Prediction of Breast Cancer Risk Based on Profiling With Common Genetic Variants


Published in:
Journal of the National Cancer Institute

Document Version:
Publisher's PDF, also known as Version of record

Queen's University Belfast - Research Portal:
Link to publication record in Queen's University Belfast Research Portal

Publisher rights
Copyright 2015 the authors. This is an open access article published under a Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution and reproduction in any medium, provided the author and source are cited.

General rights
Copyright for the publications made accessible via the Queen's University Belfast Research Portal is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

Take down policy
The Research Portal is Queen's institutional repository that provides access to Queen's research output. Every effort has been made to ensure that content in the Research Portal does not infringe any person's rights, or applicable UK laws. If you discover content in the Research Portal that you believe breaches copyright or violates any law, please contact openaccess@qub.ac.uk.

Download date: 11. Jan. 2019
Prediction of Breast Cancer Risk Based on Profiling With Common Genetic Variants


Received: March 28, 2014; Revised: December 1, 2014; Accepted: January 26, 2015

© The Author 2015. Published by Oxford University Press.

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/3.0/), which permits unrestricted reuse, distribution, and reproduction in any medium, provided the original work is properly cited.
Breast cancer is the most common cancer among Western women, with approximately 1.67 million cases diagnosed annually worldwide (1). Strategies such as endocrine risk-reducing medication and early detection by breast cancer screening can reduce the burden of disease but have disadvantages including side effects, overdiagnosis, and increased cost (2–4). Stratification of women according to the risk of developing breast cancer could improve risk reduction and screening strategies by targeting those most likely to benefit (5–8).

Both genetic and lifestyle factors are implicated in the etiology of breast cancer. Women with a history of breast cancer in a first-degree relative are at approximately two-fold higher risk than women without a family history (9). Rare high-risk mutations particularly in the BRCA1 and BRCA2 genes explain less than 20% of the two-fold familial relative risk (FRR) (10) and account for a small proportion of breast cancer cases in the general population. Low frequency variants conferring intermediate risk, such as those in CHEK2, ATM, and PALB2, explain 2% to 5% of the FRR. Genome-wide association studies (GWAS) have led to the discovery of multiple common, low-risk variants (single nucleotide polymorphisms [SNPs]) associated with breast cancer risk (11), many of which are differentially associated by estrogen receptor (ER) status (12,13). Recently, new risk-associated variants have been identified in a large-scale replication study conducted by the Breast Cancer Association Consortium (BCAC) as part of the Collaborative Oncological Gene-Environment Study (COGS). SNPs were genotyped in over 40,000 breast cancer cases and 40,000 control women, using a custom array (iCOGS). This experiment increased the number of SNPs robustly associated with breast cancer from 27 to more than 70 and identified additional variants specific to ER-negative breast cancer (14–17).

Risks conferred by SNPs are not sufficiently large to be useful in risk prediction individually. However, the combined effect of multiple SNPs could achieve a degree of risk discrimination that is useful for population-based programmes of breast cancer prevention and early detection (8,18). In this report, we investigated the value of using all 77 breast cancer susceptibility loci identified to date for risk stratification. Previous studies of polygenic risk have assumed a log-additive model for combining SNPs; however, this assumption needs to be evaluated empirically. We first assessed whether interaction between SNP pairs could influence the joint contribution of genetic factors on disease risk by testing for all possible pair-wise interactions between SNPs. We then constructed polygenic risk scores (PRSs) to capture the combined...
effects of the 77 SNPs on overall breast cancer risk, as well as on the risk of ER-positive and ER-negative disease separately. We estimated absolute risks of developing breast cancer for different levels of the PRS, accounting for the competing risk of mortality from other causes. Effect sizes were confirmed in one large study (pKARMA) that was not part of any SNP discovery set. We discuss the degree of breast cancer risk stratification obtained in women with and without a family history of breast cancer.

Methods

Study Subjects and Genotyping

Study participants for the primary analyses (set 1) were 89,049 women of European origin participating in 41 studies in BCAC. All studies were approved by the relevant institutional review boards, and all individuals gave written informed consent. Samples were genotyped using a custom Illumina iSelect array (iCOGS) comprising 211,155 SNPs (15). For some analyses, a further 72,014 women in BCAC genotyped for the relevant SNPs in earlier experiments were included (set 2). For PRS analyses (67,054 women), studies that oversampled breast cancer cases with a family history (21,995 women) were excluded. Supplementary Tables 1–3 (available online) show study designs and numbers of breast cancer cases and control women included.

Analyses were based primarily on variants reported to be associated (at $P < 5 \times 10^{-8}$) by COGS or previous publications, with either breast cancer overall or ER-negative disease. SNPs and regions included are summarized in Supplementary Table 4 (available online).

Statistical Methods

Tests for pair-wise SNP*SNP interactions (departures from a multiplicative model) were carried out using logistic regression, with breast cancer as the outcome. The two SNPs were each coded as a categorical variable (i.e., fitting a separate parameter for heterozygous and risk-allele homozygous genotypes), while the interaction term (SNP1*SNP2) was included as continuous covariate. All analyses were adjusted for study and seven principal components (PC) to account for population substructure.

