Received: 24 July 2023 | Revised: 24 October 2023 | Accepted: 20 November 2023

DOI: 10.1002/cncr.35163

ORIGINAL ARTICLE

Validation of the RSClin risk calculator in the National Cancer Data Base

Augustin G. L. Vannier BA[1] | Asim Dhungana BS[1] | Fangyuan Zhao PhD[2] | Nan Chen MD[3] | Sarah Shubeck MD, MS[4] | Olwen M. Hahn MD[3] | Rita Nanda MD[3] | Nora T. Jaskowiak MD[4] | Gini F. Fleming MD[3] | Olufunmilayo I. Olopade MD[3] | Alexander T. Pearson MD, PhD[3] | Dezheng Huo PhD[2] | Frederick M. Howard MD[3]

[1]Pritzker School of Medicine, University of Chicago, Chicago, Illinois, USA

[2]Department of Public Health Sciences, University of Chicago, Chicago, Illinois, USA

[3]Section of Hematology/Oncology, Department of Medicine, University of Chicago, Chicago, Illinois, USA

[4]Department of Surgery, University of Chicago, Chicago, Illinois, USA

Correspondence

Dezheng Huo, Department of Public Health Sciences and Medicine, University of Chicago, 5841 S Maryland Ave, Chicago, IL 60637, USA.
Email: dhuo@bsd.uchicago.edu

Frederick M. Howard, Elwood V. Jensen Scholars Program, Section of Hematology/Oncology, Department of Medicine, University of Chicago, 5841 S Maryland Ave, Chicago, IL 60637, USA.
Email: frederick.howard@uchospitals.edu

Funding information

US Department of Defense, Grant/Award Number: BC211095P1; Susan G. Komen, Grant/Award Number: SAC 210203; National Cancer Institute, Grant/Award Numbers: 1P20-CA233307, K08CA283261; Breast Cancer Research Foundation, Grant/Award Number: BCRF-21-071

Abstract

Background: Guidelines recommend the use of genomic assays such as OncotypeDx to aid in decisions regarding the use of chemotherapy for hormone receptor–positive, HER2-negative (HR+/HER2−) breast cancer. The RSClin prognostic tool integrates OncotypeDx and clinicopathologic features to predict distant recurrence and chemotherapy benefit, but further validation is needed before broad clinical adoption.

Methods: This study included patients from the National Cancer Data Base (NCDB) who were diagnosed with stage I–III HR+/HER2− breast cancer from 2010 to 2020 and received adjuvant endocrine therapy with or without chemotherapy. RSClin-predicted chemotherapy benefit was stratified into low (<3% reduction in distant recurrence), intermediate (3%–5%), and high (>5%). Cox models were used to model mortality adjusted for age, comorbidity index, insurance, and race/ethnicity.

Results: A total of 285,441 patients were identified for inclusion from the NCDB, with an average age of 60 years and a median follow-up of 58 months. Chemotherapy was associated with improved overall survival only for those predicted to have intermediate (adjusted hazard ratio [aHR], 0.68; 95% confidence interval [CI], 0.60–0.79) and high benefit per RSClin (aHR, 0.66; 95% CI, 0.61–0.72). Consistent benefit was seen in the subset with a low OncotypeDx score (<26) and intermediate (aHR, 0.66; 95% CI, 0.53–0.82) or high (aHR, 0.71; 95% CI, 0.58–0.86) RSClin-predicted benefit. No survival benefit with chemotherapy was seen in patients with a high OncotypeDx score (≥26) and low benefit per RSClin (aHR, 1.70; 95% CI, 0.41–6.99).

Conclusions: RSClin may identify high-risk patients who benefit from treatment intensification more accurately than OncotypeDx, and further prospective study is needed.

KEYWORDS
aromatase inhibitors, breast neoplasms, chemotherapy, clinical decision-making, gene expression profiling, hormone antagonists, precision medicine

INTRODUCTION

Breast cancer is the leading cause of cancer death for women globally, with an estimated 1.7 million cases diagnosed each year.[1] Hormone receptor–positive (HR+) breast cancer is the most common subtype of breast cancer, and constitutes approximately 70% of newly diagnosed cases in the United States.[2] However, there is marked heterogeneity in outcome within HR+ breast cancer, and gene expression assays have been developed to more accurately characterize prognosis and/or chemotherapy benefit, including the OncotypeDx and MammaPrint assays.[3,4] Although genomic assays enhance clinical decision-making, they must be used in combination with a patient's clinical and demographic factors to individualize care. For example, subgroup analyses of the pivotal TAILORx, RxPONDER, and MINDACT trials have demonstrated that some premenopausal women who are not classified as "high risk" by genomic assays may still benefit from chemotherapy.[5-7] One recently described attempt to refine decision-making with the OncotypeDx assay is the RSClin prognostic risk tool.[8] RSClin incorporates age, tumor grade, tumor size, and OncotypeDx scores to provide more accurate prognostic information for patients with node-negative HR+ breast cancer. The RSClin tool provides an estimate of risk of distant metastasis at 10 years with hormonal therapy alone, as well as an estimate of chemotherapy benefit as measured by a reduction in this risk of distant metastasis when chemotherapy is added to hormonal therapy. This tool has been validated with data from a large health registry in Israel, but further real-world evaluation of the tool in larger, diverse populations is required.[8] In addition, whereas OncotypeDx has been prospectively validated to demonstrate that in women with scores over 26, chemotherapy should be offered, there is no clear cutoff for the RSClin tool for which chemotherapy is deemed necessary. Some studies have reserved chemotherapy for patients where RSClin predicts at least a 5% absolute reduction in distant recurrence with the use of chemotherapy, but additional data are needed.[9]

Although genomic recurrence assays have improved the care of women with HR+ breast cancer via the more appropriate selection of patients for chemotherapy, accumulating evidence suggests that these assays are less applicable to non-White patients. This may in part be due to the fact that the most well-studied genomic recurrence assays were developed with samples from predominantly White patients. National Surgical Adjuvant Breast and Bowel Project (NSABP) B-20, which consists of only 6% Black patients, was the largest cohort contributing samples to the development of OncotypeDx.[10] MammaPrint was developed from a single-center cohort of patients younger than 55 years of age treated at the Netherlands Cancer Institute, which likely had limited representation of non-White patients.[4] Multiple studies have demonstrated that Black women with low recurrence scores have a higher risk of recurrence than a comparable body of White patients—although it remains unclear how much is due to differences in disease biology versus socioeconomic factors.[11,12] A prior analysis of the National Cancer Data Base (NCDB) found that the prognostic accuracy of OncotypeDx was nearly halved in Black and Hispanic patients.[13] A study of 95 Black women found that nearly two thirds of patients classified as low or intermediate risk by OncotypeDx would be classified as high risk by MammaPrint—which highlights the difficulties of accurate risk stratification in this population.[14] Additionally, White women are more likely to undergo genomic testing than Black and Hispanic patients in the United States, which furthers disparities.[15,16] RSClin uses the OncotypeDx score, and therefore it is possible that this tool recapitulates the inherent disparities in OncotypeDx testing. Because RSClin was developed via a patient-specific meta-analysis of the NSABP B-14 and TAILORx trials, both of which enrolled predominantly non-Hispanic White women, it may not be optimized for application in diverse patients.

In this study, we provide a large-scale validation of the RSClin risk tool via the NCDB, which correlates tool predictions with long-term survival outcomes. Furthermore, we analyzed RSClin's performance in multiple clinical and demographic subgroups, with particular attention to accuracy in patients of varying race and ethnicity, to determine whether this novel tool can be equitably applied to all patients.

