Title

Advanced Breast Cancer Definitions by Staging System Examined in the Breast Cancer Surveillance Consortium.

Journal

Journal of the National Cancer Institute, 113(7)

ISSN

0027-8874

Authors

Kerlikowske, Karla
Bissell, Michael CS
Sprague, Brian L
et al.

Publication Date

2021-07-01

DOI

10.1093/jnci/djaa176

Peer reviewed

OXFORD

JNCI J Natl Cancer Inst (2021) 113(7): djaa176

doi: 10.1093/jnci/djaa176
First published online November 10, 2020
Article

Advanced Breast Cancer Definitions by Staging System Examined in the Breast Cancer Surveillance Consortium

Karla Kerlikowske, MD,[1,2,*] Michael C. S. Bissell, PhD,[3] Brian L. Sprague, PhD,[4] Diana S. M. Buist, PhD,[5] Louise M. Henderson, PhD,[6] Janie M. Lee, MD,[7] Diana L. Miglioretti, PhD[3,5,8]

[1]Department of Medicine and Epidemiology and Biostatistics, University of California, San Francisco, CA, USA; [2]General Internal Medicine Section, Department of Veterans Affairs, University of California, San Francisco, CA, USA; [3]Department of Public Health Sciences, University of California, Davis, CA, USA; [4]Department of Surgery and Radiology, University of Vermont, Burlington, VT, USA; [5]Kaiser Permanente Washington Health Research Institute, Kaiser Permanente Washington, Seattle, WA, USA; [6]Department of Radiology, University of North Carolina, Chapel Hill, NC, USA; [7]Department of Radiology, University of Washington, Seattle, WA, USA and [8]University of California Davis Comprehensive Cancer Center, Sacramento, CA, USA

*Correspondence to: Karla Kerlikowske, MD, San Francisco Veterans Affairs Medical Center, General Internal Medicine Section, 111A1, 4150 Clement St., San Francisco, CA 94121, USA (e-mail: Karla.Kerlikowske@ucsf.edu).

Abstract

Background: Advanced breast cancer is an outcome used to evaluate screening effectiveness. The advanced cancer definition resulting in the best discrimination of breast cancer death has not been studied in a breast imaging population. **Methods:** A total of 52 496 women aged 40-79 years participating in the Breast Cancer Surveillance Consortium diagnosed with invasive cancer were staged using the 8th edition of American Joint Committee on Cancer (AJCC) anatomic and prognostic pathologic systems and Tomosynthesis Mammographic Imaging Screening Trial (TMIST) tumor categories. We calculated the area under the receiver operating characteristic curve for predicting 5-year breast cancer death and the sensitivity and specificity for predicting 5-year breast cancer death for 3 advanced cancer classifications: anatomic stage IIB or higher, prognostic pathologic stage IIA or higher, and TMIST advanced cancer. **Results:** The area under the receiver operating characteristic curves for predicting 5-year breast cancer death for AJCC anatomic stage, AJCC prognostic pathologic stage, and TMIST tumor categories were 0.826 (95% confidence interval [CI] = 0.817 to 0.835), 0.856 (95% CI = 0.846 to 0.866), and 0.789 (95% CI = 0.780 to 0.797), respectively. AJCC prognostic pathologic stage had statistically significantly better discrimination than AJCC anatomic stage (difference = 0.030, bootstrap 95% CI = 0.024 to 0.037) and TMIST tumor categories (difference = 0.067, bootstrap 95% CI = 0.059 to 0.075). The sensitivity and specificity for predicting 5-year breast cancer death for AJCC anatomic stage IIB or higher, AJCC prognostic pathologic stage IIA or higher, and TMIST advanced cancer were 72.6%, 76.7%, and 96.1%; and 78.9%, 81.6%, and 41.1%, respectively. **Conclusions:** Defining advanced cancer as AJCC prognostic pathologic stage IIA or higher most accurately predicts breast cancer death. Use of this definition by investigators will facilitate comparing breast cancer screening effectiveness studies.

ARTICLE

Women diagnosed with advanced breast cancer have worse survival than women diagnosed with early-stage disease (1,2). The goal of breast cancer screening is to reduce the number of women diagnosed with advanced breast cancer to decrease breast cancer mortality (3). Detection of early-stage breast cancer by screening has been used as an intermediate outcome of screening effectiveness, but this measure is confounded by inclusion of tumors not likely to affect women's overall survival if left undiagnosed (4).

Advanced breast cancer is recognized as an important endpoint for evaluating breast cancer screening effectiveness (1,5-8) and has been used to evaluate differences in screening effectiveness by breast cancer risk, screening interval, and modality (1,5,6,9). The American Joint Committee on Cancer (AJCC) anatomic staging system has been the standard to characterize breast cancer at diagnosis using tumor size (T), lymph node status (N), and presence of metastatic disease (M), with advanced breast cancer often defined as stage IIB or higher (1,5-7). There is limited research on whether this staging system or threshold best discriminates groups of women with worse survival (2).

