ARTICLE

Estrogen Receptor Status and the Future Burden of Invasive and In Situ Breast Cancers in the United States

Philip S. Rosenberg, Kimberly A. Barker, William F. Anderson

Affiliations of authors: Biostatistics Branch, Division of Cancer Epidemiology and Genetics, National Cancer Institute, National Institutes of Health, Department of Health and Human Services, Bethesda, MD (PSR, KAB, WFA).

Correspondence to: Philip S. Rosenberg, PhD, Biostatistics Branch, Division of Cancer Epidemiology and Genetics, National Cancer Institute, 9609 Medical Center Drive, Room 7-E-130 MSC 9780, Bethesda, MD 20892 (e-mail: rosenbep@mail.nih.gov).

Abstract

Background: No study has predicted the future incidence rate and annual burden (number) of new cases in the United States of invasive and in situ female breast cancers stratified by the estrogen receptor (ER) status.

Methods: We constructed forecasts for women age 30 to 84 years in 2011 through 2030 using cancer incidence data from the Surveillance, Epidemiology, and End Results Program, novel age-period-cohort forecasting models, and population projections from the US Census Bureau.

Results: The total number of new tumors (invasive plus in situ) is expected to rise from 283,000 in 2011 to 441,000 in 2030 (plausible range 353,500 to 466,700 cases). The proportion of all new case patients age 70 to 84 years is expected to increase from 24.3% to 34.8%, while the proportion ages 50 to 69 years is expected to decrease from 54.7% to 43.6%. The proportion of ER-positive invasive cancers is expected to remain nearly the same at 62.6%, whereas the proportion of ER-positive in situ cancers is expected to increase from 19.1% to 28.9%. The proportion of ER-negative cancers (invasive and in situ) is expected to decrease from 16.8% to 8.6%.

Conclusions: Breast cancer overall will rise in the United States through 2030, especially for ER-positive in situ tumors among women age 70 to 84 years. In contrast, ER-negative invasive and in situ tumors will fall, for reasons that are not fully understood. These results highlight a need to optimize case management among older women, characterize the natural history of in situ cancers, and identify those factors responsible for declining ER-negative incidence.

During the next several decades, 40 million American women who were born between 1946 and 1964 (baby boomers) will face high absolute risks of postmenopausal breast cancer, currently estimated to average around 2% to 4% over 10 year periods (1). An additional 56 million women, most of them currently in their 20s and 30s, will be at substantial risk of premenopausal breast cancer, around 0.4% to 1.5% over 10-year periods. The likely future burden of breast cancer in the United States, as measured by the absolute annual numbers of new case patients, has not been thoroughly explored (2). However, quantitative forecasts could help the oncology community develop a proactive roadmap to optimize prevention and treatment strategies.

A credible forecasting methodology should reflect four key aspects of breast cancer natural history and epidemiology. First, breast cancer is biologically (3,4) and etiologically (5–9) heterogeneous within subtypes defined by the estrogen receptor and/or epithelial cell of origin. Second, organized screening mammography is in place and will not likely decrease in the near future (10). Third, in situ breast cancers are being detected more frequently because of screening (11–13). Fourth, there is considerable intergenerational heterogeneity (birth cohort effects) in the rate at which breast cancers occur (14,15).

In this study, we forecast the number of new invasive and in situ breast cancer cases in the United States through 2030 using...
nationally representative cancer surveillance data, population projections produced by the Census Bureau, and mathematical age-period-cohort (APC) models for invasive and in situ estrogen receptor-positive (ER+) and –negative (ER-) tumors in premenopausal and postmenopausal women. Our analysis addresses key questions about the future: For any given age group, will there be more breast cancer cases in the United States than now, about the same number, or less? If increases are expected, will the number of new cases rise faster than population growth, or remain near par? Finally, will the demographic and/or biological spectrum of the disease be the same in the future as it is now, or change?

