Am J Transl Res 2026;18(4):2919-2934
www.ajtr.org /ISSN:1943-8141/AJTR0170312

Original Article

Identification of chemotherapy-related risk factors for lymphedema in breast cancer patients using Lasso regression and model development

Yanxiang Guo[1], Aihu An[1], Po Li[2], Jun Li[3], Baiming Zhang[3]

[1]Breast Department II, Gansu Provincial Cancer Hospital, No. 2 Xiaohuxi East Street, Qilihe District, Lanzhou 730000, Gansu, China; [2]Department of General Surgery, Zhangye Second People's Hospital, No. 270, Linsong West Street, Ganzhou District, Zhangye 734000, Gansu, China; [3]Department of Pharmacy, Zhangye Second People's Hospital, No. 270, Linsong West Street, Ganzhou District, Zhangye 734000, Gansu, China

Received November 26, 2025; Accepted January 12, 2026; Epub April 15, 2026; Published April 30, 2026

Abstract: Objective: We aimed to identify chemotherapy-related predictors of upper-limb breast cancer-related lymphedema (BCRL) and to build and validate a clinically usable prediction model. Methods: Our multicenter study analyzed 670 breast cancer patients treated with chemotherapy from December 2018 through February 2024. We divided patients into training (381 patients, Gansu Provincial Cancer Hospital) and validation (289 patients, Zhangye Second People's Hospital) cohorts. The prediction model combined four machine learning algorithms - decision tree, random forest, support vector machine, and XGBoost-using Lasso regression as the final integrator. We employed K-fold cross-validation to prevent data leakage. Performance metrics included AUC, Brier score, calibration curves, and decision curve analysis. SHAP values helped us understand which factors mattered most. Variables showing P<0.10 in initial screening entered the multivariable model after checking for multicollinearity. Results: Eight factors showed preliminary associations with lymphedema by simple comparisons. However, comprehensive analysis accounting for overlapping effects identified six independent predictors: disease stage, complete versus limited axillary surgery, sentinel node procedure, total nodes removed, radiation treatment, and whether chemotherapy came before or after surgery. Statistical significance was strongest for stage (P<0.001) and weakest for chemotherapy timing (P=0.033). When we tested the model, discrimination remained good though slightly lower than development (AUC dropped from 0.773 to 0.713). Prediction errors stayed modest (Brier scores 0.136-0.146). Calibration was excellent since predicted probabilities matched observed rates. Clinical decision analysis suggested the model adds value when risk thresholds fall between 0-63%, peaking near 18-19% threshold. Feature importance analysis confirmed disease burden (stage, node count), and treatment aggressiveness (surgery type, radiation, chemotherapy sequence) jointly determined risk. Conclusion: We identified six factors that independently raise lymphedema risk after breast cancer treatment. The machine learning model we developed discriminates risk well enough for clinical application. It may help doctors intensify monitoring for high-risk patients while avoiding unnecessary intervention for low-risk patients.

Keywords: Breast cancer-related lymphedema, stacking ensemble model, lasso regression, independent predictors, clinical decision support

Introduction

In 2020, approximately 2.3 million women were newly diagnosed with breast cancer, making it the most frequently diagnosed malignancy worldwide [1]. With improving survival, long-term toxicities have a growing effect on patients' daily function and healthcare use. Among these, breast cancer-related lymphedema (BCRL) is a major concern, leading to limb swelling, pain, limited range of motion, and recurrent infections that impair quality of life and increase cost [2]. Reviews estimate a pooled prevalence of about 21%, but published rates vary from 6% to 40%, reflecting heterogeneity in case definition and follow-up length [3]. The extent of axillary management is a key driver of risk: SLNB is associated with substantially lower incidence than ALND [3]. Radiotherapy can add risk; in some series, axillary fields have

https://doi.org/10.62347/XCIE5727

shown lower lymphedema rates than surgical dissection [4]. Because BCRL can present early after surgery or develop gradually over years, practice should incorporate early risk stratification and longitudinal surveillance throughout perioperative and adjuvant treatment.

Evidence is firmer for surgery and radiation than for chemotherapy with respect to BCRL: differences by axillary strategy (ALND vs SLNB), node counts, and radiation field or dose are well documented [5]. By contrast, the independent effects of chemotherapy-related factors remain mixed. Studies comparing neoadjuvant and adjuvant timing, course length, regimen, and venous access have tended to be small, with residual confounding and heterogeneous case definitions or follow-up windows [6]. Analytic choices add to the problem: traditional main-effects regressions and simplified scores overlook nonlinearity and interactions, making it difficult to achieve good discrimination, calibration, and clinical net benefit in a single model. Because many models do not generalize well beyond the derivation cohort, clinical uptake has been limited. Addressing these issues requires structured, interpretable modeling that evaluates chemotherapy exposures comprehensively and yields patient-level risk tools suitable for practice.

We studied the independent associations between chemotherapy exposures and BCRL using a two-layer approach. First, we built a stacking model with four base learners (decision tree, random forest, support vector machine, XGBoost) and a Lasso meta-learner. The meta-learner was trained on out-of-fold predictions from K-fold cross-validation on the training set; a separate validation set was held out for evaluation to limit information leakage. Second, we evaluated discrimination and overall error with AUC and the Brier score, and examined calibration and clinical usefulness with calibration plots and decision curve analysis (DCA). Variables were screened at $P<0.10$ in univariable tests, then entered into multivariable models after checking multicollinearity using variance inflation factors and pairwise correlation matrices. SHAP provided global feature attribution and case-level explanations to inform postoperative risk stratification and follow-up planning.

Materials and methods

General information

We assembled a two-center retrospective cohort of 670 consecutive breast cancer patients treated with chemotherapy from December 2018 to February 2024. Gansu Provincial Cancer Hospital contributed the training cohort (n=381), and Zhangye Second People's Hospital contributed the validation cohort (n=289). Baseline profiles were comparable (all $P>0.05$; Table S1). The study followed the Declaration of Helsinki [7] and was approved by Gansu Provincial Cancer Hospital ethics committee, which waived informed consent because of the retrospective, minimal-risk design.

Sample size calculation

We prespecified chronic postoperative BCRL as a binary outcome. Based on a pooled incidence of 21.4% [3], we used $P=0.214$ to plan sample size. Following the EPV rule of 10-20 events per candidate predictor and allowing for up to nine predictors, we aimed for 90-180 events. This translated to approximately 421-841 participants. Our final cohort (N=670) implies about 143 events and an EPV of roughly 16, which is consistent with accepted thresholds for multivariable modeling.

Inclusion and exclusion criteria

Inclusion criteria: Female patients with pathologically confirmed primary breast cancer. Receipt of comprehensive treatment, with chemotherapy information available (timing, regimen, duration, and venous access device). Perioperative and follow-up data sufficient to determine BCRL status. Age and demographic information available.

Exclusion criteria: History of primary or secondary lymphedema in either upper extremity. Conditions that may confound upper-limb volume or circumference assessment (e.g., major trauma, active infection, or upper-extremity deep vein thrombosis). Missing key exposure or outcome data. Follow-up insufficient to determine outcome occurrence. Enrollment in clinical trials that could affect outcome assessment when relevant data are unavailable.

Clinical data collection

We recorded demographics and comorbidities (age, body mass index [BMI], education level, marital status, and histories of diabetes and hypertension). Tumor and surgical variables included tumor size, lymph node metastasis, stage, surgical procedure (mastectomy or breast-conserving surgery), laterality (left or right), axillary management (sentinel lymph node biopsy and/or axillary lymph node dissection), the number of lymph nodes removed, and postoperative complications. Radiotherapy, endocrine therapy, biologic therapy, chemotherapy timing (neoadjuvant/adjuvant), chemotherapy regimen (anthracycline combined with taxanes/other), chemotherapy duration (categorized by clinical pathway), chemotherapy vascular access device (PICC, IVAP, PVC). BCRL (yes/no).

Follow-up

During hospitalization, we recorded contact details, clinicopathologic variables, and the planned chemotherapy regimen. Post-discharge follow-up consisted of monthly telephone contacts throughout adjuvant therapy, followed by scheduled surveillance after treatment completion: every 3 months in years 0-2, every 6 months in years 2-5, and annually thereafter. At each contact, we logged treatment implementation and any BCRL events. We completed data retrieval in June 2025.

Measurement methods and outcome definition

BCRL ascertainment combined standardized tape measurements with symptom reporting. Using a nonelastic tape and the ulnar styloid as the reference, we measured circumferences on both arms at roughly 10 cm increments up the forearm and upper arm, yielding five sites per limb. A case was called when any site showed a between-arm difference of at least 2 cm. Severity reflected the largest difference observed: mild if less than 3 cm, moderate if 3-5 cm, and severe if more than 5 cm. For patient-reported symptoms, we administered the Chinese Norman questionnaire in clinic or by phone to record perceived asymmetry of the hand, forearm, and upper arm and upper-limb discomfort. Items were scored ("no difference", "barely noticeable", "noticeable to acquaintances", "obvious difference") and summed; totals greater than 1 were considered subjectively positive and used for screening and documentation during follow-up. Objective measurements anchored the final diagnosis, supplemented by the questionnaire and clinical judgment.

