Am J Cancer Res 2025;15(3):1081-1095
www.ajcr.us /ISSN:2156-6976/ajcr0162693

Original Article

Development and validation of a risk prediction model for lymph node metastasis in stage IA2-IIA1 cervical cancer based on laboratory parameters

Yongli Hou[1], Lili Zhang[1], Hui Wang[2], Wenhao Wang[1], Min Hao[1]

[1]Department of Obstetrics and Gynecology, The Second Hospital of Shanxi Medical University, Taiyuan 030000, Shanxi, China; [2]Reproductive Medicine Center, Shanxi Maternal and Child Health Hospital, Taiyuan 030000, Shanxi, China

Received December 15, 2024; Accepted February 17, 2025; Epub March 15, 2025; Published March 30, 2025

Abstract: Objective: To develop and validate a risk prediction model for lymph node metastasis (LNM) in stage IA2-IIA1 cervical cancer (CC) using laboratory parameters to aid in preoperative risk assessment and personalized treatment planning. Methods: A retrospective analysis was conducted on 624 patients treated between 2017 and 2023, divided into a training group (418 patients) and a validation group (206 patients). Clinical and laboratory data, including squamous cell carcinoma antigen (SCC-Ag), carcinoembryonic antigen (CEA), cancer antigen 125 (CA125), platelet count (PLT), fibrinogen (FIB), and C-reactive protein (CRP), were collected. Independent risk factors for LNM were identified using Least Absolute Shrinkage and Selection Operator (LASSO) regression. A predictive model was constructed and evaluated using receiver operating characteristic (ROC) curve analysis, decision curve analysis (DCA), and calibration curve. Results: SCC-Ag, CEA, CA125, PLT, FIB, and CRP were identified as significant predictors of LNM, with SCC-Ag demonstrating an AUC of 0.811 (sensitivity: 65.00%, specificity: 93.08%). The model achieved an AUC of 0.969 in the training group and 0.942 in the validation group, indicating robust generalizability and high predictive accuracy. DCA confirmed the model's clinical utility across a wide range of risk thresholds, and the calibration curve showed a good agreement between predicted and observed outcomes. Conclusions: This laboratory parameter-based risk prediction model is a reliable and practical tool for assessing LNM risk in stage IA2-IIA1 CC patients, supporting better clinical decision-making and reducing unnecessary interventions.

Keywords: Cervical cancer, lymph node metastasis, risk prediction model, LASSO regression, laboratory parameters

Introduction

With advancements in medical technology and increased health awareness, the global incidence of cervical cancer (CC) has shown a notable decline [1]. However, in underdeveloped regions, limited CC screening programs and insufficient HPV vaccination have resulted in persistently high incidence and mortality rates [2]. In China, unequal distribution of medical resources remains a major challenge for early diagnosis and prevention. According to the 2022 Global Cancer Statistics, CC accounts for approximately 6.6% of all new cancer cases worldwide, with an estimated 600,000 new cases and 300,000 deaths annually [3]. These alarming figures highlight the urgent need for effective diagnostic tools and strategies to manage CC, particularly in economically disadvantaged areas with limited access to advanced medical resources.

As outlined by the 2018 staging criteria from the International Federation of Gynecology and Obstetrics (FIGO), treatment options and prognostic outcomes for CC vary significantly by disease stage [4]. Early-stage CC, particularly stages IA2-IIA1, represents a critical phase where lymph node metastasis (LNM) risk plays a pivotal role in treatment planning and prognostic evaluation. Statistical evidence indicates that approximately 15%-20% of early-stage CC patients exhibit LNM, which is closely linked to poor outcomes [5]. Accurate preoperative

assessment of LNM is therefore essential for optimizing treatment strategies, selecting appropriate surgical interventions, and avoiding unnecessary lymph node dissections in low-risk patients. By minimizing surgical trauma and related complications, a precise evaluation of LNM risk can significantly improve the overall prognosis and quality of life for CC patients.

Currently, various imaging modalities, such as positron emission tomography-computed tomography (PET-CT), CT, and magnetic resonance imaging (MRI), are widely used in clinical practice to assess LNM. Among these, PET-CT is often preferred for its relatively high specificity; however, it is limited by low sensitivity and a non-negligible false-negative rate, which can affect clinical decision-making [6, 7]. To overcome these limitations, sentinel lymph node biopsy (SLNB) has been introduced as an alternative for detecting LNM. While SLNB offers improved sensitivity, it is still hampered by procedural risks, false negatives, and technical challenges [8]. In response to these issues, there has been increasing interest in laboratory biomarkers, which are simpler, cost-effective, and readily accessible. Several biomarkers, including squamous cell carcinoma antigen (SCC-Ag), carcinoembryonic antigen (CEA), cancer antigen 125 (CA125), and inflammatory markers like C-reactive protein (CRP), have shown promise in predicting LNM [9, 10]. Among these, SCC-Ag is particularly noteworthy due to its strong correlation with tumor burden and metastatic progression. Despite progress, the lack of unified standards and optimal cut-off values for these markers has limited their widespread clinical use. Therefore, further research is needed to validate these biomarkers and develop standardized protocols for their clinical application.

Focusing on stage IA2-IIA1 CC patients is particularly significant, as this stage marks the emergence of LNM risk, which becomes a decisive factor in treatment outcomes. Identifying LNM preoperatively in these patients is crucial to avoid overtreatment in low-risk cases and ensure timely intervention in high-risk cases. Given the impact of LNM on treatment strategies and prognosis, there is a growing need for reliable predictive tools that provide accurate preoperative assessments.

In recent years, nomograms, which transform complex regression models into intuitive graphical representations, have gained popularity in oncology. These tools help clinicians evaluate the influence of multiple variables on patient outcomes [11]. While some studies have developed nomogram-based models for LNM prediction in CC, many of these models lack consistency in predictive performance, and no universally accepted standard has been established [12]. This inconsistency underscores the need for further refinement and validation of predictive models.

To address these gaps, the present study aims to develop and validate a robust LNM risk prediction model for stage IA2-IIA1 CC patients. Using laboratory parameters and Least Absolute Shrinkage and Selection Operator (LASSO) regression analysis, this study integrates multiple biomarkers to create an accurate, practical, and user-friendly tool. By providing reliable preoperative assessments, the model is expected to support clinicians in optimizing personalized treatment strategies, improving prognostic outcomes, and enhancing the overall quality of care for CC patients.

Methods and materials

Sample source

This retrospective study aimed to analyze the independent risk factors for LNM in patients with stage IA2-IIA1 CC and to construct an LNM risk prediction model based on multiple laboratory parameters. The study included CC patients who received treatment at The Second Hospital of Shanxi Medical University from 2017 to 2023. The study adhered to the Helsinki Declaration [13] and was approved by the ethics committee of the Second Hospital of Shanxi Medical University.

Definition of LNM

LNM refers to the infiltration of cancer cells into pelvic or distant lymph nodes, confirmed by postoperative pathological examination. All patients underwent radical hysterectomy and pelvic lymphadenectomy, with the postoperative pathological report serving as the definitive criterion for determining LNM presence. Metastatic lymph nodes typically exhibit features such as enlargement, altered texture

(hardening), or tumor cell infiltration. The presence or absence of LNM was confirmed using pathological methods, including HE staining or immunohistochemistry. In this study, the LNM status was based on the final postoperative pathological diagnosis.

Inclusion and exclusion criteria

Inclusion criteria: (1) Patients diagnosed with stage IA2-IIA1 CC according to the FIGO 2018 staging guidelines [14]; (2) Patients who underwent radical hysterectomy and pelvic lymphadenectomy; (3) Patients who had relevant clinical examinations (e.g., imaging, blood biochemical tests) before surgery, with complete medical records and traceable data.

Exclusion criteria: (1) Individuals with other malignant neoplasms or severe comorbidities, such as cardiovascular disease, hepatic or renal insufficiency; (2) Patients with incomplete clinical data or follow-up information; (3) Patients who had previously undergone other therapeutic modalities (e.g., radiotherapy or chemotherapy) before surgery.

Sample grouping

In strict adherence to the predefined eligibility criteria, a total of 744CC patient samples were collected. Of these, 624 valid samples were retained after applying the exclusion criteria. For the construction and validation of the risk prediction model, the samples were randomly divided into a training group and a validation group in a 67% to 33% ratio. The training group, consisting of 418 samples, was used for model development and training. The validation group, consisting of 206 samples, was designated for model validation and performance evaluation.

