Zhao *et al. Breast Cancer Research* (2024) 26:148
https://doi.org/10.1186/s13058-024-01905-7

Breast Cancer Research

RESEARCH Open Access

Predicting pathologic complete response to neoadjuvant chemotherapy in breast cancer using a machine learning approach

Fangyuan Zhao[1], Eric Polley[2], Julian McClellan[2], Frederick Howard[3], Olufunmilayo I. Olopade[3] and Dezheng Huo[2,3*]

Abstract

Background For patients with breast cancer undergoing neoadjuvant chemotherapy (NACT), most of the existing prediction models of pathologic complete response (pCR) using clinicopathological features were based on standard statistical models like logistic regression, while models based on machine learning mostly utilized imaging data and/or gene expression data. This study aims to develop a robust and accessible machine learning model to predict pCR using clinicopathological features alone, which can be used to facilitate clinical decision-making in diverse settings.

Methods The model was developed and validated within the National Cancer Data Base (NCDB, 2018–2020) and an external cohort at the University of Chicago (2010–2020). We compared logistic regression and machine learning models, and examined whether incorporating quantitative clinicopathological features improved model performance. Decision curve analysis was conducted to assess the model's clinical utility.

Results We identified 56,209 NCDB patients receiving NACT (pCR rate: 34.0%). The machine learning model incorporating quantitative clinicopathological features showed the best discrimination performance among all the fitted models [area under the receiver operating characteristic curve (AUC): 0.785, 95% confidence interval (CI): 0.778–0.792], along with outstanding calibration performance. The model performed best among patients with hormone receptor positive/human epidermal growth factor receptor 2 negative (HR+/HER2-) breast cancer (AUC: 0.817, 95% CI: 0.802–0.832); and by adopting a 7% prediction threshold, the model achieved 90.5% sensitivity and 48.8% specificity, with decision curve analysis finding a 23.1% net reduction in chemotherapy use. In the external testing set of 584 patients (pCR rate: 33.4%), the model maintained robust performance both overall (AUC: 0.711, 95% CI: 0.668–0.753) and in the HR+/HER2- subgroup (AUC: 0.810, 95% CI: 0.742–0.878).

Conclusions The study developed a machine learning model (https://huolab.cri.uchicago.edu/sample-apps/pcrmodel) to predict pCR in breast cancer patients undergoing NACT that demonstrated robust discrimination and calibration performance. The model performed particularly well among patients with HR+/HER2- breast cancer, having the potential to identify patients who are less likely to achieve pCR and can consider alternative treatment strategies over chemotherapy. The model can also serve as a robust baseline model that can be integrated with smaller datasets containing additional granular features in future research.

*Correspondence:
Dezheng Huo
dhuo@bsd.uchicago.edu

Full list of author information is available at the end of the article

Keywords Breast cancer, Hormone receptor positive, Neoadjuvant chemotherapy, Pathologic complete response, Prediction model, Machine learning, Decision curve analysis

Background

Breast cancer is the most common malignancy and the second leading cause of cancer-related death among women in the US [1]. Fortunately, the mortality rates of breast cancer have been decreasing steadily since the 1990s [2], resulting from advances in early detection and treatment methods. Among these advancements, the use of neoadjuvant chemotherapy (NACT) in clinical practice has grown in particular due to its ability to downsize locally advanced and/or inoperable tumors and increase the chances of breast-conserving surgery [3]. Many randomized trials have demonstrated equivalent long-term survival benefits between adjuvant and neoadjuvant settings [4, 5]. The response to neoadjuvant treatment can be monitored, with therapies such as trastuzumab emtansine and capecitabine specifically being used for patients with residual disease post-treatment [6–8].

The optimal response to NACT is a pathologic complete response (pCR), which is also considered as an efficient surrogate endpoint of overall survival [9]. However, pCR rates can range from under 10% to over 60% depending on breast cancer receptor subtype and treatment regimen [5, 10]; studies have also shown that patients from different racial/ethnic groups experience significantly different pCR rates [11, 12]. Meanwhile, chemotherapy may also lead to unfavorable changes in patients' quality of life and physical functioning [13, 14]. Therefore, identifying patients less likely to respond well to NACT (i.e., achieve pCR) a priori and suggesting them towards alternative treatment regimens instead of chemotherapy might optimize treatment outcomes while reducing undue toxicity.

Traditionally, clinical decisions on treatment selection are based on tumor extent and receptor status, raising the need for a more robust data-driven approach [15]. There have been efforts in developing prediction models of pCR using standard statistical models like multivariable logistic regression [16, 17], while there is emerging interest in applying machine learning techniques that can potentially improve predictive performance. Besides basic clinicopathological features like tumor stage, grade and subtype [9, 18], quantitative biomarkers like estrogen receptor percentage positivity (ER%), progesterone receptor percentage positivity (PR%), human epidermal growth factor receptor 2 (HER2) immunohistochemistry semi-quantitative score, amplification of HER2 and Ki-67 scores were also shown to be associated with pCR [19–23]. Machine learning tools are particularly better in handling these quantitative features as well as more granular features like gene expression and imaging data, capturing complex patterns that extend beyond traditional linear relationships. In fact, most of the existing prediction models of pCR using machine learning utilized imaging data and/or gene expression data [24–29]. Meanwhile, very few studies have built machine learning models utilizing clinicopathological features alone (area under the receiver operating characteristic curve, AUC ranging from 0.64 to 0.88), with some of them including treatment data as predictors. Furthermore, the limited sample sizes in existing studies (ranging from 363 to 2,065), along with their single-institution settings and lack of external validation, raise concerns about their robustness and applicability [30–33].

In this study, we developed and validated a prediction model for pCR using pre-treatment clinicopathological features from the National Cancer Database (NCDB) and evaluated its performance in an external testing set. We adopted a machine learning framework in model development and compared it with logistic regression. Additionally, we examined the predictive value of quantitative features and explored methods to improve model performance across diverse patient groups with differential pCR rates. Finally, we assessed the model's potential in facilitating treatment selection in clinical practice.

Methods

Study population and data source

The model was developed using data collected from the NCDB, a nationally representative hospital-based registry covering approximately 70% of all new invasive cancer diagnoses in the U.S [34]. Within the NCDB, we identified patients diagnosed with invasive non-metastatic breast cancer from 2018 to 2020 who received NACT (i.e., received chemotherapy at least 30 days prior to surgery) and had sufficient data to be used in model development (Additional file 1 Fig. A1), randomly splitting them into a 70% training set and a 30% validation set. The study also employed data from patients enrolled in the Chicago Multiethnic Epidemiologic Breast Cancer Cohort (ChiMEC), where patients with breast cancer diagnosed or treated at the University of Chicago Hospitals were enrolled at the high-risk clinic since 1992 and the breast center since 2008, with most of them coming from the Chicago metropolitan area [35, 36]. The clinical, pathological, and treatment data of ChiMEC patients were collected via electronic medical records following the same standards and protocols as the NCDB. Within ChiMEC, we identified patients diagnosed with invasive non-metastatic breast cancer who received NACT from 2010 to 2020 as an external testing set of the model, and also identified patients who did not receive chemotherapy and

only received hormone therapy to serve as a comparison group with the patients who received NACT.

Feature selection and model development

The prediction outcome, pCR, is defined as the absence of invasive cancer in both the breast and axillary nodes, irrespective of in situ carcinoma (ypT0/Tis ypN0). Based on existing literature and data availability, the basic clinicopathological features selected for model development were age at diagnosis, clinical T and N stages, histology types, tumor grades, comorbidity index [37] and four subtypes based on hormone receptor status (HR) and the amplification of HER2: HR+/HER2- (ER+and/or PR+, HER2-), HR+/HER2+ (ER+and/or PR+, HER2+), HR-/HER2+ (ER-, PR-, HER2+) and TNBC (triple negative breast cancer; ER-, PR-, HER2-). Besides these features, the study also included quantitative biomarkers ER%, PR%, HER2 immunohistochemistry (IHC) categories, HER2 to Chromosome 17 FISH (HER2/CEP17) ratios and Ki-67 scores. Socioeconomic features including insurance type of the patient, facility type and facility location of the institution were included in the sensitivity analysis.

