ORIGINAL REPORT

Impact of a Panel of 88 Single Nucleotide Polymorphisms on the Risk of Breast Cancer in High-Risk Women: Results From Two Randomized Tamoxifen Prevention Trials

Jack Cuzick, Adam R. Brentnall, Corrinne Segal, Helen Byers, Caroline Reuter, Simone Detre, Elena Lopez-Knowles, Ivana Sestak, Anthony Howell, Trevor J. Powles, William G. Newman, and Mitchell Dowsett

Author affiliations and support information (if applicable) appear at the end of this article.

Published at jco.org on December 27, 2016.

Clinical trial information: ISRCTN91879928 and ISRCTN07027313.

Corresponding author: Jack Cuzick, Queen Mary University of London, Wolfson Institute of Preventive Medicine, Charterhouse Square, London, EC1M 6BQ, United Kingdom; e-mail: j.cuzick@qmul.ac.uk.

A B S T R A C T

Purpose
At least 94 common single nucleotide polymorphisms (SNPs) are associated with breast cancer. The extent to which an SNP panel can refine risk in women who receive preventive therapy has not been directly assessed previously.

Materials and Methods
A risk score on the basis of 88 SNPs (SNP88) was investigated in a nested case-control study of women enrolled in the International Breast Intervention Study (IBIS-I) or the Royal Marsden study. A total of 359 women who developed cancer were matched to 636 controls by age, trial, follow-up time, and treatment arm. Genotyping was done using the OncoArray. Conditional logistic regression and matched concordance indices (mC) were used to measure the performance of SNP88 alone and with other breast cancer risk factors assessed using the Tyrer-Cuzick (TC) model.

Results
SNP88 was predictive of breast cancer risk overall (interquartile range odds ratio [IQ-OR], 1.37; 95% CI, 1.14 to 1.66; mC, 0.55), but mainly for estrogen receptor–positive disease (IQ-OR, 1.44; 95% CI, 1.16 to 1.79; P for heterogeneity = .10) versus estrogen receptor–negative disease. However, the observed risk of SNP88 was only 46% (95% CI, 19% to 74%) of expected. No significant interaction was observed with treatment arm (placebo IQ-OR, 1.46; 95% CI, 1.13 to 1.87; tamoxifen IQ-OR, 1.25; 95% CI, 0.96 to 1.64; P for heterogeneity = .5). The predictive power was similar to the TC model (IQ-OR, 1.45; 95% CI, 1.21 to 1.73; mC, 0.55), but SNP88 was independent of TC (Spearman rank-order correlation, 0.012; P = .7), and when combined multiplicatively, a substantial improvement was seen (IQ-OR, 1.64; 95% CI, 1.36 to 1.97; mC, 0.60).

Conclusion
A polygenic risk score may be used to refine risk from the TC or similar models in women who are at an elevated risk of breast cancer and considering preventive therapy. Recalibration may be necessary for accurate risk assessment.

J Clin Oncol 35:743-750. © 2016 by American Society of Clinical Oncology. Licensed under the Creative Commons Attribution 4.0 License: http://creativecommons.org/licenses/by/4.0/

INTRODUCTION

Breast cancer is associated with mutations in rare but highly penetrant dominant genes such as *BRCA1* and *BRCA2* and, to a lesser extent, genes such as *CHEK2* and *PALB2*.[1] Although important for families with a history of the disease, they are too rare to be of much use for risk stratification in the general population. At least 94 common breast risk single nucleotide polymorphisms (SNPs) have been shown to be associated with breast cancer.[2-18] It is anticipated that many more

will be found in future,[1] but they are expected to be individually much less predictive. Each SNP found to date only confers a small relative risk, but together, they can usefully stratify risk in the general population.[19]

Accurate risk assessment is important for many decisions about preventive interventions, such as increased screening and preventive therapy.[20] Risk evaluators available include the Breast Cancer Risk Assessment Tool (Gail model), Tyrer-Cuzick (TC), Breast Cancer Surveillance Consortium, and the Breast and Ovarian Analysis of Disease Incidence and Carrier Estimation

Algorithm models.[21-24] SNP panels are most likely to be of use in clinical practice if they can be combined with such models; therefore, it is important to measure how much additional information they add.

Previous work in this area has been done using case-control studies of mainly postmenopausal women. Wacholder et al[25] showed that fitting a SNP10 risk score in addition to Gail model factors yielded a modest improvement. Mealiffe et al[26] found a similar result in in the Women's Health Initiative trial using a predefined SNP7 panel, and Lee et al[27] reported that a SNP51 panel added to Gail model risk factors in Asian women. Vachon et al[28] showed that a predefined SNP76 panel added information to the Breast Cancer Surveillance Consortium model in three case-control studies. Mavaddat et al[29] fitted a SNP77 panel in a large case-control analysis of women from the general population. Dite et al[30] reported that a predefined SNP77 panel was informative for women at increased risk for breast cancer as a result of classic risk factors from the Australian Breast Cancer Family Registry.

Our focus is the extent to which an SNP panel can refine risk in women already at an increased risk for breast cancer as a result of classic factors and who may receive preventive therapy and to assess whether the preventive effect of tamoxifen varied according to SNP-derived estimates of risk. A previous study of the first of these issues has been conducted by Vachon et al.[31] They used a nested case-control study within the National Surgical Adjuvant Breast and Bowel Project P-1 and P-2 prevention trials, on the basis of women receiving tamoxifen or raloxifene (no placebo), and found that a 75-SNP panel had good univariable power for predicting subsequent risk and added accuracy to the Gail model. However, they did not compare directly with women who did not receive treatment and therefore were unable to directly test for a difference between women receiving treatment or not, nor could they estimate the relative information from the SNP panel compared with classic factors.

