https://doi.org/10.1245/s10434-026-20309-9 Annals of Surgical Oncology

Using Machine Learning to Predict Breast Cancer-Related Lymphedema Following Axillary Lymph Node Dissection

Benjamin D. Wagner, MPH[1] · Jonlin Chen, MD[1] · Lillian A. Boe, PhD[2] · Francis D. Graziano, MD[1] · Arielle N. Roberts, DO, MS[1] · Andrea V. Barrio, MD[3] · Michelle R. Coriddi, MD[1] · Jonas A. Nelson, MD, MPH[1] · Babak J. Mehrara, MD[1] · Danielle H. Rochlin, MD, MPH[1]

Received: 26 May 2026 / Accepted: 16 July 2026

Abstract

Background Breast cancer-related lymphedema (BCRL) is a common and debilitating sequela of axillary lymph node dissection (ALND). Although machine learning (ML)-based prediction models have been proposed, few focus exclusively on patients undergoing ALND, and direct comparisons with traditional statistical models remain limited. This study aimed to develop accurate and clinically feasible prediction models for BCRL using supervised ML and multivariable logistic regression.

Methods Demographic and clinical data were prospectively collected from women undergoing unilateral ALND for breast cancer at Memorial Sloan Kettering Cancer Center between 2016 and 2024. Supervised ML and multivariable logistic regression models to predict BCRL were trained and internally validated. Model performance was evaluated using area under the receiver operator characteristic curve (AUC), accuracy, sensitivity, specificity, and Brier score. Shapley additive explanations were used for model interpretability.

Results A total of 474 eligible patients were included. BCRL developed in 113 (23.8%) patients at a mean ± standard deviation of 16.6 ± 7.5 months postoperatively. The highest-performing ML model (random forest) achieved an AUC of 0.83, whereas traditional multivariable logistic regression achieved an optimism-corrected AUC of 0.62. Key ML predictors of BCRL on Shapley additive explanations analysis included clinical cancer stage, body mass index, age, and neoadjuvant chemotherapy.

Conclusions ML-based models outperformed traditional logistic regression in predicting BCRL among patients undergoing ALND. These models demonstrate the potential of ML for early BCRL identification and risk stratification but also highlight the difficulties in accurately predicting BCRL development. Further research is needed to improve model predictive performance and facilitate clinical implementation.

Keywords Breast cancer-related lymphedema · Machine learning · Risk prediction · Artificial Intelligence · Axillary lymph node dissection

Breast cancer-related lymphedema (BCRL) is a common and debilitating sequela of breast cancer treatment, particularly following axillary lymph node dissection (ALND).[1] BCRL is characterized by chronic limb swelling due to lymphatic dysfunction and significantly impairs quality of life, physical function, and psychological wellbeing.[2–4] Estimates suggest that over 20% of patients undergoing ALND develop BCRL, with incidence rates exceeding 50% when combined with regional radiotherapy and/or taxane-based chemotherapy.[1,5,6] In contrast, patients undergoing sentinel lymph node biopsy (SLNB) alone exhibit markedly lower rates of BCRL, typically ranging from 3% to 5%.[7–9] Delayed diagnosis and

Meetings and Presentations: Northeastern Society of Plastic Surgeons 42nd Annual Meeting; Baltimore, MD, USA; September 27, 2025.

✉ Danielle H. Rochlin, MD, MPH
rochlind@mskcc.org

[1] Plastic and Reconstructive Surgery Service, Department of Surgery, Memorial Sloan Kettering Cancer Center, New York, NY

[2] Department of Epidemiology and Biostatistics, Memorial Sloan Kettering Cancer Center, New York, NY

[3] Breast Service, Department of Surgery, Memorial Sloan Kettering Cancer Center, New York, NY

intervention are associated with recurrent infections, chronic disability, and increased healthcare utilization.[10,11] Consequently, accurate and early identification of patients undergoing ALND and at the highest risk for BCRL is critical to preventing disease progression and reducing long-term morbidity.

Several demographic and clinical risk factors have been associated with an increased risk of BCRL following ALND, including advanced patient age,[12–14] elevated body mass index (BMI),[9,15,16] extensive lymph node removal,[5,6,9,17,18] receipt of neoadjuvant chemotherapy,[6,17,19] and receipt of regional nodal irradiation.[1,5,20,21] Racial and ethnic disparities have also been reported, with Black and Hispanic patients experiencing higher rates of BCRL than white patients.[12,22,23] Despite recognition of these risk factors, accurately predicting which individuals will develop BCRL remains a challenge. Although previous predictive models,[5,24–38] including our own group's recently published model that included patients undergoing both ALND and SLNB,[23] have aimed to address limitations such as variable exclusion, lack of standardized diagnostic criteria, and suboptimal performance, significant gaps remain. Specifically, combining ALND and SLNB patients into a single model tends to overemphasize ALND as a dominant predictor, potentially masking the impact of other critical variables that influence risk. Given that ALND patients are at highest risk of developing BCRL, focusing interventions such as immediate lymphatic reconstruction (ILR) or intensive screening strategies on this population is more practical from both a resource allocation and clinical implementation standpoint.

Most existing predictive models for BCRL have relied on traditional statistical methods such as logistic regression. Although widely used, these approaches are often ill-equipped to handle complex, high-dimensional data, limiting their ability to capture nonlinear associations across diverse clinical and patient-reported variables. Machine learning (ML) offers a promising alternative, with algorithms such as random forest and gradient boosting capable of modeling intricate interactions among multiple predictors and potentially improving predictive accuracy. Although early ML-based prediction models for BCRL have been proposed, few have been developed specifically for patients undergoing ALND, and direct comparisons with traditional statistical models remain limited.[30,39,40] The primary aim of this study was to develop traditional statistical and ML models to predict the risk of developing BCRL specifically among patients undergoing ALND. The secondary aim was to compare model performance and to identify key demographic and treatment-related risk factors for BCRL. We hypothesized that ML models would outperform traditional logistic regression in predicting BCRL risk and enable more accurate identification of high-risk ALND patients based on routinely collected clinical and demographic variables.

Methods

Database Collection

This study was approved by the institutional review board at Memorial Sloan Kettering Cancer Center, and written informed consent was obtained from all participants. Female patients with breast cancer aged ≥18 years who underwent unilateral ALND at the study institution between November 2016 and March 2024 were identified from two prospectively maintained datasets. Patients were excluded if they were male, had undergone bilateral axillary surgery, had received ILR or other lymphatic reconstructive procedures, or had <6 months of follow-up from the date of ALND.

