

Supplementary Appendix

This appendix has been provided by the authors to give readers additional information about their work.

Supplement to: Wacholder S, Hartge P, Prentice R, et al. Performance of common genetic variants in breast-cancer risk models. *N Engl J Med* 2010;362:986-93.

SUPPLEMENTARY TABLE 1. REASONS FOR EXCLUSION (HIERARCHICAL)

STUDY	CONTROL		CASE		TOTAL		
	N	%	N	%	N	%	
NURSES' HEALTH STUDY	WOMEN AGED 50-79	1127	100.00	1123	100.00	2250	100.00
	Non-White	.	0.00	.	0.00	.	0.00
	In Situ BC	.	0.00	.	0.00	.	0.00
	Missing Data on all 10 SNPs	.	0.00	.	0.00	.	0.00
	Missing Data on 1-9 SNPs	179	15.88	118	10.51	297	13.20
	TOTAL IN ANALYTIC FILE	948	84.12	1005	89.49	1953	86.80
PROSTATE LUNG COLON OVARY CANCER SCREENING TRIAL	WOMEN AGED 50-79	1074	100.00	1125	100.00	2199	100.00
	Non-White	103	9.59	104	9.24	207	9.41
	In Situ BC	.	0.00	206	18.31	206	9.37
	Missing Data on all 10 SNPs	58	5.40	68	6.04	126	5.73
	Missing Data on 1-9 SNPs	15	1.40	17	1.51	32	1.46
	TOTAL IN ANALYTIC FILE	898	83.61	730	64.89	1628	74.03
WOMEN'S HEALTH INITIATIVE OBSERVATIONAL STUDY	WOMEN AGED 50-79	2742	100.00	2747	100.00	5489	100.00
	Non-White	.	0.00	.	0.00	.	0.00
	In Situ BC	.	0.00	.	0.00	.	0.00
	Missing Data on all 10 SNPs	496	18.09	492	17.91	988	18.00
	Missing Data on 1-9 SNPs	30	1.09	50	1.82	80	1.46
	TOTAL IN ANALYTIC FILE	2216	80.82	2205	80.27	4421	80.54
AMERICAN CANCER SOCIETY CANCER PREVENTION STUDY II NUTRITION COHORT	WOMEN AGED 50-79	636	100.00	631	100.00	1267	100.00
	Non-White	7	1.10	11	1.74	18	1.42
	In Situ BC	.	0.00	132	20.92	132	10.42
	Missing Data on all 10 SNPs	105	16.51	79	12.52	184	14.52
	Missing Data on 1-9 SNPs	18	2.83	21	3.33	39	3.08
	TOTAL IN ANALYTIC FILE	506	79.56	388	61.49	894	70.56
POLISH BREAST CANCER CONTROL STUDY	WOMEN AGED 50-79	1686	100.00	1574	100.00	3260	100.00
	Non-White	.	0.00	.	0.00	.	0.00
	In Situ BC	.	0.00	82	5.21	82	2.52
	Missing Data on all 10 SNPs	1	0.06	1	0.06	2	0.06
	Missing Data on 1-9 SNPs	255	15.12	229	14.55	484	14.85
	TOTAL IN ANALYTIC FILE	1430	84.82	1262	80.18	2692	82.58
ALL	WOMEN AGED 50-79	7265	100.00	7200	100.00	14465	100.00
	Non-White	110	1.51	115	1.60	225	1.56
	In Situ BC	.	0.00	420	5.83	420	2.90
	Missing Data on all 10 SNPs	660	9.08	640	8.89	1300	8.99
	Missing Data on 1-9 SNPs	497	6.84	435	6.04	932	6.44
	TOTAL IN ANALYTIC FILE	5998	82.56	5590	77.64	11588	80.11

