

Evaluating breast cancer risk projections for Hispanic women

Matthew P. Banegas · Mitchell H. Gail · Andrea LaCroix · Beti Thompson ·
Maria Elena Martinez · Jean Wactawski-Wende · Esther M. John ·
F. Allan Hubbell · Shagufta Yasmeen · Hormuzd A. Katki

Received: 22 November 2011 / Accepted: 23 November 2011 / Published online: 7 December 2011
© Springer Science+Business Media, LLC. 2011

Abstract For Hispanic women, the Breast Cancer Risk Assessment Tool (BCRAT; “Gail Model”) combines 1990–1996 breast cancer incidence for Hispanic women with relative risks for breast cancer risk factors from non-Hispanic white (NHW) women. BCRAT risk projections have never been comprehensively evaluated for Hispanic women. We compared the relative risks and calibration of BCRAT risk projections for 6,353 Hispanic to 128,976 NHW postmenopausal participants aged 50 and older in the Women’s Health Initiative (WHI). Calibration was

assessed by the ratio of the number of breast cancers observed with that expected by the BCRAT (O/E). We re-evaluated calibration for an updated BCRAT that combined BCRAT relative risks with 1993–2007 breast cancer incidence that is contemporaneous with the WHI. Cox regression was used to estimate relative risks. Discriminatory accuracy was assessed using the concordance statistic (AUC). In the WHI Main Study, the BCRAT underestimated the number of breast cancers by 18% in both Hispanics ($O/E = 1.18$, $P = 0.06$) and NHWs

Electronic supplementary material The online version of this article (doi:10.1007/s10549-011-1900-9) contains supplementary material, which is available to authorized users.

M. P. Banegas · B. Thompson
School of Public Health, Department of Health Services,
University of Washington, Box 357660, Seattle, WA 98195,
USA

M. P. Banegas (✉) · A. LaCroix · B. Thompson
Public Health Sciences Division, Fred Hutchinson Cancer
Research Center, P.O. Box 19024, M3-B232, Seattle, WA
98109, USA
e-mail: mbanegas@fhcrc.org

M. H. Gail
Division of Cancer Epidemiology and Genetics, National Cancer
Institute, National Institutes of Health, DHHS, 6120 Executive
Blvd., EPS MSC 7244, Bethesda, MD 20892-7234, USA

M. E. Martinez
Arizona Cancer Center and Mel and Enid Zuckerman College of
Public Health, University of Arizona, 1515 N Campbell Ave,
Box 245024, Tucson, AZ 85724, USA

J. Wactawski-Wende
School of Public Health and Health Professions, Department of
Social and Preventive Medicine, University at Buffalo, Farber
Hall, Rm. 270, Buffalo, NY 14214, USA

E. M. John
Cancer Prevention Institute of California, 2201 Walnut Avenue,
Suite 300, Fremont, CA 94538, USA

E. M. John
Stanford University School of Medicine and Stanford Cancer
Institute, Stanford, CA, USA

F. A. Hubbell
School of Medicine, University of California Irvine, 264 Irvine
Hall, 1001 Health Sciences Road, Mail Code 3950, Irvine, CA
92697, USA

S. Yasmeen
School of Medicine, University of California, Davis, Lawrence J.
Ellison Ambulatory Care Center, 4860 Y St., Suite 2500,
Sacramento, CA 95817, USA

H. A. Katki (✉)
Division of Cancer Epidemiology and Genetics, National Cancer
Institute, National Institutes of Health, DHHS, 6120 Executive
Blvd. Room 8014, EPS MSC 7244, Bethesda, MD 20892-7234,
USA
e-mail: katkih@mail.nih.gov

($O/E = 1.18$, $P < 0.001$). Updating the BCRAT improved calibration for Hispanic women ($O/E = 1.08$, $P = 0.4$) and NHW women ($O/E = 0.98$, $P = 0.2$). For Hispanic women, relative risks for number of breast biopsies (1.71 vs. 1.27, $P = 0.03$) and age at first birth (0.97 vs. 1.24, $P = 0.02$) differed between the WHI and BCRAT. The AUC was higher for Hispanic women than NHW women (0.63 vs. 0.58, $P = 0.03$). Updating the BCRAT with contemporaneous breast cancer incidence rates improved calibration in the WHI. The modest discriminatory accuracy of the BCRAT for Hispanic women might improve by using risk factor relative risks specific to Hispanic women.