Measured outcomes

Primary outcome: BCRL occurrence (yes/no).

Secondary outcomes: Machine learning model discrimination and calibration performance (AUC, Brier score, ROC/calibration curves), clinical net benefit assessment (decision curve analysis threshold range and peak net benefit), and univariate and multivariable logistic regression correlation results.

Statistical analysis

Statistical analyses were performed in R 4.5.1 and SPSS 26.0. Graphic visualization and machine learning were implemented in R 4.5.1 using key packages including but not limited to glmnet, xgboost, e1071, randomForest, pROC, rms, rmda, and SHAP-related tools.

Data presentation: Counted data presented as n (%). Effect sizes (OR and 95% CI) provided when appropriate.

Counted data testing: Between-group comparisons of two independent samples used Pearson chi-square test. Fisher's exact test was applied when any cell expected frequency <5. For ordinal categorical variables (e.g., stage, treatment duration stratification), linear trend chi-square or Cochran-Armitage trend test was employed. Two-sided *P*-values were reported (α=0.05).

Data processing: Missing values were handled according to preset rules. Categorical variables underwent dummy coding, numerical variables standardized as needed. All preprocessing was fit onto the training set and applied to the validation set to prevent information leakage.

Univariate analysis: Appropriate methods selected based on variable type for between-group comparison, used to screen candidate variables (entry threshold P<0.10 for multivariable analysis). In **Table 1**, multi-category variables (e.g., tumor stage, chemotherapy duration) were analyzed using chi-square tests with the lowest category as reference, yielding cate-

Table 1. Univariate analysis comparing clinical and treatment characteristics between BCRL and Non-BCRL groups

Factor	Total	BCRL Group (n=74)	N-BCRL Group (n=308)	Test Statistic	*P*-value	OR (95% CI)
Age				0.106	0.949	
<40	34 (8.9%)	6 (8.1%)	28 (9.1%)			Reference
40-59	264 (69.1%)	51 (68.9%)	213 (69.2%)			1.117 (0.439-2.841)
≥60	84 (22.0%)	17 (23.0%)	67 (21.8%)			1.184 (0.423-3.316)
BMI				2.458	0.117	
≥24	92 (24.1%)	23 (31.1%)	69 (22.4%)			0.640 (0.366-1.121)
<24	290 (75.9%)	51 (68.9%)	239 (77.6%)			
Education level				1.060	0.303	
≥ High school	117 (30.6%)	19 (25.7%)	98 (31.8%)			1.351 (0.761-2.398)
< High school	265 (69.4%)	55 (74.3%)	210 (68.2%)			
Marital status				0.000	1.000	
Married	370 (96.9%)	72 (97.3%)	298 (96.8%)			0.828 (0.177-3.861)
Other	12 (3.1%)	2 (2.7%)	10 (3.2%)			
Diabetes history				0.150	0.698	
Yes	16 (4.2%)	2 (2.7%)	14 (4.5%)			1.714 (0.381-7.713)
No	366 (95.8%)	72 (97.3%)	294 (95.5%)			
Hypertension history				0.358	0.549	
Yes	44 (11.5%)	10 (13.5%)	34 (11.0%)			0.794 (0.373-1.691)
No	338 (88.5%)	64 (86.5%)	274 (89.0%)			
Tumor size				1.710	0.191	
≥3	238 (62.3%)	51 (68.9%)	187 (60.7%)			0.697 (0.405-1.199)
<3	144 (37.7%)	23 (31.1%)	121 (39.3%)			
Lymph node metastasis				5.116	0.024	
Yes	177 (46.3%)	43 (58.1%)	134 (43.5%)			0.555 (0.332-0.928)
No	205 (53.7%)	31 (41.9%)	174 (56.5%)			
Tumor stage				18.116	<0.001	
I	74 (19.4%)	7 (9.5%)	67 (21.8%)			Reference
II	178 (46.6%)	29 (39.2%)	149 (48.4%)			1.863 (0.777-4.466)
III	121 (31.7%)	33 (44.6%)	88 (28.6%)			3.589 (1.496-8.613)
IV	9 (2.4%)	5 (6.8%)	4 (1.3%)			11.964 (2.596-55.144)
Surgical type				0.708	0.400	
Mastectomy	350 (91.6%)	66 (89.2%)	284 (92.2%)			1.434 (0.617-3.335)
Breast-conserving surgery	32 (8.4%)	8 (10.8%)	24 (7.8%)			
Surgical side				0.283	0.595	
Left	217 (56.8%)	40 (54.1%)	177 (57.5%)			1.148 (0.690-1.912)
Right	165 (43.2%)	34 (45.9%)	131 (42.5%)			
Axillary lymph node dissection				7.345	0.007	
Yes	93 (24.3%)	27 (36.5%)	66 (21.4%)			2.105 (1.220-3.636)
No	289 (75.7%)	47 (63.5%)	242 (78.6%)			
Sentinel lymph node biopsy				4.981	0.026	
Yes	77 (20.2%)	8 (10.8%)	69 (22.4%)			0.420 (0.192-0.917)
No	305 (79.8%)	66 (89.2%)	239 (77.6%)			
Number of lymph nodes removed				8.475	0.004	
≥10	64 (16.8%)	4 (5.4%)	60 (19.5%)			4.234 (1.487-12.055)
<10	318 (83.2%)	70 (94.6%)	248 (80.5%)			
Postoperative complications				0.100	0.752	
Yes	40 (10.5%)	7 (9.5%)	33 (10.7%)			1.149 (0.487-2.709)
No	342 (89.5%)	67 (90.5%)	275 (89.3%)			

Radiotherapy				6.599	0.010	
Yes	69 (18.1%)	21 (28.4%)	48 (15.6%)			0.466 (0.258-0.842)
No	313 (81.9%)	53 (71.6%)	260 (84.4%)			
Endocrine therapy				0.012	0.913	
Yes	127 (33.2%)	25 (33.8%)	102 (33.1%)			0.970 (0.567-1.660)
No	255 (66.8%)	49 (66.2%)	206 (66.9%)			
Biologic therapy				0.014	0.907	
Yes	48 (12.6%)	9 (12.2%)	39 (12.7%)			1.047 (0.483-2.270)
No	334 (87.4%)	65 (87.8%)	269 (87.3%)			
Chemotherapy timing				6.518	0.011	
Neoadjuvant chemotherapy	37 (9.7%)	13 (17.6%)	24 (7.8%)			0.397 (0.191-0.822)
Adjuvant chemotherapy	345 (90.3%)	61 (82.4%)	284 (92.2%)			
Chemotherapy regimen				0.132	0.717	
Anthracycline plus taxane	280 (73.3%)	53 (71.6%)	227 (73.7%)			1.110 (0.631-1.955)
Other	102 (26.7%)	21 (28.4%)	81 (26.3%)			
Chemotherapy duration				12.502	0.002	
4 weeks	276 (72.3%)	52 (70.3%)	224 (72.7%)			Reference
6 weeks	99 (25.9%)	17 (23.0%)	82 (26.6%)			0.893 (0.489-1.632)
8 weeks	7 (1.8%)	5 (6.8%)	2 (0.6%)			10.769 (2.033-57.058)
Chemotherapy vascular access device				4.975	0.083	
PICC	97 (25.4%)	18 (24.3%)	79 (25.6%)			Reference
IVAP	245 (64.1%)	43 (58.1%)	202 (65.6%)			0.934 (0.508-1.717)
PVC	40 (10.5%)	13 (17.6%)	27 (8.8%)			2.113 (0.915-4.878)

Note: OR, Odds ratio; CI, Confidence interval. For ordinal multi-category variables (chemotherapy duration, chemotherapy vascular access device), ORs represent the effect per one-category increase rather than comparison with a specific reference category, which accounts for differences from category-specific ORs reported in **Table 1**. PICC, Peripherally inserted central catheter; IVAP, Implantable venous access port; PVC, Peripheral venous catheter.

Table 2. Variance inflation factor (VIF) assessment of nine candidate variables

Variable	VIF	Interpretation
Lymph node metastasis	1.018	Low multicollinearity
Tumor stage	1.035	Low multicollinearity
Axillary lymph node dissection	1.024	Low multicollinearity
Sentinel lymph node biopsy	1.032	Low multicollinearity
Number of lymph nodes removed	1.024	Low multicollinearity
Radiotherapy	1.046	Low multicollinearity
Chemotherapy timing	1.057	Low multicollinearity
Chemotherapy duration	1.025	Low multicollinearity
Chemotherapy vascular access device	1.015	Low multicollinearity

Note: VIF, Variance inflation factor; ALND, Axillary lymph node dissection; SLNB, Sentinel lymph node biopsy; CVAD, Central venous access device.

