

Value of Adding Single-Nucleotide Polymorphism Genotypes to a Breast Cancer Risk Model

Mitchell H. Gail

- Background** Adding genotypes from seven single-nucleotide polymorphisms (SNPs), which had previously been associated with breast cancer, to the National Cancer Institute's Breast Cancer Risk Assessment Tool (BCRAT) increases the area under the receiver operating characteristic curve from 0.607 to 0.632.
- Methods** Criteria that are based on four clinical or public health applications were used to compare BCRAT with BCRATplus7, which includes the seven genotypes. Criteria included number of expected life-threatening events for the decision to take tamoxifen, expected decision losses (in units of the loss from giving a mammogram to a woman without detectable breast cancer) for the decision to have a mammogram, rates of risk reclassification, and number of lives saved by risk-based allocation of screening mammography. For all calculations, the following assumptions were made: Hardy-Weinberg equilibrium, linkage equilibrium across SNPs, additive effects of alleles at each locus, no interactions on the logistic scale among SNPs or with factors in BCRAT, and independence of SNPs from factors in BCRAT.
- Results** Improvements in expected numbers of life-threatening events were only 0.07% and 0.81% for deciding whether to take tamoxifen to prevent breast cancer for women aged 50–59 and 40–49 years, respectively. For deciding whether to recommend screening mammograms to women aged 50–54 years, the reduction in expected losses was 0.86% if the ideal breast cancer prevalence threshold for recommending mammography was that of women aged 50–54 years. Cross-classification of risks indicated that some women classified by BCRAT would have different classifications with BCRATplus7, which might be useful if BCRATplus7 was well calibrated. Improvements from BCRATplus7 were small for risk-based allocation of mammograms under costs constraints.
- Conclusions** The gains from BCRATplus7 are small in the applications examined. Models with SNPs, such as BCRATplus7, have not been validated for calibration in independent cohort data. Additional studies are needed to validate a model with SNPs and justify its use.

J Natl Cancer Inst 2009;101:959–963

Gail (1) previously studied the improvement in discriminatory accuracy, which was measured as the area under the receiver operating characteristic curve (AUC), from adding genotypes from seven common single-nucleotide polymorphisms (SNPs), which had previously been associated with breast cancer, to the National Cancer Institute's Breast Cancer Risk Assessment Tool (BCRAT) for estimating invasive breast cancer risk. The AUC increased from 0.607 for BCRAT to 0.632 for the model that added the seven SNPs (BCRATplus7). In an editorial accompanying Gail (1), Pepe and Janes (2) recommended that BCRAT and BCRATplus7 be compared by calculating the expected proportions of women with risks above a critical risk threshold. Such a threshold could be established by balancing expected risks and benefits, as described previously (3–6). In this study, I compare BCRAT with BCRATplus7 with respect to criteria that are based on risks and benefits.

Methods and Risk Models

Gail (1) defined the absolute risk model for BCRAT as $r_1(X_1)$, where X_1 includes age at menarche, age at first live birth, number of previous breast biopsy examinations, and the total number of

first-degree relatives (ie, mother and sisters) with breast cancer. On a short interval over which competing risks can be ignored (1), BCRATplus7 has risk $r(X_1, X_2) = rr_1(X_1)rr_2(X_2)k$, where rr_1 and rr_2 are relative risks, X_2 is the joint genotype of the seven SNPs, and k is the ratio of the age-specific breast cancer incidence rate from the National Cancer Institute's Surveillance, Epidemiology and End Results (SEER) Program to the average value of $rr_1(X_1)rr_2(X_2)$.

A risk model should be well calibrated, that is, within each subgroup as defined by its risk factors, the observed number of cases of breast cancers diagnosed has an expected value given by summing the model risks over individuals in the subgroup (5). I assume

Affiliation of author: Biostatistics Branch, Division of Cancer Epidemiology and Genetics, National Cancer Institute, Bethesda, MD.