Keywords Hispanic · Breast cancer · Risk prediction · Risk assessment · BCRAT

Abbreviations

BCRAT Breast Cancer Risk Assessment Tool
 NHW Non-Hispanic white
 SEER Surveillance, epidemiology, and end results
 WHI Women's Health Initiative

Introduction

Breast cancer is the most commonly diagnosed cancer and leading cause of cancer-related death among Hispanic women in the United States (US) [1, 2]. The Breast Cancer Risk Assessment Tool (BCRAT), also known as the “Gail model” [3], estimates a woman's risk of developing invasive breast cancer over a defined period of time, given her age and risk factor profile [3]. For Hispanic women, the BCRAT combines 1990–1996 breast cancer incidence rates from the Surveillance, Epidemiology, and End Results (SEER) program for Hispanic women with relative risks for breast cancer risk factors from non-Hispanic white (NHW) women in the Breast Cancer Detection Demonstration Project [3]. Several case-control studies suggest that relative risks for family history, reproductive, and other factors may differ between Hispanic and NHW women [4–8], calling into question the relative risk assumptions of the BCRAT for Hispanics.

Although BCRAT risk projections have been extensively validated for NHW women [9–12], we are unaware of any study that has comprehensively examined the performance of the BCRAT in a cohort of US Hispanic women. In particular, prospective follow-up is required to evaluate *calibration*, i.e., the similarity between the number of breast cancers that develop in a population with that expected by the BCRAT [13]. Adequate calibration is crucial to ensure the

validity of BCRAT-based risk thresholds used in clinical practice, such as the American Society of Clinical Oncology guideline that women with a 5-year breast cancer risk greater than 1.66% may benefit from using tamoxifen or raloxifene to prevent breast cancer [14]. Calibration may be adversely affected by changes in US breast cancer incidence, which increased through the 1990's, peaked around 2002, and has since declined [15]. For instance, the BCRAT for NHW women, which uses SEER breast cancer incidence from 1983 to 1987, underestimates breast cancer incidence in many cohorts established in the 1990s [16]. In particular, the Women's Health Initiative (WHI), when combining women of all races and ethnicities, found 20% underestimation by the BCRAT [17]. It has been shown that calibration can be improved by using breast cancer incidence rates contemporaneous to the study cohort [16].

The WHI is one of the few prospective studies with a large cohort of US Hispanic women. Since the Hispanic BCRAT uses relative risks from NHW women, we compared the relative risks, calibration, and discriminatory accuracy of the BCRAT for 6,353 Hispanic participants to 128,976 NHW postmenopausal participants aged 50 and older in the WHI. We also assessed whether updating BCRAT breast cancer incidence rates to SEER rates contemporaneous with the WHI study period (1993–2007) improved calibration.

Methods

Study design

The design of the WHI study has been previously described [18–20]. In brief, the WHI was a national, longitudinal health study composed of a set of randomized clinical trials (CT) and an observational study (OS). We used data on 6,353 Hispanic and 128,976 NHW postmenopausal women aged 50–79 without a history of breast cancer or mastectomy (bilateral or unilateral) at enrollment who were followed through March 2005 (WHI Main Study Period). BCRAT risk factors were obtained from the enrollment questionnaire. Mammograms and clinical breast exams were obtained at least biennially for CT participants but not necessarily for OS participants [20]. All reported invasive breast cancers were adjudicated locally and again centrally by physician adjudicators [4, 21].

Breast Cancer Risk Assessment Tool

The BCRAT [3, 12] estimates women's absolute risk of developing invasive breast cancer using age, age at first live birth, age at menarche, number of first-degree relatives

with breast cancer, number of breast biopsies (which is also modified by age), and presence of atypical hyperplasia on a previous breast biopsy. Information on atypical hyperplasia was unavailable in the WHI. When information on a risk factor is missing for a woman, BCRAT imputes the safest level of that risk factor. For both Hispanic and NHW women, relative risks for the model risk factors are based on NHW women in the Breast Cancer Detection Demonstration Project [3]. For NHW women, the BCRAT is calibrated to 1983–1987 SEER invasive breast cancer incidence rates for white (not NHW) women. For Hispanic women, the BCRAT is calibrated to 1990–1996 SEER invasive breast cancer incidence rates for Hispanic women. We also updated the BCRAT (“Updated BCRAT”) by combining BCRAT relative risks with SEER breast cancer incidence from 1993 to 2007 for Hispanic and NHW women, respectively, to ensure that breast cancer incidence rates overlap in time with the WHI.