TABLE 2. ODDS RATIO ESTIMATES FROM RISK MODELS

NON-GENETIC MODEL					GENETIC-VARIANT COUNT MODEL				INDIVIDUAL-VARIANT MODEL				INCLUSIVE MODEL			
					NUMBER OF RISK-CONFERRING ALLELES											
	STATUS	ODDS RATIO	LOWER 95% CONFIDENCE INTERVAL	UPPER 95% CONFIDENCE INTERVAL	STATUS	ODDS RATIO	LOWER 95% CONFIDENCE INTERVAL	UPPER 95% CONFIDENCE INTERVAL	STATUS	ODDS RATIO	LOWER 95% CONFIDENCE INTERVAL	UPPER 95% CONFIDENCE INTERVAL	STATUS	ODDS RATIO	LOWER 95% CONFIDENCE INTERVAL	UPPER 95% CONFIDENCE INTERVAL
AGE AT MENARCHE	14+yr (referent)				--	--	--	--	--	--	--	--	14+yr (referent)			
	12-13yr	1.01	0.93	1.11	--	--	--	--	--	--	--	--	12-13yr	1.02	0.94	1.12
	7-11yr	1.16	1.04	1.30	--	--	--	--	--	--	--	--	7-11yr	1.15	1.02	1.29
NUMBER OF BIOPSIES	0 (referent)				--	--	--	--	--	--	--	--	0 (referent)			
	1	1.58	1.43	1.75	--	--	--	--	--	--	--	--	1	1.56	1.41	1.73
	2+	1.31	1.09	1.58	--	--	--	--	--	--	--	--	2+	1.28	1.06	1.55
AGE AT FIRST LIVE BIRTH	<20 yr (referent)				--	--	--	--	--	--	--	--	<20 yr (referent)			
	20-24 yr	0.98	0.88	1.09	--	--	--	--	--	--	--	--	20-24 yr	0.97	0.86	1.08
	25-29 or No Births	1.25	1.11	1.40	--	--	--	--	--	--	--	--	25-29 or No Births	1.22	1.09	1.38
	30+ yr	1.38	1.18	1.61	--	--	--	--	--	--	--	--	30+ yr	1.36	1.16	1.59
NUMBER OF FIRST-DEGREE RELATIVES WITH BREAST CANCER	0 (referent)				--	--	--	--	--	--	--	--	0 (referent)			
	1	1.39	1.25	1.54	--	--	--	--	--	--	--	--	1	1.35	1.21	1.50
	2+	1.52	1.12	2.06	--	--	--	--	--	--	--	--	2+	1.44	1.06	1.96
NUMBER OF RISK CONFERRING ALLELES	--	--	--	--	0-6 (referent)				--	--	--	--	--	--	--	--
	--	--	--	--	7-8	1.23	1.07	1.42	--	--	--	--	--	--	--	--
	--	--	--	--	9-10	1.62	1.42	1.85	--	--	--	--	--	--	--	--
	--	--	--	--	11-12	2.27	1.96	2.62	--	--	--	--	--	--	--	--
	--	--	--	--	13+	2.90	2.37	3.55	--	--	--	--	--	--	--	--
RS1045485 (2q, CASP8)	--	--	--	--	--	--	--	--	0 (referent)				0 (referent)			
	--	--	--	--	--	--	--	--	1	0.89	0.82	0.98	1	0.90	0.82	0.99
	--	--	--	--	--	--	--	--	2	0.69	0.52	0.91	2	0.69	0.52	0.93