Model variables included preoperative patient characteristics to facilitate risk counseling before surgery. Variable selection was guided by exploratory ML analyses, univariate logistic regression, and expert clinical input. The final variables included age at ALND, BMI, self-reported race, smoking status, type of breast surgery (lumpectomy or mastectomy), receipt of neoadjuvant chemotherapy, clinical cancer stage, and preexisting medical comorbidities. Comorbidities included hypertension, diabetes mellitus, hyperlipidemia, and a history of psychiatric diagnosis. A psychiatric history was defined by diagnoses classified in the International Classification of Diseases (ICD), Tenth Revision (codes beginning with F) or Ninth Revision (codes 290–319.99), including anxiety, mood, substance-related, somatic, sexual/gender identity, personality, behavioral and neurodevelopmental, psychotic, sleep, and eating disorders. Only psychiatric diagnoses documented prior to ALND were included. These comorbidities were included to account for systemic health conditions and psychosocial factors that may influence immune function, wound healing, postoperative recovery, and adherence to lymphedema prevention behaviors.

These data were used to develop and evaluate predictive models for BCRL risk using both traditional statistical and ML approaches. To ensure logistic regression model stability and avoid overfitting, the number of included variables was limited based on the number of BCRL events, following the convention of at least 10 events per predictor. Variables contributing minimally to model performance or introducing collinearity were excluded to optimize interpretability and generalizability. The same predictor variables were used across traditional statistical and ML models to ensure consistency in performance comparisons.

Outcome to Predict

The primary outcome of this study was the development of BCRL, defined as a relative volume change (RVC) of

≥10% occurring at least 6 months following ALND. Bilateral limb volume measurements were recorded at baseline (prior to ALND) and every 6 months for up to 3 years postoperatively. Limb volumes were obtained using either perometry (Perometer 350 NT, Pero-System, Wuppertal, Germany) or manual circumferential measurements.[41] For manual assessments, limb circumferences were recorded at 4 cm intervals from the ulnar styloid to 40–44 cm proximally using a non-elastic measuring tape, and limb volumes were calculated using the truncated cone formula.[42] RVC was calculated using the following validated formula: RVC = [(A2U1)/(U2A1)] − 1, where A1 and A2 represent volumes of the affected arm at baseline and follow-up, and U1 and U2 represent corresponding volumes of the unaffected arm.[43]

Traditional Statistical Model Development and Evaluation

Multivariable logistic regression models were developed to predict the risk of BCRL after ALND. Results are presented as odds ratios (ORs) with 95% confidence intervals (CIs). Model discrimination was evaluated using the area under the receiver operating characteristic (ROC) curve (AUC), first calculated from internal validation and then corrected for optimism through 1000 bootstrap resamples.[44] Calibration was assessed using calibration plots comparing predicted and observed event probabilities.

ML Model Development and Evaluation

Supervised ML models were developed using the same dataset and variables as the logistic regression model. Data were preprocessed with one-hot encoding of categorical variables and standard scaling of numeric variables. The dataset was split into training, validation, and testing sets in a 60:20:20 ratio, as previously described.[45] Using the training and validation sets, five ML models (XGBoost, AdaBoost, LightGBM, random forest, and logistic regression) were developed using stratified 10-fold cross-validation.[46–49] Briefly, random forest constructs an ensemble of decision trees trained on bootstrapped subsamples and averages their predictions, offering robustness to overfitting and strong performance on tabular clinical data. XGBoost and LightGBM are gradient boosting frameworks that iteratively build decision trees to minimize residual prediction errors, with LightGBM optimized for computational efficiency on large datasets. AdaBoost is a boosting approach that reweights misclassified samples across successive weak learners, whereas logistic regression serves as a linear baseline with high interpretability but limited capacity to capture nonlinear interactions. These models were selected to represent a range of explainability, complexity, and inference time.[50] To address class imbalance, the synthetic minority oversampling technique was applied prior to model training.[51] Hyperparameter tuning was performed using Optuna optimization during cross-validation.[52]

Model performance was assessed on the independent test set using accuracy, AUC, sensitivity, and specificity. Model calibration was evaluated using the Brier score.[45] Point estimates and 95% CIs were calculated for all metrics using bootstrap resampling. AUCs were compared across models using the Wald test, and ROC curves were generated to visualize performance.[53] To enhance interpretability, Shapley additive explanations (SHAP) values were computed for the top-performing decision tree-based model.[54]

Statistical Analysis

Data were analyzed from February to August 2025. Demographic and clinical characteristics were summarized using descriptive statistics. Continuous variables were presented as means ± standard deviations (SD), whereas categorical variables were reported as frequencies and percentages. These variables were summarized both overall and by BCRL diagnosis. Continuous variables across BCRL diagnosis were compared using Welch's two-sample *t*-tests, and categorical variables were compared using Pearson's chi-squared tests and Fisher's exact tests. All ML analyses were conducted using Python version 3.8 (Python Software Foundation) with Scikit-learn, Optuna, and SHAP libraries. Traditional statistical analyses were performed in R version 4.3.2 (R Foundation for Statistical Computing). A two-sided *P* value <0.05 was considered statistically significant.

Results

Patient Characteristics

A total of 474 female patients with breast cancer treated with ALND were included for analysis, with a mean ± SD age of 48.4 ± 11.2 years and a mean BMI of 26.4 ± 5.6 kg/m^2 (Table 1). Of the total cohort, 300 (63.3%) patients self-identified as white, 82 (17.3%) as Black, 48 (10.1%) as Asian, and 44 (9.3%) as another race (including American Indian and Alaska Native, other race, or unknown). The mean ± SD follow-up time from ALND to the last recorded follow-up was 24.4 ± 8.8 months. Overall, 113 patients (23.8%) developed BCRL, with a mean ± SD time to diagnosis of 16.6 ± 7.5 months following ALND.

