BRCA1 and BRCA2, these susceptibility variants mainly have weak effects and contribute to small increases in breast cancer risk individually. It is commonly agreed that testing single susceptibility genes is not useful for prediction of breast cancer risk, but the question remains whether combining susceptibility loci in risk models could accurately identify women with markedly higher breast cancer risks. Genetic risk models could be useful in the general population to select at risk women for screening or preventive interventions, such as intensified mammography or MRI screening, the use of chemopreventive agents (tamoxifen or raloxifene), and surgical interventions such as oophorectomy and mastectomy. The predictive performance of genetic risk models has been investigated in simulation studies (5–8) and in many empirical studies, including studies on bladder, prostate, and breast cancers (9–12).

Whether genetic risk models will potentially be used in clinical or public health practice to select women at increased risk of breast cancer, first and foremost depends on the availability of an intervention that needs to be targeted. Pharoah and colleagues (13) proposed that the starting age of mammography screening can be varied according to women's genetic risks, rather than recommending screening to all women aged 50 years and more. Accordingly, women at higher genetic risk will have their first mammogram at a younger age and women at lower risk at a higher age than average. An alternative strategy is to vary the frequency of

Authors' Affiliations: Departments of [1]Epidemiology, [2]Public Health, and [3]Internal Medicine, Erasmus University Medical Center, Rotterdam, The Netherlands; and [4]Division of Cancer Control and Population Sciences, National Cancer Institute, Bethesda, Maryland

Corresponding Author: A.C.J.W. Janssens, Department of Epidemiology, Erasmus University Medical Center, PO box 2040, 3000 CA Rotterdam, The Netherlands. Phone: 31-10-7044232; Fax: 31-10-7044657. E-mail: a.janssens@erasmusmc.nl

doi: 10.1158/1055-9965.EPI-10-0329

mammograms: women at higher risk will undergo screening more frequently than women at lower risk.

Such differentiation of screening programs requires a risk model that can identify women at increased risk among all women in the population (14). For evaluations at the population level, this predictive performance of a test is assessed as the accuracy of tests to discriminate between women who will develop breast cancer and those who will not. The discriminative accuracy is generally expressed as the area under the receiver operating characteristics curve (AUC). Compared with the measures of reclassification that have been proposed (15–17), AUC is an overall measure of the predictive ability and reclassification a measure of the clinical relevance. Assessment of reclassification only is meaningful when clinically established risk thresholds are available.

Several simulation studies have demonstrated that to accurately classify individuals at high risk of disease, a risk model should either at least include a few genetic variants with strong effects or a high number of variants with small effects (6, 8, 18). We investigated the extent to which breast cancer risk in the general population can be predicted using genetic risk models. In a simulation study, we constructed risk models on the basis of currently identified polymorphisms and hypothetical variants to investigate the future potential in terms of AUC. For this purpose, we reviewed all meta-analyses of genetic association studies on breast cancer. Additionally, we investigated the magnitude of the odds ratios (OR) of 1 to 100 hypothetical polymorphisms that would be needed to achieve similar discriminative accuracy as available breast cancer risk prediction models (19–24).

Methods

Modeling strategy

We used a modeling procedure that has been developed and published previously (6), and which has also been used by others (8). Briefly, the procedure creates a data set with information on genotypes and disease status for a population of 10,000 women. The data set is constructed in such a way that the ORs and frequencies of the genotypes and the disease risk match the specified values, which are obtained from the literature. Predicted breast cancer risks are calculated using Bayes' theorem, which states that the posterior odds of breast cancer for each woman is obtained by multiplying the prior odds by the likelihood ratio (LR) of their genotype status on all polymorphisms. The prior odds is calculated from the baseline population breast cancer risk (p) using the formula $p/(1 - p)$. Under the assumption of independent genetic effects, that is, no linkage disequilibrium (LD) between the genetic variants, the LR is obtained by multiplying the LRs of all individual genotypes that are included in the risk model (25). The LRs of the genotypes of each single genetic variant are calculated from a genotype by disease status contingency table (6). This table is constructed from the frequencies and ORs of the genotypes and the population breast cancer risk. The table can also be constructed from allele frequencies and per-allele ORs when Hardy–Weinberg equilibrium is assumed for the distribution of the genotypes. The frequencies and ORs all are specified as study parameters and varied between the simulation scenarios (see below). The posterior odds are converted into breast cancer risks using the formula odds/(1 + odds). Our model does not include gene–gene and gene–environment interaction, because so far there is no strong empirical evidence for this.

Discriminative accuracy

The discriminative accuracy is the extent to which test results can discriminate between women who will develop breast cancer and those who will not (26). The AUC is the probability that the test correctly identifies the woman who will develop the disease from a pair of whom 1 will be affected and 1 will remain unaffected, and ranges from 0.5 (total lack of discrimination) to 1.0 (perfect discrimination). The AUC was obtained as the c-statistics by the R function *somers2*, which is available in the Hmisc library of R software. All simulations were repeated 20 times to obtain robust estimates of the AUC. In each repetition, the OR of each published variant will be obtained as a new random value from the 95% confidence interval (CI), assuming a normal distribution around the point estimate of the OR. All results are presented as averages of the repeated simulations with 95% CIs. Despite that we take a random OR for the published variants at each new simulation, the CIs of the ORs are small because it is theoretically possible to derive the ORs by a formula and due to the fact that a sample size of 10,000 women was used for this calculation. We choose a simulation approach, because the formula is extremely complex when different genetic models (per allele, dominant/recessive, and per genotype) are considered at the same time. Analyses were performed using R software (version 2.6.1; ref. 27).

Simulation scenarios

Three different scenarios were considered. In each scenario, genotypes and breast cancer status were simulated for 10,000 women, assuming a breast cancer lifetime risk of 10%. It is noted that this lifetime risk varies between populations (14), but our primary outcome measure, AUC, is independent of the magnitude of disease risk, that is, modeling a lower or higher risk will give similar results. The first scenario assessed the AUC for a risk model based on all polymorphisms that were investigated in published meta-analyses (see below), and for a risk model based on statistically significant polymorphisms only. The second scenario assessed the expected AUC when 2 to 5 times as many statistically significant polymorphisms with the same distribution of ORs and genotype frequencies would be known. The final scenario investigated the magnitude of the per-allele ORs of 1 to 100 polymorphisms that need to be added to the risk model to obtain AUCs similar to those of available breast cancer risk prediction models. Table 1 shows that the

Table 1. Area under the receiver operating characteristic curve of breast cancer risk models

Author	Risk factors in model	Population for model development	Population for validation	AUC (95% CI)
Clauset al. (21)	Age and family history (number and type[a] of first/second-degree relatives with breast cancer and their age at onset)	Population-based case–control data: 4,730 breast cancer cases and 4,688 matched controls[b]	1,933 women attending a family history evaluation and screening program; 5.27-year follow-up risk (34)	0.716 (0.648–0.784)
Costatino et al. (22)	Age, age at menarche, age at first live birth, number of previous breast biopsies, number of first-degree relatives with breast cancer, atypical hyperplasia	2,852 breast cancer cases and 3,146 controls; from population-based screening program	82,109 women in Nurses Health Study; 5-year risk (35)	0.58 (0.56–0.60)
			10,031 women from the Florence EPIC case–control study; 5-year risk (36)	0.588 (0.564–0.631)
			1,933 women attending a family history evaluation and screening program; 5.27-year follow-up risk (34)	0.735 (0.666–0.803)
Antoniou et al. (19)	Age, BRCA1/2 mutation status and family history (breast or ovarian cancer in relatives and their age at onset)	1,484 breast cancer cases from population-based screening program and 156 high-risk families.	NA	NA
Tyrer et al. (24)	Age at menarche, parity, age at first child birth, age at menopause, atypical hyperplasia, lobular carcinoma in situ, height, body mass index, presence of BRCA1/2 mutation, and family history (breast/ovarian cancer in first/second/third-degree relatives and their age at onset)	National cancer incidence rates, published risk figures, and cohort of daughters of patients with breast cancer	1,933 women attending a family history evaluation and screening program; 5.27-year follow-up risk (34)	0.762 (0.700–0.824)