Multicollinearity and correlation: Variance inflation factors assessed multicollinearity with correlation matrix visualization. Variables with acceptable VIF range proceeded to subsequent analysis.

Multivariable analysis: Binary logistic regression was used to construct the clinical baseline model, reporting odds ratios with 95% confidence intervals. Statistical significance was determined by two-sided testing with α=0.05.

gory-specific odds ratios. In **Table 3** logistic regression, these variables were modeled as continuous ordinal variables to assess overall dose-response trends, yielding per-unit ORs that differ in interpretation from pairwise comparisons. Minor *P*-value discrepancies between chi-square and Wald tests (e.g., lymph node metastasis: 0.024 vs 0.025) reflect differences in test statistics rather than data inconsistencies.

Machine learning and ensemble: Gansu Provincial Cancer Hospital samples formed the training set (n=381), and Zhangye Second People's Hospital samples served as a completely independent external validation set (n=289) that was not involved in any model training or tuning process. Four base learners were implemented: decision tree (rpart), random forest (randomForest), support vector machine with radial basis function kernel (e1071), and

Table 3. Univariate and multivariable logistic regression analysis of BCRL

Variable	Univariate Analysis			Multivariable Analysis		
	OR value	*P* value	95% CI	OR value	*P* value	95% CI
Lymph node metastasis	0.555	0.025	0.332-0.928	0.576	0.054	0.327-1.005
Tumor stage	2.036	<0.001	1.424-2.912	2.226	<0.001	1.534-3.295
Axillary lymph node dissection	2.105	0.007	1.220-3.636	2.012	0.023	1.096-3.663
Sentinel lymph node biopsy	0.420	0.029	0.192-0.917	0.365	0.017	0.150-0.793
Number of lymph nodes removed	4.234	0.007	1.487-12.054	4.662	0.005	1.76-16.214
Radiotherapy	0.466	0.011	0.258-0.842	0.375	0.003	0.195-0.729
Chemotherapy timing	0.397	0.013	0.191-0.822	0.408	0.033	0.180-0.950
Chemotherapy duration	1.390	0.183	0.856-2.257			
Chemotherapy vascular access device	1.353	0.179	0.871-2.101			

Note: OR, Odds ratio; CI, Confidence interval. For axillary lymph node dissection and sentinel lymph node biopsy, ORs are presented as "Yes versus No" with "No" as the reference category. For ordinal multi-category variables (chemotherapy duration, chemotherapy vascular access device), ORs represent the effect per one-category increase rather than comparison with a specific reference category, which accounts for differences from category-specific ORs reported in **Table 1**.

XGBoost (xgboost). Hyperparameter tuning was performed using 5-fold stratified cross-validation with grid search optimization (5×5 resolution), maximizing AUC as the objective. The tuned hyperparameter ranges were: decision tree (complexity parameter cp: 0.01-0.1, minimum bucket size: 5-20), random forest (mtry: 2 to total feature count, node size: 1-50), SVM (cost: 0.001-32, gamma: 0.01-16), and XGBoost (learning rate eta: 0.01-0.5, number of boosting rounds: 200-500, maximum tree depth: 1-4). For stacking ensemble, the first layer generated out-of-fold predictions using 3-fold cross-validation, ensuring that predictions for each training sample were made by models that had not seen that sample during training, thereby preventing data leakage. The second layer employed Lasso regression (glmnet package) as the meta-learner, with the regularization parameter lambda selected by internal cross-validation using the lambda.min criterion (the lambda value minimizing cross-validation error). Features with non-zero Lasso coefficients were retained in the final model. All analyses were performed in R version 4.5.1 using the mlr3 machine learning framework.

Performance evaluation and clinical utility: AUC assessed discrimination. Brier score evaluated overall prediction error, supplemented by ROC and calibration curve visualization. DeLong method was applied when necessary to compare AUC differences. Decision curve analysis evaluated clinical net benefit, reporting threshold ranges and peak values with net benefit. Feature importance and SHAP provided global and individual-level model interpretation.

Results

BCRL-related factors identified by univariate analysis

Factors significantly associated with BCRL included: lymph node metastasis (P=0.024), tumor stage (P<0.001), axillary lymph node dissection (P=0.007), sentinel lymph node biopsy (P=0.026), number of lymph nodes removed (P=0.004), radiotherapy (P=0.010), chemotherapy timing (P=0.011), and chemotherapy duration (P=0.002) (see **Table 1**). Other variables including age, BMI, education level, marital status, diabetes history, hypertension history, tumor size, surgical type, surgical side, postoperative complications, endocrine therapy, biologic therapy, chemotherapy regimen, and chemotherapy vascular access device showed no significant differences with BCRL (P>0.05) (see **Table 1**).

Multicollinearity and candidate variable correlation assessment

All nine candidate variables entering subsequent regression demonstrated low multicollinearity, with variance inflation factors within acceptable ranges (see **Table 2**). Correlation heatmap revealed generally weak correlations among variables, with only a few paired variables reaching significance: tumor stage and

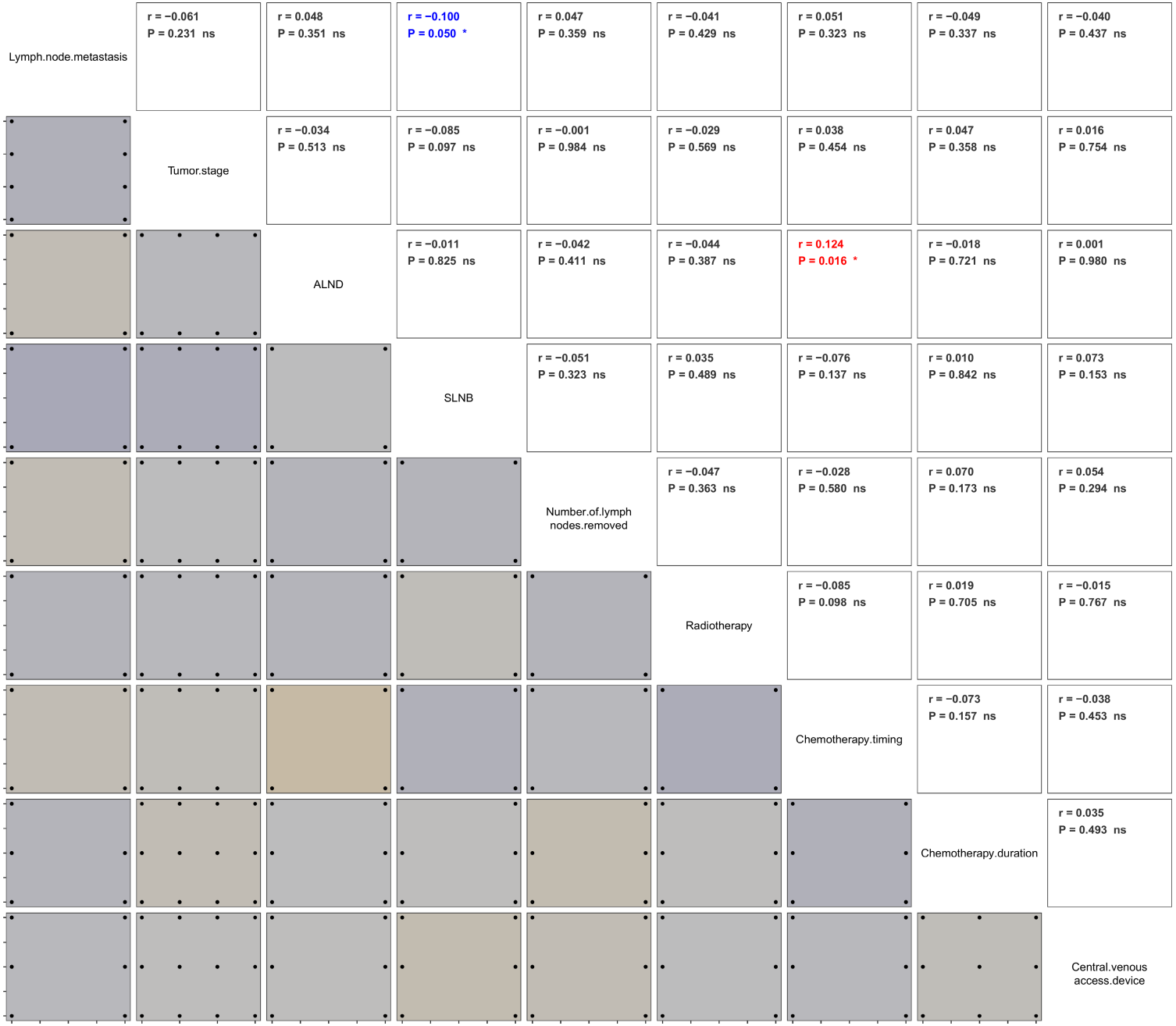
Lymph.node.metastasis
r = −0.061
P = 0.231 ns
r = 0.048
P = 0.351 ns
r = −0.100
P = 0.050 *
r = 0.047
P = 0.359 ns
r = −0.041
P = 0.429 ns
r = 0.051
P = 0.323 ns
r = −0.049
P = 0.337 ns
r = −0.040
P = 0.437 ns
Tumor.stage
r = −0.034
P = 0.513 ns
r = −0.085
P = 0.097 ns
r = −0.001
P = 0.984 ns
r = −0.029
P = 0.569 ns
r = 0.038
P = 0.454 ns
r = 0.047
P = 0.358 ns
r = 0.016
P = 0.754 ns
ALND
r = −0.011
P = 0.825 ns
r = −0.042
P = 0.411 ns
r = −0.044
P = 0.387 ns
r = 0.124
P = 0.016 *
r = −0.018
P = 0.721 ns
r = 0.001
P = 0.980 ns
SLNB
r = −0.051
P = 0.323 ns
r = 0.035
P = 0.489 ns
r = −0.076
P = 0.137 ns
r = 0.010
P = 0.842 ns
r = 0.073
P = 0.153 ns
Number.of.lymph nodes.removed
r = −0.047
P = 0.363 ns
r = −0.028
P = 0.580 ns
r = 0.070
P = 0.173 ns
r = 0.054
P = 0.294 ns
Radiotherapy
r = −0.085
P = 0.098 ns
r = 0.019
P = 0.705 ns
r = −0.015
P = 0.767 ns
Chemotherapy.timing
r = −0.073
P = 0.157 ns
r = −0.038
P = 0.453 ns
Chemotherapy.duration
r = 0.035
P = 0.493 ns
Central.venous access.device