Correspondence to: Mitchell H. Gail, MD, PhD, Biostatistics Branch, Division of Cancer Epidemiology and Genetics, National Cancer Institute, 6120 Executive Blvd, Rm 8032, Bethesda, MD 20892-7244 (e-mail: gailm@mail.nih.gov).

See "Funding" and "Notes" following "References."

DOI: 10.1093/jnci/djp130

Published by Oxford University Press 2009.

CONTEXT AND CAVEATS

Prior knowledge

BCRATplus7 is a risk assessment model for breast cancer that was developed by adding genotypes from seven single-nucleotide polymorphisms (SNPs) that had previously been associated with breast cancer to the National Cancer Institute's Breast Cancer Risk Assessment Tool (BCRAT).

Study design

Risks and benefits that were based on the thresholds for clinical decisions as assessed by BCRATplus7 were compared with those as assessed by BCRAT.

Contribution

Only small improvements in the benefit for deciding whether to take tamoxifen to prevent breast cancer, to recommend screening mammograms, to reclassify breast cancer risk, and to allocate access to screening mammograms under costs constraints.

Implications

Additional studies are required to validate the BCRATplus7 model in an independent cohort of women with available data on the seven genotypes and other risk factors for breast cancer.

Limitations

The BCRATplus7 model was constructed by using published data on SNPs and data from BCRAT. Various assumptions were made during the development of BCRATplus7, including that BCRATplus7 was well calibrated and that there were no interactions among SNPs or between SNPs and other risk factors.

From the Editors

that both BCRAT and BCRATplus7 are well calibrated, although only BCRAT has been validated with independent data (7,8). There is a special need to check the calibration of BCRATplus7 because BCRATplus7 was constructed by synthesizing data from different sources (1) rather than by constructing the model from individuals with measurements available for all SNPs and risk factors in BCRAT. BCRATplus7 is a refinement of BCRAT because for each level of X_1 , there are multiple further risk categorizations that are based on X_2 . BCRAT need not be well calibrated in subgroups that are defined by both X_1 and X_2 , although it is well calibrated in groups defined by X_1 alone.

Results

Deciding to Take Tamoxifen to Prevent Breast Cancer

The decision to take tamoxifen to prevent breast cancer is difficult because the reductions in risks of invasive breast cancer and hip fracture are offset by increased risks of endometrial cancer, stroke, and pulmonary embolism (9,10). For example, among white women aged 50–59 years, rates per 10^5 women-years of these five life-threatening events with no tamoxifen use are 246.6 invasive breast cancers, 101.6 hip fractures, 81.4 endometrial cancers, 110.0 strokes, and 50.0 pulmonary embolisms, respectively, and rates with tamoxifen use are 125.8 ($=0.51 \times 246.6$), 55.9 ($=0.55 \times 101.6$), 326.4 ($=4.01 \times 81.4$), 174.9 ($=1.59 \times 110.0$), and 150.5 ($=3.01 \times 50.0$) (9,10). A total of 589.6 events among 100000 women not

using tamoxifen are expected each year; 833.5 events are expected among women using tamoxifen. A perfectly discriminating risk model would recommend tamoxifen only to the 246.6 women destined to get invasive breast cancer, resulting in 469.7 expected events. This number was obtained by considering the 246.6 women who will get breast cancer and the remaining 99753.4 ($=100000 - 246.6$) women separately. The first group is expected to contribute 125.8 ($=0.51 \times 246.6$) breast cancers and 1.7 [$=(246.6/100000) \times (833.5 - 125.8)$] other events. The second group is expected to contribute 342.2 [$=(99753.4/100000) \times (589.6 - 246.6)$] non-breast cancer events.

To use a risk model, the optimal breast cancer risk threshold for deciding whether or not to give tamoxifen should be determined. The expected events averted by tamoxifen for women with a breast cancer risk of r is $r(1 - 0.51) + 101.6(1 - 0.55) + 81.4(1 - 4.01) + 110.0(1 - 1.59) + 50.0(1 - 3.01) = 0.49r - 364.7$.