Table 1 Patient demographic and clinical data for the study cohort with univariable comparison by breast cancer-related lymphedema (BCRL) diagnosis

Characteristic	Overall N = 474	BCRL N = 113	No BCRL N = 361	*P* value
Age at ALND, years	48.4 ± 11.2	50.3 ± 11.4	47.8 ± 11.1	0.048*
BMI, kg/m^2	26.35 ± 5.63	27.88 ± 5.92	25.87 ± 5.46	0.002*
Race[a]				0.011*
Asian	48 (10.1)	8 (7.1)	40 (11.1)	
Black or African American	82 (17.3)	31 (27.4)	51 (14.1)	
Other	44 (9.3)	10 (8.8)	34 (9.4)	
White	300 (63.3)	64 (56.6)	236 (65.4)	
Ethnicity				0.8
Not Hispanic	424 (89.5)	100 (88.5)	324 (89.8)	
Hispanic	39 (8.2)	10 (8.8)	29 (8.0)	
Unknown	11 (2.3)	3 (2.7)	8 (2.2)	
Hypertension	125 (26.4)	41 (36.3)	84 (23.3)	0.006*
Hyperlipidemia	80 (16.9)	28 (24.8)	52 (14.4)	0.010*
Diabetes mellitus	46 (9.7)	17 (15.0)	29 (8.0)	0.028*
Psychiatric history	173 (36.5)	40 (35.5)	133 (36.8)	0.8
Breast surgery				0.3
Lumpectomy	98 (20.7)	27 (23.9)	71 (19.7)	
Mastectomy	376 (79.3)	86 (76.1)	290 (80.3)	
LNs removed	19.0 ± 8.4	19.4 ± 9.4	18.9 ± 8.0	0.6
LNs positive	3.6 ± 6.6	4.2 ± 6.9	3.4 ± 6.5	0.3
Neoadjuvant chemotherapy	299 (63.1)	85 (75.2)	214 (59.3)	0.002*
Adjuvant chemotherapy	218 (46.0)	38 (33.6)	180 (49.9)	0.003*
Neoadjuvant hormone therapy	28 (5.9)	6 (5.3)	22 (6.1)	0.8
Adjuvant hormone therapy	371 (78.3)	81 (71.7)	290 (80.3)	0.052
Adjuvant radiotherapy	419 (88.4)	109 (96.5)	310 (85.9)	0.002*
Smoking status				0.6
Never smoker	389 (82.1)	92 (81.4)	297 (82.3)	
Current/Former smoker	58 (12.2)	16 (14.2)	42 (11.6)	
Unknown	27 (5.7)	5 (4.4)	22 (6.1)	
Clinical cancer stage				0.14
I–II	255 (53.8)	52 (46.0)	203 (56.2)	
III–IV	202 (42.6)	56 (49.6)	146 (40.4)	
Unknown	17 (3.6)	5 (4.4)	12 (3.3)	

Data are presented as N (%) or mean ± standard deviation unless otherwise indicated

ALND, axillary lymph node dissection; BMI, body mass index; LN, lymph node

[a]Race data were self-reported. The other race category included American Indian and Alaksa Native, other race, patient refused to disclose, or unknown

*Statistically significant value

Traditional Statistical Model Performance

The multivariable logistic regression model demonstrated modest discriminative performance, with an original AUC of 0.68 (95% CI 0.62–0.73). Upon resampling-based bootstrap correction, the optimism-adjusted AUC was 0.62 (95% CI 0.59–0.72). Factors independently associated with the development of BCRL included Black race (OR 1.82; 95% CI 1.01–3.27; $P = 0.046$) and receipt of neoadjuvant chemotherapy (OR 1.83; 95% CI 1.11–3.08; $P = 0.021$) (Supplemental Digital Content, Table 1). The calibration plot for the logistic regression model is presented in Fig. 1.

ML Model Performance and Interpretability

ML models demonstrated variable discriminative performance, with AUCs ranging from 0.50 to 0.83. Random forest was the highest-performing ML model, achieving an AUC of 0.83 (95% CI 0.60–0.84) (Table 2). ROCs for each model are shown in Fig. 2. Accuracy, sensitivity, specificity,

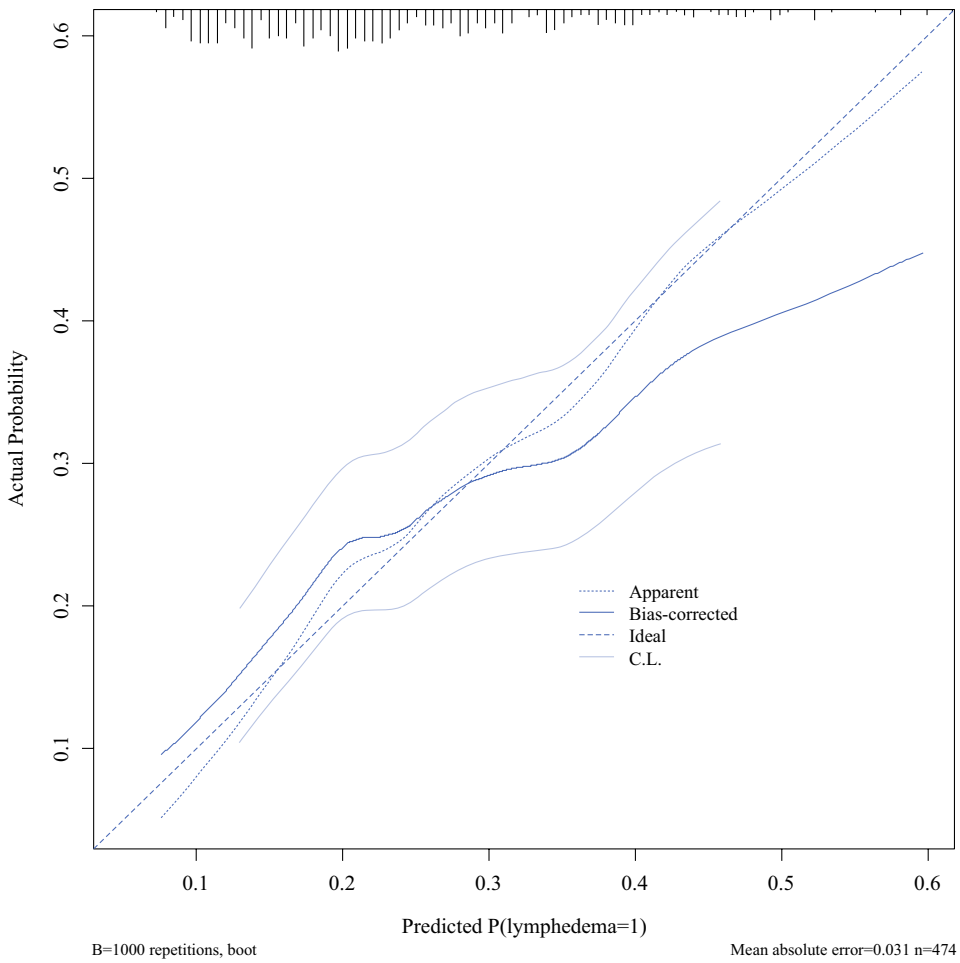

Fig. 1 Calibration curve for the multivariable logistic regression breast cancer-related lymphedema prediction model

Table 2 Evaluation metrics of all machine learning (ML) models for prediction of breast cancer-related lymphedema

