

Published in final edited form as:

*Stat Med.* 2011 May 10; 30(10): 1090–1104. doi:10.1002/sim.4187.

## Personalized estimates of breast cancer risk in clinical practice and public health

**Mitchell H. Gail**

Biostatistics Branch, Division of Cancer Epidemiology and Genetics, National Cancer Institute, Bethesda, MD, USA

### Abstract

This paper defines absolute risk and some of its properties, and presents applications in breast cancer counseling and prevention. For counseling, estimates of absolute risk give useful perspective and can be used in management decisions that require weighing risks and benefits, such as whether or not to take tamoxifen to prevent breast cancer. Absolute risk models are also useful in designing intervention trials to prevent breast cancer and in assessing the potential reductions in absolute risk of disease that might result from reducing exposures that are associated with breast cancer. In these applications, it is important that the risk model be well calibrated, namely that it accurately predict the numbers of women who will develop breast cancer in various subsets of the population. Absolute risk models are also needed to implement a “high risk” prevention strategy that identifies a high risk subset of the population and focuses intervention efforts on that subset. The limitations of the high risk strategy are discussed, including the need for risk models with high discriminatory accuracy, and the need for less toxic interventions that can reduce the threshold of risk above which the intervention provides a net benefit. I also discuss the potential use of risk models in allocating prevention resources under cost constraints. High discriminatory accuracy of the risk model, in addition to good calibration, is desirable in this application, and the risk assessment should not be expensive in comparison with the intervention.

### Keywords

absolute risk; allocation of prevention resources; breast cancer; calibration; crude risk; cumulative incidence; discriminatory accuracy; disease prevention; designing disease prevention trials; high risk prevention strategy; risk versus benefit

## 1. Introduction

This paper is based on a lecture in honor of Professor Peter Armitage entitled “Personalized Estimates of Disease Risk in Clinical Practice and Public Health,” which was given at the Medical Research Council in Cambridge, England in November, 2009. There has been a recent surge of research to develop models to project the risk of cancers and other diseases, to validate such models, to improve the discriminatory accuracy of such models, and to define criteria for evaluating the success of such models. In this paper, I reflect on the usefulness of such models for breast cancer incidence, both for the individual who is faced with a clinical decision, such as whether or not to take tamoxifen to prevent breast cancer, and for applications in public health at the population level. Thus I examine the role of risk models within the broader framework of strategies to prevent cancer [1]. I use breast cancer

risk projection and prevention as the primary example in this paper, although the ideas apply more generally. Despite expectations raised by the prospect of “personalized medicine” based on genetic and molecular data [2], available breast cancer risk prediction models have very limited ability to discriminate between those who will develop breast cancer, and those who will not [3–6]. Given this limitation, the value of current breast cancer risk prediction models and their role in clinical medicine and public health depend largely on the properties of interventions that are available to prevent breast cancer. Models for the risk of disease recurrence or death following a breast cancer diagnosis play an important role in patient management, but are not the subject of this paper.

In Section 2, I define absolute risk, which is sometimes called cumulative risk or crude risk, and I review types of models for projecting breast cancer risk. I also define the important concept of the distribution of risk in the population and related quantities, such as measures of discriminatory accuracy of the model and calibration of the model. Section 3 discusses uses of the model for advising the individual patient, such as weighing the risks and benefits of an intervention. In Section 4, I turn to some applications of absolute risk models in public health, including designing prevention trials, assessing the potential of certain preventive activities for reducing risk in the population, implementing a prevention strategy based on “high risk” subsets of the population, which is exemplified by the use of tamoxifen to prevent breast cancer, and allocating limited resources for preventive activities, such as screening a population with mammography. Given the inherently limited potential for cancer prevention based on a “high risk” strategy that focuses on only a small subset of the population at highest risk, I speculate on the potential for general population prevention strategies in the Discussion (Section 5).

## 2. Some definitions and notation

### 2.1 Absolute risk

“Absolute risk” is the probability that a person with a given set of risk factors and free of the disease of interest at age  $a$  will develop disease before a subsequent age  $a+\tau$ , where  $\tau$  is the duration of the interval over which risk is projected. Absolute risk is sometimes called “crude risk” in the competing risks literature and “cumulative incidence” in the statistics literature, although the latter term is also occasionally used for the integrated cause-specific hazard and for the “pure” probability of disease by age  $a+\tau$  that would arise if no competing causes of mortality were present. To avoid confusion, I use the term “absolute risk”. Some investigators use time since accrual rather than age as the time scale. In this case, absolute risk is defined as the probability that a person with a given set of risk factors and free of the disease of interest at the time of accrual will develop disease at or before a subsequent duration of follow-up,  $\tau$ . An example is a model for cardiovascular events based on the Framingham Study [7]. In this paper, I use the age scale, however.

To illustrate the difference between absolute risk and pure risk, consider a population of 1000 women, all aged 60 years, and all with the same combination of risk factors (Table I). We assume that these women are followed over time, and, for simplicity, that there is no loss to follow-up. Some of the women die from non-breast cancer causes, however. Thus, of the 1000 women aged 60 years, 17 developed incident breast cancer and 44 died of other causes before age 65, leaving  $1000-17-44=939$  women at risk at age 65. Such events are likewise observed in the age intervals 65–69 and 70–74. An estimate of the crude probability (or absolute risk) of developing breast cancer by age 75 is simply  $(17+20+22)/1000 = 0.059$  or 5.9%. Notice that this absolute risk is reduced by the chance of dying of other causes. For comparison, we use actuarial methods, analogous to the Kaplan Meier method, to estimate the “pure” probability of developing breast cancer if deaths from other causes could be eliminated. This calculation treats deaths from competing causes as independent censoring

events, and assumes that such deaths occur at the end of each interval. Then the “pure risk” of breast cancer is estimated as  $1 - (1 - 17/1000)(1 - 20/939)(1 - 22/856) = 1 - 0.937 = 0.0627 = 6.3\%$ . The pure risk of breast cancer is larger than the absolute risk, because other causes of death are hypothetically eliminated. Note that we needed to assume that these deaths acted like “independent censoring” to compute pure risk, whereas absolute risk can be computed directly from the observable data as in Table I and requires no such assumptions [8]. A scientist might want to focus on pure risk to understand the effects of an intervention on a particular outcome, like breast cancer incidence, regardless of its effects on competing causes of mortality. However, for clinical purposes, absolute risk is more pertinent, because a patient is always subject to other causes of mortality. In advising a 40-year old patient without breast cancer, for example, it makes little sense to ask or answer the question: “What would be your chance of developing breast cancer by age 60 if you had no risk of dying of non-breast cancer causes during this period?” Yet it is common practice to present pure risks in the clinical and epidemiological literature. In part this may be due to the widespread availability of software for computing Kaplan-Meier survival curves, rather than corresponding curves of absolute risk. The difference between absolute and pure risk is typically small for short projection intervals, such as five years. For long projection intervals such as calculations of “lifetime risk” (e.g. to age 90), pure risks can substantially exceed the more relevant absolute risk, however. Actuarial calculations [9] similar to those above were used to estimate lifetime absolute risk of cardiovascular disease for various subgroups defined by risk factors [10].

## 2.2 Covariate modeling of absolute risk

If one had large numbers of women with each possible combination of risk factors, one could estimate absolute risk very simply as in Table 1, with a slight adaptation for loss during follow-up or censoring at end of follow-up [9]. Usually there will be insufficient data for such analyses, and one needs to use regression methods to allow for multivariate risk modeling. Let  $c=1$  if the event is breast cancer (the cause of interest) and  $c=2$  if the event is death from non-breast cancer causes, and let  $T$  be the age at which the first of these events occurs. Then the absolute risk of breast cancer by age  $a+\tau$  for a woman who is well at age  $a$  with covariates  $X$  is  $\Pr(a \leq T < a+\tau, c=1 | T \geq a, X)$ . For simplicity, I assume that  $X$  remains fixed for the duration of the projection, but the formulas below generalize if  $X$  varies over time and if  $X(t)$  is assumed to be known over the duration of the projection. Two approaches are commonly used to incorporate covariates into regression models of absolute risk. One approach is to model the cause-specific hazard functions [8],

$$h_i\{t|X\} = \lim_{\varepsilon \rightarrow 0} \Pr(t \leq T < t+\varepsilon, c=i | T \geq t, X) / \varepsilon. \quad (1)$$

