An empirical comparison of meta-analyses of published gene-disease associations versus consortium analyses

*A. Cecile J. W. Janssens, PhD[1], Angela M. González-Zuloeta Ladd, MD, PhD[2],
Sandra López-Léon, MD, PhD[2], John P. A. Ioannidis, MD, PhD[3,4], Ben A. Oostra, PhD[2],
Muin J. Khoury, MD, PhD[5], and Cornelia M. van Duijn, PhD[2]*

Purpose: Consortia of investigators currently compile sufficiently large sample sizes to investigate the effects of low-risk susceptibility genetic variants. It is not clear how the results obtained by consortia compare with those derived from meta-analyses of published studies. **Methods:** We performed meta-analyses of published data for 16 genetic polymorphisms investigated by the Breast Cancer Association Consortium, and compared sample sizes, heterogeneity, and effect sizes. PubMed, Web of Science, and Human Genome Epidemiology Network databases were searched for breast cancer case-control association studies. **Results:** We found that meta-analyses of published data and consortium analyses were based on substantially different data. Published data by nonconsortium teams amounted on average to 26.9% of all available data (range 3.0 –50.0%). Both approaches showed statistically significant decreased breast cancer risks for *CASP8 D302H*. The meta-analyses of published data demonstrated statistically significant results for five other genes and the consortium analyses for two other genes, but the strength of this evidence, evaluated on the basis of the Venice criteria, was not strong. **Conclusions:** Because both approaches identified the same gene out of 16 candidates, the methods can be complimentary. The expense and complexity of consortium-based studies should be considered vis-à-vis the potential methodological limitations of synthesis of published studies. *Genet Med* 2009:11(3):153–162.

Key Words: *breast neoplasm, genetics, association, epidemiologic methods, meta-analysis, bias, collaborations, methods*

In recent years, many large-scale consortia have been initiated to identify susceptibility genes of complex disorders.[1–6] There are several perceived advantages to the consortium approach for studying genetic associations. Consortia often lead to larger sample sizes than do meta-analysis of published studies and thus to sufficient statistical power to demonstrate significant effects of weak susceptibility genes. They provide access to unpublished "negative" data, which reduces problems of publication or selective reporting bias.[7] Moreover, they can facilitate the harmonization of criteria for the selection of cases and controls and the standardization of genotype technology, which can reduce between-study heterogeneity in effect estimates by eliminating methodological differences between studies. Although these are straightforward arguments supporting the building of large consortia,[6,8] it remains to be investigated whether the results of analyses from consortia yield the acclaimed advantages over meta-analyses of available published data and whether there are differences in the results obtained by these two approaches.

To empirically investigate differences between the analyses of consortia and meta-analyses of published data, we performed meta-analyses of published data for 16 genetic polymorphisms that were investigated by the Breast Cancer Association Consortium (BCAC).[9,10] The consortium reported genetic association analyses on 16 genetic polymorphisms, including individual-level published and unpublished data contributed by up to 20 research groups. We compared the amount of collected data and the results between the two approaches and examined the presence of potential biases in the meta-analyses of published studies. We also applied the Venice criteria, a set of criteria recently developed by a consensus of the Human Genome Epidemiology Network for grading the epidemiological strength of cumulative evidence on genetic associations.[11] We wanted to determine whether the inferences derived with the two approaches were similar or divergent.

METHODS

Data sources

PubMed, Web of Science, and Human Genome Epidemiology Network literature databases were searched for breast cancer case-control studies on 16 polymorphisms investigated by the BCAC and published in the *Journal of the National Cancer Institute* and *Nature Genetics* (Table 1).[9,10] For 12 polymorphisms, the consortium collected, cleaned, and synthesized the available data from the participating teams and found no evidence for any significant association.[9] For another four polymorphisms, which had a *P* value of <0.10 upon the collection and synthesis of available data from consortium-participating teams (*CASP8 D302H, IGFBP3 C(-202)A, SOD2 V16A, TGFB1 L10P*), samples were further genotyped in other studies of the consortium.[10] For these four polymorphisms, we considered the results from the more comprehensive latter article.

The search strategy was based on the keywords *breast cancer* combined with the name of the gene without specifying the polymorphism. In addition, we searched for other articles on breast cancer written by the researchers of the consortium as listed in the appendix of their article.[9] Reference lists of all retrieved publications were screened for additional studies. Finally, previously published meta-analyses on these associations

Table 1 Meta-analyses of published data for polymorphisms investigated by the Breast Cancer Association Consortium

Gene	Variant	No. studies	No. controls	No. cases	Heterozygotes			Homozygotes			Per allele			References
					OR	95% CI	I^2	OR	95% CI	I^2	OR	95% CI	I^2	
ADHC1	*I350V*	4	2390	2130	1.00	0.87, 1.15	14	**0.83**	**0.69, 1.00**	0	0.94	0.86, 1.02	0	47–50
AURKA	*F31I*	9	9011	7294	1.04	0.96, 1.13	0	**1.28**	**1.06, 1.54**	46[a]	**1.10**	**1.01, 1.19**	58[a]	12, 51–58
BRCA2	*N372H*	8	14,387	14,065	1.04	0.99, 1.09	0	1.04	0.93, 1.16	19	1.03	0.99, 1.06	0	13, 59–64
CASP8	*D302H*	3	3591	3288	**0.85**	**0.76, 0.95**	0	**0.62**	**0.42, 0.89**	0	**0.83**	**0.75, 0.92**	0	65, 66
ERCC2	*D312N*	12	7821	9414	0.98	0.86, 1.11	63[b]	0.84	0.68, 1.05	70[c]	0.94	0.84, 1.05	77[c]	67–76
IGFBP3	*C(-202)A*	9	12,294	9774	1.01	0.94, 1.09	13	1.00	0.90, 1.11	32	1.01	0.96, 1.06	30	77–82
LIG4	*D501D T/C*	3	4113	4520	0.95	0.87, 1.05	0	0.90	0.59, 1.36	59[a]	0.95	0.87, 1.04	9	13, 83, 84
PGR	*V660L*	9	11,646	10,652	1.04	0.95, 1.13	33	1.03	0.73, 1.46	56[a]	1.02	0.92, 1.13	60[b]	14, 15, 85–91
SOD2	*V16A*	12	11,141	9991	1.03	0.95, 1.10	11	1.04	0.92, 1.17	40[a]	1.01	0.96, 1.07	28	28, 29, 31, 92–100
TGFB1	*L10P*	17	16,308	9331	1.02	0.93, 1.12	41[a]	0.98	0.86, 1.12	46[a]	1.00	0.94, 1.06	46[a]	101–113
TP53	*R72P*	14	8218	7569	1.03	0.91, 1.17	61[b]	1.04	0.87, 1.25	48[a]	1.03	0.94, 1.13	63[c]	61, 64, 114–125
XRCC1	*R399Q*	21	14,479	13,320	1.05	0.98, 1.12	21	1.07	0.96, 1.19	32[a]	1.04	0.99, 1.10	47[a]	16, 68, 69, 126–142
XRCC2	*R188H*	7	9723	10,427	0.98	0.89, 1.08	24	1.03	0.62, 1.70	40	1.00	0.89, 1.11	47[a]	13, 30, 63, 83, 143, 144
XRCC3	*5′UTR A/G*	4	6563	6303	**1.12**	**1.00, 1.24**	48	0.96	0.81, 1.15	0	1.06	0.98, 1.14	26	13, 83, 84
XRCC3	*IVS7-14*	4	6682	6270	0.92	0.82, 1.03	58[a]	**0.87**	**0.75, 1.00**	33	0.93	0.85, 1.02	64[a]	13, 83, 84
XRCC3	*T241M*	15	13,370	14,255	1.01	0.95, 1.07	7	**1.16**	**1.04, 1.28**	30	**1.06**	**1.01, 1.12**	39[a]	13, 16, 18, 30, 63, 69, 83, 128, 129, 131, 140, 144, 145

[a]$P < 0.10$ for the Q statistic
[b]$P < 0.01$ for the Q statistic
[c]$P < 0.001$ for the Q statistic. Nominally statistically significant associations ($P < 0.05$) are shown in bold type.
OR, odds ratio; CI, confidence interval; I^2, between-study heterogeneity (%).

(including meta-analyses on *AURKA F31I*,[12] *BRCA2 N372H*,[13] *PGR V660L*,[14,15] *XRCC1 R399Q*,[16] and *XRCC3 T241M*[13,17,18]) were also scrutinized to ensure that we had missed no studies. These meta-analyses were updated and reanalyzed with data from the primary studies. Searches were updated until July 2007.