Table 1. Observed and expected numbers of statistically significant pair-wise tests for SNP*SNP interaction at $P < .01$

<table>
<thead>
<tr>
<th>Type of breast cancer</th>
<th>Case-control analyses</th>
<th>Case-only analyses†</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>OBS</td>
<td>OBS/EXP</td>
</tr>
<tr>
<td>All SNPs¶</td>
<td>n = 3080 SNP pairs</td>
<td></td>
</tr>
<tr>
<td>All breast cancers</td>
<td>44</td>
<td>1.43</td>
</tr>
<tr>
<td>ER-positive</td>
<td>43</td>
<td>1.40</td>
</tr>
<tr>
<td>ER-negative</td>
<td>35</td>
<td>1.13</td>
</tr>
<tr>
<td>Unlinked SNPs¶</td>
<td>n = 2556 SNP pairs</td>
<td></td>
</tr>
<tr>
<td>All breast cancers</td>
<td>35</td>
<td>1.37</td>
</tr>
<tr>
<td>ER-positive</td>
<td>38</td>
<td>1.49</td>
</tr>
<tr>
<td>ER-negative</td>
<td>30</td>
<td>1.17</td>
</tr>
</tbody>
</table>

† 46,450 breast cancer cases and 42,599 control women were included in the analysis of all breast cancers. 27,074 breast cancer cases were included in the analysis of ER-positive disease and 7413 breast cancer cases were included in the analysis of ER-negative disease. $n =$ number of single nucleotide polymorphism (SNP) pairs tested; OBS = number of tests observed with $P_{\text{interaction}} < .01$; OBS/EXP = number of tests observed with $P_{\text{interaction}} < .01$ divided by the number of positive tests expected by chance, given the number of SNP pairs tested; SNP = single nucleotide polymorphism.

‡ Only results of SNP pairs not strongly associated in the control population ($P_{\text{interaction}} > .01$ in control-only analyses) were included in the counts.

§ $P$ value for difference between observed and expected numbers of tests, assuming each test is independent and that, under the null hypothesis, the observed number of statistically significant tests follows a poisson distribution. The statistical test was two-sided.

¶ Some SNPs were linked, as described in the Supplementary Methods (available online).

|| Only the most statistically significant SNP from each group of linked SNPs were included in these analyses.

To investigate the association between breast cancer risk and the combined effects of 77 SNPs, a PRS was derived for each individual using the formula:

$$\text{PRS} = \beta_1 x_1 + \beta_2 x_2 + \ldots + \beta_n x_n$$

where $\beta_i$ is the per-allele log odds ratio (OR) for breast cancer associated with the minor allele for SNP $k$, and $x_k$ the number of alleles for the same SNP (0, 1, or 2), and $n = 77$ is the total number of SNPs. Thus, the PRS summarizes the combined effect of the SNPs, ignoring departures from a multiplicative model (18). SNPs and corresponding odds ratios used in derivation of PRSs are summarized in Supplementary Table 4 (available online).

Logistic regression models were used to estimate the odds ratios for breast cancer by percentile of the PRS, with the middle quintile category (40th to 60th percentile) as the reference. Observed odds ratios for breast cancer by percentile of the PRS were compared with predicted odds ratios under a multiplicative polygenic model of inheritance. Modification of the PRS by age or by family history of breast cancer in a first-degree relative was evaluated by fitting additional interaction terms in the model. All tests of statistical significance were two-sided. The thresholds for statistical significance are indicated below.

The absolute risk of overall breast cancer, ER-positive and ER-negative breast cancer for individuals in each risk category, was calculated taking into account the competing risk of dying from other causes apart from breast cancer. Approximate confidence limits for the absolute risk were derived from the variance-covariance matrix of the log (relative risk) parameters in the logistic regression analysis. Detailed methods are provided in Supplementary Methods (available online).

Results

Pairwise Multiplicative SNP*SNP Interaction Analyses

Data on 46,450 breast cancer cases and 42,599 controls from 41 studies were included in the interaction analyses...
(Supplementary Table 3, available online). There was no strong evidence for interaction between any particular SNP pair after Bonferroni correction (Supplementary Tables 5–6, available online). Plots of expected vs observed log_{10}P values for SNP*SNP interaction tests showed slight departure from the null hypothesis of multiplicative effects (Supplementary Figure 1, A and B, available online), and the number of statistically significant interactions with P_{interaction} values of less than .01 was larger than expected by chance (Table 1). To investigate whether there was an excess of synergistic or antagonistic interactions, the direction of the interaction term relative to the main effects was examined for SNP pairs with P_{interaction} values of less than .01. For case-control analyses, 47% of interactions were synergistic and 53% antagonistic, and for case-only analyses 53% were synergistic and 46% antagonistic. These proportions were not statistically significantly different from the null expectation (P > .05). Meta-analysis of SNPs*SNPs interaction test results from the iCOGS dataset with those from 72,014 additional women (Supplementary Table 7, available online). Given that no SNP pair showed strong evidence for departure from the multiplicative model, subsequent analyses were based on a PRS that included the main effects of SNPs but no SNP*SNP interaction terms.

**Association Between PRS and Breast Cancer Risk**

As predicted by the polygenic, multiplicative model, the number of breast cancer risk alleles and the 77-SNP PRS approximated a normal distribution for both breast cancer cases and control women (Figure 1). The odds ratios for developing breast cancer by percentiles of the PRS, compared with women in the middle quintile (40th to 60th percentile) are shown in Figure 2A. The observed odds ratios were similar to the odds ratios predicted under a polygenic multiplicative model; the 95% confidence interval (CI) included the predicted odds ratio at all points except the 80th to 90th percentile (Figure 2A; Supplementary Table 8, available online). For women in the lowest 1% of the PRS distribution, the estimated odds ratio compared with women in the middle quintile was 0.32 (95% CI = 0.25 to 0.40). By contrast, for women in the highest 1% of the PRS distribution, the estimated OR compared with women in the middle quintile was 3.36 (95% CI = 2.95 to 3.83, P = 7.5 x 10^{-6}). When P-were derived separately for ER-positive and ER-negative disease, the corresponding odds ratios were 3.73 (95% CI = 3.24 to 4.30) and 2.80 (95% CI = 2.26 to 3.46), respectively (Figure 2B and C). The log OR per unit standard deviation of the PRS was 0.44 (95% CI = 0.42 to 0.46) for overall breast cancer, 0.49 (95% CI = 0.47 to 0.51) for ER-positive disease, and 0.37 (95% CI = 0.34 to 0.40) for ER-negative disease (Table 3). A validation analysis including only one large study (pKARMA) that was not part of any SNP discovery analyses found similar odds ratio estimates to those in the remaining studies, except for the 60th to 80th and 90th to 95th percentiles, for which estimates were higher in pKARMA (Table 4; Supplementary Table 9, available online). The log OR per unit SD was also similar for pKARMA alone (log OR per unit SD = 0.4). The associations between PRS and breast cancer in different age groups are summarized in Table 3 and Supplementary Figure 2 (available online). There was a statistically significant interaction between PRS and age, the association between PRS and breast cancer risk decreasing with age (Table 3).