Figure 1. Pairwise correlation matrix with significance annotation for nine candidate variables. Note: ALND, Axillary lymph node dissection; SLNB, Sentinel lymph node biopsy; CVAD, Central venous access device.

sentinel lymph node biopsy (P<0.05), axillary lymph node dissection, and chemotherapy timing (P=0.016). All other pairwise correlations showed no significant differences (P>0.05) (see **Figure 1**). Based on these assessments, all variables with P<0.10 were incorporated into multivariable regression modeling (see **Table 2** and **Figure 1**).

Univariate and multivariable logistic regression analysis

Univariate analysis identified factors significantly associated with upper limb lymphedema: lymph node metastasis (P=0.025), tumor stage (P<0.001), axillary lymph node dissection (P=0.007), sentinel lymph node biopsy (P=0.029), number of lymph nodes removed (P=0.007), radiotherapy (P=0.011), and chemotherapy timing (P=0.013) (see **Table 3**). Chemotherapy duration and chemotherapy vascular access device showed no significance (P>0.05) (see **Table 3**). In multivariable analysis, tumor stage (P<0.001), axillary lymph node dissection (P=0.023), sentinel lymph node biopsy (P=0.017), number of lymph nodes removed (P=0.005), radiotherapy (P=0.003), and chemotherapy timing (P=0.033) remained independent correlates (see **Table 3**). Specifically, axillary lymph node dissection was associated with increased BCRL risk (OR=2.012, 95% CI: 1.096-3.663), consistent with the extensive disruption of lymphatic drainage pathways during this procedure. In contrast, sentinel lymph node biopsy was associated with reduced BCRL risk (OR=0.365, 95% CI: 0.150-0.793), reflecting its minimally invasive nature that better preserves lymphatic function compared to complete axillary dissection. Lymph node metastasis did not reach significance after adjustment (P=0.054) (see **Table 3**).

Structure and variable importance of four base models

The four machine learning models demonstrated generally consistent feature importance rankings. Tumor stage occupied a central position across all models, followed by number of lymph nodes removed, axillary lymph node dissection (ALND), lymph node metastasis, radiotherapy, and chemotherapy timing among chemotherapy/treatment-related factors (see **Figure 2**). The decision tree used tumor stage as the primary split feature, with subsequent divisions mainly driven by radiotherapy, number of lymph nodes removed, and chemotherapy timing, indicating these variables provided discrimination even in single models (see **Figure 2**). Permutation importance results from random forest, SVM (RBF), and XGBoost mutually confirmed that treatment and pathologic burden-related variables contributed substantially to model discrimination, providing foundation for subsequent stacking and Lasso selection (see **Figure 2**).

Feature selection and coefficient direction of lasso meta-model

Cross-validation selected the optimal penalty parameter where error was minimized and stable, effectively controlling model complexity (see **Figure 3**). Coefficient paths showed that as penalty strength increased, most feature coefficients gradually shrank to zero, with only a few retained at optimal penalty (see **Figure 3**). At lambda.min, the Lasso meta-model retained prediction probabilities from base models (RF, SVM, DT) and key clinical variables (tumor stage, chemotherapy timing, number of lymph nodes removed, radiotherapy). Tumor stage (III, IV) and chemotherapy/treatment-related variables exhibited positive coefficients, while RF and DT prediction probabilities showed negative coefficients and SVM prediction probability displayed a positive coefficient (see **Figure 3**).

Discrimination, calibration, and clinical net benefit of stacking model

In both training and validation sets, the stacking model demonstrated good discriminative ability. Training set ROC curves showed favorable discrimination (AUC=0.773), with validation set AUC=0.713. Only mild performance decline occurred upon external validation, suggesting robust generalizability suitable for individual risk identification in clinical settings (see **Figure 4A**, **4D**). Calibration curves in both training and validation sets closely approximated

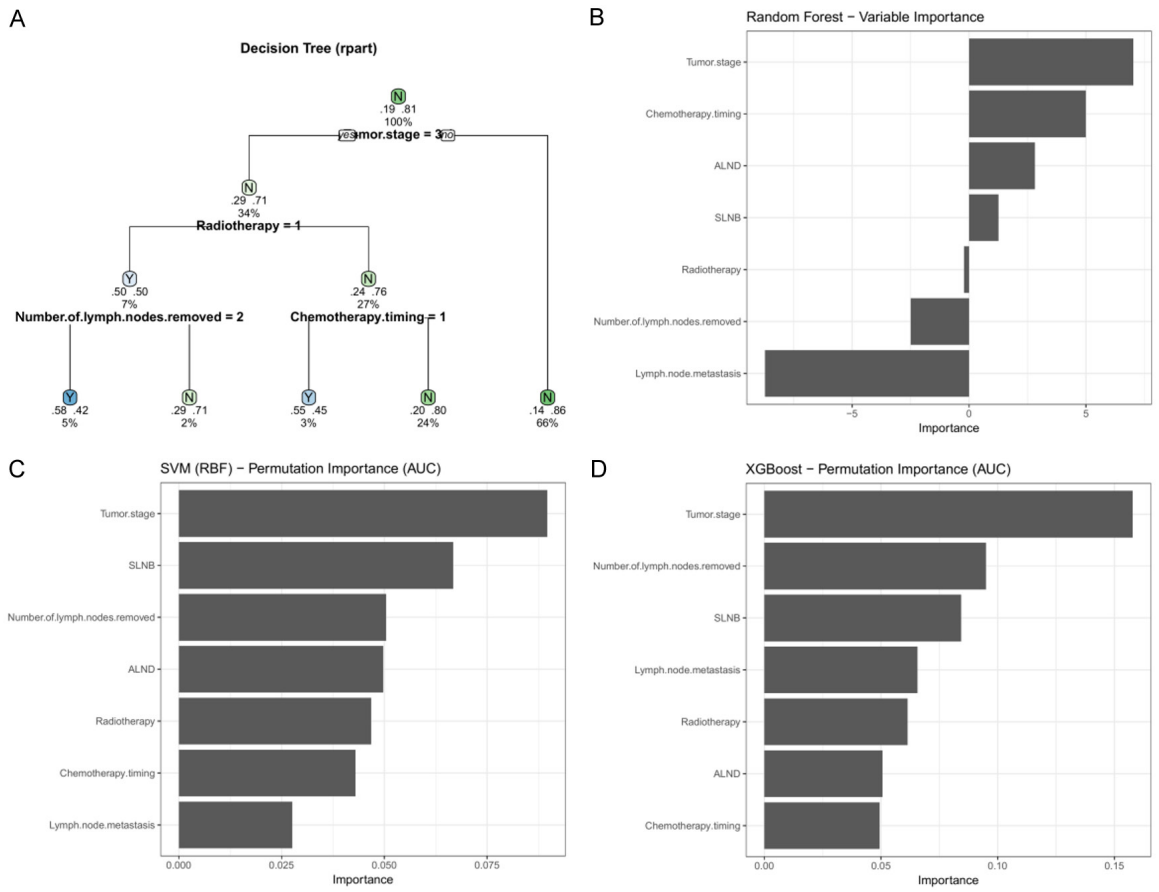

Figure 2. Structure and variable importance of four machine learning models. A: Decision tree structure diagram showing primary split features and pathways. B: Variable importance bar chart from random forest. C: Permutation importance (AUC) bar chart from SVM (RBF). D: Permutation importance (AUC) bar chart from XGBoost. Note: SVM, Support vector machine; RBF, Radial basis function; AUC, Area under the receiver operating characteristic curve; XGBoost, Extreme gradient boosting; ALND, Axillary lymph node dissection; SLNB, Sentinel lymph node biopsy.

the ideal line, with Brier scores of 0.136 and 0.146 respectively, indicating low overall prediction error without obvious systematic overestimation or underestimation. Observed and predicted values were generally consistent within risk stratification intervals, supporting use of probability outputs for clinical grading and follow-up strategy formulation (see **Figure 4B**, **4E**).