Clinical data collection

The clinical data of patients were retrospectively collected through analysis of their electronic medical records. The data included the following parameters: age, gender (male or female), body mass index (BMI), LNM status (with or without), menopausal status (yes or no), HPV infection status (positive or negative), differentiation status (poorly differentiated or moderately/well differentiated), maximum tumor diameter (≥ 2 cm or < 2 cm), pathological type (squamous cell carcinoma or other types), depth of mesenchymal infiltration (≥ 1/2 or < 1/2), and pathological stage (stage IA2-IB, or IIA1). The laboratory parameters included SCC-Ag (ng/ml), carcinoembryonic antigen (CEA, ng/mL), cancer antigen 125 (CA125, U/mL), neutrophil count (Neu, $\times 10^9$/L), platelet count (PLT, $\times 10^9$/L), neutrophil-lymphocyte ratio (NLR), platelet-lymphocyte ratio (PLR), fibrinogen (FIB, g/L), albumin (Alb, g/L), and C-reactive protein (CRP, mg/L). These parameters were measured using standardized testing methods and dedicated equipment and were used in subsequent data analysis.

Laboratory testing

Laboratory data were obtained from peripheral blood samples collected from patients before surgery. All samples were processed within 24 hours of collection. Routine blood tests, tumor markers, and other relevant biomarkers were measured. Routine blood tests were performed using a fully automatic blood analyzer (Mindray BC-6800, Mindray, Shenzhen, China). Tumor markers, including SCC-Ag, CEA, and CA125, were assayed using chemiluminescent immunoassay methodology (Roche Cobas e411, Roche Diagnostics, Switzerland). Other indices, such as FIB, CRP, and Alb, were measured using a Cobas 6000 automatic biochemical analyzer (Roche Diagnostics, Switzerland). All tests were conducted in strict accordance with standard operating procedures (SOP) to ensure the accuracy and reliability of the data.

Model construction

LASSO regression modeling: LASSO regression is a method specifically designed for variable selection and regularization. It is highly effective in identifying independent risk factors closely associated with LNM. In this study, LASSO regression analysis was employed to identify the most prognostically significant factors from a wide range of variables. The cv.glmnet function was used for model training, with the following parameter settings: the independent variable matrix (x) and dependent variable vector (y) were included in the regression analysis. The regression type was set to family = "binomial", suitable for binary classification tasks. LASSO regression (L1 regularization) was set with alpha = 1, and a 10-fold cross-validation strategy (nfolds = 10) was applied to evaluate the model's generalization ability.

Nomogram construction: Based on the regression coefficients of the LASSO model, a nomogram was constructed to visually represent the contribution of each variable to the risk of LNM. By converting regression coefficients into nomogram scores, clinicians can easily calculate the risk score for individual patients, providing valuable guidance for clinical decision-making.

Outcome measurement

Primary outcome: The development of a risk prediction model for LNM in stage IA2-IIA1 CC.

Secondary outcome: An in-depth evaluation of the model's stability and reliability, using multiple evaluation metrics, including the receiver operating characteristic (ROC) curve, decision curve analysis (DCA), and calibration curve.

Statistical analysis

Statistical analyses were performed using SPSS 25.00 for data processing and descriptive statistics. Initially, a normality test was conducted using the Shapiro-Wilk test to determine whether continuous variables followed a normal distribution. Variables conforming to a normal distribution were expressed as mean ± standard deviation (SD) and compared using an independent-sample t-test. Non-normally distributed variables were expressed as median (interquartile range, IQR) and analyzed using the Mann-Whitney U test. Categorical variables were presented as frequencies (n) and percentages (%), with inter-group comparisons performed using the chi-square test or Fisher's exact test, as appropriate. Spearman's rank correlation was used to assess monotonic relationships between variables.

For model construction, R version 4.3.3 was employed. LASSO regression, specifically the cv.glmnet function, was used for feature selection to identify independent risk factors associated with LNM. The regularization parameters were optimized by cross-validation. The model's predictive ability was assessed using ROC curves, DCA, and calibration curves. The DeLong test was used to compare the area under the curve (AUC) between different groups. Statistical significance was defined as $P < 0.05$ to ensure the reliability and validity of the results.

Results

Comparison of baseline data between training and validation groups

No statistically significant differences were observed between the training and validation groups across all variables. Specifically, characteristics such as age (P = 0.750), BMI (P = 0.417), LNM status (P = 0.819), menopausal status (P = 0.576), HPV infection status (P = 0.518), differentiation status (P = 0.852), maximum tumor diameter (P = 0.502), pathological type (P = 0.540), depth of mesenchymal infiltration (P = 0.601), and pathological stage (P = 0.761) showed no significant differences. Hematological markers, including SCC-Ag (P = 0.555), CEA (P = 0.446), CA125 (P = 0.851), Neu (P = 0.552), and PLT (P = 0.979), also did not differ significantly between the groups. Similarly, inflammation-related indicators, such as NLR (P = 0.917), PLR (P = 0.700), FIB (P = 0.919), Alb (P = 0.755), and CRP (P = 0.964), showed no statistically significant differences (**Table 1**).

Additionally, when comparing the clinical and laboratory parameters of patients with and without LNM, several significant differences were identified. The LNM group had higher rates of positive HPV infection (P < 0.001), lower differentiation status (P < 0.001), deeper mesenchymal infiltration (P < 0.001), and non-squamous carcinoma (P = 0.033). However, no significant differences were found in factors such as age (P = 0.136), BMI (P = 0.867), menopausal status (P = 0.636), maximum tumor diameter (P = 0.252), or pathological stage (P = 0.656) (Table S1). Laboratory indices, such as SCC-Ag (P < 0.001), CEA (P < 0.001), CA125 (P < 0.001), Neu (P < 0.001), PLT (P < 0.001), PLR (P < 0.001), FIB (P < 0.001), and CRP (P < 0.001), were significantly elevated in the LNM group. In contrast, no significant differences were observed in lymphocyte count (P = 0.348), NLR (P = 0.144), or Alb (P = 0.487) between the groups (Table S2).

Comparison of baseline data between LNM and non-LNM patients in training group

Table 2 presents a comparison of baseline data between patients with and without LNM in the training group. No significant differences were found in age, BMI, menopausal status,

Table 1. Comparison of baseline data between training and validation groups

Variable	Total	Training group (n = 418)	Validation group (n = 206)	Statistic	P
Age	52.679±10.117	52.770±10.034	52.495±10.305	0.319	0.750
Body mass index	24.005±2.994	24.073±2.979	23.866±3.028	0.812	0.417
Lymph node metastasis					
With	151	100 (23.92%)	51 (24.76%)	0.052	0.819
Without	473	318 (76.08%)	155 (75.24%)		
Pausimenia					
Yes	246	168 (40.19%)	78 (37.86%)	0.313	0.576
No	378	250 (59.81%)	128 (62.14%)		
HPV infection					
Positive	531	353 (84.45%)	178 (86.41%)	0.417	0.518
Negative	93	65 (15.55%)	28 (13.59%)		
Differentiation degree					
Poorly differentiated	200	135 (32.3%)	65 (31.55%)	0.035	0.852
Moderately to well differentiated	424	283 (67.7%)	141 (68.45%)		
Maximum tumor diameter					
≥ 2 cm	444	301 (72.01%)	143 (69.42%)	0.452	0.502
< 2 cm	180	117 (27.99%)	63 (30.58%)		
Pathological type					
Squamous cell carcinoma	558	376 (89.95%)	182 (88.35%)	0.375	0.540
Others	66	42 (10.05%)	24 (11.65%)		
Depth of mesenchymal infiltration					
≥ 1/2	333	220 (52.63%)	113 (54.85%)	0.274	0.601
< 1/2	291	198 (47.37%)	93 (45.15%)		
Pathological stage					
IA2-IB	402	271 (64.83%)	131 (63.59%)	0.093	0.761
IIA1	222	147 (35.17%)	75 (36.41%)		
SCC-Ag (ng/ml)	4.46 [3.05, 5.99]	4.58 [3.05, 6.09]	4.35 [3.03, 5.73]	0.590	0.555
CEA (ng/mL)	38.882±10.143	39.100±10.317	38.441±9.792	0.763	0.446
CA125 (U/mL)	10.14 [8.70, 11.95]	10.13 [8.70, 11.91]	10.19 [8.73, 12.08]	0.187	0.851
Neu ($\times 10^9$/L)	3.43 [2.75, 4.15]	3.42 [2.80, 4.15]	3.45 [2.67, 4.18]	0.595	0.552
PLT ($\times 10^9$/L)	196.00 [158.00, 232.00]	196.50 [158.00, 231.00]	195.50 [159.25, 233.00]	0.026	0.979
Lym ($\times 10^9$/L)	1.772±0.431	1.782±0.421	1.753±0.452	0.792	0.429
NLR	1.95 [1.46, 2.51]	1.95 [1.50, 2.49]	1.94 [1.41, 2.54]	0.104	0.917
PLR	109.22 [85.26, 138.97]	110.39 [85.27, 137.67]	108.83 [85.26, 144.52]	0.385	0.700
FIB (g/L)	2.437±0.536	2.436±0.552	2.440±0.502	-0.102	0.919
Alb (g/L)	41.954±3.703	41.921±3.625	42.020±3.864	-0.312	0.755
CRP (mg/L)	2.77 [1.67, 4.05]	2.77 [1.68, 4.09]	2.82 [1.65, 4.04]	0.045	0.964