In 2016, the AJCC's eighth edition prognostic pathologic staging system was released and included anatomic staging

Received: June 1, 2020; **Revised:** August 10, 2020; **Accepted:** October 28, 2020

elements plus tumor grade and estrogen receptor (ER), progesterone receptor (PR), and HER2 status. The addition of these prognostic factors appears to better discriminate breast cancer survival than the anatomic staging system (2). The Tomosynthesis Mammographic Imaging Screening Trial (TMIST) is actively recruiting and randomly assigning women to undergo 2-dimensional digital mammography or 3-dimensional digital breast tomosynthesis, which uses advance breast cancer as an intermediate outcome to compare effectiveness. The TMIST protocol defines advanced breast cancer according to combinations of tumor size; ER, PR, and HER2 status; and tumor spread (8). One study (2) evaluated survival outcomes for AJCC eighth edition anatomic vs pathologic prognostic staging in a single-institution cohort and state cancer registry, but none to our knowledge has evaluated these 3 staging systems and various advanced cancer definitions either in the same population or in a breast imaging population.

This study compares the AJCC eighth edition anatomic and prognostic pathologic stages and TMIST tumor categories in the Breast Cancer Surveillance Consortium (BCSC) cohort of women undergoing breast imaging. We sought to determine which staging systems and advanced breast cancer definitions result in the best discrimination of breast cancer death for evaluating screening effectiveness.

Methods

Study Setting and Data Sources

We used data from the BCSC's 6 mammography registries (https://www.bcsc-research.org), whose populations are comparable with the US population (10-12). We included prospectively collected data, including women's characteristics and radiology information from community and academic radiology facilities. Breast cancer diagnoses and tumor characteristics were obtained by linking women to pathology databases; regional Surveillance, Epidemiology, and End Results programs; and state tumor registries. Vital status was obtained from Surveillance, Epidemiology, and End Results programs; state tumor registries; and state death tapes. Registries and a central statistical coordinating center received institutional review board approval for active or passive consenting processes or a waiver of consent to enroll participants, link data, and perform analyses. All procedures were Health Insurance Portability and Accountability Act compliant, and registries and the coordinating center received a federal certificate of confidentiality and other protections for the identities of women, physicians, and facilities.

Participants

We studied women aged 40-79 years with an incident invasive breast cancer diagnosed from January 2005 to December 2017. Women with a personal history of ductal carcinoma in situ were excluded.

Measures, Definitions, Outcomes

Demographic and breast health history information were obtained on self-administered questionnaires completed at each mammogram. A total of 82.0% of women had undergone a breast imaging examination at a BCSC facility before their diagnosis, 15.3% after their diagnosis, and 2.7% had no breast imaging record in the BCSC.

Invasive breast cancers were classified according to the eighth edition AJCC anatomic staging system based on tumor size (T), the presence of lymph node involvement (N), and the presence or absence of distant metastasis (M) using pathologic values first and clinical values only when pathologic values were not available (13). If the eighth edition T, N, or M pathological or clinical codes were missing, prior editions were used. AJCC eighth edition prognostic pathologic stage was defined according to TNM stage; plus tumor grade; and ER, PR, and HER2 status (2). For anatomic and prognostic pathologic stages, we examined 2 thresholds each to define advanced breast cancer: stages IIA or higher and stages IIB or higher.

For TMIST analyses, we classified each tumor according to the TMIST definition of advanced breast cancer selecting the category with the worst survival with categories from worst to best survival as follows: 1) cancers that spread from the breast to a distant organ in the body, 2) cancers that spread from the breast to at least 1 nearby lymph node, 3) tumors of at least 20 mm, 4) tumors greater than 10 mm and triple negative, 5) tumors greater than 10 mm and HER2 positive, and 6) tumors not classified into any of the 5 categories.

We also examined staging systems by breast cancer mode of detection to examine for differences in thresholds for evaluating screening effectiveness for the 27 321 women with a mammogram in the BCSC around the time of diagnosis. Screen-detected cancer was defined as an invasive cancer diagnosed within 12 months of a final positive Breast Imaging Reporting and Data System assessment of 3, 4, or 5 on a screening mammogram. Interval invasive cancer was defined as an invasive cancer diagnosed within 12 months of a final negative assessment of Breast Imaging Reporting and Data System 1 or 2 on a screening mammogram or having a diagnostic mammogram around the time of diagnosis with a prior mammogram (before the diagnostic mammogram that detected the cancer) within the past 12-27 months (14). A clinically detected invasive cancer was defined as one where there was no screening mammogram 12 months before the diagnostic mammogram, or no breast imaging examination was more than 27 months before the diagnostic mammogram that detected the cancer.

Statistical Analysis

We used descriptive statistics to characterize the study population according to anatomic stage IIA or lower, IIB or higher, or unknown.