Methods

Breast Cancer Case and Population Data

We obtained single-year breast cancer case and population data for calendar years 1992 through 2011 and ages 30 to 84 years at diagnosis from the Surveillance, Epidemiology, and End Results (SEER) 13 Registries Database (November 2013 Submission) using SEER*Stat Version 8.1.5 (www.seer.cancer.gov/seerstat). SEER 13 covers approximately 13.4% of the US population and includes: Atlanta, Connecticut, Detroit, Hawaii, Iowa, New Mexico, San Francisco-Oakland, Seattle-Puget Sound, Utah, Los Angeles, San Jose-Monterey, Rural Georgia, and the Alaska Native Tumor Registry. Tumors were classified by ICD-O-3 Site and Morphology Behavior codes of “malignant” or “in situ,” and an Extent of Disease ER status of “positive,” “negative,” or “unknown.” Incidence rates (malignant or invasive and in situ) were corrected for missing ER data (16) using a validated methodology (17). We excluded 2011 data from the APC models to account for any residual delays in case reporting during the most recent available year (18) and to minimize the impact of changes in ER assay cutpoints following the 2010 guideline recommendations from the American Society of Clinical Oncology/College of American Pathologists for immunohistochemical testing of the ER (19).

In December 2012, the US Census Bureau released updated population projections of the entire US resident population by age for both sexes for 2012 through 2060 (www.census.gov/population/projections/data/national/2012). These projections are based on 2011 population estimates and assumptions about future births, deaths, and net international migration. For consistency, we assumed that population counts for 2011 were the same as for the modeled counts for 2012.

The burden or absolute number of new cancer cases equals the cancer incidence rate (cases per 100 000 person-years) times the total number of 100 000 person-years at risk. We estimated the future burden of new breast cancer cases by single year of age by multiplying APC incidence rate forecasts obtained from SEER 13 (described below) by the corresponding Census Bureau’s female population projections for the entire United States.

APC Forecasting Models and Statistical Analyses

For this study we developed a refined version of our previous APC forecasting model (16,20). Technical details are provided in the Supplementary Materials (available online). In brief, the models calculate expected future birth cohort–specific incidence rates by age, by multiplying estimates of the longitudinal age incidence in a reference cohort by the rate ratio between specific cohorts and the reference cohort. In each future year, a new birth cohort must be accounted for in the projections. To do so, our previous model fitted a single regression line to the logarithm of the observed cohort rate ratios vs cohort, and extended the log-linear fit. Our refined model fits a JoinPoint piecewise log-linear regression model (21). JoinPoint fits one or more connected line segments to the data, thereby accounting for any changes in the slope of the curve. If the trend in younger vs older cohorts is statistically significantly different, our new procedure extrapolates from the younger cohorts. If only one segment is needed for JoinPoint regression, our new adaptive procedure is equivalent to our previous procedure.

Using this approach, we modeled 16 subsets of breast tumors defined by positive vs negative ER status (a correlate of tumor biology), invasive vs in situ behavior, and four age group (two premenopausal groups age 30-39 and 40-49 years, and two postmenopausal groups age 50-69 and 70-84 years). Screening mammography is recommended for women ages 50 to 74 years by the US Preventive Services Task Force (22) and starting at age 40 years by the American Cancer Society (23); the age group-specific models allowed the APC model parameters to implicitly reflect screening effects in the population. To help characterize the observed incidence, we age-standardized the rates per 100 000 person-years (ASRs) within age groups using the 2000 US Standard Population, and we then used JoinPoint analysis to highlight observed trends in the summary ASRs. To help summarize the projected incidence and burden, we also calculated summary estimates of future annual percentage changes (EAPCs). See the Supplementary Materials (available online) for details.

The forecasting models can include or exclude data from earlier calendar years. For purposes of projecting into the future, how far backwards to include presents a bias-variance tradeoff. In ancillary analysis, we observed that 80% to 90% of in situ tumors were missing ER status from 1992 to 2002. The proportion of missing data declined to 68% in 2003 and then fell from 47% in 2004 to 20% in 2010. Therefore, for in situ cancers, we excluded the earlier years and carried out APC modeling using corrected ER+ and ER- incidence rates for 2004 to 2010.

For each APC model, we assessed goodness of fit based on the square root of the usual overdispersion parameter $\sigma$. Values near 1.0 indicate good agreement between the observed and fitted rates (24). We also constructed heat maps of residuals by age vs period and cohort, to screen for systematic lack-of-fit, and we examined how well the predicted number of new cases in 2011 agreed with the observed number based on SEER 13 rates for 2011.