Usually one takes  $h_i\{t|X\} = h_{i0}(t)rr_i(t; X)$ , where  $h_{i0}(t)$  is a baseline hazard rate for women at the reference level of risk factors,  $X_0$ , and  $rr_i(t; X)$  is the corresponding relative hazard at age  $t$ . The argument  $t$  is used to indicate that relative risk may vary over time because of interaction of age with covariates, even if  $X$  is fixed. If  $X(t)$  varies, then the relative risk may also change because of the changing covariates. In many cases,  $rr_i(t; X) = \exp(\beta_i X)$  is assumed constant. In some applications, it is assumed that competing causes of mortality do not depend on measured covariates. Then the age-specific mortality rate from non-breast cancer causes,  $h_2(t)$ , can be obtained from national mortality rates, for example, and the absolute risk is given by

$$R(a, \tau, X) \equiv \Pr(a \leq T < a + \tau, c = 1 | T \geq a, X) = \int_a^{a+\tau} h_{10}(t) rr_1(t; X) \exp \left[ - \int_a^t \{h_{10}(u) rr_1(u; X) + h_2(u)\} du \right] dt. \quad (2)$$

Equation (2) is derived by noting that the exponential term is the probability that a woman who is well at age  $a$  will remain free of breast cancer and alive until age  $t$ , at which time she has an instantaneous risk of developing breast cancer,  $h_{10}(t) rr_1(t; X) dt$ . By integrating, we add up all these instantaneous probabilities to obtain the total absolute risk. Note that the effect of competing risk of mortality is to reduce the absolute risk, because increasing  $h_2(t)$  decreases the exponential factor in the integrand of equation (2).

Cause-specific hazard models given by equation (2) have some advantages. Survival analysts are accustomed to modeling the effects of covariates on hazards, and it is useful to know whether the effect of an increase in a covariate is to increase or decrease the underlying hazards. Second, the ingredients in equation (2) can be obtained from many types of data samples. Standard survival methods can be used to estimate these elements from full cohort data [8]. In particular, proportional hazards analyses of cohort data yield relative risk estimates, and the Breslow-estimator yields baseline cumulative hazards corresponding to  $h_{10}(t)$  and  $h_2(t)$ . Nested case-control data [11] or case-cohort data [12] yield the required relative risks. A nice feature of the case-cohort design is that it yields very intuitive estimates of the baseline cumulative hazards by reweighting the risk sets for the Breslow estimator [12]. A similar reweighting is used to estimate baseline cumulative hazards from nested case-control data [11]. Sometimes there is no identifiable underlying cohort, but case-control data are available to estimate the relative risk  $rr_1(t; X)$  compared to a reference level  $X_0$  corresponding to lowest possible risk. These data can be used to estimate the attributable risk for women of age  $t$ . Recall that the attributable risk is  $AR(t) = \{h_1^*(t) - h_{10}(t)\} / h_1^*(t)$ , where  $h_1^*(t)$  is the age-specific risk of breast cancer in the general population, which represents the “composite” age-specific hazard of a mixture of individuals with different risk factors, and  $h_{10}(t)$  would be the risk if all members of the population had the risk associated with the reference level,  $X_0$ . If one knows the distribution of risk factors in the general population, one can estimate  $AR(t)$  by expressing  $h_1^*(t)$  as an average over the risk factor distribution of  $rr_1(t; X) h_{10}(t)$  and noticing that  $h_{10}(t)$  cancels out in the computation of  $AR(t)$ . Alternatively, the cases in the case-control study may be representative of all cases that arise in the population. If so, the Bruzzi formula [13] can be used to compute  $AR(t)$ . Registries, such as the National Cancer Institute’s Surveillance Epidemiology and End Results (SEER) Program, yield “composite” age-specific hazard estimates for the general population,  $h_1^*(t)$ . Inverting the formula that defines  $AR(t)$ , one can estimate the needed baseline hazard from  $h_{10}(t) = \{1 - AR(t)\} h_1^*(t)$ . This method has been widely used to estimate absolute risk by combining case-control and registry data, as for example in [14].

A second approach models  $\Pr(a \leq T < a + \tau, c = 1 | T \geq a, X)$  itself as a function of baseline covariates  $X$  [15], rather than through their effects on underlying hazards. Non-standard software is required, but is available for cohort data ([www.cran.r-project.org/web/packages/cmprsk/cmprsk.pdf](http://www.cran.r-project.org/web/packages/cmprsk/cmprsk.pdf)). An advantage of this approach is that one can appreciate the effect of  $X$  on the absolute risk itself. However, if an increase in  $X$  results in an increase in the absolute risk, one does not know whether  $X$  increases the hazard of cause 1 or decreases the hazard of competing mortality.

Several risk models are available for breast cancer, some of which were recently compared [16,17]. For example, the National Cancer Institute has a website that includes the Breast Cancer Risk Assessment Tool (BCRAT) (<http://www.cancer.gov/bcrisktool/>), which

predicts absolute breast cancer risk based on age, age at menarche, age at first live birth, number of mother or sisters with breast cancer, number of previous benign breast biopsies, and whether atypical hyperplasia was present on any biopsy. BCRAT was based on the relative risk model in [18], but the baseline hazard for breast cancer was obtained by combining SEER data with an estimate of attributable risk [19]. Some other widely used models for breast cancer risk projection rely solely on age and family history of breast cancer, as they are based on an assumption that breast cancer is, at least in part, an autosomal dominant disease [20–22]. Another model includes family history data, information on previous *in situ* breast cancer, and other factors [23].

### 2.3 The distribution of risk in a population, and criteria for assessing risk models

Suppose covariates have a distribution  $F_X$  in the general population of age  $a$ , and that the absolute risk  $R(a, \tau, X)$  maps baseline risk factors  $X$  into probabilities on  $[0,1]$  for fixed  $a$  and  $\tau$ . (I sometime suppress arguments and write  $R(X)$  or  $R$ .) Thus,  $F_X$  induces a distribution of absolute risks  $F(r) = \Pr(R \leq r) = \int_{x: R(a, \tau, x) \leq r} dF_X(x)$ , as discussed in [24]. The distribution  $F$  [24] and its inverse [25] are useful descriptors of risks. Knowing  $F$ , one can also calculate various criteria for assessing the validity and usefulness of risk models [24,26].

**2.3.1 Calibration**—One important criterion is model calibration. Suppose the model  $R$  has been developed and is now applied to an independent sample. If the distribution of risk for this model is  $F$  in the new sample of  $N$  subjects, then the expected number of cases in a subset of the population defined by covariates,  $X \in S$ , is  $\int_{x \in S} NR(x) dF_X(x)$ . This expected count can be compared with the observed number of events as a test of calibration in various subsets,  $S$ . If the subsets  $S$  are chosen to include all values of  $X$  that lead to risks  $R(X)$  in categories defined by deciles of the distribution  $F$ , then the expected number of events falling in such a category is given by the previous expression, or equivalently by

$$N \int_{\xi_{0.1(i-1)}}^{\xi_{0.1i}} r dF(r), \text{ where } \xi_{0.1(i-1)} \text{ and } \xi_{0.1i} \text{ for } i=1, 2, \dots, 10 \text{ are successive deciles of } F.$$

Comparisons of observed with expected counts in deciles are often used to assess model calibration [27].

