

ARTICLE

Performance of the Breast Cancer Risk Assessment Tool Among Women Aged 75 Years and Older

Mara A. Schonberg, Vicky W. Li, A. Heather Eliassen, Roger B. Davis, Andrea Z. LaCroix, Ellen P. McCarthy, Bernard A. Rosner, Rowan T. Chlebowski, Thomas E. Rohan, Susan E. Hankinson, Edward R. Marcantonio, Long H. Ngo

Affiliations of authors: Division of General Medicine and Primary Care, Department of Medicine, Harvard Medical School, Beth Israel Deaconess Medical Center, Boston, MA (MAS, VWL, RBD, EPM, ERM, LHN); Department of Epidemiology, Harvard School of Public Health, Boston, MA and Channing Division of Network Medicine, Brigham and Women's Hospital, Harvard Medical School, Harvard School of Public Health, Boston, MA (AHE, BAR, SEH); Division of Epidemiology, Family and Preventive Medicine, University of California San Diego, La Jolla, CA (AZL); Los Angeles Biomedical Research Institute at Harbor-UCLA Medical Center, Torrance, CA (RTC); Department of Epidemiology & Population Health, Albert Einstein College of Medicine, Bronx, NY (TER); Department of Biostatistics and Epidemiology, University of Massachusetts, Amherst, MA (SEH).

Correspondence to: Mara A. Schonberg, MD, MPH, Beth Israel Deaconess Medical Center, 1309 Beacon, Office 219, Brookline, MA 02446 (e-mail: mschonbe@bidmc.harvard.edu).

Abstract

Background: The Breast Cancer Risk Assessment Tool (BCRAT, "Gail model") is commonly used for breast cancer prediction; however, it has not been validated for women age 75 years and older.

Methods: We used Nurses' Health Study (NHS) data beginning in 2004 and Women's Health Initiative (WHI) data beginning in 2005 to compare BCRAT's performance among women age 75 years and older with that in women age 55 to 74 years in predicting five-year breast cancer incidence. BCRAT risk factors include: age, race/ethnicity, age at menarche, age at first birth, family history, history of benign breast biopsy, and atypia. We examined BCRAT's calibration by age by comparing expected/observed (E/O) ratios of breast cancer incidence. We examined discrimination by computing c-statistics for the model by age. All statistical tests were two-sided.

Results: Seventy-three thousand seventy-two NHS and 97 081 WHI women participated. NHS participants were more likely to be non-Hispanic white (96.2% vs 84.7% in WHI, $P < .001$) and were less likely to develop breast cancer (1.8% vs 2.0%, $P = .02$). E/O ratios by age in NHS were 1.16 (95% confidence interval [CI] = 1.09 to 1.23, age 57–74 years) and 1.31 (95% CI = 1.18 to 1.45, age ≥ 75 years, $P = .02$), and in WHI 1.03 (95% CI = 0.97 to 1.09, age 55–74 years) and 1.10 (95% CI = 1.00 to 1.21, age ≥ 75 years, $P = .21$). E/O ratio 95% confidence intervals crossed one among women age 75 years and older when samples were limited to women who underwent mammography and were without significant illness. C-statistics ranged between 0.56 and 0.58 in both cohorts regardless of age.

Conclusions: BCRAT accurately predicted breast cancer for women age 75 years and older who underwent mammography and were without significant illness but had modest discrimination. Models that consider individual competing risks of non-breast cancer death may improve breast cancer risk prediction for older women.

Women age 75 years and older are the fastest growing segment of the US population, and breast cancer incidence increases with age (1). However, none of the randomized trials evaluating

mammography screening included women age 75 years and older. Therefore, it is not known if screening helps these women live longer (2). Ideally, screening decisions would consider an

Received: February 9, 2015; Revised: June 17, 2015; Accepted: October 20, 2015

© The Author 2015. Published by Oxford University Press. All rights reserved. For Permissions, please e-mail: journals.permissions@oup.com.

older woman's breast cancer risk, life expectancy, and preferences (3,4). While there are tools to help estimate life expectancy (5) and elicit patient preferences (6), little is known about late-life breast cancer risk factors.

The Breast Cancer Risk Assessment Tool (BCRAT) is the most commonly used breast cancer risk prediction model in primary care; however, its performance among women age 75 years and older is not known (7–9). BCRAT was developed from a statistical model known as the Gail model, which was developed using data from the Breast Cancer Detection Demonstration Project (BCDDP), a study that recruited 280 000 women age 35 to 74 between 1973 and 1980 (10,11). BCRAT considers a woman's age, race/ethnicity, age at menarche, age at first live birth, number of first-degree relatives with breast cancer, number of benign breast biopsies, and history of atypical hyperplasia. However, several of these risk factors (eg, age at menarche) affect women's estrogen levels relatively early in life and may not be important for late-life breast cancer risk (4). Therefore, we aimed to assess BCRAT's performance among women age 75 years and older compared with that in postmenopausal women age 55 to 74 years participating in the Nurses' Health Study (NHS) and Women's Health Initiative (WHI). We chose these cohorts because they include many women age 75 years and older and have captured necessary data on breast cancer risk factors. Our goal was to determine BCRAT's accuracy for assessing breast cancer risk when deciding whether or not to screen women age 75 years and older for breast cancer.

Methods

BCRAT

BCRAT estimates the probability that a woman will develop invasive breast cancer in five years. To calculate risk, BCRAT considers a woman's baseline hazard of developing breast cancer based on her age and race using age (categorized into 5-year age groups) and race/ethnic-specific population breast cancer incidence rates from the National Cancer Institute (NCI)'s Surveillance, Epidemiology, and End Results (SEER) program. BCRAT uses SEER incidence rates for non-Hispanic whites from 1983 to 1987, for non-Hispanic Blacks from 1994 to 1998, for Hispanics from 1990 to 1996, and for Asians from 1998 to 2002. The model then adjusts a woman's risk by considering whether risk factors included in the model are present or absent and the relative risk estimate of each risk factor. The model also considers the amount of risk that can be explained by the risk factors included (attributable risk). Finally, BCRAT considers a woman's age based risk of death using National Center for Health Statistics data (7). All women in an age group are considered to have the same risk of death regardless of their health. The source data (Supplementary Table 1, available online) and formula for BCRAT may be found in the Supplementary Materials (available online).

Data

We describe briefly below each cohort used in our analyses (12–14); detailed descriptions are in the Supplementary Materials (available online). NHS is a longitudinal study of 121 700 female nurses, age 30 to 55 years in 1976, from 11 of the most populous US states (14). At baseline and in biennial follow-ups, participants provide detailed lifestyle and medical history information through mailed questionnaires. WHI is a multicenter study that recruited 161 808 postmenopausal US

women age 50 to 79 years in up to three clinical trials (WHI-CTs) or an observational study (WHI-OS) from 1993 to 1998 and initially followed women through March 2005. The majority of participants (82% of WHI-CT participants and 73% of WHI-OS participants) agreed to an extension study (WHI-ES; n = 115 400) through September 2010. We chose WHI-ES participants for our analyses because many had aged past 75 years and most had stopped using hormone therapy, which is typical of current practice (15).