ML algorithm	Accuracy	AUC	Sensitivity	Specificity	Brier score
Random forest	0.75 (0.66–0.84)	0.83 (0.60–0.84)	0.43 (0.02–0.84)	0.92 (0.94–1.00)	0.17 (0.10–0.24)
Logistic regression	0.75 (0.66–0.84)	0.81 (0.65–0.86)	0.39 (0.29–0.49)	0.86 (0.79–0.93)	0.25 (0.16–0.34)
LightGBM	0.68 (0.59–0.78)	0.67 (0.46–0.72)	0.22 (0.13–0.30)	0.83 (0.76–0.91)	0.21 (0.13–0.30)
AdaBoost	0.59 (0.49–0.69)	0.66 (0.49–0.75)	0.57 (0.47–0.67)	0.60 (0.50–0.70)	0.23 (0.14–0.31)
XGBoost	0.74 (0.65–0.83)	0.50 (0.47–0.74)	0.26 (0.17–0.35)	0.89 (0.83–0.95)	0.21 (0.13–0.30)

Data in parentheses are 95% confidence intervals

AUC, area under the receiver operating characteristic curve

and Brier scores varied across ML models (Table 2). Model interpretability was evaluated using SHAP values derived from top-performing decision tree models, which quantified the relative contribution of individual predictors to model output (Fig. 3). Clinical cancer stage, BMI, age at ALND, and receipt of neoadjuvant chemotherapy were the most influential features driving model predictions.

Discussion

In this study, we developed and internally validated a clinically feasible prediction model for BCRL among women undergoing ALND using both traditional logistic regression and ML approaches. The ML-based model demonstrated

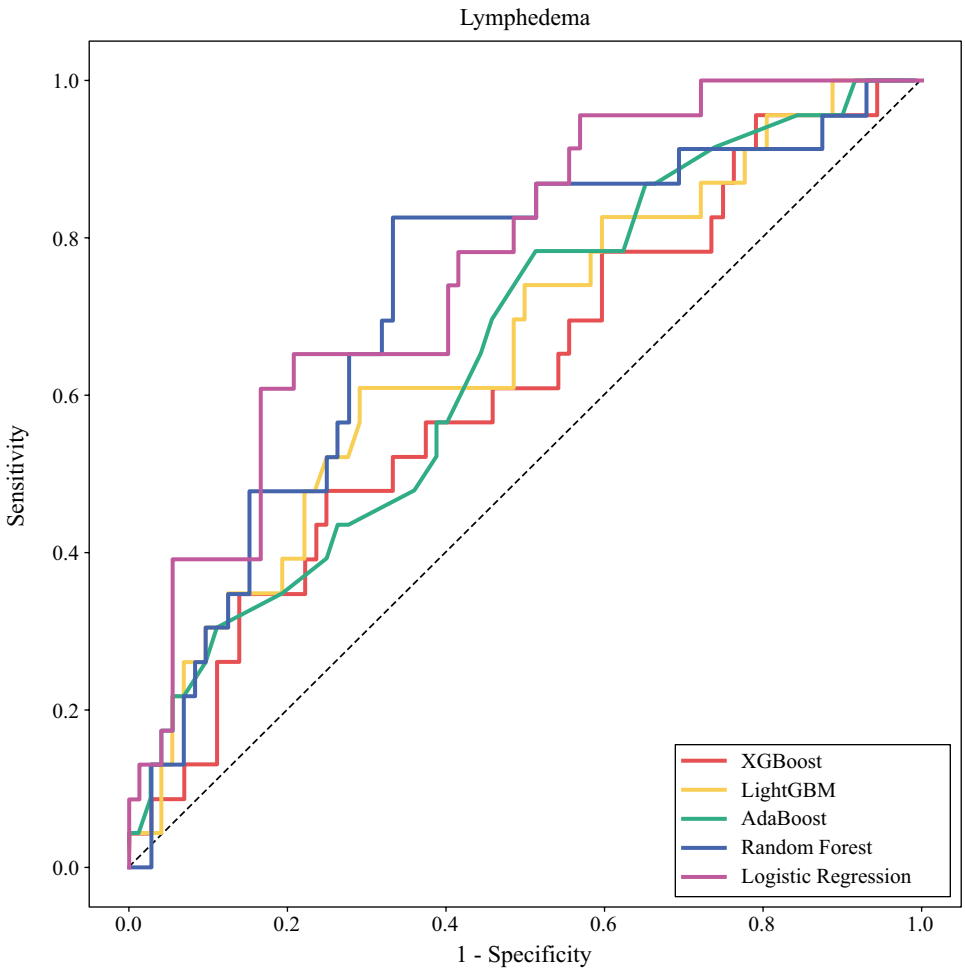

Fig. 2 Receiver operator characteristic curves across all machine learning algorithms for breast cancer-related lymphedema prediction

strong discriminative performance and outperformed traditional logistic regression, achieving an AUC of 0.83 (95% CI 0.60–0.84). These results compare favorably with previously published BCRL prediction models, which have reported AUCs ranging from 0.68 to 0.89 on internal validation or testing.[23,34] Our findings highlight the potential of ML to enhance predictive accuracy for complex postoperative complications such as BCRL, in which nonlinear associations, variable interactions, and multifactorial drivers of disease may limit the performance of conventional statistical approaches.[55,56]

Numerous BCRL prediction models have been published,[5,23–38] yet many exhibit limitations that constrain clinical utility. Common issues include reliance on retrospective data via ICD-9 and ICD-10 diagnosis codes[37] or subjective symptom reports to define lymphedema.[5,24] Several existing models also incorporate variables that are harder to obtain in clinical settings, such as mammographic breast density,[27] proportion of lymphatic flow above the axillary vein,[26] or the number of chemotherapy cycles delivered to the affected arm.[38] Additionally, many models omit key clinical predictors that have previously been established to impact BCRL risk, such as race and ethnicity.[12,22] While our group's prior model addressed several of these shortcomings,[23] it included both SLNB and ALND patients, which results in ALND status dominating as a predictor and potentially masking the influence of other clinically relevant variables. In contrast, the current model was developed exclusively for patients undergoing ALND, the population at highest risk for developing BCRL. Narrowing our predictive modeling to this high-risk group not only enhanced the identification of critical risk factors but also made targeted screening and intervention strategies more feasible from a resource allocation perspective. Importantly, our model uses standardized, objective definitions of lymphedema and relies solely on variables available preoperatively, making it suitable for risk counseling and shared decision-making before surgery. The RVC threshold of ≥10% used in this study is one of the most widely adopted volumetric criteria in prospective surveillance programs and clinical trials, supporting its use as a standardized, reproducible outcome measure.