**2.3.2 Discriminatory accuracy**—A second criterion for evaluating risk models is discriminatory accuracy, often measured as the area under the receiver operating characteristic curve (AUC). This is the probability that a randomly selected case will have a higher projected risk than that of a randomly selected non-case. AUC is usually estimated by comparing risks  $R$  in a sample of cases with those in a sample of non-cases [28]. Thus AUC conveys no information about the actual probability of disease, unlike comparisons of observed with expected numbers of events. As described in [24], the AUC can also be calculated theoretically if the distribution  $F$  is known and the model  $R(a, \tau, X)$  is perfectly calibrated. In that case, one can calculate the distribution of risk in cases,

$$G(r) \equiv P(R \leq r | Y=1) = \mu^{-1} \int_0^r u dF(u), \text{ where } \mu = P(Y=1) = \int_0^1 r dF(r). \text{ The dummy variable } u \text{ is}$$

used instead of  $r$  to avoid confusion when  $r$  is an argument. Likewise, the distribution of risk

$$\text{in non-cases is } F_{nc}(r) \equiv P(R \leq r | Y=0) = (1 - \mu)^{-1} \int_0^r (1 - u) dF(u). \text{ Hence,}$$

$$\text{AUC} = \int_0^1 \{1 - G(r)\} dF_{nc}(r), \text{ which is also the area under a plot of } 1 - G(t) \text{ versus } 1 - F_{nc}(t) \text{ as the risk threshold } t \text{ varies.}$$

A closely related plot of great interest in public health is the plot of  $1 - G(t)$  versus  $1 - F(t)$  (Figure 1) as the risk threshold  $t$  varies. Figure 1 is a plot of the probability that a randomly selected case will have a risk above  $t$  (ordinate) against the probability that a randomly selected member of the general population will have a risk above  $t$ . The area under this curve is the chance that a randomly selected case will have a higher risk than a randomly selected member of the general population. Note that the area under this curve is not exactly the same as the previously defined AUC, because the distribution of risk in non-cases is not in general the same as the distribution of risk in the general population. However, for a rare disease, or for a more common disease on a short time interval, such as five years, the two areas are nearly equal. In what follows, we use the term AUC to be the area in under the plot of  $1 - G(t)$  versus  $1 - F(t)$ . In Figure 1, the equiangular line corresponds to a model with no discriminatory accuracy (AUC=0.5). The bold solid curve corresponds to the BCRAT model with AUC=0.607. If seven genetic markers (single nucleotide polymorphisms) that have been associated with breast cancer are added to the BCRAT model (dotted curve)[3], the AUC increases to 0.632. Adding mammographic density, instead, increases the AUC to 0.654 (dashed curve locus). A hypothetical model (very bold solid curve) with AUC 0.75 is also shown. The AUC values in Figure 1 measure how spread out the risks are in the population.

Figure 1 is also of great public health interest because it measures the extent to which risk is concentrated in a small portion of the population [29]. For example, suppose the risk threshold  $t$  is so large that only 10% of the general population have risks exceeding it, corresponding to the abscissa value 0.1. For the model with AUC=0.75, about 30% of cases exceed that threshold, indicating concentration of cases in the portion of the general population with risks in the highest decile. In contrast, if the BCRAT model is used, only about 15% of cases are concentrated there (Figure 1).

Recently, a number of other criteria have been proposed to evaluate risk models [25,30–32]. One important theme has been evaluating the models in the context of specific clinical or public health applications, where costs and benefits can be assessed and one can determine how much the use of a risk model reduces expected losses [24,26,33–35]. This calculus depends primarily on the risk distribution  $F$ . I take this approach in Sections 3 and 4.

### 3. Counseling individual patients

#### 3.1 General perspective

One use of risk models is to provide an objective measure of risk that can give the patient a realistic perspective. Many women have unrealistic notions of their risks that can lead to poor management decisions. A woman with an exaggerated estimate of risk might take a drastic preventive action, such as prophylactic mastectomy, that is not warranted by her true risk. As another example, consider the controversy regarding whether women in their forties, in whom the absolute 5 year risk of breast cancer tends to be small, should have screening mammography. Even though screening mammography is thought to reduce mortality by the same proportion in such women as in older women [36], the absolute number of breast cancer deaths saved in the younger women is small and must be weighed against the aggravation, anxiety and complications that result from false positive mammographic findings. Thus, although there is consensus that women in their fifties should be screened, the U.S. Prevention Services Task Force (USPSTF) recently stated [36]: “The USPSTF recommends against routine screening mammography in women aged 40 to 49 years. The decision to start regular, biennial screening mammography before the age of 50 years should be an individual one and take into account patient context, including the patient’s values regarding specific benefits and harms.” The USPSTF analysis ignored risk factors for breast cancer apart from age. Although age is a dominant risk factor, other factors



influence risk considerably among women in their forties. Indeed, there are many women in their forties whose absolute risks of breast cancer exceed that of a 50 year old women without risk factors [37]. For example, women in their forties with two affected first-degree relatives and women with atypical hyperplasia on a breast biopsy have higher risks than 50 year old women without risk factors, as do many women in their forties with combinations of weaker risk factors. According to the BCRAT model, a 40 year-old white woman with two affected relatives has a 1.8% chance of developing breast cancer in five years, which is larger than the average risk among 50 year-old women, 1.3%. Such a 40 year-old woman stands to gain at least as much from regular mammographic screening as the 50 year-old woman, for whom screening is widely recommended. Thus, risk assessment can provide a useful perspective for management decisions.

### 3.2 Weighing risks and benefits

Estimates of absolute risk can be used more formally to weigh the risks and benefits of an intervention to prevent breast cancer if that intervention has side effects that increase the risks of other adverse health outcomes. The weighing of risks and benefits is made possible by comparing the absolute risks of the various health outcomes, in the absence and presence of the intervention.

For example, tamoxifen not only prevents breast cancer, but it also causes certain adverse events, such as stroke and endometrial cancer, and the effects of tamoxifen on the absolute risks of the various outcomes can be assessed to determine whether the benefits from tamoxifen outweigh the risks [38]. Data from the Breast Cancer Prevention Trial (BCPT) [39] (Table II) show that tamoxifen reduces the risk of invasive breast cancer and hip fractures by nearly half, compared to placebo, but increases the risk of endometrial cancer, especially in women aged 50 years or older, and the risks of stroke and pulmonary embolism. These are all called “life-threatening events.” Tamoxifen cuts the risk of *in situ* breast cancer in half, but increases the risk of deep vein thrombosis by 60%. These latter two events are called “severe events”.

How can a woman decide whether to take tamoxifen to prevent breast cancer in view of these various effects of the drug? One aid to decision is a presentation [38] that allows the woman to consider what would happen to a population of 10,000 women just like her over the next five years (Table III). Suppose that a forty year-old white woman with a uterus has an estimated five-year risk of invasive breast cancer of 2%, based on an assessment of factors like family history, age at menarche, age at first live birth and history of previous benign biopsies. In a population of 10,000 such women followed for five years in the absence of tamoxifen, one would expect 200 breast cancers, 2 hip fractures, 10 endometrial cancers, 22 strokes, and 7 pulmonary emboli (Table III). If, instead, tamoxifen is given to the entire population, 97 invasive breast cancers and 1 hip fracture will be prevented (Table III), but 16 additional endometrial cancers will be expected, as well as 13 additional strokes and 15 additional pulmonary emboli. The net number of life-threatening events prevented will be  $97+1-16-13-15=54$ . Similarly, the net number of severe events prevented will be  $53-15=38$  (Table III). Thus tamoxifen appears to reduce not only the net numbers of life-threatening events, but also the net numbers of severe events, and this intervention should be considered.

A presentation like Table III allows each woman to apply her own values to the various types of events and reach a summary decision. However, it may be helpful to form an index to obtain a single number as a guide to management. One index assigns a weight 1 to each life-threatening event and a weight 0.5 to severe events to produce a net count of “equivalent life-threatening events” prevented [38]. In the previous example, the net benefit index or net number of equivalent life-threatening events prevented was  $54+0.5 \times 38=73$ . If one accepts

such an index, one can make generalizations about which types of women might benefit from taking tamoxifen, namely those for whom the index is demonstrably positive [38]. Table IV gives data from reference [38]. White women in the age range 40–49 have positive net benefit indices that increase with increasing breast cancer risk, because more breast cancer is prevented in those with higher breast cancer risk (Table IV). White women aged 50–59 years have less favorable net benefit indices because their baseline risks of endometrial cancer and stroke are higher and because tamoxifen has a higher relative risk for endometrial cancer in older women (Table II). Indeed, for such women with a 2% invasive breast cancer risk, the net benefit index is  $-75$ , indicating that risks outweigh benefits. For black women, a similar pattern is observed, but the net benefit indices are smaller because black women tend to have higher stroke rates than white women of the same age. Reference [38] gives methods to assess the probability that the estimated net benefit index indeed exceeds zero; these methods take variability of the relative risks in Table II into account.

From this analysis, it is clear that there is no single breast cancer risk threshold that should be used to determine whether a woman should take tamoxifen. Rather, the threshold for a particular woman should depend on her risks of the other events in Table II and be high enough to insure that the benefits of breast cancer risk reduction outweigh the risks of adverse events. The concept of absolute risk is central to this type of reasoning, because it allows one to weigh the effects of the intervention on the absolute risks of the various relevant health outcomes. In reference [38], the absolute risks of events like stroke were only assumed to depend on age and race. Decision-making based on indices of this type might be improved by having more detailed absolute risk models for the various health outcomes.

