Am J Cancer Res 2026;16(4):1289-1311
www.ajcr.us /ISSN:2156-6976/ajcr0169504

Original Article

Development and internal-external validation of a nomogram for predicting postoperative 30-day malnutrition risk in cervical cancer patients: a retrospective cohort study

Fanyan Wei[1], Yuxia Du[2], Haifeng Hu[3], Wenxia Li[3], Beiyu Tuo[3]

[1]The 3rd Department of Gynecology, Northwest Women's and Children's Hospital, No. 1616, Yanxiang Road, Yanta District, Xi'an 710000, Shaanxi, China; [2]Department of Obstetrics and Gynecology, Xi'an International Medical Center, No. 777, Xitai Road, Gaoxin District, Xi'an 710000, Shaanxi, China; [3]Department of Oncology, Yan'an Traditional Chinese Medicine Hospital, No. 26, Xuanyuan Avenue, New District, Yan'an 716000, Shaanxi, China

Received October 23, 2025; Accepted February 9, 2026; Epub April 15, 2026; Published April 30, 2026

Abstract: Malnutrition at 30 days after surgery is common in cervical cancer patients and may adversely affect long-term outcomes. This retrospective study developed and validated an interpretable predictive nomogram for early identification of postoperative malnutrition (NRS-2002 ≥3) in patients with FIGO stage IB-IIA cervical cancer undergoing radical surgery. A total of 784 patients were included and divided into a training cohort (n=431), an internal validation cohort (n=180), and an independent external cohort (n=173). Clinical, sociodemographic, treatment-related, and laboratory variables were collected, and predictors were screened using univariate and multivariate logistic regression. Model discrimination, calibration, and clinical utility were assessed using the area under the receiver operating characteristic curve (AUC), Brier score, calibration analysis, decision curve analysis, and DeLong testing against albumin (ALB) alone. The final model incorporated eight independent predictors: age, body mass index (BMI), marital status, presence of a caregiver, parenteral nutrition support, lymph node metastasis, FIGO stage, and ALB. The nomogram achieved AUCs of 0.756, 0.742, and 0.807 in the training, internal validation, and external validation cohorts, respectively, with Brier scores ranging from 0.1859 to 0.2143, and showed stable net benefit across a wide threshold probability range (0.04-0.99). In the pooled sample, the nomogram significantly outperformed ALB alone (P<0.001). SHapley Additive exPlanations (SHAP) analysis enhanced interpretability and identified ALB, age, and lymph node metastasis as the most influential features driving predictions. Among 725 patients (92.5%) with follow-up data, 77 deaths (10.6%) occurred, and survival analyses demonstrated that unmarried status (HR=2.21), lymph node metastasis (HR=4.74), higher FIGO stage (HR=5.15), poor differentiation (HR=2.12), and higher risk scores (HR=1.88) were independently associated with worse overall survival, whereas human papillomavirus positivity was protective (HR=0.63). These findings suggest that the proposed nomogram provides accurate and explainable prediction of postoperative malnutrition and may support early risk stratification as well as long-term prognostic assessment in cervical cancer patients.

Keywords: Cervical cancer, malnutrition, nomogram, machine learning, SHAP, external validation, overall survival

Introduction

Cervical cancer remains a leading cause of cancer-related death among women worldwide, with hundreds of thousands of new cases and deaths reported annually [1]. While screening and treatment for cervical cancer have improved, outcomes remain inconsistent. Prognosis depends on tumor stage, histological characteristics, and treatment approach, but nutritional status has emerged as an important independent predictor of survival [2, 3]. Malnutrition is common in patients with cervical cancer, and its causes are rarely singular. Over time, tumor-induced wasting leads to weight loss and muscle loss. Treatments also result in reduced intake and increased energy requirements: surgery can lead to short-term fasting and stress; radiation therapy often causes inflammation of the oral or intestinal

https://doi.org/10.62347/LMSM8392

mucosa; and chemotherapy often causes nausea and loss of taste, reducing food intake [4]. Stress and low mood also have an impact; anxiety and depression can reduce appetite and make people lose the desire to prepare food [5]. The consequences are not limited to weight loss; postoperative infections are more common, wound healing time is longer, hospital stays are longer, and costs are higher [6]. Malnourished patients often struggle to tolerate chemotherapy and radiotherapy, forcing clinicians to reduce dosages or postpone treatment cycles, sometimes leading to premature termination and poorer outcomes [7]. Later, patients remain weak, experience slow recovery, and suffer from a decline in quality of life [8]. Studies have repeatedly shown that patients with nutritional risks before surgery or during radiotherapy have lower survival rates in the cervical cancer population [9].

Early identification of patients at risk of malnutrition is crucial, as delayed intervention can diminish the effectiveness of any intervention. Most clinics rely on inadequate tools. Nutritional Risk Screening 2002 and Subjective Global Assessment are widely used, but both require trained staff, are time-consuming to administer at the bedside, and exhibit significant differences between assessors [10, 11]. They often provide warnings too late, when patients' weight and physical function have already declined, making nutritional intervention insufficient to fully reverse this decline [12]. Single laboratory tests, while seemingly simple, often fail to address the underlying issues. Serum albumin levels can change with inflammation, fluid shifts, and organ function, and therefore their results are often unreliable when used alone [13]. Studies of cervical cancer patients undergoing surgery have also reached similar conclusions: a single checklist or a single biomarker can miss patients who actually have problems but appear stable on written examinations [8]. A more effective system should begin upon admission, continue throughout hospitalization, and keep pace with treatment progress. Screening should be rapid, repeatable, and easily taught. Results should be updated as clinical conditions change so that the healthcare team can initiate oral supplements, immunonutritional support, or counseling at the appropriate time, rather than after two treatment cycles. Standardized scoring can reduce the training burden, narrow the variability among assessors, and make the process feasible in daily practice.