Note: LNM, lymph node metastasis; SCC-Ag, squamous cell carcinoma antigen; CEA, carcinoembryonic antigen; CA125, cancer antigen 125; Neu, neutrophil count; PLT, platelet count; Lym, lymphocyte count; NLR, neutrophil-lymphocyte ratio; PLR, platelet-lymphocyte ratio; FIB, fibrinogen; Alb, albumin; CRP, C-reactive protein.

maximum tumor diameter, pathological type, pathological stage, lymphocyte count (Lym), or Alb between patients with and without LNM (*P*-values: 0.673, 0.379, 0.373, 0.800, 0.059, 0.779, 0.363, and 0.920, respectively). However, patients with LNM exhibited significantly elevated levels in several indicators, including HPV infection status (P = 0.007), differentiation status (P < 0.001), depth of mesenchymal infiltration (P < 0.001), SCC-Ag (P < 0.001), CEA (P < 0.001), CA125 (P < 0.001), Neu (P < 0.001), PLT (P < 0.001), PLR (P < 0.001), FIB (P < 0.001), and CRP (P < 0.001), compared to patients without LNM. These factors may be implicated in the occurrence of LNM (**Table 2**).

ROC curve characteristics of laboratory parameters associated with LNM

In this study, we evaluated the efficacy of several laboratory indicators in predicting LNM using ROC curve analysis. First, each measurement was binarized based on its cutoff value to standardize the data format, followed by LASSO regression using the binomial method. ROC analysis revealed that SCC-Ag had the

Table 2. Comparison of baseline data between patients with and without lymph node metastasis in training group

Variable	Total	LNM group (n = 100)	Non-LNM group (n = 318)	Statistic	P
Age	52.770±10.034	53.140±10.954	52.654±9.742	-0.422	0.673
Body mass index	24.073±2.979	24.302±2.954	24.001±2.987	-0.881	0.379
Pausimenia					
Yes	168	44 (44%)	124 (38.99%)	0.793	0.373
No	250	56 (56%)	194 (61.01%)		
HPV infection					
Positive	353	93 (93%)	260 (81.76%)	7.318	0.007
Negative	65	7 (7%)	58 (18.24%)		
Differentiation degree					
Poorly differentiated	135	49 (49%)	86 (27.04%)	16.772	< 0.001
Moderately to well differentiated	283	51 (51%)	232 (72.96%)		
Maximum tumor diameter					
≥ 2 cm	301	73 (73%)	228 (71.7%)	0.064	0.800
< 2 cm	117	27 (27%)	90 (28.3%)		
Pathological type					
Squamous cell carcinoma	376	85 (85%)	291 (91.51%)	3.567	0.059
Others	42	15 (15%)	27 (8.49%)		
Depth of mesenchymal infiltration					
≥ 1/2	220	79 (79%)	141 (44.34%)	36.659	< 0.001
< 1/2	198	21 (21%)	177 (55.66%)		
Pathological stage					
IA2-IB	271	66 (66%)	205 (64.47%)	0.079	0.779
IIA1	147	34 (34%)	113 (35.53%)		
SCC-Ag (ng/ml)	4.792±2.523	7.205±3.143	4.033±1.694	-12.987	< 0.001
CEA (ng/mL)	39.100±10.317	44.792±10.194	37.310±9.702	-6.645	< 0.001
CA125 (U/mL)	10.382±2.435	12.427±2.406	9.739±2.064	-10.901	< 0.001
Neu ($\times 10^9$/L)	3.491±1.152	3.824±1.688	3.386±0.900	-3.358	< 0.001
PLT ($\times 10^9$/L)	196.467±55.878	227.590±69.512	186.679±46.868	-6.715	< 0.001
Lym ($\times 10^9$/L)	1.782±0.421	1.815±0.430	1.771±0.418	-0.910	0.363
NLR	1.95 [1.50, 2.49]	2.11 [1.46, 2.72]	1.91 [1.51, 2.39]	1.330	0.184
PLR	110.39 [85.27, 137.67]	124.24 [97.30, 160.69]	106.03 [83.01, 130.61]	4.094	< 0.001
FIB (g/L)	2.436±0.552	2.651±0.684	2.368±0.486	-4.587	< 0.001
Alb (g/L)	41.921±3.625	41.889±3.390	41.931±3.701	0.100	0.920
CRP (mg/L)	2.77 [1.68, 4.09]	4.16 [2.25, 6.67]	2.57 [1.57, 3.49]	5.703	< 0.001

Note: LNM, lymph node metastasis; SCC-Ag, squamous cell carcinoma antigen; CEA, carcinoembryonic antigen; CA125, cancer antigen 125; Neu, neutrophil count; PLT, platelet count; Lym, lymphocyte count; NLR, neutrophil-lymphocyte ratio; PLR, platelet-lymphocyte ratio; FIB, fibrinogen; Alb, albumin; CRP, C-reactive protein.

highest predictive accuracy for LNM, with an AUC of 0.811, sensitivity of 65.00%, and specificity of 93.08%. Other markers, such as CA125 (AUC = 0.801) and CEA (AUC = 0.696), also showed relatively strong predictive performance, with sensitivities of 80.00% and 55.00%, respectively. Additionally, PLT, FIB, and CRP exhibited predictive potential, although with lower AUC values (**Figure 1**; **Table 3**).

Correlation analysis of characteristic variables

Spearman's correlation analysis was performed to examine the relationships among various clinical indicators. The results showed a significant positive correlation between PLR and PLT ($r = 0.432$, $P < 0.001$), as well as between SCC-Ag and CRP ($r = 0.306$, $P < 0.001$). Since LASSO regression automatically selects features and mitigates the impact of redundant variables through the L1 regularization term, we chose to retain all relevant variables without excluding any data (**Figure 2**).

LASSO regression model and confusion matrix analysis

The LASSO Regression Model and Confusion Matrix Analysis was constructed using the 1 standard error (1 SE) approach. At a lambda

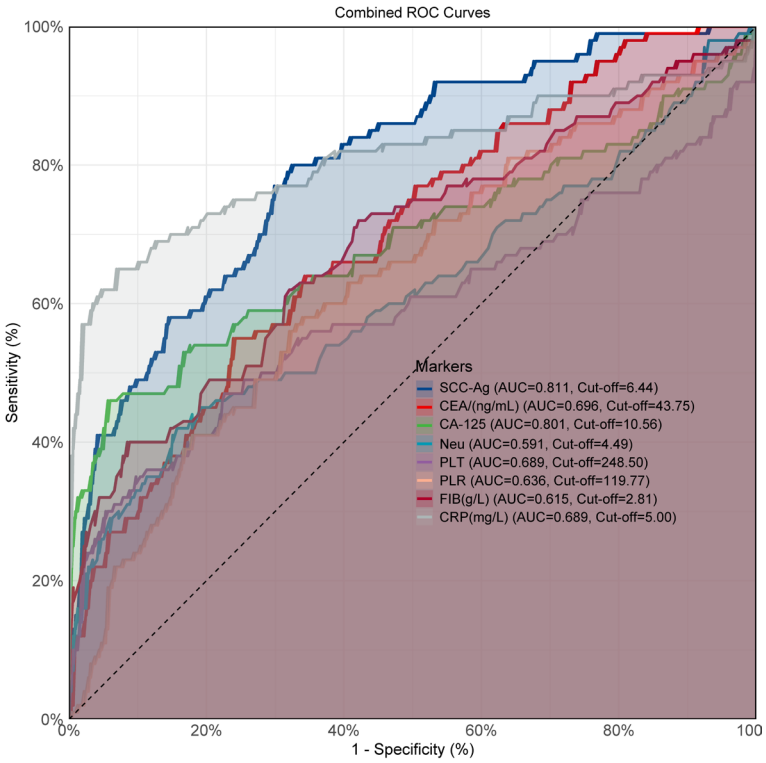

Figure 1. ROC curve analysis diagram. The ROC curve graph delineates the predictive capabilities of diverse laboratory indicators associated with lymph node metastasis, including SCC-Ag, CEA, CA125, Neu, PLT, PLR, FIB, and CRP. Note: ROC, receiver operating characteristic; AUC, area under the curve; SCC-Ag, squamous cell carcinoma antigen; CEA, carcinoembryonic antigen; CA125, cancer antigen 125; Neu, neutrophil count; PLT, platelet count; PLR, platelet-lymphocyte ratio; FIB, fibrinogen; CRP, C-reactive protein.

value of 0.0231 (indicated by the blue dots in **Figure 3A**), nine variables were selected: SCC-Ag, CEA, CA125, Neu, PLT, FIB, CRP, differentiation, and depth of mesenchymal infiltration (**Figure 3B**). Based on the regression coefficients of these selected variables, the risk model was formulated as follows:

Risk Score = 15.1419538 - 2.237099526 × SCC-Ag (ng/ml) -0.429975303 × CEA (ng/mL) - 1.46315786 × CA125 (U/mL) - 1.071651798 × Neu (×10^9L) - 1.108481595 × PLT (×10^9L) - 0.602152848 × FIB (g/L) - 1.683026492 × CRP (mg/L) - 0.388277279 × Differentiation - 0.740627285 × Depth of mesenchymal infiltration. During model validation, a confusion matrix (**Figure 3C**) was used to assess the model's performance, illustrating the accuracy of predictions across various classification categories, including the number of correctly predicted positive and negative cases and misclassifications. This information is crucial for understanding the model's predictive capacity and for guiding risk assessment in clinical applications.