Within commonly used clinical threshold ranges, the model achieved higher net benefit compared to "treat all" or "treat none" strategies. In the training set, net benefit existed within the 0-63% threshold probability range, peaking at 18.6%. The validation set similarly showed net benefit within the 0-63% range, with a peak of 19.0%. This indicates that in screening and follow-up scenarios aimed at early identification and intervention, the model can reduce unnecessary interventions while capturing more truly high-risk individuals. When thresholds exceed 63%, the model's relative advantage diminishes, requiring comprehensive judgment combining clinical preferences and resource allocation (see **Figure 4C**, **4F**). At the optimal threshold of 19.1% determined by Youden index, the model achieved sensitivity of 64.9%, specificity of 73.7%, positive predictive value (PPV) of 37.2%, and negative predictive value (NPV) of 89.7% in the training set. These metrics remained stable in external validation (sensitivity 64.9%, specificity 72.7%, PPV 37.0%, NPV 89.4%), demonstrating robust generalizability. The high NPV (approximately 90%) indicates that patients classified as low-risk have a strong probability of remaining lymphedema-free, supporting the model's utility for risk stratification and follow-up planning.

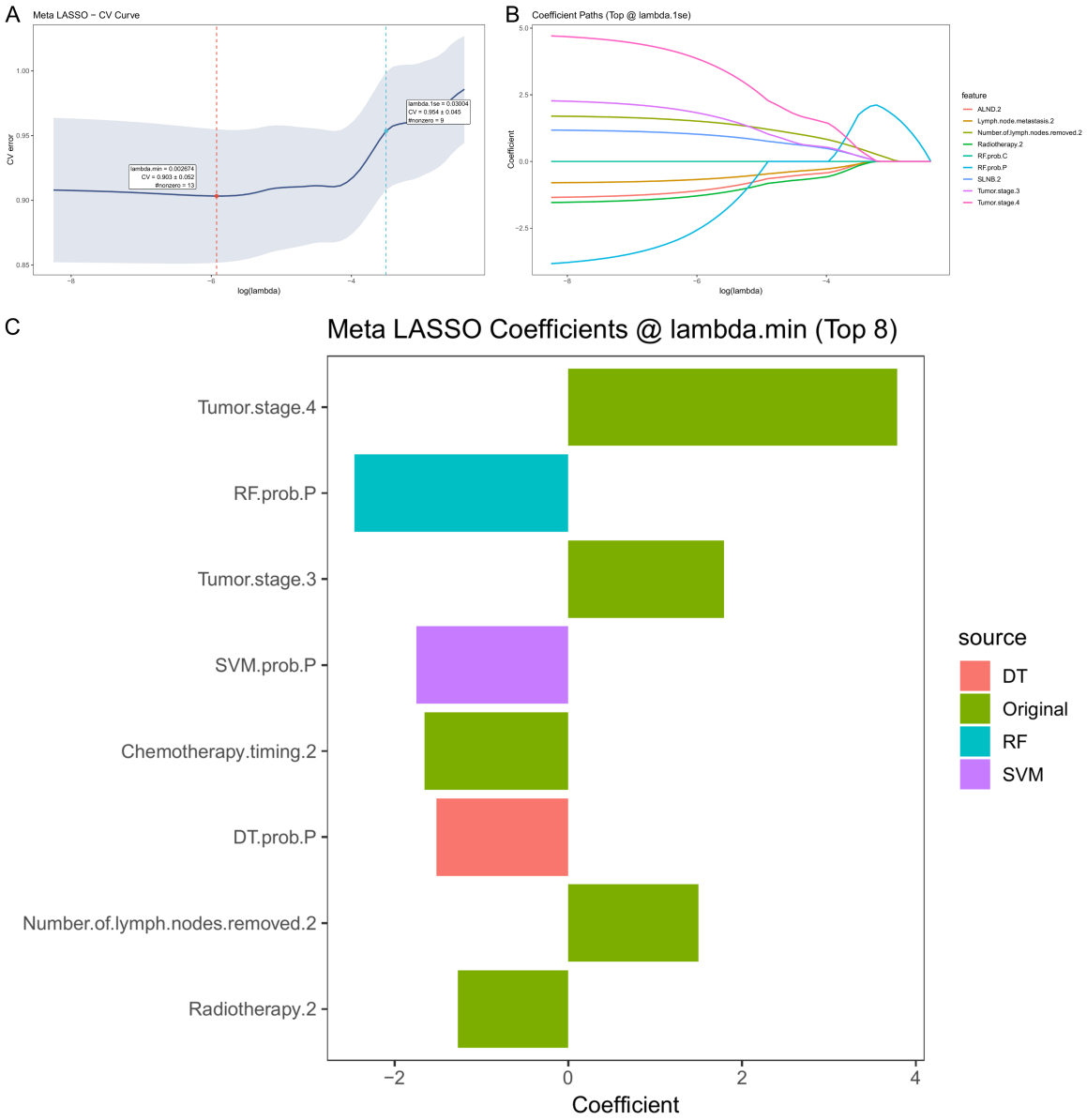

Figure 3. Cross-validation curve, coefficient path, and feature coefficients of lasso meta-model. A: Cross-validation error curve with optimal penalty parameter location. B: Lasso coefficient path plot showing coefficient trajectories with varying penalty. C: Top coefficient bar chart for features selected at lambda.min, including base model prediction probabilities and clinical variables. Note: LASSO, Least absolute shrinkage and selection operator; CV, Cross-validation; DT, Decision tree; RF, Random forest; SVM, Support vector machine.

Individual-level SHAP interpretation (cases 1 and 100)

SHAP weights across the full sample showed tumor stage exerted the greatest influence on model output, followed by number of lymph nodes removed, axillary lymph node dissection (ALND), sentinel lymph node biopsy (SLNB), radiotherapy, and chemotherapy timing. High values (or high-risk categories) generally pushed predictions upward, while low values (or low-risk categories) pulled predictions downward. Variable influence direction aligned with clinical mechanisms (see **Figure 5A**).

Case 1 achieved an individual prediction score of 0.362. Major positive contributions came from higher tumor stage, greater number of lymph nodes removed, and chemotherapy timing, with radiotherapy and SLNB providing addi-