This quantity must be positive for there to be an expected benefit from tamoxifen. Solving $0.49r - 364.7 = 0$ leads to the optimal threshold, $r^* = 364.7/0.49 = 774.3$, which is 3.14 times the risk for a typical white woman aged 50–59 years. This threshold minimizes the expected number of events, regardless of which risk projection model is used (5). For BCRATplus7, the expected number of events is

$$L_2 = \sum_{(X_1, X_2): r(X_1, X_2) > r^*} P(X_1, X_2) \{0.51r(X_1, X_2) + c_1\} + \sum_{(X_1, X_2): r(X_1, X_2) \leq r^*} P(X_1, X_2) \{r(X_1, X_2) + c_2\}, \quad [1]$$

where $P(X_1, X_2) = P(X_1)P(X_2)$ [see Gail (1)], and where $c_1 = 707.7$ and $c_2 = 343.0$. From Equation 1, the expected number of events for BCRATplus7, L_2 , was 587.8. For BCRAT, the expected number of events is

$$L_1 = \sum_{(X_1, X_2): r(X_1) > r^*} P(X_1, X_2) \{0.51r(X_1, X_2) + c_1\} + \sum_{(X_1, X_2): r(X_1) \leq r^*} P(X_1, X_2) \{r(X_1, X_2) + c_2\}. \quad [2]$$

From Equation 2, the expected number of events for BCRAT, L_1 , was 588.2. Hence, BCRATplus7 predicted only 0.4 fewer expected events than BCRAT, a reduction of 0.07%. A perfect model would yield a 20.1% reduction. The percentage of the improvement from using a perfect model compared with giving no tamoxifen that was achieved by using BCRAT was 1.17% and by using BCRATplus7 was 1.50%. For white women aged 40–49 years, the expected numbers of events were 241.4 if all women did not use tamoxifen and 251.8 if all women used tamoxifen; the optimal threshold risk r^* was 177.8 (Table 1). BCRATplus7 reduced the expected numbers of life-threatening events by 0.81% compared with BCRAT. A perfect model would yield a 29.0% reduction. The percentage of the improvement from using a perfect model compared with giving no tamoxifen that was achieved by using BCRAT was 12.1% and by using BCRATplus7 was 14.6% (Table 1). Thus, neither BCRAT nor BCRATplus7 offered much benefit to white women aged 50–59 years because only women at very high levels of breast cancer risk would stand to benefit from tamoxifen [see table 10 in Gail et al. (10)] and because neither model is discriminating enough to pick out only those women destined to get

Table 1. Expected numbers of life-threatening events per 100 000 woman-years for various risk models and strategies for deciding whether or not to take tamoxifen†

Decision rule	Ages 40–49 y		Ages 50–59 y	
	Threshold risk, r^*	No. of expected events‡	Threshold risk, r^*	No. of expected events‡
No tamoxifen	Infinite	241.4	Infinite	589.6
BCRAT risk > r^*	177.8	232.1	774.3	588.2
BCRATplus7 risk > r^*	177.8	230.2	774.3	587.8
Perfect model risk > r^*	177.8	164.8	774.3	469.7
All get tamoxifen	0	251.8	0	833.5

† BCRAT = Breast Cancer Risk Assessment Tool; BCRATplus7 = Breast Cancer Risk Assessment Tool plus genotypes from seven single-nucleotide polymorphisms; RR = relative risk.

‡ Expected numbers of events for white women aged 40–49 years are based on the following RRs from tamoxifen and event rates for women without tamoxifen: invasive breast cancer (RR = 0.51 and rate = 156.6), hip fracture (RR = 0.55 and rate = 3.8), endometrial cancer (RR = 2.53 and rate = 21.0), stroke (RR = 1.59 and rate = 45.0), and pulmonary embolism (RR = 3.01 and rate = 15.0) (9,10). For white women aged 50–59 years, the data are the same, except for endometrial cancer (RR = 4.01 and rate = 81.4) and for the rates for invasive breast cancer (rate = 246.6), hip fracture (rate = 101.6), stroke (rate = 110.0), and pulmonary embolism (rate = 50.0). Confidence intervals on RRs are available in (9) and (10). Invasive breast cancer rates are from appendix table 1 in (7); for a description of how other rates were estimated, see Gail et al. (10).

breast cancer. The gains were better for white women aged 40–49 years, but BCRATplus7 still offered only a slight improvement over BCRAT.