Sample

NHS

Participants entered our study the month they returned their 2004 questionnaire (NHS measures dates in months). Questionnaires could be returned through May 2006 (4.5% of participants returned their 2004 questionnaire in 2006), so entry into the study varied. We excluded women who were alive in 2004 but did not complete a 2004 questionnaire (n = 12 539) because ascertainment of breast cancer may be incomplete for these women. We chose to begin following NHS participants in 2004 because this time period is similar to that of WHI-ES and is two years after publication of WHI's estrogen plus progesterone (E+P) clinical trial results that found that use of E+P increased breast cancer risk (16). Breast cancer risk associated with E+P was found to rapidly decline within two years of discontinuation (15). Participants ranged in age from 57 to 86 years at study entry.

WHI-ES

Participants entered our study the day they consented to participate in WHI-ES. Participants were between age 55 and 91 years at study entry. In primary analyses, we excluded women with a history of cancer (except nonmelanomatous skin cancer) in both cohorts (n = 9394 in NHS, n = 9812 in WHI) because NHS does not confirm second diagnoses of cancer.

Outcomes

We followed participants for up to five years or until they developed invasive breast cancer or died, whichever came first. All WHI breast cancer cases were confirmed by pathology report. For NHS, we also included self-reported breast cancers (12% of cases) because validation studies have found that self-reported breast cancers in NHS are accurate (99% are confirmed when medical records are obtained) (17). Detailed definitions of covariates and outcome variables are in the Supplementary Materials (available online).

Risk Factors

We used available data on BCRAT risk factors (age, race/ethnicity, age at menarche, age at first birth, family history, and history of benign breast biopsy). Because NHS does not provide information on Asian ethnicity subgroups, to determine baseline breast cancer incidence for Asians we randomly assigned Asian ethnicity for the 0.83% of NHS participants that were Asian using 1970 Census data (NHS began in 1976). While BCRAT follows women through 89.999 years, it does not compute five-year probabilities for women older than age 85 years at entry. Therefore, we reclassified women older than age 85 years at study entry as being 85 (17 NHS and 1253 WHI-ES participants were reclassified).

To further compare NHS and WHI-ES participants, we used data on participant mammography use in the past two years, body mass index, oophorectomies, hormone therapy use, and history of significant illness (including diabetes, myocardial infarction, emphysema, congestive heart failure, stroke, or peripheral artery disease). In NHS, all of these conditions were confirmed through follow-up of participants and/or review of their medical records except for emphysema and heart failure. In WHI-ES, all conditions were physician-adjudicated with medical records except for emphysema and diabetes. WHI-ES does not assess reasons for undergoing mammography; therefore, we present receipt of any mammogram in the past two years.

Statistical Analyses

We used chi-square tests to compare characteristics between NHS and WHI-ES. Tests of statistical significance were two-sided, and a *P* value of less than .05 was considered to be statistically significant. To examine BCRAT's performance, we measured BCRAT's calibration and discrimination within our cohorts stratified by age (55–74 vs 75+) (18). Calibration assesses whether a model's predicted probabilities are accurate. Discrimination assesses how well a model distinguishes between individuals who do or do not experience the event of interest. A model that discriminates well will assign higher risk values to those who develop the outcome of interest (18). To assess calibration, we compared the expected (*E*) number of breast cancers based on BCRAT estimates (calculated using BCRAT's SAS macro) (7) to the observed number (*O*) in each cohort overall and within deciles of risk stratified by age (55–74, 75+ years). To determine deciles of risk, we ordered the probabilities given by BCRAT for each age group within a cohort and categorized these probabilities into deciles. Within each decile, we took the average probability of risk and multiplied this probability by the sample size within each decile to determine the expected number of breast cancers (*E*). We calculated 95% confidence intervals (CIs) of *E/O* ratios using the Poisson variance for the logarithm of the observed number of cases (19). We further assessed calibration using the Hosmer-Lemeshow (HL) chi-square test. An *E/O* ratio close to 1 (meaning the model's estimates of risk matches the actual risk) and a nonsignificant HL-test statistic indicate good calibration (20). To assess BCRAT's discrimination, we used Rosner and Glynn's methods to determine the *c*-statistic or area under the receiver operating characteristic curve and its standard error (21). This area ranges from 0.5 (no discrimination) to 1.0

(perfect discrimination) (18). To test whether *E/O* ratio estimates and *c*-statistics differed by age within cohort, we used the normal approximation *z*-test. Our primary analyses were limited to participants with complete data on BCRAT risk factors.

Sensitivity Analyses

Because previous studies found that BCRAT performs better among women who undergo mammography, we repeated our analyses limiting our sample to women who underwent mammography in the past two years (10,22,23). Because we were interested in BCRAT's performance among women age 75 years and older and comorbidity increases with age, we repeated our analyses excluding women with significant illnesses (defined above). In addition, we repeated our analyses limiting our sample to women who were recently screened and were without significant illness. We also repeated our analyses using SEER incidence rates from 2006 to 2010 for whites, blacks, and Hispanics because these data matched our study period. Because previous studies found that BCRAT performs better when predicting estrogen receptor-positive (ER+) breast cancer; we reassessed BCRAT's discrimination using ER+ breast cancer as our outcome (24). We also repeated our analyses including all participants regardless of missing data.

In addition, we examined BCRAT's performance among women who originally participated in WHI-OS and WHI-CT separately. Although BCRAT considers history of atypical hyperplasia, information on atypia was only available for 1155 NHS participants. However, atypia was captured for 38 218 WHI-CT participants during the trial and we repeated our analyses among these women adding history of atypia (25). In addition, we repeated our analyses including women in WHI-ES with a history of cancer (excluding breast cancer).

To examine if qualitative differences in *E/O* ratios and *c*-statistics resulting from our sensitivity analyses were statistically significant, we used bootstrapping to estimate the standard error of the difference and the *z*-statistics from which we computed the *P* values.

Relative Risks

We present the relative risks associated with each BCRAT risk factor from NHS and WHI-ES calculated using multivariable logistic regression. All analyses were completed using SAS statistical software, version 9.3 (SAS Institute Inc., NC).