Study selection

Genetic association studies were selected if they compared female breast cancer patients with controls from the general population in a case-control design and were reported in English. Studies were excluded when (1) the data were reused in a larger study on the same polymorphism; (2) genotype distributions in controls showed nominally statistically significant ($P < 0.05$) violations of the Hardy-Weinberg equilibrium (HWE); and (3) the article had incomplete reporting of genotype frequencies. HWE was recalculated from the original data using the χ^2 test. HWE testing is known to have low power, even with considerable sample sizes,[19] therefore using lower P-value thresholds would miss a lot of considerable deviations from equilibrium. Two investigators (A.M.G-.Z.L. and S.L.-L.) performed the study selection and information extraction independently in duplicate, and discrepancies were discussed with a third researcher (A.C.J.W.J.).

Data synthesis

Summary odds ratios (ORs) and 95% confidence intervals (CIs) were calculated with random-effects models. Summary ORs under the random-effects models were estimated for three

genetic models (homozygotes, heterozygotes, and per-allele comparisons), using the method of DerSimonian and Laird.[20] In the absence of between-study heterogeneity, the random and fixed-effects models coincide, whereas in the presence of heterogeneity, the basic assumption of fixed effects is violated. For this reason, we gave precedence to random-effects estimates. The statistical significance of the between-study heterogeneity was evaluated with the χ^2-based Q statistic that is considered significant for $P < 0.10$. The degree of heterogeneity between the study results was assessed by the I^2 statistic,[21] which measures the amount of inconsistency between studies that is beyond chance, taking values from 0 to 100%. Values of <25% suggest little heterogeneity, 25–50% suggests moderate heterogeneity, and >50% means large heterogeneity. In the presence of few studies, both the Q and I^2 have large uncertainty and thus should be evaluated with caution.[22]

Comparison of meta-analyses of published data versus consortium analyses

The two approaches were compared with regard to the amount of available data (number of cases and controls combined) and the results obtained. Meta-analyses were performed according to the three genetic models. All comparisons of the published versus the consortium analyses were based on comparison of the same genetic models.

We estimated Spearman correlation coefficients of the ORs and recorded whether the two ORs differed by more than

1.10-fold in any meta-analysis. Similarly, we estimated Spearman correlation coefficients for the I^2 estimates of the two approaches. We also report on whether there was discrepancy in the presence of nominal statistical significance ($P < 0.05$) with the two approaches for any meta-analyses. Such discrepancies should be interpreted cautiously, because a difference in the level of statistical significance does not mean that the difference is beyond chance. This has been pointed out repeatedly in meta-epidemiological research.[23,24] Moreover, we did not formally estimate the statistical significance of the difference in the effect sizes with the two approaches, because a large amount of the data are shared by the published and the consortium datasets, thereby overtly violating the assumption of independence.

Bias diagnostics

We also performed a number of tests to assess whether there is demonstrable potential for bias in the meta-analyses of published data that yielded nominally statistically significant results ($P < 0.05$) by random-effects calculations and that also did not have large between-study heterogeneity. First, we evaluated whether the results lost their nominal significance when the first-published study was excluded.[25] Second, we evaluated whether small studies yielded larger estimates of genetic effects than larger studies. When this is the case, it could signal publication bias, other selective reporting biases, or other biases, or it could be due to genuine heterogeneity between small and large studies. We used a modified regression test as proposed by Harbord et al.[26] that has an appropriate Type I error at $P = 0.10$. Last, we applied a test that determines whether the number of studies with nominally statistically significant results exceeds the expected number of such "positive" studies under plausible assumptions for the effect size in each meta-analysis.[27] This is an indication that there is bias in favor of publishing studies that show formally statistically significant results. The test is considered to be significant for $P < 0.10$. Data were analyzed using Review Manager, version 4.2 (Cochrane Collaboration, Oxford, UK) and Intercooled STATA 8.2 (College Station, TX).

Grading the strength of the cumulative epidemiological evidence

For all meta-analyses, we applied a grading system, the Venice criteria, for the strength of the epidemiological evidence. Details are published elsewhere.[11] Briefly, each meta-analyzed association was graded on the basis of amount of evidence, consistency of replication, and protection from bias. For amount of evidence, the grade was A when the total number of minor alleles of cases and controls combined in the meta-analyses exceeded 1000, B when it was between 100 and 1000, and C when it was <100. For consistency of replication, point estimates of I^2 exceeding 50% received Grade C, I^2 of 25–50% received Grade B, and I^2 <25% received Grade A when the result of the meta-analysis was nominally statistically significant; nonsignificant meta-analyses always received C for this criterion.[11]

For protection from bias, the guidelines propose to consider potential sources of bias at the level of individual studies, including errors in phenotypes, genotypes, and confounding (population stratification), and at the level of meta-analysis, including publication and other selective reporting biases.[11] Grade A is given when there is probably no bias that can affect the presence of the association, B when there is no demonstrable bias but important information is missing for its appraisal, and C when there is demonstrable potential or clear bias that can

invalidate the association. Protection from bias was investigated only for those meta-analyses that had not already received C grades for amount of data or replication consistency. We considered that meta-analyses based on consortium data were adequately protected from bias since publication and selective reporting are not a concern, and we considered that efforts made by the consortium should also considerably alleviate concerns for other major biases, even if bias cannot be totally excluded; thus, the grade of consortium meta-analyses was A in this regard. Meta-analyses of published analyses received Grade A when the OR deviated more than 1.15-fold from the null (>1.15 or <0.87) and no evidence for bias was demonstrated; a C grade was given when the summary genetic effect was of smaller magnitude (such small effects may easily be generated by even modest selective reporting biases or other biases) or other signs of bias were present, as described above.[11]

As suggested by the Venice criteria,[11] epidemiological evidence for significant association was rated as strong if the meta-analysis received three A grades, moderate if it received any B grade but not any C grade; and weak if it received any C grade in any of the three criteria.

RESULTS

Meta-analyses of published data

Of the 115 potentially eligible publications, five were excluded because they had incomplete reporting of genotype frequencies and one was excluded because the genotyping was performed on tumor tissue. The remaining 109 publications comprised 168 data sets addressing the 16 polymorphisms. Four datasets were excluded because the data had been reused in a larger study on the same polymorphism, two were excluded because the gene was not polymorphic in the specific data set, and 11 were excluded because the distributions of genotypes in controls violated HWE. None of the datasets were included in the analyses of the consortium. A total of 151 sets of data were thus considered in our 16 meta-analyses.

The meta-analyses showed nominally statistically significant lower risk of breast cancer for *CASP8 302H* carriers, the *VV* genotype of *ADHC1 I350V*, and the *GG* genotype of *XRCC3 IVS7-14* (Table 1). A nominally statistically significant higher risk of breast cancer was found for homozygous carriers of the *AURKA F31I* and *XRCC3 T241M* minor alleles and for heterozygous carriers of the *XRCC3 5'UTR G* allele. (For the respective results of the consortium analyses, see Table 4 in Ref. 9 and Table 1 in Ref. 10.)