Statistical analyses

We compared the BCRAT relative risks (RRs) to those estimated for Hispanic and NHW women in the WHI, separately, using Cox proportional hazards models. We assessed the discriminatory accuracy of the models with the concordance statistic, or area-under-the-curve (AUC) statistic [22]. Using the BCRAT and the Updated BCRAT, we computed each woman’s absolute risk of developing invasive breast cancer from enrollment through 2005. Projections were limited to age 90, since the BCRAT does not project risk past that age. We summed absolute risks over women in each risk factor category i , and also overall, to calculate the expected count (E_i), which was compared with the corresponding observed number of women with incident invasive breast cancer, O_i . For each category, we calculated an observed/expected (O/E) ratio and 95% confidence interval (CI) with a lower limit of $(O/E)\exp(-1.96 \times O^{-1/2})$ and upper limit of $(O/E)\exp(+1.96 \times O^{-1/2})$. Analyses were done separately for the BCRAT and the Updated BCRAT.

Results

Compared with NHW women in the WHI, Hispanics women were 3.3 years younger at baseline, 4.4 years younger at breast cancer diagnosis, reported less family history of breast cancer, and fewer breast biopsies (Table 1). Relative risks (Table 2) for Hispanics in the WHI differed from those in the BCRAT for number of breast biopsies (RR = 1.71 vs. 1.27, $P = 0.03$) and age at first live birth (RR = 0.97 vs. 1.24, $P = 0.02$). For NHW women, the BCRAT differed for number of first-degree relatives with breast cancer (RR = 1.31 vs. 2.61, $P < 0.001$), age at first live birth

(RR = 1.13 vs. 1.24, $P < 0.001$), and the interaction between family history and age at first live birth (RR = 1.01 vs. 0.83, $P < 0.001$). The only relative risk estimate that significantly differed between NHW and Hispanic women in the WHI was number of breast biopsies (RR = 1.27 vs. 1.71, $P = 0.03$). The concordance statistic (AUC) of the BCRAT was 0.63 [95% CI: 0.582–0.676] for Hispanic women and 0.58 [95%: 0.566–0.583] for NHW women in the WHI ($P = 0.03$).

The BCRAT underestimated the number of breast cancer diagnoses among Hispanics by 18% ($O/E = 1.18$, 95% CI = 0.99–1.40; $P = 0.06$) (Table 3). Underestimation occurred for both the CT and OS, as well as in most BCRAT risk factor categories (Supplemental Table 1). For NHW women, the BCRAT also underestimated the number of breast cancer diagnoses by 18% ($O/E = 1.18$, 95% CI = 1.14–1.21; $P < 0.001$).

The age-adjusted 1990–1996 SEER breast cancer incidence rate for Hispanic women used by the BCRAT is 7.7% lower than the SEER breast cancer incidence rate for Hispanic women during the WHI study period 1993–2007 (217 vs. 235/100,000 women/year; $P < 0.001$) (Fig. 1). The age-adjusted 1983–1987 SEER breast cancer incidence rate for white women used by the BCRAT is 14.5% lower than the SEER breast cancer incidence rate for NHW women during the WHI study period 1993–2007 (336 vs. 393/100,000 women/year; $P < 0.001$). The Updated BCRAT, using 1993–2007 SEER incidence rates, was well-calibrated for both Hispanic women (O/E ratio = 1.08, 95% CI = 0.91–1.28; $P = 0.4$) and NHW women ($O/E = 0.98$, 95% CI = 0.96–1.01; $P = 0.2$) (Table 3). The Updated BCRAT showed improved calibration in nearly all risk factor categories for Hispanic and NHW women, except for NHW women with first-degree relatives with breast cancer (Supplemental Table 1).

Discussion

In this comprehensive evaluation of the BCRAT for Hispanic women using prospective cohort data, we found that the BCRAT underestimated the number of invasive breast cancers by 18% for both Hispanic and NHW women. The calibration of the BCRAT improved greatly by updating the older SEER breast cancer incidence rates used by BCRAT to rates contemporaneous with the WHI (1993–2007). BCRAT relative risk estimates for number of breast biopsies and age at first live birth differed with estimates for Hispanic women from the WHI. The discriminatory accuracy of the BCRAT for Hispanic women was higher than for NHW women, but remained only modest.