Prediction models of pCR were developed using both logistic regression and machine learning. The machine learning algorithm employed is the SuperLearner, which uses cross-validation to form an ensemble of multiple candidate machine learning models that can optimize the final performance [38, 39]. The candidate machine learning models included: the mean predictor, logistic regression, Lasso regressions with all two-way interactions, elastic net regularization ('glmnet'), Bayesian generalized linear regression ('bayesglm'), Multivariate Adaptive Regression Splines ('earth'), Random Forest ('ranger', 'caret'), K-Nearest Neighbors ('knn'), and Gradient Boosted Decision Trees ('XGBoost'), with different hyper-parameter settings respectively. The model was first developed through 10-fold cross validation in the training set, and later evaluated in the validation and external testing sets.

Statistical analysis

The model's discrimination capacity was measured by AUC, with their 95% confidence intervals (CIs) computed with 2000 stratified bootstrap replicates [40] and compared by the DeLong's method [41]. Model calibration was illustrated through calibration graphs and measured using the Brier score [42], the Integrated Calibration Index (ICI) [43] and the intercept and slope of the calibration curve after locally estimated scatterplot smoothing. Decision curve analysis (DCA) was used to estimate the net reduction in intervention when applying the model in clinical decision-making [44]. Different cut-off thresholds were evaluated with the corresponding specificity and sensitivity of the model computed, enabling the selection of an optimal threshold to be used in practice. Missing data in the quantitative biomarkers were handled using multiple imputation by chained equations (MICE), implemented through the 'mice' package in R [45]. Missing values were imputed by performing regression imputation in a stepwise manner, where each missing variable is modeled as a function of the other variables. Imputation rules were established in the training set and subsequently applied to the validation and testing sets to prevent data leakage. This approach ensured the model's robustness and could accommodate missingness during future implementation, enhancing the model's accessibility [46]. Kaplan-Meier graphs and Cox proportional hazards models were used to examine the overall survival and recurrence-free survival of patients, as well as estimating the adjusted Hazard Ratios (aHRs). *P*-values were 2-sided with significance level of 5%. Statistical analyses were conducted using the R Statistical Software (v4.3.1; R Core Team 2023) and the STATA18 software (StataCorp, College Station, TX).

Results

We identified 56,209 patients with breast cancer who underwent NACT in the NCDB, and approximately 34% of them achieved pCR (Table 1). Patients with HR+/HER2- breast cancer had the lowest pCR rate (14.7%), significantly lower than that of HR+/HER2+ (40.0%), HR-/HER2+ (65.1%) and TNBC (38.8%) patients. Among racial/ethnic groups, Non-Hispanic Black ("Black") patients reported the lowest pCR rate (32.5%). In ChiMEC, the study identified 584 patients with breast cancer who received NACT (pCR rate: 33.4%) as the external testing set, where patients with HR+/HER2- breast cancer also had the lowest pCR rate (20.1%) compared to the other subtypes (Additional file 1 Table A1).

Model comparison

Using the basic clinicopathological features, the logistic regression model achieved an AUC of 0.739 (95% CI: 0.731–0.747), while the machine learning model had a slightly better AUC of 0.746 (95% CI: 0.738–0.753). After incorporating the quantitative biomarkers as predictors, the model's discrimination performance significantly improved, with an AUC of 0.781 (95% CI: 0.774–0.788) for logistic regression and an AUC of 0.785 (95% CI: 0.778–0.792) for the machine learning model (Table 2). All the models exhibited robust calibration, with their calibration curves' slopes close to 1 and intercepts close to 0, along with low Brier scores and low ICIs (Additional file 1 Table A2). The machine learning model, which integrates both basic and quantitative biomarkers, was chosen as the final model.

Table 1 Baseline characteristics of patients receiving NACT in the training and validation sets in NCDB (2018–2020)

Features No. (%)		Training Set (*n* = 39479)	Validation Set (*n* = 16730)	*P*[a]
pCR	No	26,003 (65.9)	11,067 (66.2)	0.51
	Yes	13,476 (34.1)	5663 (33.8)	
Age at Diagnosis, mean (SD)		54.5 (12.7)	54.5 (12.8)	0.42
Race / Ethnicity	Non-Hispanic White	26,128 (66.2)	11,029 (65.9)	0.77
	Non-Hispanic Black	6383 (16.2)	2675 (16.0)	
	Hispanic	3960 (10.0)	1720 (10.3)	
	Asian	2159 (5.5)	946 (5.7)	
	Other	849 (2.2)	360 (2.2)	
Subtype	HR+/HER2-	12,817 (32.5)	5481 (32.8)	0.71
	HR+/HER2+	9540 (24.2)	4043 (24.2)	
	HR-/HER2+	4153 (10.5)	1787 (10.7)	
	TNBC	12,969 (32.9)	5419 (32.4)	
Clinical T-Stage	I	9223 (23.4)	3909 (23.4)	0.10
	II	21,574 (54.6)	8970 (53.6)	
	III	5765 (14.6)	2556 (15.3)	
	IV	2917 (7.4)	1295 (7.7)	
Clinical N-Stage	0	24,096 (61.0)	10,069 (60.2)	0.059
	≥ 1	15,383 (39.0)	6661 (39.8)	
Tumor Grade	1	1694 (4.3)	734 (4.4)	0.33
	2	13,444 (34.1)	5598 (33.5)	
	3	24,341 (61.7)	10,398 (62.2)	
Histological Type	Ductal	35,088 (88.9)	14,907 (89.1)	0.91
	Lobular	1772 (4.5)	754 (4.5)	
	Ductal and Lobular	654 (1.7)	273 (1.6)	
	Mucinous	138 (0.3)	55 (0.3)	
	Papillary	22 (0.1)	11 (0.1)	
	Inflammatory	1067 (2.7)	448 (2.7)	
	Metaplasia	373 (0.9)	144 (0.9)	
	Others	365 (0.9)	138 (0.8)	
Charlson/Deyo Score	0	33,582 (85.1)	14,132 (84.5)	0.18
	1	4410 (11.2)	1930 (11.5)	
	≥ 2	1487 (3.8)	668 (4.0)	
ER Status	Negative	18,357 (46.5)	7729 (46.2)	0.51
	Positive	21,122 (53.5)	9001 (53.8)	
ER (% Positivity)	median (IQR)	0 (0, 90)	0 (0, 90)	0.69
	missing (%)	4805 (12.2)	2110 (12.6)	
PR Status	Negative	22,568 (57.2)	9512 (56.9)	0.50
	Positive	16,911 (42.8)	7218 (43.1)	
PR (% Positivity)	median (IQR)	0 (0, 45)	0 (0, 41)	0.93
	missing (%)	570 (1.4)	252 (1.5)	
HER2 Status	Negative	25,786 (65.3)	10,900 (65.2)	0.71
	Positive	13,693 (34.7)	5830 (34.8)	
HER2 IHC Categories	0	10,301 (26.1)	4365 (26.1)	0.85
	1+	7988 (20.2)	3419 (20.4)	
	2+	6886 (17.4)	2930 (17.5)	
	3+	10,429 (26.4)	4403 (26.3)	
	missing	3875 (9.8)	1613 (9.6)	
HER/CEP17 Ratio	median (IQR)	1.4 (1.1, 2.5)	1.4 (1.1, 2.5)	0.33
	missing (%) [b]	27,201 (68.9)	11,542 (69.0)	

Table 1 (continued)

Features No. (%)		Training Set (*n* = 39479)	Validation Set (*n* = 16730)	*P*[a]
Ki-67 score	median (IQR)	46 (25, 72)	45 (25, 72)	0.97
	missing (%)	19,210 (48.7)	8144 (48.7)	