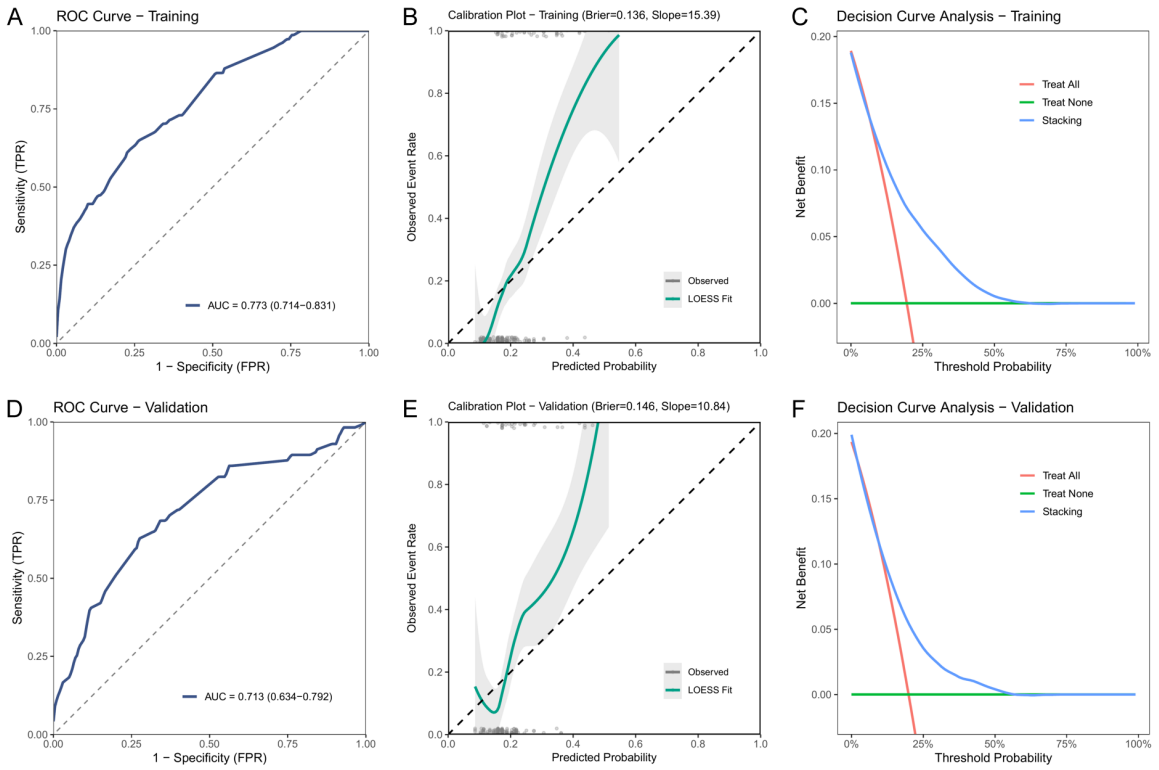

Figure 4. ROC, calibration, and decision curve analysis of stacking model in training and validation sets. A: Training set ROC curve (AUC=0.773). B: Training set calibration curve (Brier=0.136). C: Training set decision curve showing net benefit within 0-63% threshold range, with peak at 18.6%. D: Validation set ROC curve (AUC=0.713). E: Validation set calibration curve (Brier=0.146). F: Validation set decision curve showing net benefit within 0-63% threshold range, with peak at 19.0%. Note: ROC, Receiver operating characteristic; AUC, Area under the curve; Brier, Brier score; DCA, Decision curve analysis; NB, Net benefit.

tional upward push. Limited downward pull from low-risk values resulted in net upward movement, explaining its placement in the higher-risk range (see **Figure 5B**).

Case 100 received an individual prediction score of 0.190. Low-risk chemotherapy timing, fewer lymph nodes removed, ALND-related values, and lower tumor stage collectively formed major negative contributions. Only minor upward pulls from a few factors were insufficient to offset the overall downward trend, explaining its placement in the lower-risk range (see **Figure 5C**).

Discussion

In a two-center retrospective cohort with standardized follow-up led by dedicated case managers and supported by a disease-management system, we examined whether chemotherapy-related factors independently relate to postoperative BCRL. We trained a stacking model that combined four base learners (decision tree, random forest, support vector machine, XGBoost) with a Lasso meta-learner. Independent correlates included tumor stage, ALND, SLNB, the number of nodes removed, radiotherapy, and the timing of chemotherapy. Discrimination and overall error were acceptable (AUC 0.773/0.713; Brier 0.136/0.146 for training/validation). Decision-curve analysis indicated positive net benefit for threshold probabilities between 0% and 63%, with peaks at approximately 18.6% and 19.0%. SHAP profiles supported a risk pattern driven by pathologic burden together with treatment exposure, observable at the cohort and patient levels. The model provides a practical basis for risk stratification and follow-up planning and helps fill the evidence gap around chemotherapy-related risk.

Large-scale evidence links axillary management and radiotherapy to BCRL risk: in 65,543 patients, Kuruvilla et al. [8] identified both fac-

Figure 5. SHAP-based feature importance across full sample and individual-level interpretation for two cases. A: SHAP distribution across full sample showing impact of high/low values on prediction direction and magnitude. B: Case 1 (lymphedema): Bar chart of positive and negative contributions of each feature to individual prediction. C: Case 100 (non-lymphedema): Bar chart of positive and negative contributions of each feature to individual prediction.

tors as strong predictors, and dose to the axillary lateral thoracic artery vessel junction (ALTJ) region was associated with BCRL in a dosimetric study [9]. Signals from chemotherapy are less uniform. Aoishi et al. [10] reported higher risk with docetaxel (HR 3.790, 95% CI 1.413-10.167, P=0.0081) but no increase with neoadjuvant timing, whereas a prospective cohort by Montagna et al. [11] observed higher BCRL after neoadjuvant therapy than after immediate surgery (29.3% vs 11.1%, P=0.01). Risk modeling that combined lymph-node burden with chemotherapy and radiotherapy reached similar conclusions: Kim et al. [12] reported 5-year incidences of 19% and 38% for patients with two versus three risk factors, aligning with a pattern in which pathologic burden and treatment exposures both shape risk.

In our data, the stacking ensemble exceeded each individual learner. Prior work shows comparable ranges: an XGBoost model reported a validation AUC of 0.89 [13], and in a head-to-head comparison of nine methods, logistic regression was optimal (AUC 0.87), consistent with the competitiveness of a penalized-logistic (Lasso) meta-learner [14]. A nomogram developed in a China-based cohort reported AUCs of 0.99 in training and 0.89 in validation [15]. Cross-study AUCs are not strictly compa-

rable because of differences in case definitions and cohorts, but together these reports support the feasibility of BCRL risk prediction across settings.

We embedded risk stratification into routine follow-up managed by case managers and tracked through a hospital registry with preset contact intervals, enabling timely outcome ascertainment. Consistent with Hing et al. [16], our workflow combined patient-reported symptoms, standardized tape measurements, and clinician review. Decision-curve analysis indicated net benefit for threshold probabilities up to 63%. As a practical example, a predicted risk above roughly 40% can trigger intensified surveillance and early rehabilitation (progressive resistance training, complex decongestive therapy), whereas lower-risk patients remain on the standard schedule. In a China-based cohort, a 1,054-patient nomogram provided 5-year risk estimates at preoperative and early/late postoperative time points [17], which supports risk assessment that adapts over time.

The independent correlates in this study map to concrete changes in care pathways. For surgery, axillary lymph node dissection was associated with a two-fold increase in BCRL risk (OR=2.012, 95% CI: 1.096-3.663), whereas sentinel lymph node biopsy demonstrated a protective effect (OR=0.365, 95% CI: 0.150-0.793). These findings strongly support the preference for sentinel node biopsy over complete axillary dissection when oncologically appropriate, since SLNB preserves lymphatic drainage while providing adequate staging information. Prior work underscores this gradient: the level of lymph node removal was the dominant predictor of BCRL in Martínez-Jaimez et al. [18], and IPN dissection carried the highest odds of severe lymphedema (OR 7.76, 95% CI 3.87-15.54) in Liu et al. [19]. For radiotherapy, dose-volume sparing is relevant: radiotherapy was the strongest independent predictor of late BCRL in one analysis (OR 2.12, 95% CI 1.22-3.67, P=0.007) [20]. Chemotherapy-related signals point to timing and duration as modifiable touchpoints: plan activity guidance during the neoadjuvant phase, standardize device care (PICC/IVAP/PVC), and reinforce adherence. SHAP profiles indicated that different combinations of predictors can yield comparable risk ("equivalent risk combinations"), which allows tailoring interventions to the mix of risks seen in a given patient. Integrating the model into the EMR to trigger alerts, follow-up intensification, or rehabilitation referrals at predefined thresholds enables threshold-adaptive CDS.

Methodologically, this study employed leak-prevention strategies and multi-metric evaluation within a dual-center training/external validation framework, ensuring robustness and translational potential. Lin et al. [21] externally validated a 5-factor risk model showing good discriminative ability (sensitivity 0.83, specificity 0.89, accuracy 0.88). Variable screening used $P<0.10$ entry threshold, with multicollinearity controlled by VIF and correlation matrices (VIF approaching 1), and EPV approximately 15.9 meeting stability requirements. Four base learners generated OOF predictions through K-fold cross-validation, with second-layer Lasso completing stacking to prevent information leakage and overfitting. Literature [22] employed Cox-LASSO combined with seven survival machine learning algorithms, similar to our framework. Wei et al. [23] compared six algorithms with logistic regression performing best (AUC 0.889) and deployed a web application. Performance was assessed by AUC, Brier score, and ROC/calibration curves. DCA quantified clinical utility, superior to "treat all" or "treat none" across broad thresholds. Li et al. [24] reported a Chinese multicenter nomogram model with AUC 0.728/0.710, demonstrating net benefit at 10%-80% thresholds. SHAP provided global and individual interpretability consistent with mechanisms.

BCRL incidence in this study aligns generally with international literature, though heterogeneity across studies reveals systematic effects of methodologic differences on outcome estimation. DiSipio et al.'s [3] meta-analysis of 72 studies reported overall incidence of 16.6% (95% CI 13.6-20.2), but prospective cohort subgroup analysis showed this rose to 21.4%, suggesting retrospective designs may underestimate true disease burden. Ahn et al. [25] reported 12.2% overall incidence in a Korean nationwide cohort, while the modified radical mastectomy subgroup reached 28.5%, highlighting the decisive role of surgical extent.