Deciding to Have Screening Mammography

Most health-care providers agree that women aged 50–59 years should be screened with mammography, and the American Cancer Society and National Cancer Institute have recommended mammography for women aged 40–49 years as well. The main argument against mammography in young women is that the expected costs, including unnecessary biopsy examinations and other medical procedures, outweigh the expected benefits of detecting breast cancer if breast cancer prevalence is too small.

As previously described (5) and with numerical values calculated below, let C_{00} be the loss from recommending no mammography to a woman without detectable breast cancer, C_{10} be the loss from recommending mammography to a woman without detectable breast cancer, C_{01} be the loss from recommending no mammography to a woman with detectable breast cancer, and C_{11} be the loss from recommending mammography to a woman with detectable breast cancer. Because the prevalence threshold for recommending mammography depends only on ratios of these costs, the units of cost can be arbitrary. For this analysis, I set the unit of cost C_{10} to be equal to 1, that is, the cost of giving a mammogram to a woman who does not need it. I set C_{00} equal to 0 because there is no loss from not giving a mammogram to a woman without detectable breast cancer. I assume that C_{11} equals $0.75C_{01}$ because mammographic screening reduces mortality from breast cancer by approximately 25% in women aged 50 years or older (11). The optimal breast cancer prevalence threshold for recommending mammography (5,6) is

$$p^* = (C_{10} - C_{00}) / (C_{10} - C_{00} + C_{01} - C_{11}) = 1 / (1 + 0.25C_{01}). \quad [3]$$

On the basis of the consensus that woman aged 50–54 years should receive mammography, I chose p^* to be the prevalence of breast cancer detectable by routine screening mammography in women aged 50–54 years, which is estimated as 313×10^{-5} , the cancer detection rate from the Breast Cancer Surveillance

Consortium (http://breastscreening.cancer.gov/data/performance/screening/rate_age.html). To relate projected prevalence, p , to the projected incidence rate, r , I took $p = 1.3 \times r$ for white women because age-specific cancer detection rates were nearly equal to 1.3 times SEER incidence rates for white women throughout the age range of 45–74 years. Because SEER rates for white women aged 50–54 years were $r^* = 241.4 \times 10^{-5}$, I used $p^* = 1.3 \times 241.4 \times 10^{-5} = 313.8 \times 10^{-5}$. From Equation 3, $C_{01} = 1270.6$ and $C_{11} = 953.0$. The expected loss for recommending mammography if the projected prevalence exceeds p^* is

$$L = C_{11}\mu(\text{sens}) + C_{01}\mu(1 - \text{sens}) + C_{10}(1 - \mu)(1 - \text{spec}) + C_{00}(1 - \mu)(\text{spec}), \quad [4]$$

where μ is the average prevalence, sens is the probability that a woman who has detectable breast cancer will have a projected prevalence exceeding p^* , and spec is the probability that a woman who does not have detectable breast cancer will have a projected prevalence no greater than p^* . For women aged 50–54 years, the expected loss for a perfect model (sens = spec = 1) was 2.991, which was 2.991 times the loss from giving a mammogram to a woman who does not need it. For BCRAT, the expected loss was 3.834, and for BCRATplus7, the expected loss was 3.801. Thus, the expected loss from BCRATplus7 was 0.86% smaller than that from BCRAT, and the expected loss with a perfect model was 22.0% smaller. If one used the prevalence in women aged 40–44 years as the threshold, then $p^* = 1.3 \times 118.9 \times 10^{-5}$, the implied costs would increase to $C_{01} = 2583.8$ and $C_{11} = 1937.9$, and with this lower prevalence threshold, the expected losses for white women aged 50–54 years for BCRAT, BCRATplus7, and a perfect model would be, respectively, 7.078, 7.071, and 6.081. The improvements over BCRAT were 0.10% for BCRATplus7 and 14.1% for a perfect model.