A family history of breast cancer in one of more affected first-degree relatives was reported by 18.5% of breast cancer cases and 11.1% of control women. The odds ratio for family history was attenuated from 1.81 to 1.68 (12.6% attenuation) after adjusting for the PRS (Table 2). At younger ages (<40 years), there was less attenuation (from 2.90 to 2.76, 4.6% attenuation) (Table 2). The joint effects of the PRS and family history were largely consistent with a multiplicative model (P_{interaction} = .34 for the interaction between the PRS and family history; data not shown); however, we observed a stronger effect of family history for women at the lowest 1% of the PRS (Supplementary Table 10, available online).

The discriminative accuracy of the PRS, as measured by the C-statistic, was 0.622 (95% CI = 0.619 to 0.627); discrimination was worse for ER-negative disease (0.42) than for ER-positive disease (0.49). For both sexes and ER status, the interaction term was of similar magnitude to the main effects of the PRS. For overall breast cancer, the C-statistic, was 0.622 (95% CI = 0.619 to 0.627); discrimination was worse for ER-negative disease (0.42) than for ER-positive disease (0.49). For both sexes and ER status, the interaction term was of similar magnitude to the main effects of the PRS.

**Interaction Between PRS and Family History**

Table 2. Odds ratio for family history of breast cancer in first-degree relatives: unadjusted and adjusted by PRS and stratified by age

<table>
<thead>
<tr>
<th>Age group*</th>
<th>OR* (95% CI)</th>
<th>Adjusted by PRS</th>
<th>% attenuation†</th>
</tr>
</thead>
<tbody>
<tr>
<td>All subjects</td>
<td>1.81 (1.69 to 1.93)</td>
<td>1.68 (1.56 to 2.86)</td>
<td>12.6%</td>
</tr>
<tr>
<td>&lt;40 y</td>
<td>2.90 (2.07 to 4.07)</td>
<td>2.76 (1.96 to 3.89)</td>
<td>4.6%</td>
</tr>
<tr>
<td>40–60 y</td>
<td>1.88 (1.71 to 2.08)</td>
<td>1.72 (1.56 to 1.90)</td>
<td>14.1%</td>
</tr>
<tr>
<td>≥60 y</td>
<td>1.63 (1.47 to 1.82)</td>
<td>1.53 (1.37 to 1.70)</td>
<td>13.0%</td>
</tr>
</tbody>
</table>

* Odds ratio for developing breast cancer for women with a family history of breast cancer in a first-degree relative compared with women without a family history, adjusting for study and seven principal components. 21,865 breast cancer cases and 15,830 control women provided family history information. CI = confidence intervals; PRS = polygenic risk score; OR = odds ratio.
† Percent attenuation on log scale.

**Table 3. Association between PRS and breast cancer risk in different age groups**

<table>
<thead>
<tr>
<th>Age group*</th>
<th>log OR† (95% CI)</th>
<th>log OR (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>All ages</td>
<td>0.44 (0.42 to 0.46)</td>
<td>0.49 (0.47 to 0.51)</td>
</tr>
<tr>
<td>&lt;40 y</td>
<td>0.46 (0.38 to 0.53)</td>
<td>0.56 (0.47 to 0.65)</td>
</tr>
<tr>
<td>40–49 y</td>
<td>0.46 (0.42 to 0.50)</td>
<td>0.53 (0.48 to 0.57)</td>
</tr>
<tr>
<td>50–59 y</td>
<td>0.48 (0.45 to 0.51)</td>
<td>0.54 (0.50 to 0.57)</td>
</tr>
<tr>
<td>≥60 y</td>
<td>0.41 (0.38 to 0.43)</td>
<td>0.44 (0.41 to 0.47)</td>
</tr>
</tbody>
</table>

Interaction between PRS and age

<table>
<thead>
<tr>
<th>Interaction OR† (95% CI)</th>
<th>Interaction OR (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.98 (0.96 to 0.99)</td>
<td>0.97 (0.95 to 0.98)</td>
</tr>
</tbody>
</table>

* OR per 10 years for interaction between PRS and age.
† log OR for association between the PRS coded as a continuous variable and breast cancer risk (per unit SD of the PRS)
In this report, we evaluated the degree of breast cancer risk stratification that can be attained in women of European ancestry using data for 77 common genetic variants, summarized as a PRS. Our results show that the PRS stratifies breast cancer risk in women without family history and refines genetic risk in women with a family history of breast cancer.

The PRS we used (sum of the minor alleles weighted by the per-allele log OR) is the most efficient, assuming that SNP odds ratios combine multiplicatively (i.e., no interactions on a log-additive scale) (18). Evaluation of pairwise SNP interactions showed that this was a reasonable assumption. Although no individual interactions could be established, we observed an excess of multiplicative interactions at P less than .01. This could be the result of underlying population stratification not accounted for by principal components adjustment or reflect the presence of multiple interactions too weak to be established individually. A recent study also found no evidence for interactions among SNPs with weaker evidence for main effects (19). Although we did not test for higher order interactions among SNPs, consistency between empirical and predicted odds ratios assuming multiplicative effects suggests that across all possible multivariable interactions the overall effect is close to multiplicative.