Model variables included demographic, clinical, and treatment-related factors previously associated with BCRL risk, such as age,[12–14] BMI,[9,15,16] and receipt of neoadjuvant

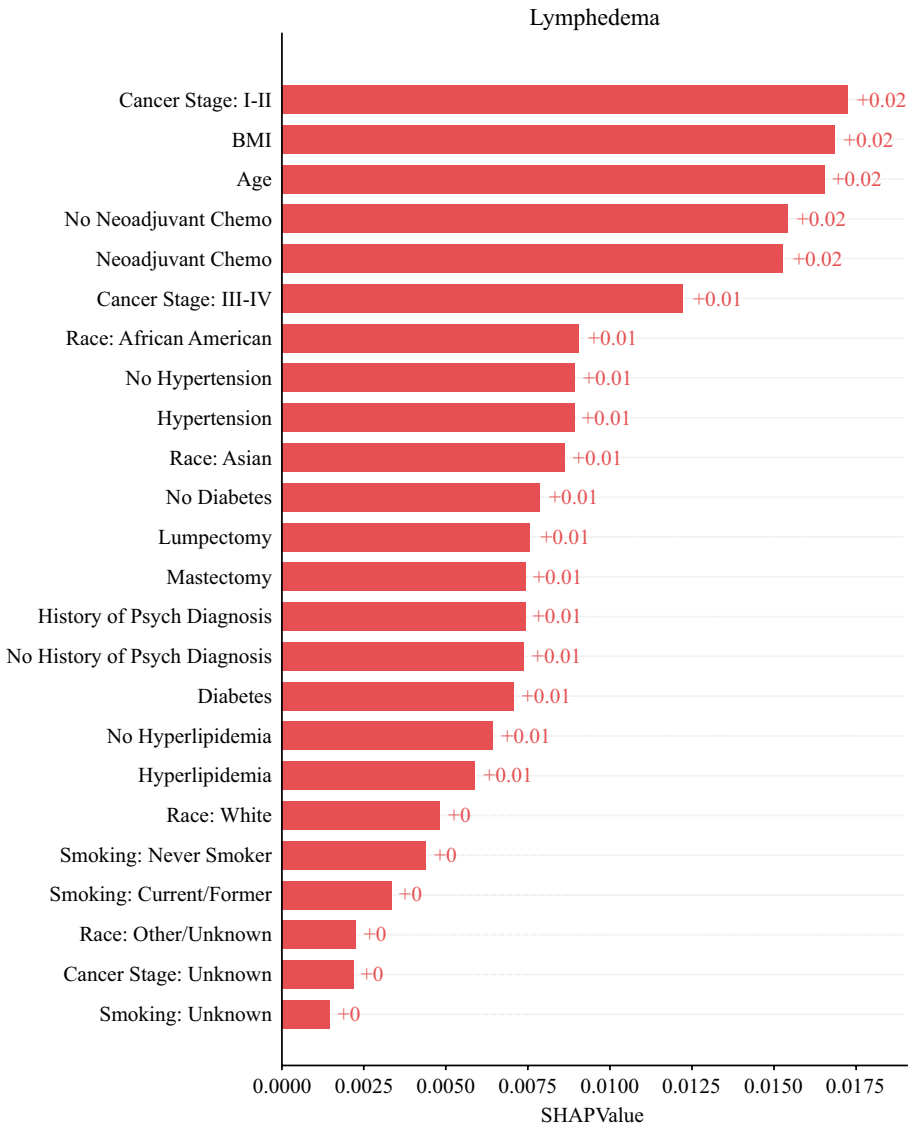

Fig. 3 Feature importance of top performing decision tree machine learning breast cancer-related lymphedema prediction model. Features are ranked in descending order of importance by mean absolute Shapley additive explanation (SHAP) values. BMI, body mass index.

chemotherapy.[6,12,15,19] SHAP analysis identified clinical cancer stage, BMI, age, and receipt of neoadjuvant chemotherapy as the strongest contributors to model predictions, consistent with prior evidence. Notably, Black race was independently associated with an increased risk of BCRL in the multivariable logistic regression model, aligning with our group's previous work and other studies demonstrating racial disparities in lymphedema.[12,22,23,57] These differences likely reflect a combination of biological susceptibility and structural inequalities in care.[58–60] To our knowledge, apart from the model presented in this study, our previously published models remain the only BCRL risk prediction tools that incorporate race and ethnicity as independent predictors and report race-specific risk estimates.[23] By including racially diverse predictors, our models may help reduce, rather than perpetuate, disparities in lymphedema prevention and management.

Although logistic regression remains a valuable tool for hypothesis testing and clinical inference, ML offers distinct advantages for modeling high-dimensional, nonlinear interactions that may otherwise go undetected.[61] The integration of ML and traditional statistical approaches in our study enhanced both model predictive performance and interpretability, allowing for individualized risk estimation and greater transparency regarding the influence of specific predictors. The application of ML to BCRL prediction remains nascent, and this study represents one of the first efforts to demonstrate both ML and logistic regression models using identical input variables within a well-defined ALND cohort. Prior ML efforts have typically involved heterogeneous patient populations, relied on administrative codes or self-reported symptoms to define lymphedema, and often lacked direct comparisons with traditional methods.[30,39,40,62–65] In addition, many models incorporated variables not easily

available in clinical settings, thereby limiting clinical applicability. Although early ML models have shown promise in predicting surgical risks,[66,67] a comprehensive and clinically feasible model for real-time BCRL prediction following ALND has not yet been established.

Although AUC is the most widely reported performance metric in prior BCRL predictive models, we evaluated several additional measures to provide a more comprehensive assessment of model performance.[68] These included accuracy, sensitivity, specificity, and Brier score, each offering complementary insights into model discrimination and calibration. The top-performing ML prediction model achieved an AUC of 0.83 and an overall accuracy of 0.75, with high specificity and modest sensitivity. This pattern indicates that the model performs well in correctly identifying patients who do not develop BCRL, thereby minimizing false-positive classifications, although at the expense of reduced sensitivity for detecting true cases. The Brier score of 0.17 confirmed acceptable calibration. Together, these results emphasize the importance of multidimensional model evaluation when developing tools intended to guide clinical decision-making. Depending on clinical priorities, models may be optimized to favor sensitivity in preventive contexts or specificity when resource constraints necessitate targeted surveillance.