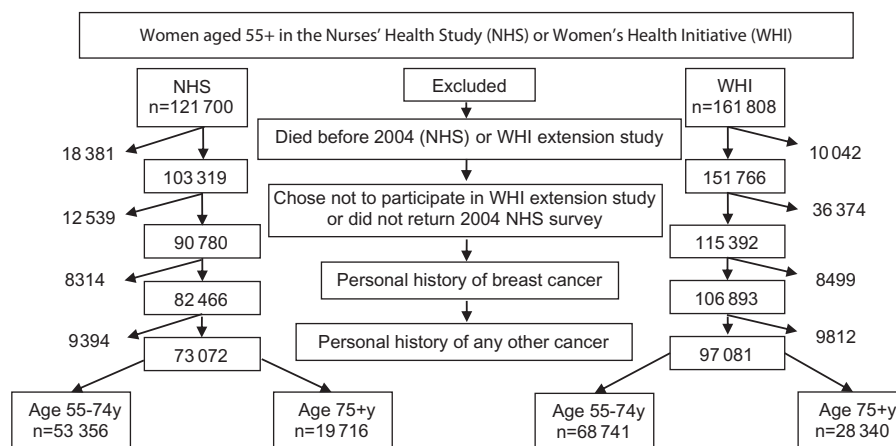


Figure 1. Sample population.

Table 1. Baseline characteristics, overall and by age, among Nurses' Health Study (n = 73 072) and the Women's Health Initiative Extension Study (n = 97 081) participants*†

Breast cancer risk assessment tool risk factors	NHS*			WHI-ES*		
	Overall (n = 73 072)	57–74‡ y (n = 53 356)	75+ y (n = 19 716)	Overall (n = 97 081)	55–64 y (n = 68 741)	75+ y (n = 28 340)
Age, mean (SD), y	70 (7)	66 (4.8)	79 (2.4)	71 (6.8)	67 (4.5)	79 (3.1)
Race						
Non-Hispanic white, %	96.2	96.1	96.3	84.7	83.0	88.8
Non-Hispanic black, %	1.8	1.8	1.7	7.4	8.3	5.1
Hispanic, %	0.9	0.9	0.9	3.8	4.5	2.4
Asian, %	0.9	0.8	0.9	2.1	2.2	1.9
Hawaiian/Pacific Islander, %	0.0	0.0	0.0	0.2	0.2	0.1
Native American, %	0.2	0.2	0.2	1.0	1.1	0.7
Unknown, %	0.0	0.0	0.0	0.8	0.7	1.0
Age at menarche, y						
≤11, %	22.2	23.3	19.4	22.1	23.7	18.2
12–13, %	57.4	57.7	56.4	55.2	54.9	55.9
≥14, %	19.7	18.3	23.3	22.4	21.1	25.5
Unknown	0.7	0.7	0.9	0.3	0.3	0.4
Age at first birth, y						
≤19, %	0.7	0.8	0.5	12.2	14.2	7.4
20–24, %	48.9	53.2	37.4	38.9	40.4	35.4
25–29, %	35.0	32.8	40.9	21.9	20.0	26.4
≥30, %	8.5	6.7	13.4	7.3	6.3	9.7
Nulliparous	5.1	4.9	5.9	2.5	2.7	1.9
Unknown	1.7	1.7	1.8	17.2	16.4	19.3
Number of biopsies						
0, %	73.6	72.6	76.2	71.9	71.9	72.0
1, %	23.2	23.8	21.5	17.5	17.5	17.4
2+, %	3.2	3.6	2.4	10.6	10.6	10.6
First-degree relatives with history of breast cancer						
0, %	82.3	83.3	79.6	86.7	87.3	85.1
1, %	15.6	15.1	17.2	12.1	11.7	13.3
2+, %	2.1	1.7	3.1	1.2	1.0	1.6
Other covariates of interest						
Mammogram in past 2 y						
Yes, %	73.2	75.2	67.8	84.6	86.3	80.4
No, %	9.5	7.8	13.9	15.2	13.5	19.4
Unknown, %	17.4	17.0	18.4	0.2	0.2	0.2
Mammogram in past 2 y excluding unknowns, %	88.5	90.6	83.0	84.7	86.4	80.6
Postmenopausal hormone therapy use						
Never, %	30.6	28.2	37.3	40.8	36.2	51.8
Past-estrogen alone	18.9	16.6	25.0	36.9	39.0	31.9
Past-estrogen plus progesterone	29.4	33.6	18.0	16.4	18.5	11.4
Past user, unknown type	4.2	2.5	9.0	0.0	0.0	0.0
Current estrogen alone user	11.3	12.3	8.5	5.5	5.8	4.8
Current estrogen plus progesterone user	5.0	6.2	1.8	0.4	0.5	0.2
Current user, unknown type	0.6	0.6	0.5	0.0	0.0	0.0
Body mass index, kg/m ² §						
<20, %	6.41	4.94	10.38	3.3	3.0	3.9
20–24, %	37.13	35.34	41.98	30.6	29.6	32.9
25–29, %	33.66	34.14	32.35	34.9	33.9	37.2
30+, %	22.66	25.45	15.10	30.4	32.6	25.0
Unknown	7.8	7.7	7.8	0.9	0.9	1.0
Oophorectomy, %						
At least 1 ovary intact	72.93	74.29	69.26	80.9	81.3	80.0
Bilateral oophorectomy	25.15	24.05	28.13	18.3	18.1	18.8
At least 1 ovary removed	1.92	1.66	2.61	0.8	0.6	1.2
Significant illnesses						
Diabetes, %	11.1	10.4	13.0	9.9	9.7	10.3
Emphysema, %	7.7	6.5	10.9	5.8	5.2	7.2
Myocardial infarction, %	2.6	1.8	4.5	3.4	2.4	5.9
Peripheral artery disease, %	0.8	0.6	1.3	1.8	1.4	2.7

Table 1. Continued

	NHS*			WHI-ES*		
	Overall (n = 73 072)	57–74‡ y (n = 53 356)	75+ y (n = 19 716)	Overall (n = 97 081)	55–64 y (n = 68 741)	75+ y (n = 28 340)
Breast cancer risk assessment tool risk factors						
Stroke, %	2.2	1.5	4.2	2.2	1.6	3.7
Congestive heart failure, %	3.1	1.9	6.4	2.2	1.4	4.0
Number of significant illnesses from above						
0, %	77.9	81.0	69.2	79.5	81.9	73.7
1, %	17.9	15.9	23.2	16.7	15.2	20.5
2+, %	4.3	3.1	7.5	3.8	2.9	5.9
Outcomes¶						
Breast cancer diagnosed during study, %	1.8	1.8	1.8	2.0	2.0	1.9
Died during study, %	6.8	3.6	15.3	5.4	3.1	11.0

* Nurses' Health Study included participants that completed the 2004 questionnaire. Women's Health Initiative Extension Study began in 2005. NHS = Nurses' Health Study; WHI-ES = Women's Health Initiative Extension Study.

† All comparisons between NHS and WHI-ES overall were statistically significant at $P < .001$ using chi-square statistics, except the P value for the difference between breast cancer incidence was $P = .02$.

‡ The youngest women in NHS at study entry were age 57 years.

§ Body mass index was based on nurse self-report in NHS and was measured in WHI.

¶ In NHS, diabetes, myocardial infarction, peripheral artery disease, and stroke were confirmed by participants and/or adjudicated by review of their medical records. Congestive heart failure, myocardial infarction, peripheral artery disease, and stroke were physician adjudicated with medical records in WHI-ES.