Notably, the independent contributions of race and social determinants to BCRL risk have re-

ceived increasing attention. Both Montagna et al. [11] and Rochlin et al. [5] confirmed Black race and Hispanic ethnicity as independent risk factors (OR 3.88 and 3.01), with race variable inclusion significantly improving model discrimination (AUC from 0.75 to 0.86). However, Siotos et al. [26] also identified Medicaid insurance type as a risk factor, suggesting "race" may be a composite variable integrating genetic susceptibility, body composition differences, healthcare accessibility, and socioeconomic status. This study did not include race and socioeconomic variables, which have limited impact in relatively homogeneous Chinese populations but warrant attention in future multicenter or international collaborative research.

Systematic reviews by Shen et al. [27] and Lin et al. [28] surveyed 21-22 BCRL models and reported wide AUC ranges (0.601-0.965) alongside high risk of bias; only two models had external validation (pooled AUC 0.70 and 0.80). Marked drops from training to validation were common, indicating overfitting. The most consistent predictors across reviews-radiotherapy, preoperative BMI, node count removed, and chemotherapy-are consistent with the independent correlates in our analysis. Methodologically, we addressed several pain points by using dual-center external validation and an out-of-fold stacking ensemble, and by reporting clinical utility (DCA) and interpretability (SHAP). Our performance gap from training to validation (AUC 0.773 to 0.713) was modest. Nevertheless, reviewers conclude that current models are not yet suitable for guideline-level use. We add granular evidence on chemotherapy timing and outline a route to clinical decision support by system integration, but confirm that moving from prediction to risk-guided treatment will require prospective randomized trials.

For clinical implementation, we propose a risk-stratified management pathway based on model predictions. At the recommended threshold of approximately 20%, patients can be classified into high-risk (predicted probability ≥20%) and low-risk (<20%) groups. High-risk patients (approximately 30% of the population based on our data) should receive intensified surveillance including monthly circumference measurements during the first postoperative year, early referral to lymphedema specialists, prophylactic interventions such as compression garments and range-of-motion exercises, and patient education on early warning signs. Low-risk patients can follow standard surveillance protocols with measurements every 3-6 months. The model's high NPV (approximately 90%) provides reassurance that low-risk classifications are reliable, reducing unnecessary interventions while maintaining safety. For clinical translation, the model could be developed into a web-based calculator or integrated into electronic medical record systems to provide real-time risk assessment at key decision points: preoperatively for surgical planning, immediately postoperatively for rehabilitation referral, and during adjuvant therapy for monitoring intensification. Future work should prospectively validate whether risk-stratified management improves outcomes.

This study has certain limitations. The retrospective dual-center design may introduce selection bias. BCRL determination primarily relied on circumference measurement, possibly causing misclassification risk. Chemotherapy granular data was insufficient. Sample event number was limited with a class imbalance. Future multicenter prospective studies should incorporate multimodal data including volumetrics and bioelectrical impedance, optimize models within causal frameworks with external validation, and embed models in clinical decision support systems to drive BCRL management toward proactive prevention.

Conclusion

Based on dual-center retrospective data, this study constructed a stacking ensemble model for predicting BCRL risk. Analysis confirmed tumor stage, axillary lymph node dissection, radiotherapy, and chemotherapy timing as independent risk factors.

Acknowledgements

This work was supported by the Natural Science Foundation of Gansu Province (25JRRA769).

Disclosure of conflict of interest

None.

Address correspondence to: Baiming Zhang, Department of Pharmacy, Zhangye Second People's Hospital, No. 270, Linsong West Street, Ganzhou District, Zhangye 734000, Gansu, China. E-mail: 13993640981@163.com

References

[1] Sung H, Ferlay J, Siegel RL, Laversanne M, Soerjomataram I, Jemal A and Bray F. Global Cancer Statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin 2021; 71: 209-249.

[2] Donahue PMC, MacKenzie A, Filipovic A and Koelmeyer L. Advances in the prevention and treatment of breast cancer-related lymphedema. Breast Cancer Res Treat 2023; 200: 1-14.

[3] DiSipio T, Rye S, Newman B and Hayes S. Incidence of unilateral arm lymphoedema after breast cancer: a systematic review and meta-analysis. Lancet Oncol 2013; 14: 500-515.

[4] Donker M, van Tienhoven G, Straver ME, Meijnen P, van de Velde CJ, Mansel RE, Cataliotti L, Westenberg AH, Klinkenbijl JH, Orzalesi L, Bouma WH, van der Mijle HC, Nieuwenhuijzen GA, Veltkamp SC, Slaets L, Duez NJ, de Graaf PW, van Dalen T, Marinelli A, Rijna H, Snoj M, Bundred NJ, Merkus JW, Belkacemi Y, Petignat P, Schinagl DA, Coens C, Messina CG, Bogaerts J and Rutgers EJ. Radiotherapy or surgery of the axilla after a positive sentinel node in breast cancer (EORTC 10981-22023 AMAROS): a randomised, multicentre, open-label, phase 3 non-inferiority trial. Lancet Oncol 2014; 15: 1303-1310.

[5] Rochlin DH, Barrio AV, McLaughlin S, Van Zee KJ, Woods JF, Dayan JH, Coriddi MR, McGrath LA, Bloomfield EA, Boe L and Mehrara BJ. Feasibility and clinical utility of prediction models for breast cancer-related lymphedema incorporating racial differences in disease incidence. JAMA Surg 2023; 158: 954-964.

[6] Sharifi N and Ahmad S. Breast cancer-related lymphedema: a critical review on recent progress. Surg Oncol 2024; 56: 102124.

[7] Wen B, Zhang G, Zhan C, Chen C and Yi H. The 2024 revision of the declaration of Helsinki: a modern ethical framework for medical research. Postgrad Med J 2025; 101: 371-382.

[8] Kuruvilla AS, Krajewski A, Li X, Yang J, Mulay SR, Agha SM, Kohli HK, Bellis RM, Tannous HJ and Shroyer ALW. Risk factors associated with postmastectomy breast cancer lymphedema: a multicenter retrospective analysis. Ann Plast Surg 2022; 88 Suppl 3: S239-S245.

[9] Jarm T, Besic N, Arnez RC, But-Hadzic J and Ratosa I. Breast cancer related lymphedema and shoulder mobility following radiotherapy. Strahlenther Onkol 2026; 202: 187-195.

[10] Aoishi Y, Oura S, Nishiguchi H, Hirai Y, Miyasaka M, Kawaji M, Shima A and Nishimura Y. Risk factors for breast cancer-related lymphedema: correlation with docetaxel administration. Breast Cancer 2020; 27: 929-937.

[11] Montagna G, Zhang J, Sevilimedu V, Charyn J, Abbate K, Gomez EA, Mehrara B, Morrow M and Barrio AV. Risk factors and racial and ethnic disparities in patients with breast cancer-related lymphedema. JAMA Oncol 2022; 8: 1195-1200.

[12] Kim M, Kim SW, Lee SU, Lee NK, Jung SY, Kim TH, Lee ES, Kang HS and Shin KH. A model to estimate the risk of breast cancer-related lymphedema: combinations of treatment-related factors of the number of dissected axillary nodes, adjuvant chemotherapy, and radiation therapy. Int J Radiat Oncol Biol Phys 2013; 86: 498-503.

[13] Sun Y, Xia X and Liu X. Predictive modeling of breast cancer-related lymphedema using machine learning algorithms. Gland Surg 2024; 13: 2243-2252.

[14] Du J, Yang J, Yang Q, Zhang X, Yuan L and Fu B. Comparison of machine learning models to predict the risk of breast cancer-related lymphedema among breast cancer survivors: a cross-sectional study in China. Front Oncol 2024; 14: 1334082.

[15] Jiang Q, Hu H, Liao J, Li ZH and Tan J. Development and validation of a nomogram for breast cancer-related lymphedema. Sci Rep 2024; 14: 15602.

[16] Hing JX, Chua YN, Tan PT, Tan MSL, Mok CW, Seet MYL, Lin ZC, Seah CM, Lee WP and Tan SM. Defining breast cancer-related lymphedema (BCRL) prevalence and risk factors: a pragmatic approach to lymphedema surveillance. Ann Acad Med Singap 2024; 53: 80-89.

[17] Bevilacqua JL, Kattan MW, Changhong Y, Koifman S, Mattos IE, Koifman RJ and Bergmann A. Nomograms for predicting the risk of arm lymphedema after axillary dissection in breast cancer. Ann Surg Oncol 2012; 19: 2580-2589.

[18] Martínez-Jaimez P, Armora Verdú M, Forero CG, Álvarez Salazar S, Fuster Linares P, Monforte-Royo C and Masia J. Breast cancer-related lymphoedema: risk factors and prediction model. J Adv Nurs 2022; 78: 765-775.