Our findings suggest two potential improvements for the BCRAT. Although breast cancer incidence rates in 2007

Table 1 Distribution of BCRAT risk factors and breast cancer outcomes in the WHI

Characteristic	Hispanic (<i>n</i> = 6,353)		Non-Hispanic white (<i>n</i> = 128,976)		<i>P</i>
BCRAT risk factors ^a					
Age at baseline, years (mean [95% CI])	60.20 [60.04, 60.37]		63.51 [63.47, 63.55]		<.001
Age at menarche, years	<i>N</i>	%	<i>N</i>	%	
<12	1,560	24.8	27,718	21.6	<.001
12–13	3,040	48.2	71,953	56.0	
≥14	1,701	27.0	28,848	22.4	
Age at first live birth, years					
<20	1,087	17.6	14,603	11.5	<.001
20–24	1,869	30.2	51,016	40.0	
25–29/Nulliparous	2,773	44.9	52,436	41.1	
≥30	453	7.3	9,484	7.4	
Number of first-degree relatives with breast cancer					
0	5,318	89.9	103,979	85.1	<.001
1	529	8.9	16,471	13.5	
≥2	71	1.2	1,696	1.4	
Number of breast biopsies					
0	4,909	83.2	95,075	78.9	<.001
1	689	11.7	18,955	15.6	
≥2	304	5.1	7,949	6.5	
OS Participants	3,479	54.8	73,485	57.0	<.001
CT Participants	2,874	45.2	55,491	43.0	<.001
HT	1,536	53.4	22,006	39.7	
Non-HT	1,338	46.6	33,485	60.3	
Breast cancer outcomes and follow-up time in WHI main study					
Number of Invasive breast cancers	130		4,713		
Age at diagnosis, years (mean [95% CI])	63.8 [62.7, 65.0]		68.3 [68.1, 68.5]		<.001
Follow-up time, years (mean [95% CI])	7.57 [7.53, 7.62]		8.12 [8.11, 8.13]		<.001

Note: CT clinical trial, OS observational study, HT hormone therapy trial, 95% CI 95% confidence interval. *P* value for differences among categorical variables are from χ^2 test and for differences among continuous variables are from *t* test

^a Information on presence of atypical hyperplasia was not available in the WHI data

Table 2 Comparison of RR Estimates from the BCRAT and the WHI

BCRAT risk category	BCRAT	WHI Hispanic		WHI non-Hispanic White	
	RR	RR [95% CI]	<i>P</i> [†]	RR [95% CI]	<i>P</i> [‡]
Age at menarche	1.10	1.08 [0.85, 1.37]	0.905	1.07 [1.03, 1.11]	0.230
Number of breast biopsies	1.27	1.71 [1.31, 2.24]	0.031	1.27 [1.22, 1.33]	0.994
Age at first live birth (AFB)	1.24	0.97 [0.78, 1.20]	0.024	1.13 [1.09, 1.18]	<0.001
Number of first-degree relatives with breast cancer (FDR)	2.61	2.16 [1.13, 4.13]	0.571	1.31 [1.15, 1.49]	<0.001
AFB * FDR interaction	0.83	0.84 [0.55, 1.29]	0.925	1.01 [0.94, 1.09]	<0.001

Parameter estimates are Relative Risks (RRs), based on follow-up through the end of WHI Main study only; RR estimates based on comparisons to the referent category for each variable: Age at menarche (≥14 years old); Number of breast biopsies (0 biopsies); Age at first live birth (<20 years old); Number of first-degree relatives with breast cancer (0 relatives); AFB * FDR (<20 years old and 0 relatives)

[†] Test of difference between WHI Hispanic and Gail Hispanic parameter estimates

[‡] Test of difference between WHI White and Gail White parameter estimates

returned close to the rates observed in 1990, if breast cancer incidence rates increase again in the near future, as expected [15], then serious consideration should be given to recalibrating the BCRAT to more recent rates. Another advantage of recalibrating the BCRAT for NHW women is

that it would then conform to rates more specific to NHWs, rather than the 1983–1987 rates which included Hispanics as “white women.”