## 4. Some applications of absolute risk in public health

### 4.1 Designing prevention trials

Models of absolute risk are useful for designing prevention trials for two reasons. First, the statistical power of such trials depends on the number of events, such as incident breast cancer, that develop during the trial. The expected number of such events is the sum over trial participants of their absolute risks. Thus, the BCRAT model was used to plan the size and duration of the BCPT trial of tamoxifen [39] and the Study of Tamoxifen and Raloxifene (STAR) trial [40]. In both trials, BCRAT predicted observed numbers of events well, an indication of good calibration. High discriminatory accuracy is not needed for such sample size calculations.

There are also ethical reasons to consider the use of absolute risk models to define eligibility criteria for intervention trials. For example, in the BCPT, it was decided that women under age 60 years old should have a five year risk of invasive breast cancer of at least 1.66%, which is that of an average 60 year old woman, to participate. The reasoning was that younger women with lower invasive breast cancer risks would not stand to gain much from tamoxifen and should therefore not be eligible to participate.

### 4.2 Assessing the effects of modifications of the risk factor distribution on the absolute risk of disease in the population and in subgroups of the population

The average absolute risk over a defined time period in a given age group the population can be obtained by averaging the absolute risks over the distribution of risk factors,  $F_X$ , in that age group. If it is possible to intervene to reduce exposure to a particular risk factor or risk factors, a new distribution,  $F_X^*$ , is induced, and one can calculate the reduction in average absolute risk in the population that might result from this modification of the exposure



distribution. The reduction in absolute risk can be more informative than the analogous fractional reduction in risk. One can also calculate the reduction in absolute risks from such modification within subsets of the population, such as among women with a positive family history of breast cancer. A risk model need not have high discriminatory accuracy to be useful for such calculations, but it must be well calibrated. Of course, calculations such as these only give an indication of the types of intervention effects that might result from modifying risk factor distributions. These calculations do not take into account the feasibility of modifying the risk factor distribution, nor can they anticipate the actual effects of the interventions chosen to cause the modification in risk factors. The actual effects may be quite different from the associations estimated from observational studies, and data from observational studies are typically used to define absolute risk models.

### 4.3 Implementing a “high risk” prevention strategy

Absolute risk models can be used to define high risk subsets of the population that might benefit from a preventive intervention. Geoffrey Rose [1] distinguished between the “population strategy” of disease prevention and the “high risk strategy.” If one is recommending a very safe intervention, such as an advertising campaign to urge members of the public to consume less salt, one can apply the intervention to the entire population. If everyone in the population took up this suggestion and had a small drop in systolic blood pressure of two millimeters of mercury, more heart attacks would be prevented than by a “high risk” strategy that involved identifying those members of the population who had the highest blood pressure and treating them vigorously [1]. The general population strategy is more effective, because many heart attacks occur among people with normal or even low blood pressure, even though the risk of heart attack increases with blood pressure. Moreover, risk assessment is not needed to implement a general population strategy. Thus, when feasible, a general population strategy is to be recommended.

There are two main reasons why the “high risk strategy” has a place in preventive medicine. The most important reason is that available interventions, such as the use of tamoxifen to prevent breast cancer, carry risks. Surgical interventions, such as oophorectomy or mastectomies to prevent breast cancer, are effective [41,42], but have adverse effects, and are therefore appropriate only for subsets of women at very high risk, for whom benefits exceed risks. A second reason for implementing a “high risk strategy” is economic. The intervention may be so expensive that it can only be made available to those at highest risk. In both these settings, a well calibrated model with high discriminatory accuracy can improve performance, compared to a model with less discriminatory accuracy, because the more risk is concentrated into a small subset of the population at highest risk, the more effective the “high risk strategy” will be.

The use of tamoxifen to prevent breast cancer illustrates the need for a “high risk strategy” to avoid exposing many women in the population to a potentially toxic intervention. In the absence of tamoxifen, one would expect to see 589.6 life-threatening events in one year in a population of 100,000 white women with a uterus and aged 50–59 years old (Table V). If tamoxifen were given to all these women, invasive breast cancers and hip fractures would be cut nearly in half, but the increases in endometrial cancer, stroke and pulmonary embolism would more than compensate, resulting in an increase to 833.5 expected life-threatening events, a public health calamity. Clearly, tamoxifen must be directed to a high risk subgroup if it is to be useful in breast cancer prevention.

To get some idea of how high the risk of breast cancer,  $R$ , must be, one needs to compute the expected reduction in life-threatening events per  $10^5$  women per year from tamoxifen. From the relative risks and the baseline risks in the absence of tamoxifen in Tables V, the expected reduction in life-threatening events is

$$R(1-0.51)+101.6(1-0.55)+81.4(1-4.01)+110.0(1-1.59)+50.0(1-3.01)=0.49R-363.7.$$

The first term on the left corresponds to the expected reduction in invasive breast cancers, the second to the reduction in hip fractures, the third to the increase in endometrial cancers, the fourth to the increase in strokes, and the fifth to the increase in pulmonary emboli. Solving for the invasive breast cancer risk that makes this expectation positive, we obtain the threshold value  $r^* = 774.3 \times 10^{-5}$ . If tamoxifen is given to a woman with this risk or higher, the expected reduction in life-threatening events is positive. This threshold is optimal for any well calibrated risk model. For the BCRAT model, whose AUC = 0.607, only 1% of white women aged 50–59 years have invasive breast cancer risks above  $774.3 \times 10^{-5}$ , and most of the breast cancers arise among women with lower risks. Thus, the high risk strategy based on BCRAT is expected to have a limited impact on breast cancer prevention. In fact, giving tamoxifen to this high risk group reduces the expected number of life-threatening events per year in 100,000 women by only 1.4, from 589.6 to 588.2. If a slightly more discriminating model with AUC = 0.632 is used, that includes the factors in BCRAT and 7 SNPs, the number of life-threatening events per year with the same risk threshold is reduced by 1.8, from 589.6 to 587.8 [26]. Thus, neither of these models produces a large reduction in life-threatening events from tamoxifen. If a perfectly discriminating breast cancer risk model were available, much larger risk reductions could be achieved, because all women destined to develop breast cancer in the absence of tamoxifen would have risks above the threshold, and all women not so destined would have risks below the threshold. Thus, only those 246.6 women destined to develop breast cancer would be given tamoxifen, and because this is such a small number of women, there would be few adverse events [26], resulting in a net reduction of 119.9 life-threatening events, from 589.6 to 469.7.

This example illustrates the limitations of the “high risk strategy” for an intervention with adverse effects. Because most risk models have limited discriminatory accuracy, the potential for prevention is small if the high risk group is small. There are two ways to prevent more disease with a high risk strategy. One is to improve the discriminatory accuracy of risk models. This has been frustratingly hard, because discrete risk factors must have very high relative risks to achieve high discriminatory accuracy [43], and continuous risk factors need to have much larger variation in associated risk than has been found, for example, from polygenic models of SNP effects [44]. Studies of risk models for cardiovascular disease often report higher AUC values, but these models also include age as a regression factor, and these studies include age effects in the calculation of AUC. The AUC values reported for breast cancer are age-specific and address how different the distributions of risk are in cases and non-cases of the same age. Because age is a strong risk factor for many chronic diseases, discriminatory accuracy will tend to be higher if one compares risks among cases and non-cases of various ages.

Earlier studies of breast and cervical cancer showed that high risk subgroups contained fewer than half the women with breast cancers [45] or with cervical cancers or advanced precancerous cervical lesions [46]. However, identification of specific carcinogenic strains of human papillomavirus has led to the definition of high risk groups containing up to 90% of cervical cancers or advanced precancerous cervical lesions in some populations [47].

A second approach to improve the preventive value of the high risk strategy is to find interventions with fewer adverse effects. For example, the drug raloxifene also prevents breast cancer and causes less endometrial cancer than tamoxifen [40], and other chemopreventive agents are in development [48]. A less toxic intervention may be advantageously given to a larger “high risk group”, thus preventing more disease.

#### 4.4 Allocation of public health resources under cost constraints

Disease prevention efforts are often constrained by limited resources, and one way to allocate those resources is to direct them to those who stand to benefit most, who are usually those at highest risk of the disease. For example, an expert panel convened by the American Cancer Society recommended [49] magnetic resonance imaging to screen for breast cancer only in women with an estimated lifetime risk of “approximately 20–25% or greater.” It is possible to take a formal approach to maximizing the population health benefit of a preventive action, subject to cost constraints, by using the distribution of risk in the population for a give risk model, and its associated Lorenz curve of risks [50].