MATERIALS AND METHODS

Study design and data source

This was a retrospective cohort study evaluating the RSClin prognostic risk calculator in patients with nonmetastatic, node-negative HR+/HER2− breast cancer. Patients were identified via the NCDB, a large hospital-based registry maintained jointly by the American College of Surgeons and American Cancer Society, which captures approximately 70% of new cancer diagnoses in the United States.[17] This study was determined to be exempt from review by the University of Chicago institutional review board. No informed consent was obtained because we did not use any identifying patient information. This study followed the Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) reporting guideline.[18]

Study population and covariates

We included patients diagnosed with node-negative HR+/HER2− invasive breast carcinoma from January 1, 2010, through December 31, 2020, who were treated with adjuvant hormonal therapy with or

without chemotherapy. We excluded patients with ductal carcinoma in situ or metastatic disease. Covariates analyzed included age, sex, race/ethnicity, Charlson/Deyo comorbidity index, insurance status, grade, tumor size, and recurrence score result; patients with unknown results for these variables were excluded (Figure S1). For this patient cohort, each unique combination of grade, tumor size, recurrence score, and age was iteratively passed to the online (https://rsclin.genomichealth.com/#/app/home) RSClin calculator to generate predictions of distant disease-free survival and absolute benefit of chemotherapy (which is listed as a reduction in distant recurrence rate compared to hormonal therapy alone) at 10 years. Because the specific hormonal therapy used is not recorded in the NCDB, the RSClin estimates of recurrence and chemotherapy benefit were calculated compared to treatment with aromatase inhibitors unless otherwise stated.

Statistical analysis

Statistical analysis was performed with Python version 3.8.8 (Python Software Foundation, Wilmington, Delaware). Early studies of breast cancer chemotherapy demonstrated benefits of as low as 2%–3% in postmenopausal women,[19] whereas clinicians may prefer to recommend chemotherapy for patients when the benefit is >5%.[9,20] Thus, we evaluated subgroups where RSClin-predicted chemotherapy benefit was <3% (low benefit), 3%–5% (intermediate benefit), and >5% (high benefit). Decisions to administer chemotherapy in the NCDB are also likely confounded by patient performance status as well as socioeconomic factors.[21,22] We performed inverse probability of treatment weighting and calculated propensity scores with a logistic regression model to correct for the influence of age, race/ethnicity, comorbidity index, and insurance status on chemotherapy treatment decisions.[23] Additionally, because age and comorbidity index are strongly associated with noncancer mortality in this population, adjusted hazard ratio (aHR) estimates incorporated these features as covariates to better describe the association of RSClin predictions with breast cancer–specific survival.

The primary outcome was the aHR for overall survival (OS) in patients who received and did not receive chemotherapy for subgroups separated by RSClin-predicted chemotherapy benefit. Secondary outcomes included assessing the prognostic accuracy of RSClin for patients treated with hormonal therapy alone, evaluating the predictive accuracy within clinical and demographic subgroups, and evaluating outcomes in patients with discordant RSClin and OncotypeDx risk categories. Prognostic accuracy of RSClin was compared to the input variables of grade, tumor size, and recurrence score with the Harrell concordance index (C-index).[24] We used restricted cubic spline regression to model chemotherapy benefit as a function of RSClin predictions on a continuous basis.[25] All statistical testing was done at the $\alpha = .05$ significance level, and no correction for multiple hypothesis testing was performed given the exploratory nature of this analysis.

RESULTS

We identified 285,441 patients within the NCDB who met our inclusion criteria. The average patient age was 59.5 years, the mean recurrence score was 16.3, and 41,789 patients (14.6%) received chemotherapy (Tables 1 and S1). All patients received endocrine therapy. The median follow-up in this cohort was 58.3 months, and 11,706 patients (4.6%) died during this follow-up. The vast majority of patients were female ($n = 238,447$; 99.3%), although a few male patients met eligibility criteria ($n = 1994$; 0.7%); male patients had higher grade, larger tumors with higher OncotypeDx scores (Table-S2). The breakdown of race/ethnicity was similar to what was seen in the TAILORx trial[26]: the majority (235,804; 82.6%) of patients were non-Hispanic White, 21,937 (7.7%) were non-Hispanic Black, 13,788 (4.8%) were of Hispanic ethnicity (of any race), and 11,735 (4.1%) were Asian/Pacific Islander. The average predicted risk of distant recurrence at 10 years (via RSClin) was 7.0% with aromatase inhibitors alone, and the average predicted reduction in distant recurrence with chemotherapy was 2.5%. To further characterize the distribution of RSClin predictions in the NCDB population, histograms were generated for the risk of distant recurrence (Figure 1). Notably, non-Hispanic White patients had a lower predicted risk of recurrence and a lower predicted chemotherapy benefit than other racial/ethnic subgroups (Figure S2). For example, if a threshold of 5% absolute benefit per RSClin prediction was used to select patients for chemotherapy, this would result in treatment of 11% of non-Hispanic White, 12% of Hispanic, 13% of Asian/Pacific Islander, and 16% of non-Hispanic Black patients. This is a direct reflection of heterogeneity in the input variables for the RSClin calculator. Consistent with existing studies, non-Hispanic Black patients in this NCDB cohort were younger, with higher grade, larger tumors and higher recurrence scores compared to their non-Hispanic White counterparts (Table S3).

Next, RSClin-predicted risk of distant recurrence on hormonal therapy alone was compared to observed OS in the NCDB cohort. The 243,652 patients treated with hormonal therapy alone were stratified by RSClin-predicted distant recurrence risk in 5% increments (Figure S3; Table S4). Notably, survival declined monotonically with increasing predicted risk. Similar findings were seen in patients younger than 50 years of age, regardless of whether treatment with aromatase inhibitors or tamoxifen was selected for the RSClin calculator (Figure S4; Tables S5 and S6). To determine whether RSClin accurately incorporated clinical and demographic features to predict prognosis, we computed the C-index for a Cox model[24] for OS with RSClin recurrence risk and compared this to Cox models via tumor grade, size, and recurrence score (Table S7). RSClin alone achieved a C-index of 0.574, similar to the C-index of 0.570, for a model with grade, size, and recurrence score. These models demonstrated more prognostic accuracy for OS than recurrence score alone (C-index of 0.543). However, in absolute terms, RSClin predictions remained only a modest predictor of survival (a C-index of 0.5 is the performance that would be obtained via random

TABLE 1 Characteristics of included patients.