Reclassification of Risk

Cook (12) suggested that two risk models should be compared by cross-classifying individuals on their risks from both models. From such tables, the proportion of individuals that falls below some risk threshold of one model and above the threshold of another model can be calculated. The joint distribution of the risks from BCRAT and BCRATplus7 is

$$\begin{aligned} & \text{Prob}(\text{BCRAT risk} \leq u_1, \text{BCRATplus7 risk} \leq u_2) \\ &= \sum_{(X_1, X_2): r_1(X_1) \leq u_1, r(X_1, X_2) \leq u_2} P(X_1)P(X_2), \end{aligned} \quad [5]$$

from which I calculated the values in Table 2 for white women aged 50–59 years. BCRAT predicted that 38.0% + 44.0% = 82.0% of the population had risks of less than 1.5% (Table 2). However, 0.6% + 6.0% + 0.9% + 0.1% = 7.6% of the population had a BCRAT risk of less than 1.5% and a BCRATplus7 risk of 1.5% or more. If BCRATplus7 is assumed to be a well-calibrated refinement of BCRAT, $100 \times 7.6/82.0 = 9.27\%$ of those classified as having BCRAT risk of less than 1.5% would have a risk above that level. Likewise, of the 17.9% of women with a BCRAT risk of 1.5% or more, 21.2% would be reclassified as having risk of less than 1.5%. If the threshold risk of interest was 2%, 4.14% of the women with BCRAT risk of less than 2% would be reclassified as having a risk of 2% or greater, and 31.7% of the women with BCRAT risk of 2% or more would be reclassified as having a risk of less than 2%. Janes et al. (13) argued that the focus should be on the values in the margins of Table 2 to determine which fraction of the population falls into the highest and lowest risk categories, rather than on reclassification rates. However, if BCRATplus7 were a well-calibrated refinement of BCRAT, then these reclassification rates could be interpreted as the proportions of women who are incorrectly categorized by BCRAT.

Allocating Mammography Under Cost Constraints

If there are insufficient resources to give mammograms to all women in a population and the costs of providing a questionnaire to all women and of estimating breast cancer risk from data in the questionnaire are negligible, then one could rank women in decreasing order of risk and give mammograms only to the women at highest risk, until the money runs out. The distribution of risk factors in the general population induces (5) the distribution of risk r in the general population, F , as well as the distribution of risk in women with disease,

$$\text{Prob}(\text{risk} \leq x \mid \text{woman has disease}) \equiv F_d(x) = \int_0^x r dF(r) / \int_0^1 r dF(r).$$

The area under a curve of $1 - F_d(x)$ plotted against $1 - F(x)$ approximates the AUC (5). If we assume that the prevalence of detectable breast cancer is proportional to breast cancer risk, such

curves can be used to calculate the proportion of women with detectable breast cancer who are among the proportion p of the general population that is at highest risk (14). We computed the risk distributions F and F_d separately for BCRAT and BCRATplus7 as described previously (1). From these distributions, the 50% of women in the general population with highest risks, and by assumption the highest breast cancer prevalence, includes 62.1% of the women with prevalent breast cancer when ranked by BCRAT and 67.6% of women with prevalent breast cancer when ranked by BCRATplus7. Thus, if there were only enough money to give screening mammography to half the population and if risk assessment costs were negligible, more than half the prevalent detectable breast cancers would be found by screening the 50% of women at highest risk.