¶ Participants were followed for five years.

Results

Sample

We included 73 072 NHS (27.0% ≥ 75 years) and 97 081 WHI-ES participants (29.2% ≥ 75 years). Because of the large sample sizes, even small differences between NHS and WHI-ES were statistically significant (Table 1). Fewer NHS participants (1.7%) were missing data on age at first birth than WHI-ES participants (17.2%, $P < .001$). More NHS participants (17.4%) were missing data on recent mammography use than WHI-ES participants (0.2%, $P < .001$); these NHS participants completed a shorter version of the 2004 questionnaire that did not assess mammography use. Otherwise, there were small differences in missing data between cohorts on sample characteristics. WHI-ES participants were more racially/ethnically diverse (96.2% vs 84.7%, $P < .001$), younger at first birth, more likely to have had a breast biopsy, and to have a BMI of 30 kg/m² or more; they were less likely to have a family history of breast cancer and to have ever used hormone replacement therapy. NHS participants were more likely to have significant illness (22.1% vs 20.5%, $P < .001$) and to die during follow-up (6.8% vs 5.4%, $P < .001$; 15.3% vs 11.0% among women ≥ 75 years). Fewer NHS participants were diagnosed with breast cancer than WHI-ES participants (1.8% vs 2.0%, $P = .02$).

Calibration

Calibration graphs (Figure 2) demonstrate that BCRAT accurately predicted breast cancer risk in WHI-ES except for participants categorized as being at the highest risk, where BCRAT over-predicted breast cancer risk. Supplementary Table 2 (available online) presents E/O ratios and their 95% confidence intervals for each decile of risk stratified by age and cohort. BCRAT was more likely to overpredict breast cancer risk in NHS, particularly for women categorized as being at higher risk and among women age 75 years and older (Table 2). E/O ratios on average by age in WHI-ES were 1.03 (95% CI = 0.97 to 1.09) for women age 55 to 74 years and 1.10 (95% CI = 1.00 to 1.21) for women age 75 years and older ($P = .21$). In NHS, E/O ratios on average by age were 1.16 (95% CI = 1.09 to 1.23) for women age 57 to 74 years

and 1.31 (95% CI = 1.18 to 1.45) for women age 75 years and older ($P = .02$ for age comparison). Hosmer-Lemeshow tests were statistically significantly different from 1 among NHS participants and among WHI-ES participants age 55 to 74 years. The P value for the Hosmer-Lemeshow test among women age 75 years and older in WHI-ES was .05.

In both cohorts, the 95% confidence intervals of the E/O ratios crossed one for women age 75 years and older when we limited our sample to women who underwent mammography and were without significant illness (E/O ratios decreased from 1.30, 95% CI = 1.17 to 1.45, to 1.13, 95% CI = 0.98 to 1.29, $P = .12$ in NHS and from 1.10, 95% CI = 1.00 to 1.21, to 1.04, 95% CI = 0.92 to 1.16 in WHI-ES, $P = .42$). Modifying our samples did not improve E/O ratios for women age 55 to 74 years. Using 2006 to 2010 SEER incidence rates led to statistically significantly increased E/O ratios among women age 55 to 74 years in both cohorts but did not statistically change E/O ratios among women age 75 years and older.

Within WHI-ES, BCRAT's calibration was similar between WHI-OS and WHI-CT participants (Supplementary Table 4, available online). Among WHI-CT participants, including information on history of atypia did not change BCRAT's calibration (Supplementary Table 4, available online). Including women with a history of cancer also did not change calibration in WHI-ES (Supplementary Table 8, available online).

Discrimination

BCRAT c-statistics ranged between 0.56 and 0.58 in both cohorts regardless of age and regardless of which SEER breast cancer incidence rates were used. None of the sensitivity analyses led to statistically significant improvements in c-statistics. However, c-statistics tended to improve among women age 75 years and older when the outcome was limited to ER+ breast cancer (eg, in NHS, c-statistic improved from 0.57 to 0.60, $P = .21$).

Relative Risks

RRs for the interaction between family history and age at first birth tended to be higher in the BCRAT development cohort

Table 2. Breast Cancer Risk Assessment Tool's calibration (expected/observed breast cancer cases and 95% confidence intervals) among Nurses' Health Study (n = 71 293) and the Women's Health Initiative Extension Study (n = 79 611) participants

Calibration	NHS*			WHI-ES*			P age comparison
	Overall (n = 71 293)	57–74 y† (n = 52 111)	75+ y (n = 19 182)	Overall (n = 79 611)	Age 55–74 y (n = 57 009)	Age 75+ y (n = 22 602)	
Primary analyses (expected/observed ratios)‡,§							
BCRAT expected/observed ratios (95% CI)	1.20 (1.13 to 1.26)	1.16 (1.09 to 1.23)	1.31 (1.18 to 1.45)	1.05 (1.00 to 1.10)	1.03 (0.97 to 1.09)	1.10 (1.00 to 1.21)	.21
Hosmer-Lemeshow P	<.0001	<.0001	<.0001	<.0001	<.0001	.051	
Sensitivity analyses							
Limited to women who underwent mammography in past 2 y (NHS = 52 380, WHI = 67 457)	1.18 (1.11 to 1.25)	1.17 (1.09 to 1.25)	1.21 (1.07 to 1.36)	1.04 (0.98 to 1.10)	1.03 (0.96 to 1.10)	1.07 (0.97 to 1.18)	.49
Limited to women without significant illness§ (NHS = 55 567, WHI = 63 439)	1.19 (1.12 to 1.26)	1.17 (1.09 to 1.26)	1.24 (1.09 to 1.40)	1.05 (0.99 to 1.11)	1.04 (0.97 to 1.11)	1.07 (0.96 to 1.19)	.60
Limited to women who underwent mammography in past 2 y and without significant illness, (NHS = 41 887, WHI = 54 255)	1.17 (1.09 to 1.26)	1.19 (1.10 to 1.29)	1.13 (0.98 to 1.30)	1.03 (0.97 to 1.10)	1.03 (0.96 to 1.11)	1.04 (0.92 to 1.16)	.97
Sensitivity analyses (changing SEER incidence)							
Using SEER incidence rates from 2006–2010 (NHS = 71 293, WHI = 79 611)	1.27 (1.20 to 1.34)	1.25¶ (1.17 to 1.33)	1.33 (1.20 to 1.48)	1.12 (1.07 to 1.18)	1.12¶ (1.06 to 1.19)	1.12 (1.02 to 1.23)	.97

* Nurses' Health Study included participants alive in 2004. Women's Health Initiative Extension Study began in 2005. CI = confidence interval; E/O = expected/observed; NHS = Nurses' Health Study; SEER = Surveillance, Epidemiology, and End Results; WHI-ES = Women's Health Initiative Extension Study.

† The youngest women in NHS in 2004 were age 57 years.