		Missing	Overall	Chemotherapy	No chemotherapy	*p*
Patients, No.			285,441	41,789	243,652	
Age, mean (SD), years		0	59.5 (10.5)	55.4 (10.7)	60.2 (10.2)	<.001
Sex, No. (%)	Female	0	283,447 (99.3)	41,501 (99.3)	241,946 (99.3)	.828
	Male		1994 (0.7)	288 (0.7)	1706 (0.7)	
Race/ethnicity, No. (%)	Asian	0	11,735 (4.1)	1918 (4.6)	9817 (4.0)	<.001
	Hispanic		13,788 (4.8)	2218 (5.3)	11,570 (4.7)	
	Native American		814 (0.3)	92 (0.2)	722 (0.3)	
	Non-Hispanic Black		21,937 (7.7)	3922 (9.4)	18,015 (7.4)	
	Non-Hispanic White		235,804 (82.6)	33,437 (80.0)	202,367 (83.1)	
	Other		1363 (0.5)	202 (0.5)	1161 (0.5)	
Insurance, No. (%)	Government/uninsured	0	116,661 (40.9)	12,776 (30.6)	103,885 (42.6)	<.001
	Private/managed		168,780 (59.1)	29,013 (69.4)	139,767 (57.4)	
Charlson/Deyo score, No. (%)	0	0	239,780 (84.0)	35,778 (85.6)	204,002 (83.7)	<.001
	≥1		45,661 (16.0)	6011 (14.4)	39,650 (16.3)	
Grade, No. (%)	1	0	84,015 (29.4)	4686 (11.2)	79,329 (32.6)	<.001
	2		161,435 (56.6)	20,468 (49.0)	140,967 (57.9)	
	3		39,991 (14.0)	16,635 (39.8)	23,356 (9.6)	
Histologic subtype, No. (%)	Ductal	0	220,154 (77.1)	34,799 (83.3)	185,355 (76.1)	<.001
	Ductal and lobular		18,094 (6.3)	2116 (5.1)	15,978 (6.6)	
	Lobular		38,316 (13.4)	3873 (9.3)	34,443 (14.1)	
	Other		8877 (3.1)	1001 (2.4)	7876 (3.2)	
Tumor size, median (IQR), mm		0	15.0 (10.0–20.0)	17.0 (12.0–24.0)	14.0 (10.0–20.0)	<.001
Tumor size, No. (%)	≤20 mm	0	217,857 (76.3)	27,236 (65.2)	190,621 (78.2)	<.001
	21–50 mm		63,181 (22.1)	13,641 (32.6)	49,540 (20.3)	
	>50 mm		4403 (1.5)	912 (2.2)	3491 (1.4)	
Receptor status, No. (%)	ER+ PR+	6682	254,280 (91.2)	31,968 (79.1)	222,312 (93.3)	<.001
	ER+ PR−		24,251 (8.7)	8336 (20.6)	15,915 (6.7)	
	ER− PR+		228 (0.1)	108 (0.3)	120 (0.1)	
ER%, mean (SD)		185,652	93.2 (11.1)	90.3 (14.8)	93.4 (10.7)	<.001
PR%, mean (SD)		185,582	68.4 (34.8)	51.5 (38.4)	69.8 (34.1)	<.001
Ki67%, mean (SD)		237,615	16.4 (14.6)	27.0 (19.9)	15.4 (13.6)	<.001
OncotypeDx score, No. (%)	High (26+)	0	33,468 (11.7)	24,281 (58.1)	9187 (3.8)	<.001
	Intermediate (11–25)		180,769 (63.3)	16,693 (39.9)	164,076 (67.3)	
	Low (0–10)		71,204 (24.9)	815 (2.0)	70,389 (28.9)	
Last contact, mean (SD)		32,554	62.5 (32.2)	73.0 (32.1)	60.6 (31.8)	<.001
Vital status, No. (%)	Alive	32,551	241,184 (95.4)	36,791 (93.9)	204,393 (95.6)	<.001
	Deceased		11,706 (4.6)	2379 (6.1)	9327 (4.4)	
RSClin-predicted percentage risk of distant recurrence at 10 years with AI alone, mean (SD)		0	7.0 (7.6)	16.6 (12.3)	5.4 (4.8)	<.001

TABLE 1 (Continued)

		Missing	Overall	Chemotherapy	No chemotherapy	*p*
RSClin-predicted benefit of chemotherapy (absolute percentage reduction in distant recurrence), No. (%)	High (>5%)	0	31,873 (11.2)	22,196 (53.1)	9677 (4.0)	<.001
	Intermediate (3%–5%)		27,090 (9.5)	10,113 (24.2)	16,977 (7.0)	
	Low (<3%)		226,478 (79.3)	9480 (22.7)	216,998 (89.1)	

Note: All demographic factors other than sex were associated with receipt of chemotherapy.

Abbreviations: AI, aromatase inhibitors; ER, estrogen receptor; IQR, interquartile range; PR, progesterone receptor.

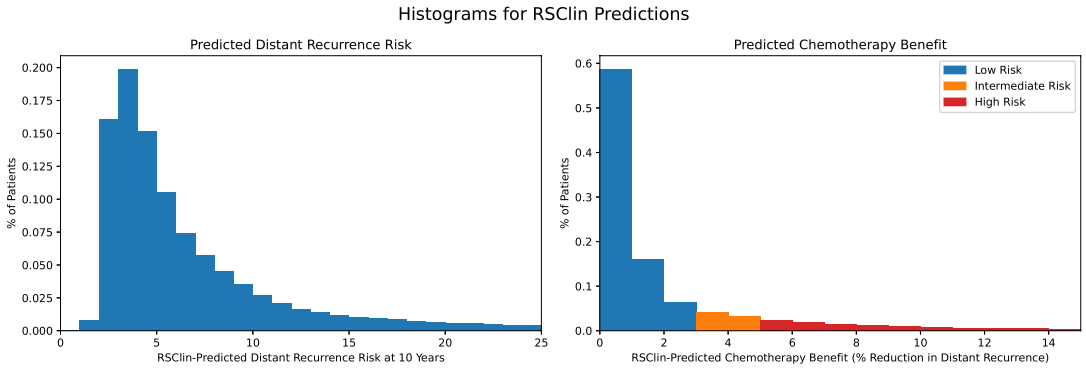

FIGURE 1 Histograms for distributions of RSClin predictions. Predictions of risk of distant recurrence on aromatase inhibitors (*Left*) and absolute chemotherapy benefit as the percentage reduction in distant recurrence (*Right*) are illustrated for the overall study population. For this study, RSClin-predicted chemotherapy benefit was stratified into low (<3% reduction in distant recurrence), intermediate (3%–5%), and high (>5%); these divisions are illustrated above.

chance), and age is a more important factor because models incorporating age achieve a C-index of 0.708–0.719.

We next assessed the accuracy of RSClin by comparing survival with and without chemotherapy in patients with low (<3%), intermediate (3%–5%), and high (>5%) predicted chemotherapy benefit (Figure 2; Table S8). No survival advantage was seen with patients undergoing chemotherapy with low RSClin-predicted benefit (aHR, 0.99; 95% confidence interval [CI], 0.89–1.11; $p = .91$). Significant survival benefit was seen with chemotherapy in patients with intermediate (aHR, 0.68; 95% CI, 0.60–0.79; $p < .01$) and high (aHR, 0.66; 95% CI, 0.61–0.72; $p < .01$) RSClin-predicted benefit. Similar findings were seen in the subset of patients younger than 50 years of age (Figure S5) when calculating RSClin with a presumption of aromatase inhibitor therapy (Table S9), but there was only a trend toward chemotherapy benefit that did not reach statistical significance when calculations were made presuming tamoxifen therapy (Table S10). We also quantified the relationship of RSClin-predicted chemotherapy benefit versus actual survival benefit via a restricted cubic spline model (Figure 3). The survival advantage of chemotherapy became significant when the predicted benefit was at least 3%.

Because the correlation between RSClin predictions and survival may depend on several clinical and demographic factors, we repeated this analysis in additional subgroups (Figure 4). Notably, no patient subgroups experienced a significant survival benefit with chemotherapy when RSClin-predicted benefit was low (<3%). Most subgroups demonstrated a survival benefit when RSClin-predicted benefit was high (>5%). The magnitude of benefit of chemotherapy was smaller in patients with lobular histology compared to ductal histology, and did not reach significance even in the subset with high RSClin-predicted benefit (aHR, 0.80; 95% CI, 0.60–1.07). In general, similar findings were seen across racial and ethnic subgroups. However, we observed no trend toward benefit in Hispanic patients regardless of RSClin predictions, even those predicted to have a high RSClin-predicted benefit (aHR, 1.19; 95% CI, 0.66–2.14).