Predictive modeling plays a critical role in identifying patients with breast cancer at highest risk for developing lymphedema and guiding both primary and secondary prevention efforts. The single predictive model developed in this study may be particularly useful for estimating individualized risk before surgery and identifying patients who could benefit from preventive interventions such as ILR or early referral to complex decongestive therapy. Although ILR has demonstrated efficacy and cost effectiveness in selected patients,[69–71] it is resource intensive and may not be appropriate for all. Risk stratification based on individualized preoperative estimates could help prioritize those most likely to benefit while reducing unnecessary interventions. Given the simplicity and accessibility of the model's inputs, it could be incorporated into clinical workflows, electronic health records, or decision-support platforms to enable real-time BCRL risk prediction and support personalized surgical planning.

Despite encouraging results, predicting BCRL remains inherently challenging, particularly in the context of evolving treatment paradigms and multidisciplinary decision-making. Even within a uniform ALND population, individual risk is shaped by a complex interplay of demographic, biological, and treatment-related factors, many of which are not known until after surgery.[72] Although rates of ALND have declined with the broader adoption of targeted axillary dissection, SLNB, and nodal radiotherapy, this trend underscores rather than diminishes the value of accurate risk stratification in the patients who do require ALND, who remain at highest risk for BCRL. Whether models developed in ALND cohorts can be adapted for lower-risk surgical populations, including those receiving neoadjuvant systemic therapy, is an important question for future investigation. The wide CIs observed for the AUCs, particularly those from ML-based estimates, suggest a substantial degree of unexplained variability in model performance. These findings highlight the need for external validation in larger, multisite cohorts to improve generalizability, refine calibration, and optimize performance. Future models may benefit from integrating more granular data sources, such as genomic profiles, lab work, imaging, or patient-reported outcomes.

Study Limitations

This study has several limitations. It was conducted at a single high-volume cancer center, which may limit the generalizability of findings to other institutions and patient populations. The sample size was modest, with a limited number of BCRL events, potentially reducing statistical power and contributing to imprecision in performance estimates. Nevertheless, this remains one of the largest cohorts to date in which BCRL was diagnosed using prospectively collected arm volume measurements, making it comparatively robust relative to prior studies. External validation was not performed because no independent cohort with prospectively collected limb volume measurements using the same diagnostic criteria was available. In addition, the current model was designed to predict BCRL development as a binary outcome and does not account for disease severity, which may have important implications for risk stratification and the tailoring of preventive interventions. These results warrant further investigation and underscore the complexity of BCRL risk prediction in the context of modern, multimodal cancer treatment. Future work should focus on validating these models in larger, more diverse patient cohorts to enhance model performance and support broader clinical application.

Conclusions

We developed and internally validated a clinically relevant prediction model for BCRL using both traditional logistic regression and ML approaches. The ML-based model outperformed logistic regression in predicting BCRL among patients undergoing ALND, demonstrating the potential of ML approaches to enhance early risk stratification and individualized prediction. By incorporating readily available, patient-facing variables, the model provides a practical tool to support personalized prevention strategies, perioperative counseling, and postoperative surveillance. At the same time, these findings highlight the challenges

inherent to accurately predicting BCRL development. External validation in larger, multi-institutional cohorts is warranted to improve model performance and facilitate broader clinical implementation.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1245/s10434-026-20309-9.

Disclosure This research was supported in part by the National Institutes of Health through grants R01HL111130, R01CA278599, R21AR081076, and R21AG076132 awarded to Dr. Mehrara; the Plastic Surgery Foundation Research Fellowship Grant awarded to Dr. Chen; the Chanel Endowment for Survivorship Research and the Manhasset Women's Coalition Against Breast Cancer awarded to Dr. Barrio; and the Cancer Center Support Grant P30CA008748 that supports the research infrastructure at Memorial Sloan Kettering Cancer Center. Dr. Nelson has served as a consultant for RTI Surgical. Dr. Mehrara has received investigator-initiated grants from Pfizer, Regeneron, Atyr, and Integra; royalties from PureTech and Elsevier; and serves as a consultant for Mediflix Corp. Dr. Barrio has received speaking honoraria from MSD and Novartis. The other authors have no disclosures to report.