As an example, suppose there is not enough money to support a program of screening mammography for an entire population, such as white women aged 50–59 years. Recent data [36] indicate that mammographic screening reduces deaths from breast cancer by 14% in this age range. Whatever the actual reduction in mortality is if mammography is offered to all women, it would be only half as much if there were only enough money to screen half the women, chosen at random. However, if women were first ranked by their risks of breast cancer, and if mammograms were then allocated to women in decreasing order of risk until the money ran out, more lives might be saved. In this example, mammography is regarded as an intervention, and it is not to be confused with the “risk assessment”, which refers to gathering the data needed for a risk model such as BCRAT and obtaining the resulting risk estimate. If “risk assessment” is too expensive, there will not be enough money left over for the intervention.

The optimal risk assessment strategy will depend on the ratio,  $k$ , of the cost of risk assessment to the cost of the intervention, and on  $h$ , the ratio of the money available to the money required to give mammograms to the entire population. The optimal strategy will also depend on the Lorenz curve of the distribution of risks in the population,  $F$ . The Lorenz curve  $L(p)$  is the proportion of the total population risk possessed by the portion of the

population with risks below the  $p^{\text{th}}$  quantile of  $F$ . To be precise,  $L(p) = \mu^{-1} \int_0^{\xi_p} r dF(r) = G(\xi_p)$ , where  $\xi_p = F^{-1}(p)$ . From this relationship, it can be seen that  $1 - L(1-p) = 1 - G(\xi_{1-p})$  is the proportion of cases that develop in the proportion  $p$  of the population at highest risk. Thus the Lorenz curve conveniently captures the degree to which cases are concentrated in those at highest risk. In discussing Figure 1 (Section 2.3.2), I described a similar concentration of cases. Indeed, the Lorenz curves can be obtained from Figure 1 by reflections about its ordinate and abscissa.

The optimal risk allocation strategy depends on three parameters. The first is the fraction of the entire population that will be offered risk assessment,  $g$ . The second is the proportion  $p$  of those offered risk assessment who will be offered mammograms in descending order of risk. The third is the proportion of the population who will not be offered risk assessment but who will receive mammography at random,  $m$ .

To determine the optimal strategy, consider the number of breast cancer deaths saved by a strategy  $(g, p, m)$ . This number is

$$N\mu - Ng \int_0^{\xi_{1-p}} r dF(r) - Ng(1-p) \int_{\xi_{1-p}}^1 r dF(r) - N(1-g)\mu m(1-p) - N(1-g)(1-m)\mu = N\mu\rho[g\{1 - L(1-p)\} + (1-g)m], \quad (3)$$

where  $N$  is the size of the population,  $\mu$  is the average absolute risk, and  $\rho$  is the fractional reduction in mortality from mammography. For example, if mammography reduces mortality by 14%, then  $\rho = 0.14$ , and the effect of mammography is to decrease mortality in the population by the factor  $(1 - \rho) = 0.84$ . The third term in the top line of equation (3) represents the deaths among the  $pNg$  women whose risks were assessed risks and whose risks were among the proportion  $p$  with highest risks; these women received mammograms. The fourth term represents deaths among the  $Nm$  women who received mammograms at random. The ratio of the numbers of lives saved under the strategy  $(g, p, m)$  to the number who would be saved if all women got mammograms,  $N\mu\rho$ , is

$$[g\{1 - L(1 - p)\} + (1 - g)m]. \quad (4)$$

I call the quantity in expression (4) the “fraction of lives saved.” The optimal strategy  $(g, p, m)$  maximizes equation (4) subject to the cost constraints

$$gk + gp + (1 - g)m \leq h, \quad (5)$$

in which the unit of cost is the amount of money required to give mammograms to the entire population. If the cost ratio  $k$  exceeds about 0.20, risk assessment is too expensive to be useful, and all mammograms should be allocated at random, namely  $g=0$  and  $m=1$  [50].

On the assumption that the cost ratio is only  $k=0.02$ , which is plausible for a questionnaire-based risk assessment like BCRAT, the optimal strategy is to assess the risk in all subjects and use the remaining money to give mammograms to those at highest risk until the money runs out ( $g=1, m=0$ ). For BCRAT, this strategy yields a fraction of lives saved of 0.632, which is 26.4% better than random allocation of mammograms, for which the fraction of lives saved is 0.500 (Table VI). If the model BCRAT+7SNPs that also includes SNPs is used, its higher AUC results in a fraction of lives saved of 0.667, an improvement of 33.4% over random allocation. A perfect model would assure that the small number of women destined to develop breast cancer would receive mammograms, resulting in a fraction of lives saved of 1.000. At the present time, however, the costs of obtaining DNA and measuring SNPs are too high to employ in risk-based allocation, and there is no perfectly discriminating model.

These calculations indicate the potential for allocating resources based on absolute risk models. The example is unrealistic in some regards, however. First, in countries like the U.S., cost constraints on provision of mammographic screening are not compelling, and insurance plans often provide for mammographic screening. Implicit in equation (3) is the assumption that the women who accept the offer of risk assessment are representative of the entire population. Another implicit assumption is that women will not obtain mammograms unless it is paid for through the screening strategy outlined. Further work is needed to refine this approach to allow for self-selection and for alternative sources of access. Nonetheless, many of the parameters employed, such as the Lorenz curve of risk in the population and certain cost ratios will be important in fashioning an effective allocation program.

## 5. Discussion

I have tried to indicate where models of absolute breast cancer risk are useful in clinical management and disease prevention. In addition to giving important general perspective to patients, such models can be used to weigh risks and benefits of a preventive outcome

formally (Section 3). In the context of disease prevention, these models are useful for designing intervention trials (Section 4.1), and for assessing the potential absolute reductions in disease risk that might result from reductions in modifiable exposures in the population (Section 4.2). These four applications do not require high discriminatory accuracy, but do depend on good model calibration. Absolute risk models are needed to implement “high risk strategies” for disease prevention, for which good discriminatory accuracy is needed, in addition to good calibration, in order to achieve large reductions in disease incidence (Section 4.3). Absolute risk models may also play a role in the productive allocation of prevention resources under cost constraints (Section 4.4). Here again, good discriminatory accuracy is advantageous, and risk assessment should be inexpensive in comparison with the cost of intervention.

This paper has discussed the use of absolute risk models for predicting disease risk and for disease prevention, but models of absolute risk are also very important in patient management following disease onset. For example, if a 65 year old man has a recent diagnosis of prostate cancer with favorable prognostic indicators, he may have a small absolute risk of dying of prostate cancer, because competing causes of mortality are likely to kill him first. Such a patient might be well advised to defer surgical or radiation treatment unless later indications, such as increasing prostate-specific antigen levels, suggest a need for active intervention. Thus the absolute risk of disease-specific mortality can influence management decisions following disease onset.

The disappointing results of the “high risk strategy” for breast cancer prevention reflect the twin weaknesses of risk models with limited discriminatory accuracy, and interventions that are too toxic to benefit all but those at highest breast cancer risk. As mentioned previously, it is very hard to increase discriminatory accuracy [43,51] unless strong risk factors are available. The percentage of the area that is radiographically dense on a mammogram, called mammographic density, is a strong risk factor that has been used in absolute risk models [52], as has a related radiographic feature, the Breast Imaging Reporting Data System (BIRADS) [53]. When added to BCRAT, mammographic density increased the AUC by 0.047 [52], which is nearly twice as much as the increase from adding SNPs (Figure 1) [3]. Although further gains could be expected from adding SNPs to a model with mammographic density, the gains are likely to be modest [44]. Another strong risk factor is pathology identified in breast biopsies [54], but such information is only available from women with biopsies. However, unless a nearly painless way can be found to obtain strongly predictive pathology or biomarker data from women in the general population, the discriminatory accuracy of available models will remain a limiting factor in the “high risk” prevention strategy.