Finally, we evaluated treatment patterns and outcomes in patients on the basis of RSClin and recurrence score risk categories. In our NCDB cohort, chemotherapy was administered for 73% of patients with a high-risk OncotypeDx recurrence score (≥26; $n = 33,468$) and 25% of patients with high-end intermediate scores (21–25; $n = 38,351$; Figure S6). Similarly, chemotherapy was used in 70% of patients with high ($n = 31,873$) and 37% of patients with intermediate ($n = 27,090$) predicted chemotherapy benefit per RSClin. Whereas only 1.4% ($n = 476$) of patients with a high-risk OncotypeDx recurrence score had a low RSClin-predicted benefit, 21% ($n = 6316$) of patients with high and 75% ($n = 13,640$) of patients with intermediate RSClin-predicted benefit had OncotypeDx recurrence scores of lower than 26.

In these discordant cases where survival status was available, we found that RSClin was more predictive of the impact of chemotherapy (Figure 5; Table S11). In the small subset with patients with a

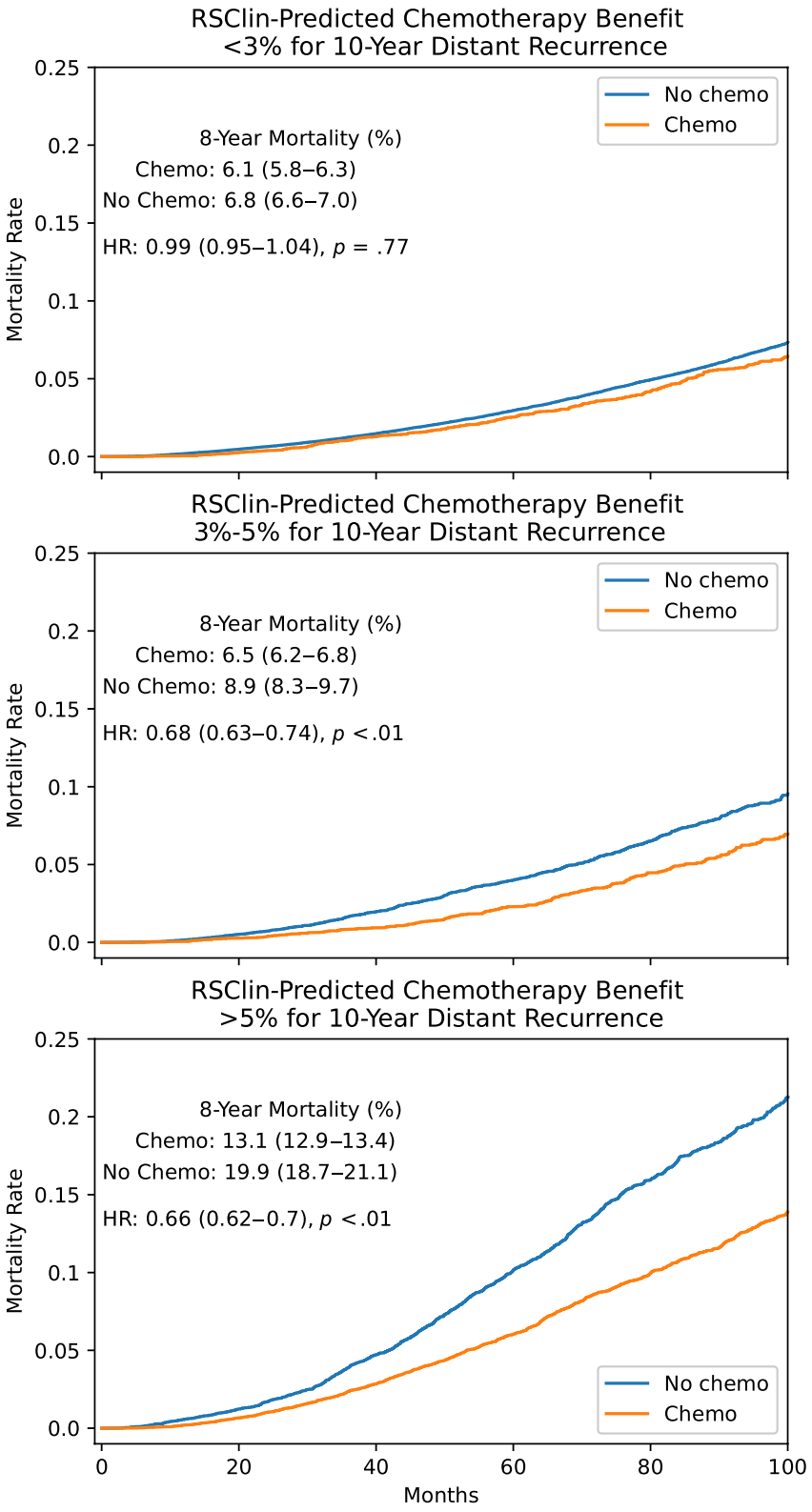

FIGURE 2 Correlation of RSClin predictions and mortality rate in patients treated with or without chemotherapy. Mortality rate is compared between patients treated with and without chemotherapy within three subgroups of RSClin-predicted chemotherapy benefit (reported as the absolute reduction in 10-year distant recurrence rate with chemotherapy). A significant survival benefit is seen starting at the 3%–5% subgroup. Inverse probability of treatment weighting is applied to mitigate the effect of patient fitness on treatment decisions, and hazard ratios are listed for chemotherapy benefit in a multivariable Cox model that also incorporates age and comorbidity index. HR indicates hazard ratio.

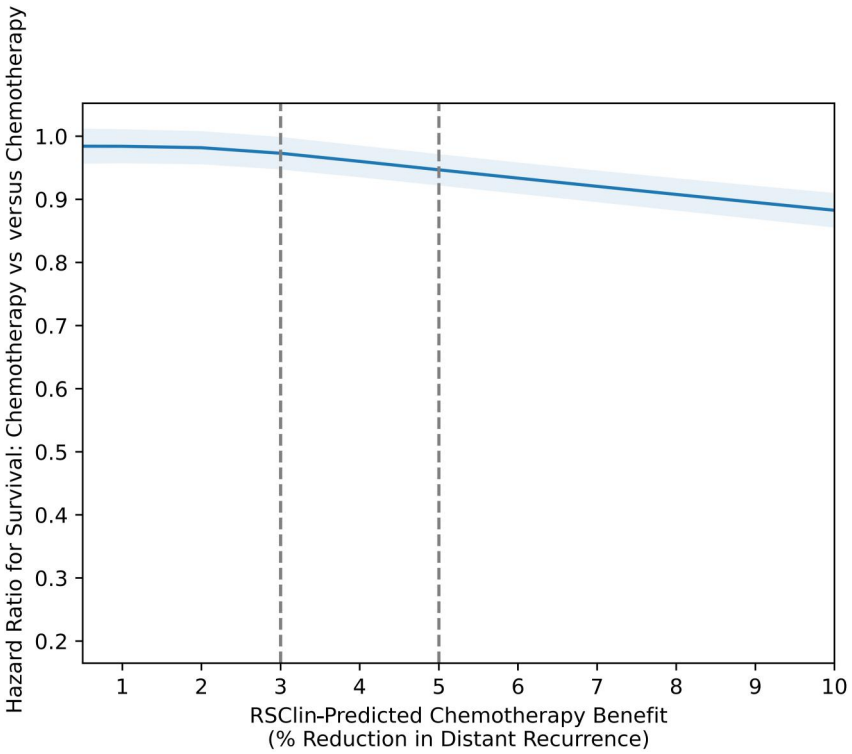

FIGURE 3 Continuous prediction of survival benefit with chemotherapy as a function of RSClin-predicted benefit. Relationship was modeled with a cubic spline regression with four knots, with inverse probability of treatment weighting and covariates of age and comorbidity index included in the model. A significant overall survival benefit with chemotherapy was seen when RSClin-predicted chemotherapy benefit was at least 3%.