Comparisons between meta-analyses and BCAC analyses

Included studies and sample size

The mean number of patients and controls was 18,289 (range 4,520–28,452) in the meta-analyses of published data and 23,140 (range 12,013–37,633) in the consortium analyses (Fig. 1A; Fig. 1B shows the respective data limited to breast cancer cases only). Figure 2 shows that, by and large, the subjects in the meta-analyses were different from those in the consortium analyses. The consortium had access to unpublished data of, on average, 13,797 persons per polymorphism (range 2,213–30,774), which amounted to 42.7% of all available data (range 9.0–76.2%). Data were most often unpublished for *ADH1C I350V* (76.2% of all available data were unpublished) and most often published for *PGR V660L* (9.0% unpublished). On average, the consortium did not use data of 8,945 cases and controls (range 1,031–19,177) that were published by others. The latter
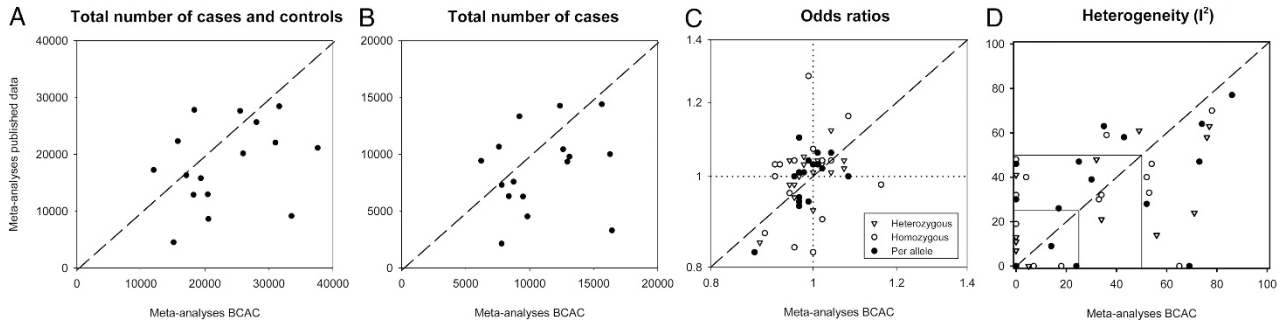
Fig. 1. Total number of cases and controls, odds ratios, and heterogeneity of meta-analyses of published data compared with meta-analyses of the Breast Cancer Association Consortium.[9] Summary odds ratios of the consortium were obtained from random-effects meta-analyses.
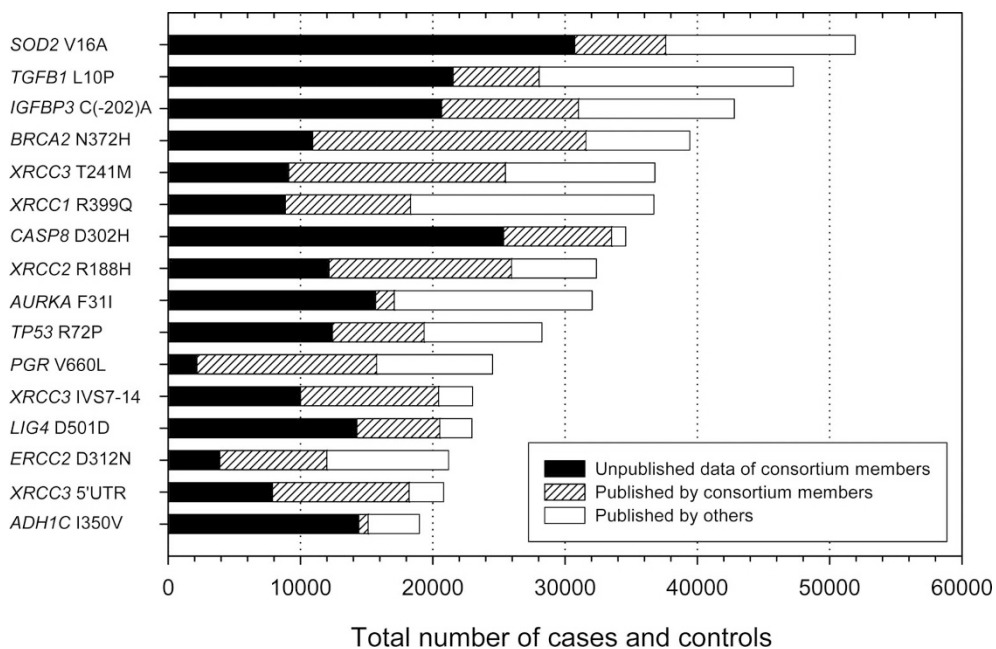

Fig. 2. Total number of cases and controls included in the meta-analyses of published data and in the analyses of the Breast Cancer Association Consortium.[9]

comprised 47.8% of all published data (range 11.3–91.5%) and 26.9% of all available data (range 3.0–50.0%). For two polymorphisms (*AURKA F31I* and *XRCC1 R399Q*), almost half of all available data were published by research groups that were not part of the consortium. For example, for *XRCC1 R399Q*, the combined studies of the consortium included 18,339 cases and controls, whereas an additional 18,354 were available from other studies.

Estimates of effect

Differences between the ORs of the meta-analyses and the consortium analyses are presented in Figure 1C. Eleven of the meta-analyses of published data and eight of the consortium analyses reached nominal statistical significance. The three meta-analyses for *CASP8* were the only ones that reached statistical significance in both approaches. For 39 of the 48 comparisons, the ORs of the meta-analyses of published data were <10% higher or lower than the ORs of the consortium analyses. The largest differences were found for the analyses of

homozygous carriers of the *TGFB1 P* alleles (OR 1.16 [95% CI: 1.08–1.25] in the consortium analysis vs. OR 0.98 [95% CI: 0.86–1.12] in the meta-analysis), *ADH1C V* alleles (OR 1.00 [95% CI: 0.83–1.21] vs. OR 0.83 [95% CI: 0.69–1.00]), and *AURKA I* alleles (OR 0.99 [95% CI: 0.78–1.27] vs. OR 1.28 [95% CI: 1.06–1.54]). However, in all of these analyses, the 95% CIs overlapped. The OR of *AURKA* homozygotes was not statistically significant when the analysis was restricted to whites (OR 1.30 [95% CI: 0.98–1.72]). The Spearman correlation coefficient for the OR estimates between the two approaches was 0.44 ($P = 0.002$).

Heterogeneity

Of the 48 meta-analyses of published data, 18 (37.5%) showed no or low heterogeneity ($I^2 \leq 25\%$), 19 (39.6%) showed moderate heterogeneity, and 11 (22.9%) had large estimated between-study heterogeneity (Fig. 1D). The latter 11 meta-analyses concerned six genes. Figure 1D compares I^2 estimates of meta-analyses of published data with that available

Table 2 Grading of epidemiological evidence of association using the criteria of the Human Genome Epidemiology Network (HuGENet)[11]

Gene	Variant	Heterozygotes			Homozygotes			Per allele		
		Amount of data	Replication consistency	Protection from bias	Amount of data	Replication consistency	Protection from bias	Amount of data	Replication consistency	Protection from bias
ADHC1	I350V	A (A)	C (C)	– (–)	B (A)	A (C)	C (–)	A (A)	C (C)	– (–)
AURKA	F31I	A (A)	C (C)	– (–)	A (B)	B (C)	A (–)	A (A)	C[a] (C)	– (–)
BRCA2	N372H	A (A)	C (C)	– (–)	A (A)	C (C)	– (–)	A (A)	C (C)	– (–)
CASP8	D302H	A (A)	A (A)	A (A)	B (B)	A (A)	A (A)	A (A)	A (A)	A (A)
ERCC2	D312N	A (A)	C (C)	– (–)	A (A)	C (C)	– (–)	A (A)	C (C)	– (–)
IGFBP3	C(-202)A	A (A)	C (C)	– (–)	A (A)	C (A)	– (A)	A (A)	C (A)	– (A)
LIG4	D501D T/C	A (A)	C (C)	– (–)	B (B)	C (C)	– (–)	A (A)	C (C)	– (–)
PGR	V660L	A (A)	C (C)	– (–)	B (B)	C (C)	– (–)	A (A)	C (C)	– (–)
SOD2	V16A	A (A)	C (C)	– (–)	A (A)	C (C)	– (–)	A (A)	C (C)	– (–)
TGFB1	L10P	A (A)	C (A)	– (A)	A (A)	C (A)	– (A)	A (A)	C (A)	– (A)
TP53	R72P	A (A)	C (C)	– (–)	A (A)	C (C)	– (–)	A (A)	C (C)	– (–)
XRCC1	R399Q	A (A)	C (C)	– (–)	A (A)	C (C)	– (–)	A (A)	C (C)	– (–)
XRCC2	R188H	A (A)	C (C)	– (–)	B (B)	C (C)	– (–)	A (A)	C (C)	– (–)
XRCC3	5′ UTR A/G	A (A)	B (C)	C (–)	B (B)	C (C)	– (–)	A (A)	C (C)	– (–)
XRCC3	IVS7-14	A (A)	C (C)	– (–)	A (A)	B (C)	C (–)	A (A)	C (C)	– (–)
XRCC3	T241M	A (A)	C (C)	– (–)	A (A)	B (C)	C (–)	A (A)	B (C)	C (–)

Grade of the evidence from meta-analysis of published data (grade of the evidence from consortium analysis). See text for interpretation of the grades. A grade of C for replication consistency was given because of no association, except for the AURKA per-allele analyses, which had high heterogeneity. A grade of C for protection from bias was given because the association was no longer nominally statistically significant after exclusion of the first study.

for the consortium analyses. The Spearman correlation coefficient for I^2 estimates was 0.47 ($P = 0.001$). Large estimates of between-study heterogeneity was observed in 14 (31%) of the 45 meta-analyses of the consortium (I^2 statistics could not be calculated for the *PGR* analyses of the consortium because the data were not provided). In nine of the 45 meta-analyses, the consortium had large estimates of between-study heterogeneity where this was not seen in the meta-analysis of published data, whereas the opposite situation was observed in four meta-analyses.