As a sensitivity analysis, we developed a complementary forecasting model based on the cross-sectional (rather than the longitudinal) age incidence curve and the period (rather than the cohort) rate ratio curve (25). See the Supplementary Materials (available online) for details. If JoinPoint analysis of the period rate ratio curve determines that the observed trend in more recent vs past calendar periods is statistically significantly different, the alternative model extrapolates from the recent periods (rather than the younger cohorts). For both models, 95% confidence intervals (CIs) incorporated uncertainty about APC model parameters and JoinPoint cohorts. All calculations were carried out using MATLAB version 14 (26).

Downloaded from https://academic.oup.com/jnci/article-abstract/107/9/djv159/897479 by guest on 20 November 2018
Results

Of 505,727 cancers included in our analysis, 357,226 were invasive ER+ and 100,614 were invasive ER- tumors diagnosed between 1992 and 2010 with $2.11 \times 10^8$ woman-years of follow-up in SEER 13. An additional 40,342 were in situ ER+ and 7,345 were in situ ER- tumors diagnosed between 2004 and 2010 with $8.28 \times 10^7$ woman-years of follow-up. APC models were successfully fitted to the observed data for all 16 subsets of breast tumors, i.e., there was negligible overdispersion and no evidence of lack of fit (Supplementary Figure 1, available online). APC age effects are shown in Supplementary Figure 2 (available online), cohort effects in Supplementary Figure 3 (available online), and period effects in Supplementary Figure 4 (available online).

The incidence of invasive ER+ tumors (rate per 100,000 woman-years) between 1992 and 2010 had distinct patterns by age (Figure 1A, left-hand side). As previously noted (13,27), incidence peaked circa 2000 among postmenopausal women in age groups older than or equal to 50 years, but there was no pronounced peak among premenopausal women. JoinPoint analysis showed that incidence has been increasing by around 1% to 2% per year in recent years in all age groups (Supplementary Table 1, available online). The projected incidence (Figure 1A, right-hand side) is also expected to increase substantially between 2011 and 2030 (Supplementary Table 2, available online).

The absolute number of new ER+ invasive tumors per year is also expected to increase between 2011 and 2030 (Figure 1B; Supplementary Table 3, available online). Importantly, as the baby boomer cohort ages, new cases among women age 70 to 84 years is expected to increase by 4.0% per year, from 47,800 cases in 2011 to 95,300 cases in 2030. The number of new cases among women age 50 to 69 years is expected to increase by 1.6% per year, from 98,000 cases in 2011 to 125,700 cases in 2030.

In each age group, the incidence of in situ ER+ tumors observed between 2004 and 2010 (Figure 1C, left-hand side) increased by around 1.5% to 2.5%/year (Supplementary Table 1, available online). The projected incidence (Figure 1C, right-hand side) is expected to increase substantially between 2011 and 2030 in each age group except 30- to 39-year-olds (Supplementary Table 2, available online). New cases of in situ ER+ tumors are also expected to increase between 2011 and 2030 among women age 40+ years (Figure 1D; Supplementary Table 3, available online). Among women age 70 to 84, new cases are expected to increase by 7.9% per year, from 10,700 cases in 2011 to 43,200 cases in 2030. In all ages combined, the number is expected to increase from 53,900 to 127,400 cases.

As previously reported (16), the incidence of invasive ER- tumors (Figure 2A, left-hand side) has been decreasing from 1992 to 2010 in each age group (Supplementary Table 1, available online). The projected incidence between 2011 and 2030 (Figure 2A, right-hand side) is also expected to decrease in each age group (Supplementary Table 2, available online). Interestingly, among women age 70 to 84 years, the incidence of invasive ER- tumors is expected to decrease by around 0.7% per year (Supplementary Table 2, available online), but because of population growth in this age group the corresponding burden (Figure 2B) is expected to increase by around 2% per year (Supplementary Table 3, available online). In all age groups combined, new cases of invasive ER- tumors are expected to decrease between 2011 and 2030, from 40,000 cases in 2011 to 34,000 cases in 2030.