	--	--	--	--	--	--	--	--	--	0 (referent)				0 (referent)			
RS13281615 (8q, unknown)	--	--	--	--	--	--	--	--	--	1	1.14	1.05	1.24	1	1.14	1.04	1.24
	--	--	--	--	--	--	--	--	--	2	1.36	1.22	1.51	2	1.35	1.21	1.50
RS13387042 (2q, unknown)	--	--	--	--	--	--	--	--	--	0 (referent)				0 (referent)			
	--	--	--	--	--	--	--	--	--	1	0.80	0.73	0.87	1	0.80	0.74	0.88
	--	--	--	--	--	--	--	--	--	2	0.70	0.63	0.78	2	0.71	0.64	0.79
RS2981582 (10q, FGFR2)	--	--	--	--	--	--	--	--	--	0 (referent)				0 (referent)			
	--	--	--	--	--	--	--	--	--	1	1.18	1.09	1.28	1	1.17	1.08	1.28
	--	--	--	--	--	--	--	--	--	2	1.60	1.44	1.79	2	1.58	1.42	1.77
RS3803662 (10q, TOX3)	--	--	--	--	--	--	--	--	--	0 (referent)				0 (referent)			
	--	--	--	--	--	--	--	--	--	1	1.16	1.07	1.25	1	1.15	1.06	1.24
	--	--	--	--	--	--	--	--	--	2	1.44	1.26	1.66	2	1.42	1.23	1.63
RS3817198 (11p, LSPI)	--	--	--	--	--	--	--	--	--	0 (referent)				0 (referent)			
	--	--	--	--	--	--	--	--	--	1	1.04	0.96	1.12	1	1.03	0.95	1.12
	--	--	--	--	--	--	--	--	--	2	1.18	1.04	1.33	2	1.15	1.02	1.31
RS889312 (16q, MAP3K1)	--	--	--	--	--	--	--	--	--	0 (referent)				0 (referent)			
	--	--	--	--	--	--	--	--	--	1	1.05	0.98	1.14	1	1.05	0.97	1.13
	--	--	--	--	--	--	--	--	--	2	1.10	0.96	1.26	2	1.10	0.96	1.27
RS7716600 (5p, unknown)	--	--	--	--	--	--	--	--	--	0 (referent)				0 (referent)			
	--	--	--	--	--	--	--	--	--	1	1.11	1.03	1.20	1	1.12	1.03	1.21
	--	--	--	--	--	--	--	--	--	2	1.46	1.23	1.72	2	1.47	1.24	1.74
RS11249433 (1p, unknown)	--	--	--	--	--	--	--	--	--	0 (referent)				0 (referent)			
	--	--	--	--	--	--	--	--	--	1	1.23	1.13	1.34	1	1.24	1.14	1.34
	--	--	--	--	--	--	--	--	--	2	1.30	1.16	1.45	2	1.30	1.16	1.45
RS999737 (14q, RAD51L1)	--	--	--	--	--	--	--	--	--	0 (referent)				0 (referent)			
	--	--	--	--	--	--	--	--	--	1	0.91	0.85	0.99	1	0.92	0.85	0.99
	--	--	--	--	--	--	--	--	--	2	0.67	0.56	0.80	2	0.69	0.58	0.82

TABLE 3. COMPARISON OF MODELS USING AREA UNDER THE CURVE (AUC) AND INTEGRATED DISCRIMINATION IMPROVEMENT (IDI)

		Area under the curve ¹	Integrated Discrimination Improvement ² (annual risk of breast cancer)			
			<u>Comparison Model</u>			
<u>Model #</u>	<u>Base Model</u>		<u>Non-Genetic</u>	<u>Genetic-Variant Count</u>	<u>Genetic Individual-Variant</u>	<u>Inclusive</u>
		-				
1	<u>Demographic</u>	<i>53.4%</i>	<i>0.016%</i>	<i>0.019%</i>	<i>0.023%</i>	<i>0.038%</i>
2	<u>Non-Genetic</u>	<i>58.0%</i>	-	<i>0.003%</i>	<i>0.007%</i>	<i>0.022%</i>
3	<u>Genetic-Variant Count</u>	<i>58.9%</i>	-	-	<i>0.004%</i>	<i>0.019%</i>
4	<u>Genetic Individual-Variant</u>	<i>59.7%</i>	-	-	-	<i>0.015%</i>
5	<u>Inclusive</u>	<i>61.8%</i>	-	-	-	-

¹ The range of the standard errors of the AUC is from 0.30% to 0.51%.