Women were followed up from their first primary invasive breast cancer diagnosis until the earliest of the following: death from breast cancer, death from other causes, end of complete vital status capture, or 10 years. We estimated 5- and 10-year breast cancer survival overall and by mode of detection using the Kaplan–Meier estimator. Five-year survival was considered high if more than 95% as was observed for stage I anatomic or prognostic stage breast cancers (2). We estimated the sensitivity, specificity, and positive predictive value (PPV) using Bayes' rule for each dichotomized classification of advanced breast cancer (IIA or higher, IIB or higher, TMIST advanced cancer) for predicting death within 5 years (see the Supplementary Methods, available online, for additional details). The time-dependent area under the receiver operating characteristic curve (AUC) for each staging classification was computed using each stage for predicting 5-year breast cancer death; 8 categories

ARTICLE

for AJCC staging systems and 6 (nonadvanced plus 5 advanced categories) for TMIST ordered from worst to best survival. Confidence intervals for performance measures and AUCs were calculated using a nonparametric bootstrap (15) with 10 000 iterations with estimated means stabilizing within 4 decimal places in the first 5000 iterations. Differences between AUCs were statistically significantly different if the bootstrap confidence intervals (CIs) excluded zero.

Primary analyses were based on all available cancers with complete component data for each staging system. We excluded cancers from each staging system for which we could not calculate that stage irrespective of the availability of components to calculate stage for other staging systems. As a sensitivity analysis, we followed the methods and recommendations used by Howlader (16) to use multiple imputation chained equations (17) to impute 50 datasets for missing values of the primary staging component variables: T; N; ER, PR, and HER2 status; and tumor grade. AJCC staging variables and TMIST category were then rederived using the observed and imputed components only when missing (see the Supplementary Methods and Supplementary Tables 1 and 2, available online, for imputation model details).

Survival estimates and proportions were calculated in SAS 9.4 (SAS Institute Inc., Cary, NC). Performance measures, AUCs, and bootstrapped confidence intervals were calculated using the "survival" and "survivalROC" packages in R version 3.5.1 (18-20).

Results

Study Population Characteristics

We included 52 496 women aged 40-79 years (median = 60 years) who had an incident invasive cancer of whom 67.4% were non-Hispanic White, 12.4% non-Hispanic Black, 7.3% Asian, 4.4% Hispanic or Latina, and 8.5% mixed or other (Table 1). Women with anatomic stage IIB or higher were more likely to be aged 40-59 years and Black and have interval or clinically detected breast cancers compared with women with anatomic stage IIA or lower (Table 1).

Staging Systems Accuracy to Predict 5-Year Breast Cancer Death

The AUCs for predicting 5-year breast cancer death for anatomic and prognostic pathologic stages were 0.826 (95% CI = 0.817 to 0.835) and 0.856 (95% CI = 0.846 to 0.866), respectively (Table 2). Prognostic pathologic stage had statistically significantly better discrimination than anatomic stage in the main analysis (difference = 0.030, bootstrap 95% CI = 0.024 to 0.037) and sensitivity analysis based on multiple imputation of missing data (difference = 0.019, bootstrap 95% CI = 0.014 to 0.025) (Supplementary Table 3, available online). AUCs were consistently higher for prognostic stage when stratified by age and race (Supplementary Table 4, available online). The AUC for predicting 5-year breast cancer death for the TMIST 6 tumor categories was 0.789 (95% CI = 0.780 to 0.797). The AJCC classification systems had statistically significantly higher AUCs than the TMIST 6 tumor categories (difference for anatomic = 0.037, bootstrap 95% CI = 0.031 to 0.043; difference for pathologic = 0.067, bootstrap 95% CI = 0.059 to 0.075).

The AUCs for predicting 10-year breast cancer death for anatomic and prognostic pathologic stages were 0.784 (95% CI = 0.773 to 0.796) and 0.799 (95% CI = 0.787 to 0.813), respectively (Supplementary Table 5, available online), and for the TMIST 6 tumor categories was 0.749 (95% CI = 0.740 to 0.758). AUCs remained statistically significantly different across staging systems over the majority of the 10-year follow-up period with nonoverlapping bootstrapped confidence intervals (Supplementary Figure 1, available online).

Advanced cancer definitions, that is, anatomic stage IIB or higher and prognostic pathologic stage IIA or higher, influenced the sensitivity (72.6% vs 76.7%), specificity (78.9% vs 81.6%), and PPV (18.1% vs 21.0%) of predicting 5-year breast cancer death, respectively (Table 2). Values were statistically significantly higher for prognostic stage IIA or higher compared with anatomic stage IIB or higher for sensitivity (difference = 4.1%, bootstrap 95% CI = 2.4% to 5.8%), specificity (difference = 2.7%, bootstrap 95% CI = 2.3% to 3.0%), and PPV (difference = 2.9%, bootstrap 95% CI = 2.4% to 3.4%). Specificity for predicting 5-year breast cancer death improved with prognostic pathologic vs anatomic staging systems whether the advanced breast cancer thresholds were set at IIA or higher or IIB or higher, whereas sensitivity was highest for anatomic IIA or higher and lowest for prognostic pathologic IIB or higher. The TMIST tumor categories for advanced breast cancer had the highest sensitivity (96.1%) and the lowest specificity (41.1%) and PPV (9.9%) (Table 3). The proportions of cancers defined as anatomic stage IIB or higher, prognostic pathologic stage IIA or higher, and TMIST advanced cancer were 24.2%, 21.9%, and 61.2%, respectively (Tables 2 and 3).