An alternative approach that relies less strongly on the risk model is to use interventions with fewer adverse effects. Such interventions can be beneficial for women with lower breast cancer risks, and if the intervention is safe enough, breast cancer risk assessment may not be needed at all. Primarily on the basis of observational studies, Cummings et al. [55] noted that “exercise, weight reduction, low-fat diet, and reduced alcohol intake were associated with decreased risk of breast cancer” in most studies. A randomized trial in the Women’s Health Initiative [56] indicated a 9% reduction in the incidence of invasive breast cancer in the low-fat diet group, compared to placebo, which almost attained statistical significance (95% confidence interval (CI): -1% to 17% risk reduction). Because these lifestyle changes are thought to be safe, Cummings et al. concluded that they “should be recommended regardless of (breast cancer) risk” for post-menopausal women. The Women’s Health Initiative randomized trial showed an increase in invasive breast cancer incidence of 26% (95% CI: 0 to 59%) in post-menopausal women receiving hormone replacement with estrogen and progestin, compared to placebo [57]. Since that trial, sales of

this type of hormone replacement therapy have dropped, and breast cancer incidence rates have declined [58]. Avoidance of this exposure is another safe way to reduce breast cancer risk, based on evidence from the randomized trial.

As more is learned about the etiology of breast cancer, there may be other safe preventive interventions that are even more effective. If a viral agent were found to be important, safe vaccines might be developed. Promising early results suggest this will be an effective approach for human papillomavirus [59], which causes cervical cancer, and for hepatitis B [60], which causes liver cancer. Armitage and Doll studied cancer incidence trends and proposed a multistage model of carcinogenesis to explain the exponential increase in cancer incidence with age for many cancers [61]. Subsequent studies in cancer biology have identified several changes that a cell must undergo in order to become cancerous [62]. If safe interventions could be found that reduced the rates of several of these transitions, the compound effect could be to substantially decrease breast cancer incidence rates and the strong dependence of incidence on age. Migration studies indicate that Asian women, whose rates of breast cancer incidence were low in Asia, have much higher rates in the U.S. a generation or two after migrating [63], and in Shanghai, breast cancer incidence rates have been increasing rapidly during a period of changing dietary and lifestyle patterns [64]. These facts suggest that lifestyle exposures play a prominent role in breast cancer carcinogenesis and may offer safe opportunities for prevention.

In summary, models of absolute risk currently have a useful but limited role in counseling and in breast cancer prevention. Efforts to increase discriminatory accuracy can expand that role, but increased success in disease prevention will depend on safer and more effective interventions that may or may not need to be used in conjunction with risk models.

## Acknowledgments

This work was supported by the Intramural Research Program of the Division of Cancer Epidemiology and Genetics, National Cancer Institute, National Institutes of Health. I would like to thank Dr. Ruth Pfeiffer and the reviewers and editor for helpful suggestions. I would also like to thank Professor Peter Armitage, whose writings and leadership are an inspiration.

## References

1. Rose, GA. The Strategy of Preventive Medicine. Oxford University Press; Oxford: 1992.
2. Feero WG, Guttmacher AE, Collins FS. The genome gets personal -Almost. *Jama-Journal of the American Medical Association*. 2008; 299:1351–1352.
3. Gail MH. Discriminatory accuracy from single-nucleotide polymorphisms in models to predict breast cancer risk. *Journal of the National Cancer Institute*. 2008; 100:1037–1041. [PubMed: 18612136]
4. Pharoah PDP, Antoniou AC, Easton DF, Ponder BAJ. Polygenes, risk prediction, and targeted prevention of breast cancer. *New England Journal of Medicine*. 2008; 358:2796–2803. [PubMed: 18579814]
5. 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. *Journal of the National Cancer Institute*. 2001; 93:358–366. [PubMed: 11238697]
6. Wacholder S, Hartge P, Prentice R, Garcia-Closas M, Feigelson HS, Diver WR, Thun MJ, Cox DG, Hankinson SE, Kraft P, Rosner B, Berg CD, Brinton LA, Lissowska J, Sherman ME, Chlebowski R, Kooperberg C, Jackson RD, Buckman DW, Hui P, Pfeiffer R, Jacobs KB, Thomas GD, Hoover RN, Gail MH, Chanock SJ, Hunter DJ. Performance of common genetic variants in breast-cancer risk models. *New England Journal of Medicine*. 2010; 362:986–993. [PubMed: 20237344]
7. Wilson PWF, D'Agostino RB, Levy D, Belanger AM, Silbershatz H, Kannel WB. Prediction of coronary heart disease using risk factor categories. *Circulation*. 1998; 97:1837–1847. [PubMed: 9603539]

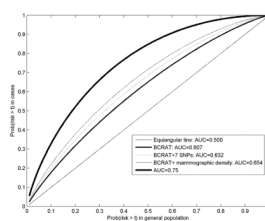


8. Prentice RL, Kalbfleisch JD, Peterson AV, Flournoy N, Farewell VT, Breslow NE. Analysis of failure times in the presence of competing risks. *Biometrics*. 1978; 34:541–554. [PubMed: 373811]
9. Gaynor JJ, Feuer EJ, Tan CC, Wu DH, Little CR, Straus DJ, Clarkson BD, Brennan MF. On the use of cause-specific failure and conditional failure probabilities-examples from clinical oncology data. *Journal of the American Statistical Association*. 1993; 88:400–409.
10. Lloyd-Jones DM, Leip EP, Larson MG, D'Agostino RB, Beiser A, Wilson PWF, Wolf PA, Levy D. Prediction of lifetime risk for cardiovascular disease by risk factor burden at 50 years of age. *Circulation*. 2006; 113:791–798. [PubMed: 16461820]
11. Langholz B, Borgan O. Estimation of absolute risk from nested case-control data (vol 53, pg 767, 1997). *Biometrics*. 2003; 59:451–451.
12. Self SG, Prentice RL. Asymptotic-distribution theory and efficiency results for case cohort studies. *Annals of Statistics*. 1988; 16:64–81.
13. Bruzzi P, Green SB, Byar DP, Brinton LA, Schairer C. Estimating the population attributable risk for multiple risk-factors using case-control data. *American Journal of Epidemiology*. 1985; 122:904–913. [PubMed: 4050778]
14. Gail MH, Costantino JP, Pee D, Bondy M, Newman L, Selvan M, Anderson GL, Malone KE, Marchbanks PA, McCaskill-Stevens W, Norman SA, Simon MS, Spirtas R, Ursin G, Bernstein L. Projecting individualized absolute invasive breast cancer risk in African American women. *Journal of the National Cancer Institute*. 2007; 99:1782–1792. [PubMed: 18042936]
15. Fine JP, Gray RJ. A proportional hazards model for the subdistribution of a competing risk. *Journal of the American Statistical Association*. 1999; 94:496–509.
16. Amir E, Freedman OC, Seruga B, Evans DG. Assessing Women at High Risk of Breast Cancer: A Review of Risk Assessment Models. *Journal of the National Cancer Institute*. 2010; 102:680–691. [PubMed: 20427433]
17. Gail MH, Mai PL. Comparing Breast Cancer Risk Assessment Models. *Journal of the National Cancer Institute*. 2010; 102:665–688. [PubMed: 20427429]
18. Gail MH, Brinton LA, Byar DP, Corle DK, Green SB, Schairer C, Mulvihill JJ. Projecting individualized probabilities of developing breast cancer for white females who are being examined annually. *Journal of the National Cancer Institute*. 1989; 81:1879–1886. [PubMed: 2593165]
19. Costantino JP, Gail MH, Pee D, Anderson S, Redmond CK, Benichou J, Wieand HS. Validation studies for models projecting the risk of invasive and total breast cancer incidence. *Journal of the National Cancer Institute*. 1999; 91:1541–1548. [PubMed: 10491430]
20. Antoniou AC, Cunningham AP, Peto J, Evans DG, Lalloo F, Narod SA, Risch HA, Eyfjord JE, Hopper JL, Southey MC, Olsson H, Johannsson O, Borg A, Passini B, Radice P, Manoukian S, Eccles DM, Tang N, Olah E, Anton-Culver H, Warner E, Lubinski J, Gronwald J, Gorski B, Tryggvadottir L, Syrjakoski K, Kallioniemi OP, Eerola H, Nevanlinna H, Pharoah PDP, Easton DF. The BOADICEA model of genetic susceptibility to breast and ovarian cancers: updates and extensions. *British Journal of Cancer*. 2008; 98:1457–1466. [PubMed: 18349832]
21. Berry DA, Iversen ES, Gudbjartsson DF, Hiller EH, Garber JE, Peshkin BN, Lerman C, Watson P, Lynch HT, Hilsenbeck SG, Rubinstein WS, Hughes KS, Parmigiani G. BRCAPRO validation, sensitivity of genetic testing of BRCA1/BRCA2, and prevalence of other breast cancer susceptibility genes. *Journal of Clinical Oncology*. 2002; 20:2701–2712. [PubMed: 12039933]
22. Claus EB, Risch N, Thompson WD. Autosomal-dominant inheritance of early-onset breast-cancer -implications for risk prediction. *Cancer*. 1994; 73:643–651. [PubMed: 8299086]
23. Tyrer J, Duffy SW, Cuzick J. A breast cancer prediction model incorporating familial and personal risk factors (vol 23, pg 1111, 2004). *Statistics in Medicine*. 2005; 24:156–156.
24. Gail MH, Pfeiffer RM. On criteria for evaluating models of absolute risk. *Biostatistics*. 2005; 6:227–239. [PubMed: 15772102]
25. Pepe MS, Feng Z, Huang Y, Longton G, Prentice R, Thompson IM, Zheng Y. Integrating the predictiveness of a marker with its performance as a classifier. *American Journal of Epidemiology*. 2008; 167:362–368. [PubMed: 17982157]
26. Gail MH. Value of adding single-nucleotide polymorphism genotypes to a breast cancer risk model. *Journal of the National Cancer Institute*. 2009; 101:959–963. [PubMed: 19535781]