Bias diagnostics

The consortium checked for *CASP8* and *TGFB1* whether the first positive study could have contributed to the statistically significant finding and found that the results remained the same.[10] Exclusion of the first-published study resulted in loss of the nominal significance in five of the nine meta-analyses of published data that had nominally statistically significant results. For the remaining four meta-analyses (comparison of homozygotes for *AURKA* and the three *CASP8* meta-analyses), the modified regression test showed no significant difference between small and larger studies and the excess test showed no statistically significant excess of "positive" studies.

Grading of cumulative epidemiological evidence

Table 2 shows the grading for amount of data, replication consistency, and protection from bias in each of the meta-analyses of published data versus the respective consortium meta-analyses. Forty-four of the 48 meta-analyses of published

data and 40 of the 48 consortium analyses had at least one C grade and were therefore considered "weak" evidence for association. Eight meta-analyses of the consortium were evaluated as showing "strong" evidence for association (homozygous and per-allele meta-analyses of *IGFBP3* and all meta-analyses for *CASP8* and *TGFB1*); of those, only the three meta-analyses of *CASP8* were also deemed to show "strong" evidence in the analyses of published data. Conversely, one meta-analysis of published data (homozygotes comparison for *AURKA*) was deemed to show "moderate" evidence but was "weak" in the consortium data.

Combining meta-analyses of published data with consortium analyses

The association between *CASP8* and breast cancer risk was still statistically significant when all available data were combined. The OR for homozygous carriers of the rare allele was 0.74 (95% CI: 0.61–0.89, I^2 12%), for heterozygous carriers was 0.89 (95% CI: 0.85–0.94, I^2 0%), and for the per-allele effect was 0.88 (95% CI: 0.84–0.92, I^2 0%). The meta-analysis of *TGFB1* heterozygotes yielded a nominally statistically significant effect (OR 1.05, 95% CI: 1.00–1.11, I^2 22%), but the homozygotes and per-allele meta-analyses were no longer statistically significant (homozygotes OR 1.05, 95% CI: 0.96–1.15, I^2 43% and per-allele OR 1.03, 95% CI: 0.99–1.07, I^2 40%). In addition, the *IGFBP3* meta-analyses did not reach nominal statistical significance when the consortium data were combined with data published by nonmembers. The meta-anal-

yses for the *AURKA* homozygotes yielded borderline significance with large between-study heterogeneity (OR 1.15, 95% CI: 0.99–1.34, I^2 52%). None of the other meta-analyses were statistically significant.

DISCUSSION

Our meta-analyses of published studies and the consortium analyses both identified *CASP8* as a breast cancer susceptibility gene and found insufficient evidence that the other genes are associated with risk of breast cancer. Some discrepancies did occur, however, especially for whether nominal statistical significance is claimed for specific comparisons. Summary effect sizes tended to correlate, between-study heterogeneity was not larger or smaller on average with either of the two approaches, and final inferences on the grading of the cumulative evidence agreed in the large majority of the meta-analyses. Exceptions to these statements did exist, and in these cases, the consideration of all data (both from consortia and other published evidence) may be most useful.

Before interpreting the results of this study, one methodological issue needs to be addressed. We should acknowledge that we are using statistical significance thresholds that are lenient, as traditionally used for candidate gene variants, but it is possible that even for candidate gene variants much lower *P* values may be needed to reach high credibility levels. Moreover, we adopt an approach where we examine statistical significance separately for each particular OR (i.e., separately for heterozygotes and for homozygotes), rather than global testing of the null across all genotypes. Although this approach is suitable for comparing the results of the meta-analyses and the consortium analyses, it is not the most conservative for the identification of susceptibility variants with high credibility.

The differences between the two approaches that we found may be explained by the fact that the meta-analyses and the consortium analyses, by and large, were based on different data collected in different populations. Research groups that did not participate in the consortium contributed about half of all published data and more than a quarter of all available data. Large-scale studies, such as the Shanghai Breast Cancer Study ($n = 1193$),[28] the Carolina Breast Cancer Study ($n = 2045$),[29] a nested case-control study within the Nurses Health Study ($n = 1004$),[30] and the Long Island Breast Cancer Study Project ($n = 1052$),[31] as well as many smaller studies with approximately 500 patients each, were not participating in the consortium. For example, the *AURKA F31I* meta-analyses had only one overlapping study in the consortium versus published data.[12] Although this study showed a decreased breast cancer risk for carriers of the I allele, all eight other published studies found increased risks, yielding a nominally significant per-allele summary OR of 1.10 (95% CI: 1.01–1.19). For *ADH1C I350V*, all four published studies showed nonsignificant lower breast cancer risks for carriers of the V allele, whereas the consortium included two unpublished studies that demonstrated increased risks, the largest of which (4317 patients) was statistically significant. Given the diversity of studies included by the consortium, there is no methodological reason why these published studies were not included.

The meta-analyses of published data and the BCAC analyses led to different conclusions about the association with breast cancer risk for seven of the 16 polymorphisms (12 of 48 meta-analyses) when based on nominal significance testing. However, the differences in the summary ORs were generally minor. When ORs are in the order of 1.1, a minor difference can make the difference between a finding being statistically signif-

icant or not. However, with proper attention to issues of heterogeneity and protection from bias, the final inferences may end up being similar with both approaches. On the other hand, interpreting meta-analyses without attention to these issues and looking only at the statistical significance may lead to erroneous inferences. Substantial differences in breast cancer risk were found in the comparison of homozygous carriers of the *TGFB1*, *ADH1C*, and *AURKA* alleles, but even in these, the 95% CIs overlapped with the two approaches. For four genes, meta-analyses of published data (*ADH1C* and the three polymorphisms in *XRCC3*) were no longer nominally statistically significant after exclusion of the first-published study, hence three differences remain. We found a nominally significantly increased risk of breast cancer for homozygous carriers of the *AURKA* I alleles, and the consortium found nominally significantly increased risks associated with *TGFB1* (all genetic models) and borderline association with *IGFBP3* (per-allele model only), but they do not exclude that these might be a false-positive findings.[10] The *AURKA* meta-analysis was not statistically significant in whites. When consortium data were combined with data published by others, none of these meta-analyses retained nominal statistical significance.

Contrary to our expectations, we found no systematic differences in the degree of between-study heterogeneity between the two approaches. The consortium approach did not consistently reduce between-study heterogeneity. This may be explained by the fact that research groups in the BCAC had not applied the same study methodology, that is, used the same criteria and definitions for the selection of cases and controls and the same technology for the genotyping. The consortium was established retrospectively after the individual studies were designed and conducted. Given the diversity in studies included by the consortium, for analogy, we opted to be all-inclusive in our analyses and to consider studies regardless of menopausal status, unilateral versus bilateral breast cancer, familial versus sporadic cases, screened versus unscreened populations, hospital-based versus population controls, type of genotype platform, and different ethnicities. However, in some settings, consortia may wish to restrict participation to teams that fulfill certain criteria or moreover exclude teams that fail to pass some quality checks (e.g., in genotyping or phenotyping accuracy). For example, in a consortium analysis in Parkinson disease genetics, 7 of 18 teams were excluded from the analyses because of significant deviation in Hardy-Weinberg or genotyping error of 10% or higher upon central genotyping check.[32] It remains to be demonstrated whether consortia show reduced heterogeneity when their members have harmonized and standardized their methodology on a prospective basis,[33] such as is done by the Consortium on the Genetics of Schizophrenia and the Type 1 Diabetes Genetics Consortium.[2,3]

Previous studies have tried to compare the results of meta-analyses based on consortia and individual-level data with meta-analyses of published data. Most of these studies pertain to data from clinical trials and refer to single topics for which the evidence has been synthesized with these two different approaches.[34–37] Data from genetic association studies are sparse.[38,39] Comparisons for meta-analyses involving clinical trials exhibit some additional issues that are uncommon for genetic associations, such as the intricacies of time-to-event analyses, in which individual-level consortia have a clear advantage.[40,41] Consortia also have a clear advantage for the more efficient and reliable investigation of subgroup analyses and effect modifications,[42] but this was not an issue in the data that we analyzed.