Machine learning offers a way to address these shortcomings. Models do not have to rely on a single score or a single analyte but can extract signals from routine variables such as age, body size, tumor and treatment details, vital signs, complete blood counts, and biochemical indicators. The interactions between these input variables are beyond the grasp of simple rules, and nonlinearity is the norm rather than the exception [14, 15]. Recent oncology reports have shown that such models achieve or exceed the accuracy of traditional tools in classifying patient nutritional risks. In this study, we set out to build and test a model for cervical cancer patients scheduled for surgery. We collected a retrospective cohort, selected clinically relevant and data-stable predictors, and trained the model against predefined discrimination and calibration targets. We then performed internal validation and added external testing to examine the model's generalizability while preventing overfitting. We used the SHapley Additive exPlanations (SHAP) method to show which variables increased and decreased risk in each patient. These results are easily accessible at the clinical level. Our goal is practical: to provide clinicians with an objective screening method that encourages timely nutritional care and helps patients adhere to treatment. Achieving this should improve patient outcomes.

Materials and methods

Data source

We conducted a retrospective cohort study at Northwest Women's and Children's Hospital, Xi'an International Medical Center and Yan'an Traditional Chinese Medicine Hospital, including cervical cancer patients who underwent surgery between January 2015 and January 2022. A total of 784 patients met the inclusion criteria. These patients were divided into three cohorts: a training set (n=431), an internal validation set (n=180), and an external validation set (n=173). The protocol was approved by the Medical Ethics Committee of Yan'an Traditional Chinese Medicine Hospital, exempting informed consent, and all procedures complied with the Declaration of Helsinki. Since the core of this study is the development of a predictive model