Nomogram construction and variable correlation analysis

A Nomogram was constructed based on the LASSO regression model to predict LNM risk. The regression coefficients of individual variables were used to evaluate their contribution to the risk score. Based on the analysis, variables were categorized into three groups: strongly correlated, moderately correlated, and weakly correlated. SCC-Ag, CEA, and CA125 were strongly correlated with LNM and had the greatest impact on the model's predictive capacity. PLT, FIB, and CRP were moderately correlated and contributed moderately to the prediction. Neu, differentiation status, and depth of mesenchymal infiltration were weakly correlated, making a minor contribution to the risk score (**Figure 4**).

ROC curve, calibration curve, and DCA of the training group model

This study comprehensively evaluated the performance of the training group model, demonstrating its strong predictive ability. First, the ROC curve (**Figure 5A**) illustrated the model's discriminatory capacity across various thresholds. The AUC value provided a quantitative measure, indicating the model's high classification power. The calibration curve (**Figure 5B**) assessed the agreement between the predicted and observed values, with a goodness-of-fit test yielding a chi-square value of less than 0.001 and a *P*-value of 1, suggesting a favorable calibration. Additionally, the C-index was 0.969 (95% CI: 0.951-0.987), further confirming the model's accuracy in distinguishing different risk levels ($P < 0.001$). Lastly, the DCA curve (**Figure 5C**) demonstrated that the model provided benefits across a probability range of 0% to 97%, with the maximum benefit reaching

Table 3. ROC curve characteristics of laboratory parameters related to lymph node metastasis

Marker	AUC	95% CI	Specificity	Sensitivity	Youden index	Cut off
SCC-Ag (ng/ml)	0.811	0.750-0.873	93.08%	65.00%	58.08%	6.44
CEA (ng/mL)	0.696	0.638-0.755	76.10%	55.00%	31.10%	43.745
CA125 (U/mL)	0.801	0.752-0.851	67.61%	80.00%	47.61%	10.565
Neu ($\times 10^9$/L)	0.591	0.515-0.667	90.25%	35.00%	25.25%	4.49
PLT ($\times 10^9$/L)	0.689	0.624-0.754	91.51%	40.00%	31.51%	248.5
PLR	0.636	0.571-0.700	66.04%	58.00%	24.04%	119.768
FIB (g/L)	0.615	0.544-0.686	84.28%	42.00%	26.28%	2.815
CRP (mg/L)	0.689	0.618-0.760	94.34%	46.00%	40.34%	5

Note: ROC, receiver operating characteristic; AUC, area under the curve; SCC-Ag, squamous cell carcinoma antigen; CEA, carcinoembryonic antigen; CA125, cancer antigen 125; Neu, neutrophil count; PLT, platelet count; PLR, platelet-lymphocyte ratio; FIB, fibrinogen; CRP, C-reactive protein.

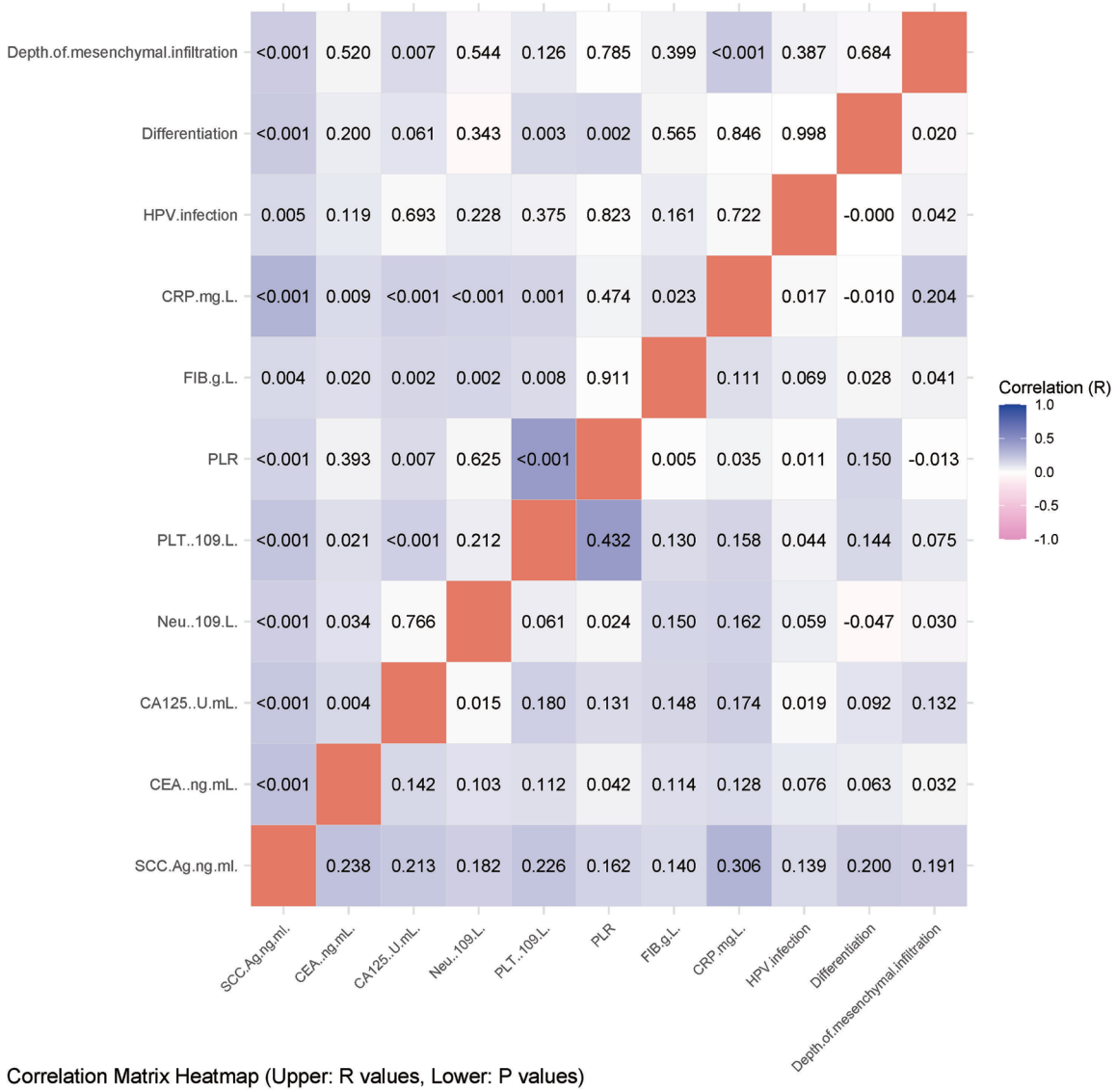

Figure 2. Spearman's test correlation matrix for variables associated with lymph node metastasis. Note: HPV, human papillomavirus; CRP, C-reactive protein; FIB, fibrinogen; PLR, platelet-lymphocyte ratio; PLT, platelet count; Neu, neutrophil count; CA125, cancer antigen 125; CEA, carcinoembryonic antigen; SCC-Ag, squamous cell carcinoma antigen.