²IDI represents improvement in discrimination from using the comparison model instead of the base model. When the base model is the Demographic or the Comparison model is inclusive, the p-values for IDI are all below 1 in a billion. The p-values for the three IDIs comparing the other three models are less than 0.0001, except that the p-value for the comparison between the Genetic-Variant Count and Non-Genetic models is 0.053.

TABLE 4. CROSS-CLASSIFICATION OF CASES INTO QUINTILES OF ESTIMATED ANNUAL BREAST CANCER RISK ACCORDING TO INCLUSIVE AND NON-GENETIC MODELS

INCLUSIVE MODEL

RANGE OF ANNUAL ESTIMATED BREAST CANCER RISK

NON-GENETIC MODEL		Quintile										TOTAL
		1		2		3		4		5		
RANGE OF ANNUAL ESTIMATED BREAST CANCER RISK		MIN	MAX	MIN	MAX	MIN	MAX	MIN	MAX	MIN	MAX	
MINIMUM	MAXIMUM	0.000 %	0.367 %	0.367 %	0.434 %	0.434 %	0.498 %	0.498 %	0.575 %	0.575 %	1.000 %	
...	0.367%	9.8%		3.7%		1.0%		0.2%		0.0%		14.6%
0.367%	0.434%	5.2%		8.3%		6.3%		2.8%		0.7%		23.3%
0.434%	0.498%	1.4%		4.7%		6.9%		6.2%		2.8%		21.9%
0.498%	0.575%	0.2%		1.7%		3.6%		6.9%		8.9%		21.3%
0.575%	...	0.0%		0.1%		0.5%		2.9%		15.3%		18.9%
TOTAL		16.6%		18.4%		18.3%		18.9%		27.7%		100.0%

Supplementary Appendices

Appendix 1: Description of Studies

WHI OS. The Women's Health Initiative (WHI) Observational Study (OS) cohort enrolled 93,676 postmenopausal women between 1993 and 1998.¹ Women ineligible for, or not interested in, the clinical trials were given the opportunity to enroll in the observational study (OS). As of September 2005, the OS cohort had been followed for an average of 7.5 years.

CPS II. The American Cancer Society Cancer Prevention Study II Nutrition Cohort (CPS II) was established in 1992. The cohort includes over 86,000 men and 97,000 women from 21 U.S. states who completed a mailed questionnaire in 1992.² At baseline, the cohort was 97% white and the median age of participants was 63 (range: 40-92). Starting in 1997, follow-up questionnaires have been sent to surviving cohort members every other year to update exposure information and to ascertain occurrence of new cases of cancer. A >90% response rate has been achieved for each follow-up questionnaire. From 1998 to 2001, blood samples were collected from a subgroup of 39,376 cohort members.

PBCS. The Polish Breast Cancer Case-Control Study (PBCS) is a population-based case-control study in Warsaw and Łódź, Poland, conducted between 2000 and 2003.³

PLCO. The Prostate Lung Colon Ovary Cancer Screening (PLCO) trial cohort consists of women randomized into the intervention arm of the trial—a randomized multicenter trial to investigate whether screening for prostate, lung, colorectal, and ovarian cancer will reduce cancer-specific incidence and mortality.⁴ Details of the study have been described elsewhere. Briefly, women aged 55 to 74 years were recruited between November 1993 and July 2001 in 10 U.S. centers. Participants with a personal history of one of the four PLCO cancers, a recent history of screening procedures for one of the cancers, or who were currently undergoing treatment for any cancer, except non-melanoma skin cancer were excluded from the trial. After approval by the institutional review boards of the U.S. National Cancer Institute and each of the participating centers, each eligible participant provided written informed consent. Women randomized to the intervention arm underwent periodic cancer screening tests, including chest X-ray, flexible sigmoidoscopy, digital rectal examination, cancer antigen 125 screening, and transvaginal ultrasound. Women randomized to the control arm were instructed to follow their usual medical practice.