[19] Liu X, Sun K, Yang H, Xia L, Lu K, Meng X and Li Y. Risk factors for the development of severe breast cancer-related lymphedema: a retrospective cohort study. BMC Cancer 2023; 23: 361.

[20] Bahanesteh A, Yarmohammadi H, Shahshenas S, Dalil D, Soltanipur M, Taheri H and Sheikhi Z. Risk factors for severity of breast cancer-related lymphedema. Radiat Oncol 2025; 20: 81.

[21] Lin C, Su J, Wu AJ, Lin N, Hossack MS, Shi W, Xu W, Liu FF and Kwan JYY. External validation of a 5-factor risk model for breast cancer-related lymphedema. JAMA Netw Open 2025; 8: e2455383.

[22] Shen A, Ye J, Zhao H, Qiang W, Zhao H, Huang Y, Zhou Y, Wang Y, Li X, Zhang Z, Bian J, Zhang L, Wu P, Wang Y and Lu Q. Risk factors and prediction model of breast cancer-related lymphoedema in a Chinese cancer centre: a prospective cohort study protocol. BMJ Open 2024; 14: e089769.
[23] Wei X, Lu Q, Jin S, Li F, Zhao Q, Cui Y, Jin S, Cao Y and Fu MR. Developing and validating a prediction model for lymphedema detection in breast cancer survivors. Eur J Oncol Nurs 2021; 54: 102023.
[24] Li MM, Wu PP, Qiang WM, Li JQ, Zhu MY, Yang XL and Wang Y. Development and validation of a risk prediction model for breast cancer-related lymphedema in postoperative patients with breast cancer. Eur J Oncol Nurs 2023; 63: 102258.
[25] Ahn HR, Jeong HE, Jeong C, Kang SY, Jung SH, Youn HJ and Kim JS. Incidence and risk factors of breast cancer-related lymphedema in Korea: a nationwide retrospective cohort study. Int J Surg 2024; 110: 3518-3526.
[26] Siotos C, Arnold SH, Seu M, Lunt L, Ferraro J, Najafali D, Damoulakis G, Vorstenbosch J, Mehrara BJ, Antony AK, Shenaq DS and Kokosis G. Breast cancer-related lymphedema: a comprehensive analysis of risk factors. J Surg Oncol 2024; 130: 1521-1531.
[27] Shen A, Wei X, Zhu F, Sun M, Ke S, Qiang W and Lu Q. Risk prediction models for breast cancer-related lymphedema: a systematic review and meta-analysis. Eur J Oncol Nurs 2023; 64: 102326.
[28] Lin Q, Yang T, Yongmei J and Die YM. Prediction models for breast cancer-related lymphedema: a systematic review and critical appraisal. Syst Rev 2022; 11: 217.

Table S1. Comparison of baseline characteristics between training and validation groups

Factor	Total	Validation Group (n=288)	Training Group (n=382)	Test Statistic	*P*-value	OR (95% CI)
Age				3.396	0.183	
<40	51 (7.6%)	17 (5.9%)	34 (8.9%)			Reference
40-59	480 (71.6%)	216 (75.0%)	264 (69.1%)			1.636 (0.890-3.010)
≥60	139 (20.7%)	55 (19.1%)	84 (22.0%)			1.310 (0.667-2.570)
BMI				0.336	0.562	
≥24	167 (24.9%)	75 (26.0%)	92 (24.1%)			0.901 (0.633-1.282)
<24	503 (75.1%)	213 (74.0%)	290 (75.9%)			
Education level				0.554	0.457	
≥ High school	213 (31.8%)	96 (33.3%)	117 (30.6%)			0.883 (0.636-1.225)
< High school	457 (68.2%)	192 (66.7%)	265 (69.4%)			
Marital status				0.228	0.633	
Married	647 (96.6%)	277 (96.2%)	370 (96.9%)			1.224 (0.532-2.816)
Other	23 (3.4%)	11 (3.8%)	12 (3.1%)			
Diabetes history				0.387	0.534	
Yes	31 (4.6%)	15 (5.2%)	16 (4.2%)			0.796 (0.387-1.637)
No	639 (95.4%)	273 (94.8%)	366 (95.8%)			
Hypertension history				0.796	0.372	
Yes	71 (10.6%)	27 (9.4%)	44 (11.5%)			1.258 (0.759-2.086)
No	599 (89.4%)	261 (90.6%)	338 (88.5%)			
Tumor size				0.099	0.753	
≥3	414 (61.8%)	176 (61.1%)	238 (62.3%)			1.052 (0.768-1.441)
<3	256 (38.2%)	112 (38.9%)	144 (37.7%)			
Lymph node metastasis				0.245	0.621	
Yes	316 (47.2%)	139 (48.3%)	177 (46.3%)			0.926 (0.681-1.257)
No	354 (52.8%)	149 (51.7%)	205 (53.7%)			
Tumor stage				0.633	0.889	
I	132 (19.7%)	58 (20.1%)	74 (19.4%)			Reference
II	318 (47.5%)	140 (48.6%)	178 (46.6%)			1.003 (0.667-1.510)
III	204 (30.4%)	83 (28.8%)	121 (31.7%)			0.875 (0.562-1.363)
IV	16 (2.4%)	7 (2.4%)	9 (2.4%)			0.992 (0.349-2.824)
Surgical type				0.000	0.984	
Mastectomy	614 (91.6%)	264 (91.7%)	350 (91.6%)			0.994 (0.572-1.728)
Breast-conserving surgery	56 (8.4%)	24 (8.3%)	32 (8.4%)			
Surgical side				0.055	0.815	
Left	378 (56.4%)	161 (55.9%)	217 (56.8%)			1.037 (0.762-1.412)
Right	292 (43.6%)	127 (44.1%)	165 (43.2%)			
Axillary lymph node dissection				0.160	0.689	
Yes	167 (24.9%)	74 (25.7%)	93 (24.3%)			0.931 (0.654-1.324)
No	503 (75.1%)	214 (74.3%)	289 (75.7%)			
Sentinel lymph node biopsy				0.293	0.588	
Yes	140 (20.9%)	63 (21.9%)	77 (20.2%)			0.902 (0.620-1.312)
No	530 (79.1%)	225 (78.1%)	305 (79.8%)			
Number of lymph nodes removed				0.788	0.375	
≥10	105 (15.7%)	41 (14.2%)	64 (16.8%)			1.212 (0.792-1.856)
<10	565 (84.3%)	247 (85.8%)	318 (83.2%)			
Postoperative complications				1.190	0.275	
Yes	63 (9.4%)	23 (8.0%)	40 (10.5%)			1.348 (0.787-2.306)
No	607 (90.6%)	265 (92.0%)	342 (89.5%)			
Radiotherapy				0.055	0.814	
Yes	119 (17.8%)	50 (17.4%)	69 (18.1%)			1.049 (0.703-1.567)
No	551 (82.2%)	238 (82.6%)	313 (81.9%)			

Endocrine therapy				0.094	0.760	
Yes	226 (33.7%)	99 (34.4%)	127 (33.2%)			0.951 (0.688-1.314)
No	444 (66.3%)	189 (65.6%)	255 (66.8%)			
Biologic therapy				0.512	0.474	
Yes	79 (11.8%)	31 (10.8%)	48 (12.6%)			1.191 (0.737-1.925)
No	591 (88.2%)	257 (89.2%)	334 (87.4%)			
Chemotherapy timing				0.000	0.987	
Neoadjuvant chemotherapy	65 (9.7%)	28 (9.7%)	37 (9.7%)			0.996 (0.594-1.669)
Adjuvant chemotherapy	605 (90.3%)	260 (90.3%)	345 (90.3%)			
Chemotherapy regimen				0.012	0.912	
Anthracycline plus taxane	490 (73.1%)	210 (72.9%)	280 (73.3%)			1.020 (0.722-1.439)
Other	180 (26.9%)	78 (27.1%)	102 (26.7%)			
Chemotherapy duration				4.566	0.102	
4 weeks	490 (73.1%)	214 (74.3%)	276 (72.3%)			Reference
6 weeks	161 (24.0%)	62 (21.5%)	99 (25.9%)			0.808 (0.561-1.163)
8 weeks	19 (2.8%)	12 (4.2%)	7 (1.8%)			2.211 (0.856-5.712)
Chemotherapy vascular access device				1.416	0.493	
PICC	182 (27.2%)	85 (29.5%)	97 (25.4%)			Reference
IVAP	419 (62.5%)	174 (60.4%)	245 (64.1%)			0.810 (0.571-1.150)
PVC	69 (10.3%)	29 (10.1%)	40 (10.5%)			0.827 (0.473-1.448)

Note: OR, Odds ratio; CI, Confidence interval; PICC, Peripherally inserted central catheter; IVAP, Implantable venous access port; PVC, Peripheral venous catheter.