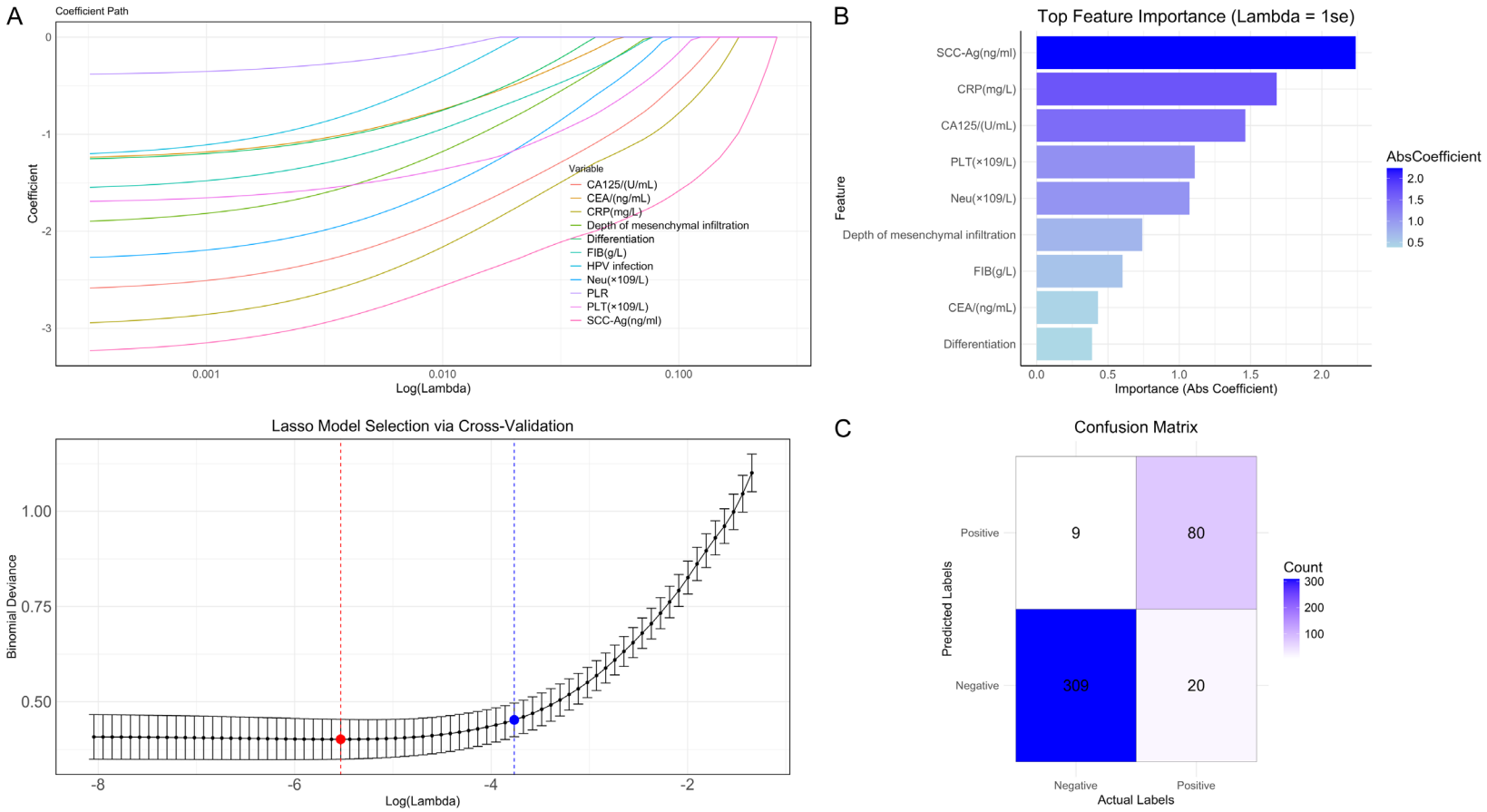

Figure 3. LASSO regression model construction. A. Lambda selection diagram of LASSO regression. B. The selected nine variables. C. Confusion matrix diagram. Note: LASSO, Least Absolute Shrinkage and Selection Operator.

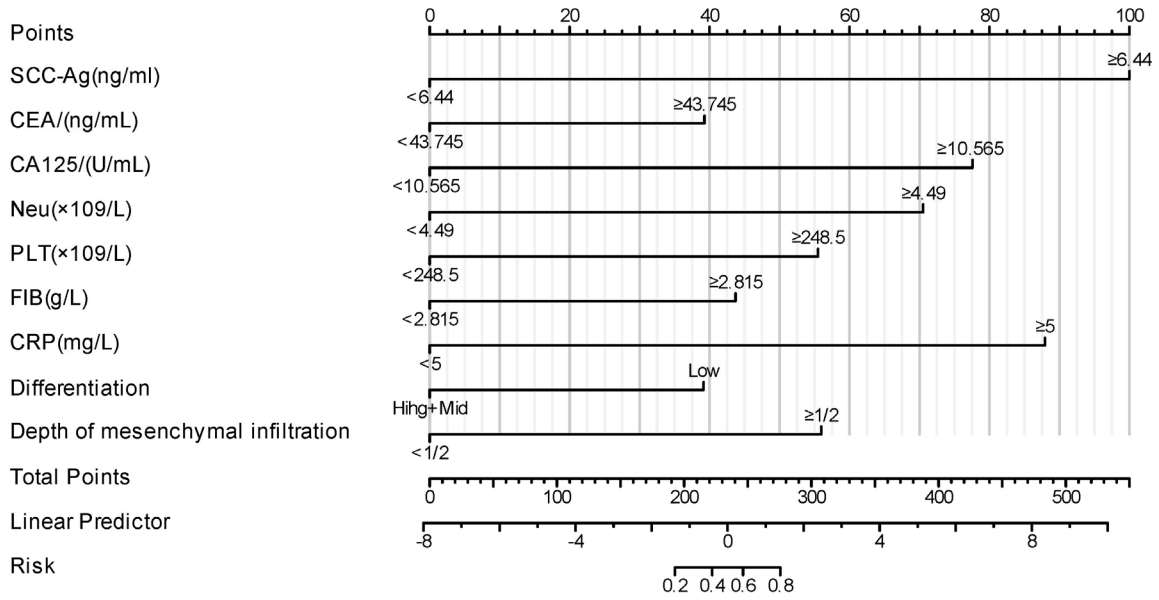

Figure 4. Nomogram construction diagram. The Nomogram diagram depicts the influence of each variable on the lymph node metastasis risk score. The variables are divided into strongly correlated, correlated, and weakly correlated classifications in accordance with their contributions to the prediction.

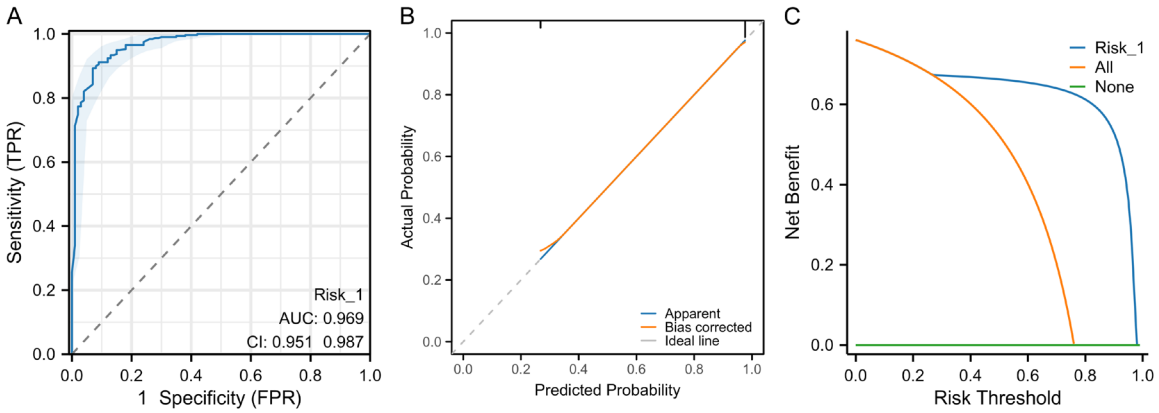

Figure 5. ROC curve, calibration curve, and DCA of the training group model. A. ROC curve of the training group model, employed to evaluate the discriminatory power of the model. B. Calibration curve, utilized for assessing the calibration accuracy of the model. C. DCA curve, applied to evaluate the clinical benefits of the model under different decision thresholds. Note: ROC, receiver operating characteristic; DCA, decision curve analysis.

76.07%. Together, these findings highlight the exceptional predictive performance of the model in the training group, underscoring its clinical value.