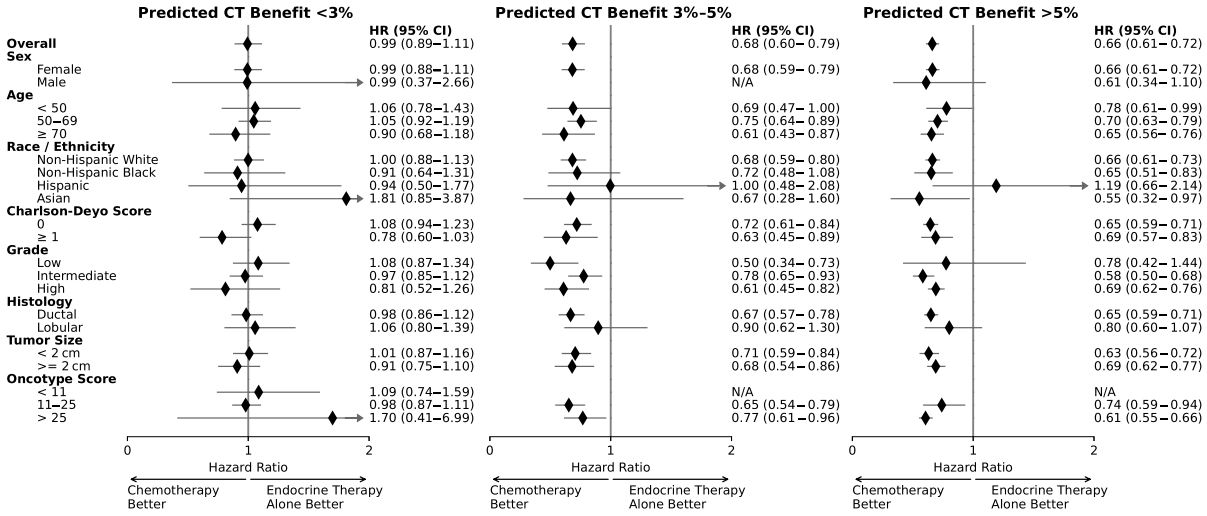

FIGURE 4 Survival benefit of chemotherapy, stratified by RSClin-predicted benefit and clinical/demographic factors. The analysis is pictured with inverse probability of treatment weighting, and models also incorporated age and comorbidity index as variables. No clear subgroup clearly benefitted from chemotherapy when RSClin-predicted benefit was <3%, although there was a trend toward benefit in patients with grade 3 tumors. Conversely, nearly all subgroups demonstrated a clear survival advantage with chemotherapy when RSClin-predicted benefit was >5%. Notable exceptions include lobular tumors and Hispanic patients. CT indicates chemotherapy. Subgroups with fewer than 10 patients where hazard ratio estimates could not be reliably generated are marked with 'N/A'.

high-risk OncotypeDx recurrence score with low RSClin-predicted benefit, there was no discernable difference in survival associated with treatment with chemotherapy (aHR, 1.70; 95% CI, 0.41–6.99; $p = .46$). All these patients had grade 1 tumors, with an OncotypeDx recurrence score no greater than 28 and tumor size of less than 1.3 cm. By contrast, patients with a low OncotypeDx score and high

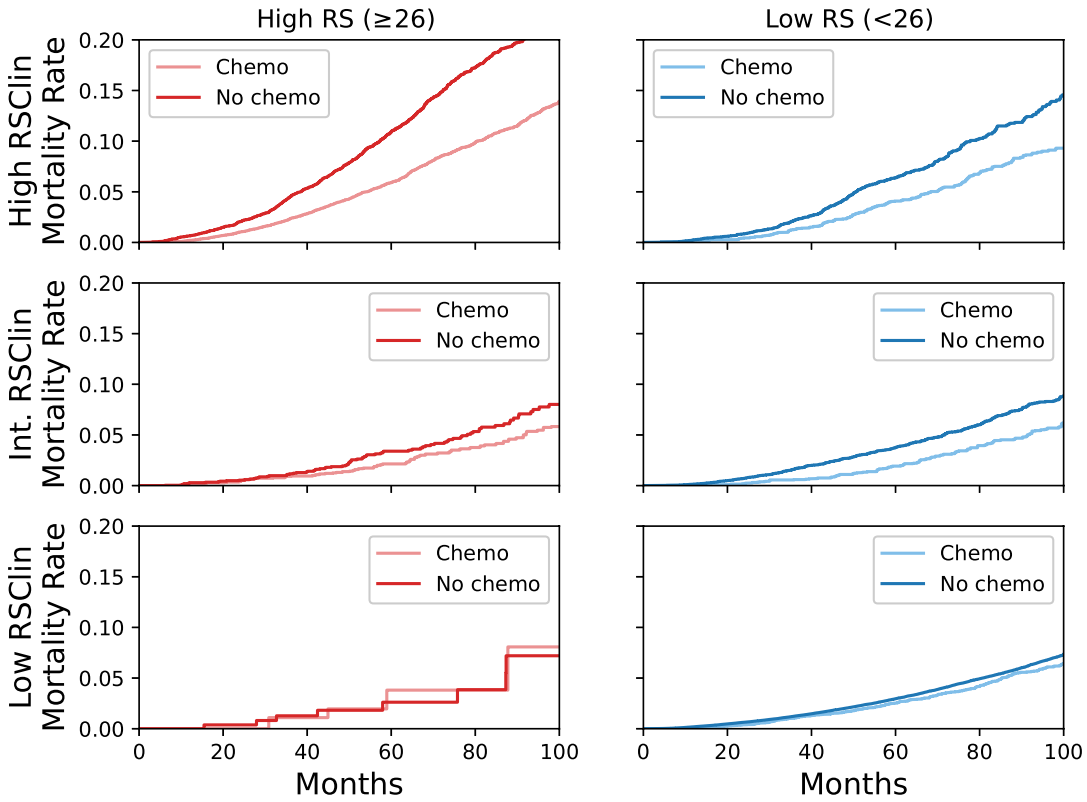

FIGURE 5 Mortality rate in patients in discordant RSClin and OncotypeDx risk groups. High RSClin is defined as a predicted absolute chemotherapy benefit of >5%, intermediate as between 3% and 5%, and low risk as less than 3%. High OncotypeDx risk is defined as a score of 26 or higher. Inverse probability of treatment weighting is applied to survival curves. Int. indicates intermediate; RS, RSClin.

(aHR, 0.71; 95% CI, 0.58–0.86; $p < .01$) or intermediate (aHR, 0.66; 95% CI, 0.53–0.82; $p < .01$) RSClin-predicted benefit had prolonged OS with chemotherapy. Notably, the intermediate RSClin group had a median OncotypeDx score of 23 (interquartile range, 20–24). Similar results were seen in patients aged 50 years or older (Table S12; Figure S7) and those younger than 50 years of age (Table S13), which suggests that this effect is not dependent on menopausal status. Similar to the overall population, patients younger than 50 years of age, with a low OncotypeDx score but a high or intermediate RSClin-predicted benefit, showed improved OS with chemotherapy. There was only a trend toward chemotherapy benefit in patients with concordant high RSClin-predicted chemotherapy benefit and OncotypeDx scores (aHR, 0.78; 95% CI, 0.57–1.08; $p = .14$), which may not have reached significance in part because of the imbalanced treatment decisions in this cohort (only 13%, or $n = 627$, received endocrine therapy alone). The results in patients aged 50 years or older mirrored the overall population, with an advantage to chemotherapy exclusively seen in patients with intermediate or high predicted benefit per RSClin.