A consortium has designed a custom Illumina (San Diego, CA) array (OncoArray), with more than half a million SNPs, including most known breast-cancer risk SNPs. In this article, we present the results from an SNP score on the basis of this assay, using a nested case-control study of women recruited into the International Breast Intervention Study (IBIS-I)[32] and Royal Marsden[33] randomized tamoxifen prevention trials. The SNPs used were those for which the breast cancer risk has been previously validated and are directly available or have close surrogates on the OncoArray. The main objectives were to evaluate the performance of the SNP panel using a population at increased risk as a result of classic risk factors (including a family history or prior proliferative benign tissue diagnosis); compare and combine this with risk estimates from the TC model,[22] which uses classic phenotypic factors; and to conduct subgroup analyses to assess the performance by treatment allocation (tamoxifen or placebo) and estrogen-receptor (ER) status of the tumors that subsequently occurred.

MATERIALS AND METHODS

Patients

Women were recruited into the IBIS-I and Royal Marsden (Marsden) trials, as previously described[32,33] (Data Supplement). Both trials were double blind with women randomly assigned to receive tamoxifen (20 mg/day) or placebo for 5 years (IBIS-I) or 5 to 8 years (Marsden). Subsequent

cancers were ascertained by clinic visits during the treatment period and thereafter by questionnaire, cancer registry follow-up, and/or clinic visits. All women provided informed consent, including for use of their samples in future research. The trials are registered at www.controlled-trials.com as ISRCTN91879928 (IBIS-I) and ISRCTN07027313 (Marsden).

Specimen Characteristics

Blood samples were taken at baseline from all women in IBIS-I and were stored at −70°C. The baseline blood samples for Marsden were destroyed by a fire, and subsequent blood samples were collected where possible and stored at −70°C. However, they were not available for 38 cases, where tissue samples from formalin-fixed paraffin-embedded blocks were used. They were included only in a sensitivity analysis.

Assay Methods

We used 10 μL of extracted DNA from IBIS-I (concentration 100 ng/μL) and Marsden (50 ng/μL) samples on thirteen 96-well plates, including a separate plate for the formalin-fixed paraffin-embedded block samples. Two samples with known genotypes were used as internal controls for each plate. Assays were carried out blindly at Genome Quebec (Montreal, Canada), which was clinical service provider–certified by Illumina; the Illumina OncoArray was used, and the Illumina HTS (high-throughput sequencing) protocol was rigorously followed.

Study Design

The primary end point was diagnosis of invasive breast cancer or ductal carcinoma in situ. The study was designed to use all cases from the trials with material available for genotyping and had more than 99% power to detect the expected polygenic score at a 5% level.[34] The Marsden trial was a pilot for the IBIS-I trial, and the cohorts had similar characteristics. They were combined to increase precision and to provide greater power for subgroup analyses, as in previous analyses.[35] Case-control matching was by study, age at baseline (± 2 years), treatment arm, and follow-up time, with two controls per case in IBIS-I and one in Marsden. IBIS-I recruited from 1992 to 2001, and Marsden recruited from 1986 to 1996. The end of follow-up for the current analysis was 2014 and 2010, respectively; median follow-up was 16.5 years for IBIS-I and 18.4 years for Marsden.

A polygenic score was used to provide an overall risk estimate.[26] Three SNPs were excluded because they had no surrogate SNP on the OncoArray ($r^2 <$ .8 from the 1000 Genomes Project Central European panel[36]) and three SNPs failed quality control. We calculated the normalized odds ratio for each of the three SNP genotypes (no risk alleles, 1 risk allele, and 2 risk alleles) from published per-allele odds ratios, assuming independence and normalizing by an assumed risk allele frequency so that the average risk was unity.[16-18] An overall SNP risk score for each woman (SNP88) was formed by multiplying the odds ratios for each of her 88 genotypes together (Data Supplement). The odds ratios for the TC risk[22] (v7.02) at entry to each study were obtained by dividing the 10-year predicted odds for a person by that from the average population of a woman of the same age (Data Supplement).

Analysis Methods

SNPs that failed in more than 2% of the samples were excluded, as well as samples that failed for more than 10 SNPs. Hardy-Weinberg equilibrium for each SNP in cases and controls was tested by assessing the observed number of homozygotes against expected using a binomial distribution. Sporadic assay failures in SNP88 were assigned the population risk of 1.0. Conditional logistic regression was used for the logarithm of the SNP88 score and TC risk, alone and in combination, assuming independence. A bivariable model was then used to assess the added value of the SNP88 score to the TC risk model. Interaction tests were used to examine subgroups. SNP88 and TC calibration was assessed primarily by regressing observed risks on expected risks. Calibration of SNP88 was further explored by deciles of the SNP88 in controls and for TC by the ratio of observed to expected cancers diagnosed in

the entire placebo arm of the IBIS-I trial, overall and by decile (TC in the complete Marsden cohort was unavailable). Concordance indices for matched sets were used as a secondary measure of discrimination.[34] The distribution of absolute 10-year risk was estimated by weighting cases and controls to reflect a 6% 10-year risk for the whole trial.[32,33] An 8% 10-year risk cut point was chosen because this is the threshold recommended in the United Kingdom for offering preventive therapy.[20] The reclassification of cases and controls about this cut point from TC was assessed using an SNP score that was recalibrated to the data. A sensitivity analysis was undertaken to impute risk from failed assays in the 88 SNPs by applying the Beagle algorithm (Data Supplement).[37] A sensitivity analysis considered when tissue samples that did not fail more than 10 SNPs were included.