ROC curve, calibration curve, and DCA of the validation group model

In the validation cohort, the model's performance was thoroughly assessed, and the results confirmed its high accuracy. The ROC curve (**Figure 6A**) effectively depicted the model's classification efficacy across different thresholds, with the AUC value reflecting its pronounced discriminatory capacity. The calibration curve (**Figure 6B**) examined the alignment between predicted and observed values, with a goodness-of-fit test showing a chi-square value less than 0.001 (P = 1), indicating excellent calibration. The C-index was 0.942 (95% CI: 0.905-0.980), further supporting the model's strong ability to distinguish among various risk levels (P < 0.001). The DCA curve (**Figure 6C**) revealed that the model was beneficial within a probability range of 0% to 94%,

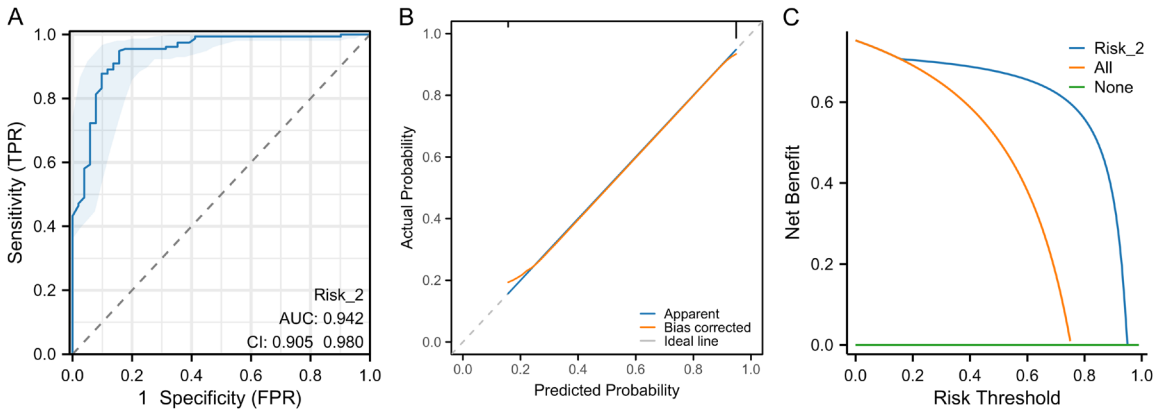

Figure 6. ROC curve, calibration curve, and DCA of the validation group model. A. ROC curve of the validation group model, employed to evaluate the discriminatory power of the model. B. Calibration curve, utilized for assessing the calibration accuracy of the model. C. DCA curve, applied to evaluate the clinical benefits of the model at different decision thresholds. Note: ROC, receiver operating characteristic; DCA, decision curve analysis.

Table 4. Comparison of ROC curve characteristics of training group and validation group

Marker	Risk_1	Risk_2
AUC	0.969	0.942
CI_lower_upper	0.951-0.987	0.905-0.980
Specificity	89.31%	94.84%
Sensitivity	93.00%	84.31%
Youden_index	82.31%	79.15%
Cut_off	0.21	0.257
Accuracy	90.19%	92.23%
Precision	93.00%	84.31%
F1_Score	81.94%	84.31%

Note: ROC, receiver operating characteristic; DCA, decision curve analysis.

with a maximum benefit of 75.24%. Collectively, these results demonstrate that the model performed excellently in the validation group as well, confirming its substantial clinical application potential.

Comparison of ROC curve characteristics between training group and validation group

Table 4 presents a comparison of the ROC curve characteristics for the models in the training group (Risk_1) and the validation group (Risk_2). The AUC for the training group was 0.969 (95% CI: 0.951-0.987), while the validation group had an AUC of 0.942 (95% CI: 0.905-0.980). This difference suggests that the model in the training group had slightly better predictive performance. The sensitivity of the training group was 93.00%, compared to 84.31% in the validation group. In contrast, the validation group had a higher specificity of 94.84%, compared to 89.31% in the training group. The Youden index was 82.31% for the training group and 79.15% for the validation group, showing a minor difference between the two. The accuracy rates were 90.19% for the training group and 92.23% for the validation group. Precision and F1 scores were 93.00% and 81.94%, respectively, for the training group, and 84.31% and 84.31% for the validation group. To assess the statistical significance of the AUC difference between the two groups, the DeLong test was conducted. The results showed a D value of 1.249, degrees of freedom (df) of 304.93, and a *P*-value of 0.213. These findings indicate no statistically significant difference in AUC between the training and validation groups (**Table 4**).

Discussion

CC is the most prevalent malignancy among women globally, particularly in developing countries, where it has high incidence and mortality rates [15]. For patients with stage IA2-IIA1 CC, early intervention often results in favorable outcomes. However, the presence of LNM significantly worsens prognosis, making LNM a critical factor influencing survival and treatment decisions. In this study, we developed and validated an LNM risk prediction model based on laboratory parameters (SCC-Ag, CEA, CA125, etc.) and clinicopathological features. The results show that this model can accurately predict LNM risk in patients with stage IA2-

IIA1 CC, aiding in preoperative decision-making and the tailoring of personalized treatment plans.

We employed the LASSO regression model to identify key prognostic factors. The advantage of the LASSO method is its ability to handle high-dimensional data efficiently. By using L1 regularization, it controls variable selection and reduces the risk of overfitting. Laboratory markers such as SCC-Ag, CEA, and CA125 were identified as independent prognostic factors with significant effects on LNM in CC. These markers offer valuable prognostic information for LNM, enhancing the accuracy of preoperative LNM assessments.

The nine key variables identified in this study - SCC-Ag, CEA, CA125, Neu, PLT, FIB, CRP, differentiation status, and depth of mesenchymal infiltration - encompass tumor markers, inflammatory and immune indices, platelet-related parameters, and histopathological features [16-18]. These factors not only shed light on the potential mechanisms underlying LNM in CC but also provide new clinical insights. Tumor markers, such as SCC-Ag, CEA, and CA125, directly reflect tumor burden and invasiveness. SCC-Ag promotes tumor cell migration by inducing epithelial-mesenchymal transition (EMT) [19, 20], a process that allows tumor cells to become more invasive and metastatic. Similarly, CEA facilitates metastasis by modulating inflammation and immune evasion [21], highlighting its role in the metastatic cascade. CA125, on the other hand, alters the extracellular matrix, enhancing tumor cell invasiveness [22, 23], emphasizing the critical interaction between tumor cells and the surrounding stroma during metastasis.

Inflammatory and immune-related indices, such as Neu, CRP, and FIB, underscore the role of the tumor-associated inflammatory microenvironment in LNM formation. Tumor-associated neutrophils (TANs) promote tumor invasion via cytokine release [24], while CRP enhances metastatic potential by activating the complement system, bridging inflammation and tumor invasion. FIB facilitates metastasis by protecting tumor cells and promoting angiogenesis [25], creating favorable conditions for distant implantation. Platelets play a complex role in the metastatic microenvironment. By interacting with tumor cells, platelets form a protective barrier, shielding tumor cells from immune clearance and promoting angiogenesis through the release of vascular growth factors [26]. This not only enhances tumor cell survival but also supports distant metastasis.

Histopathological features, including differentiation status and depth of mesenchymal infiltration, directly reflect tumor aggressiveness. Poorly differentiated tumors are often associated with higher malignancy, while deeper mesenchymal infiltration indicates that the tumor has breached local barriers, entering the lymphatic and vascular systems and significantly increasing the risk of LNM [16-18]. Furthermore, depth of mesenchymal infiltration is strongly correlated with parametrial involvement, further emphasizing its role in tumor spread and invasion [28]. These findings align with the study by Yang et al. [27], which showed an increasing trend in the conditional survival rate of high-risk patients with LNM over time, highlighting the importance of early diagnosis and risk stratification. Early intervention in patients with poor differentiation and deep mesenchymal infiltration could significantly improve prognosis.

Previous studies have highlighted the potential of integrating clinical, imaging, and laboratory characteristics into predictive models to enhance the accuracy of LNM prediction in CC. Prognostic models, such as the Nomogram, have proven particularly effective in this regard. For example, the Nomogram developed by Deng et al. [29], which incorporated variables such as age, histological type, tumor grade, tumor size, and FIGO staging, demonstrated strong discriminatory ability, achieving an AUC of 0.723. Its performance was consistent across both training and validation datasets, underscoring the utility of combining multiple clinical parameters for accurate LNM prediction. Similarly, Dong et al. [30] integrated three-dimensional power Doppler ultrasonography (3D-PDU) parameters with clinical features to construct a Nomogram that achieved an AUC of 0.845, effectively predicting both LNM and lymphovascular space invasion (LVSI). These studies underscore the importance of multidimensional models in refining risk stratification and diagnostic precision, which aligns with the approach adopted in the current research.

Building on these advancements, Wenzel et al. [31] introduced multiple laboratory parameters to address the limitations of traditional single-factor prediction models. Their study emphasized the value of integrating laboratory markers, which provide dynamic insights into tumor biology and systemic inflammation, into predictive frameworks. This approach aligns with the methodology used in the current study, where biomarkers such as SCC-Ag, CEA, and CA125 were incorporated to enhance predictive performance. These markers not only reflect tumor burden and metastatic potential but also capture systemic and microenvironmental factors contributing to LNM, thereby enhancing the predictive model's accuracy and applicability.