DISCUSSION

In this study, we further validated the RSClin prognostic risk tool in a large cohort of patients in the NCDB. We demonstrated that increasing RSClin-predicted recurrence correlates with decreased survival in patients treated with hormonal therapy alone and identifies a subset of patients who benefit from chemotherapy. Our subgroup analysis by clinical and demographic factors provides insight into the validity of RSClin in different patient populations, with notable heterogeneity in patients with lobular breast cancer or high-grade tumors and in patients of varying race/ethnicity. Several studies have suggested that chemotherapy may be less impactful in lobular tumors, even with high-risk OncotypeDx scores,[27,28] and our subgroup analysis similarly found that lobular tumors had a trend toward lower chemotherapy benefit (compared to ductal tumors) for each category of RSClin-predicted risk.

In our study, we consistently found a survival benefit for patients with an intermediate (3%–5%) and high (>5%) predicted chemotherapy benefit per RSClin, although the absolute difference in mortality was more clinically meaningful in the latter group, which reached 6.8% at 8 years, and the survival advantage was more consistent across subgroups in the high–predicted benefit group. In our NCDB cohort, 21% of node-negative patients would have an intermediate or high predicted benefit per RSClin, which is much larger than the 12% who have a high-risk OncotypeDx score (≥26). However, the TAILORx trial has also suggested that chemotherapy may benefit some patients with intermediate-risk OncotypeDx scores, with a trend toward lower recurrence rates in women aged 50 years or younger with scores of 16–25, especially in patients with high–intermediate scores of 21–25.[6] These findings in TAILORx may in part be due to chemotherapy-induced menopause, and the use of

ovarian suppression was likely low in our study because of the time frame. Thus, our findings should not prompt the use of chemotherapy for all patients with an elevated risk per RSClin. Nonetheless, our study demonstrates that outcomes are suboptimal with hormonal therapy alone, and this tool may prove helpful to select patients for treatment intensification with ovarian suppression,[29] CDK4/6 inhibitors,[30] or bisphosphonates,[31] with chemotherapy reserved for patients at highest risk. RSClin may also serve to deescalate therapy—in the small group of patients with low risk via RSClin but a high OncotypeDx score, no overall survival advantage was seen with chemotherapy (Figure 5, *Bottom Left*).

Importantly, we found that RSClin, like OncotypeDx, may be less prognostic in some non-White patients. There is clear heterogeneity in RSClin risk predictions in racial/ethnic subgroups, which is reflective of underlying differences in model inputs. It is well described that Black women are more likely to have high-risk recurrence score results,[12] high-grade tumors, and low estrogen receptor expression.[32] In our study, we demonstrated that regardless of the cutoff for RSClin predictions, chemotherapy will be recommended more frequently for non-White patients—in particular, Black women would be nearly 50% more likely than White women to be recommended to undergo chemotherapy with a cutoff of 5% predicted chemotherapy benefit. However, the survival benefit of chemotherapy in our study only reached significance for the highest RSClin risk category in Black women and was not seen for any risk category in Hispanic patients, which suggests it may not stratify patient risk in these groups. Similarly, an analysis of the TAILORx trial found that Black patients with intermediate recurrence scores had a 60% higher distant recurrence rate than White patients but did not experience a chemotherapy benefit.[26] Black women with breast cancer are 40% more likely to die of their disease than White women[2]; there are myriad contributors to this gap, including differences in disease biology, early detection, treatment delays,[33,34] and access to treatment.[21,22] It is important that diverse patient data are used to drive prognostic testing and model development to ensure that these disparities are not perpetuated given the biologic differences in breast cancer in Black and Hispanic patients (weaker ER expression,[35,36] differences in HER2 expression,[33,37] and a higher proportion of GATA3 and TP53 mutations[38-40]).

Our study described the performance of the RSClin calculator in the largest clinical cancer registry in the world, which allows for careful assessment of performance in a variety of subgroups. However, there are several limitations to the use of the NCDB. Most notably, only OS data are collected, which limits direct validation of RSClin predictions, which are measured in terms of distant recurrence rates. OS is moderately correlated with disease-free survival in breast cancer,[41] and likely serves as a reasonable proxy for distant recurrence rates. The duration of follow-up also limits the correlation of OS to distant recurrence rates because there is a several year delay between the development of metastasis and death from breast cancer. Nonetheless, the large sample size enables detection of minute differences in survival, which may compound further follow-up. Another limitation of our study is the lack of specific treatment annotations in the NCDB. We conducted much of our analysis presuming treatment with aromatase inhibitors because the type of endocrine therapy is not specified in the NCDB, but many premenopausal women included in this study may have received tamoxifen. However, in women younger than 50 years of age, estimating risk with tamoxifen only demonstrated a nonsignificant trend toward chemotherapy benefit in the RSClin risk categories defined in our analysis. Selecting tamoxifen treatment raises the RSClin predictions of chemotherapy benefit, which may assign more lower risk women to the intermediate/high-benefit categories and may explain this discrepancy. The thresholds defined with aromatase inhibitor treatment as an input for RSClin still identified younger women with a significant benefit of chemotherapy, and thus may still be reasonably applied to identify younger women for treatment intensification. Additionally, the NCDB does not provide annotations for menopausal status, which would be particularly important given that ovarian function suppression may be the prime driver of chemotherapy benefit in intermediate-risk node-negative patients. Although we performed an age-based analysis (separating patients older and younger than 50 years of age) and found largely similar results, the variable onset of menopause limits our confidence in this finding. Additionally, we were unable to include the specific chemotherapy regimen received or the dose intensity, which may account for some of the racial disparities in chemotherapy benefit because non-White patients are more likely to receive nonstandard regimens and are less likely to complete planned therapy.[42,43] Finally, limited comorbidity information is included in the NCDB. We attempted to control for the impact of comorbidities and other demographic factors in chemotherapy administration via inverse probability of treatment weighting, but in our subgroup analysis, older patients experienced a greater magnitude of chemotherapy benefit than younger patients when RSClin predicted a high benefit, which may reflect the selection of relatively healthier older patients for chemotherapy.

In conclusion, the RSClin prognostic risk tool accurately predicts OS with hormonal therapy in the NCDB—performing as well as a single metric or in combination with recurrence score, grade, and tumor size. Increased mortality was seen in node-negative patients treated with hormonal therapy alone when RSClin predicted a chemotherapy benefit of at least 3%, although this mortality difference was most consistent across subgroups when predicted chemotherapy benefit was more than 5%. Further study is needed to clarify optimal strategies in such patients, which may include ovarian suppression, chemotherapy, or other emerging adjuvant therapies.

AUTHOR CONTRIBUTIONS

Augustin G. L. Vannier: Conceptualization, methodology, data curation, formal analysis, and writing–original draft. **Asim Dhungana**: Data curation and formal analysis. **Fangyuan Zhao**: Data curation and formal analysis. **Nan Chen**: Writing–review and editing. **Sarah Shubeck**: Writing–review and editing. **Olwen M. Hahn**: Writing–review and editing. **Rita Nanda**: Writing–review and editing. **Nora T. Jaskowiak**: Writing–review and editing. **Gini F. Fleming**: Writing–review and editing. **Olufunmilayo I. Olopade**: Writing–review and editing.