(Continued on the following page)

Table 1. Area under the receiver operating characteristic curve of breast cancer risk models (Cont'd)

Author	Risk factors in model	Population for model development	Population for validation	AUC (95% CI)
Barlow et al. (20)	Premenopausal: age, prior breast procedure, first-degree family history of breast cancer, breast density; Postmenopausal: age, ethnicity, body mass index, age at birth first child, prior breast procedure, first-degree family history of breast cancer, current hormone therapy use, surgical menopause, previous mammographic outcome, breast density	Women participating in a mammography screening program (>million mammograms)	251,789 Women participating in a mammography screening program (>million mammograms), 5.3-year follow-up risk (37)	0.66 (0.651–0.669)
Gail et al. (23)	Age at menarche, number of affected mother or sisters and number of previous benign biopsy examinations	African American women participating in the Women's CARE Study: 1,622 cases and 1,647 controls	14,059 African American women who entered the Women's Health Initiative without a prior history of breast cancer, 7.57-year follow-up risk (23)	0.555 (0.535–0.575)

Abbreviation: NA, not available.

[a]Type of relative indicates whether mother, sister, or aunts were affected.

[b]Matched for geographic region and 5-year age category.

AUC of currently available breast cancer risk models ranges from 0.555 to 0.762. Therefore, we investigated AUC thresholds of 0.70, 0.75, or 0.80. The ORs of hypothetical variants needed to reach these thresholds were obtained for different frequencies of the risk alleles.

Literature search

PubMed and HuGE Navigator were searched for meta-analyses on genetic association studies on breast cancer published before May 2010. The PubMed search strategy was based on the keywords "breast cancer," "meta-analysis" in combination with "gene," "polymorphism," or "allele." Meta-analyses were selected if they were based on genetic association studies that applied a case–control design, included women only, focused on breast cancer risk and were written in English. Meta-analyses were excluded when the reported data were reused in a larger study on the same polymorphism. Summary ORs and genotype frequencies were extracted for all genetic models reported in this article, which could refer to per-allele analyses, comparisons of dominant or recessive effects, or comparisons of homozygous and heterozygous carriers with noncarriers. Summary ORs of the total population were extracted, because not all meta-analysis stratified their data for factors such as age or ethnicity. Therefore, all findings will predominantly apply to European and European American women.

For the simulation study, we calculated the AUC for a prediction model based on all polymorphisms and for a model based on statistically significant polymorphisms. For the latter, we selected polymorphisms for which at least 1 comparison yielded a statistically significant result. When multiple comparisons were statistically significant, we preferred the OR of the homozygous/heterozygous comparisons as the comparison of our first choice, the dominant/recessive comparisons as our second choice and the per-allele analysis as our last choice. Statistical significance was based on the nominal P value ($P < 0.05$) of the OR per comparison. For the statistically nonsignificant polymorphisms that were included in the simulation study, we also preferred the OR and genotypes for homozygous/heterozygous comparisons over the other comparisons. Because the modeling approach assumes independent effects of the polymorphisms in the risk model, we examined whether polymorphisms were in LD. When $R^2 > 0.40$, only the polymorphism with the lowest P value was included in the simulation analyses.

Results

Of the 217 retrieved articles from the literature, 107 met the inclusion criteria (Fig. 1). These articles described 199 meta-analyses on 103 polymorphisms in 70 genes. Ninety-six meta-analyses were excluded because the data had been reused in a larger meta-analysis on the same polymorphism and 3 because polymorphisms were in LD. Of the 96 polymorphisms, in 70 genes that were eligible for inclusion in the simulation analysis, 41 showed statistically significant results for at least 1 genetic model. Eleven of the 96 variants and 5 of the 41 statistically significant variants were in slight LD ($R^2 < 0.40$) with other variants in the model.

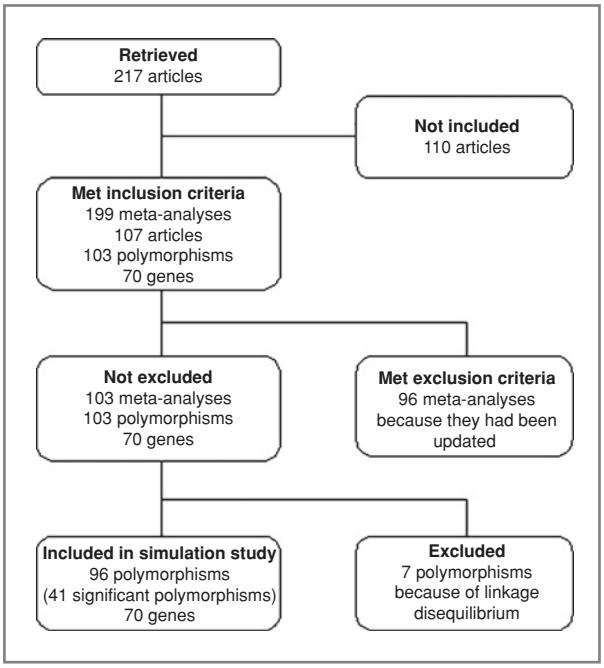

Figure 1. Selection of polymorphisms for inclusion in the simulation study.

The AUC was 0.68 for the risk model based on all 96 published polymorphisms and 0.67 for the risk model based on the 41 polymorphisms that were significantly associated with breast cancer risk (Table 2). When 2, 3, 4, or 5 times as many polymorphisms with the same distribution of ORs and genotype frequencies would be identified and included, AUCs were 0.73, 0.77, 0.80, and 0.82, respectively. As a comparison, the AUCs for the combinations of the 7 polymorphisms investigated by Gail et al. (5) and Pharoah et al. (13) were both 0.55.

Table 3 shows the magnitude of the ORs and genotype frequencies of genetic variants that would be needed in addition to the original 41 genetic risk variants to obtain AUCs of 0.70, 0.75, and 0.80. The table shows that to achieve an AUC of 0.70, the minimal OR of 5 additional genetic variants should be 1.5 (95% CI: 1.4–1.5) when their risk genotype frequencies are 0.30. To achieve an AUC of 0.75 with 20 to 100 additional genetic variants, the minimal ORs ranged from 1.2 to 2.1 depending on the frequencies of the risk genotypes. These values were 1.3 to 2.7 to achieve an AUC of 0.80.

Discussion

This study investigated to what degree genetic risk models can predict breast cancer in a general population setting. We estimated that the AUC would be 0.68 when all 96 polymorphisms investigated were included and