Figure 1. Observed and projected incidence of invasive and in situ estrogen receptor (ER)-positive breast tumors in Surveillance, Epidemiology, and End Results (SEER) 13 and corresponding forecasts of cancer burden in the entire United States. A) Observed and projected incidence of invasive ER+ tumors per 100,000 woman-years in SEER 13. B) Predicted burden of invasive ER+ tumors in the United States (number of newly diagnosed cases per year) by age group and overall. C) Observed and projected incidence of in situ ER+ tumors per 100,000 woman-years in SEER 13. D) Predicted burden of in situ ER+ tumors in the entire United States. In each panel, circles show point estimates for each year. Shaded bands show point-wise 95% confidence limits. Vertical reference line separates observed from forecast period. ER = estrogen receptor.
The incidence of in situ ER- tumors (Figure 2C, left-hand side) between 2004 and 2010 decreased statistically significantly among women age 30 to 69 years, and was stable among women age 70 to 84 years (Supplementary Table 1, available online). The projected incidence between 2011 and 2030 (Figure 2C, right-hand side) is also expected to decrease (Supplementary Table 2, available online). In all age groups combined, new cases of in situ ER- tumors are expected to decrease from 7700 in 2011 to 3800 in 2030 (Figure 2D).

Frequency plots show the projected numbers of new cases by single-year of age at diagnosis in 2011 and 2030 (Figure 3). Among women age 70 to 84 years, new cases of invasive ER+ tumors are expected to increase by 100% (Figure 3A), while new cases of in situ ER+ tumors could be as much as 300% higher (Figure 3B), although the confidence limits are broad. By way of comparison, the population projection for women age 70 to 84 years increases by 85%, from 13 million to 24 million. Hence, it is plausible to expect that new cases of in situ tumors among women age 70 to 84 years may increase considerably faster than the population. In contrast, among women age 50 to 69 years, new cases of invasive ER- tumors are expected to decline by 43% (Figure 3C) and by 73% for in situ ER- tumors (Figure 3D). These decreases greatly outpace the corresponding female population projections, which increase by 5%, from 39 million to 41 million.

Figure 4 presents a summary of cases in 2011 vs 2030 by age group (Figure 4, A and B) and tumor type (Figure 4, C and D) via scaled pie charts in which the area of each pie is proportional to the number of cases (Supplementary Figure 5, available online, presents these data using bar charts). The total number of breast cancers is expected to increase from 283 000 to 441 000. The proportion of cases age 70 to 84 years is expected to increase from 24.3% to 34.8%, while the proportion of cases age 50 to 69 years is expected to decrease from 54.7% to 43.6%. The proportion of ER+ invasive cancers is expected to remain about the same, from 64.1% to 62.6%. The proportion of ER+ in situ cancers is expected to increase from 19.1% to 28.9%, and the combined proportion of ER- invasive and in situ tumors is expected to decrease from 16.8% to 8.6%.

In our sensitivity analyses, the alternative and base models produced broadly similar forecasts; for 2011, the observed numbers of new cases that were not included in the APC models (ER+ and ER- invasive and in situ cancers by single-year of age) fell within the confidence bands of both forecasts (Supplementary Figures 6 and 7, available online). The alternative models also revealed two key uncertainties: ER- invasive incidence decreased more rapidly, and ER+ in situ incidence increased more slowly among women age 70 to 84 years. The alternative model yielded a 2030 total of 390 600 cases (95% CI = 353 500 to 427 800) vs 441 400 for the base model (95% CI = 416 100 to 466 700). A plausible range for the 2030 total combining both models is 353 500 to 466 700 cases, a 25% to 65% increase in cases overall compared with 2011.

Discussion
Screening mammography has been well-accepted in the United States during our study period (10), despite lingering concerns about the potential of overdiagnosis (28). Therefore, it seems unlikely that screening will become less prevalent in the future. This motivated us to forecast the entire spectrum of breast neoplasms, including in situ tumors that are almost entirely detected via screening mammography (29), especially ER+ tumors (30). We also anticipate that screening will become more sensitive in the future (31–33). Hence, our current projections for in situ lesions could be conservative.