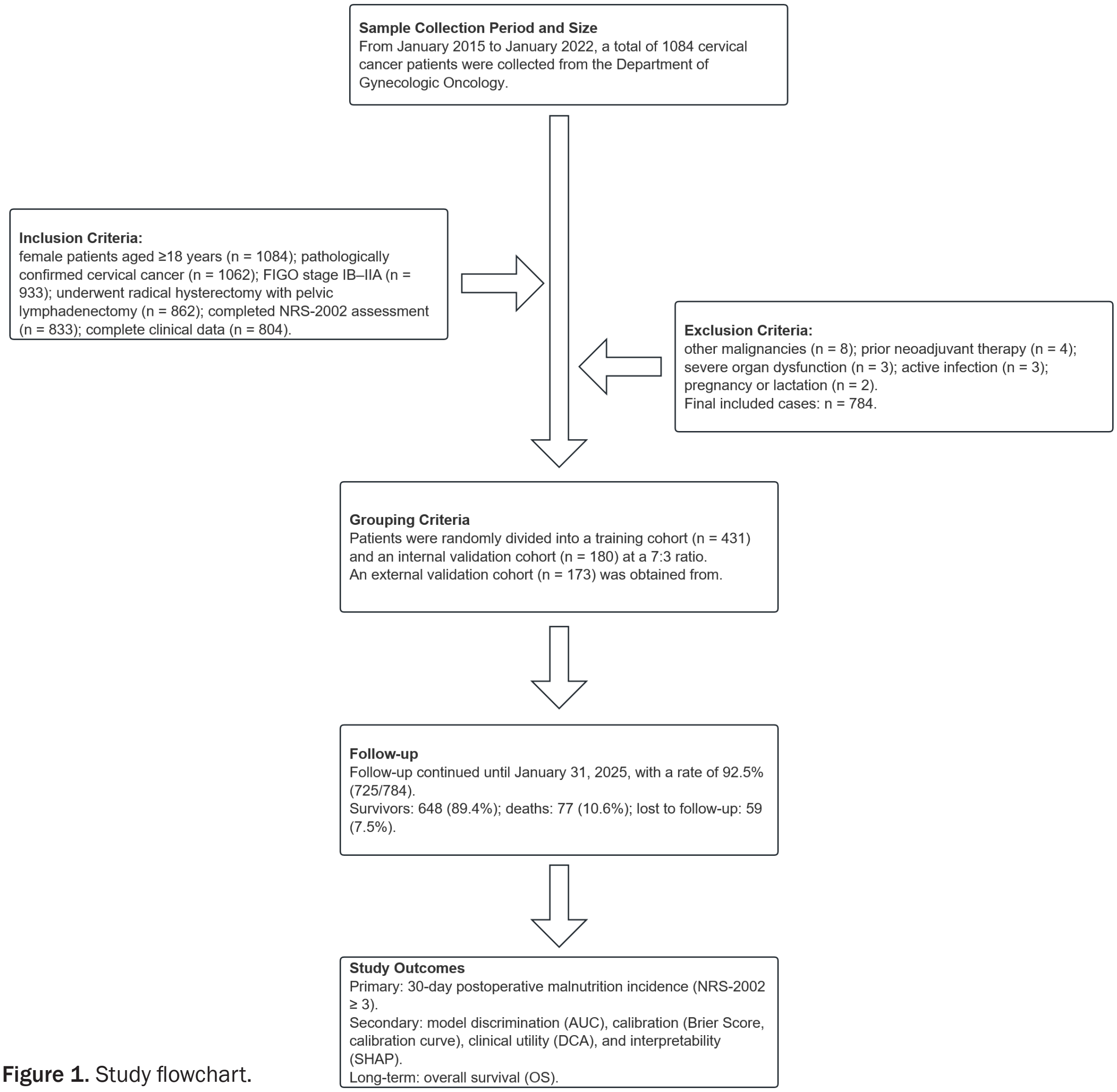

Figure 1. Study flowchart.

rather than the estimation of a single proportion, we used the variable per event (EPV) principle to determine the sample size rather than a single sample rate calculation. According to previous literature, the recommended EPV for model development is 10-20 [16]. Given that the incidence of postoperative malnutrition in cervical cancer patients is reported to be about 40%, and there are 8 candidate variables, when EPV=20, the calculation formula yields the minimum sample size N=8 × 20/0.40=400 cases ($N_{total} = \frac{k \times EPV}{P_{event}}$). The actual cohort of 784 patients significantly exceeded this threshold, ensuring statistical robustness of model development, internal, and external validations. Samples were randomly assigned in a 7:3 ratio to the training set (n=431) and the internal validation set (n=180), with 173 cases in the external set used to assess generalizability (**Figure 1**).

Inclusion and exclusion criteria

Inclusion criteria included: female patients aged ≥18 years; pathologically confirmed cervical cancer; FIGO stage IB or IIA; first radical hysterectomy and pelvic lymph node dissection; completion of the Nutritional Risk Screening 2002 (NRS-2002) assessment [17, 18]; and complete clinical data.

Exclusion criteria included: concomitant malignancies; preoperative neoadjuvant chemotherapy or radiotherapy; severe heart, liver, or kidney dysfunction or other systemic diseases;

preoperative acute/chronic infections; pregnancy or lactation; missing >20% of clinical or follow-up data; incomplete preoperative NRS-2002; refusal to participate or loss to follow-up.

Data collection

Clinical data and laboratory indicators were obtained from the hospital's electronic medical record (EMR) system and outpatient follow-up records. All data extraction was performed independently by two trained researchers, with a third researcher resolving any discrepancies. Collected variables included demographic information (age, body mass index [BMI], marital status, place of residence, education level), disease-related factors (FIGO stage, pathological type, differentiation degree, deep stromal invasion, lymphovascular invasion, lymph node metastasis, surgical margin status), treatment details (surgical procedure, preoperative parenteral nutrition support, human papillomavirus [HPV] status), postoperative complications (lymphedema, pain), functional status (Eastern Cooperative Oncology Group performance status [ECOG-PS]), and social support (presence or absence of caregivers). Preoperative fasting (8-12 hours) venous blood samples were tested by the hospital laboratory, including albumin (ALB). In this study, some cervical cancer patients received parenteral nutrition (PN) within 48 hours after surgery. The main reasons were as follows: (1) the gastrointestinal tract was not fully recovered and could not take in enough nutrition through oral/enteral routes; (2) there was potential nutritional risk or hypoalbuminemia, requiring short-term intravenous supplementation; (3) the doctor decided according to the patient's overall condition (e.g., advanced age, low BMI, gastrointestinal symptoms, comorbidities).

Laboratory tests

ALB was measured using a fully automated biochemical analyzer (Cobas 8000 c702 module, Roche Diagnostics, Basel, Switzerland) and matching test kits (Roche Diagnostics, Basel, Switzerland) via the bromocresol green (BCG) colorimetric method, with a reference range of 40-55 g/L. Human papillomavirus (HPV) detection was performed using real-time quantitative PCR (Cobas 4800 HPV system, Roche Diagnostics, Basel, Switzerland).

Nutritional risk assessment

Nutritional risk screening using the NRS-2002 was conducted at the outpatient follow-up visit 30 days post-surgery. The NRS-2002 score integrates the degree of nutritional impairment, disease severity, and age-adjusted factors, with a score range of 0-7 [17, 18]. According to the criteria, patients with a score ≥3 were classified as malnourished, and those with a score <3 were classified as nutritionally normal. Routine single nutritional indicators (such as ALB) were collected routinely through fasting blood sampling upon admission; general information was retrieved from the EMR. All assessments were performed using standardized procedures by uniformly trained clinical nurses to ensure consistency.

Follow-up

Long-term survival follow-up was conducted on enrolled patients via inpatient EMR, outpatient/inpatient records, telephone, or WeChat. Follow-up content included survival status, recurrence/metastasis, and the date of last visit. Follow-up frequency: every 3 months for the first 2 years post-surgery, every 6 months thereafter for 5 years, and annually thereafter. The primary follow-up methods were outpatient visits and telephone follow-ups; for patients who did not adhere to the schedule, nurses would contact them supplementarily via telephone/WeChat until their status was confirmed or they were lost to follow-up. The endpoint event was the date of last follow-up or the date of death; for patients without an event, the cutoff date was the date of last follow-up. The cutoff date was January 31, 2025. Of the 784 patients, 725 (92.5%) were successfully followed up, and 59 (7.5%) were lost to follow-up. Among the patients successfully followed up, 648 (89.4%) survived, and 77 (10.6%) died, indicating a high follow-up completion rate. The median follow-up time was calculated using the inverse Kaplan-Meier method, and overall survival (OS) was assessed using survival curves.