RESULTS

A CONSORT diagram of the source of patients is shown in Figure 1. A total of 1,276 women (483 cases, 793 controls) from both trials were initially selected for the case-control study. However, insufficient DNA was available in 169 samples, and the initial design was adjusted to ensure that all cases had matched controls. Twenty-five samples (2.3%) with more than 10 failures were excluded (Data Supplement), in addition to all tissue sample results, because tissue samples were more likely to have more than 10 SNP failures. This led to a primary analysis sample of 995 women (359 cases, 636 controls).

The sample characteristics of participants are listed in Table 1. Briefly, 41% were randomly assigned to tamoxifen and 59% to placebo; the imbalance reflects the lower proportion of cancers in women treated with tamoxifen. The majority of cancers were ER-positive (74%). At trial entry the median age was 50 years, 46% of participants were postmenopausal, and the median body mass index was 25.5 kg/m². The characteristics in the two trials were broadly similar (Data Supplement).
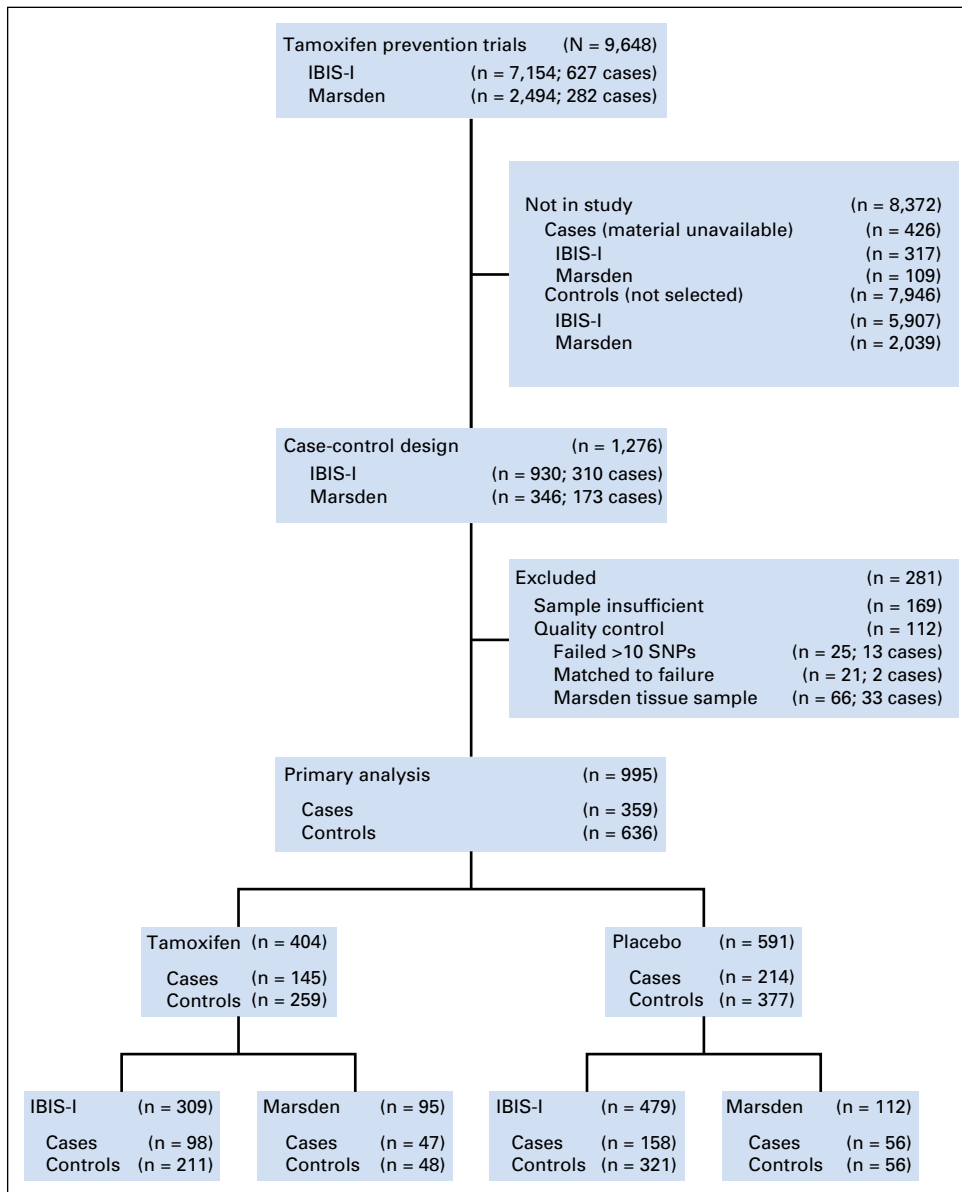

Fig 1. CONSORT diagram. IBIS-I, International Breast Intervention Study; Marsden, Royal Marsden study; SNPs, single nucleotide polymorphisms.