References

1. Naoum GE, Roberts S, Brunelle CL, et al. Quantifying the Impact of Axillary Surgery and Nodal Irradiation on Breast Cancer-Related Lymphedema and Local Tumor Control: Long-Term Results From a Prospective Screening Trial. *J Clin Oncol*. 2020;38(29):3430–8.
2. Taghian NR, Miller CL, Jammallo LS, O'Toole J, Skolny MN. Lymphedema following breast cancer treatment and impact on quality of life: a review. *Crit Rev Oncol Hematol*. 2014;92(3):227–34.
3. Togawa K, Ma H, Smith AW, et al. Self-reported symptoms of arm lymphedema and health-related quality of life among female breast cancer survivors. *Sci Rep*. 2021;11(1):10701.
4. Jørgensen MG, Toyserkani NM, Hansen FG, Bygum A, Sørensen JA. The impact of lymphedema on health-related quality of life up to 10 years after breast cancer treatment. *NPJ Breast Cancer*. 2021;7(1):70.
5. Byun HK, Chang JS, Im SH, et al. Risk of Lymphedema Following Contemporary Treatment for Breast Cancer: An Analysis of 7617 Consecutive Patients From a Multidisciplinary Perspective. *Ann Surg*. 2021;274(1):170–8.
6. Armer JM, Ballman KV, McCall L, et al. Factors Associated With Lymphedema in Women With Node-Positive Breast Cancer Treated With Neoadjuvant Chemotherapy and Axillary Dissection. *JAMA Surg*. 2019;154(9):800–9.
7. McLaughlin SA, Wright MJ, Morris KT, et al. Prevalence of lymphedema in women with breast cancer 5 years after sentinel lymph node biopsy or axillary dissection: objective measurements. *J Clin Oncol*. 2008;26(32):5213–9.
8. Galimberti V, Cole BF, Viale G, et al. Axillary dissection versus no axillary dissection in patients with breast cancer and sentinel-node micrometastases (IBCSG 23–01): 10-year follow-up of a randomised, controlled phase 3 trial. *Lancet Oncol*. 2018;19(10):1385–93.
9. DiSipio T, Rye S, Newman B, Hayes S. Incidence of unilateral arm lymphoedema after breast cancer: a systematic review and meta-analysis. *Lancet Oncol*. 2013;14(6):500–15.
10. Whitworth P, Vicini F, Valente SA, et al. Reducing rates of chronic breast cancer-related lymphedema with screening and early intervention: an update of recent data. *J Cancer Surviv*. 2024;18(2):344–51.
11. Rafn BS, Christensen J, Larsen A, Bloomquist K. Prospective Surveillance for Breast Cancer-Related Arm Lymphedema: A Systematic Review and Meta-Analysis. *J Clin Oncol*. 2022;40(9):1009–26.
12. Montagna G, Zhang J, Sevilimedu V, et al. Risk Factors and Racial and Ethnic Disparities in Patients With Breast Cancer-Related Lymphedema. *JAMA Oncol*. 2022;8(8):1195–200.
13. Thomis S, Devoogdt N, Bechter-Hugl B, Fourneau I. Early Disturbance of Lymphatic Transport as a Risk Factor for the Development of Breast-Cancer-Related Lymphedema. *Cancers (Basel)*. 2023;15(6):1774.
14. Ren Y, Kebede MA, Ogunleye AA, et al. Burden of lymphedema in long-term breast cancer survivors by race and age. *Cancer*. 2022;128(23):4119–28.
15. Koelmeyer LA, Gaitatzis K, Dietrich MS, et al. Risk factors for breast cancer-related lymphedema in patients undergoing 3 years of prospective surveillance with intervention. *Cancer*. 2022;128(18):3408–15.
16. Wang L, Chen H, Li Y, et al. Body mass index increases the risk of breast cancer-related lymphedema at 6–18 months after surgery: a retrospective study. *Support Care Cancer*. 2023;31(5):278.
17. Fu J, Chen R, He L, et al. Factors affecting lymphedema after neoadjuvant chemotherapy and axillary dissection in female breast cancer patients: a retrospective cohort study based on the Chinese population. *Front Oncol*. 2024;14:1436748.
18. Shen A, Lu Q, Fu X, et al. Risk factors of unilateral breast cancer-related lymphedema: an updated systematic review and meta-analysis of 84 cohort studies. *Support Care Cancer*. 2022;31(1):18.
19. Jung SY, Shin KH, Kim M, et al. Treatment factors affecting breast cancer-related lymphedema after systemic chemotherapy and radiotherapy in stage II/III breast cancer patients. *Breast Cancer Res Treat*. 2014;148(1):91–8.
20. Shaitelman SF, Chiang YJ, Griffin KD, et al. Radiation therapy targets and the risk of breast cancer-related lymphedema: a systematic review and network meta-analysis. *Breast Cancer Res Treat*. 2017;162(2):201–15.
21. Chandra RA, Miller CL, Skolny MN, et al. Radiation therapy risk factors for development of lymphedema in patients treated with regional lymph node irradiation for breast cancer. *Int J Radiat Oncol Biol Phys*. 2015;91(4):760–4.
22. Black DM, Jiang J, Kuerer HM, Buchholz TA, Smith BD. Racial disparities in adoption of axillary sentinel lymph node biopsy and lymphedema risk in women with breast cancer. *JAMA Surg*. 2014;149(8):788–96.
23. Rochlin DH, Barrio AV, McLaughlin S, et al. Feasibility and Clinical Utility of Prediction Models for Breast Cancer-Related Lymphedema Incorporating Racial Differences in Disease Incidence. *JAMA Surg*. 2023;158(9):954–64.
24. Kim M, Kim SW, Lee SU, et al. 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(3):498–503.
25. Gross JP, Whelan TJ, Parulekar WR, et al. Development and Validation of a Nomogram to Predict Lymphedema After Axillary Surgery and Radiation Therapy in Women With Breast Cancer From the NCIC CTG MA.20 Randomized Trial. *Int J Radiat Oncol Biol Phys*. 2019;105(1):165–73.
26. Yuan Q, Hou J, Zhou R, et al. Development and Validation of an Intraoperative Nomogram to Predict Breast Cancer-Related Lymphedema Based on the Arm Lymphatics Distribution. *Ann Surg Oncol*. 2021;28(12):7319–28.