Abbreviations pCR, pathologic complete response; HR, hormone receptor; HER2, human epidermal growth factor receptor 2; TNBC, triple-negative breast cancer; SD, standard deviation; IQR, interquartile range; ER, estrogen receptor; PR, progesterone receptor; IHC, immunohistochemistry; HER/CEP17, HER2 to Chromosome 17 FISH

[a] *P*-values comparing the 70% training set and 30% validation set were estimated using t-tests for age at diagnosis; Wilcoxon Rank-Sum tests for clinical T-stage, tumor grade, Charlson/Deyo score, ER%, PR%, HER2 IHC categories, HER/CEP17 ratio and Ki-67 score; and χ2 tests for the other categorical variables

[b] In clinical practice, the HER2/CEP17 ratio is typically assessed only in tumors that are scored as HER2 IHC 2+, resulting in a considerable amount of missing here

Table 2 Discrimination of the logistic regression and machine learning models with and without quantitative features

AUC (95% CI)	Logistic Regression [a]		Machine Learning	
	Basic Model [b]	Quantitative Model [c]	Basic Model	Quantitative Model
Overall[d] (*n* = 16730)	0.739 (0.731–0.747)	0.781 (0.774–0.788)	0.746 (0.738–0.753)	0.785 (0.778–0.792)
HR+/HER2- (*n* = 5481)	0.749 (0.732–0.767)	0.811 (0.795–0.827)	0.756 (0.739–0.773)	0.817 (0.802–0.832)
HR+/HER2+ (*n* = 4043)	0.612 (0.595–0.630)	0.744 (0.729–0.760)	0.623 (0.605–0.640)	0.751 (0.736–0.766)
HR-/HER2+ (*n* = 1787)	0.558 (0.530–0.586)	0.616 (0.588–0.644)	0.603 (0.576–0.631)	0.640 (0.613–0.668)
TNBC (*n* = 5419)	0.649 (0.634–0.663)	0.654 (0.639–0.668)	0.647 (0.632–0.662)	0.654 (0.639–0.669)

Abbreviations AUC, area under the receiver operating characteristic curve; CI, confidence interval; HR, hormone receptor; HER2, human epidermal growth factor receptor 2; TNBC, triple-negative breast cancer

[a] Details of the logistic regression models can be found in eTable 13

[b] The basic models included the basic features (i.e. age at diagnosis, clinical T and N stages, histology types, tumor grades and comorbidity index)

[c] The quantitative models included both the basic features (i.e. age at diagnosis, clinical T and N stages, histology types, tumor grades and comorbidity index) and the quantitative features (i.e. ER%, PR%, HER2 IHC categories, HER2/CEP17 ratios and Ki-67 scores

[d] The AUC of each model was estimated among the 30% hold-out validation set overall and within each breast cancer subtype, the 95% CIs of the AUCs were calculated using the 'pROC' package in R

Final model's performance among different subgroups

The final model's performance varied across different breast cancer subtypes (Table 2), performing the best for patients with HR+/HER2- diseases (AUC: 0.817, 95% CI: 0.802–0.832). Incorporating the quantitative biomarkers as predictors significantly improved the model's discrimination performance in all subtypes except for patients with TNBC. The permutation feature importance graphs (Fig. 1) illustrated that these quantitative features were among the most important to the model's performance within their respective subtypes (e.g., ER% and PR% for the HR+subtypes, HER2 IHC categories and HER2/CEP17 ratio for the HER2+subtypes).

In assessing the final model's performance across different racial/ethnic groups, we found that it displayed consistent discriminatory ability among HR-/HER2+and TNBC patients, with no significant racial/ethnic disparity observed (Fig. 2). For the HR+/HER2- subgroup, the AUC for Black patients was approximately 5% lower than the other racial/ethnic groups, although not reaching statistical significance (Additional file 1 Table A3). Notably, in the HR+/HER2+subgroup, the AUC for Black patients was significantly lower than that for other racial/ethnic groups, with about 10% difference (0.699 vs. 0.765). Nevertheless, the model showed great calibration across all the racial/ethnic groups (Fig. 3).

Sensitivity analysis

To address the final model's varying performance across subtypes, we examined whether subtype-specific models had improved performance. We found their performance closely mirrored that of the final model (Additional file 1 Table A4). To test the validity of the imputation methods, we also fitted subtype-specific models taking a complete case analysis approach, i.e., only including patients without any missing values. Compared with the complete case analysis, the final model fitted in the imputed dataset performed similarly in the HR+/HER2- and TNBC patients, although losing some prediction power in the two HER2+subtypes, likely due to the high proportion of missing in HER2/CEP17 ratio (Additional file 1 Table A5).

To account for disparities in treatment, we trained and validated a model within patients who received NACT for 16–28 weeks. This range (10-90th percentile) excluded patients with unusually short or long treatment durations, indicating possible non-adherence or treatment delays. Nevertheless, the corresponding model showed comparable performance with the final model both in general as well as among different patient groups (Additional file 1 Table A6). To address the impact of socioeconomic determinants, we also trained and validated a model including both the basic and quantitative clinicopathological features, as well as racial/ethnic group, insurance type, facility type and facility location as

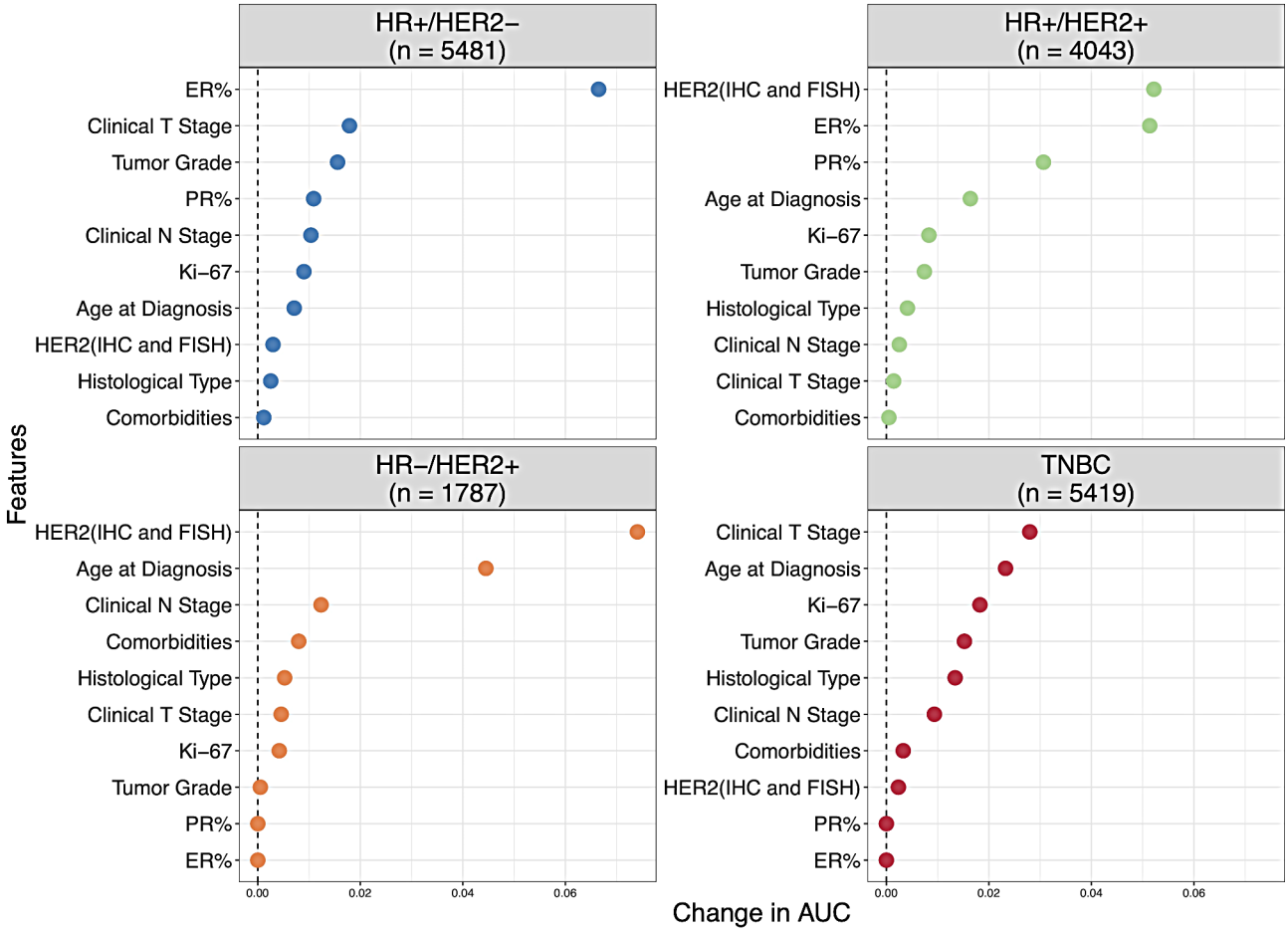

Fig. 1 Permutation feature importance of the final model in different breast cancer subtypes

predictors, yet there was negligible improvement in the model's AUCs (Additional file 1 Table A7).