Alexander T. Pearson: Writing–review and editing. **Dezheng Huo**: Conceptualization, methodology, data curation, formal analysis, and writing–review and editing. **Frederick M. Howard**: Conceptualization, methodology, data curation, and writing–original draft.

ACKNOWLEDGMENTS

Frederick M. Howard received funding support from National Institutes of Health/National Cancer Institute (NIH/NCI) Grant K08CA283261 and the Cancer Research Foundation. Dezheng Huo, Alexander T. Pearson, and Frederick M. Howard received support from the Department of Defense (BC211095P1). Dezheng Huo, Rita Nanda, and Olufunmilayo I. Olopade received support from the NIH/NCI (1P20-CA233307). Olufunmilayo I. Olopade received support from Susan G. Komen (SAC 210203). Dezheng Huo and Olufunmilayo I. Olopade received support from the Breast Cancer Research Foundation (BCRF-21-071). These funding organizations did not have a role in the design and conduct of the study; the collection, management, analysis, and interpretation of the data; the preparation, review, and approval of the manuscript; or the decision to submit the manuscript for publication.

CONFLICT OF INTEREST STATEMENT

Nan Chen reports consulting/advisory work for Stemline Therapeutics, Seagen, Olema Oncology, and Eli Lilly. Olwen M. Hahn reports consulting/advisory work for Pfizer. Rita Nanda reports contracted research with Arvinas, AstraZeneca, Celgene, Corcept Therapeutics, Genentech/Roche, Immunomedics, Merck, OBI Pharma, Odonate Therapeutics, Oncosec, Pfizer, Seattle Genetics, and Taiho and consulting fees from AstraZeneca, BeyondSpring, Cardinal Health, Daiichi Sankyo, Fujifilm, General Electric, Immunomedics/Gilead, Infinity, iTeos, Merck, Moderna, Novartis, OBI Pharma, Oncosec, Pfizer, Sanofi, Seagen, and Stemline Therapeutics. Gini F. Fleming reports consulting/advisory work for Compugen, Sermonix, Molecular Templates, Roche, Astellas, Iovance, Pfizer, AstraZeneca, K-Group Beta, and Corcept Therapeutics. Olufunmilayo I. Olopade reports ownership interest in 54Gene, CancerIQ, and Tempus and financial interest in Color Genomics, Healthy Life for All Foundation, and Roche/Genentech. Alexander T. Pearson reports consulting work for AbbVie. The other authors declare no conflicts of interest.

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are available from the American College of Surgeons/Commission on Cancer. Restrictions apply to the availability of these data, which were used under license for this study. Data are available from https://www.facs.org/quality-programs/cancer-programs/national-cancer-database/. The code used for analysis in this article and for the generation of relevant figures and tables is available at https://github.com/fmhoward/NCDBRSClin.

ORCID

Rita Nanda https://orcid.org/0000-0001-5248-0876
Olufunmilayo I. Olopade https://orcid.org/0000-0002-9936-1599
Frederick M. Howard https://orcid.org/0000-0002-6039-1141

REFERENCES

1. Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin*. 2018;68(6):394-424. doi:10.3322/caac.21492
2. DeSantis CE, Ma J, Gaudet MM, et al. Breast cancer statistics, 2019. *CA Cancer J Clin*. 2019;69(6):438-451. doi:10.3322/caac.21583@10.1002/(ISSN)2045-7634.oncology-breast-cancer
3. Paik S, Shak S, Tang G, et al. A multigene assay to predict recurrence of tamoxifen-treated, node-negative breast cancer. *N Engl J Med*. 2004;351(27):2817-2826. doi:10.1056/NEJMoa041588
4. van't Veer LJ, Dai H, van de Vijver MJ, et al. Gene expression profiling predicts clinical outcome of breast cancer. *Nature*. 2002;415(6871):530-536. doi:10.1038/415530a
5. Piccart M, van 't Veer LJ, Poncet C, et al. 70-gene signature as an aid for treatment decisions in early breast cancer: updated results of the phase 3 randomised MINDACT trial with an exploratory analysis by age. *Lancet Oncol*. 2021;22(4):476-488. doi:10.1016/S1470-2045(21)00007-3
6. Sparano JA, Gray RJ, Ravdin PM, et al. Clinical and genomic risk to guide the use of adjuvant therapy for breast cancer. *N Engl J Med*. 2019;380(25):2395-2405. doi:10.1056/NEJMoa1904819
7. Kalinsky K, Barlow WE, Meric-Bernstam F, et al. Abstract GS3-00: first results from a phase III randomized clinical trial of standard adjuvant endocrine therapy (ET) +/- chemotherapy (CT) in patients (pts) with 1-3 positive nodes, hormone receptor-positive (HR+) and HER2-negative (HER2-) breast cancer (BC) with recurrence score (RS) < 25: SWOG S1007 (RxPonder). *Cancer Res*. 2021;81(suppl 4):GS3-00. doi:10.1158/1538-7445.SABCS20-GS3-00
8. Sparano JA, Crager MR, Tang G, Gray RJ, Stemmer SM, Shak S. Development and validation of a tool integrating the 21-gene recurrence score and clinical-pathological features to individualize prognosis and prediction of chemotherapy benefit in early breast cancer. *J Clin Oncol*. 2021;39(6):557-564. doi:10.1200/JCO.20.03007
9. Zimmerman BS, Berwick S, Kessler AJ, et al. Application of RSClin to guide treatment recommendations for premenopausal patients with early-stage hormone-positive breast cancer and intermediate risk oncotype recurrence scores. *J Clin Oncol*. 2021;39(15)(suppl):e12511. doi:10.1200/JCO.2021.39.15_suppl.e12511
10. Fisher B, Dignam J, Emir B, et al. Tamoxifen and chemotherapy for lymph node-negative, estrogen receptor-positive breast cancer. *J Natl Cancer Inst*. 1997;89(22):1673-1682. doi:10.1093/jnci/89.22.1673
11. Collin LJ, Yan M, Jiang R, et al. Oncotype DX recurrence score implications for disparities in chemotherapy and breast cancer mortality in Georgia. *NPJ Breast Cancer*. 2019;5:32. doi:10.1038/s41523-019-0129-3
12. Hoskins KF, Danciu OC, Ko NY, Calip GS. Association of race/ethnicity and the 21-gene recurrence score with breast cancer-specific mortality among US women. *JAMA Oncol*. 2021;7(3):370-378. doi:10.1001/jamaoncol.2020.7320
13. Ibraheem A, Olopade OI, Huo D. Propensity score analysis of the prognostic value of genomic assays for breast cancer in diverse populations using the National Cancer Data Base. *Cancer*. 2020;126(17):4013-4022. doi:10.1002/cncr.32956
14. Robinson PA, Tsai ML, Lo SS, et al. Genomic risk classification by the 70-gene signature and 21-gene assay in African American early-stage breast cancer patients. *J Clin Oncol*. 2021;39(15)(suppl):e12568. doi:10.1200/JCO.2021.39.15_suppl.e12568
15. Davis BA, Aminawung JA, Abu-Khalaf MM, et al. Racial and ethnic disparities in Oncotype DX test receipt in a statewide population-based study. *J Natl Compr Cancer Netw*. 2017;15(3):346-354. doi:10.6004/jnccn.2017.0034
16. Press DJ, Ibraheem A, Dolan ME, Goss KH, Conzen S, Huo D. Racial disparities in omission of Oncotype DX but no racial disparities in