NHS. The Nurses' Health Study (NHS) is a longitudinal study of 121,700 women enrolled in 1976.⁵ The CGEMS⁶ (Cancer Genetic Marker of Susceptibility) participants are from a nested, case-control study derived from 32,826 participants who provided a blood sample between 1989 and 1990 and were free of diagnosed breast cancer at blood collection and followed for incident disease until June 1, 2004. Cancer follow-up in the NHS was conducted by personal mailings and searches of the National Death Index. It is estimated that the percentage of true cancers in the cohort captured by this system is greater than 98%.⁷ Permission was requested from all participants diagnosed with cancer to review medical records to confirm the diagnoses and obtain additional information on tumor histology, staging, and other characteristics. Informed consent was obtained from all participants. The study was approved by the Institutional Review Board of the Brigham and Women's Hospital, Boston, MA, USA.

Appendix 2: Description of Ascertainment

WHI OS. Initial reports of breast cancer were ascertained by self-administered annual questionnaires. Breast cancer occurrences were confirmed by medical record and pathology report review by physician-adjudicators at the local clinical centers. All cases were subsequently classified centrally at the coordinating center by using the National Cancer Institute's Surveillance, Epidemiology, and End Results (SEER) coding system.⁸ A complete description of the WHI outcomes adjudication process is in reference⁹. Pathologic confirmation is standard in WHI. Occasionally a cancer is diagnosed clinically because histologic confirmation is not available.

Ninety-seven percent of self-reported breast cancer cases were confirmed locally. Subsequent central adjudication confirmed > 99% of the locally confirmed breast cancer cases, and 98% of the invasive breast cancer cases. Invasive breast cancers only are included in this project.

CPS II. Breast cancer cases are initially identified through self-report on questionnaires or identification in the National Death Index. Incident cancers are verified through medical records, state cancer registries, or death certificates. Similarly, eligible cases were required to be post-menopausal at diagnosis and needed to have sufficient sample for DNA extraction. Subjects were not eligible if they had a history of another cancer other than non-melanoma skin cancer prior to their diagnosis of breast cancer.

PBCS. Incident cases from 2000 to 2003 were identified through a rapid identification system in participating hospitals covering ~ 90% of all eligible cases; and through periodic checks against the cancer registries in Warsaw and Łódź to assure complete identification of cases. 79% of eligible cases agreed to personal interviews; 84% of interviewed cases provided DNA.

PLCO. Study subjects were sent a mail-in health survey annually asking whether they had been diagnosed with cancer, and if so, the type of cancer. Incident breast cancer cases were ascertained through self-report in the annual survey, state cancer registries, death certificates, physician reports, and next-of-kin reports for deceased individuals. Pathology reports were sought for all cases and 73% of ascertained breast cancer cases were confirmed through medical record review. The results from sub-analyses excluding unconfirmed cases (27% of total cases) and in situ cases (13% of total cases) did not differ from the results for all cases; therefore, we included all ascertained cases in our final analyses to increase the statistical power.

NHS. All study participants who were menopausal at blood draw with a confirmed diagnosis of invasive breast cancer and had sufficient blood sample available for DNA extraction at the time of case and control selection were included as cases in the CGEMS project.

Appendix 3: Description of Control Selection

WHI OS. The WHI OS set of breast cancer cases and controls used in CGEMS were obtained from two previously selected breast cancer core studies. All participants were ineligible if they had an inadequate volume of available DNA, or refused genetic testing. One of the core studies also excluded participants with a history of breast cancer reported at baseline. Invasive breast cancer cases were matched to one control on age at screening, enrollment date, race/ethnicity, hysterectomy status at baseline, and history of breast cancer at baseline.