‡ We compared the expected number of breast cancers based on Breast Cancer Risk Assessment Tool (BCRAT) estimates (calculated using the BCRAT SAS macro) to the observed number in each cohort stratified by age (55–74, 75+ years) (7). To determine the 95% CI of the E/O ratios, we used the Poisson variance of the logarithm of the observed number of cases (19). To test whether E/O ratio estimates differed by age within cohort, we used the normal approximation z-test.

§ Significant illnesses included: history of diabetes, myocardial infarction, congestive heart failure, stroke, emphysema, and peripheral artery disease.

|| For our primary analyses, we used SEER breast cancer incidence rates from 1983–1987 for whites, from 1994–1998 for blacks, from 1990–1996 for Hispanics, and from 1988–2002 for Asians. In sensitivity analyses, we used SEER incidence rates from 2006–2010 for whites, blacks, and Hispanics.

¶ This represents a statistically significant change (P < .05) from the results of the primary analyses. To examine if qualitative differences in E/O ratios resulting from our sensitivity analyses were statistically different than in primary analyses, we used bootstrapping to estimate the standard error of the difference and the z-statistics from which we computed the P values.

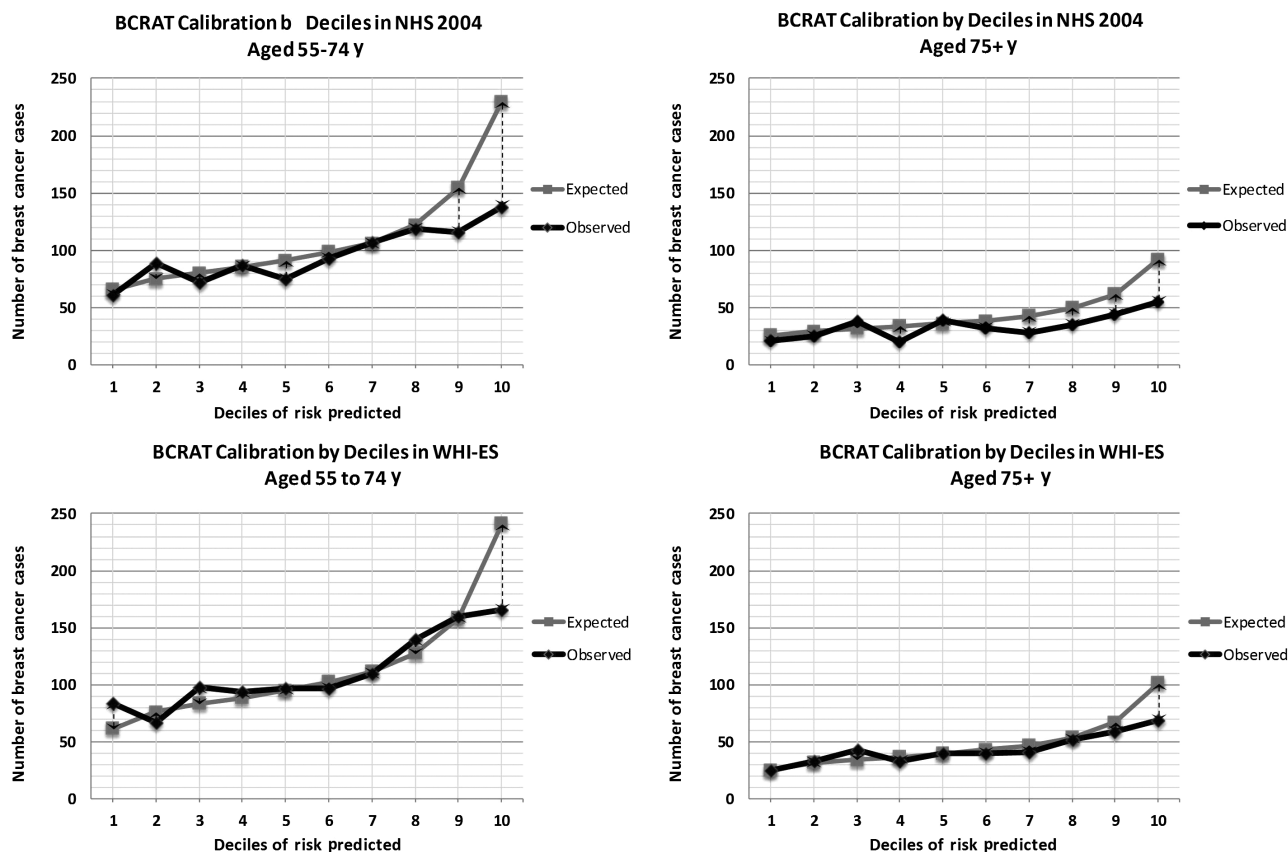


Figure 2. Calibration by decile of risk and by age among participants in the Nurses' Health Study ($n = 73\,072$) and the Women's Health Initiative Extension Study ($n = 97\,081$).

(Breast Cancer Demonstration and Detection Project, BCDDP) (10) than in NHS or WHI-ES (Table 4). Age at menarche was not a statistically significant breast cancer risk factor for WHI-ES or for NHS participants age 75 years and older.

Discussion

In WHI-ES, BCRAT accurately predicted five-year breast cancer risk for women age 55 to 74 years and overpredicted breast cancer risk by 10% on average among participants age 75 years and older. In NHS, BCRAT overpredicted breast cancer risk by 16% on average among women age 57 to 74 years and by 31% on average among women age 75 years and older. In both cohorts, BCRAT was most likely to overpredict breast cancer among women categorized as being at higher risk. BCRAT's prediction accuracy improved among women age 75 years and older when the samples were limited to women who underwent mammography and were without significant illness, suggesting that BCRAT may be most appropriate to use among older women with these characteristics. BCRAT's discrimination was modest in both cohorts and age groups, with c -statistics ranging between 0.56 and 0.58.

BCRAT's performance has previously been tested among NHS participants using data from: 1) 1976 (26), 2) 1982 (26), and 3) 1992 (27) (when participants were age 30–55, 36–61, and 45–71 years, respectively). In the first two analyses, the model was used to predict invasive and noninvasive breast cancer and baseline breast cancer incidence was estimated using BCDDP data. In the third analysis, BCRAT was used to predict only invasive breast cancer and 1987 SEER incidence rates were used to estimate baseline breast cancer incidence. In the first two

analyses, BCRAT was found to overestimate breast cancer risk by 33%, which was attributed to greater use of mammography by BCDDP participants than NHS participants. Because mammography may find breast cancer before symptoms develop and may even find some cancers that would never have caused problems, mammography may lead to an increased estimated risk among those who are screened (28,29). In the third analysis, BCRAT was found to underestimate breast cancer by 6%, which was attributed to greater use of mammography by NHS participants than among SEER participants. Similar to our findings, BCRAT's c -statistic in the third analysis was 0.58 (95% CI = 0.56 to 0.60) (27).