The 77-SNP PRS was associated with a larger effect than previously reported for a 10-SNP PRS (20). For example, our odds ratio for breast cancer for women in the highest compared with the middle quintile was 1.82 (95% CI = 1.73 to 1.90) vs 1.44 (95% CI = 1.35 to 1.53) for the 10-SNP PRS (20). A potential concern is that the PRS was constructed using iCOGS data that were, in part, the basis for discovery of many of the loci. This could lead to some upward bias in the odds ratio estimates (winner’s curse); however, analyses based on a large study (pKARMA) that was not part of any discovery set obtained similar estimates indicating that any winner’s curse effect is likely to be small.

There has been little evidence of differences by age in the per-allele odds ratio for individual SNPs. However, we observed a small but statistically significant decrease in odds ratio for PRS with increasing age. As expected, the odds ratio for family history was reduced after adjustment for the PRS. This attenuation (~12.6%) was consistent with the estimated fraction of the two-fold FRR explained by the 77-SNPs under a polygenic risk model (15). The joint effects of PRS and family history were consistent with a multiplicative model. A stronger FRR was observed for women at the lowest percentile of the PRS, but this was based on small numbers and requires confirmation. The degree of attenuation of the family history odds ratio was lower below age 40 years, as a result of the higher FRR at young ages, suggesting that rarer genetic variants may be more important at young ages.

We calculated the absolute risk of developing breast cancer for women at different levels of genetic risk according to the PRS. The lifetime risk for women below the first and above the 99th percentile of the PRS was 3.5% (95% CI = 2.6% to 4.4%) and 29.0% (95% CI = 24.9% to 33.5%), respectively, (Figure 3A). For the lowest and highest quintiles of the PRS, the risk was 5.3% (95% CI = 5.1% to 5.7%) and 17.2% (95% CI = 16.1% to 18.1%), respectively (data not shown). The corresponding risks of developing ER-positive disease were 4.1% and 15.7% for women in the lowest and highest quintiles, respectively, of the ER-positive PRS (averaged over all ER-negative PRS categories), whereas the highest lifetime risk for ER-negative disease was 2.4% (women in the highest quintile of ER-negative PRS and average ER-positive risk) (Figure 3). Lifetime risk of breast cancer for women in the lowest and highest quintiles of the PRS were 5.2% and 16.6% for a woman without family history and 8.6% and 24.4% for a woman with a first-degree family history of breast cancer (Figure 4).

We estimated the 10-year absolute risk of breast cancer at different ages and evaluated the age at which women at different levels of the PRS reach a threshold of 2.4%, which corresponds to the average 10-year risk of breast cancer for women age 47 years. This threshold was reached at 32 years for women whose PRS is above the 99th percentile of the PRS, and 57 years for women in the 20th to 40th percentiles of the PRS, and was never reached for women in lower percentiles (Figure 3D). As expected, lifetime risks were higher, and the ages at which the 2.4% threshold was reached were lower for women with a family history of breast cancer (Figure 4).

**Discussion**

In this report, we evaluated the degree of breast cancer risk stratification that can be attained in women of European ancestry using data for 77 common genetic variants, summarized as a PRS. Our results show that the PRS stratifies breast cancer risk in women without family history and refines genetic risk in women with a family history of breast cancer.
would, however, depend on the acceptability of risk-based surveillance, together with health economic considerations.

Prediction of subtype-specific breast cancer should also be informative for prevention (4). Recently updated NICE guidelines include recommendations to use endocrine treatments (tamoxifen and raloxifene) for primary prevention of breast cancer for women at moderate to high risk (21). These guidelines are based on risk of overall breast cancer for women with a family history of breast cancer. However, because these drugs prevent only ER-positive tumors, risk estimates incorporating the ER-positive PRS could better define the subset of women most likely to benefit. Our sample was derived from studies in Europe, North America, and Australia and restricted to women of European origin. While the results should be widely applicable in these populations, additional studies will be required to develop and validate genetic profiles for other populations, in particular Asian and African populations, where SNP associations, background incidence rates and distribution of tumour characteristics are substantially different.

Our analysis summarized family history in terms of a single binary variable, but familial risk of breast cancer also depends on the number of affected and unaffected relatives and their ages. Risk prediction algorithms that combine full family history data with a polygenic component perform better than simpler models (22). It is possible to incorporate the current PRS into family-history based models for breast cancer, such BOADICEA, to improve genetic risk prediction (23).

The COGS project includes the largest set of breast cancer studies with both phenotype and genotype information, and our analysis utilized by far the largest number of SNPs with confirmed associations with breast cancer, including all SNPs discovered to date. Further refinement of the risk stratification should be possible through incorporating additional SNPs exhibiting evidence for association, but not at formal genome-wide
Figure 2. Association between the polygenic risk score (PRS) and breast cancer risk in women of European origin for (A) all breast cancers, (B) estrogen receptor (ER)-positive disease, and (C) ER-negative disease. Odds ratios are for different percentiles of the PRS relative to the middle quintile (40% to 60%) of the PRS. Odds ratios and 95% confidence intervals are shown. Regular lines denote the observed estimates, and dotted lines the theoretical estimates under a multiplicative polygenic model with a standard deviation of the PRS of 0.45 for all breast cancer, 0.50 for ER-positive breast cancer, and 0.38 for ER-negative breast cancer, as derived from the estimated effect sizes and allele frequencies/haplotype frequencies for each locus. PRS = polygenic risk score.
statistical significance, together with variants in genes conferring intermediate or high risk.

The risk discrimination provided by the genetic profile, summarised in the PRS and family history, should be further improved by combining, with lifestyle risk factors, benign breast disease, and mammographic density. Although we did not consider lifestyle factors explicitly in this dataset, other large studies have found no good evidence for interactions between common susceptibility SNPs and lifestyle factors for breast cancer, suggesting that SNPs generally combined multiplicatively. Darabi et al. estimated a C-statistic of 0.60 for lifestyle risk factors including mammographic density. By comparison, we estimated the C-statistic for the PRS to be 0.62. Assuming that the multiplicative model is correct, the C-statistic would increase to 0.66 with the addition of the lifestyle risk factors. If modifiable risk factors and the PRS act multiplicatively, targeting public health interventions to women at higher genetic risk should result in a larger absolute risk reduction. For example, the decision to prescribe hormone replacement therapy might be guided by the PRS (available online), considering other risk factors, will be required to validate the overall accuracy of risk prediction. Such a comprehensive risk prediction model will need to be explicitly tested for in future studies.