CPS II. Controls were selected to match to cases by single year of age, gender, ethnicity, and date of sample collection. All controls are postmenopausal at the time of case diagnosis. All controls were selected from individuals who were cancer-free (except for non-melanoma skin cancer) at the time of diagnosis of each case.

PBCS. Randomly selected from population lists of all residents of Poland, stratified and frequency matched to cases by case city and age in 5 year categories. 69% of eligible controls agreed to personal interview; 94% of interviewed controls provided DNA.

PLCO. Women eligible for control selection had at least one annual study update (ASU), a baseline questionnaire, no prior history of breast cancer, signature, and blood available. Controls were then frequency matched to cases by age at entry (55-59, 60-64, 65-69, 70-74) and year of entry (pre-9/30/1997 vs. post-10/1/1997).

NHS. Controls were not diagnosed with breast cancer during follow-up, nor were they matched to cases based on age at diagnosis, blood collection variables (time of day, season, and year of blood collection, as well as

recent (<3 months) use of postmenopausal hormones), ethnicity (all cases and controls are self-reported Caucasians), and menopausal status (all cases were postmenopausal at diagnosis). Menopausal status was defined according to self-reported information on whether a woman's regular periods had ceased.¹⁰

Appendix 4: Age Adjustment

In WHI OS, CPS II and PBCS, we adjusted for attained age (age at diagnosis for cases; age of matched case for controls) as categorical (in decades) and continuous variables. For cases and controls in NHS and PLCO, we adjusted for categorical and continuous age at entry to cohort and continuous entry year.

Appendix 5: Assessment of Interaction

To use the Tukey one-degree-of- freedom method,^{11, 12} we calculated the sums of the products of the regression coefficients and regression variables from the BCRAT factors and from the 10 SNPs. We calculated the likelihood ratio from adding a single interaction variable, calculated as the product of the two sums, less their means. We also assigned cases and controls to levels of the two sums based on quintiles of the controls, and compared the fit of models from adding the 16 interaction variables.

Appendix 6: Calibration and estimates of absolute risk

The average risk for women in stratum j is

$$B_j = c_j \sum_{i=1}^{N_j} R(v_i) / N_j,$$

where we use j to index age intervals and obtained rates from the National Cancer Institute's Surveillance, Epidemiology and End Results (SEER) registry for the US cohorts, and cancer registries in Warsaw and Warsaw and Łódź, Poland for the Polish Breast Cancer Study. The right hand side of the equation above is the product of the calibration constant c_j and the average of individual relative risks $R_i = R(v_i)$, where v_i is a set of factors that affects risk. Because both sides of the equation represent average risk in stratum j , the calibration constant is $c_j = B_j / \overline{R_j}$, where $\overline{R_j} = \sum_{i=1}^{N_j} R(v_i) / N_j$, with summation over all the women in the stratum. We now can calculate the calibrated absolute risk as

$$A_i = c_j R(v_i) = B_j R_i / \overline{R_j}$$

for woman i among the N_j women in demographically-defined (age and cohort) stratum $j(i) = S_i$ with crude (average) risk in stratum j known to be B_j from a population-based registry. Note that the model is calibrated within a stratum. Summing over all the women in stratum j , we see that

$$\sum A_i = \sum B_j (R_i / \overline{R_j}) = B_j \sum R_i / \overline{R_j} = B_j N_j.$$

In calculations of absolute risk for individuals, B_j corresponds to single years of age rather than 10-year age strata. Thus, the calibration constants are different for each single year.

These calculations ignore age matching in some studies. They also ignore competing risks and are appropriate for short time intervals, such as five years or less.

Appendix 7: Sampling Fractions and Weights

We calculated sampling fractions for all studies by demographic stratum to reconstruct a hypothetical cohort using weights equal to the reciprocal of the sampling fraction of the case or control. The sampling fraction for cases and controls in a demographic stratum are the stratum-specific proportions of cases included in the analysis among all cases diagnosed in the cohort and of cohort members in the analysis among all cohort members, respectively. Sampling fractions reflect exclusion for non-participation or missing data. The sampling fractions for controls are very small, reflecting the economical designs of the studies.