Potential in clinical decision-making

To evaluate the model's utility in clinical decision-making, we assessed its potential to identify patients with HR+/HER2- breast cancer who might be suitable candidates to forgo chemotherapy. To accommodate different clinical judgments on the optimal prediction threshold to choose, we have proposed a range of reasonable thresholds from 3 to 15% (Table 3).

Within the reasonable range of thresholds, DCA showed that the quantitative machine learning model provided the most net benefit compared with the other fitted models (Fig. 4). Here we chose 7% as the suggested threshold since it offered approximately 90% sensitivity. With 7% threshold, the model achieved a net reduction in intervention of 23.1% among HR+/HER2- patients. Over 40% of the HR+/HER2- patients would have a predicted pCR probability lower than 7%, while only 3.6% or fewer for patients in the other subtypes (Additional file 1 Table A8).

To further examine the model's potential in clinical decision-making compared with existing tools, we identified a subset of 1266 h+/HER2- patients in the NCDB with Oncotype Dx score available (Additional file 1 Table A9). In this subset, we found that the final model, developed using the entire dataset, demonstrated more robust discrimination (AUC: 0.735) compared to the model developed within the subset, even when incorporating Oncotype Dx and all other predictors (AUC: 0.684) (Additional file 1 Table A10). Integrating Oncotype Dx score as an additional predictor with the final model only marginally improved the model's performance (AUC: 0.736). This is probably because the predicted values from the final model were strongly correlated with the Oncotype Dx scores (Pearson's correlation coefficient: 0.63, $P<0.001$).

The final model demonstrated comparable discrimination performance in the external testing set, with an overall AUC of 0.711 (95% CI: 0.668–0.753). Similarly, the model performed best for the HR+/HER2- subgroup, achieving an AUC of 0.810 (95% CI: 0.742–0.878). With the selected threshold of 7%, the model achieved a

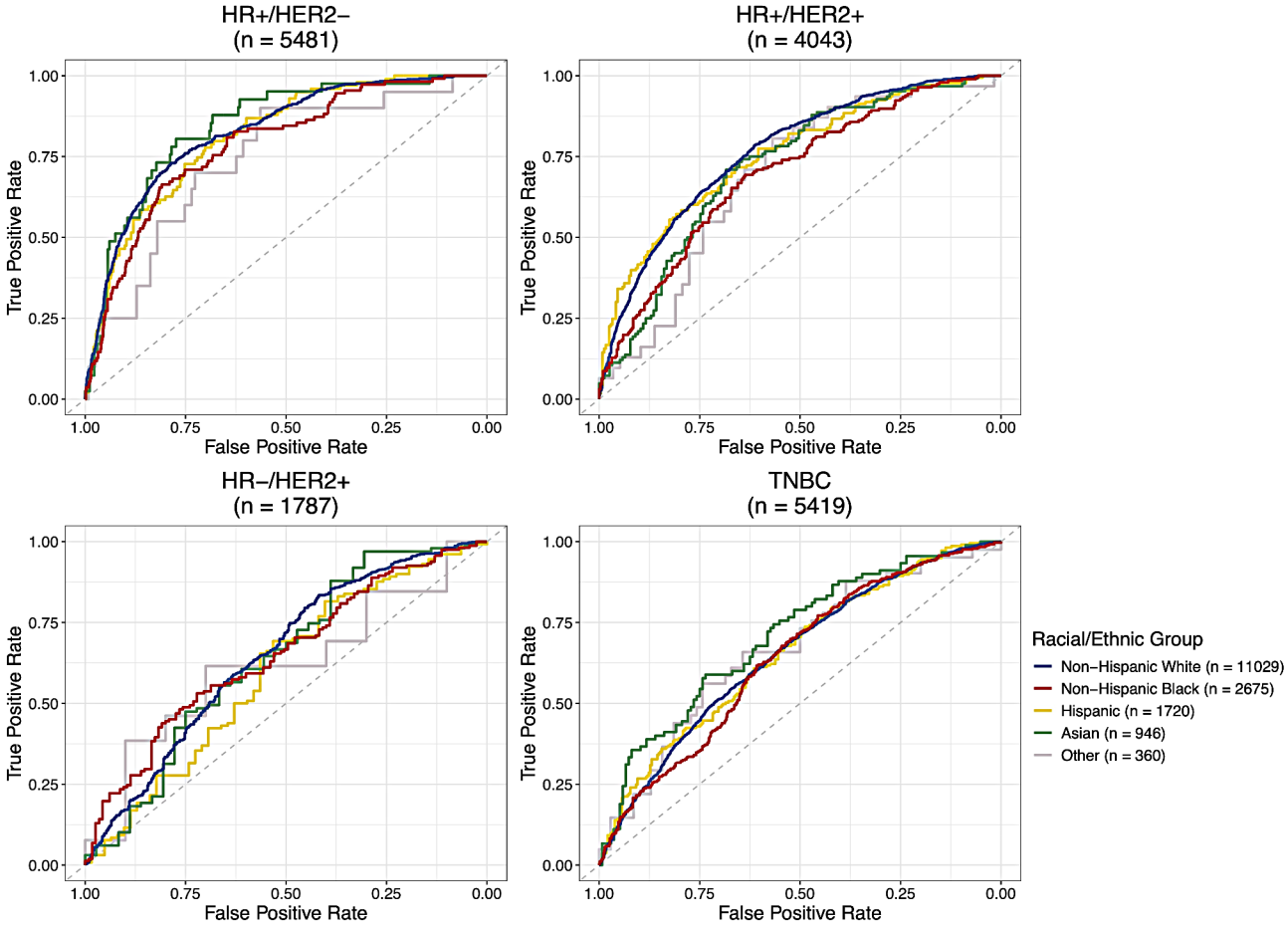

Fig. 2 Receiver operating characteristic curves of final model across different racial/ethnic groups and subtypes in validation set

sensitivity of 92.3% and specificity of 46.7% among the HR+/HER2- patients, selecting 38.7% of them who might be eligible to spare chemotherapy (Additional file 1 Table A11). Within the HR+/HER2- patients in ChiMEC, we identified the 70 patients who had a predicted pCR probability lower than 7% and received NACT, and compared to the 408 patients who did not receive chemotherapy (i.e., only underwent hormone therapy) (Additional file 1 Table A12). No statistically significant difference were found in overall survival and recurrence-free survival between these chemotherapy recipients and non-recipients (Additional file 1 Fig. A2), with their corresponding aHRs being 1.05 (95% CI: 0.43–2.54) and 1.10 (95% CI: 0.52–2.31) adjusting for age at diagnosis, race/ethnicity, clinical T and N stages, grades and comorbidities.

Discussion

In this study, we developed and validated a prediction model of pCR following NACT using data from 56,209 patients in the NCDB (2018–2020). The final machine learning model showed strong discrimination and calibration performance in the validation set, achieving an AUC of 0.785 overall and an AUC of 0.817 for HR+/HER2- subtypes.