Second, discriminatory accuracy may improve by using relative risks specific to each racial/ethnic group. The

Table 3 BCRAT and updated BCRAT observed/expected ratios in the WHI

	Observed breast cancers	BCRAT		Updated BCRAT ^a	
		Expected breast cancers ^b	O/E ratio	Expected breast cancers ^b	O/E ratio
Hispanics					
Main study (<i>n</i> = 6,353)	130	110	1.18 [0.99, 1.40]	120	1.08 [0.91, 1.28]
Non-Hispanic Whites					
Main study (<i>n</i> = 128,976)	4,713	4,009	1.18 [1.14, 1.21]	4,788	0.98 [0.96, 1.01]

^a For Hispanics, BCRAT is calibrated to 1990–1996 SEER Hispanic women rates and Updated BCRAT is calibrated to 1993–2007 SEER Hispanic women rates. For NHWs, BCRAT is calibrated to 1983–1987 SEER White women rates and Updated BCRAT is calibrated to 1993–2007 SEER non-Hispanic White women rates

^b Expected cancers are those estimated by the BCRAT and updated BCRAT, respectively

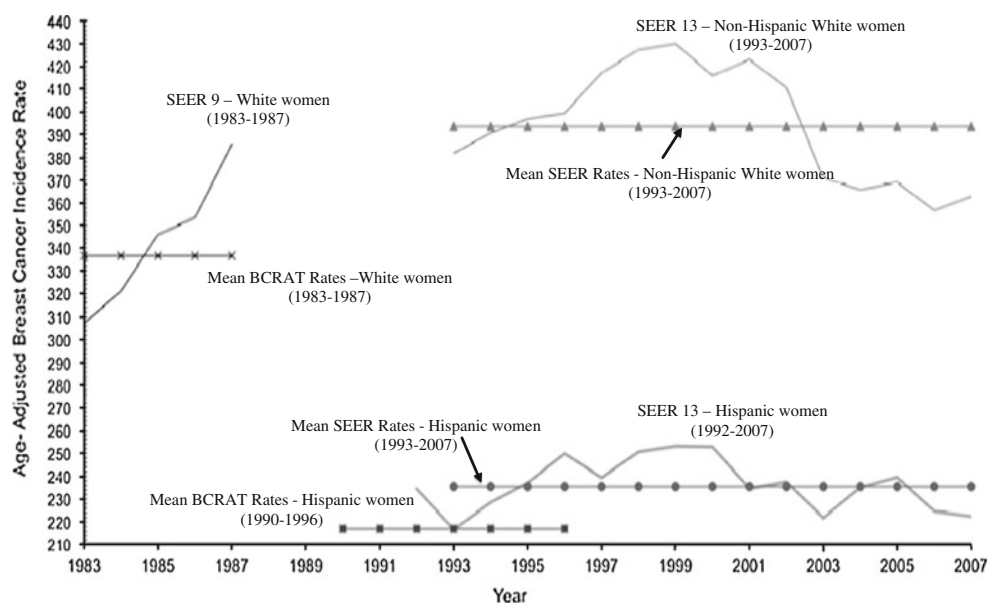


Fig. 1 Age-adjusted SEER Invasive Breast Cancer Incidence Rates over the BCRAT and WHI time periods. Solid lines show SEER incidence rates of invasive breast cancer for the years used in BCRAT that are available in SEER (1983–1987 for White women and 1992–1996 for Hispanic women), as well as for those years available in SEER that cover the WHI study period (1993–2007). Dotted lines show the mean incidence rate of invasive breast cancer over the time periods used in the BCRAT for White women and Hispanic women,

and for SEER 1993–2007 for both NHW and Hispanic women. Mean rates were estimated using age-specific invasive breast cancer incidence rates, among women aged 50 and older, based on 5-year age categories (i.e. 50–54 years of age, 55–59 years of age, 60–64 years of age, etc.) as used in the BCRAT. Rates are based on SEER 9 and 13 registries. All rates are per 100,000 and age adjusted to the 2000 US Standard Population rates among women age 50 and older

relative risks for number of breast biopsies may differ between Hispanic and NHW women. Previous case-control studies [4–8] have reported that several breast cancer risk factors may have different effects in Hispanic women compared with NHW women. Furthermore, Hispanic women likely have unique breast cancer risk factors, such as migration history, degree of acculturation, Hispanic origin, and ancestral genetic admixture [23–25] that are not considered in the current BCRAT. For NHW women, the BCRAT relative risk for family history is much larger than that estimated in our data, but estimates from our data are

consistent with those from most cohorts [16, 26] and meta-analyses [27, 28].