Costs of assessing risk may not be negligible. In Gail (15), I calculated the optimal risk-based allocation strategy by assuming that beta distributions approximate the distributions of risk F and F_d for BCRAT and BCRATplus7, that risk assessment costs were 2% of the cost of a mammogram, and that there was only enough money to give mammograms to half the population. For BCRAT, the optimal strategy was to give the entire population a risk assessment and then to give mammograms to the 48% of women with highest ranked risks, thus achieving 63.2% of the mortality benefit that would have been achieved by giving mammography to the entire population. For BCRATplus7, this optimal strategy achieved 66.7% of the maximum mortality benefit from mammography. Consequently, BCRATplus7 gave an improvement of 5.5% [$=100(66.7 - 63.2)/63.2$] compared with BCRAT. This comparison ignores the costs of obtaining DNA and genotyping for BCRATplus7, however. Currently, commercial testing for the SNPs included in BCRATplus7 costs more than mammography. The costs would need to drop to less than 20% of the cost of a mammogram before risk assessment with SNPs could be used to increase the mortality benefit by allocating mammography to women at highest risk (15).

Discussion

This article has presented criteria, other than AUC, to assess how much adding information from seven SNPs could enhance the usefulness of BCRAT. Two examples concerned the use of an

Table 2. Distribution in percent of joint risk categories defined by five-year probabilities of invasive breast cancer in white women aged 50–59 years from BCRAT and from BCRATplus7*

Categories of 5-year risk from BCRAT	Probability of invasive breast cancer, %					Total
	Categories of 5-year risk from BCRATplus7					
	<1.0%	1.0% to <1.5%	1.5% to <2.0%	2.0% to <2.5%	≥2.5%	
<1.0%	29.4	8.0	0.6	0.0	0.0	38.0
1.0% to <1.5%	15.4	21.6	6.0	0.9	0.1	44.0
1.5% to <2.0%	0.2	3.0	3.7	1.9	0.9	9.7
2.0% to <2.5%	0.0	0.6	1.8	1.6	1.3	5.3
≥2.5%	0.0	0.0	0.2	0.4	2.3	2.9
Total	45.0	33.2	12.3	4.8	4.6	99.9

* BCRAT = Breast Cancer Risk Assessment Tool; BCRATplus7 = Breast Cancer Risk Assessment Tool plus genotypes from seven single-nucleotide polymorphisms.

optimal risk threshold for a decision to take a medical action, that is, to use tamoxifen or to have a mammogram. These analyses indicated that the decrease in expected losses from the use of BCRATplus7 rather than BCRAT was small. Much more discriminating models are needed to reduce the expected losses substantially. Nonetheless, a better decision can be made in some women by using BCRATplus7 instead of BCRAT, if it is shown to be true that BCRATplus7 is well calibrated (Table 2). Ranking the risks of women may be useful for allocating limited medical resources. For this purpose, BCRATplus7 is modestly better than BCRAT, but the comparison ignores the additional costs of obtaining DNA and genotyping.

While this article was in revision, four new SNPs were reported to be associated with breast cancer (16,17). These SNPs had a geometric mean odds ratio per allele of 1.11, compared with 1.15 for the SNPs in BCRATplus7. The AUC for a model with all 11 SNPs was 0.585, which is less than that for BCRAT, 0.607. Adding these four SNPs to produce the model BCRATplus11 with the methods used to construct BCRATplus7, I found that the AUC was increased from 0.632 for BCRATplus7 to 0.637 for BCRATplus11. Under the assumptions that BCRATplus11 is well calibrated, the numbers of life-threatening events in 1 year among 100 000 women aged 50–59 years who decided whether to take tamoxifen on the basis of risk models were reduced from 587.63 for BCRATplus7 to 587.62 for BCRATplus11. The losses from mammographic screening with the threshold prevalence of detectable breast cancer of women aged 50–54 years decreased from 3.801 with BCRATplus7 to 3.792 for BCRATplus11. These calculations indicate that the improvements from adding the additional four SNPs are small.

This study had several limitations. The model BCRATplus7 was constructed by synthesizing data from the literature on SNPs with data from BCRAT under the assumptions as described previously (1). The assumptions included Hardy–Weinberg equilibrium at each SNP locus, linkage disequilibrium across SNP loci, additive effects of disease-associated alleles on a logistic scale, independence of SNPs from factors in BCRAT, and additive effects of genotypes with BCRAT log odds effects on a logistic scale. Under these assumptions, I assumed that BCRATplus7 was well calibrated, and in particular that there were no interactions on a logistic scale among the SNPs or between SNPs and factors in BCRAT. Ideally, these assumptions and the calibration of BCRATplus7 would be tested in cohorts of women with data on SNPs and the factors in BCRAT who are followed prospectively to determine breast cancer incidence. If there are positive correlations among the SNPs and factors in BCRAT, then the improvement from BCRATplus7 compared with that from BCRAT would be less (1).