We observed a significant improvement in the model's discrimination performance upon integrating the quantitative clinicopathological features. This improvement was especially pronounced in the HR+and HER2+subgroups, which exhibited a broad spectrum of values for features like ER% positivity, PR% positivity, and HER2/CEP17 ratios. Previous studies have suggested that using quantitative ER% and PR% values, rather than binary positive/negative categories, might provide additional prognostic value in survival and predictive value of pCR [12, 47]. Furthermore, the specific cutoff percentage used to categorize tumors as ER/PR-positive has still been a topic under debate [48, 49]. Therefore, it is sound to treat ER and PR as continuous features in the model.

To address missing values in these quantitative features, the model incorporated a rigorous imputation method. Although the considerable amount of missing data for HER2/CEP17 ratio and Ki-67 diluted their predictive power, sensitivity analysis suggested that this loss was not substantial. Moreover, in recognition that these quantitative features might be unavailable in real-world

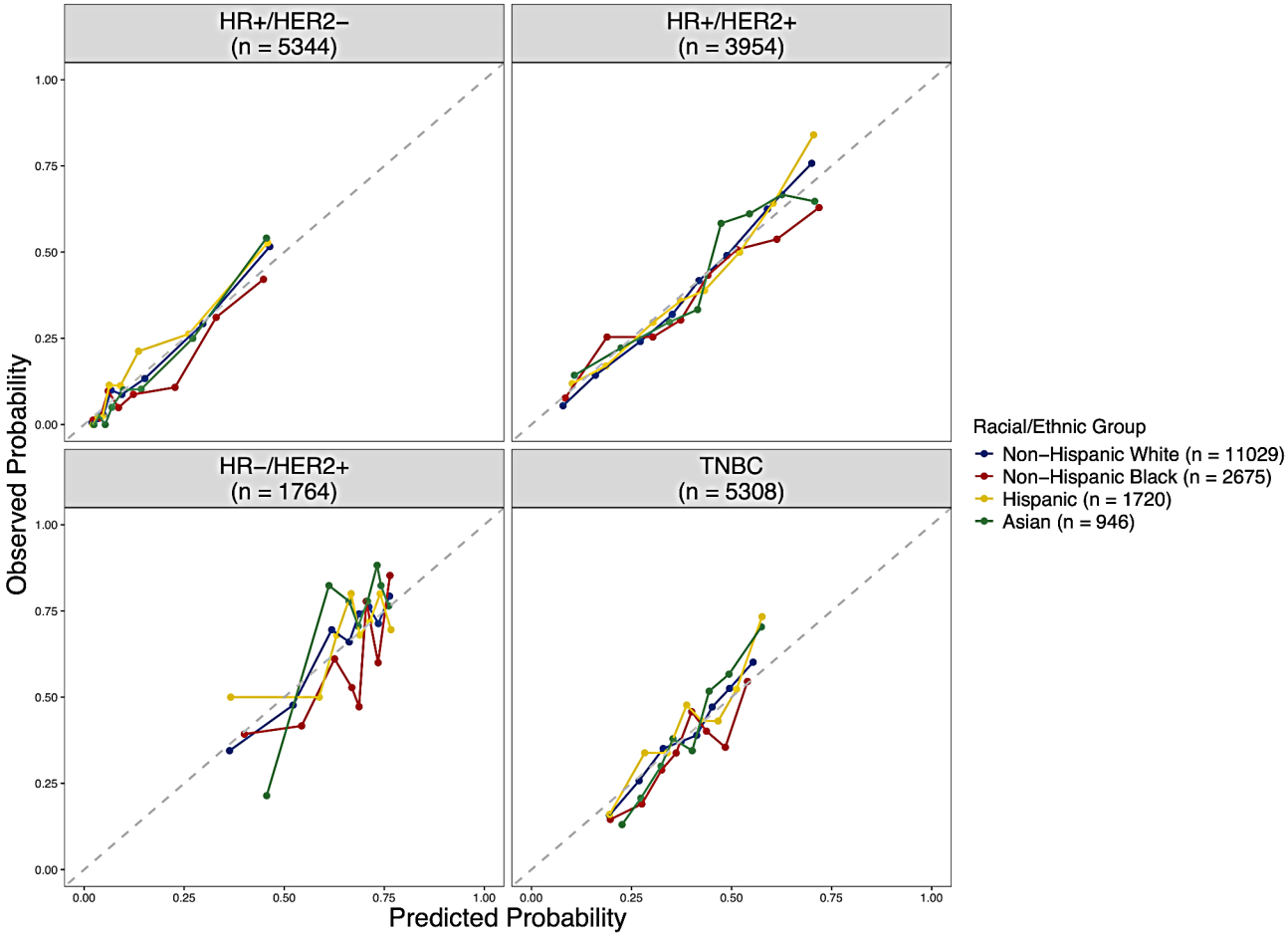

Fig. 3 Calibration plots of the final model across different racial/ethnic groups and subtypes in validation set. * The calibration plots of patients from "Other" racial/ethnic groups were not shown because of their limited sample size

Table 3 Performance Metrics of the final model for different prediction thresholds among HR+/HER2- patients

Threshold [a]	Sensitivity (%)	Specificity (%)	PPV (%)	NPV (%)	Patients waiving chemotherapy (%) [b]	Net reduction in intervention (%) [c]
3%	99.1	14.5	16.7	99.0	12.5	8.1
5%	96.9	34.1	20.2	98.5	29.5	20.4
7%	90.5	48.8	23.4	96.7	43.0	23.1
10%	82.3	62.8	27.6	95.4	56.2	30.1
15%	73.7	74.7	33.4	94.3	67.5	41.8

Abbreviations HR, hormone receptor; HER2, human epidermal growth factor receptor 2; PPV, positive predictive value; NPV, negative predictive value

[a]Having a predicted pCR probability lower than the threshold is predicted non-pCR, whereas greater than the threshold would be predicted pCR. The prediction model used here is the quantitative machine learning model

[b]Treatment options other than chemotherapy might be considered if the patients' predicted pCR probability estimated by the quantitative machine learning prediction model is less than the selected threshold

[c]Net reduction in intervention is calculated by $\text{Specificity} \times (1 - \text{Prevalence}) - (1 - \text{Sensitivity}) \times \text{Prevalence} \times \frac{Threshold}{1-Threshold}$ based on Decision Curve Analysis, where Prevalence is the rate of pCR among HR+/HER2- patients (14.7%)

data as well, allowing missing values enable the model to represent a wider patient population and can be applied in low-resource settings where IHC measurements may be challenging to perform [50], or when biomarkers like Ki-67 are not available in practice [51].

A key motivation for this study was to apply the model to facilitate clinical decision-making. Notably, the model performed best for the HR+/HER2- subtype in both the validation set (AUC: 0.817) and the testing set (AUC: 0.810). Given that this subgroup also had the lowest rate of achieving pCR (14.7%) and had emerging alternative treatment options aside from chemotherapy [52–55], we assessed the potential of applying the model to identify HR+/HER2- patients who might not benefit significantly