It is important to revisit the relative advantages and limitations of consortium-based approaches to genetic associations compared with meta-analyses of published literature. A major advantage of a consortium approach is the access to unpublished data, which reduces problems of publication and selective reporting bias that may affect the integration of published literature.[8,38] On the other hand, meta-analysis of published literature may provide a relatively inexpensive and readily available approach to the synthesis of a large amount of data on genetic associations. Our study showed that sample size and between-study heterogeneity may not differ between the two approaches. The value of consortia may lie predominantly in the improvement of the design and conduct of gene-disease association studies through greater consensus about definitions, criteria, techniques, and procedures in the data collection and analyses, which prospectively facilitates future meta-analyses.[8] Variability between techniques and procedures may remain even after standardization and training.[43] Most important, differences in the summary ORs and inferences between the two approaches may largely depend on which studies are included in the analyses. Finally, with the rapid growth of genetic association investigations, it will become increasingly common to have several different consortia working on the genetics of the same disease. This is already the case in various diseases, including Parkinson disease, breast cancer, and Type 2 diabetes.[6] Collaboration will particularly be important for rare tumors and less common diseases where it may not be possible to amass many thousands of participants with the disease of interest, uncertainty in the effects will be considerably larger and minor methodological biases could more easily create false positives and false negatives. In general, meta-analyses of data from all consortia[44,45] and from additional investigators outside their confines will remain useful for integrating the total evidence. Transparent availability of all consortium and other data are essential in this regard.[46]

In conclusion, in the rapidly emerging field of genetic associations for common diseases, both meta-analyses of published literature and consortium-based analyses may provide important information on genetic associations. We performed meta-analyses of published studies for polymorphisms that were investigated by one consortium that in retrospect combined the data collected by consortium members. Our analyses need to be replicated for the findings of other consortia and for other diseases beyond breast cancer to make a more generalizable statement about all genetic association studies. Nevertheless, despite the methodological issues concerning published studies of genetic associations, we have shown here that consortia and meta-analyses of published data may offer complementary insights.

ACKNOWLEDGMENTS

REFERENCES

1. Chenevix-Trench G, Milne RL, Antoniou AC, Couch FJ, Easton DF, Goldgar DE; CIMBA. An international initiative to identify genetic modifiers of cancer risk in BRCA1 and BRCA2 mutation carriers: the Consortium of Investigators of Modifiers of BRCA1 and BRCA2 (CIMBA). *Breast Cancer Res* 2007;9:104.
2. Calkins ME, Dobie DJ, Cadenhead KS, et al. The Consortium on the Genetics of Endophenotypes in Schizophrenia: model recruitment, assessment, and endophenotyping methods for a multisite collaboration. *Schizophr Bull* 2007;33:33–48.
3. Rich SS, Concannon P, Erlich H, et al. The Type 1 Diabetes Genetics Consortium. *Ann N Y Acad Sci* 2006;1079:1–8.
4. Schaid DJ, Chang BL. Description of the International Consortium For Prostate Cancer Genetics, and failure to replicate linkage of hereditary prostate cancer to 20q13. *Prostate* 2005;63:276–290.
5. Wellcome Trust Case Control Consortium. Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. *Nature* 2007;447:661–678.
6. Ioannidis JP, Bernstein J, Boffetta P, et al. A network of investigator networks in human genome epidemiology. *Am J Epidemiol* 2005;162:302–304.
7. Ioannidis JP, Trikalinos TA, Khoury MJ. Implications of small effect sizes of individual genetic variants on the design and interpretation of genetic association studies of complex diseases. *Am J Epidemiol* 2006;164:609–614.
8. Seminara D, Khoury MJ, O'Brien TR, et al. The emergence of networks in human genome epidemiology: challenges and opportunities. *Epidemiology* 2007;18:1–8.
9. Breast Cancer Association Consortium. Commonly studied single-nucleotide polymorphisms and breast cancer: results from the Breast Cancer Association Consortium. *J Natl Cancer Inst* 2006;98:1382–1396.
10. Cox A, Dunning AM, Garcia-Closas M, et al. A common coding variant in CASP8 is associated with breast cancer risk. *Nat Genet* 2007;39:352–358.
11. Ioannidis JP, Boffetta P, Little J, et al. Assessment of cumulative evidence on genetic associations: interim guidelines. *Int J Epidemiol* 2008;37:120–132.
12. Fletcher O, Johnson N, Palles C, et al. Inconsistent association between the STK15 F31I genetic polymorphism and breast cancer risk. *J Natl Cancer Inst* 2006;98:1014–1018.
13. Garcia-Closas M, Egan KM, Newcomb PA, 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;119:376–388.
14. Johnatty SE, Spurdle AB, Beesley J, et al. Progesterone receptor polymorphisms and risk of breast cancer: results from two Australian breast cancer studies. *Breast Cancer Res Treat* 2008;109:91–99.
15. Pooley KA, Healey CS, Smith PL, et al. Association of the progesterone receptor gene with breast cancer risk: a single-nucleotide polymorphism tagging approach. *Cancer Epidemiol Biomarkers Prev* 2006;15:675–682.
16. Zhang Y, Newcomb PA, Egan KM, et al. Genetic polymorphisms in base-excision repair pathway genes and risk of breast cancer. *Cancer Epidemiol Biomarkers Prev* 2006;15:353–358.
17. Han J, Tranah GJ, Hankinson SE, Samson LD, Hunter DJ. Polymorphisms in O6-methylguanine DNA methyltransferase and breast cancer risk. *Pharmacogenet Genomics* 2006;16:469–474.
18. Lee SA, Lee KM, Park SK, et al. Genetic polymorphism of XRCC3 Thr241Met and breast cancer risk: case-control study in Korean women and meta-analysis of 12 studies. *Breast Cancer Res Treat* 2007;103:71–76.
19. Salanti G, Amountza G, Ntzani EE, Ioannidis JP. Hardy-Weinberg equilibrium in genetic association studies: an empirical evaluation of reporting, deviations, and power. *Eur J Hum Genet* 2005;13:840–848.
20. DerSimonian R, Laird N. Meta-analysis in clinical trials. *Control Clin Trials* 1986;7:177–188.
21. Higgins JPT, Thompson SG, Deeks JJ, Altman DG. Measuring inconsistency in meta-analyses. *BMJ* 2003;327:557–560.
22. Ioannidis JP, Patsopoulos N, Evangelou E. Uncertainty in heterogeneity estimates in meta-analysis. *BMJ* 2007;914–916.
23. Ioannidis JP, Cappelleri JC, Lau J. Issues in comparisons between meta-analyses and large trials. *JAMA* 1998;279:1089–1093.
24. Sterne JA, Jüni P, Schulz KF, Altman DG, Bartlett C, Egger M. Statistical methods for assessing the influence of study characteristics on treatment effects in 'meta-epidemiological' research. *Stat Med* 2002;21:1513–1524.
25. Ioannidis JP, Ntzani EE, Trikalinos TA, Contopoulos-Ioannidis DG. Replication validity of genetic association studies. *Nat Genet* 2001;29:306–309.
26. Harbord RM, Egger M, Sterne JA. A modified test for small-study effects in meta-analyses of controlled trials with binary endpoints. *Stat Med* 2006;25:3443–3457.
27. Ioannidis JP, Trikalinos TA. An exploratory test for an excess of significant findings. *Clinical Trials* 2007;4:245–253.
28. Cai Q, Shu XO, Wen W, 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;6:647–655.
29. Millikan RC, Player J, de Cotret AR, et al. Manganese superoxide dismutase Ala-9Val polymorphism and risk of breast cancer in a population-based case-control study of African Americans and whites. *Breast Cancer Res* 2004;6:264–274.
30. Han J, Hankinson SE, Ranu H, De Vivo I, Hunter DJ. Polymorphisms in DNA double-strand break repair genes and breast cancer risk in the Nurses' Health Study. *Carcinogenesis* 2004;25:189–195.
31. Gaudet MM, Gammon MD, Santella RM, et al. MnSOD Val-9Ala genotype,

pro- and anti-oxidant environmental modifiers, and breast cancer among women on Long Island, New York. *Cancer Causes Control* 2005;16:1225–1234.

32. Maraganore DM, de Andrade M, Elbaz A, et al. Collaborative analysis of alpha-synuclein gene promoter variability and Parkinson disease. *JAMA* 2006;296:661–670.

33. Ioannidis JP, Gwinn M, Little J, et al. A road map for efficient and reliable human genome epidemiology. *Nat Genet* 2006;38:3–5.

34. Stewart LA, Parmar MK. Meta-analysis of the literature or of individual patient data: is there a difference? *Lancet* 1993;341:418–422.