Appendix 8: Overfitting

Overfitting is the tendency of models to perform better in the data from which they were derived. We assessed overfitting by comparing estimated AUC's for the key models in the random 70% of the data used for the fit and in the remainder of the data which was not used for the estimation. All models were fit using a 70% simple random sample (3,914 cases and 4,199 controls) of the 11,588 women available for analysis. We found differences of less than 1.8 percentage points between estimates of AUC's and of less than 0.7 percentage points in differences between AUC's of pairs of models when fitting and evaluating models with the full data, compared to fitting the model in 70% of the data and evaluating its performance in the other 30% (data not shown). We report results from the full data.

Appendix 9: Precision and Testing of Differences in Model Performance

We evaluated the standard error of AUC and its significance by fitting each model 1,000 times based on bootstrap resampling with replacement of cases and controls. The significance of the IDI, which is calculated from the calibrated estimates of absolute risk, is calculated using the standard error estimator in Pencina et al.¹³

REFERENCES

1. Langer RD, White E, Lewis CE, Kotchen JM, Hendrix SL, Trevisan M. The Women's Health Initiative Observational Study: baseline characteristics of participants and reliability of baseline measures. *Ann Epidemiol* 2003;13:S107-21.
2. Calle EE, Rodriguez C, Jacobs EJ, et al. The American Cancer Society Cancer Prevention Study II Nutrition Cohort: rationale, study design, and baseline characteristics. *Cancer* 2002;94:2490-501.
3. 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;119:376-88.
4. Hayes RB, Reding D, Kopp W, et al. Etiologic and early marker studies in the prostate, lung, colorectal and ovarian (PLCO) cancer screening trial. *Control Clin Trials* 2000;21:349S-55S.
5. Willett WC, Stampfer MJ, Colditz GA, Rosner BA, Hennekens CH, Speizer FE. Moderate alcohol consumption and the risk of breast cancer. *N Engl J Med* 1987;316:1174-80.
6. Thomas G, Jacobs KB, Kraft P, et al. A multistage genome-wide association study in breast cancer identifies two new risk alleles at 1p11.2 and 14q24.1 (RAD51L1). *Nat Genet* 2009;41:579-84.
7. Tworoger SS, Eliassen AH, Sluss P, Hankinson SE. A prospective study of plasma prolactin concentrations and risk of premenopausal and postmenopausal breast cancer. *J Clin Oncol* 2007;25:1482-8.
8. Surveillance, Epidemiology, and End Results (SEER) Program. SEER 17 Incidence and Mortality, 2000-2003, with Kaposi Sarcoma and Mesothelioma. National Cancer Institute, 2007. (Accessed at www.seer.cancer.gov/seerstat.)
9. Curb JD, McTiernan A, Heckbert SR, et al. Outcomes ascertainment and adjudication methods in the Women's Health Initiative. *Ann Epidemiol* 2003;13:S122-8.
10. Colditz GA, Rosner B. Cumulative risk of breast cancer to age 70 years according to risk factor status: data from the Nurses' Health Study. *Am J Epidemiol* 2000;152:950-64.
11. Tukey JW. One degree of freedom for non-additivity. *Biometrics* 1949;5:232-42.
12. Chatterjee N, Kalaylioglu Z, Moslehi R, Peters U, Wacholder S. Powerful multilocus tests of genetic association in the presence of gene-gene and gene-environment interactions. *Am J Hum Genet* 2006;79:1002-16.
13. Pencina MJ, D'Agostino RB, Sr., D'Agostino RB, Jr., Vasan RS. Evaluating the added predictive ability of a new marker: from area under the ROC curve to reclassification and beyond. *Stat Med* 2008;27:157-72; discussion 207-12.