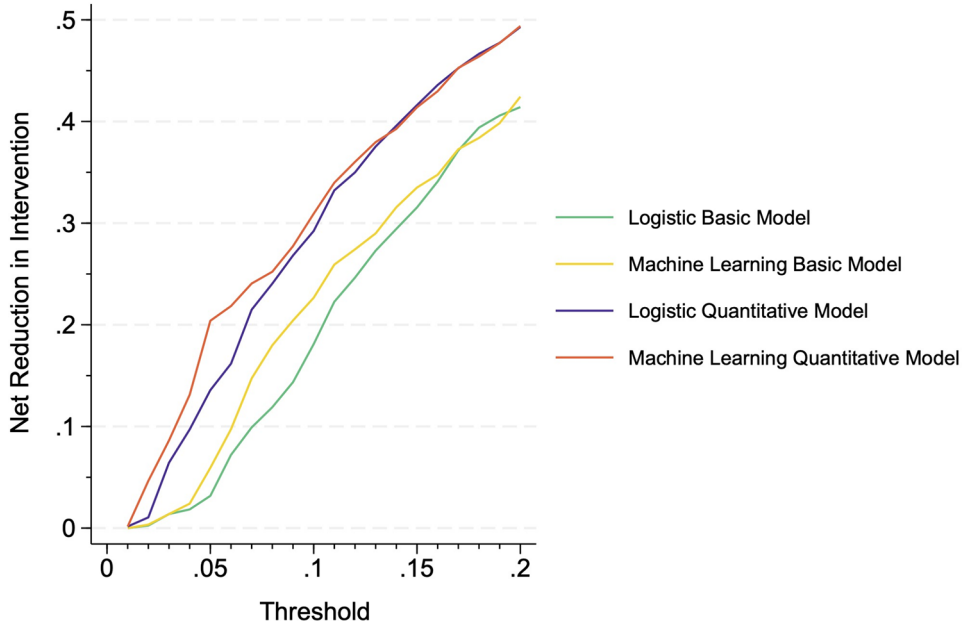

Fig. 4 Net reduction in intervention among HR+/HER2- patients in the validation set using decision curve analysis

from chemotherapy. Setting the prediction threshold at 7%, the model can achieve a sensitivity of 90.5% and a specificity of 48.8%. DCA results showed that the quantitative machine learning model had the highest net reduction in intervention potential compared to logistic regression models, achieving a 23.1% net reduction in chemotherapy rate with the selected 7% threshold. In other words, 23.1% of the chemotherapy can be spared without overlooking any HR+/HER2- patient who could have achieved pCR.

Furthermore, we found that HR+/HER2- patients with a low predicted pCR probability, as determined by our model, had very similar survival outcomes regardless of receiving chemotherapy or not. This observation indicated a potential lack of meaningful long-term survival improvement from chemotherapy for this subset of patients. In the adjuvant setting, gene expression-based assays like Oncotype Dx, MammaPrint and PAM50 have been used to identify HR+/HER2- patients who could avoid chemotherapy [53, 56]. What sets our model apart is that it only utilized common clinicopathological features, enhancing the model's accessibility in practice, yet still demonstrating superior performance compared with using Oncotype Dx to predict pCR (AUC: 0.767) [57].

Although the model performed particularly well among the HR+subtypes, it did not perform equally for the different racial/ethnic groups, with notably lower AUCs among Black patients. This performance gap remained after controlling for treatment duration differences and integrating additional socioeconomic factors into the model. To improve the model's performance across diverse patient populations, it may be beneficial to include more granular biomarkers that can capture the diseases' heterogeneity more effectively, including gene expression signatures like Oncotype Dx, HER2DX or other genomic and transcriptomic features [57–60]. Our sensitivity analysis, conducted on a subset of NCDB patients with Oncotype Dx scores, highlighted the model's potential as a baseline framework for further research. Developed on a large dataset, the model can robustly capture the predictive power of the clinicopathological features. Notably, datasets with more granular features (gene expression and imaging data), while potentially enhancing predictive performance, are often limited in size and at risk of overfitting. Thus, integrating our robust baseline model with these nuanced, yet small, datasets through data fusion offers a promising approach to optimize model performance in future studies. The ability of machine learning models to handle complex non-linear relationships and high-dimensional data also offers great research potential.

The major limitation of the study is the lack of granular quantitative features like gene expression signatures within the NCDB. Nevertheless, as previously discussed, our model can potentially serve as a robust baseline to incorporate these features in subsequent studies. Another limitation is our reliance on a retrospectively matched control group to simulate the model's utility in clinical decision-making. Prospective validation in a randomized clinical trial setting is needed to confirm the model's efficacy in identifying HR+/HER2- patients who could spare chemotherapy without compromising long-term survival benefits.

Conclusions

Utilizing a large, contemporary sample from the NCDB, this study developed a machine learning model to predict pCR following NACT that showed robust discrimination and calibration capabilities. The model performed best among the HR+/HER2- subgroup, and may potentially facilitate clinical decision-making through identifying HR+/HER2- patients unlikely to achieve pCR who can consider alternative treatment strategies over chemotherapy. The model can be implemented in diverse settings and can serve as a robust baseline model for future research.

Abbreviations

NACT	Neoadjuvant chemotherapy
pCR	Pathologic complete response
HR	Hormone receptor
ER	Estrogen receptor
PR	Progesterone receptor
HER2	Human epidermal growth factor receptor 2
IHC	Immunohistochemistry
HER/CEP17	HER2 to Chromosome 17
TNBC	Triple-negative breast cancer
NCDB	NATIONAL Cancer Data Base
ChiMEC	Chicago Multiethnic Epidemiologic Breast Cancer Cohort
AUC	Area under the receiver operating characteristic curve
CI	Confidence interval
SD	Standard deviation
IQR	Interquartile range
ICI	Integrated Calibration Index
aHR	Adjusted Hazard Ratio
DCA	Decision curve analysis

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s13058-024-01905-7.

Supplementary Material 1: Additional file 1.docx (including Table A1-A13, Fig. A1 and Fig. A2)

Author contributions

FZ: Conceptualization; Formal Analysis; Software; Writing – original draft; Writing – review and editing; EP: Conceptualization; Software; Writing – review and editing; JM: Software; Writing – review and editing; FH: Conceptualization; Writing – review and editing; OIO: Conceptualization; Funding Acquisition; Supervision; Writing – review and editing; DH: Conceptualization; Funding Acquisition; Supervision; Formal Analysis; Writing – review and editing. All authors read and approved the final manuscript.

Funding

This research was partially supported by the Breast Cancer Research Foundation (BCRF-23-071), National Cancer Institute (P20CA233307), the Department of Defense (W81XWH2210791), and Susan G Komen for the Cure (TREND21675016). The funding source played no role in the conceptualization, design, data collection, analysis, decision to publish, or preparation of the manuscript.

Data availability

The final machine learning model is accessible online: https://huolab.cri.uchicago.edu/sample-apps/pcrmodel/. The details of the logistic regression models are presented in Additional file 1 Table A13. De-identified data and the algorithms of the model will be made available to interested researchers upon reasonable request to the corresponding author, Dezheng Huo (dhuo@bsd.uchicago.edu). All requests must comply with the guidelines of the Institutional Review Board at the University of Chicago.

Declarations

Ethics approval and consent to participate

The Institutional Review Board at the University of Chicago granted a waiver status for the use of NCDB data in this study because no protected health information was reviewed, and the analysis was retrospective using de-identified data. The NCDB is a joint project of the Commission on Cancer of the American College of Surgeons and the American Cancer Society. The American College of Surgeons and the Commission on Cancer have not verified and are not responsible for the analytic or statistical methodology employed, or the conclusions drawn from these data by the investigators. The study protocol in ChiMEC was approved by the Institutional Review Board at the University of Chicago, and all participants provided their written informed consent.

Consent for publication

Not applicable.

Competing interests

OIO has disclosed financial relationships with CancerIQ, HealthWell Solutions, Tempus; research funding from Ayala Pharmaceuticals, Cepheid, Color Genomics, Novartis, and Roche/Genentech. The other authors, FZ, EP, JM, FH, and DH, declare no financial or non-financial competing interests.