Table 2. Meta-analyses of genetic susceptibility variants for breast cancer risk

Gene	Variant	No. of studies	No. of cases	No. of controls	Referent variant	Comparison 1				Comparison 2				Reference[b]
						Associated variant	OR	95% CI	Frequency (%)	Associated variant	OR	95% CI	Frequency (%)	
1p11	**rs11249433 (T/C)**	9	9,335	10,263	TT	TC	**1.16**	**1.09–1.24**	46.7	CC	**1.30**	**1.19–1.41**	16.0	31
2q35	**rs13387042 (G/A)**	6	4,533	17,513	GG	GA	**1.11**	**1.03–1.20**	50.0[a]	AA	**1.44**	**1.30–1.58**	24.7[a]	2
3p24	**rs4973768 (C/T)**	32	30,256	34,063	CC	CT	**1.12**	**1.08–1.17**	49.7[a]	TT	**1.23**	**1.17–1.29**	22.1[a]	30
5p12	**rs2067980 (A/G)**	9	9,391	10,309	AA	AG	**1.08**	**1.02–1.15**	26.6	GG	**1.29**	**1.09–1.52**	2.7	31
5p12	**rs7716600 (C/A)**	9	9,400	10,321	CC	CA	**1.10**	**1.04–1.17**	34.3	AA	**1.28**	**1.13–1.45**	4.9	31
6q25	**rs2046210 (G/A)**	2	6,472	3,962	GG	GA	**1.36**	**1.24–1.49**	45.7	AA	**1.59**	**1.40–1.82**	13.6	32
8q24	**rs13281615 (T/C)**	20	21,860	22,578	TT	TC	**1.06**	**1.01–1.11**	48.0[a]	CC	**1.18**	**1.10–1.25**	16.0[a]	1
Chr 9	**rs1011970 (C/A)**	2	12,253	12,000	CC	CA	**1.07**	**1.01–1.13**	27.9	AA	**1.29**	**1.12–1.50**	2.8	4
Chr 10	**rs2380205 (G/A)**	2	12,235	11,961	GG	GA	0.95	0.90–1.01	49.3	AA	**0.89**	**0.82–0.95**	18.3	4
Chr 10	**rs10995190 (G/A)**	2	12,261	12,000	GG	GA	**0.84**	**0.79–0.89**	25.6	AA	**0.83**	**0.69–1.00**	2.1	4
Chr 10	**rs704010 (G/A)**	2	12,222	11,992	GG	GA	**1.11**	**1.05–1.17**	46.9	AA	**1.13**	**1.04–1.21**	16.0	4
Chr 11	**rs614367 (G/A)**	2	12,114	11,967	GG	GA	**1.16**	**1.10–1.23**	24.4	AA	**1.27**	**1.10–1.47**	3.0	4
14q24	**rs999737 (T/C)**	9	9,395	10,298	TT	TC	**0.94**	**0.88–0.99**	35.5	CC	**0.70**	**0.62–0.80**	6.1	31
17q23	**rs6504950 (G/A)**	33	30,470	33,302	GG	GA	**1.04**	**1.01–1.09**	39.4[a]	AA	**1.12**	**1.05–1.21**	7.3[a]	30
ADH1C	I350V	5	7,805	7,320	II	IV	0.97	0.85–1.10	48.0	VV	1.00	0.83–1.21	17.1	38
AKAP9	**M463I (G/T)**	7	9,523	13,770	GG	GT	**1.08**	**1.02–1.15**	46.4	TT	**1.17**	**1.08–1.27**	13.4	39
AR	**CAG repeat**	7	6,622	7,160	SS	SL	**0.61**	**0.38–0.97**	47.6	LL	**0.62**	**0.39–0.97**	30.1	40
AURKA	T91A	10	14,361	17,780	AA + TA	TT	0.99	0.96–1.01	17.5					41
BID	rs8190315 (A/G)	11	13,664	14,193	AA	AG	0.99	0.87–1.14	0.03					42
BRCA2	N372H	22	22,515	22,388	NN	NH	1.01	0.97–1.05	39.2	HH	1.05	0.97–1.13	7.5	43/44
CASP8	**D302H**	15	16,909	17,654	DD	DH	**0.89**	**0.85–0.94**	22.8	HH	**0.74**	**0.61–0.89**	33.0	45
CASP10	rs13010627 (G/A)	28	26,917	30,429	GG	GA	1.03	0.98–1.09	12.2	AA	0.94	0.72–1.22	0.4	42
CCND1	**G870A**	7	5,371	5,336	GG	GA	**1.12**	**1.01–1.23**	49.7	AA	**1.18**	**1.06–1.32**	21.1	46
CDKN1A	rs1801270 (C/A)	21	22,109	29,127	CC	CA	1.03	0.96–1.11	12.0	AA	1.32	0.99–1.76	0.3	47/48
CHEK2	**1100delC**	12	18,329	18,580	CC	delC	**2.40**	**1.80–3.20**	0.5					49
COMT	G472A	41	25,627	34,222	GG	GA	0.99	0.93–1.04	47.8	AA	0.96	0.88–1.04	23.8	50
CYP17	MspA1	15	4,227	4,730	A1/A1	A1/A2 + A2A2	0.98	0.89–1.07	64.2					51/52
CYP19	R264C (C/T)	8	2,355	3,592	C	T	1.06	0.93–1.21	9.4					53
CYP19	3-bp del/ins	10	7,720	7,454	Ins	Del	1.00	0.93–1.08	29.2					53
CYP19	$TTTA_n$	14	4,198	4,644		$TTTA_{12}$	0.91	0.72–1.15	5.1					53
						$TTTA_{10}$	**1.52**	**1.12–2.06**	2.4					53
CYP1A1	Ile462Val (A/G)	13	9,552	9,320	AA	AG	1.01	0.82–1.23	21.4	GG	1.04[c]	0.61–1.76	2.6	54
CYP1A1	T3801C	23	10,520	14,567	CC	TC	0.95	0.79–1.14	27.6	TT	0.93	0.72–1.19	66.7	55

(Continued on the following page)

Table 2. Meta-analyses of genetic susceptibility variants for breast cancer risk (Cont'd)

Gene	Variant	No. of studies	No. of cases	No. of controls	Referent variant	Comparison 1				Comparison 2				Reference[b]
						Associated variant	OR	95% CI	Frequency (%)	Associated variant	OR	95% CI	Frequency (%)	
CYP1A1	Thr461Asp (C/A)	3	2,245	1,139	CC	CA	1.12	0.87–1.43	8.0	AA	0.95	0.20–4.49	0.2	56
CYP1A2	A(−164)C	9	7,580	10,020	AA	AC	1.02	0.92–1.13	42.8	CC	1.17	0.83–1.64	9.2	57
CYP1B1	G119T	11	10,715	11,678	GG	GT	0.99	0.90–1.10	43.6	TT	0.93	0.73–1.20	9.4	58/59
CYP1B1	Arg48Gly	10	11,321	13,379	AA	AG	0.93	0.81–1.08	42.9	GG	0.82	0.61–1.10	8.4	58/59
CYP1B1	Asn453Ser (A/G)	12	11,630	14,053	AA	AG	0.96	0.91–1.02	29.3	GG	0.98	0.85–1.14	3.5	58/59
CYP1B1	**Val432Leu (G/C)**	13	4,167	3,187	GG	GC	**1.6**	**1.1–2.3**	46.9	CC	1.1	0.8–1.4	31.0	60
ECCR4	rs744154 (G/C)	27	25,743	29,074	GG	GC	1.00	0.97–1.04	39.9	CC	0.97	0.91–1.04	7.5	42
eNOS	**G894T**	11	4,665	4,842	GG	GT	1.00	0.91–1.10	33.1	TT	**1.22**	**1.02–1.44**	6.4	61
eNOS	**T(−786)C**	3	1,856	1,470	CC	CT	0.74	0.51–1.05	27.5	TT	**0.60**	**0.42–0.86**	68.4	62
ERCC2	A751C	32	14,545	15,352	AA	AC + CC	1.13	1.02–1.24	63.4					63/64
ERCC2	G312A	24	16,254	14,006	GG + GA	AA	0.83	0.65–1.05	35.5					63/64
ESR1	G351A	7	3,555	10,924	GG	GA	0.94	0.71–1.25	45.6	AA	1.07	0.90–1.27	16.5	65
ESR1	**rs3020314 (T/C)**	21	25,034	29,460	TT	TC	**1.05**	**1.02–1.09**	43.3	CC	**1.06**	**1.02–1.09**	9.4	66
ESR2	rs3020450 (C/T)	5	5,789	7,761	CC	CT	1.01	0.94–1.08	43.0	TT	0.95	0.85–1.06	11.0	67
ESR2	rs1256031 (A/G)	5	5,789	7,761	AA	AG	0.99	0.92–1.07	49.0	GG	0.93	0.85–1.02	22.0	67
ESR2	rs1256049 (G/A)	8	11,652	15,726	GG	GA	1.04	0.96–1.12	12.6	AA	1.00	0.82–1.22	1.5	68
ESR2	**rs4986938 (A/G)**	9	10,837	16,021	AA	AG	**0.94**	**0.90–1.00**	44.4	GG	0.94	0.87–1.02	13.9	68
FAS	**G(−1377)A**	3	2,396	2,442	GG	GA	**1.18**	**1.04–1.35**	25.2[a]	AA	**1.29**	**1.00–1.67**	2.2[a]	69
FAS	A(−670)G	3	2,386	2,430	AA	AG	1.01	0.89–1.16	49.8[a]	GG	1.00	0.85–1.18	21.9[a]	69
FGFR2	**rs2981582 (C/T)**	11	40,292	51,598	CC	CT	**1.18**	**1.09–1.27**	47.0	TT	**1.48**	**1.35–1.61**	14.4	70/71
GPX1	C198T	6	5,509	6,542	CC + CT	TT	**1.04**	**0.92–1.18**	10.1					72
GSTM1	**Deletions**	41	14,207	15,281	Present	Null	**1.10**	**1.05–1.15**	56.5					73
GSTP1	Ile105Val (A/G)	10	2,163	2,282	AA	AG		Not reported		GG	1.04	0.87–1.25	2.6	74
GSTT1	**Deletions**	15	4,873	5,245	Present	Null	**1.11**	**1.01–1.22**	49.3					74
HER2	**I655V (G/A)**	27	11,504	12,538	GG	GA	**1.05**	**1.00–1.12**	30.3	AA	1.15	0.92–1.43	4.6	75
hOGG1	C326G	11	6,804	6,725	CC	CG	0.99	0.91–1.07	34.1	GG	1.07	0.94–1.20	4.8	76/77
HSD17B1	A312G	9	13,987	17,066	AA	AG	0.97	0.92–1.02	49.5	GG	0.96	0.90–1.03	22.6	78
IGFBP3	**A(−202)C**	27	33,557	45,254	AA	AC	1.03	0.99–1.07	49.7	CC	**1.06**	**1.02–1.11**	24.1	79/80/81
IGF-I	(CA) 19 repeats	7	3,828	8,999	Absent	Present	1.03	0.90–1.17	57.9					82
IL-B1	rs16944 (C/T)	4	1,744	1,342	CC	CT	0.91	0.77–1.07	46.3	TT	0.91	0.60–1.37	19.9	83
LIG4	D501D (T/C)	8	8,933	9,874	TT	CT	0.96	0.90–1.02	28.6	CC	1.02	0.80–1.30	3.3	38