chemotherapy receipt following completed Oncotype DX test results. *Breast Cancer Res Treat*. 2018;168(1):207-220. doi:10.1007/s10549-017-4587-8
17. Bilimoria KY, Stewart AK, Winchester DP, Ko CY. The National Cancer Data Base: a powerful initiative to improve cancer care in the United States. *Ann Surg Oncol*. 2008;15(3):683-690. doi:10.1245/s10434-007-9747-3
18. von Elm E, Altman DG, Egger M, Pocock SJ, Gøtzsche PC, Vandenbroucke JP. Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. *BMJ*. 2007;335(7624):806-808. doi:10.1136/bmj.39335.541782.AD
19. Early Breast Cancer Trialists' Collaborative Group. Polychemotherapy for early breast cancer: an overview of the randomised trials. *Lancet*. 1998;352(9132):930-942.
20. Stiggelbout AM, de Haes JC, van de Velde CJ. Adjuvant chemotherapy in node negative breast cancer: patterns of use and oncologists' preferences. *Ann Oncol*. 2000;11(5):631-633. doi:10.1023/a:1008379628579
21. Daly B, Olopade OI. A perfect storm: how tumor biology, genomics, and health care delivery patterns collide to create a racial survival disparity in breast cancer and proposed interventions for change. *CA Cancer J Clin*. 2015;65(3):221-238. doi:10.3322/caac.21271
22. Howard FM, Olopade OI. Epidemiology of triple-negative breast cancer: a review. *Cancer J*. 2021;27(1):8-16. doi:10.1097/PPO.0000000000000500
23. Austin PC, Stuart EA. Moving towards best practice when using inverse probability of treatment weighting (IPTW) using the propensity score to estimate causal treatment effects in observational studies. *Stat Med*. 2015;34(28):3661-3679. doi:10.1002/sim.6607
24. Harrell FE, Califf RM, Pryor DB, Lee KL, Rosati RA. Evaluating the yield of medical tests. *JAMA*. 1982;247(18):2543-2546, doi:10.1001/jama.247.18.2543
25. Desquilbet L, Mariotti F. Dose-response analyses using restricted cubic spline functions in public health research. *Stat Med*. 2010;29(9):1037-1057. doi:10.1002/sim.3841
26. Albain KS, Gray RJ, Makower DF, et al. Race, ethnicity, and clinical outcomes in hormone receptor-positive, HER2-negative, node-negative breast cancer in the randomized TAILORx trial. *J Natl Cancer Inst*. 2021;113(4):390-399. doi:10.1093/jnci/djaa148
27. Kizy S, Huang JL, Marmor S, Tuttle TM, Hui JYC. Impact of the 21-gene recurrence score on outcome in patients with invasive lobular carcinoma of the breast. *Breast Cancer Res Treat*. 2017;165(3):757-763. doi:10.1007/s10549-017-4355-9
28. Chen XH, Zhang WW, Wang J, et al. 21-gene recurrence score and adjuvant chemotherapy decisions in patients with invasive lobular breast cancer. *Biomarkers Med*. 2019;13(2):83-93. doi:10.2217/bmm-2018-0396
29. Francis PA, Pagani O, Fleming GF, et al. Tailoring adjuvant endocrine therapy for premenopausal breast cancer. *N Engl J Med*. 2018;379(2):122-137. doi:10.1056/NEJMoa1803164
30. Johnston SRD, Harbeck N, Hegg R, et al. Abemaciclib combined with endocrine therapy for the adjuvant treatment of HR+, HER2-node-positive, high-risk, early breast cancer (monarchE). *J Clin Oncol*. 2020;38(34):3987-3998. doi:10.1200/JCO.20.02514
31. Eisen A, Somerfield MR, Accordino MK, et al. Use of adjuvant bisphosphonates and other bone-modifying agents in breast cancer: ASCO-OH (CCO) guideline update. *J Clin Oncol*. 2022;40(7):787-800. doi:10.1200/JCO.21.02647
32. Warner ET, Tamimi RM, Hughes ME, et al. Racial and ethnic differences in breast cancer survival: mediating effect of tumor characteristics and sociodemographic and treatment factors. *J Clin Oncol*. 2015;33(20):2254-2261. doi:10.1200/JCO.2014.57.1349
33. Shubeck S, Zhao F, Howard FM, Olopade OI, Huo D. Response to treatment, racial and ethnic disparity, and survival in patients with breast cancer undergoing neoadjuvant chemotherapy in the US. *JAMA Netw Open*. 2023;6(3):e235834. doi:10.1001/jamanetworkopen.2023.5834
34. Zhao F, Miyashita M, Hattori M, et al. Racial disparities in pathological complete response among patients receiving neoadjuvant chemotherapy for early-stage breast cancer. *JAMA Netw Open*. 2023;6(3):e233329. doi:10.1001/jamanetworkopen.2023.3329
35. Purrington KS, Gorski D, Simon MS, et al. Racial differences in estrogen receptor staining levels and implications for treatment and survival among estrogen receptor positive, HER2-negative invasive breast cancers. *Breast Cancer Res Treat*. 2020;181(1):145-154. doi:10.1007/s10549-020-05607-4
36. Patel TA, Colon-Otero G, Bueno Hume C, Copland JA, Perez EA. Breast cancer in Latinas: gene expression, differential response to treatments, and differential toxicities in Latinas compared with other population groups. *Oncologist*. 2010;15(5):466-475. doi:10.1634/theoncologist.2010-0004
37. Peiffer DS, Zhao F, Chen N, et al. Clinicopathologic characteristics and prognosis of ERBB2-low breast cancer among patients in the National Cancer Database. *JAMA Oncol*. 2023;9(4):500-510. doi:10.1001/jamaoncol.2022.7476
38. Ademuyiwa FO, Tao Y, Luo J, Weilbaecher K, Ma CX. Differences in the mutational landscape of triple negative breast cancer in African Americans and Caucasians. *Breast Cancer Res Treat*. 2017;161(3):491-499. doi:10.1007/s10549-016-4062-y
39. Pitt JJ, Riester M, Zheng Y, et al. Characterization of Nigerian breast cancer reveals prevalent homologous recombination deficiency and aggressive molecular features. *Nat Commun*. 2018;9(1):4181. doi:10.1038/s41467-018-06616-0
40. Miyashita M, Bell JSK, Wenric S, et al. Molecular profiling of a real-world breast cancer cohort with genetically inferred ancestries reveals actionable tumor biology differences between European ancestry and African ancestry patient populations. *Breast Cancer Res*. 2023;25(1):58. doi:10.1186/s13058-023-01627-2
41. Ng R, Pond GR, Tang PA, MacIntosh PW, Siu LL, Chen EX. Correlation of changes between 2-year disease-free survival and 5-year overall survival in adjuvant breast cancer trials from 1966 to 2006. *Ann Oncol*. 2008;19(3):481-486. doi:10.1093/annonc/mdm486
42. Griggs JJ, Culakova E, Sorbero MES, et al. Social and racial differences in selection of breast cancer adjuvant chemotherapy regimens. *J Clin Oncol*. 2007;25(18):2522-2527. doi:10.1200/JCO.2006.10.2749
43. Green AK, Aviki EM, Matsoukas K, Patil S, Korenstein D, Blinder V. Racial disparities in chemotherapy administration for early stage breast cancer: a systematic review and meta-analysis. *Breast Cancer Res Treat*. 2018;172(2):247-263. doi:10.1007/s10549-018-4909-5

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Vannier AGL, Dhungana A, Zhao F, et al. Validation of the RSClin risk calculator in the National Cancer Data Base. *Cancer*. 2023;1-11. doi:10.1002/cncr.35163