Author details

[1]Key Laboratory of Carcinogenesis and Translational Research (Ministry of Education), Laboratory of Molecular Oncology, Peking University Cancer Hospital & Institute, Beijing, China
[2]Department of Public Health Sciences, University of Chicago, Chicago, IL, USA
[3]Department of Medicine, Section of Hematology and Oncology, University of Chicago, Chicago, IL, USA

Received: 20 May 2024 / Accepted: 16 October 2024
Published online: 29 October 2024

References

1. Siegel RL, Miller KD, Fuchs HE, Jemal A. Cancer statistics, 2022. CA Cancer J Clin. 2022;72(1):7–33.
2. DeSantis CE, Ma J, Gaudet MM, Newman LA, Miller KD, Goding Sauer A, Jemal A, Siegel RL. Breast cancer statistics, 2019. CA Cancer J Clin. 2019;69(6):438–51.
3. Hayes DF, Schott AF. Neoadjuvant chemotherapy: what are the benefits for the patient and for the Investigator? J Natl Cancer Inst Monogr. 2015;2015(51):36–9.
4. Korde LA, Somerfield MR, Carey LA, Crews JR, Denduluri N, Hwang ES, Khan SA, Loibl S, Morris EA, Perez A, et al. Neoadjuvant chemotherapy, endocrine therapy, and targeted therapy for breast Cancer: ASCO Guideline. J Clin Oncol. 2021;39(13):1485–505.
5. Early Breast Cancer Trialists', Collaborative G. Long-term outcomes for neoadjuvant versus adjuvant chemotherapy in early breast cancer: meta-analysis of individual patient data from ten randomised trials. Lancet Oncol. 2018;19(1):27–39.
6. Matuschek C, Jazmati D, Bolke E, Tamaskovics B, Corradini S, Budach W, Krug D, Mohrmann S, Ruckhaberle E, Fehm T et al. Post-neoadjuvant treatment strategies in breast Cancer. Cancers (Basel) 2022, 14(5).
7. Masuda N, Lee SJ, Ohtani S, Im YH, Lee ES, Yokota I, Kuroi K, Im SA, Park BW, Kim SB, et al. Adjuvant capecitabine for breast Cancer after preoperative chemotherapy. N Engl J Med. 2017;376(22):2147–59.
8. von Minckwitz G, Huang CS, Mano MS, Loibl S, Mamounas EP, Untch M, Wolmark N, Rastogi P, Schneeweiss A, Redondo A, et al. Trastuzumab Emtansine for residual invasive HER2-Positive breast Cancer. N Engl J Med. 2019;380(7):617–28.
9. Cortazar P, Zhang L, Untch M, Mehta K, Costantino JP, Wolmark N, Bonnefoi H, Cameron D, Gianni L, Valagussa P, et al. Pathological complete response and long-term clinical benefit in breast cancer: the CTNeoBC pooled analysis. Lancet. 2014;384(9938):164–72.
10. Schmid P, Cortes J, Dent R, Pusztai L, McArthur H, Kummel S, Bergh J, Denkert C, Park YH, Hui R, et al. Event-free survival with Pembrolizumab in Early Triple-negative breast Cancer. N Engl J Med. 2022;386(6):556–67.

11. 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.
12. Zhao F, Miyashita M, Hattori M, Yoshimatsu T, Howard F, Kaneva K, Jones R, Bell JSK, Fleming GF, Jaskowiak N, 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.
13. Kayl AE, Meyers CA. Side-effects of chemotherapy and quality of life in ovarian and breast cancer patients. Curr Opin Obstet Gynecol. 2006;18(1):24–8.
14. Jim HS, Phillips KM, Chait S, Faul LA, Popa MA, Lee YH, Hussin MG, Jacobsen PB, Small BJ. Meta-analysis of cognitive functioning in breast cancer survivors previously treated with standard-dose chemotherapy. J Clin Oncol. 2012;30(29):3578–87.
15. Bertsimas D, Wiberg H. Machine learning in Oncology: methods, applications, and challenges. JCO Clin Cancer Inf. 2020;4:885–94.
16. Keskin S, Muslumanoglu M, Saip P, Karanlik H, Guveli M, Pehlivan E, Aydogan F, Eralp Y, Aydiner A, Yavuz E, et al. Clinical and pathological features of breast cancer associated with the pathological complete response to anthracycline-based neoadjuvant chemotherapy. Oncology. 2011;81(1):30–8.
17. Kantor O, Sipsy LM, Yao K, James TA. A predictive model for Axillary Node Pathologic Complete response after neoadjuvant chemotherapy for breast Cancer. Ann Surg Oncol. 2018;25(5):1304–11.
18. Goorts B, van Nijnatten TJ, de Munck L, Moossdorff M, Heuts EM, de Boer M, Lobbes MB, Smidt ML. Clinical tumor stage is the most important predictor of pathological complete response rate after neoadjuvant chemotherapy in breast cancer patients. Breast Cancer Res Treat. 2017;163(1):83–91.
19. Greenwell K, Hussain L, Lee D, Bramlage M, Bills G, Mehta A, Jackson A, Wexelman B. Complete pathologic response rate to neoadjuvant chemotherapy increases with increasing HER2/CEP17 ratio in HER2 overexpressing breast cancer: analysis of the National Cancer Database (NCDB). Breast Cancer Res Treat. 2020;181(2):249–54.
20. Dieci MV, Griguolo G, Bottosso M, Tsvetkova V, Giorgi CA, Vernaci G, Michieletto S, Angelini S, Marchet A, Tasca G, et al. Impact of estrogen receptor levels on outcome in non-metastatic triple negative breast cancer patients treated with neoadjuvant/adjuvant chemotherapy. NPJ Breast Cancer. 2021;7(1):101.
21. Landmann A, Farrugia DJ, Zhu L, Diego EJ, Johnson RR, Soran A, Dabbs DJ, Clark BZ, Puhalla SL, Jankowitz RC, et al. Low estrogen receptor (ER)-Positive breast Cancer and neoadjuvant systemic chemotherapy: is response similar to typical ER-Positive or ER-Negative disease? Am J Clin Pathol. 2018;150(1):34–42.
22. Tao M, Chen S, Zhang X, Zhou Q. Ki-67 labeling index is a predictive marker for a pathological complete response to neoadjuvant chemotherapy in breast cancer: a meta-analysis. Med (Baltim). 2017;96(51):e9384.
23. Peiffer DS, Zhao F, Chen N, Hahn OM, Nanda R, Olopade OI, Huo D, Howard FM. Clinicopathologic characteristics and prognosis of ERBB2-Low breast Cancer among patients in the National Cancer Database. JAMA Oncol. 2023;9(4):500–10.
24. Cain EH, Saha A, Harowicz MR, Marks JR, Marcom PK, Mazurowski MA. Multivariate machine learning models for prediction of pathologic response to neoadjuvant therapy in breast cancer using MRI features: a study using an independent validation set. Breast Cancer Res Treat. 2019;173(2):455–63.
25. Gass P, Lux MP, Rauh C, Hein A, Bani MR, Fiessler C, Hartmann A, Haberle L, Pretscher J, Erber R, et al. Prediction of pathological complete response and prognosis in patients with neoadjuvant treatment for triple-negative breast cancer. BMC Cancer. 2018;18(1):1051.
26. Qu YH, Zhu HT, Cao K, Li XT, Ye M, Sun YS. Prediction of pathological complete response to neoadjuvant chemotherapy in breast cancer using a deep learning (DL) method. Thorac Cancer. 2020;11(3):651–8.
27. Howard FM, He G, Peterson JR, Pfeiffer JR, Earnest T, Pearson AT, Abe H, Cole JA, Nanda R. Highly accurate response prediction in high-risk early breast cancer patients using a biophysical simulation platform. Breast Cancer Res Treat. 2022;196(1):57–66.
28. Ren Z, Pineda FD, Howard FM, Hill E, Szasz T, Safi R, Medved M, Nanda R, Yankeelov TE, Abe H, et al. Differences between Ipsilateral and Contralateral Early Parenchymal Enhancement kinetics Predict response of breast Cancer to Neoadjuvant Therapy. Acad Radiol. 2022;29(10):1469–79.
29. Ren Z, Pineda FD, Howard FM, Fan X, Nanda R, Abe H, Kulkarni K, Karczmar GS. Bilateral asymmetry of quantitative parenchymal kinetics at ultrafast DCE-MRI predict response to neoadjuvant chemotherapy in patients with HER2 + breast cancer. Magn Reson Imaging. 2023;104:9–15.
30. Basmadjian RB, Kong S, Boyne DJ, Jarada TN, Xu Y, Cheung WY, Lupichuk S, Quan ML, Brenner DR. Developing a prediction model for pathologic complete response following neoadjuvant chemotherapy in breast Cancer: a comparison of Model Building approaches. JCO Clin Cancer Inf. 2022;6:e2100055.
31. Kim JY, Jeon E, Kwon S, Jung H, Joo S, Park Y, Lee SK, Lee JE, Nam SJ, Cho EY, et al. Prediction of pathologic complete response to neoadjuvant chemotherapy using machine learning models in patients with breast cancer. Breast Cancer Res Treat. 2021;189(3):747–57.
32. Meti N, Saednia K, Lagree A, Tabbarah S, Mohebpour M, Kiss A, Lu FI, Slodkowska E, Gandhi S, Jerzak KJ, et al. Machine learning frameworks to Predict Neoadjuvant Chemotherapy response in breast Cancer using clinical and pathological features. JCO Clin Cancer Inf. 2021;5:66–80.
33. Jung JJ, Kim EK, Kang E, Kim JH, Kim SH, Suh KJ, Kim SM, Jang M, Yun B, Park SY, et al. Development and External Validation of a machine learning model to predict pathological complete response after neoadjuvant chemotherapy in breast Cancer. J Breast Cancer. 2023;26(4):353–62.
34. Mallin K, Browner A, Palis B, Gay G, McCabe R, Nogueira L, Yabroff R, Shulman L, Facktor M, Winchester DP, et al. Incident cases captured in the National Cancer Database compared with those in U.S. Population Based Central Cancer registries in 2012–2014. Ann Surg Oncol. 2019;26(6):1604–12.
35. Zhao F, Copley B, Niu Q, Liu F, Johnson JA, Sutton T, Khramtsova G, Sveen E, Yoshimatsu TF, Zheng Y, et al. Racial disparities in survival outcomes among breast cancer patients by molecular subtypes. Breast Cancer Res Treat. 2021;185(3):841–9.
36. Shang L, Hattori M, Fleming G, Jaskowiak N, Hedeker D, Olopade OI, Huo D. Impact of post-diagnosis weight change on survival outcomes in Black and White breast cancer patients. Breast Cancer Res. 2021;23(1):18.
37. Deyo RA, Cherkin DC, Ciol MA. Adapting a clinical comorbidity index for use with ICD-9-CM administrative databases. J Clin Epidemiol. 1992;45(6):613–9.
38. van der Laan MJ, Polley EC, Hubbard AE. Super learner. Stat Appl Genet Mol 2007, 6.
39. Super Learner. In Prediction [https://biostats.bepress.com/ucbbiostat/paper266]
40. Robin X, Turck N, Hainard A, Tiberti N, Lisacek F, Sanchez JC, Muller M. pROC: an open-source package for R and S + to analyze and compare ROC curves. BMC Bioinformatics. 2011;12:77.
41. DeLong ER, DeLong DM, Clarke-Pearson DL. Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach. Biometrics. 1988;44(3):837–45.
42. Rufibach K. Use of Brier score to assess binary predictions. J Clin Epidemiol. 2010;63(8):938–9. author reply 939.
43. Austin PC, Steyerberg EW. The Integrated Calibration Index (ICI) and related metrics for quantifying the calibration of logistic regression models. Stat Med. 2019;38(21):4051–65.
44. Vickers AJ, van Calster B, Steyerberg EW. A simple, step-by-step guide to interpreting decision curve analysis. Diagn Progn Res. 2019;3:18.
45. van Buuren S, Groothuis-Oudshoorn K. Mice: multivariate imputation by chained equations in R. J Stat Softw. 2011;45(3):1–67.
46. Sisk R, Sperrin M, Peek N, van Smeden M, Martin GP. Imputation and missing indicators for handling missing data in the development and deployment of clinical prediction models: a simulation study. Stat Methods Med Res. 2023;32(8):1461–77.
47. Ma H, Lu Y, Marchbanks PA, Folger SG, Strom BL, McDonald JA, Simon MS, Weiss LK, Malone KE, Burkman RT, et al. Quantitative measures of estrogen receptor expression in relation to breast cancer-specific mortality risk among white women and black women. Breast Cancer Res. 2013;15(5):R90.
48. Yi M, Huo L, Koenig KB, Mittendorf EA, Meric-Bernstam F, Kuerer HM, Bedrosian I, Buzdar AU, Symmans WF, Crow JR, et al. Which threshold for ER positivity? A retrospective study based on 9639 patients. Ann Oncol. 2014;25(5):1004–11.
49. Allison KH, Hammond MEH, Dowsett M, McKernin SE, Carey LA, Fitzgibbons PL, Hayes DF, Lakhani SR, Chavez-MacGregor M, Perlmutter J, et al. Estrogen and progesterone receptor testing in breast Cancer: ASCO/CAP Guideline Update. J Clin Oncol. 2020;38(12):1346–66.
50. Kimambo AH, Vuhahula EA, Mwakigonja AR, Ljung BM, Zhang L, Van Loon K, Ng DL. Evaluating estrogen receptor immunohistochemistry on cell blocks from breast Cancer patients in a low-resource setting. Arch Pathol Lab Med. 2021;145(7):834–41.
51. Brown J, Scardo S, Method M, Schlauch D, Misch A, Picard S, Hamilton E, Jones S, Burris H, Spigel D. A real-world retrospective study of the use of