In previous reports, improvement in risk discrimination by genomic profiling over that conferred by known risk factors was not substantial, although better discrimination was obtained for certain subgroups of women. Previous analyses, however, were based on a much smaller set of SNPs than included in this report. This study provides precise empirical estimates of the combined effects of multiple SNPs and the level of risk stratification possible. These estimates may inform the debate on public health utility and implementation of the PRS in clinical practice. Our work suggests that the PRS, particularly when used in combination with other risk factors, could help identify subsets of women at different levels of risk, for whom management would differ. The PRS may facilitate early detection of cancers in younger women and, importantly, identify individuals at risk of specific subtypes of breast cancer. Finally, there is potential for a stronger impact in modifying environmental factors will need to be explicitly tested for in future studies.

Figure 3. Cumulative and 10-year absolute risks of developing breast cancer for women of European origin by percentiles of the polygenic risk score (PRS). Cumulative absolute risk of developing breast cancer for (A) all breast cancers, (B) estrogen receptor (ER)–positive disease, and (C) ER-negative disease by percentiles of the PRS, and 10-year absolute risk of developing breast cancer for (D) all breast cancers, (E) ER-positive disease, and (F) ER-negative disease. Note different scales and PRS categories in the different panels. The red line shows the 2.4% risk threshold corresponding to the risk for women age 47 years who were eligible for screening, calculated as described in the Supplementary Methods (available online). Absolute risks were calculated using the PRS relative risks estimated as described in the Supplementary Methods (available online), and breast cancer incident rates and mortality from other causes obtained from the UK National Office for Statistics. For subtype-specific disease, the absolute risk for women in a particular PRS category for ER-positive disease and another PRS category for ER-negative disease were calculated. Information on proportions of tumors by ER status was obtained from the West Midlands Registry.
algorithm could provide a powerful basis for stratified breast cancer prevention programs.

**Funding**

This work was supported by Cancer Research-UK (grant numbers C1287/A10118, C1287/A12014) and the European Community’s Seventh Framework Programme (223175 [HEALTH-F2-2009-223175]) (COGS). Genotyping of the iCOGS array was funded by the European Union (HEALTH-F2-2009-223175), Cancer Research UK (C1287/A10710), the Canadian Institutes of Health Research (CIHR) for the “CIHR Team in Familial Risks of Breast Cancer” program, and the Ministry of Economic Development, Innovation and Export Trade of Quebec (PSR-SIIRI-701). This work was also supported by Breakthrough Breast Cancer funding (to MGC). Analysis was supported in part by the National Institutes of Health Post-Genome Wide Association initiative (1U19CA148065 (DRIVE) and 1U19CA148537 (ELLIPSE)). Laboratory infrastructure was funded by Cancer Research UK (C8197/A10123). This work was also supported by the Government of Canada through Genome Canada and the Canadian Institutes of Health Research and the Ministère de l’enseignement supérieur, de la recherche, de la science et de la technologie du Québec through Génome Québec for the PERSPECTIVE project. Breast Cancer Association Consortium meetings were funded by the European Union European Cooperation in Science and Technology (COST) programme (BM0606).

The Australian Breast Cancer Family Study (ABCFS), Northern California Breast Cancer Family Registry (NC-BCFR) and Ontario Familial Breast Cancer Registry (OFBCR) studies were supported by the US National Cancer Institute (UM1 CA164920). The ABCFS was also supported by the National Health and Medical Research Council of Australia, the New South Wales Cancer Council, the Victorian Health Promotion Foundation (Australia), and the Victorian Breast Cancer Research Consortium. JLH is a National Health and Medical Research Council (NHMRC) Australia Fellow and a Victorian Breast Cancer Research Consortium Group Leader. MCS is an NHMRC Senior Research Fellow and a Victorian Breast Cancer Research Consortium Group Leader. JLH and MCS are both group leaders of the Victoria Breast Cancer Research Consortium. The content of this manuscript does not necessarily reflect the views or policies of the US National Cancer Institute or any of the collaborating centers in the Breast Cancer Family Registry (BCFR), nor does mention of trade names, commercial products,
or organizations imply endorsement by the US Government or the BCFR.

The Amsterdam Breast Cancer Study (ABCS) was supported by the Dutch Cancer Society (NKI 2007–3839, 2009 4363) and BBMRI-NL, which is a Research Infrastructure financed by the Dutch government (NWO 184-021-007).

The Australian Breast Cancer Tissue Bank (ABCTB) study was supported by the National Health and Medical Research Council of Australia, The Cancer Institute New South Wales and the National Breast Cancer Foundation.

The Bavarian Breast Cancer Cases and Controls (BBCC) study was partly funded by ELAN-Fond of the University Hospital of Erlangen.

The British Breast Cancer Study (BBCS) was funded by Cancer Research UK and Breakthrough Breast Cancer and acknowledges National Health Service funding to the National Institutes for Health Research Biomedical Research Centre, and the National Cancer Research Network (NCRN). The BBCS GWAS received funding from The Institut National de cancer.

The Breast Cancer In Galway Genetic Study (BIGGS) was supported by National Institutes for Health Research Comprehensive Biomedical Research Centre, Guy’s & St. Thomas’ NHS Foundation Trust in partnership with King’s College London, United Kingdom (ES), and the Oxford Biomedical Research Centre (IT).

The Breast Cancer Study of the University Clinic Heidelberg (BSUCH) was supported by the Dietmar-Hopp Foundation, the Helmholtz Society, and the German Cancer Research Center (DKFZ).