Breast Cancer Survival by Staging System and Mode of Detection

Five-year survival was similar and greater than 95% for stages IA and IB across AJCC staging systems and declined below 95% for anatomic stage IIB or higher and prognostic pathologic stage IIA or higher (Table 2). The overall 5-year survival for anatomic stage IIB or higher was 81.9% and for prognostic pathologic stage IIA or higher was 79.0%.

Five-year survival was 99.4% for tumors classified as TMIST nonadvanced breast cancer. Five-year survival for tumors classified as TMIST advanced cancer was 90.1% overall and for all TMIST advanced tumor subgroups was at least 94.9%, except those with positive lymph nodes and stage IV breast cancer (Table 3). Tenfold cross-validation of the ordering of the 5-year survival curves for the 6 TMIST categories did not change the AUC (mean = 0.789).

When evaluating staging systems by mode of cancer detection, 5-year survival declined below 95% for anatomic screen-detected stage IIB or higher, interval cancer stage IIA or higher, and clinically detected stage IB or higher; and for prognostic pathologic screen-detected stage IIA or higher, and both interval or clinically detected stage IB or higher (Table 4). For the TMIST tumor categories, 5-year survival declined below 95% for screen-detected lymph node–positive cancers; for interval or clinically detected cancers that were at least 20 mm, greater than 10 mm, and triple negative, or lymph node positive; and for stage IV cancer irrespective of mode of detection (Table 3).

The sensitivity to predict breast cancer death was higher for clinically detected breast cancer than either screen-detected or interval cancer for both anatomic and prognostic pathologic staging systems whether advanced cancer was defined as IIA or higher or as IIB or higher (Table 4). The PPV for predicting 5-year breast cancer death was higher for interval and clinically detected breast cancer than screen-detected breast cancer for

Table 1. Summary of study population by AJCC eighth edition anatomic stage advanced breast cancer definition

Characteristics	AJCC anatomic stage IIA or lower No. (%[a])	AJCC anatomic stage IIB or higher No. (%[a])	AJCC anatomic stage unknown No. (%[a])
Total No.	38 001	12 113	2382
Age at diagnosis, y			
40-49	6444 (17.0)	2744 (22.7)	536 (22.5)
50-59	10 948 (28.8)	3930 (32.4)	692 (29.1)
60-69	12 680 (33.4)	3468 (28.6)	674 (28.3)
70-79	7929 (20.9)	1971 (16.3)	480 (20.2)
Race, ethnicity			
White, non-Hispanic	26 091 (73.7)	7681 (67.9)	1590 (71.1)
Black, non-Hispanic	4258 (12.0)	1875 (16.6)	374 (16.7)
Asian	2838 (8.0)	868 (7.7)	132 (5.9)
Native American	107 (0.3)	32 (0.3)	5 (0.2)
Hispanic or Latina	1554 (4.4)	637 (5.6)	111 (5.0)
Mixed or other	564 (1.6)	211 (1.9)	25 (1.1)
Missing	2589 (6.8)	809 (6.7)	145 (6.1)
Mode of cancer detection			
Screen detected	11 441 (59.5)	1795 (28.2)	944 (54.6)
Interval cancer	5193 (27.0)	2416 (37.9)	544 (31.5)
Clinically detected	2590 (13.5)	2157 (33.9)	241 (13.9)
Missing	18 777 (49.4)	5745 (47.4)	653 (27.4)

[a]All percentages are percent among nonmissing except missing, which is percent of total. AJCC = American Joint Committee on Cancer.

Table 2. Five-year probability of breast cancer survival by AJCC eighth edition staging system and accuracy of predicting death within 5 years of diagnosis[a]

Measures	AJCC anatomic stage (N = 50 114)	AJCC prognostic pathologic stage (N = 48 049)
Stage, proportion % (% survival)		
IA	50.5 (98.8)	65.5 (98.9)
IB	2.8 (96.9)	13.4 (95.3)
IIA	22.6 (95.7)	8.5 (90.5)
IIB	10.6 (92.1)	2.8 (88.1)
IIIA	6.3 (88.1)	3.5 (83.6)
IIIB	1.4 (73.0)	1.4 (79.2)
IIIC	2.3 (77.0)	1.0 (60.7)
IV	3.6 (45.2)	3.9 (45.1)
IIA or higher	46.8 (88.5)	21.9 (79.0)
IIB or higher	24.2 (81.9)	13.3 (71.6)
AUC (95% CI) for all stages	0.826 (0.817 to 0.835)	0.856 (0.846 to 0.866)
Accuracy of predicting breast cancer death for stage IIA or higher, test measure % (95% CI)		
Sensitivity	88.9 (87.5 to 90.2)	76.7 (74.9 to 78.5)
Specificity	56.0 (55.5 to 56.4)	81.6 (81.3 to 82.0)
PPV	11.5 (11.0 to 12.0)	21.0 (20.2 to 21.9)
Accuracy of predicting breast cancer death for stage IIB or higher, test measure % (95% CI)		
Sensitivity	72.6 (70.7 to 74.4)	62.9 (60.9 to 64.9)
Specificity	78.9 (78.6 to 79.3)	89.9 (89.6 to 90.1)
PPV	18.1 (17.4 to 18.9)	28.4 (27.2 to 29.7)