Ki-67 testing and treatment patterns in patients with HR+, HER2- early breast cancer in the United States. BMC Cancer. 2022;22(1):502.
52. Harbeck N, Rastogi P, Martin M, Tolaney SM, Shao ZM, Fasching PA, Huang CS, Jaliffe GG, Tryakin A, Goetz MP, et al. Adjuvant abemaciclib combined with endocrine therapy for high-risk early breast cancer: updated efficacy and Ki-67 analysis from the monarchE study. Ann Oncol. 2021;32(12):1571–81.
53. Harbeck N, Burstein HJ, Hurvitz SA, Johnston S, Vidal GA. A look at current and potential treatment approaches for hormone receptor-positive, HER2-negative early breast cancer. Cancer. 2022;128(Suppl 11):2209–23.
54. Twelves C, Bartsch R, Ben-Baruch NE, Borstnar S, Dirix L, Tesarova P, Timcheva C, Zhukova L, Pivot X. The place of Chemotherapy in the Evolving Treatment Landscape for patients with HR-positive/HER2-negative MBC. Clin Breast Cancer. 2022;22(3):223–34.
55. Akhade A, Van Wambeke S, Gyawali B. CDK 4/6 inhibitors for adjuvant therapy in early breast cancer-Do we have a clear winner? *Ecancermedicalscience* 2022, 16:ed124.
56. Jacobson A. Benefits of Adjuvant Chemotherapy Differ by Menopausal Status in Women with HR+/HER2- early breast Cancer, 1–3 positive nodes, and a low recurrence score. Oncologist. 2022;27(Suppl 1):S15–6.
57. Freeman JQ, Shubeck S, Howard FM, Chen N, Nanda R, Huo D. Evaluation of multigene assays as predictors for response to neoadjuvant chemotherapy in early-stage breast cancer patients. NPJ Breast Cancer. 2023;9(1):33.
58. Sammut SJ, Crispin-Ortuzar M, Chin SF, Provenzano E, Bardwell HA, Ma W, Cope W, Dariush A, Dawson SJ, Abraham JE, et al. Multi-omic machine learning predictor of breast cancer therapy response. Nature. 2022;601(7894):623–9.
59. Prat A, Guarneri V, Pascual T, Braso-Maristany F, Sanfeliu E, Pare L, Schettini F, Martinez D, Jares P, Griguolo G, et al. Development and validation of the new HER2DX assay for predicting pathological response and survival outcome in early-stage HER2-positive breast cancer. EBioMedicine. 2022;75:103801.
60. Villacampa G, Tung NM, Pernas S, Pare L, Bueno-Muino C, Echavarria I, Lopez-Tarruella S, Roche-Molina M, Del Monte-Millan M, Marin-Aguilera M, et al. Association of HER2DX with pathological complete response and survival outcomes in HER2-positive breast cancer. Ann Oncol. 2023;34(9):783–95.

Publisher's note