Table 1. Sample Characteristics

Characteristic	Case	Control	P*
Total, No.	359	636	
Treatment, No.			
Placebo	214	377	
Tamoxifen	145	259	
ER status, No.			
Positive	265	466	
Negative	64	117	
Unknown	30	53	
Median age, years (IQR)	50 (45-54)	49 (46-54)	.9
Age at menarche, years, No. (%)			
< 12	81 (23)	152 (24)	.9
12	67 (19)	105 (17)	
> 1	206 (58)	374 (59)	
Unknown	5	5	
Parity			
Nulliparous, No. (%)	61 (17)	85 (13)	.19
Age at first birth, years (IQR)	24.0 (21.0-27.0)	24.0 (21.0-27.0)	.7
Unknown	6	10	
Menopausal status, No. (%)			
Premenopausal	188 (52)	305 (48)	.059
Perimenopausal	18 (5)	27 (4)	
Postmenopausal	153 (43)	304 (48)	
BMI, kg/m^2			
Median (IQR)	25.7 (22.6-28.7)	25.4 (23.0-29.4)	.5
Unknown, No.	11	8	
No. of first-degree relatives affected (%)			
0	39 (11)	55 (9)	.40
1	261 (73)	482 (76)	
> 2	59 (16)	99 (16)	
Median 10-year TC OR (IQR)	2.46 (1.91-3.21)	2.20 (1.78-2.81)	< .001
Median SNP88 score (IQR)	1.13 (0.81-1.62)	1.00 (0.72-1.40)	< .001

Abbreviations: BMI, body mass index; ER, estrogen receptor; IQR, interquartile range; OR, odds ratio; SNP, single nucleotide polymorphism; TC, Tyrer-Cuzick model, compared with general population.
*P from conditional logistic regression.

Quality control of individual SNPs was assessed before calculating an SNP score. One SNP failed all samples (rs6678914), but the percentage of failed samples was less than 2% for all but two of the others, rs4808801 and rs2236007 (Data Supplement), which were excluded, leaving 88 SNPs for our primary analysis. The distribution of homozygotes satisfied the Hardy-Weinberg equilibrium in controls ($P > .067$ for each SNP; Data Supplement), and overall, the observed and expected number of homozygotes in controls were 36,177 and 36,176, respectively ($P = 1.0$). For cases, it was 20,349 and 20,443, respectively ($P = .28$).

As expected, the SNPs were mostly uncorrelated. However, three pairs had $r^2 > .1$: rs554219 and rs7591516 (0.36); rs2363956 and rs8170 (0.25); and rs12662670 and rs2046210 (0.15). These are in line with expectations from the 1000 Genomes Project Central European panel. We made no adjustment for these small correlations in the SNP score.

SNP88 was predictive in the overall sample (interquartile range odds ratio [IQ-OR], 1.37; 95% CI, 1.14 to 1.66; matched concordance index, 0.55; $P < .001$), mainly for ER-positive disease (IQ-OR, 1.44; 95% CI, 1.16 to 1.79) versus ER-negative disease (IQ-OR, 0.99; 95% CI, 0.61 to 1.61; $P = .10$). No significant interaction was observed by treatment arm, but there were slightly stronger effects in untreated women (IQ-OR, 1.46; 95% CI, 1.13 to 1.87) compared with those receiving tamoxifen (IQ-OR, 1.25; 95% CI, 0.96 to 1.64; $P = .5$). The SNP score was nonsignificantly more

predictive in IBIS-I than Marsden ($P = .21$; Table 2; Data Supplement), and so was TC ($P = .6$).

There was almost no correlation between the TC model and the SNP score (Spearman coefficient, 0.012; $P = .7$; Fig 2). SNP88 added significant independent information to TC ($P < .001$), and when combined multiplicatively, a substantial increase was seen in the predictive power (Table 2). The matched concordance index increased from 0.55 to 0.60, and the fit ($\chi_1^2 = 28.6$) was close to that for a full bivariable model ($\chi_2^2 = 30.0$). These results suggest that TC and SNP88 are largely independent and may be combined multiplicatively.

Calibration of the SNP score was assessed by comparing the model calculated odds ratios with those predicted by the SNP88 score in population deciles (Fig 2C). The observed SNP score risk was, in general, closer to unity than expected, and the log-odds ratio for SNP88 was estimated to be 46% (95% CI, 19% to 74%) of expected (Data Supplement). Calibration of the SNP score was slightly better for ER-positive cancers, where risk was 56% (95% CI, 23% to 88%) of expected. After allowance for TC in a bivariable model, the observed SNP88 risk was 49% (95% CI, 21% to 77%) of expected.

The TC model risk distribution was not significantly different from observed in this case-control study (slope, 71%; 95% CI, 37% to 106%; Data Supplement). In addition, TC was well calibrated for absolute risk in the IBIS-I placebo arm, where 363 cancers were

Table 2. Discrimination of the SNP Score Alone and in Combination with the TC Model, and in Subgroups Based on Trial, ER Status, and Treatment Arm With Tests for Heterogeneity

