Absolute Risk Models for Subtypes of Breast Cancer
Mitchell H. Gail, William F. Anderson, Montserrat Garcia-Closas, Mark E. Sherman

Statistical models developed to predict the absolute risks of breast cancer subtypes may help identify women who could benefit from specific preventive interventions and improve estimates of total breast cancer risk. Chlebowski et al. (1) present data from the observational study and clinical trial cohorts of the Women’s Health Initiative (WHI) to evaluate and improve absolute risk prediction models for estrogen receptor (ER)-positive and -negative invasive breast cancer among women aged 50–79 years. They evaluate how well the Gail model [model 2 in Costantino et al. (2)] predicts the numbers of breast cancers in the combined observational study and clinical trial cohorts (i.e., calibration) and the Gail model’s discriminatory accuracy, expressed as the area under the receiver operating curve (AUC), in the clinical trial population. They examine whether models with many risk factors can improve discriminatory accuracy and propose parsimonious models. Although they found that the AUC for the Gail model for all invasive cancer was 0.60, which is in line with previous validation studies (2,3), they reported AUC values of 0.58 for ER-positive breast cancer but only 0.50 for ER-negative breast cancer. The similarity in AUC values for ER-positive disease and for all invasive breast cancers may reflect the fact that most invasive breast cancer in women aged 50–79 years is ER positive (see Fig. 1).

Affiliations of authors: 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 Plaza South, Room 8032, Bethesda, MD 20892-7244 [e-mail: gailm@mail.nih.gov].
DOI: 10.1093/jnci/djm228
Published by Oxford University Press 2007.
The epidemiologic literature supports several findings of Chlebowski et al. (1). Age-specific incidence rates for ER-positive invasive breast cancer increase continuously with age, although at a slower rate after age 50 years (Fig. 1), whereas the rates for ER-negative breast cancer flatten or fall after age 50 years (4–6). Thus, the finding that age is not associated with ER-negative invasive breast cancer in the combined observational study and clinical trial populations is consistent with population incidence rates.

Previous studies (7,8) indicated that the protective effects of reproductive risk factors (such as delayed age at menarche, increased parity, early age at birth of first child, and early menopause) tend to be stronger for ER-positive than for ER-negative breast cancer. Chlebowski et al. (1) found stronger protective effects for ER-positive than for ER-negative breast cancer for parity and early menopause, but they noted that protective effects from early age at birth of first child and delayed age at menarche were similar for both.

Chlebowski et al. (1) found an odds ratio of only 1.12 for one or more affected first-degree relatives for ER-negative disease and of 1.44 for ER-positive breast cancer. This finding contrasts with those from 12 studies [supplement to Althuis et al. (7)] showing that larger odds ratios were related to family history for ER-negative disease than found by Chlebowski et al. (1); these odds ratios exceeded the corresponding odds ratios for ER-positive disease in six of the 12 comparisons. For ER-negative breast cancer, these odds ratios ranged from 1.2 to 5.7, with a median of 1.7, a value also near that in Colditz et al. (9). Because family history is the strongest determinant of risk in the Gail model, the unusually low odds ratio for family history in WHI data could explain, at least in part, the poor discriminatory accuracy found for the Gail model in ER-negative disease in WHI. Applying the Gail model to data from a population-based study in Poland (10), we found that the AUC was slightly higher in ER-positive than in ER-negative women with ages comparable to those of the WHI population (Gail MH, Anderson WF, Garcia-Closas M, Sherman ME, Richesson D: unpublished data), but the disparity was not nearly as great as reported by Chlebowski et al. (1) Thus, further assessments of the discriminatory accuracy of the Gail model are warranted for ER-negative disease.

Some analysts advocate adding numerous weak risk factors to improve AUC. Chlebowski et al. (1) found that including all the factors in their table 2 increased the AUC by 0.03 or less to 0.61, 0.62, and 0.53, respectively, for all, ER-positive, and ER-negative breast cancer. Chen et al. (11) found that adding a single strong risk factor, mammographic density, increased the average age-specific AUC by 0.05.

Chlebowski et al. (1) found that the Gail model underestimated the observed numbers of invasive breast cancers in WHI [table 3 in Chlebowski et al. (1)]. The observed (O) to expected (E) ratio was 

\[
\frac{O}{E} = \frac{3236}{2562} = 1.26 \quad \text{(95% confidence interval [CI] = 1.22 to 1.31)}
\]

For comparison, the O/E ratio in women aged 50 years or older was 

\[
\frac{843}{788.28} = 1.07 \quad \text{(95% CI = 1.00 to 1.15)}
\]

in data from 1992 to 1997 from the Nurses Health Study (3). From National Cancer Institute’s Surveillance, Epidemiology, and End Results Program [SEER 13, 1998–2002 (12)], except Alaska, we calculated a standardized incidence ratio of 1.06 (95% CI = 1.02 to 1.10) for WHI, which was based on the age and race/ethnicity distributions from the observational study and clinical trial cohorts. Thus, the WHI population has slightly higher rates of invasive breast cancer than SEER. Chlebowski et al. (1) comment that “more comprehensive mammography use” and “common use of percutaneous core biopsy specimens” may have “contributed to a more comprehensive ascertainment,” which could explain why the Gail model underestimated risk in women aged 50 years or older in the WHI.

Good statistical absolute risk models for breast cancer subtypes may be useful for improving estimates of total invasive breast cancer risk. If a preventive intervention only affects a certain subtype of breast cancer, then a risk model for that subtype could improve the assessment of benefits from intervention. Chlebowski et al. (1) proposed such models for ER-positive breast cancer. Independent studies are needed to see if these models predict observed ER-positive breast cancer incidence well. Although Chlebowski et al. (1) seem to emphasize the “Gail model threshold of 1.7%” in discussing these models, no single threshold should be used to decide whether a woman should take a preventive intervention like tamoxifen. For example, a 40-year-old woman with a 5-year risk of 2.0% might be advised to take tamoxifen because she has very small risks of adverse effects like stroke, pulmonary emboli, or endometrial cancer, whereas a 60-year-old woman with the same 2.0% risk would
probably not be advised to take tamoxifen because the risks from adverse events outweigh the benefits [tables 10 and 11 in Gail et al. (13)].

Chlebowski et al. (1) have presented useful and important results that illustrate the promise and difficulty of estimating absolute risk in subtypes of breast cancer. Additional studies of this type [see also Colditz et al. (9)] are needed in women aged 50 years or older and in younger women. The WHI data can also be valuable for developing or validating absolute risk models for other health outcomes, such as stroke and pulmonary emboli, that are needed to weigh risks and benefits.

References