Outcome measures: Primary outcome measures: incidence of malnutrition at 30 days post-surgery (proportion of patients with NRS-2002 ≥3 points) and model discrimination ability (assessed by area under the ROC curve (AUC)). Secondary outcome measures: model indicators (sensitivity, specificity, positive predictive value (PPV), negative predictive value (NPV), F1 score, Youden index, accuracy); calibration (Brier score, calibration curve); utility (decision curve analysis (DCA), net benefit, interventions avoided per 100 patients); interpretability (SHAP value, feature ranking). In addition, overall survival, as a long-term prognostic indicator, was assessed using a Cox proportional hazards model to evaluate the impact of malnutrition and related factors.

Statistical analysis

Analysis was performed in R version 4.3.4 (Vienna, Austria, R Foundation for Statistical Computation), using packages including caret, pROC, rms, ggplot2, SHAPforxgboost, dplyr, survival, and survminer. A two-sided *p*-value <0.05 was considered statistically significant. Continuous variables were expressed as mean ± standard deviation ($\bar{x}\pm s$) or median [Q1, Q3]. Categorical variables were expressed as n (%). Between-group comparisons: independent samples t-tests or Mann-Whitney U tests were used for continuous variables; χ^2 tests or Fisher's exact test were used for categorical variables. Correlation analysis was performed using Pearson or Spearman correlation analysis; multicollinearity was assessed using the variance inflation factor (VIF >10 indicates severe multicollinearity). Risk factors for malnutrition were identified using univariate logistic regression (P<0.05). Statistically significant variables were then included in a multivariate forward stepwise (Wald) regression model for prediction, with odds ratio (OR) and 95% confidence interval (CI) as outputs. Model performance evaluation metrics included the AUC, (>0.7 indicating good discriminative ability), Brier score (<0.2 indicating good calibration), calibration consistency, DCA to assess the net benefit of threshold specificity. The DeLong test was used to compare the AUC of RISK and ALB; eigenvalue contribution and model interpretability were assessed using SHAP. The training set/internal validation set was randomly stratified in a 7:3 ratio to match the incidence of malnutrition. The external validation set was obtained from independent research centers to assess the model's generalizability. Samples with >20% missing data were excluded; samples with <5% missing data were processed using multiple imputation. OS analysis: all patients were included, defined as the time from the date of surgery to death from any cause (censored at the last follow-up or loss to follow-up). Survival rates were estimated using the Kaplan-Meier method and compared using the log-rank test. Independent influencing factors (HR, 95% CI) were identified using univariate and multivariate Cox proportional hazards regression models, and variables with P<0.05 were included in the multivariate analysis. The proportional hazards hypothesis was validated using Schoenfeld residuals. The effects of malnutrition, ALB, HPV, FIGO stage, lymph node status, and differentiation degree on OS were assessed. The goodness of fit of the models was evaluated using the likelihood ratio test, Wald test, and Score test.

Results

Analysis of clinical characteristics and malnutrition risk in cervical cancer patients

Of the 784 patients, 310 (39.54%) experienced postoperative malnutrition (NRS-2002 ≥3), while 474 (60.46%) had normal nutritional status. This study analyzed the association between clinical characteristics and the risk of malnutrition in cervical cancer patients. There were no significant differences in the incidence of malnutrition among the training set, validation set (split in a 7:3 ratio), and external validation set (P>0.05, see **Table 1**). Within each group, there were no significant differences in age, body mass index (BMI), marital status, presence or absence of caregivers, PN support, lymph node metastasis, postoperative lymphedema, postoperative pain, ECOG-PS, FIGO stage, pathological type, differentiation degree, surgical margin status, deep stromal invasion, lymphovascular invasion, surgical procedure, HPV status, place of residence, and education level (all P>0.05, **Table 1**).

Analysis of clinical characteristics and malnutrition in the training set of cervical cancer patients

In the training set, malnutrition was defined as a NRS-2002 score ≥3; a score <3 indicated normal nutritional status. Univariate analysis revealed significant associations between malnutrition and age (P<0.001), BMI (P<0.001), marital status (P<0.001), presence of caregivers (P<0.001), receipt of PN support (P<0.001), lymph node metastasis (P<0.001), FIGO stage (P=0.012), and ALB level (P<0.001). In con-

Table 1. Analysis of clinical characteristics and their association with malnutrition risk in cervical cancer patients