Result	IQ-OR	95% CI	LR-χ^2	P	mC	mC 95% CI
Effects by SNP88	1.37	1.14 to 1.66	11.3	< .001	0.55	0.51 to 0.60
Effects by TC alone	1.45	1.21 to 1.73	17.6	< .001	0.55	0.51 to 0.60
Effects by combined TC × SNP88 (df = 1)*	1.64	1.36 to 1.97	28.6*	< .001	0.60	0.55 to 0.64
			$\Delta\chi^2_1$ = 12.4†	< .001		
SNP88 effects by subgroup						
Trial						
IBIS-I	1.45	1.17 to 1.78	12.6	< .001	0.57	0.52 to 0.62
Marsden	1.11	0.73 to 1.69	0.2	.6	0.49	0.39 to 0.58
Heterogeneity				.21		
ER status						
ER-positive	1.44	1.16 to 1.79	11.6	< .001	0.56	0.50 to 0.61
ER-negative	0.99	0.61 to 1.61	0.0	1.0	0.52	0.42 to 0.63
Heterogeneity				.10		
Treatment						
Placebo	1.46	1.13 to 1.87	8.9	.003	0.57	0.52 to 0.63
Tamoxifen	1.25	0.96 to 1.64	2.8	.094	0.52	0.45 to 0.60
Heterogeneity				.5		
TC effects by subgroup						
Trial						
IBIS-I	1.44	1.19 to 1.73	15.3	< .001	0.55	0.50 to 0.60
Marsden	1.28	0.94 to 1.73	2.6	.11	0.56	0.46 to 0.66
Heterogeneity				.6		
ER status						
ER-positive	1.33	1.10 to 1.60	9.2	.002	0.55	0.49 to 0.60
ER-negative	1.94	1.22 to 3.07	8.8	.003	0.59	0.48 to 0.68
Heterogeneity				.23		
Treatment						
Placebo	1.54	1.21 to 1.96	13.3	< .001	0.55	0.49 to 0.61
Tamoxifen	1.31	1.02 to 1.68	4.6	.032	0.56	0.50 to 0.63
Heterogeneity				.6		
Combined TC × SNP88 effects by subgroup						
Trial						
IBIS-I	1.68	1.37 to 2.05	27.5	< .001	0.61	0.56 to 0.66
Marsden	1.33	0.92 to 1.92	2.3	.13	0.54	0.45 to 0.64
Heterogeneity				.28		
ER status						
ER-positive	1.62	1.31 to 1.99	21.9	< .001	0.60	0.54 to 0.65
ER-negative	1.63	0.96 to 2.76	3.5	.063	0.58	0.46 to 0.69
Heterogeneity				.6		
Treatment						
Placebo	1.79	1.39 to 2.31	22.3	< .001	0.62	0.56 to 0.68
Tamoxifen	1.43	1.10 to 1.88	7.2	.007	0.56	0.49 to 0.63
Heterogeneity				.36		

Abbreviations: ER, estrogen receptor; IBIS-I, International Breast Intervention Study; IQ-OR, odds ratio difference between 25th and 75th percentile in controls; LR-χ^2, log likelihood-ratio χ^2 statistic; Marsden, Royal Marsden study; mC, matched concordance indices; SNP, single nucleotide polymorphism; TC, Tyrer-Cuzick.
*TC × SNP88, combined multiplicatively (df = 1).
†$\Delta\chi^2$ is for improvement when SNP88 is added to TC in a bivariable model.

diagnosed compared with 361 expected, and calibration was maintained by decile (Hosmer-Lemeshow χ^2_8 = 13.7; P = .09; Data Supplement).

When reweighted to the original population, the percentage of women with 10-year predicted risk above 8% if untreated was 18% for TC and increased to 21% if recalibrated SNP88 was added. Table 3 lists the number of cases and controls who were reclassified when recalibrated SNP88 was added to TC. It shows that although the number in the ≥ 8% group was not substantially increased, those who were upgraded were at a significantly higher risk than were those who were downgraded (OR, 2.67; 95% CI, 1.12 to 6.60; P = .015).

No significant interactions between SNP88 and any subgroups were observed (all P > 0.1; Data Supplement). In sensitivity

analyses, there was little difference when missing data were imputed (Data Supplement) or when the analysis included all tissue samples that met the quality control thresholds (Data Supplement).

DISCUSSION

A polygenic SNP score on the basis of common low penetrance genetic polymorphisms provides additional risk information for women who are already at an increased risk for the disease as a result of classic factors. SNP88 was independent of the TC risk model, and no loss of relative predictive power was seen when it was used in combination with TC. However, the SNP panel seems
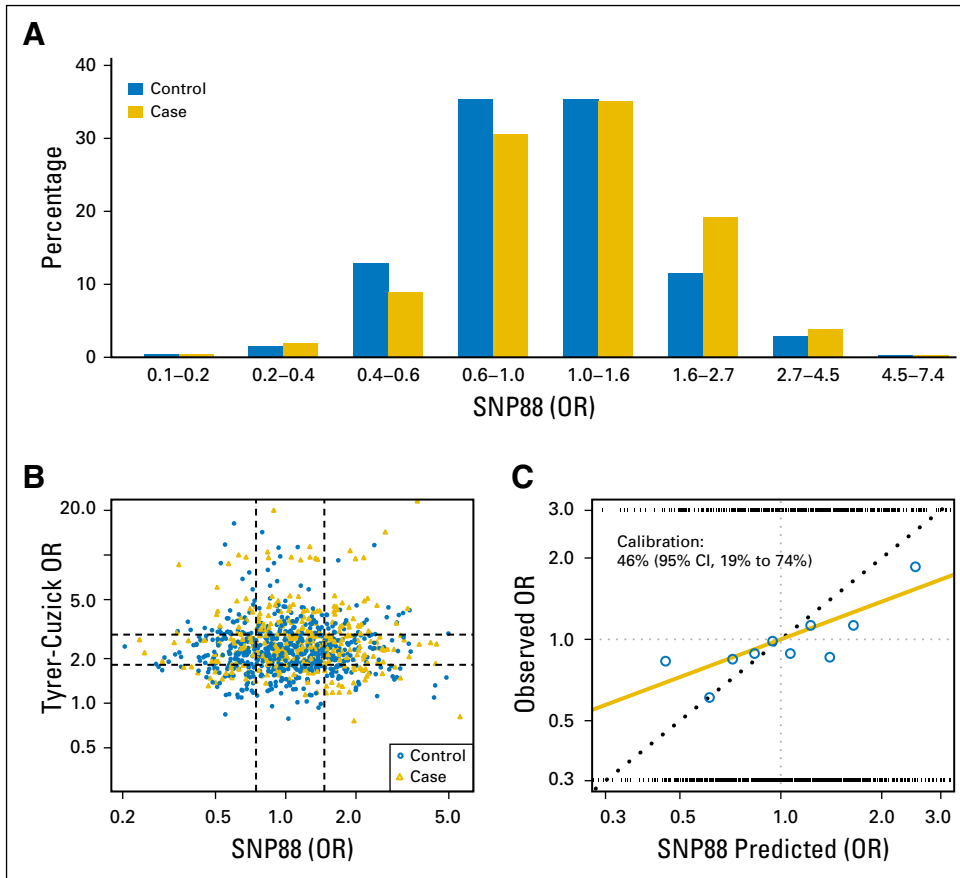
Fig 2. (A) Distribution of 88 single nucleotide polymorphisms risk score (SNP88) in cases and controls. (B) SNP88 and Tyrer-Cuzick odds ratios (ORs) for cases and controls; Spearman coefficient 0.012, $P = .7$, where each interquartile range in controls is represented by the dashed lines. (C) logistic regression fit (gold line) of observed versus predicted ORs from the SNP score in both arms, blue points are estimates from each decile of SNP88 in controls, and the rug plot at the top and bottom gives the observed SNP88 in cases and controls, respectively.