The CECILE Breast Cancer Study (CECILE) was funded by Fondation de France, Institut National du Cancer (INCa), Ligue Nationale contre le Cancer, Ligue contre le Cancer Grand Ouest, Agence Nationale de Sécurité Sanitaire (ANSES), Agence Nationale de la Recherche (ANR).

The Copenhagen General Population Study (CGPS) was supported by the Chief Physician Johan Boserup and Lise Boserup Fund, the Danish Medical Research Council, and Herlev Hospital.

The Spanish National Cancer Centre Breast Cancer Study (CNIO-BCS) was supported by the Genome Spain Foundation, the Red Temática de Investigación Cooperativa en Cáncer, and by grants from the Asociación Española Contra el Cáncer and the Red Temática de Investigación Cooperativa en Cáncer, and State task of the Ministry of Education (RUS08/017) and by the Ministry of Education and Science of the Russian Federation (number 14.574.21.0026), agreement dated June 17, 2014, a unique identifier agreement (RFMEFI57414X0026), the Russian Foundation for Basic Research (14-04-31169 mol.a) and State task of the Ministry of Education and Science of the Russian Federation (310-14).

The Karolinska Breast Cancer Study (KARBAC) was supported by the regional agreement on medical training and clinical research (ALF) between Stockholm County Council and Karolinska Institutet, the Swedish Cancer Society, the Gustav V. Jubilee foundation, and the Bert von Kantzows foundation.

The Kuopio Breast Cancer Project (KBCP) was supported by the special Government Funding (EVO) of Kuopio University Hospital grants, Cancer Fund of North Savo, the Finnish Cancer Organizations, the Academy of Finland, and by the strategic funding of the University of Eastern Finland.

The kConFab study was supported by the National Breast Cancer Foundation and previously by the National Health and Medical Research Council, the Queensland Cancer Fund, the Cancer Councils of New South Wales, Victoria, Tasmania, and South Australia, and the Cancer Foundation of Western Australia. The kConFab Clinical Follow Up Study was funded by the National Health and Medical Research Council (145684, 288704, 454508). RB was a Cancer Institute NSW Fellow.

The A0CS study was supported by the United States Army Medical Research and Materiel Command (DAMD17-01-0729), the Cancer Council of Tasmania, the Cancer Foundation of Western Australia and the National Health and Medical Research Council (199600), and a National Health and Medical Research Council grant to GCT.
The Leuven Multidisciplinary Breast Centre (LMBC) study is supported by the “Stichting tegen Kanker” (232–2008, 196–2010) and by the FWO and the KULPFPV/10/016-SymBioSysII to DL.

The Mammary Carcinoma Risk Factor Investigation (MARI) study was supported by the Deutsche Krebshilfe e.V. (70-2892-BR I, 106332, 108253, 108419), the Hamburg Cancer Society, the German Cancer Research Center (DKFZ), and the Federal Ministry of Education and Research (BMBF) Germany (01KH0402).

The Milan Breast Cancer Study Group (MBCSG) was supported by grants from the Italian Association for Cancer Research (AIRC) and by funds from the Italian citizens who allocated the 5/1000 share of their tax payment in support of the Fondazione IRCCS Istituto Nazionale Tumori, according to Italian laws (INT-Institutional strategic projects “sx1000”).

The Mayo Clinic Breast Cancer Study (MCBCS) was supported by the National Institutes of Health (CA122340, CA128978), a National Institutes of Health Specialized Program of Research Excellence (SPORE) in Breast Cancer (CA116201), the Breast Cancer Research Foundation, the Komen Race for the Cure, and by a generous gift from the David F. and Margaret T. Grohne Family Foundation and the Ting Tsung and Wei Fong Chao Foundation.

The Melbourne Collaborative Cohort Study (MCPCS) cohort recruitment was funded by VicHealth and Cancer Council Victoria. The MCPCS was further supported by Australian National Health and Medical Research Council (209057, 251553, 504711) and by infrastructure provided by Cancer Council Victoria.

The Multi-ethnic Cohort (MEC) was supported by the National Institutes of Health (CA63464, CA54281, CA098758, and CA132839).

The Memorial Sloan-Kettering Cancer Center (MSKCC) was supported by grants from the Breast Cancer Research Foundation and the Robert and Kate Niehaus Clinical Cancer Genetics Initiative.

The work of Montreal Gene-Environment Breast Cancer Study (MTLGBCS) was supported by the Quebec Breast Cancer Foundation, the Canadian Institutes of Health Research for the “CIHR Team in Familial Risks of Breast Cancer” program (CRN-87521), and the Ministry of Economic Development, Innovation and Export Trade (PSR-SIIIRI-701). MG received an Investigator Award from the CIHR and a Health Scholar Award from the Fonds de la recherche en santé du Québec. J Simard is chairholder of the Canada Research Chair in Oncogenetics.

The Norwegian Breast Cancer Study (NBCS) was supported by grants from the Norwegian Research council (155218/V40, 175240/S10 to ALBD, FUGE-NPR 181600/V11) to VNK and a Swizz Bridge Award to ALBD.

The Nashville Breast Health Study (NBHS) was supported by National Institutes of Health (grant R01CA100374).

The Nurses Health Study (NHS) was funded by National Institutes of Health (PO1 CA87969), and this project was (in part) supported by the Genetic Associations and Mechanisms in Oncology (GAME-ON) Network (U19 CA148065).

The Oulu Breast Cancer Study (OBCS) was supported by research grants from the Finnish Cancer Foundation, the Academy of Finland, the University of Oulu, and the Oulu University Hospital.

The Leiden University Medical Centre Breast Cancer Study (ORIGO) was supported by the Dutch Cancer Society (RUL 1997-1505) and the Biobanking and Biomolecular Resources Research Infrastructure (BBMRI-NL CP16).

The NCI Polish Breast Cancer Study (PBCS) was supported by the Intramural Research Programs of the Division of Cancer Epidemiology and Genetics and the Center for Cancer Research of the National Cancer Institute.