(Continued on the following page)

Table 2. Meta-analyses of genetic susceptibility variants for breast cancer risk (Cont'd)

Gene	Variant	No. of studies	No. of cases	No. of controls	Referent variant	Comparison 1				Comparison 2				Reference[b]
						Associated variant	OR	95% CI	Frequency (%)	Associated variant	OR	95% CI	Frequency (%)	
LSP1	**rs3817198 (T/C)**	20	21,860	22,578	TT	TC	**1.06**	**1.02–1.11**	42.0[a]	CC	**1.17**	**1.08–1.25**	9.0[a]	1
MAP3K1	**rs889312 (A/C)**	20	21,860	22,578	AA	AC	**1.13**	**1.09–1.18**	40.3[a]	CC	**1.27**	**1.19–1.36**	7.8[a]	1
MDM2	T309G	5	5,191	3,834	TT	TG	1.04	0.95–1.13	45.3	GG	0.90	0.80–1.02	15.0	84
MTHFR	A1298C	20	12,170	15,685	A	C	0.99	0.94–1.04	26.6					85
MTHFR	**C677T**	41	16,480	22,388	C	T	**1.04**	**1.01–1.07**	33.6					85
MTRR	A66G	6	6,084	6,756	AA + AG	GG	1.00	0.91–1.09	21.2					86
NAT2	NAT2	20	7,479	8,612	Rapid	Slow	1.02	0.95–1.08	54.3					87
NBS1	**G8360C**	10	4,452	5,665	GG	GC	0.97	0.85–1.11	43.0	CC	**0.75**	**0.74–0.98**	10.5	88
NBS1	I171V (A/G)	3	2,954	2,531	AA	AG	1.05	0.64–1.74	1.0					89
PGR	G331A	10	13,702	14,726	GG	GA	1.06	0.89–1.27	9.4	AA	0.94	0.57–1.56	0.3	90
PGR	rs1042838 (G/T)	24	23,129	27,507	GG	GT	1.02	0.98–1.06	25.8	TT	1.04	0.93–1.17	2.4	42
RAD51	G135C	9	13,241	13,203	GG	CG	0.94	0.88–1.01	14.0	CC	1.02	0.65–1.60	2.5	91
SOD2	V16A	32	26,022	32,426	VV	VA	1.02	0.98–1.06	24.2	AA	1.01	0.93–1.08	1.9	92/93
TGFB1	**L10P**	30	20,401	27,416	LL	LP	**1.05**	**1.00–1.09**	51.7	PP	1.05	0.97–1.13	26.8	79/94
TGFBR1	*6A	3	1,533	2,066	9A/9A	6A/9A	1.20[a]	0.73–1.96	12.5	6A/6A	2.21	0.29–16.90	0.2	95
TNF	rs361525 (G/A)	28	28,333	31,901	GG	GA	0.99	0.94–1.05	9.3	AA	1.14	0.85–1.52	0.3	42
TNRC9	**rs3803662 (C/T)**	20	21,860	22,578	CC	CT	**1.23**	**1.18–1.29**	37.5[a]	TT	**1.39**	**1.26–1.45**	6.3[a]	1
TP53	G72C	17	12,226	10,782	GC + CC	GG	1.13	0.98–1.31	47.4					96
UGT1A1	TA repeat	7	5,746	8,365	7/7	7/6	0.90	0.80–1.01	40.1	6/6	0.90	0.80–1.01	49.9	97
VDR	Apa1	4	1,138	1,198	AA	Aa + aa	1.06	0.73–1.54	69.3					98
VDR	Bsm1	14	5,586	7,943	BB	Bb + bb	0.97	0.83–1.13	83.4					98
VDR	**Fok1**	8	5,284	7,500	FF + Ff	ff	**1.14**	**1.03–1.26**	14.6					98
VDR	Taq1	10	4,459	5,485	TT	Tt + tt	1.02	0.94–1.11	64.5					98
WDR79	**R68G C/G)**	2	2,692	3,367	CC	CG	1.08	0.95–1.23	21.7	GG	**1.60**	**1.04–2.47**	1.2	99
XPC	L939G (A/C)	7	3,073	3,048	A	C	0.96	0.86–1.07	35.4					100
XPC	A499V (C/T)	2	408	426	C	T	1.06	0.80–1.41	33.2					100
XRCC1	R194W (C/T)	21	10,465	10,888	CC	CT	0.96	0.89–1.04	15.7	TT	0.83	0.66–1.04	82.6	101
XRCC1	R280H (G/A)	9	6,165	5,806	GG	GA	1.08	0.95–1.22	1.2	AA	1.64	0.85–3.16	87.8	101

(Continued on the following page)

Table 2. Meta-analyses of genetic susceptibility variants for breast cancer risk (Cont'd)

Gene	Variant	No. of studies	No. of cases	No. of controls	Referent variant	Comparison 1				Comparison 2				Reference[b]
						Associated variant	OR	95% CI	Frequency (%)	Associated variant	OR	95% CI	Frequency (%)	
XRCC1	**R399Q (A/G)**	40	21,467	22,766	GG	GA	1.00	0.96–1.05	44.2	AA	**1.12**	**1.02–1.23**	**43.9**	101
XRCC2	R188H (G/A)	16	18,341	19,028	RR	RH	0.93	0.86–1.02	14.3	HH	0.90	0.69–1.18	0.6	102
XRCC3	5'UTR (A/G)	6	8,343	9,703	AA	AG	1.04	0.95–1.13	31.9	GG	0.95	0.82–1.10	4.7	38
XRCC3	IVS5-14 (A/G)	5	10,537	10,970	AA	AG	1.00	0.88–1.13	43.7	GG	0.90	0.81–1.10	11.2	38
XRCC3	T241M (C/T)	9	12,365	13,138	TT	TC	1.04	0.98–1.14	45.3	CC	1.08	0.98–1.19	13.5	38

NOTE: Printed in bold are the polymorphisms for which at least 1 genetic model showed statistically significant association with breast cancer risk. These polymorphisms were included in the modeling study.