[a]AJCC = American Joint Committee on Cancer; AUC = area under the receiver operating characteristic curve; CI = confidence interval; PPV = positive predictive value.

both anatomic and prognostic pathologic stages and for the TMIST classification of advanced breast cancer (Tables 3 and 4). The AUC for predicting 5-year breast cancer death was highest for interval cancers defined by prognostic pathologic stage (Tables 3 and 4).

Discussion

Advanced breast cancer is an important outcome because it is a surrogate for breast cancer mortality (2) and used to evaluate screening effectiveness (1,5-8). We examined staging systems and thresholds for defining advanced breast cancer and found AJCC prognostic pathologic stage was more accurate than both anatomic stage and the TMIST tumor classification for predicting breast cancer death. Additionally, we found AJCC prognostic pathologic stage IIA or higher as the staging system and threshold that has the best balance of sensitivity and specificity for predicting breast cancer death and thus may be the best choice when evaluating breast cancer screening effectiveness or supplemental imaging strategies.

Table 3. Five-year probability of breast cancer survival by TMIST advanced breast cancer definition and mode of detection and accuracy of predicting 5-year breast cancer death

Measures	All cancers (N = 45 366)	Screen detected (n = 11 902)	Interval cancer (n = 7026)	Clinically detected (n = 4469)
Tumor characteristics, proportion % (% survival)				
Nonadvanced cancer	38.8 (99.4)	53.4 (99.4)	28.1 (99.8)	15.0 (98.2)
Advanced cancer[a]	61.2 (90.1)	46.6 (93.9)	71.9 (88.8)	85.0 (83.0)
Tumor size 10-20 mm and HER2+	2.2 (99.3)	2.4 (99.2)	2.1 (99.1)	1.5 (100.0)
Tumor size 10-20 mm and triple negative	2.0 (96.1)	2.3 (96.8)	2.1 (94.4)	0.9 (92.6)
Tumor size >20 mm	20.8 (94.9)	16.1 (95.1)	24.4 (94.8)	27.4 (92.7)
Lymph node positive	32.6 (91.0)	24.6 (94.4)	39.1 (88.7)	45.8 (85.2)
Stage IV	3.6 (45.0)	1.1 (46.7)	4.2 (45.3)	9.4 (39.3)
AUC (95% CI)[b]	0.789 (0.780 to 0.797)	0.764 (0.740 to 0.788)	0.776 (0.758 to 0.794)	0.749 (0.727 to 0.770)
Accuracy of predicting breast cancer death, test measure % (95% CI)				
Sensitivity	96.1 (95.2 to 97.0)	89.2 (85.4 to 92.7)	99.2 (98.3 to 99.8)	98.1 (96.8 to 99.3)
Specificity	41.1 (40.6 to 41.6)	54.8 (53.9 to 55.7)	30.5 (29.4 to 31.7)	17.3 (16.0 to 18.5)
PPV	9.9 (9.5 to 10.3)	6.1 (5.4 to 6.9)	11.2 (10.2 to 12.2)	17.0 (15.7 to 18.4)

[a]Defined as at least 1 of the following: tumor size 10-20 mm and HER2+; or tumor size 10-20 mm and triple negative; or tumor size greater than 20 mm; or lymph node positive; or stage IV breast cancer. AUC = area under the receiver operating characteristic curve; CI = confidence interval; PPV = positive predictive value; TMIST = Tomosynthesis Mammographic Imaging Screening Trial.
[b]AUCs for TMIST 6 tumor categories overall and for TMIST 6 tumor categories by mode of breast cancer detection.

When selecting a staging system and threshold for defining advanced breast cancer, investigators need to consider how the information will be used. For example, if the goal is to identify women at increased risk of breast cancer death who may benefit from earlier detection of aggressive tumors, having a definition of advanced cancer that maximizes sensitivity for identifying breast cancer death may be most important when evaluating supplemental screening effectiveness to ensure the highest proportion of women are identified who could potentially have the opportunity to reduce the chance of dying from breast cancer. Maximizing specificity also is an important consideration so women who have a low likelihood of breast cancer death will not experience high numbers of false-positive tests through supplemental screening. If the TMIST advanced breast cancer definition were used to study effectiveness of supplemental screening, a large proportion of women would undergo supplemental imaging who are not at high risk of breast cancer death. As another example, when evaluating primary prevention interventions, using AJCC prognostic pathologic stage to maximize specificity may be prioritized to minimize recommending medications with high risks of side effects to women who have a low likelihood of breast cancer death.