The Karolinska Mammography Project for Risk Prediction of Breast Cancer (pKARMA) study was supported by Märit and Hans Rausings Initiative Against Breast Cancer and Cancer Risk Prediction Center, a Linneus Centre (contract 70867902) financed by the Swedish Research Council.

The Prospective Study of Outcomes in Sporadic Versus Hereditary Breast Cancer (POSH) study was supported by Cancer Research UK (A7572, A11699, C22524).

The Rotterdam Breast Cancer Study (RBCS) was funded by the Dutch Cancer Society (DDHK 2004–3124, DDHK 2009–4318).

The Singapore and Sweden Breast Cancer Study (SASBAC) was supported by the Agency for Science, Technology and Research of Singapore (A*STAR), the National Institutes of Health, and the Susan G. Komen Breast Cancer Foundation.

The Sheffield Breast Cancer Study (SBCS) was supported by Yorkshire Cancer Research (S295, S299, S305PA) and the Sheffield Experimental Cancer Medicine Centre.

The Southern Community Cohort Study (SCCS) is funded by National Institutes of Health (R01 CA092447). The Arkansas Central Cancer Registry is fully funded by a grant from the National Program of Cancer Registries, Centers for Disease Control and Prevention (CDC).

Study of Epidemiology and Risk factors in Cancer Heredity (SEARCH) was funded by a programme grant from Cancer Research UK (C490/A10124), the UK National Institute for Health Research Biomedical Research Centre at the University of Cambridge, and a Cancer Research UK grant (C8197/A10123) to AMD.

The Städtisches Klinikum Karlsruhe Deutsches Krebsforschungszentrum Study (SKKDKFZS) was supported by the Deutsches Krebsforschungszentrum (DKFZ).

The IHCC-Szczecin Breast Cancer Study (SZBCS) was supported by grant (PBZ_KBN_122/PO5/2004).

The Triple Negative Breast Cancer Consortium Study (TNBCC) was supported by the National Institutes of Health (CA128978) and the National Institutes of Health Specialized Program of Research Excellence in Breast Cancer (CA116201), the Breast Cancer Research Foundation, a generous gift from the David F. and Margaret T. Grohne Family Foundation and the Ting Tsung and Wei Fong Chao Foundation.

The UCIBCS was supported by the National Institutes of Health, National Cancer Institute grant CA-58860 and the Lon V Smith Foundation grant LVS-39420.
The UK Breakthrough Generations Study (UKBGS) was funded by Breakthrough Breast Cancer and the Institute of Cancer Research (ICR). ICR acknowledges NHS funding to the Royal Marsden Hospital/ICR National Institutes for Health Research Biomedical Research Centre.

The US Three State Study (US3SS) was supported by Massachusetts (R01CA47305 to KME), Wisconsin (R01 CA47147 to PAN), and New Hampshire (R01CA69664 to LTE) centers and Intramural Research Funds of the National Cancer Institute, Department of Health and Human Services.

The US Radiologic Technologists Study (USRRT) was funded by the Intramural Research Program of the Division of Cancer Epidemiology and Genetics, National Cancer Institute, National Institutes of Health, US Department of Health and Human Services.

Epidemiology Unit (WMCIU) for providing data on breast cancer incidence by ER status for 2010. As of 1st April 2013, WMCIU are conducting the study, data collection, and revising the manuscript. The remaining authors led individual studies and contributed to the design of the study, data collection, and revising the manuscript. The sponsors had no role in the study design, data collection, and revising the manuscript.

We extend our thanks to the many women who generously took part in these studies. We also thank all the researchers, nurses, clinicians, technicians, and administrative staff who have enabled this work to be carried out. In particular, we thank: Andrew Lee, Ed Dicks, and the staff of the Centre for Genetic Epidemiology Laboratory, the staff of the CNIO genotyping unit, Sylvie LaBoissière and Frederic Robidoux and the staff of the Copenhagen DNA laboratory, and Julie M. Cunningham, Sharon A. Windebank, Christopher A. Hilker, Jeffrey Meyer and the staff of Mayo Clinic Genotyping Core Facility.

The authors would also like to thank the West Midlands Cancer Intelligence Unit (WMCIU) for providing data on breast cancer incidence by ER status for 2010. As of 1st April 2013, WMCIU are part of Public Health England. The ABCFS study thanks Maggie Angelakos, Judi Maskiell, and Gillian Dite. The OFBCR study thanks Teresa Selander and Nayana Weerasooriya. The ABCS study acknowledges Sanquin Research. ABCTB Investigators include Christine Clarke, Rosemary Balleine, Robert Baxter, Stephen Braye, Jane Carpenter, Jane Dahlstrom, John Forbes, Soon Lee, Debbie Marsh, Adrienne Morey, Nirmala Pathmanathan, Rodney Scott, Allan Spigelman, Nicholas Wilcken, and Desmond Yip. The BBCS study thanks Sonja Oser, Silke Landrith, and Matthias Rübben. The BBCS study thanks Eileen Williams, Elaine Ryder-Mills, and Kara Sargus. The BIGGS study thanks Niall McInerney, Gabrielle Colleran, Andrew Rowan, and Angela Jones.

The BIGGS study thanks Niall McInerney, Gabrielle Colleran, Kara Sargus.

Notes

Author contributions: NM, DFE, and MG analyzed data relating to this manuscript and drafted the initial manuscript. DFE coordinated the BCAC and led the iCOGS genotyping, PH led the COGS collaboration. KM performed statistical analyses of the iCOGS data, with assistance from JT. JD provided bioinformatics support. MKB and QW coordinated the BCAC database. The remaining authors led individual studies and contributed to the design of the study, data collection, and revising the manuscript. The sponsors had no role in the study design, in the collection, analysis, or interpretation of data, in the writing of the report, or in the decision to submit the paper for publication.