[a]Genotype frequencies are calculated from the allele frequencies assuming Hardy–Weinberg equilibrium.

[b]When 2 references are given, the first is the source for the odds ratios, the second for the genotype frequencies. When 1 reference is given, both were obtained from the same meta-analysis.

[c]Calculated with Review Manager 5.0 using the raw data provided in this article.

Table 3. Minimal odds ratios needed to obtain AUCs of 0.70–0.80 in addition to the 41 statistically significant genetic susceptibility variants (AUC = 0.67)

Risk allele frequency	Number of extra genetic variants	AUC 0.70	AUC 0.75	AUC 0.80
0.05	1	3.2 (2.9–3.6)	10.7 (9.6–11.8)	51.6 (46.8–56.4)
	5	2.0 (1.8–2.1)	3.6 (3.5–3.7)	5.8 (5.7–6.0)
	20	1.5 (1.4–1.6)	2.1 (2.0–2.1)	2.7 (2.7–2.8)
	50	1.3 (1.3–1.4)	1.6 (1.6–1.7)	2.0 (2.0–2.0)
	100	1.3 (1.2–1.3)	1.5 (1.4–1.5)	1.7 (1.7–1.7)
0.30	1	2.0 (1.9–2.2)	4.0 (3.9–4.1)	7.1 (6.9–7.3)
	5	1.5 (1.4–1.5)	2.0 (2.0–2.1)	2.7 (2.7–2.8)
	20	1.3 (1.2–1.3)	1.5 (1.5–1.5)	1.8 (1.7–1.8)
	50	1.2 (1.2–1.2)	1.3 (1.3–1.3)	1.5 (1.5–1.5)
	100	1.1 (1.1–1.2)	1.2 (1.2–1.2)	1.3 (1.3–1.3)
0.50	1	2.0 (1.9–2.1)	4.0 (3.8–4.2)	8.7 (8.4–9.0)
	5	1.4 (1.4–1.5)	2.0 (2.0–2.0)	2.7 (2.7–2.8)
	20	1.2 (1.2–1.3)	1.5 (1.4–1.5)	1.7 (1.7–1.7)
	50	1.2 (1.1–1.2)	1.3 (1.3–1.3)	1.4 (1.4–1.4)
	100	1.1 (1.1–1.1)	1.2 (1.2–1.2)	1.3 (1.3–1.3)

NOTE: Odds ratios are presented as mean (95% CI) of 20 simulations each.
The obtained CIs from the simulation were quite narrow partly due to using a sample size of 10,000 women for this calculation. Therefore, the meaning of these CIs is distinct from the CIs of the AUCs that correspond with existing breast cancer risk prediction models as presented in Table 1, where uncertainty in the CIs is due to variation in the population, sampling, and estimation.

0.67 when only statistically significant polymorphisms were considered. These AUC values are comparable to current breast cancer risk prediction models.

Before further interpreting the public health relevance of our findings, 2 methodologic issues need to be disclosed. First, we assumed that genetic variants inherited independently and that the combined effect of the genetic variants on disease risk followed a multiplicative risk model of independent effects (i.e., no statistical interaction terms were included in the model). Although so far no studies have demonstrated gene–gene interactions with breast cancer risk in general populations, it is still possible that these will be discovered in future studies in larger populations. However, gene–gene interactions only improve the breast cancer risk predictions if their effect sizes are substantially high (e.g., OR > 5). When interaction effects are smaller, their effects on the predictive accuracy will be comparable with that of single gene effects, because by definition their frequencies are lower.

Second, we attempted to compare the AUC of the genetic risk models with observed values for available breast cancer risk prediction models, but this comparison should be made with caution. It is important that risk prediction models should be validated in populations that are representative for the population in which the risk model ultimately is applied (28), and that the models address the same time horizon for the risk prediction. We considered application of the risk model in a general population, but many of the available breast cancer risk models were not evaluated in the general population constructed in a similar population. Some have validated risk models in women who have at least 1 affected relative (Table 1), but not in women who have no positive family history. Finally, we modeled lifetime breast cancer risk, where most published models predicted 5-year risks. In principle, this is a valid comparison, because AUC is independent of disease risk, and we assumed that the effects of the polymorphisms do not vary over time, which so far seems to be a reasonable assumption. The comparison may be weakened if the existing models would have had higher AUC when longer follow-up time had been investigated. This is well plausible, but the effect on AUC is unclear because intermediate risk factors may better predict short-term risk of disease and other risk factors may require long follow-up time to demonstrate their effects. Yet, the breast cancer risk models have AUC values between 0.55 and 0.76, and this range is comparable with risk prediction models based on nongenetic risk factors for other diseases (29).

Of the 96 polymorphisms that have been reviewed in meta-analyses or investigated in GWAS, 41 showed statistically significant associations with breast cancer risk (Table 2). These include the 18 polymorphisms that have been identified in recent GWAS (1, 3, 4, 30–32). The heterozygous OR of the polymorphisms identified in GWAS ranged from 0.84 to 1.36. When future GWAS will identify polymorphisms with per-allele ORs around 1.1, theoretically the predictive ability of the genetic risk model can be improved beyond that of existing models.

Yet, even such small improvements still require the discovery of hundreds of new variants (6), or the discovery of the true causal variants, which may have stronger effects than the current variants included. Another avenue is to improve breast cancer prediction by combining genetic with nongenetic risk factors. Breast cancer prediction may be further improved when nongenetic risk factors are unrelated to the genetic factors, such as age, lifestyle, and dietary factors, and even the presence of affected relatives, which in nonfamilial breast cancer is unlikely explained by the low-risk susceptibility genes (33). It should not be expected that risk prediction is markedly improved when nongenetic risk factors are potential intermediate factors in a biological pathway linking the genetic factors to breast cancer (29), such as benign breast disease, personal history of breast cancer, and hormonal factors. On the basis of current knowledge of breast cancer risk factors, it is likely that risk prediction models solely based on nongenetic factors will perform better than models based on common single-nucleotide polymorphism (SNP) alone, as ORs for SNPs tend to be smaller than ORs based on nongenetic factors. Investigating the combined predictive performance of genetic and nongenetic factors is of interest to investigate whether available prediction models can be further improved. Yet, also existing prediction models will only be markedly improved when a larger number of susceptibility variants can be added (5).

What level of predictive performance is required for practical implementation of the risk model depends on the intended use. When prediction models are used to make decisions at the individual patient level, higher AUCs are required compared to when models are used to implement population-based prevention or therapeutic strategies. Whether the risk predictions are used as a strategy to determine the age of starting or interval of mammography screening, they will require that the risk model can accurately identify women at increased risk among all women regardless of family history of breast cancer. Other interventions such as MRI screening, the use of chemopreventive agents, lifestyle behavioral changes, and surgical interventions could also be considered. Yet, the more detail that is desired for risk stratification, the better the prediction model should perform. Specifying age of entry in the mammography screening program by year, as suggested by Pharoah and colleagues, requires better predictive performance than specifying age of entry in broader age categories, for example, enter screening at ages 45, 50, or 55 years. What level of AUC is required for this application is a question to be investigated in future modeling studies.

In conclusion, our analyses show that prediction of breast cancer risk based on low susceptibility variants theoretically can achieve similar predictive performance to existing breast cancer risk models, and can even improve prediction of disease when more variants are being discovered. Whether this predictive performance is sufficient for implementation of the risk models in mass prevention programs is ultimately determined by the intended use of the test and the performance of the interventions, weighing the benefits, harms, and costs.

Disclosure of Potential Conflicts of Interest

No potential conflicts of interest were disclosed.