We studied a large, diverse population-based cohort of women diagnosed with an incident breast cancer. We were able to evaluate mode of breast cancer detection, an important predictor of mortality, because information was available on breast imaging around the time of diagnosis for more than one-half of the cohort. We present results by mode of detection to show how definitions of advanced breast cancer may differ depending on the composition of the study population. For example, the sensitivity for predicting 5-year breast cancer death was higher for interval and clinically detected breast cancer than screen-detected breast cancer, whereas specificity for predicting 5-year breast cancer death was highest for screen-detected breast cancer. Notably, our results should not be used to compare survival by mode of detection, because cancer case survival is influenced by lead time, length time, and overdiagnosis biases in these analyses (21). However, within each mode of detection category, a threshold for predicting breast cancer death could be selected to evaluate screening effectiveness strategies for that mode of detection.

Our study cohort is women who were diagnosed with breast cancer within the BCSC, which includes women in 6 states with a similar distribution of ages and races and ethnicities as the California Cancer Registry (CCR). Our results are consistent with CCR analyses (2) that show AJCC prognostic pathologic staging was better at predicting breast cancer death than anatomic stage. We extend the CCR results by showing our findings are consistent across age and racial and ethnic groups. However, our results may not be comparable with study populations that differ in the distribution of age and/or race or other potential confounders compared with BCSC and CCR populations. Validation of our findings in additional populations is warranted.

We build on the Weiss et al. (2) analysis by reporting results for the TMIST advanced breast cancer definition, which includes approximately 40% breast cancers that if found early may be more treatable. TMIST is a large, randomized trial designed to compare the effectiveness of screening digital mammography vs digital breast tomosynthesis, with a primary goal of testing whether there will be a lower rate of advanced breast cancer in women undergoing screening breast tomosynthesis (8). We found the TMIST definition of advanced breast cancer had the highest sensitivity and the lowest specificity for predicting 5-year breast cancer death, likely because a high proportion of treatable cancers with good survival are categorized as advanced cancer. A recent study reported that a higher proportion of advanced breast cancers defined by the TMIST definition were detected with tomosynthesis vs digital mammography (32.6% [76 of 233] vs 25.0% [9 of 36], respectively), with a higher proportion of lymph node–positive cancers contributing to this difference (22).

The Weiss study excluded women with missing prognostic stage (2). In a sensitivity analysis, we imputed receptor status when missing because a prior study found that missing receptor status was associated with worse prognosis and led to overestimation of survival without imputation (16); however, our results based on the complete case and imputed results were similar

Table 4. Five-year probability of breast cancer survival by AJCC eighth edition staging system and mode of detection and accuracy of predicting 5-year breast cancer death[a]

	AJCC anatomic stage			AJCC prognostic pathologic stage		
Measures	Screen detected (n = 13 236)	Interval cancer (n = 7608)	Clinically detected (n = 4746)	Screen detected (n = 12 279)	Interval cancer (n = 6910)	Clinically detected (n = 4193)
Stage, proportion % (% survival)						
IA	64.6 (99.0)	39.3 (98.7)	24.7 (97.7)	77.9 (99.0)	55.2 (99.1)	42.0 (97.5)
IB	2.8 (98.5)	3.1 (95.4)	2.1 (93.0)	11.2 (96.2)	16.5 (94.1)	17.7 (93.7)
IIA	19.0 (96.3)	25.8 (94.5)	27.7 (93.7)	5.6 (93.7)	11.3 (88.6)	12.9 (86.8)
IIB	7.0 (94.0)	13.8 (91.6)	17.9 (87.4)	1.4 (89.6)	4.1 (86.7)	5.2 (84.8)
IIIA	3.9 (90.6)	8.7 (86.7)	10.8 (82.9)	1.8 (88.7)	4.7 (79.3)	6.9 (81.3)
IIIB	0.3 (75.9)	1.6 (74.8)	3.7 (68.8)	0.6 (81.9)	1.8 (81.3)	3.1 (76.2)
IIIC	1.3 (79.8)	3.3 (72.1)	3.8 (72.6)	0.5 (70.2)	1.6 (61.4)	1.7 (54.2)
IV	1.1 (47.8)	4.4 (43.6)	9.3 (39.1)	1.1 (48.2)	4.8 (43.6)	10.5 (39.1)
IIA or higher	32.6 (92.6)	57.6 (87.0)	73.1 (81.6)	11.4 (85.1)	29.2 (77.4)	40.9 (71.2)
IIB or higher	13.6 (87.6)	31.8 (81.0)	45.4 (74.3)	5.8 (77.6)	17.7 (70.2)	27.8 (63.7)
AUC (95% CI)	0.780 (0.751 to 0.807)	0.814 (0.793 to 0.834)	0.803 (0.781 to 0.825)	0.788 (0.757 to 0.820)	0.863 (0.844 to 0.882)	0.834 (0.810 to 0.856)
Accuracy of predicting breast cancer death stage IIA or higher, test measure % (95% CI)						
Sensitivity	77.0 (72.3 to 81.8)	91.7 (89.0 to 94.2)	94.9 (92.8 to 96.7)	55.0 (49.4 to 60.6)	80.6 (76.8 to 84.3)	84.6 (81.2 to 87.8)
Specificity	68.9 (68.1 to 69.7)	45.5 (44.3 to 46.7)	30.4 (29.0 to 31.9)	90.0 (89.4 to 90.5)	75.4 (74.4 to 76.4)	66.1 (64.6 to 67.7)
PPV	7.4 (6.5 to 8.3)	13.0 (11.9 to 14.2)	18.4 (16.9 to 19.9)	14.9 (12.8 to 17.0)	22.6 (20.6 to 24.7)	28.8 (26.5 to 31.2)
Accuracy of predicting breast cancer death for stage IIB or higher, test measure % (95% CI)						
Sensitivity	54.3 (48.9 to 59.8)	74.3 (70.3 to 78.2)	82.6 (79.2 to 85.9)	42.4 (37.1 to 47.9)	65.1 (60.6 to 69.5)	71.9 (67.8 to 75.9)
Specificity	87.7 (87.2 to 88.3)	72.0 (70.9 to 73.1)	60.7 (59.1 to 62.2)	95.3 (94.9 to 95.7)	86.4 (85.6 to 87.3)	79.4 (78.0 to 80.7)
PPV	12.4 (10.7 to 14.2)	19.0 (17.3 to 20.9)	25.7 (23.6 to 27.9)	22.4 (19.1 to 25.8)	29.8 (27.0 to 32.7)	36.3 (33.3 to 39.4)