Some limitations need to be considered when interpreting our findings. Our analysis was limited to post-menopausal women aged 50 or older. There were only 130 Hispanic women with breast cancer, although our findings for Hispanic women are consistent with those from the much larger sample of NHW women. Though the WHI was a national study, conducted by 40 Clinical Centers in 24 states and the District of Columbia, Hispanic participants may not be representative of the US Hispanic population.

Finally, women in the WHI CT underwent mammography at least biennially, which may have increased the number of breast cancers detected over that expected in the general population with less frequent mammography.

Further development of a more comprehensive model for Hispanic women, that considers breast cancer risk factors distinct to Hispanic women and is validated in a large cohort of US Hispanic women, is warranted.

Acknowledgments This research was supported, in part, by the National Cancer Institute Biobehavioral Cancer Prevention and Control Training Program (grant number R25CA092408) at the University of Washington and the Intramural Research Program of the National Institutes of Health/National Cancer Institute. The Women's Health Initiative (WHI) program is funded by the National Heart, Lung, and Blood Institute, National Institutes of Health, U.S. Department of Health and Human Services through contracts N01WH22110, 24152, 32100-2, 32105-6, 32108-9, 32111-13, 32115, 32118-32119, 32122, 42107-26, 42129-32, and 44221. The authors thank the WHI investigators and staff for their dedication, and the study participants for making the program possible. A listing of WHI investigators can be found at http://www.whiscience.org/publications/WHI_investigators_shortlist.pdf.

Conflicts of interest The authors have no conflicts of interest to disclose.