However, accurate LNM prediction extends beyond diagnostic utility to treatment implications. Matsuo et al. [32] revealed that, despite the widespread use of concurrent chemoradiotherapy in early-stage CC patients with positive lymph nodes, its impact on survival rates remains limited. This finding underscores the need for precise risk stratification to avoid overtreatment and reduce unnecessary treatment burdens, particularly for low-risk patients. Misclassification of low-risk patients can lead to unnecessary lymph node dissection, increased surgical trauma, and the adverse effects of overtreatment, emphasizing the value of highly accurate predictive models in clinical decision-making. In this context, the current study leverages a comprehensive approach by integrating laboratory parameters, clinical features, and histopathological characteristics to construct a robust predictive framework. This multidimensional methodology not only addresses the limitations of traditional single-variable models but also aligns with prior evidence supporting the importance of combining clinical and laboratory data. By improving predictive precision, this model has the potential to guide individualized treatment strategies, minimizing unnecessary interventions for low-risk patients while ensuring timely and appropriate care for high-risk patients.

The CC LNM prediction model proposed in this study has significant clinical potential. Early diagnosis of LNM is crucial for treatment decisions. By predicting LNM risk preoperatively, physicians can more effectively assess the need for lymph node dissection and tailor treatment regimens to the specific condition of the patient. For high-risk patients, comprehensive preoperative treatments can be planned, while low-risk patients can avoid unnecessary lymph node dissections, thereby reducing surgical trauma and minimizing complications. Furthermore, the model's simplicity and ease of use make it highly applicable in clinical practice. Laboratory parameters such as SCC-Ag, CEA, and CA125, which are widely used in clinical settings, are simple to measure and cost-effective. Therefore, this model not only supports decision-making in expert teams at large hospitals but can also be applied in resource-limited regions, improving diagnostic and treatment outcomes for CC in primary care settings. Previous studies [29, 30, 33], including this one, have suggested that models combining laboratory parameters and clinicopathological features enhance LNM prediction accuracy. Future research should aim to optimize these models and validate their clinical value in large-scale, multi-center studies.

Although the prediction model developed in this study demonstrates relatively high accuracy in predicting LNM in CC, several limitations remain. First, the retrospective design may introduce selection bias. Despite efforts to select representative samples, retrospective studies inherently carry this risk. Second, the relatively small sample size may limit the model's generalizability. Although cross-validation was employed to assess the model, large-scale, multi-center prospective validation is necessary. Third, laboratory parameters may be influenced by individual differences and detection methods. Future research should optimize laboratory testing procedures to improve accuracy and consistency. To further improve the model's reliability, future studies should expand the sample size, conduct multi-center prospective research, and incorporate other clinical features (such as imaging and genomic data) to provide more accurate tools for predicting and treating CC.

In conclusion, the LNM risk prediction model for stage IA2-IIA1 CC developed in this study, based on laboratory parameters, demonstrates high predictive accuracy and provides valuable auxiliary decision-making support for clinicians. Despite its limitations, the model represents a significant advancement in the clinical management of CC, laying the foundation for further refinement and enhancement in future research.

Acknowledgements

This work was supported by Shanxi Provincial Science and Technology Cooperation and Exchange Key National Science and Technology Cooperation Project: Folic acid-mediated m6A methylation modification promotes pri-miR-125b maturation and targets YAP to regulate macrophage differentiation in cervical cancer (No.: 202304041101013), Shanxi Provincial Natural Science Research General Project: Based on single-cell sequencing technology, the mechanism of exosome delivery of miR125b regulating KLF5 on different subtypes of macrophages in the occurrence of cervical cancer (No.: 202203021221273) and Shanxi Provincial Youth Science Research Project: LncRNAAL137782.1 as ceRNA regulates ferroptosis in cervical cancer through miR-125b/TAZ axis (No.: 202403021212222).

Disclosure of conflict of interest

None.

Address correspondence to: Min Hao, Department of Obstetrics and Gynecology, The Second Hospital of Shanxi Medical University, No. 382, Wuyi Road, Xinghualing District, Taiyuan 030000, Shanxi, China. Tel: +86-13934569668; E-mail: 2yuanhaomin@163.com

References

[1] Lei J, Arroyo-Mühr LS, Lagheden C, Eklund C, Nordqvist Kleppe S, Elfström M, Andrae B, Sparén P, Dillner J and Sundström K. Human papillomavirus infection determines prognosis in cervical cancer. J Clin Oncol 2022; 40: 1522-1528.

[2] Volesky-Avellaneda KD, Laurie C, Tsyruk-Romano O, El-Zein M and Franco EL. Human papillomavirus detectability and cervical cancer prognosis: a systematic review and meta-analysis. Obstet Gynecol 2023; 142: 1055-1067.

[3] Bray F, Laversanne M, Sung H, Ferlay J, Siegel RL, Soerjomataram I and Jemal A. Global cancer statistics 2022: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin 2024; 74: 229-263.

[4] Wright JD, Matsuo K, Huang Y, Tergas AI, Hou JY, Khoury-Collado F, St Clair CM, Ananth CV, Neugut AI and Hershman DL. Prognostic performance of the 2018 international federation of gynecology and obstetrics cervical cancer staging guidelines. Obstet Gynecol 2019; 134: 49-57.

[5] Hu Y, Han Y, Shen Y, Chen J, Chen Y, Chen Y, Tang J, Xue M, Hong L, Cheng W, Wang D, Liang Z, Wang Y, Zhang Q, Xing H, Zhang Y, Yi C, Yu Z, Chen Y, Cui M, Ma C, Yang H, Li R, Long P, Zhao Y, Qu P, Tao G, Yang L, Wu S, Liu Z, Yang P, Lv W, Xie X, Ma D, Wang H and Li K. Neoadjuvant chemotherapy for patients with international federation of gynecology and obstetrics stages IB3 and IIA2 cervical cancer: a multicenter prospective trial. BMC Cancer 2022; 22: 1270.

[6] Dhomps A, Trecourt A and Tordo J. Cervix abscess mimicking cervical cancer explored with 18 F-FDG PET/CT and MRI. Clin Nucl Med 2023; 48: e237-e238.

[7] Liu S, Li R, Liu Q, Sun D, Yang H, Pan H, Wang L and Song S. Radiomics model of 18F-FDG PET/CT imaging for predicting disease-free survival of early-stage uterine cervical squamous cancer. Cancer Biomark 2022; 33: 249-259.

[8] Alonso-Espías M, Gorostidi M, García-Pineda V, Gracia M, Diestro MD, Siegrist J, Zapardiel I and Hernández A. Survival impact of sentinel lymph node biopsy in patients with early-stage cervical cancer. Reprod Sci 2024; 31: 3066-3073.

[9] Dyikanov D, Zaitsev A, Vasileva T, Wang I, Sokolov AA, Bolshakov ES, Frank A, Turova P, Golubeva O, Gantseva A, Kamysheva A, Shpudeiko P, Krauz I, Abdou M, Chasse M, Conroy T, Merriam NR, Alesse JE, English N, Shpak B, Shchetsova A, Tikhonov E, Filatov I, Radko A, Bolshakova A, Kachalova A, Lugovykh N, Bulahov A, Kilina A, Asanbekov S, Zheleznyak I, Skoptsov P, Alekseeva E, Johnson JM, Curry JM, Linnenbach AJ, South AP, Yang E, Morozov K, Terenteva A, Nigmatullina L, Fastovetz D, Bobe A, Balabanian L, Nomie K, Yong ST, Davitt CJH, Ryabykh A, Kudryashova O, Tazearslan C, Bagaev A, Fowler N, Luginbuhl AJ, Ataullakhanov RI and Goldberg MF. Comprehensive peripheral blood immunoprofiling reveals five immunotypes with immunotherapy response characteristics in patients with cancer. Cancer Cell 2024; 42: 759-779, e712.

[10] Zhang Q, Wang Z, Tang H, Zhang B, Yue C, Gao J and Ying C. Serum CYFRA21-1 and SCC-Ag levels in women during pregnancy and their diagnostic value for cervical cancer. Cancer Treat Res Commun 2024; 38: 100786.

[11] Yang S, Liu C, Li C and Hua K. Nomogram predicting lymph node metastasis in the early-stage cervical cancer. Front Med (Lausanne) 2022; 9: 866283.

[12] Jiang K, Ai Y, Li Y and Jia L. Nomogram models for the prognosis of cervical cancer: a SEER-based study. Front Oncol 2022; 12: 961678.

[13] Bierer BE. Declaration of Helsinki-revisions for the 21st century. JAMA 2024; 333: 18-19.