[a]AJCC = American Joint Committee on Cancer; AUC = area under the receiver operating characteristic curve; CI = confidence interval; PPV = positive predictive value.

and support the use of AJCC prognostic pathologic stage if receptor status and tumor grade are available. This suggests if receptor status and grade are missing, values can be imputed using multiple imputations. Our results support the recommendations of the AJCC manual in that countries where receptor status and tumor grade information are not available, AJCC anatomic stage IIB or higher had modestly lower accuracy compared with AJCC prognostic pathologic stage IIA or higher and can be used to evaluate effective screening strategies (23).

Advanced breast cancer is an important outcome because it is a surrogate for breast cancer mortality and thus often used as an intermediate outcome to evaluate breast cancer screening strategies. Staging systems and advanced cancer definitions used to evaluate the effectiveness of various screening strategies may vary depending on the preferences and thresholds for an intervention's benefits and harms. Comparing results across breast imaging studies of screening effectiveness will be facilitated by investigators using similar advanced breast definitions cancer and staging systems.

Funding

This research work was supported by the National Institutes of Health (P01CA154292). The Breast Cancer Surveillance Consortium additionally supported data collection for this research with funding from the National Institutes of Health (U54CA163303), the Patient-Centered Outcomes Research Institute (PCORI) award (PCS-1504–30370), and the Agency for Health Research and Quality (R01 HS018366-01A1) and by the University of Vermont Cancer Center with funds generously awarded by the Lake Champlain Cancer Research Organization (grant #032800). Cancer and vital status data collection was supported by several state public health departments and cancer registries (http://www.bcsc-research.org/work/acknowledgement).

Notes

Role of the funder: The funder had no role in the design of the study; the collection, analysis, and interpretation of the data; the writing of the manuscript; and the decision to submit the manuscript for publication.

Disclosures: Dr Janie Lee had a consulting agreement with GE Healthcare (2017 only). Dr Karla Kerlikowske is an unpaid consultant for STRIVE study for GRAIL. Dr Diana Buist receives honorarium from 2015-present from the Athena WISDOM Study Data Safety and Monitoring Board and Microseed Scientific Advisory Board. Dr Diana Miglioretti received honorarium and travel for participating on the Hologic scientific advisory board in May and July 2017. Dr Janie Lee received Research Grant from GE Healthcare (11/15/2016–12/31/2020). Dr Diana Miglioretti receives royalties from Elsevier. All other authors report no potential conflicts of interest.

Author contributions: Dr Miglioretti had full access to all the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis. *Concept and design:* Kerlikowske, Bissell, Buist, Miglioretti. *Acquisition, analysis, or interpretation of data:* All authors. *Drafting of the manuscript:* Kerlikowske, Bissell. *Critical revision of the manuscript for important intellectual content:* All authors. *Statistical analysis:* Bissell, Miglioretti. *Obtained funding:* Kerlikowske, Sprague, Lee, Buist, Henderson, Miglioretti. *Administrative, technical, or material support:* Kerlikowske, Miglioretti, Henderson, Sprague. *Supervision:* Kerlikowske, Miglioretti.

Acknowledgements: We thank the participating women, mammography facilities, and radiologists for the data they provided for this study.