References

1. American Cancer Society (ACS) (2007) Breast Cancer Facts & Figures 2007–2008. ACS, Atlanta
2. O'Brien K, Cokkinides V, Jemal A, Cardinez CJ, Murray T, Samuels A, Ward E, Thun MJ (2003) Cancer statistics for Hispanics, 2003. *CA Cancer J Clin* 53(4):208–226
3. Gail MH, Brinton LA, Byar DP, Corle DK, Green SB, Schairer C, Mulvihill JJ (1989) Projecting individualized probabilities of developing breast cancer for white females who are being examined annually. *J Natl Cancer Inst* 81(24):1879–1886
4. Chlebowski RT, Chen Z, Anderson GL, Rohan T, Aragaki A, Lane D, Dolan NC, Paskett ED, McTiernan A, Hubbell FA et al (2005) Ethnicity and breast cancer: factors influencing differences in incidence and outcome. *J Natl Cancer Inst* 97(6):439–448
5. Bondy ML, Spitz MR, Halabi S, Fueger JJ, Vogel VG (1992) Low incidence of familial breast cancer among Hispanic women. *Cancer Causes Control* 3(4):377–382
6. Hines LM, Risendal B, Slattery ML, Baumgartner KB, Giuliano AR, Sweeney C, Rollison DE, Byers T (2010) Comparative analysis of breast cancer risk factors among Hispanic and non-Hispanic white women. *Cancer* 116(13):3215–3223
7. Risendal B, Hines LM, Sweeney C, Slattery ML, Giuliano AR, Baumgartner KB, Curtin K, Byers TE (2008) Family history and age at onset of breast cancer in Hispanic and non-Hispanic white women. *Cancer Causes Control* 19(10):1349–1355
8. Sweeney C, Baumgartner KB, Byers T, Giuliano AR, Herrick JS, Murtaugh MA, Slattery ML (2008) Reproductive history in relation to breast cancer among Hispanic and non-Hispanic white women. *Cancer Causes Control* 19(4):391–401
9. Bondy ML, Lustbader ED, Halabi S, Ross E, Vogel VG (1994) Validation of a breast cancer risk assessment model in women with a positive family history. *J Natl Cancer Inst* 86(8):620–625
10. Spiegelman D, Colditz GA, Hunter D, Hertzmark E (1994) Validation of the Gail et al. model for predicting individual breast cancer risk. *J Natl Cancer Inst* 86(8):600–607
11. Rockhill B, Spiegelman D, Byrne C, Hunter DJ, Colditz GA (2001) Validation of the Gail et al. model of breast cancer risk prediction and implications for chemoprevention. *J Natl Cancer Inst* 93(5):358–366
12. Costantino JP, Gail MH, Pee D, Anderson S, Redmond CK, Benichou J, Wieand HS (1999) Validation studies for models projecting the risk of invasive and total breast cancer incidence. *J Natl Cancer Inst* 91(18):1541–1548
13. Gail MH, Pfeiffer RM (2005) On criteria for evaluating models of absolute risk. *Biostatistics* 6(2):227–239
14. Visvanathan K, Chlebowski RT, Hurley P, Col NF, Ropka M, Collyar D, Morrow M, Runowicz C, Pritchard KI, Hagerty K et al (2009) American society of clinical oncology clinical practice guideline update on the use of pharmacologic interventions including tamoxifen, raloxifene, and aromatase inhibition for breast cancer risk reduction. *J Clin Oncol* 27(19):3235–3258
15. Anderson WF, Katki HA, Rosenberg PS (2011) Incidence of breast cancer in the United States: current and future trends. *J Natl Cancer Inst* 103(18):1397–1402
16. Schonfeld SJ, Pee D, Greenlee RT, Hartge P, Lacey JV Jr, Park Y, Schatzkin A, Visvanathan K, Pfeiffer RM (2010) Effect of changing breast cancer incidence rates on the calibration of the Gail model. *J Clin Oncol* 28(14):2411–2417
17. Chlebowski RT, Anderson GL, Lane DS, Aragaki AK, Rohan T, Yasmeen S, Sarto G, Rosenberg CA, Hubbell FA (2007) Predicting risk of breast cancer in postmenopausal women by hormone receptor status. *J Natl Cancer Inst* 99(22):1695–1705
18. Ritenbaugh C, Patterson RE, Chlebowski RT, Caan B, Fels-Tinker L, Howard B, Ockene J (2003) The Women's Health Initiative Dietary Modification trial: overview and baseline characteristics of participants. *Ann Epidemiol* 13(9 Suppl):S87–S97
19. Hays J, Hunt JR, Hubbell FA, Anderson GL, Limacher M, Allen C, Rossouw JE (2003) The Women's Health Initiative recruitment methods and results. *Ann Epidemiol* 13(9 Suppl):S18–S77
20. Design of the Women's Health Initiative clinical trial and observational study. The Women's Health Initiative Study Group (1998) *Control Clin Trials* 19(1):61–109
21. Curb JD, McTiernan A, Heckbert SR, Kooperberg C, Stanford J, Nevitt M, Johnson KC, Proulx-Burns L, Pastore L, Criqui M et al (2003) Outcomes ascertainment and adjudication methods in the Women's Health Initiative. *Ann Epidemiol* 13(9 Suppl):S122–S128
22. Wieand S, Gail MH, James BR, James KL (1989) A family of nonparametric statistics for comparing diagnostic markers with paired or unpaired data. *Biometrika* 76(3):585–592
23. John EM, Phipps AI, Davis A, Koo J (2005) Migration history, acculturation, and breast cancer risk in Hispanic women. *Cancer Epidemiol Biomarkers Prev* 14(12):2905–2913
24. John EM, Sangaramoorthy M, Phipps AI, Koo J, Horn-Ross PL (2010) Adult body size, hormone receptor status, and premenopausal breast cancer risk in a multiethnic population: the San Francisco Bay Area breast cancer study. *Am J Epidemiol* 173(2):201–216
25. Miranda PY, Wilkinson AV, Etzel CJ, Zhou R, Jones LA, Thompson P, Bondy ML (2011) Policy implications of early onset breast cancer among Mexican-origin women. *Cancer* 117(2):390–397

26. Pharoah PD, Day NE, Duffy S, Easton DF, Ponder BA (1997) Family history and the risk of breast cancer: a systematic review and meta-analysis. *Int J Cancer* 71(5):800–809
27. Collaborative Group on Hormonal Factors in Breast Cancer (2001) Familial breast cancer: collaborative reanalysis of individual data from 52 epidemiological studies including 58,209 women with breast cancer and 101,986 women without the disease. *Lancet* 358(9291):1389–1399
28. Wacholder S, Hartge P, Prentice R, Garcia-Closas M, Feigelson HS, Diver WR, Thun MJ, Cox DG, Hankinson SE, Kraft P et al (2010) Performance of common genetic variants in breast-cancer risk models. *N Engl J Med* 362(11):986–993