35. Pignon JP, Bourhis J. Meta-analysis of chemotherapy in head and neck cancer: individual patient data vs literature data. *Br J Cancer* 1995;72:1062–1063.

36. Jeng GT, Scott JR, Burmeister LF. A comparison of meta-analytic results using literature vs individual patient data. Paternal cell immunization for recurrent miscarriage. *JAMA* 1995;274:830–836.

37. Steinberg KK, Smith SJ, Stroup DF, et al. Comparison of effect estimates from a meta-analysis of summary data from published studies and from a meta-analysis using individual patient data for ovarian cancer studies. *Am J Epidemiol* 1997;145:917–925.

38. Ioannidis JP, Rosenberg PS, Goedert JJ, O'Brien TR. Commentary: meta-analysis of individual participants' data in genetic epidemiology. *Am J Epidemiol* 2002;156:204–210.

39. Ioannidis JP, Ralston SH, Bennett ST, et al. Differential genetic effects of ESR1 gene polymorphisms on osteoporosis outcomes. *JAMA* 2004;292:2105–2114.

40. Vale CL, Tierney JF, Stewart LA. Effects of adjusting for censoring on meta-analyses of time-to-event outcomes. *Int J Epidemiol* 2002;31:107–111.

41. Tierney JF, Stewart LA. Investigating patient exclusion bias in meta-analysis. *Int J Epidemiol* 2005;34:79–87.

42. Berlin JA, Santanna J, Schmid CH, Szczech LA, Feldman HI. Individual patient- versus group-level data meta-regressions for the investigation of treatment effect modifiers: ecological bias rears its ugly head. *Stat Med* 2002;21:371–387.

43. Swerdlow NR, Sprock J, Light GA, et al. Multi-site studies of acoustic startle and prepulse inhibition in humans: initial experience and methodological considerations based on studies by the Consortium on the Genetics of Schizophrenia. *Schizophr Res* 2007;92:237–251.

44. Evangelou E, Maraganore DM, Ioannidis JP. Meta-analysis in genome-wide association datasets: strategies and application in Parkinson disease. *PLoS ONE* 2007;2:e196.

45. Ioannidis JP, Patsopoulos NA, Evangelou E. Heterogeneity in meta-analyses of genome-wide association investigations. *PLoS ONE* 2007;2:e841.

46. Manolio TA, Rodriguez LL, Brooks L, et al. New models of collaboration in genome-wide association studies: the Genetic Association Information Network. *Nat Genet* 2007;39:1045–1051.

47. Freudenheim JL, Ambrosone CB, Moysich KB, et al. Alcohol dehydrogenase 3 genotype modification of the association of alcohol consumption with breast cancer risk. *Cancer Causes Control* 1999;10:369–377.

48. Hines LM, Hankinson SE, Smith-Warner SA, et al. A prospective study of the effect of alcohol consumption and ADH3 genotype on plasma steroid hormone levels and breast cancer risk. *Cancer Epidemiol Biomarkers Prev* 2000;9:1099–105.

49. Terry MB, Gammon MD, Zhang FF, et al. ADH3 genotype, alcohol intake and breast cancer risk. *Carcinogenesis* 2006;27:840–847.

50. Visvanathan K, Crum RM, Strickland PT, et al. Alcohol dehydrogenase genetic polymorphisms, low-to-moderate alcohol consumption, and risk of breast cancer. *Alcohol Clin Exp Res* 2007;31:467–476.

51. Dai Q, Cai QY, Shu XO, et al. Synergistic effects of STK15 gene polymorphisms and endogenous estrogen exposure in the risk of breast cancer. *Cancer Epidemiol Biomarkers Prev* 2004;13:2065–2070.

52. Egan KM, Newcomb PA, Ambrosone CB, et al. STK15 polymorphism and breast cancer risk in a population-based study. *Carcinogenesis* 2004;25:2149–2153.

53. Sun T, Miao X, Wang J, et al. Functional Phe31Ile polymorphism in Aurora A and risk of breast carcinoma. *Carcinogenesis* 2004;25:2225–2230.

54. Ewart-Toland A, Dai Q, Gao YT, et al. Aurora-A/STK15 T+91A is a general low penetrance cancer susceptibility gene: a meta-analysis of multiple cancer types. *Carcinogenesis* 2005;26:1368–1373.

55. Lo YL, Yu JC, Chen ST, et al. Breast cancer risk associated with genotypic polymorphism of the mitosis-regulating gene Aurora-A/STK15/BTAK. *Int J Cancer* 2005;115:276–283.

56. Cox DG, Hankinson SE, Hunter DJ. Polymorphisms of the AURKA (STK15/Aurora Kinase) gene and breast cancer risk (United States). *Cancer Causes Control* 2006;17:81–83.

57. Tchatchou S, Wirtenberger M, Hemminki K, et al. Aurora kinases A and B and familial breast cancer risk. *Cancer Lett* 2007;247:266–272.

58. Vidarsdottir L, Bodvarsdottir SK, Hilmarsdottir H, Tryggvadottir L, Eyjord JE. Breast cancer risk associated with AURKA 91T–>A polymorphism in relation to BRCA mutations. *Cancer Lett* 2007;250:206–212.

59. Spurdle AB, Hopper JL, Chen X, et al. The BRCA2 372 HH genotype is associated with risk of breast cancer in Australian women under age 60 years. *Cancer Epidemiol Biomarkers Prev* 2002;11:413–416.

60. Ishitobi M, Miyoshi Y, Ando A, et al. Association of BRCA2 polymorphism at codon 784 (Met/Val) with breast cancer risk and prognosis. *Clin Cancer Res* 2003;9:1376–1380.

61. Menzel HJ, Sarmanova J, Soucek P, et al. Association of NQO1 polymorphism with spontaneous breast cancer in two independent populations. *Br J Cancer* 2004;90:1989–1994.

62. Cox DG, Hankinson SE, Hunter DJ. No association between BRCA2 N372H and breast cancer risk. *Cancer Epidemiol Biomarkers Prev* 2005;14:1353–1354.

63. Millikan RC, Player JS, Decotret AR, Tse CK, Keku T. Polymorphisms in DNA repair genes, medical exposure to ionizing radiation, and breast cancer risk. *Cancer Epidemiol Biomarkers Prev* 2005;14:2326–2334.

64. Baynes C, Healey CS, Pooley KA, et al. Common variants in the ATM, BRCA1, BRCA2, CHEK2 and TP53 cancer susceptibility genes are unlikely to increase breast cancer risk. *Breast Cancer Res* 2007;9:R27.

65. Frank B, Hemminki K, Wappenschmidt B, et al. Association of the CASP10 V410I variant with reduced familial breast cancer risk and interaction with the CASP8 D302H variant. *Carcinogenesis* 2006;27:606–609.

66. MacPherson G, Healey CS, Teare MD, et al. Association of a common variant of the CASP8 gene with reduced risk of breast cancer. *J Natl Cancer Inst* 2004;96:1866–1869.

67. Tang D, Cho S, Rundle A, et al. Polymorphisms in the DNA repair enzyme XPD are associated with increased levels of PAH-DNA adducts in a case-control study of breast cancer. *Breast Cancer Res Treat* 2002;75:159–166.

68. Nexo BA, Vogel U, Olsen A, et al. A specific haplotype of single nucleotide polymorphisms on chromosome 19q13.2–3 encompassing the gene RAI is indicative of post-menopausal breast cancer before age 55. *Carcinogenesis* 2003;24:899–904.

69. Forsti A, Angelini S, Festa F, et al. Single nucleotide polymorphisms in breast cancer. *Oncol Rep* 2004;11:917–922.

70. Justenhoven C, Hamann U, Pesch B, et al. ERCC2 genotypes and a corresponding haplotype are linked with breast cancer risk in a German population. *Cancer Epidemiol Biomarkers Prev* 2004;13:2059–2064.

71. Shi Q, Wang LE, Bondy ML, Brewster A, Singletary SE, Wei Q. Reduced DNA repair of benzo[a]pyrene diol epoxide-induced adducts and common XPD polymorphisms in breast cancer patients. *Carcinogenesis* 2004;25:1695–1700.

72. Kuschel B, Chenevix-Trench G, Spurdle AB, et al. Common polymorphisms in ERCC2 (Xeroderma pigmentosum D) are not associated with breast cancer risk. *Cancer Epidemiol Biomarkers Prev* 2005;14:1828–1831.

73. Lee SA, Lee KM, Park WY, et al. Obesity and genetic polymorphism of ERCC2 and ERCC4 as modifiers of risk of breast cancer. *Exp Mol Med* 2005;37:86–90.