Variable	Total	Training Set (n=431)	Validation Set (n=180)	External Validation Set (n=173)	χ^2 Value	*P* Value
Malnutrition					5.773	0.056
Yes	310 (39.54%)	161 (37.35%)	85 (47.22%)	64 (36.99%)		
No	474 (60.46%)	270 (62.65%)	95 (52.78%)	109 (63.01%)		
Age					3.767	0.152
≥60 years	285 (36.35%)	152 (35.27%)	76 (42.22%)	57 (32.95%)		
<60 years	499 (63.65%)	279 (64.73%)	104 (57.78%)	116 (67.05%)		
BMI (kg/m^2)					0.140	0.932
<18.5	125 (15.94%)	70 (16.24%)	29 (16.11%)	26 (15.03%)		
≥18.5	659 (84.06%)	361 (83.76%)	151 (83.89%)	147 (84.97%)		
Marital Status					2.550	0.279
Married	606 (77.30%)	324 (75.17%)	145 (80.56%)	137 (79.19%)		
Other	178 (22.70%)	107 (24.83%)	35 (19.44%)	36 (20.81%)		
Caregiver Presence					5.600	0.061
Family	656 (83.67%)	364 (84.45%)	141 (78.33%)	151 (87.28%)		
Other	128 (16.33%)	67 (15.55%)	39 (21.67%)	22 (12.72%)		
Parenteral Nutrition Support					1.096	0.578
Yes	188 (23.98%)	108 (25.06%)	38 (21.11%)	42 (24.28%)		
No	596 (76.02%)	323 (74.94%)	142 (78.89%)	131 (75.72%)		
Lymph Node Metastasis					1.747	0.417
Yes	383 (48.85%)	202 (46.87%)	90 (50.00%)	91 (52.60%)		
No	401 (51.15%)	229 (53.13%)	90 (50.00%)	82 (47.40%)		
Postoperative Lymphedema					2.308	0.315
Yes	246 (31.38%)	144 (33.41%)	55 (30.56%)	47 (27.17%)		
No	538 (68.62%)	287 (66.59%)	125 (69.44%)	126 (72.83%)		
Postoperative Pain					4.292	0.117
Yes	124 (15.82%)	78 (18.10%)	26 (14.44%)	20 (11.56%)		
No	660 (84.18%)	353 (81.90%)	154 (85.56%)	153 (88.44%)		
ECOG-PS					0.444	0.801
0-1	575 (73.34%)	312 (72.39%)	134 (74.44%)	129 (74.57%)		
2	209 (26.66%)	119 (27.61%)	46 (25.56%)	44 (25.43%)		
FIGO Stage					0.431	0.806
IB	400 (51.02%)	218 (50.58%)	90 (50.00%)	92 (53.18%)		
IIA	384 (48.98%)	213 (49.42%)	90 (50.00%)	81 (46.82%)		
Pathology Type					1.284	0.526
Squamous Cell Carcinoma	644 (82.14%)	356 (82.60%)	143 (79.44%)	145 (83.82%)		
Adenocarcinoma or Other	140 (17.86%)	75 (17.40%)	37 (20.56%)	28 (16.18%)		
Differentiation Grade					5.858	0.053
Poor	285 (36.35%)	171 (39.68%)	53 (29.44%)	61 (35.26%)		
Moderate/High	499 (63.65%)	260 (60.32%)	127 (70.56%)	112 (64.74%)		
Surgical Margin					0.855	0.652
Positive	53 (6.76%)	31 (7.19%)	13 (7.22%)	9 (5.20%)		
Negative	731 (93.24%)	400 (92.81%)	167 (92.78%)	164 (94.80%)		
Deep Stromal Invasion					3.237	0.198
≥1/2	510 (65.05%)	288 (66.82%)	107 (59.44%)	115 (66.47%)		
<1/2	274 (34.95%)	143 (33.18%)	73 (40.56%)	58 (33.53%)		

Lymphovascular Invasion					0.854	0.652
Yes	234 (29.85%)	128 (29.70%)	58 (32.22%)	48 (27.75%)		
No	550 (70.15%)	303 (70.30%)	122 (67.78%)	125 (72.25%)		
Surgical Approach					1.962	0.375
Laparoscopic	336 (42.86%)	184 (42.69%)	84 (46.67%)	68 (39.31%)		
Open	448 (57.14%)	247 (57.31%)	96 (53.33%)	105 (60.69%)		
HPV					4.025	0.134
Positive	546 (69.64%)	298 (69.14%)	135 (75.00%)	113 (65.32%)		
Negative	238 (30.36%)	133 (30.86%)	45 (25.00%)	60 (34.68%)		
Residence					1.841	0.398
Rural	280 (35.71%)	151 (35.03%)	60 (33.33%)	69 (39.88%)		
Urban	504 (64.29%)	280 (64.97%)	120 (66.67%)	104 (60.12%)		
Education Level					0.789	0.674
≥ High School	542 (69.13%)	300 (69.61%)	127 (70.56%)	115 (66.47%)		
< High School	242 (30.87%)	131 (30.39%)	53 (29.44%)	58 (33.53%)		
ALB					4.614	0.100
≥40	190 (24.23%)	100 (23.20%)	54 (30.00%)	36 (20.81%)		
<40	594 (75.77%)	331 (76.80%)	126 (70.00%)	137 (79.19%)		

Note: ALB, albumin; BMI, body mass index; ECOG-PS, Eastern Cooperative Oncology Group Performance Status; FIGO, International Federation of Gynecology and Obstetrics.

trast, postoperative lymphedema (P=0.705), postoperative pain (P=0.184), ECOG-PS (P=0.145), pathological type (P=0.191), differentiation degree (P=0.162), surgical margin (P=0.823), deep stromal invasion (P=0.070), lymphovascular invasion (P=0.634), surgical procedure (P=0.175), HPV status (P=0.712), place of residence (P=0.769), and education level (P=0.525) showed no significant association (all P>0.05, **Table 2**).