27. Hosmer DW, Hosmer T, leCessie S, Lemeshow S. A comparison of goodness-of-fit tests for the logistic regression model. *Statistics in Medicine*. 1997; 16:965–980. [PubMed: 9160492]
28. Pepe, MS. *The Statistical Evaluation of Medical Tests for Classification and Prediction*. Oxford University Press; New York: 2003.
29. Pharoah PDP, Antoniou A, Bobrow M, Zimmern RL, Easton DF, Ponder BAJ. Polygenic susceptibility to breast cancer and implications for prevention. *Nature Genetics*. 2002; 31:33–36. [PubMed: 11984562]
30. Cook NR. Use and misuse of the receiver operating characteristic curve in risk prediction. *Circulation*. 2007; 115:928–935. [PubMed: 17309939]
31. Cook NR. Statistical evaluation of prognostic versus diagnostic models: Beyond the ROC curve. *Clinical Chemistry*. 2008; 54:17–23. [PubMed: 18024533]
32. Pencina MJ, D'Agostino RB, Vasan RS. Evaluating the added predictive ability of a new marker: From area under the ROC curve to reclassification and beyond. *Statistics in Medicine*. 2008; 27:157–172. [PubMed: 17569110]
33. Baker SG, Cook NR, Vickers A, Kramer BS. Using relative utility curves to evaluate risk prediction. *Journal of the Royal Statistical Society Series A*. 2009
34. Pauker SG, Kassirer JP. Therapeutic decision-making -cost-benefit analysis. *New England Journal of Medicine*. 1975; 293:229–234. [PubMed: 1143303]
35. Vickers AJ, Elkin EB. Decision curve analysis: A novel method for evaluating prediction models. *Medical Decision Making*. 2006; 26:565–574. [PubMed: 17099194]
36. Calonge N, Petitti DB, DeWitt TG, Dietrich AJ, Gregory KD, Grossman D, Isham G, LeFevre ML, Leipzig RM, Marion LN, Melnyk B, Moyer VA, Ockene JK, Sawaya GF, Schwartz JS, Wilt T. Screening for Breast Cancer: US Preventive Services Task Force Recommendation Statement. *Annals of Internal Medicine*. 2009; 151:716–W236. [PubMed: 19920272]
37. Gail MH, Schairer C. Comments and response on the USPSTF recommendation on screening for breast cancer. *Annals of Internal Medicine*. 2010; 152:540. [PubMed: 20157102]
38. Gail MH, Costantino JP, Bryant J, Croyle R, Freedman L, Helzlsouer K, Vogel V. Weighing the risks and benefits of tamoxifen treatment for preventing breast cancer. *Journal of the National Cancer Institute*. 1999; 91:1829–1846. [PubMed: 10547390]
39. Fisher B, Costantino JP, Wickerham DL, Redmond CK, Kavanah M, Cronin WM, Vogel V, Robidoux A, Dimitrov N, Atkins J, Daly M, Wieand S, Tan-Chiu E, Ford L, Wolmark N. Tamoxifen for prevention of breast cancer: Report of the National Surgical Adjuvant Breast and Bowel Project P-1 study. *Journal of the National Cancer Institute*. 1998; 90:1371–1388. [PubMed: 9747868]
40. Vogel VG, Costantino JP, Wickerham DL, Cronin WM, Cecchini RS, Atkins JN, Bevers TB, Fehrenbacher L, Pajon ER, Wade JL, Robidoux A, Margolese RG, James J, Lippman SM, Runowicz CD, Ganz PA, Reis SE, McCaskill-Stevens W, Ford LG, Jordan VC, Wolmark N. Effects of tamoxifen vs raloxifene on the risk of developing invasive breast cancer and other disease outcomes -The NSABP study of tamoxifen and raloxifene (STAR) P-2 trial. *Jama-Journal of the American Medical Association*. 2006; 295:2727–2741.
41. Hartmann LC, Sellers TA, Schaid DJ, Frank TS, Soderberg CL, Sitta DL, Frost MH, Grant CS, Donohue JH, Woods JE, McDonnell SK, Vockley CW, Deffenbaugh A, Couch FJ, Jenkins RB. Efficacy of bilateral prophylactic mastectomy in BRCA1 and BRCA2 gene mutation carriers. *Journal of the National Cancer Institute*. 2001; 93:1633–1637. [PubMed: 11698567]
42. McDonnell SK, Schaid DJ, Myers JL, Grant CS, Donohue JH, Woods JE, Frost MH, Johnson JL, Sitta DL, Slezak JM, Crotty TB, Jenkins RB, Sellers TA, Hartmann LC. Efficacy of contralateral prophylactic mastectomy in women with a personal and family history of breast cancer. *Journal of Clinical Oncology*. 2001; 19:3938–3943. [PubMed: 11579114]
43. Pepe MS, Janes H, Longton G, Leisenring W, Newcomb P. Limitations of the odds ratio in gauging the performance of a diagnostic, prognostic, or screening marker. *American Journal of Epidemiology*. 2004; 159:882–890. [PubMed: 15105181]
44. Park JH, Wacholder S, Gail MH, Peters U, Jacobs KB, Chanock SJ, Chatterjee N. Estimation of effect size distribution from genome-wide association studies and implications for future discoveries. *Nature Genetics*. 42:570–U139. [PubMed: 20562874]