We extend our thanks to the many women who generously took part in these studies. We also thank all the researchers, nurses, clinicians, technicians, and administrative staff who have enabled this work to be carried out. In particular, we thank: Andrew Lee, Ed Dicks, and the staff of the Centre for Genetic Epidemiology Laboratory, the staff of the CNIO genotyping unit, Sylvie LaBoissière and Frederic Robidoux and the staff of the McGill University and Génome Québec Innovation Centre, the staff of the Copenhagen DNA laboratory, and Julie M. Cunningham, Sharon A. Windebank, Christopher A. Hilker, Jeffrey Meyer and the staff of Mayo Clinic Genotyping Core Facility.

The authors would also like to thank the West Midlands Cancer Intelligence Unit (WMCIU) for providing data on breast cancer incidence by ER status for 2010. As of 1st April 2013, WMCIU are part of Public Health England. The ABCFS study thanks Maggie Angelakos, Judi Maskiell, and Gillian Dite. The OFBCR study thanks Teresa Selander and Nayana Weerasooriya. The ABCS study acknowledges Sanquin Research. ABCTB Investigators include Christine Clarke, Rosemary Balleine, Robert Baxter, Stephen Braye, Jane Carpenter, Jane Dahlstrom, John Forbes, Soon Lee, Debbie Marsh, Adrienne Morey, Nirmala Pathmanathan, Rodney Scott, Allan Spigelman, Nicholas Wilcken, and Desmond Yip. The BBCS study thanks Sonja Oser, Silke Landrith, and Matthias Rübben. The BBCS study thanks Eileen Williams, Elaine Ryder-Mills, and Kara Sargus. The BIGGS study thanks Niall McInerney, Gabrielle Colleran, Andrew Rowan, and Angela Jones.
of Oncology (Warsaw, Poland), and Pey Chao and Michael Stagner from Information Management Services (Sliver Spring, MD) for their valuable contributions to the study.

The RBCS study thanks Petra Bos, Jannet Blom, Ellen Crepin, Anja Nieuwlait, Annette Heemsker, the Erasmus MC Family Cancer Clinic.

The SBCS study thanks Sue Higham, Helen Cramp, and Dan Connely.

Data on SCCS breast cancer cases used in this publication were provided by the Alabama Statewide Cancer Registry, Kentucky Cancer Registry (Lexington, KY, Tennessee Department of Health (Office of Cancer Surveillance), Florida Cancer Data System, North Carolina Central Cancer Registry (North Carolina Division of Public Health), Georgia Comprehensive Cancer Registry, Louisiana Tumor Registry, Mississippi Cancer Registry, South Carolina Central Cancer Registry, Virginia Department of Health (Virginia Cancer Registry), Arkansas Department of Health (Cancer Registry). Data on SCCS cancer cases from Mississippi were collected by the Mississippi Cancer Registry, which participates in the National Program of Cancer Registries (NPCR) of the Centers for Disease Control and Prevention. The contents of this publication are solely the responsibility of the authors and do not necessarily represent the official views of the CDC or the Mississippi Cancer Registry.

The SEARCH study thanks the SEARCH and EPIC-Norfolk teams.

The TNBCC study thanks the Human Genetics Sample Bank for processing of samples and providing OSU Columbus area control samples.

Consortia Membership

Australian Ovarian Cancer Study Group

David D. Bowtell, Adele C. Green, Georgia Chenevix-Trench, Anna deFazio, Dorota Gertig, Penelope M. Webb.

The Breast and Ovarian Cancer Susceptibility Collaboration (BOCS)


GENICA Network

Dr. Margarete Fischer-Bosch-Institute of Clinical Pharmacology, Stuttgart, and University of Tübingen, Germany: HB, Wing-Yee Lo, Christina Justenhoven; German Cancer Consortium (DKTK) and German Cancer Research Center (DKFZ): HB: Department of Internal Medicine, Evangelische Kliniken Bonn gGmbH, Johanniter Krankenhaus, Bonn, Germany: YDK, Christian Baisch; Institute of Pathology, Medical Faculty University of Bonn, Germany: Hans-Peter Fischer; Molecular Genetics of Breast Cancer, Deutsches Krebsforschungszentrum (DKFZ), Heidelberg, Germany: UH; Institute for Prevention and Occupational Medicine of the German Social Accident Insurance, Institute of the Ruhr University Bochum (IPA), Bochum, Germany: TB, Beate Pesch, Sylvia Rabenstein, Anne Lotz; Institute of Occupational Medicine and Maritime Medicine, University Medical Center Hamburg-Eppendorf, Germany: Volker Harth.

HEBON

Netherlands Cancer Institute, Amsterdam: Senno Verhoef, Martijn Verheus, Laura J. van’t Veer, Flora E. van Leeuwen; Erasmus Medical Center, Rotterdam: Margriet Collée, Maartje J. Hooning, Madeleine, M. A. Tilanus-Linthorst, Caroline Seynae; Leiden University Medical Center, Leiden: Christi J. van Aspenen, Juul T. Wijnen, Rob A. Tollenaar; Radboud University Nijmegen Medical Center, Nijmegen: Marjolijn J. Lijtenberg; University Medical Center Utrecht, Utrecht: Margreet G. Ausems; Amsterdam Medical Center: Cora M. Aalfs, Theo A. van Os; VU University Medical Center, Amsterdam: Johan J. P. Gille, Quينten Waisfisz; University Hospital Maastricht, Maastricht: Cees E. van Roorzaanda, Marinus J. Blok; University Medical Center Groningen University: Jan C. Oosterwijk, Annemarie H van der Hou, Marjan J. M. Mouris; the Netherlands Foundation for the detection of hereditary tumours, Leiden: Hans F. Vase.

kConFab*

Alison Trainer, Kathy Tucker, Janet Tyler, Jane Visvader, Logan Walker, Paul Waring, Robin Ward, By Warnier, Rachael Williams, Ingrid Winship, Mary Ann Young. *Peter MacCallum Cancer Center, Melbourne, Australia.

References