Grant Support

This study was financially supported by grants from the Erasmus University Medical Center Rotterdam and by the Center for Medical Systems Biology in the framework of the Netherlands Genomics Initiative (NGI). A.C.J.W. Janssens was additionally sponsored by the Vidi grant of the Netherlands Organisation for Scientific Research (NWO).

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 March 30, 2010; revised August 19, 2010; accepted August 23, 2010; published online January 6, 2011.

References

1. Easton DF, Pooley KA, Dunning AM, Pharoah PD, Thompson D, Ballinger DG, et al. Genome-wide association study identifies novel breast cancer susceptibility loci. Nature 2007;447:1087–93.
2. Hunter DJ, Kraft P, Jacobs KB, Cox DG, Yeager M, Hankinson SE, et al. A genome-wide association study identifies alleles in FGFR2 associated with risk of sporadic postmenopausal breast cancer. Nat Genet 2007;39:870–4.
3. Stacey SN, Manolescu A, Sulem P, Rafnar T, Gudmundsson J, Gudjonsson SA, et al. Common variants on chromosomes 2q35 and 16q12 confer susceptibility to estrogen receptor-positive breast cancer. Nat Genet 2007;39:865–9.
4. Turnbull C, Ahmed S, Morrison J, Pernet D, Renwick A, Maranian M, et al. Genome-wide association study identifies five new breast cancer susceptibility loci. Nat Genet 2010;42:504–7.
5. Gail MH. Discriminatory accuracy from single-nucleotide polymorphisms in models to predict breast cancer risk. J Natl Cancer Inst 2008;100:1037–41.
6. Janssens AC, Aulchenko YS, Elefante S, Borsboom GJ, Steyerberg EW, van Duijn CM. Predictive testing for complex diseases using multiple genes: fact or fiction? Genet Med 2006;8:395–400.
7. Wray NR, Goddard ME, Visscher PM. Prediction of individual genetic risk of complex disease. Curr Opin Genet Dev 2008;18:257–63.
8. Pepe MS, Gu JW, Morris DE. The potential of genes and other markers to inform about risk. Cancer Epidemiol Biomarkers Prev 2010;19:655–65.
9. Salinas CA, Koopmeiners JS, Kwon EM, FitzGerald L, Lin DW, Ostrander EA, et al. Clinical utility of five genetic variants for predicting prostate cancer risk and mortality. Prostate 2009;69:363–72.
10. 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.
11. Wu X, Lin J, Grossman HB, Huang M, Gu J, Etzel CJ, et al. Projecting individualized probabilities of developing bladder cancer in white individuals. J Clin Oncol 2007;25:4974–81.
12. Zheng SL, Sun J, Wiklund F, Smith S, Stattin P, Li G, et al. Cumulative association of five genetic variants with prostate cancer. N Engl J Med 2008;358:910–9.

13. Pharoah PD, Antoniou AC, Easton DF, Ponder BA. Polygenes, risk prediction, and targeted prevention of breast cancer. N Engl J Med 2008;358:2796–803.
14. Gail MH, Mai PL. Comparing breast cancer risk assessment models. J Natl Cancer Inst 2010;102:665–8.
15. Cook NR. Use and misuse of the receiver operating characteristic curve in risk prediction. Circulation 2007;115:928–35.
16. 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;discussion 207–12.
17. Gail MH. Value of adding single-nucleotide polymorphism genotypes to a breast cancer risk model. J Natl Cancer Inst 2009; 101:959–63.
18. Wray NR, Goddard ME, Visscher PM. Prediction of individual genetic risk to disease from genome-wide association studies. Genome Res 2007;17:1520–8.
19. Antoniou AC, Pharoah PD, McMullan G, Day NE, Stratton MR, Peto J, et al. A comprehensive model for familial breast cancer incorporating BRCA1, BRCA2 and other genes. Br J Cancer 2002;86:76–83.
20. Barlow WE, White E, Ballard-Barbash R, Vacek PM, Titus-Ernstoff L, Carney PA, et al. Prospective breast cancer risk prediction model for women undergoing screening mammography. J Natl Cancer Inst 2006;98:1204–14.
21. Claus EB, Risch N, Thompson WD. Autosomal dominant inheritance of early onset breast cancer. Implications for risk prediction. Cancer 1994;73:643–51.
22. Costantino JP, Gail MH, Pee D, Anderson S, Redmond CK, Benichou J, et al. Validation studies for models projecting the risk of invasive and total breast cancer incidence. J Natl Cancer Inst 1999;91:1541–8.
23. Gail MH, Costantino JP, Pee D, Bondy M, Newman L, Selvan M, et al. Projecting individualized absolute invasive breast cancer risk in African American women. J Natl Cancer Inst 2007;99:1782–92.
24. Tyrer J, Duffy SW, Cuzick J. A breast cancer prediction model incorporating familial and personal risk factors. Stat Med 2004;23: 1111–30.
25. Sackett DL, Tugwell P. Clinical epidemiology: a basic science for clinical medicine. Boston/Toronto: Little, Brown and Company; 1985.
26. Hanley JA, McNeil BJ. The meaning and use of the area under a receiver operating characteristic (ROC) curve. Radiology 1982;143: 29–36.
27. Ihaka GR. R: A language for data analysis and graphics. J Comput Graph Stat 1996;5:299–314.
28. Freedman AN, Seminara D, Gail MH, Hartge P, Colditz GA, Ballard-Barbash R, et al. Cancer risk prediction models: a workshop on development, evaluation, and application. J Natl Cancer Inst 2005;97:715–23.
29. Janssens AC, van Duijn CM. Genome-based prediction of common diseases: advances and prospects. Hum Mol Genet 2006;17:166–73.
30. Ahmed S, Thomas G, Ghoussaini M, Healey CS, Humphreys MK, Platte R, et al. Newly discovered breast cancer susceptibility loci on 3p24 and 17q23.2. Nat Genet 2009; 41:585–90.
31. Thomas G, Jacobs KB, Kraft P, Yeager M, Wacholder S, Cox DG, et al. A multistage genome-wide association study in breast cancer identifies two new risk alleles at 1p11.2 and 14q24.1 (RAD51L1). Nat Genet 2009;41:579–84.
32. Zheng W, Long J, Gao Y, Li C, Zheng Y, Xiang YB, et al. Genome-wide association study identifies a new breast cancer susceptibility locus at 6q25.1. Nat Genet 2009;41:324–8.
33. Janssens AC, Khoury MJ. Predictive value of testing for multiple genetic variants in multifactorial diseases: implications for the discourse on ethical, legal and social issues. Ital J Public Health 2006;3:35–41.
34. Amir E, Evans DG, Shenton A, Lalloo F, Moran A, Boggis C, et al. Evaluation of breast cancer risk assessment packages in the family history evaluation and screening programme. J Med Genet 2003;40:807–14.
35. 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.
36. Decarli A, Calza S, Masala G, Specchia C, Palli D, Gail MH. Gail model for prediction of absolute risk of invasive breast cancer: independent evaluation in the Florence-European Prospective Investigation Into Cancer and Nutrition cohort. J Natl Cancer Inst 2006;98:1686–93.
37. 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.
38. Commonly studied single-nucleotide polymorphisms and breast cancer: results from the Breast Cancer Association Consortium. J Natl Cancer Inst 2006;98:1382–96.
39. Frank B, Wiestler M, Kropp S, Hemminki K, Spurdle AB, Sutter C, et al. Association of a common AKAP9 variant with breast cancer risk: a collaborative analysis. J Natl Cancer Inst 2008;100:437–42.
40. Hao Y, Montiel R, Li B, Huang E, Zeng L, Huang Y. Association between androgen receptor gene CAG repeat polymorphism and breast cancer risk: a meta-analysis. Breast Cancer Res Treat 2010. [Epub Apr 30, 2010].
41. Sun H, Bai J, Chen F, Jin Y, Yu Y, Fu S. Lack of an association between AURKA T91A polymorphisms and breast cancer: a meta-analysis involving 32,141 subjects. Breast Cancer Res Treat 2010. [Epub May 13, 2010].
42. Gaudet MM, Milne RL, Cox A, Camp NJ, Goode EL, Humphreys MK, et al. Five polymorphisms and breast cancer risk: results from the Breast Cancer Association Consortium. Cancer Epidemiol Biomarkers Prev 2009;18:1610–6.
43. Qiu LX, Yao L, Xue K, Zhang J, Mao C, Chen B, et al. BRCA2 N372H polymorphism and breast cancer susceptibility: a meta-analysis involving 44,903 subjects. Breast Cancer Res Treat 2010. [Epub Feb 5, 2010].
44. Garcia-Closas M, Egan KM, Newcomb PA, Brinton LA, Titus-Ernstoff L, Chanock S, et al. Polymorphisms in DNA double-strand break repair genes and risk of breast cancer: two population-based studies in USA and Poland, and meta-analyses. Hum Genet 2006;119:376–88.
45. Janssens AC, Gonzalez-Zuloeta Ladd AM, Lopez-Leon S, Ioannidis JP, Oostra BA, Khoury MJ, et al. An empirical comparison of meta-analyses of published gene-disease associations versus consortium analyses. Genet Med 2009;11:153–62.
46. Lu C, Dong J, Ma H, Jin G, Hu Z, Peng Y, et al. CCND1 G870A polymorphism contributes to breast cancer susceptibility: a meta-analysis. Breast Cancer Res Treat 2009;116:571–5.
47. Qiu LX, Zhang J, Zhu XD, Zheng CL, Sun S, Wang ZH, et al. The p21 Ser31Arg polymorphism and breast cancer risk: a meta-analysis involving 51,236 subjects. Breast Cancer Res Treat 2010. [Epub Mar 27, 2010].
48. MARIE-GENICA Consortium on Genetic Susceptibility for Menopausal Hormone Therapy Related Breast Cancer Risk. Polymorphisms in the BRCA1 and ABCB1 genes modulate menopausal hormone therapy associated breast cancer risk in postmenopausal women. Breast Cancer Res Treat 2010;120:727–36.
49. Weischer M, Bojesen SE, Ellervik C, Tybjaerg-Hansen A, Nordestgaard BG. CHEK2*1100delC genotyping for clinical assessment of breast cancer risk: meta-analyses of 26,000 patient cases and 27,000 controls. J Clin Oncol 2008;26:542–8.
50. Mao C, Wang XW, Qiu LX, Liao RY, Ding H, Chen Q. Lack of association between catechol-O-methyltransferase Val108/158Met polymorphism and breast cancer risk: a meta-analysis of 25,627 cases and 34,222 controls. Breast Cancer Res Treat 2010;121:719–25.
51. Ye Z, Parry JM. The CYP17 MspA1 polymorphism and breast cancer risk: a meta-analysis. Mutagenesis 2002;17:119–26.
52. Feigelson HS, McKean-Cowdin R, Coetzee G, Stram DO, Kolonel LN, Henderson BE. Building a multigenic model of breast cancer susceptibility: CYP17 and HSD17B1 are two important candidates. Cancer Res 2001;61:785–9.