45. Farewell VT, Math B, Math M. Combined effect of breast cancer risk factors. *Cancer*. 1977; 40:931–936. [PubMed: 890675]
46. Hakama M, Pukkala E, Saastamoinen P. Selective screening -theory and practice based on high-risk groups of cervical cancer. *Journal of Epidemiology and Community Health*. 1979; 33:257–261. [PubMed: 536674]
47. Schiffman M, Khan MJ, Solomon D, Herrero R, Wacholder S, Hildesheim A, Rodriguez AC, Bratti MC, Wheeler CM, Burk RD. A study of the impact of adding HPV types to cervical cancer screening and triage tests. *Journal of the National Cancer Institute*. 2005; 97:147–150. [PubMed: 15657345]
48. Cuzick J. Chemoprevention of breast cancer. *Breast Cancer Research*. 2008; 15:10–16.
49. Saslow D, Boetes C, Burke W, Harms S, Leach MO, Lehman CD, Morris E, Pisano E, Schnall M, Sener S, Smith RA, Warner E, Yaffe M, Andrews KS, Russell CA. American Cancer Society guidelines for breast screening with MRI as an adjunct to mammography. *Ca-a Cancer Journal for Clinicians*. 2007; 57:75–89. [PubMed: 17392385]
50. Gail MH. Applying the Lorenz curve to disease risk to optimize health benefits under cost constraints. *Statistics and Its Interface*. 2009; 2:117–121. [PubMed: 19779595]
51. Park Y, Freedman AN, Gail MH, Pee D, Hollenbeck A, Schatzkin A, Pfeiffer RM. Validation of a colorectal cancer risk prediction model among white patients age 50 years and older. *Journal of Clinical Oncology*. 2009; 27:694–698. [PubMed: 19114700]
52. Chen JB, Pee D, Ayyagari R, Graubard B, Schairer C, Byrne C, Benichou J, Gail MH. Projecting absolute invasive breast cancer risk in white women with a model that includes mammographic density. *Journal of the National Cancer Institute*. 2006; 98:1215–1226. [PubMed: 16954474]
53. Tice JA, Cummings SR, Smith-Bindman R, Ichikawa L, Barlow WE, Kerlikowske K. Using clinical factors and mammographic breast density to estimate breast cancer risk: development and validation of a new predictive model. *Annals of Internal Medicine*. 2008; 148:337–347. [PubMed: 18316752]
54. Degnim AC, Visscher DW, Berman HK, Frost MH, Sellers TA, Vierkant RA, Maloney SD, Pankratz VS, de Groen PC, Lingle WL, Ghosh K, Penheiter L, Tlsty T, Melton LJ, Reynolds CA, Hartmann LC. Stratification of breast cancer risk in women with atypia: A Mayo cohort study. *Journal of Clinical Oncology*. 2007; 25:2671–2677. [PubMed: 17563394]
55. Cummings SR, Tice JA, Bauer S, Browner WS, Cuzick J, Ziv E, Vogel V, Shepherd J, Vachon C, Smith-Bindman R, Kerlikowske K. Prevention of Breast Cancer in Postmenopausal Women: Approaches to Estimating and Reducing Risk. *Journal of the National Cancer Institute*. 2009; 101:384–398. [PubMed: 19276457]
56. Prentice RL, Caan B, Chlebowski RT, Patterson R, Kuller LH, Ockene JK, Margolis KL, Limacher MC, Manson JE, Parker LM, Paskett E, Phillips L, Robbins J, Rossouw JE, Sarto GE, Shikany JM, Stefanick ML, Thomson CA, Van Horn L, Vitamins MZ, Wactawski-Wende J, Wallace RB, Wassertheil-Smoller S, Whitlock E, Yano K, Adams-Campbell L, Anderson GL, Assaf AR, Beresford SAA, Black HR, Brunner RL, Brzyski RG, Ford L, Gass M, Hays J, Heber D, Heiss G, Hendrix SL, Hsia J, Hubbell FA, Jackson RD, Johnson KC, Kotchen JM, LaCroix AZ, Lane DS, Langer RD, Lasser NL, Henderson MM. Low-fat dietary pattern and risk of invasive breast cancer -The women's health initiative randomized controlled dietary modification trial. *Jama-Journal of the American Medical Association*. 2006; 295:629–642.
57. Rossouw J, Anderson G, Prentice R, LaCroix A, Kooperberg C, Stefanick M, Jackson R, Beresford SAA, Howard B, Johnson K, Kotchen J, Ockene J. Risks and benefits of estrogen plus progestin in healthy postmenopausal women: principal results From the Women's Health Initiative randomized controlled trial. *JAMA (Chicago, Ill)*. 2002; 288:321–333.
58. Ravdin PM, Cronin KA, Howlader N, Berg CD, Chlebowski RT, Feuer EJ, Edwards BK, Berry DA. The decrease in breast-cancer incidence in 2003 in the United States. *New England Journal of Medicine*. 2007; 356:1670–1674. [PubMed: 17442911]
59. Harper DM, Franco EL, Wheeler C, Ferris DG, Jenkins D, Schuind A, Zahaf T, Innis B, Naud P, De Carvalho NS, Roteli-Martins CM, Teixeira J, Blatter MM, Korn AP, Quint W, Dubin G. Efficacy of a bivalent L1 virus-like particle vaccine in prevention of infection with human papillomavirus types 16 and 18 in young women: a randomised controlled trial. *Lancet*. 2004; 364:1757–1765. [PubMed: 15541448]

60. Chen CJ, You SL, Lin LH, Hsu WL, Yang YW. Cancer epidemiology and control in Taiwan: a brief review. *Japanese Journal of Clinical Oncology*. 2002; 32:S66–S81. [PubMed: 11959880]
61. Armitage P, Doll R. The age distribution of cancer and a multi-stage theory of carcinogenesis. *British Journal of Cancer*. 1954; 8:1–12. [PubMed: 13172380]
62. Hahn WC, Weinberg RA. Modelling the molecular circuitry of cancer. *Nature Reviews Cancer*. 2002; 2:331–341.
63. Ziegler RG, Hoover RN, Pike MC, Hildesheim A, Nomura AMY, West DW, Wuwilliams AH, Kolonel LN, Hornross PL, Rosenthal JF, Hyer MB. Migration patterns and breast-cancer risk in Asian-American women. *Journal of the National Cancer Institute*. 1993; 85:1819–1827. [PubMed: 8230262]
64. Ziegler RG, Anderson WF, Gail MH. Increasing breast cancer incidence in china: The numbers add up. *Journal of the National Cancer Institute*. 2008; 100:1339–1341. [PubMed: 18812546]



**Figure 1.**

Plots of the probability that risk exceeds a threshold  $t$  in cases versus the probability that the risk exceeds that threshold in the general population as  $t$  varies. The area under this curve (AUC) is a measure of discriminatory accuracy and is shown for the equiangular line, for three risk models, and for a hypothetical risk model with AUC=0.75. Curves were calculated assuming risks had beta distributions with the AUC values shown. The AUC values are independent of the mean risk, which was 0.02 for each of the beta distributions.

**Table I**

Life table to compare crude with pure risk of breast cancer.

Age at start of interval (years)	# At risk	# Incident breast cancers	# Deaths from other causes
60	1000	17	44
65	939	20	63
70	856	22	89
75	745	....	....



**Table II**

Effects of tamoxifen on various life-threatening and severe events, expressed as relative risks compared to placebo.\*

Type of event	Relative risk	95% Confidence interval
Life-threatening events		
Invasive breast cancer	0.51	0.39 to 0.66
Hip fracture	0.55	0.25 to 1.15
Endometrial cancer		
All women	2.53	1.35 to 4.97
Women $\geq 50$ years old	4.01	1.70–10.90
Stroke	1.59	0.93 to 2.77
Pulmonary embolism	3.01	1.15 to 9.27
Severe events		
<i>In situ</i> breast cancer	0.50	0.33 to 0.77
Deep vein thrombosis	1.60	0.91 to 2.86

\* Data from [39].

**Table III**

Health outcomes expected in 10,000 forty year old women in five years, each of whom has a uterus and a projected five-year risk of invasive breast cancer of 2 percent.\*

	Without tamoxifen	With tamoxifen	Events prevented by tamoxifen
Life-threatening events			
Invasive breast cancer	200	103	97
Hip fracture	2	1	1
Endometrial cancer	10	26	-16
Stroke	22	35	-13
Pulmonary embolus	7	22	-15
Net life-threatening events prevented			54
Severe events			
<i>In situ</i> breast cancer	106	53	53
Deep vein thrombosis	24	37	-13
Net severe events prevented			38

\* Data from [38].

**Table IV**

Net number of life-threatening equivalent events prevented by tamoxifen in 10,000 women with uteri over five years.\*

Invasive breast cancer risk	White women		Black women	
	40–49 yrs. old	50–59 yrs. old	40–49 yrs. old	50–59 yrs. old
2%	73	–75	14	–187
4%	196	38	137	–74
6%	318	149	259	37

\* Data from [38].

**Table V**

Numbers of life-threatening events in one year in a population of 100,000 white women with uteri and aged 50–59 years without tamoxifen and with tamoxifen.\*

Health outcome	Relative risk	Events without Tamoxifen	Events with Tamoxifen
Invasive Breast Cancer	0.51	246.6	125.8
Hip Fracture	0.55	101.6	55.9
Endometrial Cancer	4.01	81.4	326.4
Stroke	1.59	110	174.9
Pulmonary Embolism	3.01	50	150.5
Total		589.6	833.5

\* Data from [26].

**Table VI**

Fraction of maximum possible lives saved that is achieved by allocating mammograms according to estimated breast cancer risk when there is only enough money to give mammograms to half the population.\*

Risk assessment model	AUC	Fraction of lives saved	Percent improvement
None (random allocation)		0.500	
BCRAT	0.607	0.632	26.4%
BCRAT+7SNPs	0.632	0.667	33.4%
Perfect model	1.000	1.000	100.0%

\* Abbreviations: BCRAT=Breast Cancer Risk Assessment Tool; BCRAT+7SNPs = model that includes BCRAT and 7 single nucleotide polymorphisms that are associated with breast cancer; AUC= area under curves such as those in Figure 1. Fraction of lives saved is the ratio of lives saved with the risk assessment strategy to the lives that could have been saved if the entire population received mammography. Percent improvement is 100 times (fraction of lives saved minus 0.5)/0.5. In these calculations, the cost ratio of risk assessment to mammograms is assumed to be  $k=0.02$ . Data from [50].