Disclaimer: The views in this work are solely the responsibility of the authors and do not necessarily represent the views of the National Cancer Institute or Patient-Centered Outcomes Research Institute (PCORI), its Board of Governors or Methodology Committee. Neither the National Cancer Institute or Patient Centered Outcome Research Institute had a role in the design or conduct of the study or the reporting of results.

Data Availability

Available after study aims of funded grants are addressed and following approval by the BCSC Steering Committee (http://breastscreening.cancer.gov).

References

1. Kerlikowske K, Zhu W, Tosteson AN, et al. Identifying women with dense breasts at high risk for interval cancer: a cohort study. *Ann Intern Med*. 2015;162(10):673-681.
2. Weiss A, Chavez-MacGregor M, Lichtensztajn DY, et al. Validation study of the American Joint Committee on Cancer Eighth Edition prognostic stage compared with the anatomic stage in breast cancer. *JAMA Oncol*. 2018;4(2):203-209.
3. Nelson HD, Pappas M, Cantor A, Griffin J, Daeges M, Humphrey L. Harms of breast cancer screening: systematic review to update the 2009 U.S. Preventive Services Task Force Recommendation. *Ann Intern Med*. 2016;164(4):256-267.
4. Welch HG, Black WC. Overdiagnosis in cancer. *J Natl Cancer Inst*. 2010;102(9):605-613.
5. Kerlikowske K, Zhu W, Hubbard RA, et al. Outcomes of screening mammography by frequency, breast density, and postmenopausal hormone therapy. *JAMA Intern Med*. 2013;173(9):807-816.
6. Kerlikowske K, Sprague BL, Tosteson ANA, et al. Strategies to identify women at high risk of advanced breast cancer during routine screening for discussion of supplemental imaging. *JAMA Intern Med*. 2019;179(9):1230.
7. Miglioretti D, Zhu W, Kerlikowske K, et al. Risk of less-favorable breast tumor characteristics with biennial versus annual mammography. *JAMA Oncol*. 2015;1(8):1069-1077.
8. National Cancer Institute. TMIST (Tomosynthesis Mammographic Imaging Screening Trial). 2019 . https://www.cancer.gov/about-cancer/treatment/clinical-trials/nci-supported/tmist#goals. Accessed March 22, 2020.
9. Puliti D, Zappa M, Giorgi Rossi P, et al.; the DENSITY Working Group. Volumetric breast density and risk of advanced cancers after a negative screening episode: a cohort study. *Breast Cancer Res*. 2018;20(1):95.
10. Ballard-Barbash R, Taplin SH, Yankaskas BC, et al. Breast Cancer Surveillance Consortium: a national mammography screening and outcomes database. *Am J Roentgenol*. 1997;169(4):1001-1008.
11. Sickles EA, Miglioretti DL, Ballard-Barbash R, et al. Performance benchmarks for diagnostic mammography. *Radiology*. 2005;235(3):775-790.
12. Lehman CD, Arao RF, Sprague BL, et al. National performance benchmarks for modern screening digital mammography: update from the Breast Cancer Surveillance Consortium. *Radiology*. 2017;283(1):49-58.
13. Hortobagyi GN, Connolly JL, Edge SB, et al. Breast. In: MB Amin, S Edge, F Greene, et al., eds. *AJCC Cancer Staging Manual*. 8th ed. New York: Springer International Publishing; 2016.
14. American College of Radiology. *American College of Radiology Breast Imaging Reporting and Data System Atlas (BI-RADS® Atlas)*. Vol 5. Reston, VA: American College of Radiology; 2013.
15. Efron B, Tibshirani R. *An Introduction to the Bootstrap*. Boca Raton, FL: Chapman & Hall/CRC; 1993.
16. Howlader N, Cronin KA, Kurian AW, Andridge R. Differences in breast cancer survival by molecular subtypes in the United States. *Cancer Epidemiol Biomarkers Prev*. 2018;27(6):619-626.
17. White I, Royston P. Imputing missing covariate values for the Cox model. *Stat Med*. 2009;28(15):1982-1998.
18. Heagerty PJ. survivalROC: time-dependent ROC curve estimation from censored survival data. R package version 1.0.3. 2013. https://CRAN.R-project.org/package=survivalROC. *R package version 103* https://CRANR-projectorg/package=survivalROC. Accessed December 26, 2018.

19. Team RC. R: a language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing; 2018.
20. Therneau T. A Package for Survival Analysis in R_. R package version 3.1-11. 2020. https://CRAN.R-project.org/package=survival. Accessed April 06, 2020.
21. Black WC, Welch HG. Screening for disease. *Am J Roentgenol*. 1997;168(1):3-11.
22. Conant EF, Zuckerman SP, McDonald ES, et al. Five consecutive years of screening with digital breast tomosynthesis: outcomes by screening year and round. *Radiology*. 2020;295(2):285-293.
23. Shieh Y, Eklund M, Madlensky L, et al. Breast cancer screening in the precision medicine era: risk-based screening in a population-based trial. *J Natl Cancer Inst*. 2017;109(5):1-8.