In addition, BCRAT's performance was tested among WHI participants beginning at enrollment (between 1993–1998) (24). In that study, BCRAT was found to underestimate breast cancer by 20%, which was attributed to greater use of mammography and breast biopsies among WHI participants than in SEER. That study also found that BCRAT's c -statistic was 0.58 and improved to 0.60 when the outcome was limited to ER+ breast cancer. This finding was attributed to the fact that older women are more likely to develop ER+ breast cancer (30) and may explain why we also found that BCRAT tended to show better discrimination among women age 75 years and older when predicting ER+ breast cancer. BCRAT's performance has also been tested in other cohorts of older women. In the National Institutes of Health–AARP Diet and Health Study (recruited women age 50–71 years between 1995–1996) and the Prostate, Lung, Colorectal, and Ovarian Cancer Screening trial (recruited women age 55–74 years between 1993–2001), BCRAT underpredicted breast cancer by 13% to 14%, which was attributed to use of

Table 3. Breast Cancer Risk Assessment Tool's discrimination (c-statistics and 95% confidence intervals) among Nurses' Health Study (n = 71 293) and the Women's Health Initiative Extension study (n = 79 611) participants

Discrimination	NHS*			WHI-ES*				
	Overall (n = 71 293)	57-74 y† (n = 52 111)	75+ y (n = 19 182)	P† age comparison	Overall (n = 79 611)	55-74 y (n = 57 009)	75+ y (n = 22 602)	P† age comparison
Primary analyses‡,§	0.565 (0.549 to 0.580)	0.564 (0.546 to 0.582)	0.569 (0.538 to 0.600)	.78	0.578 (0.563 to 0.592)	0.577 (0.560 to 0.594)	0.579 (0.552 to 0.606)	.91
BCRAT c-statistic (95% CI)								
Sensitivity analyses	0.567 (0.549 to 0.585)	0.567 (0.547 to 0.588)	0.564 (0.529 to 0.599)	.86	0.582 (0.567 to 0.597)	0.581 (0.563 to 0.599)	0.585 (0.556 to 0.614)	.82
Limited to women who underwent mammography in past 2 y (NHS = 52 380, WHI = 67 457)								
Limited to women without significant illness (NHS = 55 567, WHI = 63 437)	0.563 (0.545 to 0.581)	0.563 (0.543 to 0.584)	0.563 (0.527 to 0.599)	1.00	0.571 (0.555 to 0.587)	0.573 (0.554 to 0.592)	0.563 (0.532 to 0.595)	.60
Limiting outcome to estrogen receptor-positive breast cancers (NHS = 70 295, WHI = 79 343)	0.568 (0.550 to 0.587)	0.562 (0.541 to 0.583)	0.598 (0.562 to 0.635)	.09	0.578 (0.563 to 0.594)	0.576 (0.557 to 0.594)	0.586 ² (0.556 to 0.616)	.56
Sensitivity analyses (changing SEER incidence)§								
Using SEER incidence rates from 2006-2010 (NHS = 71 293, WHI = 79 611)	0.567 (0.551 to 0.583)	0.566 (0.548 to 0.585)	0.569 (0.538 to 0.599)	.90	0.575 (0.561 to 0.590)	0.575 (0.559 to 0.592)	0.576 (0.549 to 0.603)	.99

* Nurses' Health Study included participants alive in 2004. Women's Health Initiative Extension Study began in 2005. CI = confidence interval; NHS = Nurses' Health Study; SEER = Surveillance, Epidemiology, and End Results; WHI-ES = Women's Health Initiative Extension Study.

† The youngest women in NHS in 2004 were age 57 years.

‡ To assess discrimination in each cohort, we calculated the c-statistic or area under the receiver operating characteristic (ROC) curve and its standard error (21). To test if c-statistics differed by age within cohort, we used the normal approximation z-test.

§ For our primary analyses, we used SEER breast cancer incidence rates from 1983-1987 for whites, from 1994-1998 for blacks, from 1990-1996 for Hispanics, and from 1988-2002 for Asians. In sensitivity analyses, we used SEER incidence rates from 2006-2010 for whites, blacks, and Hispanics.

|| Significant illnesses included: history of diabetes, myocardial infarction, congestive heart failure, stroke, emphysema, and peripheral artery disease.

Table 4. Relative risks and 95% confidence intervals for Breast Cancer Risk Assessment Tool risk factors from BCDDP, NHS, and WHI-ES

Risk factors	BCDDP*	NHS*,†	WHI-ES*,†	NHS*,†	WHI-ES*,†	NHS*,†	WHI-ES*,†
	Overall	Overall (n = 71 293)	Overall (n = 79 611)	55–74 y‡ (n = 52 111)	55–74 y (n = 57 009)	75+ y (n = 19 182)	75+ y (n = 22 602)
Age at menarche, y							
≤11, %	1.21	1.27 (1.07 to 1.50)	1.08 (0.93 to 1.25)	1.38 (1.14 to 1.68)	1.02 (0.86 to 1.21)	0.99 (0.72 to 1.37)	1.22 (0.92 to 1.62)
12–13, %	1.10	1.13 (1.04 to 1.22)	1.04 (0.96 to 1.12)	1.17 (1.07 to 1.30)	1.01 (0.93 to 1.10)	0.99 (0.85 to 1.17)	1.10 (0.96 to 1.27)
≥14, %	1.0	1.0	1.0	1.0	1.0	1.0	1.0
Age at first birth, y X 0,1,2+ first degree relatives							
≤19 and 0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
≤19 and 1	2.61	1.46 (1.12 to 1.91)	1.48 (1.21 to 1.81)	1.33 (0.96 to 1.84)	1.47 (1.15 to 1.86)	1.87 (1.16 to 3.00)	1.50 (1.01 to 2.22)
≤19 and 2+	6.80	2.13 (1.25 to 3.64)	2.18 (1.45 to 3.29)	1.77 (0.92 to 3.40)	2.15 (1.33 to 3.48)	3.50 (1.35, 9.02)	2.24 (1.02 to 4.91)
20–24 and 0	1.24	1.11 (1.01 to 1.22)	1.13 (1.06 to 1.21)	1.13 (1.02 to 1.26)	1.14 (1.06 to 1.23)	1.09 (0.91 to 1.30)	1.12 (0.98 to 1.27)
20–24 and 1	2.68	1.58 (1.30 to 1.90)	1.67 (1.44 to 1.92)	1.51 (1.21 to 1.89)	1.71 (1.44 to 2.01)	1.82 (1.26 to 2.62)	1.61 (1.22 to 2.14)
20–24 and 2+	5.78	2.24 (1.63 to 3.07)	2.45 (1.91 to 3.16)	2.01 (1.38 to 2.94)	2.54 (1.89 to 3.42)	3.05 (1.67 to 5.56)	2.33 (1.42 to 3.80)
25–29/nulliparous§ and 0	1.55	1.23 (1.03 to 1.48)	1.28 (1.13 to 1.46)	1.28 (1.03 to 1.59)	1.31 (1.12 to 1.52)	1.18 (0.83 to 1.69)	1.25 (0.97 to 1.60)
25–29/nulliparous and 1	2.76	1.70 (1.40 to 2.06)	1.88 (1.61 to 2.20)	1.71 (1.36 to 2.16)	1.98 (1.64 to 2.40)	1.77 (1.22 to 2.58)	1.74 (1.29 to 2.33)
25–29/nulliparous and 2+	4.91	2.34 (1.80 to 3.05)	2.75 (2.13 to 3.56)	2.29 (1.65 to 3.18)	3.00 (2.19 to 4.12)	2.65 (1.65 to 4.27)	2.41 (1.54 to 3.78)
≥30 and 0	1.93	1.37 (1.04 to 1.80)	1.46 (1.20 to 1.77)	1.46 (1.05 to 2.01)	1.49 (1.19 to 1.88)	1.29 (0.75 to 2.20)	1.40 (0.96 to 2.04)
≥30 and 1	2.83	1.83 (1.39 to 2.41)	2.12 (1.68 to 2.68)	1.95 (1.38 to 2.75)	2.30 (1.72 to 3.07)	1.72 (1.04 to 2.85)	1.87 (1.23 to 2.85)
≥30 and 2+	4.17	2.46 (1.59 to 3.79)	3.09 (2.04 to 4.68)	2.60 (1.47 to 4.60)	3.55 (2.11 to 5.96)	2.31 (1.14 to 4.67)	2.51 (1.23 to 5.09)
No. of biopsies							
0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
1	1.27	1.32 (1.21 to 1.45)	1.27 (1.19 to 1.36)	1.33 (1.20 to 1.48)	1.25 (1.15 to 1.35)	1.29 (1.07 to 1.57)	1.33 (1.18 to 1.50)
2+	1.62	1.75 (1.45 to 2.10)	1.61 (1.41 to 1.84)	1.77 (1.43 to 2.18)	1.55 (1.33 to 1.82)	1.67 (1.13 to 2.46)	1.77 (1.38 to 2.26)