Correlation and VIF analysis in the training set of cervical cancer patients

Correlation analysis in the training set showed weak overall associations among variables (absolute coefficient <0.15), with only a few statistically significant but small-effect-size correlations. Specifically, age was weakly negatively correlated with the ECOG-PS score (r=-0.131, P=0.007), indicating that older patients had slightly better performance status scores. BMI was weakly negatively correlated with marital status (r=-0.111, P=0.021), meaning that unmarried individuals tend to have lower BMIs. Postoperative lymphedema was weakly positively correlated with the FIGO stage (r=0.120, P=0.013), while FIGO stage was weakly negatively correlated with lymphovascular invasion (r=-0.119, P=0.013) and HPV status (r=-0.098, P=0.043). Pathological type was negatively correlated with surgical method (r=-0.099, P=0.040); differentiation degree was weakly negatively correlated with lymphovascular invasion (r=-0.102, P=0.035), HPV status (r=-0.095, P=0.049), and education level (r=-0.103, P=0.032), but weakly positively correlated with ALB (r=0.105, P=0.030). Marital status was also weakly negatively correlated with place of residence (r=-0.096, P=0.047). Other correlations were not statistically significant (P≥0.05). Overall, linear correlations were minimal, and no significant multicollinearity was observed (P>0.05, **Table 3** and **Figure 2**).

Univariable and multivariable logistic regression analysis in the training set of cervical cancer patients

Univariate and multivariate logistic regression were used to assess the association between clinical characteristics and the risk of malnutrition. In the univariate analysis, age, BMI, marital status, presence of a caregiver, PN support, lymph node metastasis, FIGO stage, and ALB were significantly associated with this risk - specifically, age (P<0.001), BMI (P<0.001), marital status (P<0.001), presence of a caregiver (P<0.001), PN support (P<0.001), lymph node metastasis (P=0.001), and ALB (P<0.001) were key risk factors, as was FIGO stage (P=0.013). Other factors, including postoperative lymph-

Table 2. Analysis of clinical characteristics and malnutrition in the training set of cervical cancer patients

Variable	Total	Malnourished Group (N=161)	Normal Group (N=270)	χ^2 Value	*P* Value
Age				12.879	<0.001
≥60 years	152 (35.27%)	74 (45.96%)	78 (28.89%)		
<60 years	279 (64.73%)	87 (54.04%)	192 (71.11%)		
BMI (kg/m^2)				12.038	<0.001
<18.5	70 (16.24%)	39 (24.22%)	31 (11.48%)		
≥18.5	361 (83.76%)	122 (75.78%)	239 (88.52%)		
Marital Status				15.408	<0.001
Married	324 (75.17%)	104 (64.60%)	220 (81.48%)		
Other	107 (24.83%)	57 (35.40%)	50 (18.52%)		
Caregiver Presence				12.708	<0.001
Family	364 (84.45%)	123 (76.40%)	241 (89.26%)		
Other	67 (15.55%)	38 (23.60%)	29 (10.74%)		
Parenteral Nutrition Support				11.342	<0.001
Yes	108 (25.06%)	55 (34.16%)	53 (19.63%)		
No	323 (74.94%)	106 (65.84%)	217 (80.37%)		
Lymph Node Metastasis				10.896	<0.001
Yes	202 (46.87%)	92 (57.14%)	110 (40.74%)		
No	229 (53.13%)	69 (42.86%)	160 (59.26%)		
Postoperative Lymphedema				0.143	0.705
Yes	144 (33.41%)	52 (32.30%)	92 (34.07%)		
No	287 (66.59%)	109 (67.70%)	178 (65.93%)		
Postoperative Pain				1.765	0.184
Yes	78 (18.10%)	24 (14.91%)	54 (20.00%)		
No	353 (81.90%)	137 (85.09%)	216 (80.00%)		
ECOG-PS				2.127	0.145
0-1	312 (72.39%)	110 (68.32%)	202 (74.81%)		
2	119 (27.61%)	51 (31.68%)	68 (25.19%)		
FIGO Stage				6.263	0.012
IIA	218 (50.58%)	94 (58.39%)	124 (45.93%)		
IB	213 (49.42%)	67 (41.61%)	146 (54.07%)		
Pathology Type				1.713	0.191
Squamous Cell Carcinoma	356 (82.60%)	128 (79.50%)	228 (84.44%)		
Adenocarcinoma or Other	75 (17.40%)	33 (20.50%)	42 (15.56%)		
Differentiation Grade				1.959	0.162
Poor	171 (39.68%)	57 (35.40%)	114 (42.22%)		
Moderate/High	260 (60.32%)	104 (64.60%)	156 (57.78%)		
Surgical Margin				0.050	0.823
Positive	31 (7.19%)	11 (6.83%)	20 (7.41%)		
Negative	400 (92.81%)	150 (93.17%)	250 (92.59%)		
Deep Stromal Invasion				3.294	0.070
≥1/2	288 (66.82%)	99 (61.49%)	189 (70.00%)		
<1/2	143 (33.18%)	62 (38.51%)	81 (30.00%)		
Lymphovascular Invasion				0.227	0.634
Yes	128 (29.70%)	50 (31.06%)	78 (28.89%)		
No	303 (70.30%)	111 (68.94%)	192 (71.11%)		