[14] Mohamud A, Høgdall C and Schnack T. Prognostic value of the 2018 FIGO staging system

for cervical cancer. Gynecol Oncol 2022; 165: 506-513.
[15] Kim W, Chun S and Lee SA. Changes in cervical dysplasia, carcinoma in situ, and cervical cancer after expanding the National Cancer Screening Program to younger women in Korea. Int J Qual Health Care 2023; 35: mzad024.
[16] Huang BX and Fang F. Progress in the study of lymph node metastasis in early-stage cervical cancer. Curr Med Sci 2018; 38: 567-574.
[17] Wenzel HHB, Schnack TH, Van der Aa MA, Jensen PT, Høgdall CK, Hardie AN, Falconer H and Bekkers RLM; Dutch, dANish and sweDish gynaEcoLogIcal ONcology (DANDELION) research group. Risk factors for lymph node metastasis in women with FIGO 2018 IA cervical cancer with a horizontal spread of > 7 mm. Eur J Cancer 2024; 212: 115063.
[18] Zhao J, Cai J, Wang H, Dong W, Zhang Y, Wang S, He X, Sun S, Huang Y, Huang B, Willborn KC, Jiang P and Wang Z. Region-specific risk factors for pelvic lymph node metastasis in patients with stage IB1 cervical cancer. J Cancer 2021; 12: 2624-2632.
[19] Xu X, Liu F, Zhao X, Wang C, Li D, Kang L, Liu S and Zhang X. The value of multiparameter MRI of early cervical cancer combined with SCC-Ag in predicting its pelvic lymph node metastasis. Front Oncol 2024; 14: 1417933.
[20] Jiang CZ, Zheng K, Zhang YY, Yang J, Ye H and Peng X. (18)F-FDG PET/CT semi-quantitative parameters combined with SCC-Ag in predicting lymph node metastasis in stage I-II cervical cancer. Front Oncol 2024; 14: 1278464.
[21] Konishi I, Fujii S, Nanbu Y, Nonogaki H and Mori T. Mucin leakage into the cervical stroma may increase lymph node metastasis in mucin-producing cervical adenocarcinomas. Cancer 1990; 65: 229-237.
[22] Ran C, Sun J, Qu Y and Long N. Clinical value of MRI, serum SCCA, and CA125 levels in the diagnosis of lymph node metastasis and parauterine infiltration in cervical cancer. World J Surg Oncol 2021; 19: 343.
[23] Honda KI, Miyama M, Nishii Y, Tasaka R, Nakano Y, Umesaki N, Fukuda T, Ichimura T, Yasui T and Sumi T. Metastasis of cervical cancer indicated by elevation of serum CA125 produced by mediastinal lymph nodes: a case report. J Med Case Rep 2024; 18: 112.
[24] Ning Y, Chen Y, Tian T, Gao X, Liu X, Wang J, Chu H, Zhao C, Yang Y, Lei K, Ren H and Cui Z. S100A7 orchestrates neutrophil chemotaxis and drives neutrophil extracellular traps (NETs) formation to facilitate lymph node metastasis in cervical cancer patients. Cancer Lett 2024; 605: 217288.
[25] Yang Y, Li X, Qian H, Di G, Zhou R, Dong Y, Chen W and Ren Q. C-reactive protein as a prognostic biomarker for gynecologic cancers: a meta-analysis. Comput Intell Neurosci 2022; 2022: 6833078.
[26] Xu F, Li Y, Fan L, Ma J, Yu L, Yi H, Chen X, Wei W, Wu P, Liang L, Hu H, Xing H and Wang W. Preoperative SCC-Ag and thrombocytosis as predictive markers for pelvic lymphatic metastasis of squamous cervical cancer in early FIGO stage. J Cancer 2018; 9: 1660-1666.
[27] Yang XL, Wang MM, Kou LN, Lai H and Wu DJ. Conditional survival for high-risk early-stage cervical cancer patients with lymph node metastasis after hysterectomy. Curr Probl Cancer 2021; 45: 100756.
[28] Ma C, Zhang Y, Li R, Mao H and Liu P. Risk of parametrial invasion in women with early stage cervical cancer: a meta-analysis. Arch Gynecol Obstet 2018; 297: 573-580.
[29] Deng YR, Chen XJ, Xu CQ, Wu QZ, Zhang W, Guo SQ and Li LX. A preoperative nomogram predicting risk of lymph node metastasis for early-stage cervical cancer. BMC Womens Health 2023; 23: 568.
[30] Dong S, Peng YQ, Feng YN, Li XY, Gong LP, Zhang S, Du XS and Sun LT. Based on 3D-PDU and clinical characteristics nomogram for prediction of lymph node metastasis and lymph-vascular space invasion of early cervical cancer preoperatively. BMC Womens Health 2024; 24: 438.
[31] Wenzel HHB, Hardie AN, Moncada-Torres A, Høgdall CK, Bekkers RLM, Falconer H, Jensen PT, Nijman HW, van der Aa MA, Martin F, van Gestel AJ, Lemmens VEPP, Dahm-Kähler P, Alfonzo E, Persson J, Ekdahl L, Salehi S, Frøding LP, Markauskas A, Fuglsang K and Schnack TH. A federated approach to identify women with early-stage cervical cancer at low risk of lymph node metastases. Eur J Cancer 2023; 185: 61-68.
[32] Matsuo K, Nusbaum DJ, Machida H, Huang Y, Khetan V, Matsuzaki S, Klar M, Grubbs BH, Roman LD and Wright JD. Populational trends and outcomes of postoperative radiotherapy for high-risk early-stage cervical cancer with lymph node metastasis: concurrent chemo-radiotherapy versus radiotherapy alone. Am J Obstet Gynecol 2020; 222: 484.e481-484.e415.
[33] Wang Z, Zhang S, Ma Y, Li W, Tian J and Liu T. A nomogram prediction model for lymph node metastasis in endometrial cancer patients. BMC Cancer 2021; 21: 748.

Table S1. Clinical data of LNM and non-LNM patients

Variable	Total	LNM group (n = 151)	Non-LNM group (n = 473)	Statistic	P
Age	52.679±10.117	53.748±10.640	52.338±9.931	-1.493	0.136
Body mass index	24.005±2.994	23.969±3.023	24.016±2.988	0.168	0.867
Pausimenia					
Yes	246	62	184	0.223	0.636
No	378	89	289		
HPV infection					
Positive	531	143	388	14.493	< 0.001
Negative	93	8	85		
Differentiation degree					
Poorly differentiated	200	72	128	22.348	< 0.001
Moderately to well differentiated	424	79	345		
Maximum tumor diameter					
≥ 2 cm	444	113	331	1.315	0.252
< 2 cm	180	38	142		
Pathological type					
Squamous cell carcinoma	558	128	430	4.564	0.033
Others	66	23	43		
Depth of mesenchymal infiltration					
≥ 1/2	333	118	215	49.153	< 0.001
< 1/2	291	33	258		
Pathological staging					
IA2-IB	402	95	307	0.198	0.656
IIA1	222	56	166		

Note: LNM, lymph node metastasis.

Table S2. Laboratory parameters of LNM and non-LNM patients

Indicators	Total	LNM group (n = 151)	Non-LNM group (n = 473)	Statistic	P
SCC-Ag	4.746±2.463	7.088±3.014	3.998±1.675	-15.905	< 0.001
CEA (ng/mL)	38.882±10.143	44.871±10.014	36.970±9.423	-8.833	< 0.001
CA125 (U/mL)	10.403±2.505	12.324±2.725	9.790±2.089	-12.001	< 0.001
Neu ($\times 10^9$/L)	3.471±1.152	3.757±1.674	3.380±0.910	-3.535	< 0.001
PLT ($\times 10^9$/L)	196.335±55.553	228.152±67.775	186.178±46.759	-8.537	< 0.001
Lym ($\times 10^9$/L)	1.772±0.431	1.801±0.478	1.763±0.415	-0.940	0.348
NLR	1.95 [1.46, 2.51]	2.11 [1.40, 2.82]	1.91 [1.49, 2.43]	1.459	0.144
PLR	109.22 [85.26, 138.97]	127.20 [97.42, 169.77]	105.68 [82.93, 132.78]	5.133	< 0.001
FIB (g/L)	2.437±0.536	2.611±0.639	2.382±0.486	-4.653	< 0.001
Alb (g/L)	41.954±3.703	41.771±3.759	42.012±3.687	0.695	0.487
CRP (mg/L)	2.77 [1.67, 4.05]	4.22 [2.48, 6.29]	2.56 [1.54, 3.51]	7.917	< 0.001

Note: LNM, lymph node metastasis; SCC-Ag, squamous cell carcinoma antigen; CEA, carcinoembryonic antigen; CA125, cancer antigen 125; Neu, neutrophil count; PLT, platelet count; Lym, lymphocyte count; NLR, neutrophil-lymphocyte ratio; PLR, platelet-lymphocyte ratio; FIB, fibrinogen; Alb, albumin; CRP, C-reactive protein.