* Nurses' Health Study began in 2004. Women's Health Initiative Extension Study began in 2005. BCDDP = Breast Cancer Demonstration and Detection Project (31); CI = confidence interval; NHS = Nurses' Health Study; RR = relative risk; WHI-ES = Women's Health Initiative Extension Study.

† The relative risks for Gail model risk factors were estimated from logistic regression models. The Gail model treated each variable (as it is categorized) as continuous in its model.

‡ The youngest women in NHS in 2004 were age 57 years.

§ No separate indicator was used for nulliparous women.

SEER incidence rates from 1983 to 1987 (21). BCRAT's calibration improved when 1995 to 2003 SEER incidence rates were used, which corresponded to the time period of that study. BCRAT's c-statistic in these cohorts was 0.58 to 0.59 (22,24). These studies and our data suggest that BCRAT has modest discrimination in predicting breast cancer among postmenopausal women.

Interestingly, we found BCRAT's calibration to be better in WHI-ES than NHS. This is likely because breast cancer incidence in WHI-ES (2.0%) was more similar to SEER (Supplementary Table 5, available online). Breast cancer incidence may have been lower in NHS (1.8%) because NHS participants were less likely to undergo breast biopsy, were less likely to be obese, and were more likely to undergo oophorectomy. Also, NHS participants, particularly those age 75 years and older, were more likely to

have significant illness and to die. Competing mortality risks may have prevented these women from undergoing mammography and as a result these women would have been less likely to have an early-stage breast cancer detected (Supplementary Table 6, available online). Also, 1069 NHS participants (1.5% of our sample) remained alive during our study but did not complete a follow-up questionnaire. Breast cancer may have been missed among these women. If we assume that 2.0% of these women were diagnosed with breast cancer, then our overall incidence in NHS would increase to 1.9%. Only 111 of the WHI participants that remained alive during our study period did not complete a follow-up questionnaire.

While we found that BCRAT provides accurate probabilities of five-year breast cancer risk for older women without

significant illness who undergo mammography, it has modest discrimination. Including additional breast cancer risk factors in BCRAT such as breast density has been shown to improve discrimination modestly (31); unfortunately, our data do not include adequate information on participant breast density. Considering risk factors that are important in predicting late-life breast cancer (eg, obesity) may also be needed to improve BCRAT's discrimination among older women. Also, the RRs associated with some of the risk factors in BCRAT, such as family history, may no longer be accurate. Broad use of mammography has led to increased detection of early-stage breast cancers and may have attenuated the impact of a family history of breast cancer on late-life breast cancer risk.

Our study has limitations. Family history of breast cancer and diagnosis of emphysema (for WHI-CT participants only) were obtained on average 8.5 years before start of the WHI-ES and therefore may be underestimated in WHI-ES. Although 17% of WHI-ES participants were missing data on age at first birth, when we repeated our analyses including these women our results did not change (Supplementary Table 8, available online).

In summary, BCRAT provides accurate five-year probabilities of breast cancer among women age 75 years and older who undergo mammography and are without significant illness but tends to overpredict breast cancer among postmenopausal women categorized at the highest risk of breast cancer. Discrimination of the model is modest among all postmenopausal women. Incorporating other aspects of personalized breast cancer risk including lifelong estrogen exposure (eg, from obesity) and competing mortality risks may improve model performance.

Funding

This work was supported by the National Institute on Aging at the National Institutes of Health (R01 AG041860), a National Health Service (NHS) cohort infrastructure grant (UM1 CA186107), and an NHS program project grant (P01 CA87969).

Notes

We would like to thank the participants and staff of the Nurses' Health Study for their valuable contributions as well as the following state cancer registries for their help: AL, AZ, AR, CA, CO, CT, DE, FL, GA, ID, IL, IN, IA, KY, LA, ME, MD, MA, MI, NE, NH, NJ, NY, NC, ND, OH, OK, OR, PA, RI, SC, TN, TX, VA, WA, WY. In addition, this study was approved by the Connecticut Department of Public Health (DPH) Human Investigations Committee. Certain data used in this publication were obtained from the DPH.

The Women's Health Initiative (WHI) program is funded by the National Heart, Lung, and Blood Institute, National Institutes of Health, and US Department of Health and Human Services through contracts HHSN268201100046C, HHSN268201100001C, HHSN268201100002C, HHSN268201100003C, HHSN268201100004C, and HHSN271201100004C.