Surgical Approach				1.837	0.175
Laparoscopic	184 (42.69%)	62 (38.51%)	122 (45.19%)		
Open	247 (57.31%)	99 (61.49%)	148 (54.81%)		
HPV				0.137	0.712
Positive	307 (71.23%)	113 (70.19%)	194 (71.85%)		
Negative	124 (28.77%)	48 (29.81%)	76 (28.15%)		
Residence				0.086	0.769
Rural	151 (35.03%)	55 (34.16%)	96 (35.56%)		
Urban	280 (64.97%)	106 (65.84%)	174 (64.44%)		
Education Level				0.404	0.525
≥ High School	300 (69.61%)	115 (71.43%)	185 (68.52%)		
< High School	131 (30.39%)	46 (28.57%)	85 (31.48%)		
ALB				13.119	<0.001
≥40	100 (23.20%)	22 (13.66%)	78 (28.89%)		
<40	331 (76.80%)	139 (86.34%)	192 (71.11%)		

Note: ALB, albumin; BMI, body mass index; ECOG-PS, Eastern Cooperative Oncology Group Performance Status; FIGO, International Federation of Gynecology and Obstetrics.

Table 3. VIF analysis and variable assignment in the training set of cervical cancer patients

Variable	VIF	Interpretation	Assignment Content
Age (years)	1.051	Low multicollinearity	≥60 years = 1, <60 years = 0
BMI (kg/m^2)	1.038	Low multicollinearity	<18.5=1, ≥18.5=0
Marital_status	1.077	Low multicollinearity	Married = 1, Other = 0
Caregiving_situation	1.064	Low multicollinearity	Family = 1, Other = 0
Parenteral_nutrition_support	1.051	Low multicollinearity	Yes = 1, No = 0
Lymph_node_metastasis	1.061	Low multicollinearity	Yes = 1, No = 0
Postoperative_lymphedema	1.037	Low multicollinearity	Yes = 1, No = 0
Postoperative_pain	1.045	Low multicollinearity	Yes = 1, No = 0
ECOG_PS	1.073	Low multicollinearity	0-1=1, 2=0
FIGO_staging	1.088	Low multicollinearity	IIA=1, IB=0
Pathological_type	1.039	Low multicollinearity	Squamous cell carcinoma = 1, Adenocarcinoma or other = 0
Differentiation_grade	1.056	Low multicollinearity	Poor = 1, Moderate/High = 0
Surgical_margin	1.041	Low multicollinearity	Positive = 1, Negative = 0
Deep_stromal_invasion	1.046	Low multicollinearity	≥1/2=1, <1/2=0
Lymphovascular_invasion	1.05	Low multicollinearity	Yes = 1, No = 0
Surgical_approach	1.041	Low multicollinearity	Laparoscopic = 1, Open = 0
HPV	1.041	Low multicollinearity	Positive = 1, Negative = 0
Place_of_residence	1.037	Low multicollinearity	Rural = 1, Urban = 0
Education_level	1.044	Low multicollinearity	≥ High school = 1, < High school = 0
ALB	1.063	Low multicollinearity	≥40=1, <40=0

Note: BMI, body mass index; ECOG-PS, Eastern Cooperative Oncology Group Performance Status; FIGO, International Federation of Gynecology and Obstetrics; HPV, human papillomavirus; ALB, albumin.

edema (P=0.705), postoperative pain (P=0.185), ECOG-PS (P=0.145), pathological type (P=0.192), differentiation degree (P=0.306), surgical margin (P=0.823), deep stromal invasion (P=0.137), lymphovascular invasion (P=0.634), surgical procedure (P=0.176), HPV (P=0.712), place of residence (P=0.769), and education level (P=0.525), were not statistically significant (**Table 4**). Multivariate analysis showed that age (P<0.001), BMI (P=0.007), marital status (P=0.002), presence of caregiver (P<0.001), PN support (P<0.001), lymph node metastasis (P=0.001), and ALB (P<0.001) were independent risk factors; FIGO stage (P=0.004) remained significant, while the other factors were not (see **Table 4**). Using