74. Zhang L, Zhang Z, Yan W. Single nucleotide polymorphisms for DNA repair genes in breast cancer patients. *Clin Chim Acta* 2005;359:150–155.

75. Debniak T, Scott RJ, Huzarski T, et al. XPD common variants and their association with melanoma and breast cancer risk. *Breast Cancer Res Treat* 2006;98:209–215.

76. Mechanic LE, Hemminki K, Israelsson E, et al. Polymorphisms in nucleotide excision repair genes, smoking and breast cancer in African Americans and whites: a population-based case-control study. *Carcinogenesis* 2006;27:1377–1385.

77. Rohrbacher M, Risch A, Kropp S, Chang-Claude J. The A(-336C) insulin-like growth factor binding protein-3 promoter polymorphism is not a modulator of breast cancer risk in Caucasian women. *Cancer Epidemiol Biomarkers Prev* 2005;14:289–290.

78. Wagner K, Hemminki K, Israelsson E, et al. Polymorphisms in the IGF-1 and IGFBP 3 promoter and the risk of breast cancer. *Breast Cancer Res Treat* 2005;92:133–140.

79. Al-Zahrani A, Sandhu MS, Luben RN, et al. IGF1 and IGFBP3 tagging polymorphisms are associated with circulating levels of IGF1, IGFBP3 and risk of breast cancer. *Hum Mol Genet* 2006;15:1–10.

80. Canzian F, McKay JD, Cleveland RJ, et al. Polymorphisms of genes coding for insulin-like growth factor 1 and its major binding proteins, circulating levels of IGF-I and IGFBP-3 and breast cancer risk: results from the EPIC study. *Br J Cancer* 2006;94:299–307.

81. Cheng I, Penney KL, Stram DO, et al. Haplotype-based association studies of IGFBP1 and IGFBP3 with prostate and breast cancer risk: the multiethnic cohort. *Cancer Epidemiol Biomarkers Prev* 2006;15:1993–1997.

82. Slattery ML, Baumgartner KB, Byers T, et al. Genetic, anthropometric, and lifestyle factors associated with IGF-1 and IGFBP-3 levels in Hispanic and non-Hispanic white women. *Cancer Causes Control* 2005;16:1147–1157.

83. Kuschel B, Auranen A, McBride S, et al. Variants in DNA double-strand break repair genes and breast cancer susceptibility. *Hum Mol Genet* 2002;11:1399–1407.

84. Han DF, Zhou X, Hu MB, et al. Sulfotransferase 1A1 (SULT1A1) polymorphism and breast cancer risk in Chinese women. *Toxicol Lett* 2004;150:167–177.

85. Fabjani G, Tong D, Czerwenka K, et al. Human progesterone receptor gene polymorphism PROGINS and risk for breast cancer in Austrian women. *Breast Cancer Res Treat* 2002;72:131–137.

86. Fernandez LP, Milne RL, Barroso E, et al. Estrogen and progesterone receptor gene polymorphisms and sporadic breast cancer risk: a Spanish case-control study. *Int J Cancer* 2006;119:467–471.

87. Manolitsas TP, Englefield P, Eccles DM, Campbell IG. No association of a 306-bp insertion polymorphism in the progesterone receptor gene with ovarian and breast cancer. *Br J Cancer* 1997;75:1398–1399.

88. Pearce CL, Hirschhorn JN, Wu AH, et al. Clarifying the PROGINS allele association in ovarian and breast cancer risk: a haplotype-based analysis. *J Natl Cancer Inst* 2005;97:51–59.

89. Romano A, Lindsey PJ, Fischer DC, et al. Two functionally relevant polymorphisms in the human progesterone receptor gene (+331 G/A and progins) and the predisposition for breast and/or ovarian cancer. *Gynecol Oncol* 2006;101:287–295.

90. Wang-Gohrke S, Chang-Claude J, Becher H, Kieback DG, Runnebaum IB. Progesterone receptor gene polymorphism is associated with decreased risk for breast cancer by age 50. *Cancer Res* 2000;60:2348–2350.

91. De Vivo I, Hankinson SE, Colditz GA, Hunter DJ. The progesterone receptor Val660–>Leu polymorphism and breast cancer risk. *Breast Cancer Res* 2004;6:636–639.

92. Ambrosone CB, Freudenheim JL, Thompson PA, et al. Manganese superoxide dismutase (MnSOD) genetic polymorphisms, dietary antioxidants, and risk of breast cancer. *Cancer Res* 1999;59:602–606.

93. Mitrunen K, Sillanpää P, Kataja V, et al. Association between manganese superoxide dismutase (MnSOD) gene polymorphism and breast cancer risk. *Carcinogenesis* 2001;22:827–829.

94. Egan KM, Thompson PA, Titus-Ernstoff L, Moore JH, Ambrosone CB. MnSOD polymorphism and breast cancer in a population-based case-control study. *Cancer Lett* 2003;199:27–33.

95. Knight JA, Onay UV, Wells S, et al. Genetic variants of GPX1 and SOD2 and breast cancer risk at the Ontario site of the Breast Cancer Family Registry. *Cancer Epidemiol Biomarkers Prev* 2004;13:146–149.

96. Tamimi RM, Hankinson SE, Spiegelman D, Colditz GA, Hunter DJ. Manganese superoxide dismutase polymorphism, plasma antioxidants, cigarette smoking, and risk of breast cancer. *Cancer Epidemiol Biomarkers Prev* 2004;13:989–996.

97. Bergman M, Ahnstrom M, Palmeback Wegman P, Wingren S. Polymorphism in the manganese superoxide dismutase (MnSOD) gene and risk of breast cancer in young women. *J Cancer Res Clin Oncol* 2005;131:439–444.

98. Cheng TC, Chen ST, Huang CS, et al. Breast cancer risk associated with genotype polymorphism of the catechol estrogen-metabolizing genes: a multigenic study on cancer susceptibility. *Int J Cancer* 2005;113:345–353.

99. Cebrian A, Pharoah PD, Ahmed S, et al. Tagging single-nucleotide polymorphisms in antioxidant defense enzymes and susceptibility to breast cancer. *Cancer Res* 2006;66:1225–1233.

100. Slanger TE, Chang-Claude J, Wang-Gohrke S. Manganese superoxide dismutase Ala-9Val polymorphism, environmental modifiers, and risk of breast cancer in a German population. *Cancer Causes Control* 2006;17:1025–1031.

101. Ziv E, Cauley J, Morin PA, Saiz R, Browner WS. Association between the T29–>C polymorphism in the transforming growth factor beta1 gene and breast cancer among elderly white women: The Study of Osteoporotic Fractures. *Jama* 2001;285:2859–2863.

102. Dunning AM, Ellis PD, McBride S, et al. A transforming growth factorbeta1 signal peptide variant increases secretion in vitro and is associated with increased incidence of invasive breast cancer. *Cancer Res* 2003;63:2610–2615.

103. Hishida A, Iwata H, Hamajima N, et al. Transforming growth factor B1 T29C polymorphism and breast cancer risk in Japanese women. *Breast Cancer* 2003;10:63–69.

104. Krippl P, Langsenlehner U, Renner W, et al. The L10P polymorphism of the transforming growth factor-beta 1 gene is not associated with breast cancer risk. *Cancer Lett* 2003;201:181–184.

105. Jin Q, Hemminki K, Grzybowska E, et al. Polymorphisms and haplotype structures in genes for transforming growth factor beta1 and its receptors in familial and unselected breast cancers. *Int J Cancer* 2004;112:94–99.

106. Le Marchand L, Haiman CA, van den Berg D, Wilkens LR, Kolonel LN, Henderson BE. T29C polymorphism in the transforming growth factor beta1 gene and postmenopausal breast cancer risk: the Multiethnic Cohort Study. *Cancer Epidemiol Biomarkers Prev* 2004;13:412–415.

107. Saha A, Gupta V, Bairwa NK, Malhotra D, Bamezai R. Transforming growth factor-beta1 genotype in sporadic breast cancer patients from India: status of enhancer, promoter, 5′-untranslated-region and exon-1 polymorphisms. *Eur J Immunogenet* 2004;31:37–42.

108. Kaklamani VG, Baddi L, Liu J, et al. Combined genetic assessment of transforming growth factor-beta signaling pathway variants may predict breast cancer risk. *Cancer Res* 2005;65:3454–3461.