53. Ma X, Qi X, Chen C, Lin H, Xiong H, Li Y, et al. Association between CYP19 polymorphisms and breast cancer risk: results from 10,592 cases and 11,720 controls. Breast Cancer Res Treat 2010; 122:495–501.
54. Chen C, Huang Y, Li Y, Mao Y, Xie Y. Cytochrome P450 1A1 (CYP1A1) T3801C and A2455G polymorphisms in breast cancer risk: a meta-analysis. J Hum Genet 2007;52:423–35.
55. Yao L, Yu X, Yu L. Lack of significant association between CYP1A1 T3801C polymorphism and breast cancer risk: a meta-analysis involving 25,087 subjects. Breast Cancer Res Treat 2010;122; 503–7.
56. Masson LF, Sharp L, Cotton SC, Little J. Cytochrome P-450 1A1 gene polymorphisms and risk of breast cancer: a HuGE review. Am J Epidemiol 2005;161:901–15.
57. Qiu LX, Yao L, Mao C, Yu KD, Zhan P, Chen B, et al. Lack of association of CYP1A2–164 A/C polymorphism with breast cancer susceptibility: a meta-analysis involving 17,600 subjects. Breast Cancer Res Treat 2010;122: 521–5.
58. Economopoulos KP, Sergentanis TN. Three polymorphisms in cytochrome P450 1B1 (CYP1B1) gene and breast cancer risk: a meta-analysis. Breast Cancer Res Treat 2010;122:545–51.
59. MARIE-GENICA Consortium on Genetic Susceptibility for Menopausal Hormone Therapy Related Breast Cancer Risk. Genetic polymorphisms in phase I and phase II enzymes and breast cancer risk associated with menopausal hormone therapy in postmenopausal women. Breast Cancer Res Treat 2010;119:463–74.
60. Paracchini V, Raimondi S, Gram IT, Kang D, Kocabas NA, Kristensen VN, et al. Meta- and pooled analyses of the cytochrome P-450 1B1 Val432Leu polymorphism and breast cancer: a HuGE-GSEC review. Am J Epidemiol 2007;165:115–25.
61. Hao Y, Montiel R, Huang Y. Endothelial nitric oxide synthase (eNOS) 894 G>T polymorphism is associated with breast cancer risk: a meta-analysis. Breast Cancer Res Treat 2010. [Epub Mar 19, 2010].
62. Yao L, Fang F, Zhong Y, Yu L. The association between two polymorphisms of eNOS and breast cancer risk: a meta-analysis. Breast Cancer Res Treat 2010. [Epub Mar 5, 2010].
63. Pabalan N, Francisco-Pabalan O, Sung L, Jarjanazi H, Ozcelik H. Meta-analysis of two ERCC2 (XPD) polymorphisms, Asp312Asn and Lys751Gln, in breast cancer. Breast Cancer Res Treat 2010. [Epub Apr 9, 2010].
64. Debniak T, Scott RJ, Huzarski T, Byrski T, Masojc B, van de Wetering T, et al. XPD common variants and their association with melanoma and breast cancer risk. Breast Cancer Res Treat 2006;98:209–15.
65. Gonzalez-Zuloeta Ladd AM, Vasquez AA, Rivadeneira F, et al. Estrogen receptor alpha polymorphisms and postmenopausal breast cancer risk. Breast Cancer Res Treat 2008;107:415–9.
66. Dunning AM, Healey CS, Baynes C, et al. Association of ESR1 gene tagging SNPs with breast cancer risk. Hum Mol Genet 2009;18: 1131–9.
67. Breast and Prostate Cancer Cohort, Cox DG, Bretsky P, Kraft P, Pharoah P, Albanes D, Altshuler D, et al. Haplotypes of the estrogen receptor beta gene and breast cancer risk. Int J Cancer 2008;122:387–92.
68. Yu KD, Rao NY, Chen AX, Fan L, Yang C, Shao ZM. A systematic review of the relationship between polymorphic sites in the estrogen receptor-beta (ESR2) gene and breast cancer risk. Breast Cancer Res Treat 2010. [Epub Apr 14, 2010].
69. Zhang Z, Xue H, Gong W, Wang M, Yuan L, Han S. FAS promoter polymorphisms and cancer risk: a meta-analysis based on 34 case-control studies. Carcinogenesis 2009;30:487–93.
70. Jia C, Cai Y, Ma Y, Fu D. Quantitative assessment of the effect of FGFR2 gene polymorphism on the risk of breast cancer. Breast Cancer Res Treat 2010. [Epub Apr 4, 2010].
71. Garcia-Closas M, Hall P, Nevanlinna H, Pooley K, Morrison J, Richesson DA, et al. Heterogeneity of breast cancer associations with five susceptibility loci by clinical and pathological characteristics. PLoS Genet 2008;4:e1000054.
72. Hu J, Zhou GW, Wang N, Wang YJ. GPX1 Pro198Leu polymorphism and breast cancer risk: a meta-analysis. Breast Cancer Res Treat 2010. [Epub Mar 21, 2010].
73. Yu KD, Di GH, Fan L, Wu J, Hu Z, Shen ZZ, et al. A functional polymorphism in the promoter region of GSTM1 implies a complex role for GSTM1 in breast cancer. FASEB J 2009;23:2274–87.
74. Egan KM, Cai Q, Shu XO, Jin F, Zhu TL, Dai Q, et al. Genetic polymorphisms in GSTM1, GSTP1, and GSTT1 and the risk for breast cancer: results from the Shanghai Breast Cancer Study and meta-analysis. Cancer Epidemiol Biomarkers Prev 2004;13:197–204.
75. Lu S, Wang Z, Liu H, Hao X. HER2 Ile655Val polymorphism contributes to breast cancer risk: evidence from 27 case-control studies. Breast Cancer Res Treat 2010. [Epub Apr 17, 2010].
76. Gu D, Wang M, Zhang Z, Chen J. Lack of association between the hOGG1 Ser326Cys polymorphism and breast cancer risk: evidence from 11 case-control studies. Breast Cancer Res Treat 2010; 122:527–31.
77. Zhang Y, Newcomb PA, Egan KM, Titus-Ernstoff L, Chanock S, Welch R, et al. Genetic polymorphisms in base-excision repair pathway genes and risk of breast cancer. Cancer Epidemiol Biomarkers Prev 2006;15:353–8.
78. Yao L, Cao LH, Qiu LX, Yu L. The association between HSD17B1 Ser312Gly polymorphism and breast cancer risk: a meta-analysis including 31,053 subjects. Breast Cancer Res Treat 2010. [Epub Feb 12, 2010].
79. Qiu LX, Yao L, Mao C, Chen B, Zhan P, Xue K, et al. TGFB1 L10P polymorphism is associated with breast cancer susceptibility: evidence from a meta-analysis involving 47,817 subjects. Breast Cancer Res Treat 2010. [Epub Feb 9, 2010].
80. Patel AV, Cheng I, Canzian F, Le Marchand L, Thun MJ, Berg CD, et al. IGF-1, IGFBP-1, and IGFBP-3 polymorphisms predict circulating IGF levels but not breast cancer risk: findings from the Breast and Prostate Cancer Cohort Consortium (BPC3). PLoS One 2008;3: e2578.
81. Qiu LX, Yao L, Yuan H, Mao C, Chen B, Zhan P, et al. IGFBP3 A-202C polymorphism and breast cancer susceptibility: a meta-analysis involving 33,557 cases and 45,254 controls. Breast Cancer Res Treat 2010 Aug;122:867–71. [Epub Jan 19, 2010].
82. Chen X, Guan J, Song Y, Chen P, Zheng H, Tang C, et al. IGF-I (CA) repeat polymorphisms and risk of cancer: a meta-analysis. J Hum Genet 2008;53:227–38.
83. Liu X, Wang Z, Yu J, Lei G, Wang S. Three polymorphisms in interleukin-1beta gene and risk for breast cancer: a meta-analysis. Breast Cancer Res Treat 2010. [Epub May 1, 2010].
84. Schmidt MK, Reincke S, Broeks A, Braaf LM, Hogervorst FB, Tollenaar RA, et al. Do MDM2 SNP309 and TP53 R72P interact in breast cancer susceptibility? A large pooled series from the breast cancer association consortium. Cancer Res 2007; 67:9584–90.
85. Qi X, Ma X, Yang X, Fan L, Zhang Y, Zhang F, et al. Methylenetetrahydrofolate reductase polymorphisms and breast cancer risk: a meta-analysis from 41 studies with 16,480 cases and 22,388 controls. Breast Cancer Res Treat 2010. [Epub Feb 5, 2010].
86. Hu J, Zhou GW, Wang N, Wang YJ. MTRR A66G polymorphism and breast cancer risk: a meta-analysis. Breast Cancer Res Treat 2010. [Epub Apr 22, 2010].
87. Ochs-Balcom HM, Wiesner G, Elston RC. A meta-analysis of the association of N-acetyltransferase 2 gene (NAT2) variants with breast cancer. Am J Epidemiol 2007;166:246–54.
88. Wang Z, Cui D, Lu W. NBS1 8360G > C polymorphism is associated with breast cancer risk: a meta-analysis. Breast Cancer Res Treat 2010. [Epub Feb 9, 2010].
89. Bogdanova N, Schurmann P, Waltes R, Feshchenko S, Zalutsky IV, Bremer M, et al. NBS1 variant I171V and breast cancer risk. Breast Cancer Res Treat 2008;112:75–9.
90. Yu KD, Chen AX, Shao ZM. No association between a progesterone receptor gene promoter polymorphism (+331G>A) and breast cancer risk in Caucasian women: evidence from a literature-based meta-analysis. Breast Cancer Res Treat 2010 Aug;122:853–8. [Epub Jan 14, 2010].
91. Wang Z, Dong H, Fu Y, Ding H. RAD51 135G>C polymorphism contributes to breast cancer susceptibility: a meta-analysis