Our base forecast is that the total number of new invasive and in situ breast cancers will increase from 283 000 cases in 2011 to 441 000 cases in 2030. Combining our base and alternative
forecasting models, a plausible range for the total number of new breast cancers in 2030 is 353,500 to 466,700 cases. The prior estimate by Smith et al. (2), based on more limited SEER data, falls just above our lower limit.

Notably, however, our results suggest that different subgroups of breast cancers are moving in different directions and at different trajectories. For example, in 2011, almost four of every five breast cancers (79.0%) occurred among postmenopausal women. We expect this will remain unchanged in 2030. Even so, the age distribution is expected to skew towards older postmenopausal women. Between 2011 and 2030, the proportion of cases age 50 to 69 years is expected...
to decrease, from 54.7% to 43.6%, whereas the proportion of cases age 70 to 84 years is expected to increase, from 24.3% to 34.8%.

In 2030, the proportion of invasive ER+ tumors is forecast to remain about the same as in 2011, near 63%. However, the proportion of in situ ER+ tumors will increase, from around one of every five tumors circa 2011 (19.1%) to almost one of every three tumors circa 2030 (28.9%). Hence, both the proportion as well as the absolute number of in situ ER+ tumors is forecast to be substantially greater in 2030 than 2011, especially among women age 70 to 84 years.

In contrast, both invasive and in situ ER- tumors are expected to decline. Declining ER- invasive cancers have previously been reported among women enrolled in the Kaiser Permanente Northwest health plan (34) and in the entire United States (16,35) and validated in Denmark (36), where this pattern was shown to be related to a birth-cohort effect (exposure) rather than calendar-period effect (changing ER assays, case ascertainment, etc.).

Our models quantify the future numbers of new breast cancers in realistic scenarios taking into account many of the known correlates of breast cancer incidence. Although the disease could evolve quite differently than predicted by our models, our estimates provide our best assessment of expected future challenges and opportunities in light of current data. Therefore, our methods can and should also be applied to obtain burden forecasts specific to each race and ethnic group, perhaps incorporating additional statistical methods to compensate for the smaller numbers of cases (37). Importantly, we do not believe that race/ethnic differences have confounded our estimates for all women combined. Indeed, we previously showed that the overall projected incidence trends for invasive ER+ and ER- tumors are similar in non-Hispanic whites, Hispanics, blacks, and Asian or Pacific Islanders (16). Hence, it seems unlikely that there are strong race/ethnic biases affecting our analyses of 16 subgroups of in situ tumors. Also worth considering in future studies is how increases or decreases in obesity and other established risk factors might impact the projections.

Our study has three key limitations. First, we obtained national forecasts by extrapolating from the SEER 13, but the 13 registries may not be completely representative of the whole United States. Second, our long-term forecasts may be sensitive to the accuracy of reported incidence rates in SEER 13 among younger women in recent years. Third, our forecasts of total burden reflect the aging of the large baby boomer cohort, increasing life expectancy, as well as increasing rates of ER+ tumors, but it remains unclear how much each factor is contributing to the expected increase.

Nonetheless, we expect that the total number of breast cancer cases will be unacceptably higher in 2030 than in 2011. Clearly, managing this clinical burden will present a huge challenge. Between now and then our results highlight three key opportunities. First, there is an urgent need to optimize the management of breast cancers among the increasingly larger population of women age 70 to 84 years, who tend to have less favorable outcomes than younger women and who historically have been underrepresented in clinical trials (38). Second, we need to better understand the natural histories of in situ lesions, which might identify screening and treatment protocols. Third, we need to further elucidate the exposure (etiologic) factors responsible for the declining rate of ER-negative tumors in order to develop prevention clues for the most difficult-to-treat breast cancers (9,39).

**Funding**

This research was supported entirely by the Intramural Research Program of the National Institutes of Health, National Cancer Institute (NCI), Division of Cancer Epidemiology and Genetics.

**Notes**

The study sponsor had no role in the design of the study, the collection, analysis, or interpretation of the data, the writing of the manuscript, nor the decision to submit the manuscript for publication.

Author Contributions: PSR developed the statistical methodology and software and designed the study. KAB and WFA assembled SEER 13 data. KAB and PSR assembled population projections. All authors analyzed results and wrote the paper.

The authors declare no competing financial interests.

**References**