109. Lee KM, Park SK, Hamajima N, et al. Genetic polymorphisms of TGF-beta1 & TNF-beta and breast cancer risk. *Breast Cancer Res Treat* 2005;90:149–155.

110. 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;14:1567–1570.

111. Feigelson HS, Patel AV, Diver WR, Stevens VL, Thun MJ, Calle EE. Transforming growth factor beta receptor type I and transforming growth factor beta1 polymorphisms are not associated with postmenopausal breast cancer. *Cancer Epidemiol Biomarkers Prev* 2006;15:1236–1237.

112. Scola L, Vaglica M, Crivello A, et al. Cytokine gene polymorphisms and breast cancer susceptibility. *Ann N Y Acad Sci* 2006;1089:104–109.

113. Gonzalez-Zuloeta Ladd AM, Arias-Vásquez A, Siemes C, et al. Transforming-growth factor beta(1) Leu10Pro polymorphism and breast cancer morbidity. *Eur J Cancer* 2007;43:371–374.

114. Sjalander A, Birgander R, Hallmans G, et al. p53 polymorphisms and haplotypes in breast cancer. *Carcinogenesis* 1996;17:1313–1316.

115. Weston A, Godbold JH. Polymorphisms of H-ras-1 and p53 in breast cancer and lung cancer: a meta-analysis. *Environ Health Perspect* 1997;105(suppl 4):919–926.

116. Wang-Gohrke S, Becher H, Kreienberg R, Runnebaum IB, Chang-Claude J. Intron 3 16 bp duplication polymorphism of p53 is associated with an increased risk for breast cancer by the age of 50 years. *Pharmacogenetics* 2002;12:269–272.

117. Buyru N, Tigli H, Dalay N. P53 codon 72 polymorphism in breast cancer. *Oncol Rep* 2003;10:711–714.

118. Huang XE, Hamajima N, Katsuda N, et al. Association of p53 codon Arg72Pro and p73 G4C14-to-A4T14 at exon 2 genetic polymorphisms with the risk of Japanese breast cancer. *Breast Cancer* 2003;10:307–311.

119. Mabrouk I, Baccouche S, El-Abed R, et al. No evidence of correlation between p53 codon 72 polymorphism and risk of bladder or breast carcinoma in Tunisian patients. *Ann N Y Acad Sci* 2003;1010:764–770.

120. Suspitsin EN, Buslov KG, Grigoriev MY, et al. Evidence against involvement of p53 polymorphism in breast cancer predisposition. *Int J Cancer* 2003;103:431–433.

121. Tommiska J, Eerola H, Heinonen M, et al. Breast cancer patients with p53 Pro72 homozygous genotype have a poorer survival. *Clin Cancer Res* 2005;11:5098–5103.

122. Ma H, Hu Z, Zhai X, et al. Joint effects of single nucleotide polymorphisms in P53BP1 and p53 on breast cancer risk in a Chinese population. *Carcinogenesis* 2006;27:766–771.

123. Gaudet MM, Gammon MD, Bensen JT, et al. Genetic variation of TP53, polycyclic aromatic hydrocarbon-related exposures, and breast cancer risk among women on Long Island, New York. *Breast Cancer Res Treat* 2008; 108:93–99.

124. Khadang B, Fattahi MJ, Talei A, Dehaghani AS, Ghaderi A. Polymorphism of TP53 codon 72 showed no association with breast cancer in Iranian women. *Cancer Genet Cytogenet* 2007;173:38–42.

125. Sprague BL, Trentham-Dietz A, Garcia-Closas M, et al. Genetic variation in TP53 and risk of breast cancer in a population-based case-control study. *Carcinogenesis* 2007;28:1680–1686.

126. Han J, Hankinson SE, De Vivo I, et al. A prospective study of XRCC1 haplotypes and their interaction with plasma carotenoids on breast cancer risk. *Cancer Res* 2003;63:8536–8541.

127. Moullan N, Cox DG, Angèle S, Romestaing P, Gérard JP, Hall J. Polymorphisms in the DNA repair gene XRCC1, breast cancer risk, and response to radiotherapy. *Cancer Epidemiol Biomarkers Prev* 2003;12:1168–1174.

128. Smith TR, Miller MS, Lohman K, et al. Polymorphisms of XRCC1 and XRCC3 genes and susceptibility to breast cancer. *Cancer Lett* 2003;190: 183–190.

129. Smith TR, Levine EA, Perrier ND, et al. DNA-repair genetic polymorphisms and breast cancer risk. *Cancer Epidemiol Biomarkers Prev* 2003;12:1200–1204.

130. Deligezer U, Dalay N. Association of the XRCC1 gene polymorphisms with cancer risk in Turkish breast cancer patients. *Exp Mol Med* 2004;36:572–575.

131. Figueiredo JC, Knight JA, Briollais L, Andrulis IL, Ozcelik H. Polymorphisms XRCC1-R399Q and XRCC3-T241M and the risk of breast cancer at the Ontario site of the Breast Cancer Family Registry. *Cancer Epidemiol Biomarkers Prev* 2004;13:583–591.

132. Chacko P, Rajan B, Joseph T, Mathew BS, Pillai MR. Polymorphisms in DNA repair gene XRCC1 and increased genetic susceptibility to breast cancer. *Breast Cancer Res Treat* 2005;89:15–21.

133. Dufloth RM, Costa S, Schmitt F, Zeferino LC. DNA repair gene polymorphisms and susceptibility to familial breast cancer in a group of patients from Campinas, Brazil. *Genet Mol Res* 2005;4:771–782.

134. Metsola K, Kataja V, Sillanpää P, et al. XRCC1 and XPD genetic polymorphisms, smoking and breast cancer risk in a Finnish case-control study. *Breast Cancer Res* 2005;7:987–997.

135. Patel AV Calle EE, Pavluck AL, Feigelson HS, Thun MJ, Rodriguez C. A prospective study of XRCC1 (X-ray cross-complementing group 1) polymorphisms and breast cancer risk. *Breast Cancer Res* 2005;7:1168–1173.

136. Shen J, Gammon MD, Terry MB, et al. Polymorphisms in XRCC1 modify

the association between polycyclic aromatic hydrocarbon-DNA adducts, cigarette smoking, dietary antioxidants, and breast cancer risk. *Cancer Epidemiol Biomarkers Prev* 2005;14:336–342.

137. Brewster AM, Jorgensen TJ, Ruczinski I, et al. Polymorphisms of the DNA repair genes XPD (Lys751Gln) and XRCC1 (Arg399Gln and Arg194Trp): relationship to breast cancer risk and familial predisposition to breast cancer. *Breast Cancer Res Treat* 2006;95:73–80.

138. Bu D, Tomlinson G, Lewis CM, Zhang C, Kildebeck E, Euhus DM. An intronic polymorphism associated with increased XRCC1 expression, reduced apoptosis and familial breast cancer. *Breast Cancer Res Treat* 2006;99:257–265.

139. Pachkowski BF Winkel S, Kubota Y, Swenberg JA, Millikan RC, Nakamura J. XRCC1 genotype and breast cancer: functional studies and epidemiologic data show interactions between XRCC1 codon 280 His and smoking. *Cancer Res* 2006;66:2860–2868.

140. Thyagarajan B, Anderson KE, Folsom AR, et al. No association between XRCC1 and XRCC3 gene polymorphisms and breast cancer risk: Iowa Women's Health Study. *Cancer Detect Prev* 2006;30:313–321.

141. Zhai X, Liu J, Hu Z, et al. Polymorphisms of ADPRT Val762Ala and XRCC1 Arg399Glu and risk of breast cancer in Chinese women: a case control analysis. *Oncol Rep* 2006;15:247–252.

142. Costa S, Pinto D, Pereira D, et al. DNA repair polymorphisms might contribute differentially on familial and sporadic breast cancer susceptibility: a study on a Portuguese population. *Breast Cancer Res Treat* 2007;103: 209–217.

143. Rafii S, O'Regan P, Xinarianos G, et al. A potential role for the XRCC2 R188H polymorphic site in DNA-damage repair and breast cancer. *Hum Mol Genet* 2002;11:1433–1438.

144. Webb PM, Hopper JL, Newman B, et al. Double-strand break repair gene polymorphisms and risk of breast or ovarian cancer. *Cancer Epidemiol Biomarkers Prev* 2005;14:319–323.

145. Jacobsen NR, Nexø BA, Olsen A, et al. No association between the DNA repair gene XRCC3 T241M polymorphism and risk of skin cancer and breast cancer. *Cancer Epidemiol Biomarkers Prev* 2003;12:584–585.