We would also like to thank the following WHI investigators for their help with this project: Program Office: (National Heart, Lung, and Blood Institute, Bethesda, MD) Jacques Rossouw, Shari Ludlam, Dale Burwen, Joan McGowan, Leslie Ford, and Nancy Geller; Clinical Coordinating Center: (Fred Hutchinson Cancer Research Center, Seattle, WA) Garnet Anderson, Ross Prentice, Andrea LaCroix, and Charles Kooperberg; Investigators and Academic Centers: (Brigham and Women's Hospital, Harvard Medical School, Boston, MA) JoAnn E. Manson; (MedStar Health Research Institute/Howard University, Washington, DC) Barbara V. Howard; (Stanford Prevention Research Center, Stanford,

CA) Marcia L. Stefanick; (The Ohio State University, Columbus, OH) Rebecca Jackson; (University of Arizona, Tucson/Phoenix, AZ) Cynthia A. Thomson; (University at Buffalo, Buffalo, NY) Jean Wactawski-Wende; (University of Florida, Gainesville/Jacksonville, FL) Marian Limacher; (University of Iowa, Iowa City/Davenport, IA) Robert Wallace; (University of Pittsburgh, Pittsburgh, PA) Lewis Kuller; (Wake Forest University School of Medicine, Winston-Salem, NC) Sally Shumaker.

The sponsor had no role in the design of the study; the collection, analysis, or interpretation of the data; the writing of the manuscript; or the decision to submit the manuscript for publication.

References

1. Data on Seer (Surveillance, Epidemiology and End Results) cancer statistics. <http://seer.cancer.gov>. Accessed on September 10, 2014.
2. The benefits and harms of breast cancer screening: an independent review. *Lancet*. 2012;380(9855):1778–1786.
3. Breast cancer screening in older women. American Geriatrics Society Clinical Practice Committee. *J Am Geriatr Soc*. 2000;48(7):842–844.
4. Walter LC, Schonberg MA. Screening mammography in older women: a review. *JAMA*. 2014;311(13):1336–1347.
5. Yourman LC, Lee SJ, Schonberg MA, Widera EW, Smith AK. Prognostic indices for older adults: a systematic review. *JAMA*. 2012;307(2):182–192.
6. Schonberg MA, Hamel MB, Davis RB, et al. Development and evaluation of a decision aid on mammography screening for women 75 years and older. *JAMA Intern Med*. 2014;174(3):417–424.
7. Breast Cancer Risk Assessment SAS Macro (Gail Model). <http://dceg.cancer.gov/tools/risk-assessment/bcrasasmacro>. Accessed September 10, 2014.
8. Elmore JG, Fletcher SW. The risk of cancer risk prediction: "What is my risk of getting breast cancer"? *J Natl Cancer Inst*. 2006;98(23):1673–1675.
9. Amir E, Freedman OC, Seruga B, et al. Assessing women at high risk of breast cancer: a review of risk assessment models. *J Natl Cancer Inst*. 2010;102(10):680–691.
10. Gail MH, Brinton LA, Byar DP, et al. Projecting individualized probabilities of developing breast cancer for white females who are being examined annually. *J Natl Cancer Inst*. 1989;81(24):1879–1886.
11. Baker LH. Breast Cancer Detection Demonstration Project: five-year summary report. *CA Cancer J Clin*. 1982;32(4):194–225.
12. Anderson GL, Manson J, Wallace R, et al. Implementation of the Women's Health Initiative study design. *Ann Epidemiol*. 2003;13(9):S5–S17.
13. Hays J, Hunt JR, Hubbell FA, et al. The Women's Health Initiative recruitment methods and results. *Ann Epidemiol*. 2003;13(9):S18–S77.
14. Colditz GA, Manson JE, Hankinson SE. The Nurses' Health Study: 20-year contribution to the understanding of health among women. *J Womens Health*. 1997;6(1):49–62.
15. Chlebowski RT, Kuller LH, Prentice RL, et al. Breast cancer after use of estrogen plus progestin in postmenopausal women. *N Engl J Med*. 2009;360(6):573–587.
16. Rossouw JE, Anderson GL, Prentice RL, et al. Risks and benefits of estrogen plus progestin in healthy postmenopausal women: principal results From the Women's Health Initiative randomized controlled trial. *JAMA*. 2002;288(3):321–333.
17. Colditz GA, Martin P, Stampfer MJ, et al. Validation of questionnaire information on risk factors and disease outcomes in a prospective cohort study of women. *Am J Epidemiol*. 1986;123(5):894–900.
18. Justice AC, Covinsky KE, Berlin JA. Assessing the generalizability of prognostic information. *Ann Intern Med*. 1999;130(6):515–524.
19. Daly L. Simple SAS macros for the calculation of exact binomial and Poisson Confidence limits. *Comput Biol Med*. 1992;22(5):351–361.
20. Hosmer DW, Hosmer T, Le Cessie S, Lemeshow S. A comparison of goodness-of-fit tests for the logistic regression model. *Stat Med*. 1997;16(9):965–980.
21. Rosner B, Glynn RJ. Power and sample size estimation for the Wilcoxon rank sum test with application to comparisons of C statistics from alternative prediction models. *Biometrics*. 2009;65(1):188–197.
22. Schonfeld SJ, Pee D, Greenlee RT, et al. Effect of changing breast cancer incidence rates on the calibration of the Gail model. *J Clin Oncol*. 2010;28(14):2411–2417.
23. Bondy ML, Lustbader ED, Halabi S, Ross E, Vogel VG. Validation of a breast cancer risk assessment model in women with a positive family history. *J Natl Cancer Inst*. 1994;86(8):620–625.
24. Chlebowski RT, Anderson GL, Lane DS, et al. Predicting risk of breast cancer in postmenopausal women by hormone receptor status. *J Natl Cancer Inst*. 2007;99(22):1695–1705.
25. Rohan TE, Negassa A, Chlebowski RT, et al. Conjugated equine estrogen and risk of benign proliferative breast disease: a randomized controlled trial. *J Natl Cancer Inst*. 2008;100(8):563–571.

26. Spiegelman D, Colditz GA, Hunter D, Hertzmark E. Validation of the Gail et al. model for predicting individual breast cancer risk. *J Natl Cancer Inst.* 1994;86(8):600-607.
27. Rockhill B, Spiegelman D, Byrne C, Hunter DJ, Colditz GA. Validation of the Gail et al. model of breast cancer risk prediction and implications for chemoprevention. *J Natl Cancer Inst.* 2001;93(5):358-366.
28. Schonberg MA, McCarthy EP, Davis RB, Phillips RS, Hamel MB. Breast cancer screening in women aged 80 and older: results from a national survey. *J Am Geriatr Soc.* 2004;52(10):1688-1695.
29. Cook NR, Rosner BA, Hankinson SE, et al. Mammographic screening and risk factors for breast cancer. *Am J Epidemiol.* 2009;170(11):1422-1432.
30. Schonberg MA, Marcantonio ER, Li D, et al. Breast cancer among the oldest old: tumor characteristics, treatment choices, and survival. *J Clin Oncol.* 2010;28(12):2038-2045.
31. Chen J, Pee D, Ayyagari R et al. Projecting absolute invasive breast cancer risk in white women with a model that includes mammographic density. *J Natl Cancer Inst.* 2006;98(17):1215-1225.