to be in need of recalibration, at least when applied to high-risk women in similar populations. SNP88 had less than half the expected risk on the basis of multiplying published risks for individual SNPs. In addition, the spread of risks associated with SNP88 were of a smaller magnitude than those of TC (Table 1), especially after recalibration. Thus, although the SNP score had minimal ability to increase the number of classically high-risk women to be at more than 8% 10-year risk (Table 3), it did more accurately populate this group. The SNP score also did not predict which women would differentially benefit from tamoxifen.

One possible reason for the poor calibration of the SNP score is overfitting. Using the published odds ratios and CIs, we applied a James-Stein shrinkage estimator and estimated that this might account for overfitting of approximately 5% in calibration (data not shown). Thus, it is unlikely to explain most of the difference between observed and expected risks.

In general populations, polygenic scores have demonstrated good calibration. For example, Vachon et al[28] used a SNP76 panel in a general population and found an OR per standard deviation (SD) of approximately 1.5, compared with 1.6 expected (calculated by computer simulation following Brentnall et al[19]): calibration was approximately 86% [$\log_e(1.5)/\log_e(1.6)$]. Dite et al[30] reported a SNP77 panel in women with a family history. The OR per SD was approximately 1.4, which was only approximately 70% of expected. In the context of prevention trials, Vachon et al[31] reported an odds ratio of approximately 1.8 for the difference between the top and bottom quintile of their SNP75 panel, compared with approximately 3.8 if the score had been well calibrated. Thus, overall, our results are not inconsistent with previous findings for women with an elevated risk of cancer, and it seems that recalibration of polygenic risk scores is needed for these women. A strong family history might change the relationship between SNPs

Table 3. Reclassification of the Number of Cases of the Total Number of Cases and Controls, by 8% 10-Year Risk Groups From TC Alone and When Combined With cSNP88

TC Risk	< 8%, No. (%)	TC × cSNP88 Risk ≥ 8%, No. (%)	Total, No (%)
< 8%	225 of 702 (32)	**40 of 85 (47)**	265 of 787 (34)
≥ 8 %	**11 of 44 (25)**	83 of 164 (51)	94 of 208 (45)
Total	236 of 746 (32)	123 of 249 (49)	359 of 995 (36)

NOTE. Bold cells indicate those who changed risk group. Percentages are based on the case-control population and are not weighted back to the trial population.
Abbreviations: cSNP88, recalibrated 88 single nucleotide polymorphisms risk score; TC, Tyrer-Cuzick model.

and risk because risks for each SNP are different for *BRCA1* and *BRCA2* mutation carriers than for the general population.[38,39] The percentage of *BRCA1* and *BRCA2* carriers was estimated to be 2.8% in Marsden and 1.9% in IBIS-I from the TC model; therefore, this is only a partial explanation.

A limitation of this study is that we could not assess the performance of SNPs in conjunction with mammographic breast density, which is another important component of risk.[40] Also, the findings are only applicable to similar high-risk groups of women of the same age. It is likely that these findings apply to other risk models, such as the Gail model, because SNP scores are largely independent of the factors used.[25]

The study has several strengths. First, the cancers were diagnosed prospectively from an extended follow-up period. Second, we were able to compare the performance within the tamoxifen and placebo arms directly, without indirect comparisons with other studies. Third, we were able to assess how much an SNP score adds compared with a risk model on the basis of classic factors (except for age). This was not possible in the analysis from the P1 and P2 trials,[31] because they matched on risk from classic factors. Another advantage of the current study is that the SNPs were genotyped directly (36 of 75 were imputed in Vachon et al[31]), so that a poor imputation procedure is not a possible reason for the calibration issue.

In conclusion, our findings extend earlier work to high-risk populations and indicate that SNP scores increase the accuracy of risk assessment for these women, but substantial recalibration seems to be required for accurate risk assessment.