involving 26,444 subjects. Breast Cancer Res Treat 2010. [Epub Apr 16, 2010].
92. Qiu LX, Yao L, Mao C, Chen B, Zhan P, Yuan H, et al. Lack of association between MnSOD Val16Ala polymorphism and breast cancer risk: a meta-analysis involving 58,448 subjects. Breast Cancer Res Treat 2010. [Epub Feb 9, 2010].
93. Cai Q, Shu XO, Wen W, Cheng JR, Dai Q, Gao YT, et al. Genetic polymorphism in the manganese superoxide dismutase gene, antioxidant intake, and breast cancer risk: results from the Shanghai Breast Cancer Study. Breast Cancer Res 2004;6:R647–55.
94. Shin A, Shu XO, Cai Q, Gao YT, Zheng W. Genetic polymorphisms of the transforming growth factor-beta1 gene and breast cancer risk: a possible dual role at different cancer stages. Cancer Epidemiol Biomarkers Prev 2005;14:1567–70.
95. Kaklamani VG, Hou N, Bian Y, Reich J, Offit K, Michel LS, et al. TGFBR1*6A and cancer risk: a meta-analysis of seven case-control studies. J Clin Oncol 2003;21: 3236–43.
96. Zhuo W, Zhang Y, Xiang Z, Cai L, Chen Z. Polymorphisms of TP53 codon 72 with breast carcinoma risk: evidence from 12226 cases and 10782 controls. J Exp Clin Cancer Res 2009;28:115.
97. Yao L, Qiu LX, Yu L, Yang Z, Yu XJ, Zhong Y. The association between TA-repeat polymorphism in the promoter region of UGT1A1 and breast cancer risk: a meta-analysis. Breast Cancer Res Treat 2010 Aug;122:879–82. [Epub Jan 20, 2010].
98. Tang C, Chen N, Wu M, Yuan H, Du Y. Fok1 polymorphism of vitamin D receptor gene contributes to breast cancer susceptibility: a meta-analysis. Breast Cancer Res Treat 2009;117:391–9.
99. Garcia-Closas M, Kristensen V, Langerod A, Qi Y, Yeager M, Burdett L, et al. Common genetic variation in TP53 and its flanking genes, WDR79 and ATP1B2, and susceptibility to breast cancer. Int J Cancer 2007;121:2532–8.
100. Francisco G, Menezes PR, Eluf-Neto J, Chammas R. XPC polymorphisms play a role in tissue-specific carcinogenesis: a meta-analysis. Eur J Hum Genet 2008;16:724–34.
101. Li H, Ha TC, Tai BC. XRCC1 gene polymorphisms and breast cancer risk in different populations: a meta-analysis. Breast 2009;18:183–91.
102. Yu KD, Chen AX, Qiu LX, Fan L, Yang C, Shao ZM. XRCC2 Arg188His polymorphism is not directly associated with breast cancer risk: evidence from 37,369 subjects. Breast Cancer Res Treat 2010 Aug;123:219–25. [Epub Feb 2, 2010].