In view of these uncertainties and the small improvements from BCRATplus7 in these applications, further studies are needed to validate models with SNPs and to assess how much they improve performance over simpler models.

References

1. Gail MH. Discriminatory accuracy from single-nucleotide polymorphisms in models to predict breast cancer risk. *J Natl Cancer Inst.* 2008; 100(14):1037–1041.
2. Pepe MS, Janes HE. Gauging the performance of SNPs, biomarkers, and clinical factors for predicting risk of breast cancer. *J Natl Cancer Inst.* 2008;100(14):978–979.
3. Baker SG, Cook NR, Vickers A, Kramer BS. Using relative utility curves to evaluate risk prediction (published online ahead of print April 3, 2009). *J R Stat Soc Ser A.* 2009;172(4).
4. Baker SG, Kramer BS. Peirce, Youden, and receiver operating characteristic curves. *Am Stat.* 2007;61(4):343–346.
5. Gail MH, Pfeiffer RM. On criteria for evaluating models of absolute risk. *Biostatistics.* 2005;6(2):227–239.
6. Pauker SG, Kassirer JP. Therapeutic decision-making—cost-benefit analysis. *N Engl J Med.* 1975;293(5):229–234.
7. Costantino JP, Gail MH, Pee D, et al. Validation studies for models projecting the risk of invasive and total breast cancer incidence. *J Natl Cancer Inst.* 1999;91(18):1541–1548.
8. 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.
9. Fisher B, Costantino JP, Wickerham DL, et al. Tamoxifen for prevention of breast cancer: Report of the National Surgical Adjuvant Breast and Bowel Project P-1 study. *J Natl Cancer Inst.* 1998;90(18):1371–1388.
10. Gail MH, Costantino JP, Bryant J, et al. Weighing the risks and benefits of tamoxifen treatment for preventing breast cancer. *J Natl Cancer Inst.* 1999;91(21):1829–1846.
11. Freedman DA, Petitti DB, Robins JM. On the efficacy of screening for breast cancer. *Int J Epidemiol.* 2004;33(1):43–55.
12. Cook NR. Use and misuse of the receiver operating characteristic curve in risk prediction. *Circulation.* 2007;115(7):928–935.
13. Janes H, Pepe MS, Gu W. Assessing the value of risk predictions by using risk stratification tables. *Ann Intern Med.* 2008;149(10):751–760.
14. Pharoah PDP, Antoniou A, Bobrow M, Zimmern RL, Easton DF, Ponder BAJ. Polygenic susceptibility to breast cancer and implications for prevention. *Nat Genet.* 2002;31(1):33–36.
15. Gail MH. Applying the Lorenz curve to disease risk to optimize health benefits under cost constraints. *Stat Interface.* 2009;2(2):117–121.
16. 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) [published online ahead of print March 29, 2009]. *Nat Genet.* 2009;41(5):579–584.
17. Ahmed S, Thomas G, Ghousaini M, et al. Newly discovered breast cancer susceptibility loci on 3p24 and 17q23.2. *Nat Genet.* 2009;41(5):585–590.

Funding

The Intramural Research Program of the Division of Cancer Epidemiology and Genetics, National Cancer Institute and National Institutes on Health, supported this work.

Notes

The author had full responsibility for the design of the study, the collection of data, the analysis and interpretation of the data, the decision to submit the manuscript for publication, and the writing of the manuscript.

I thank Dr Ruth M. Pfeiffer for helpful comments and Dr Diana L. Miglioretti for comments on data in the Breast Cancer Surveillance Consortium.

Manuscript received December 22, 2008; revised March 29, 2009; accepted April 6, 2009.