AUTHOR CONTRIBUTIONS

Conception and design: Jack Cuzick, William G. Newman, Mitchell Dowsett
Collection and assembly of data: Corrinne Segal, Helen Byers, Caroline Reuter, Simone Detre, Elena Lopez-Knowles, Ivana Sestak, William G. Newman
Data analysis and interpretation: Jack Cuzick, Adam R. Brentnall, Ivana Sestak, Anthony Howell, Trevor J. Powles, William G. Newman, Mitchell Dowsett
Manuscript writing: All authors
Final approval of manuscript: All authors
Accountable for all aspects of the work: All authors

REFERENCES

1. Bogdanova N, Helbig S, Dörk T: Hereditary breast cancer: Ever more pieces to the polygenic puzzle. Hereditary Cancer Clin Pract 11:12, 2013

2. Easton DF, Pooley KA, Dunning AM, et al: Genome-wide association study identifies novel breast cancer susceptibility loci. Nature 447:1087-1093, 2007

3. Cox A, Dunning AM, Garcia-Closas M, et al: A common coding variant in CASP8 is associated with breast cancer risk. Nat Genet 39:352-358, 2007

4. Stacey SN, Manolescu A, Sulem P, et al: Common variants on chromosomes 2q35 and 16q12 confer susceptibility to estrogen receptor-positive breast cancer. Nat Genet 39:865-869, 2007

5. Stacey SN, Manolescu A, Sulem P, et al: Common variants on chromosome 5p12 confer susceptibility to estrogen receptor-positive breast cancer. Nat Genet 40:703-706, 2008

6. Thomas G, Jacobs KB, Kraft P, 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 41:579-584, 2009

7. Ahmed S, Thomas G, Ghoussaini M, et al: Newly discovered breast cancer susceptibility loci on 3p24 and 17q23.2. Nat Genet 41:585-590, 2009

8. Zheng W, Long J, Gao YTT, et al: Genome-wide association study identifies a new breast cancer susceptibility locus at 6q25.1. Nat Genet 41:324-328, 2009

9. Antoniou AC, Wang X, Fredericksen ZS, et al: A locus on 19p13 modifies risk of breast cancer in BRCA1 mutation carriers and is associated with hormone receptor-negative breast cancer in the general population. Nat Genet 42:885-892, 2010

10. Turnbull C, Ahmed S, Morrison J, et al: Genome-wide association study identifies five new breast cancer susceptibility loci. Nat Genet 42:504-507, 2010

11. Haiman CA, Chen GK, Vachon CM, et al: A common variant at the TERT-CLPTM1L locus is associated with estrogen receptor-negative breast cancer. Nat Genet 43:1210-1214, 2011

12. Fletcher O, Johnson N, Orr N, et al: Novel breast cancer susceptibility locus at 9q31.2: Results of a genome-wide association study. J Natl Cancer Inst 103:425-435, 2011

13. Ghoussaini M, Fletcher O, Michailidou K, et al: Genome-wide association analysis identifies three new breast cancer susceptibility loci. Nat Genet 44:312-318, 2012

14. French JD, Ghoussaini M, Edwards SL, et al: Functional variants at the 11q13 risk locus for breast cancer regulate cyclin D1 expression through long-range enhancers. Am J Hum Genet 92:489-503, 2013

15. Bojesen SE, Pooley KA, Johnatty SE, et al: Multiple independent variants at the TERT locus are associated with telomere length and risks of breast and ovarian cancer. Nat Genet 45:371-384, e1-e2, 2013

16. Michailidou K, Hall P, Gonzalez-Neira A, et al: Large-scale genotyping identifies 41 new loci associated with breast cancer risk. Nat Genet 45:353-361, e1-e2, 2013

17. Garcia-Closas M, Couch FJ, Lindstrom S, et al: Genome-wide association studies identify four ER negative-specific breast cancer risk loci. Nat Genet 45:392-398, e1-e2, 2013

18. Michailidou K, Beesley J, Lindstrom S, et al: Genome-wide association analysis of more than 120,000 individuals identifies 15 new susceptibility loci for breast cancer. Nat Genet 47:373-380, 2015

19. Brentnall AR, Evans DG, Cuzick J: Distribution of breast cancer risk from SNPs and classical risk factors in women of routine screening age in the UK. Br J Cancer 110:827-828, 2014

20. Evans DG, Brentnall AR, Harvie M, et al: Breast cancer risk in young women in the national breast screening programme: Implications for applying NICE guidelines for additional screening and chemoprevention. Cancer Prev Res (Phila) 7:993-1001, 2014

21. Gail MH, Brinton LA, Byar DP, et al: Projecting individualized probabilities of developing breast cancer for white females who are being examined annually. J Natl Cancer Inst 81:1879-1886, 1989

22. Tyrer J, Duffy SW, Cuzick J: A breast cancer prediction model incorporating familial and personal risk factors. Stat Med 23:1111-1130, 2004

23. Tice JA, Cummings SR, Smith-Bindman R, et al: Using clinical factors and mammographic breast density to estimate breast cancer risk: Development and validation of a new predictive model. Ann Intern Med 148:337-347, 2008

24. Antoniou AC, Cunningham AP, Peto J, et al: The BOADICEA model of genetic susceptibility to breast and ovarian cancers: Updates and extensions. Br J Cancer 98:1457-1466, 2008 [Erratum: Br J Cancer 98:2015, 2008]

25. Wacholder S, Hartge P, Prentice R, et al: Performance of common genetic variants in breast-cancer risk models. N Engl J Med 362:986-993, 2010

26. Mealiffe ME, Stokowski RP, Rhees BK, et al: Assessment of clinical validity of a breast cancer risk model combining genetic and clinical information. J Natl Cancer Inst 102:1618-1627, 2010

27. Lee CP, Irwanto A, Salim A, et al. Breast cancer risk assessment using genetic variants and risk factors in a Singapore Chinese population. Breast Cancer Res 16:R64, 2014

28. Vachon CM, Pankratz VS, Scott CG, et al: The contributions of breast density and common genetic variation to breast cancer risk. J Natl Cancer Inst 107:dju397, 2015