27. Kwan JYY, Famiyeh P, Su J, et al. Development and Validation of a Risk Model for Breast Cancer-Related Lymphedema. *JAMA Netw Open*. 2020;3(11):e2024373.
28. Martínez-Jaimez P, Armora Verdú M, Forero CG, et al. Breast cancer-related lymphoedema: Risk factors and prediction model. *J Adv Nurs*. 2022;78(3):765–75.
29. Wang L, Li HP, Liu AN, et al. A Scoring System to Predict Arm Lymphedema Risk for Individual Chinese Breast Cancer Patients. *Breast Care (Basel)*. 2016;11(1):52–6.
30. Wei X, Lu Q, Jin S, et al. Developing and validating a prediction model for lymphedema detection in breast cancer survivors. *Eur J Oncol Nurs*. 2021;54:102023.
31. Liu YF, Liu JE, Zhu Y, et al. Development and validation of a nomogram to predict the risk of breast cancer-related lymphedema among Chinese breast cancer survivors. *Support Care Cancer*. 2021;29(9):5435–45.
32. Soran A, Wu WC, Dirican A, Johnson R, Andacoglu O, Wilson J. Estimating the probability of lymphedema after breast cancer surgery. *Am J Clin Oncol*. 2011;34(5):506–10.
33. Byun HK, Kim JS, Chang JS, et al. Validation of a nomogram for predicting the risk of lymphedema following contemporary treatment for breast cancer: a large multi-institutional study (KROG 20–05). *Breast Cancer Res Treat*. 2022;192(3):553–61.
34. Lin Q, Yang T, Yongmei J, Die YM. Prediction models for breast cancer-related lymphedema: a systematic review and critical appraisal. *Syst Rev*. 2022;11(1):217.
35. Li X, Huang H, Lin Q, et al. Validation of a breast cancer nomogram to predict lymphedema in a Chinese population. *J Surg Res*. 2017;210:132–8.
36. Soran A, Menekse E, Girgis M, DeGore L, Johnson R. Breast cancer-related lymphedema after axillary lymph node dissection: does early postoperative prediction model work? *Support Care Cancer*. 2016;24(3):1413–9.
37. Basta MN, Wu LC, Kanchwala SK, et al. Reliable prediction of postmastectomy lymphedema: The Risk Assessment Tool Evaluating Lymphedema. *Am J Surg*. 2017;213(6):1125-1133. e1121.
38. Bevilacqua JL, Kattan MW, Changhong Y, et al. Nomograms for predicting the risk of arm lymphedema after axillary dissection in breast cancer. *Ann Surg Oncol*. 2012;19(8):2580–9.
39. Du J, Yang J, Yang Q, Zhang X, Yuan L, 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.
40. Sun Y, Xia X, Liu X. Predictive modeling of breast cancer-related lymphedema using machine learning algorithms. *Gland Surg*. 2024;13(12):2243–52.
41. Wiser I, Mehrara BJ, Coriddi M, et al. Preoperative Assessment of Upper Extremity Secondary Lymphedema. *Cancers (Basel)*. 2020;12(1):135.
42. Brorson H, Höijer P. Standardised measurements used to order compression garments can be used to calculate arm volumes to evaluate lymphoedema treatment. *J Plast Surg Hand Surg*. 2012;46(6):410–5.
43. Ancukiewicz M, Russell TA, Otoole J, et al. Standardized method for quantification of developing lymphedema in patients treated for breast cancer. *Int J Radiat Oncol Biol Phys*. 2011;79(5):1436–43.
44. Tam S, Dong W, Adelman DM, Weber RS, Lewis CM. Risk-adjustment models in patients undergoing head and neck surgery with reconstruction. *Oral Oncol*. 2020;111:104917.
45. Karabacak M, Jagtiani P, Schupper AJ, Carr MT, Steinberger J, Margetis K. Machine Learning for Individualized Risk Estimation in Anterior Lumbar Interbody Fusion. *Neurosurg Pract*. 2024;5(3):e00099.
46. Chen T, Guestrin C. Xgboost: A scalable tree boosting system. Paper presented at: Proceedings of the 22nd ACM SIGKDD international conference on knowledge discovery and data mining. 2016 https://doi.org/10.1145/2939672.2939785.
47. Ying C, Qi-Guang M, Jia-Chen L, Lin G. Advance and prospects of AdaBoost algorithm. *Acta Automatica Sinica*. 2013;39(6):745–58.
48. Ke G, Meng Q, Finley T, et al. Lightgbm: A highly efficient gradient boosting decision tree. *Adv Neural Inf Process Syst* 2017;30.
49. Pal M. Random forest classifier for remote sensing classification. *Int J Remote Sens*. 2005;26(1):217–22.
50. Sidey-Gibbons JAM, Sidey-Gibbons CJ. Machine learning in medicine: a practical introduction. *BMC Med Res Methodol*. 2019;19(1):64.
51. Chawla NV, Bowyer KW, Hall LO, Kegelmeyer WP. SMOTE: synthetic minority over-sampling technique. *J Artif Intell Res*. 2002;16:321–57.
52. Akiba T, Sano S, Yanase T, Ohta T, Koyama M. Optuna: A next-generation hyperparameter optimization framework. Paper presented at: Proceedings of the 25th ACM SIGKDD international conference on knowledge discovery & data mining. 2019 https://doi.org/10.1145/3292500.3330701.
53. Vickers AJ, Cronin AM, Begg CB. One statistical test is sufficient for assessing new predictive markers. *BMC Med Res Methodol*. 2011;11(1):13.
54. Sundararajan M, Najmi A. The many Shapley values for model explanation. Paper presented at: International conference on machine learning. 2020 https://proceedings.mlr.press/v119/sundararajan20b.html.
55. Peng H, Long F, Ding C. Feature selection based on mutual information: criteria of max-dependency, max-relevance, and min-redundancy. *IEEE Trans Pattern Anal Mach Intell*. 2005;27(8):1226–38.
56. Delen D, Walker G, Kadam A. Predicting breast cancer survivability: a comparison of three data mining methods. *Artif Intell Med*. 2005;34(2):113–27.
57. Kwan ML, Yao S, Lee VS, et al. Race/ethnicity, genetic ancestry, and breast cancer-related lymphedema in the Pathways Study. *Breast Cancer Res Treat*. 2016;159(1):119–29.
58. Hellwege JN, Torstenson ES, Russell SB, Edwards TL, Velez Edwards DR. Evidence of selection as a cause for racial disparities in fibroproliferative disease. *PLoS One*. 2017;12(8):e0182791.
59. Diaz A, Tan K, He H, et al. Keloid lesions show increased IL-4/IL-13 signaling and respond to Th2-targeting dupilumab therapy. *J Eur Acad Dermatol Venereol*. 2020;34(4):e161–4.
60. Schmitz KH, Neuhouser ML, Agurs-Collins T, et al. Impact of obesity on cancer survivorship and the potential relevance of race and ethnicity. *J Natl Cancer Inst*. 2013;105(18):1344–54.
61. Obermeyer Z, Emanuel EJ. Predicting the Future - Big Data, Machine Learning, and Clinical Medicine. *N Engl J Med*. 2016;375(13):1216–9.
62. Fu MR, Wang Y, Li C, et al. Machine learning for detection of lymphedema among breast cancer survivors. *Mhealth*. 2018;4:17.
63. Wu X, Guan Q, Cheng ASK, et al. Comparison of machine learning models for predicting the risk of breast cancer-related lymphedema in Chinese women. *Asia Pac J Oncol Nurs*. 2022;9(12):100101.
64. Yaghoobi Notash A, Yaghoobi Notash A, Omidi Z, Haghighat S. Prediction of lymphedema occurrence in patients with breast cancer using the optimized combination of ensemble learning algorithm and feature selection. *BMC Med Inform Decis Mak*. 2022;22(1):195.
65. Trinh X-T, Chien PN, Long N-V, et al. Development of predictive models for lymphedema by using blood tests and therapy data. *Sci Rep*. 13(1):19720.
66. Chen J, Gabay A, Kim M, et al. AI Risk Prediction Tools for Alloplastic Breast Reconstruction. *Plast Reconstr Surg*. 2025;156(4):639–50.

67. Gabay AJ, Chen J, Stern CS, Gibbons C, Mehrara B, Nelson JA. Predicting Breast Reconstruction Readmission, Reoperation, and Prolonged Length of Stay: A Machine Learning Approach. *J Surg Oncol*. 2025;132(2):345–53.
68. Dankers FJWM, Traverso A, Wee L, van Kuijk SMJ. Prediction modeling methodology. In: P Kubben, M Dumontier, A Dekker, editors. Fundamentals of Clinical Data Science. Cham: Springer International Publishing; 2019. p. 101–20.
69. Johnson AR, Fleishman A, Granoff MD, et al. Evaluating the Impact of Immediate Lymphatic Reconstruction for the Surgical Prevention of Lymphedema. *Plast Reconstr Surg*. 2021;147(3):373e–81e.
70. Johnson AR, Asban A, Granoff MD, et al. Is Immediate Lymphatic Reconstruction Cost-effective? *Ann Surg*. 2021;274(6):e581–8.
71. Coriddi M, Dayan J, Bloomfield E, et al. Efficacy of Immediate Lymphatic Reconstruction to Decrease Incidence of Breast Cancer-related Lymphedema: Preliminary Results of Randomized Controlled Trial. *Ann Surg*. 2023;278(4):630–7.
72. Runowicz CD, Leach CR, Henry NL, et al. American Cancer Society/American Society of Clinical Oncology Breast Cancer Survivorship Care Guideline. *J Clin Oncol*. 2016;34(6):611–35.