29. Mavaddat N, Pharoah PD, Michailidou K, et al: Prediction of breast cancer risk based on profiling

with common genetic variants. J Natl Cancer Inst 107:djv036, 2015

30. Dite GS, MacInnis RJ, Bickerstaffe A, et al: Breast cancer risk prediction using clinical models and 77 independent risk-associated SNPs for women aged under 50 years: Australian Breast Cancer Family Registry. Cancer Epidemiol Biomarkers Prev 25:359-365, 2016

31. Vachon CM, Schaid DJ, Ingle JN, et al: A polygenic risk score for breast cancer in women receiving tamoxifen or raloxifene on NSABP P-1 and P-2. Breast Cancer Res Treat 149:517-523, 2015

32. Cuzick J, Sestak I, Cawthorn S, et al: Tamoxifen for prevention of breast cancer: Extended long-term follow-up of the IBIS-I breast cancer prevention trial. Lancet Oncol 16:67-75, 2015

33. Powles TJ, Ashley S, Tidy A, et al: Twenty-year follow-up of the Royal Marsden randomized, double-blinded tamoxifen breast cancer prevention trial. J Natl Cancer Inst 99:283-290, 2007

34. Brentnall AR, Cuzick J, Field J, et al: A concordance index for matched case-control studies with applications in cancer risk. Stat Med 34:396-405, 2015

35. Brentnall AR, Cuzick J, Byers H, et al: Relationship of ZNF423 and CTSO with breast cancer risk in two randomised tamoxifen prevention trials. Breast Cancer Res Treat 158:591-596, 2016

36. Abecasis GR, Auton A, Brooks LD, et al: An integrated map of genetic variation from 1,092 human genomes. Nature 491:56-65, 2012

37. Browning SR: Multilocus association mapping using variable-length Markov chains. Am J Hum Genet 78:903-913, 2006

38. Couch FJ, Wang X, McGuffog L, et al: Genome-wide association study in BRCA1 mutation carriers identifies novel loci associated with breast and ovarian cancer risk. PLoS Genet 9:e1003212, 2013

39. Gaudet MM, Kuchenbaecker KB, Vijai J, et al: Identification of a BRCA2-specific modifier locus at 6p24 related to breast cancer risk. PLoS Genet. 9: e1003173, 2013

40. Cuzick J, Warwick J, Pinney E, et al: Tamoxifen-induced reduction in mammographic density and breast cancer risk reduction: A nested case-control study. J Natl Cancer Inst 103:744-752, 2011

Affiliations

Jack Cuzick, **Adam R. Brentnall**, **Caroline Reuter**, and **Ivana Sestak**, Centre for Cancer Prevention, Wolfson Institute of Preventive Medicine, Queen Mary University of London; **Corrinne Segal**, The Institute of Cancer Research; **Corrinne Segal**, **Simone Detre**, **Elena Lopez-Knowles**, and **Mitchell Dowsett**, Royal Marsden Hospital; **Trevor J. Powles**, Cancer Centre London, London; **Helen Byers** and **William G. Newman**, University of Manchester and Central Manchester Foundation Trust; and **Anthony Howell**, The Christie NHS Foundation Trust, Manchester, United Kingdom.

Support

Prior Presentation

Presented at the 40th Annual Meeting of the American Society of Clinical Oncology, Chicago, IL, May 30 to June 3, 2014.

AUTHORS' DISCLOSURES OF POTENTIAL CONFLICTS OF INTEREST

Impact of a Panel of 88 Single Nucleotide Polymorphisms on the Risk of Breast Cancer in High-Risk Women: Results From Two Randomized Tamoxifen Prevention Trials

The following represents disclosure information provided by authors of this manuscript. All relationships are considered compensated. Relationships are self-held unless noted. I = Immediate Family Member, Inst = My Institution. Relationships may not relate to the subject matter of this manuscript. For more information about ASCO's conflict of interest policy, please refer to www.asco.org/rwc or ascopubs.org/jco/site/ifc.

Jack Cuzick
Consulting or Advisory Role: Becton Dickinson, Hologic, Merck
Speakers' Bureau: Myriad Genetics, Cepheid, Trovagene, Abbott Diagnostics, Merck
Research Funding: Avantis Medical Systems (Inst), Qiagen (Inst), Becton Dickinson (Inst), Abbott Diagnostics (Inst), Hologic (Inst), Trovagene (Inst), OncoHealth (Inst), Genera Biosystems (Inst), Cepheid (Inst)
Patents, Royalties, Other Intellectual Property: Royalties from Myriad Genetics to institution, of which I receive a share for development of cell cycle progression score (Prolaris), royalty payments through Cancer Research UK from licenses for commercial use of the Tyrer-Cuzick algorithm
Travel, Accommodations, Expenses: Merck

Adam R. Brentnall
Patents, Royalties, Other Intellectual Property: Royalty payments through Cancer Research UK from licenses for commercial use of the Tyrer-Cuzick algorithm

Corrinne Segal
No relationship to disclose

Helen Byers
No relationship to disclose

Caroline Reuter
No relationship to disclose

Simone Detre
No relationship to disclose

Elena Lopez-Knowles
No relationship to disclose

Ivana Sestak
No relationship to disclose

Anthony Howell
No relationship to disclose

Trevor J. Powles
No relationship to disclose

William G. Newman
No relationship to disclose

Mitchell Dowsett
Honoraria: Novartis, Pfizer, Myriad Genetics
Consulting or Advisory Role: Genoptix, Genentech, Janssen Oncology, GTx, Radius
Research Funding: Roche (Inst), AstraZeneca (Inst), Novartis (Inst), Pfizer (Inst), Radius Health (Inst)
Other Relationship: Institute of Cancer Research

© 2016 by American Society of Clinical Oncology

Acknowledgment