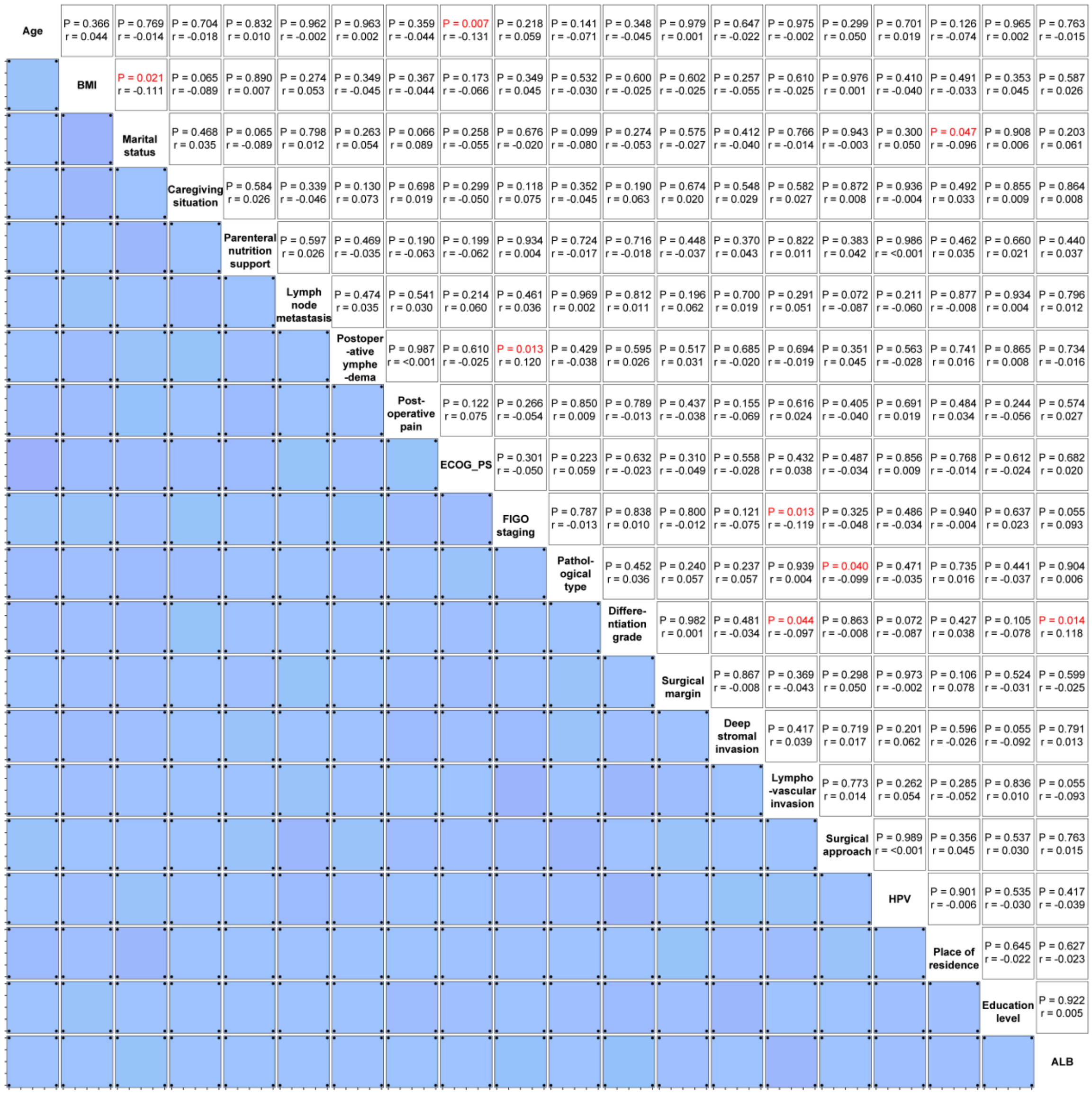

Figure 2. Correlation analysis matrix among clinical characteristics in the training group of cervical cancer patients. Note: BMI, Body Mass Index; ECOG-PS, Eastern Cooperative Oncology Group Performance Status; FIGO, International Federation of Gynecology and Obstetrics; HPV, Human Papillomavirus; ALB, Albumin.

the reference category as a baseline, the direction of the regression coefficients was clearly defined and clinically interpretable (≥60 years, BMI <18.5 kg/m^2, married, with family caregiver, receiving PN support, positive lymph nodes, FIGO stage IIA, ALB ≥40 g/L). Adults under 60 years of age with a BMI of 18.5 kg/m^2 had a lower probability of malnutrition. This suggests that older age and being underweight increase the likelihood of malnutrition. Unmarried status and non-family care significantly increase the risk, while family care reduces the risk. Patients receiving intravenous nutrition have a higher risk of malnutrition than those receiving conventional nutrition. The study suggests this finding may be due to reverse causality, as intravenous nutrition is only administered when patients are perceived as nutritionally vulnerable. Lymph node metastasis and FIGO staging increase the risk of malnutrition; albumin levels reflect the effectiveness of professional nutritionist intervention.

ROC curve analysis and performance evaluation of models in training, validation, and external validation sets

Figure 3 depicts the ROC curves and performance metrics for the training, validation, and

Table 4. Univariable and multivariable logistic regression analysis in the training set of cervical cancer patients

Characteristics	Total (N)	Univariate Analysis		Multivariable Analysis	
		OR (95% CI)	*P* Value	OR (95% CI)	*P* Value
Age	431				
≥60 years	152	Reference		Reference	
<60 years	279	0.478 (0.318-0.717)	<0.001	0.449 (0.285-0.706)	<0.001
BMI	431				
<18.5	70	Reference		Reference	
≥18.5	361	0.406 (0.241-0.682)	<0.001	0.452 (0.253-0.808)	0.007
Marital_status	431				
Married	324	Reference		Reference	
Other	107	2.412 (1.544-3.766)	<0.001	2.190 (1.332-3.600)	0.002
Caregiving_situation	431				
Family caregiver	364	Reference		Reference	
Other caregiver	67	2.567 (1.511-4.361)	<0.001	2.986 (1.631-5.468)	<0.001
Parenteral_nutrition_support	431				
Yes	108	Reference		Reference	
No	323	0.471 (0.302-0.733)	<0.001	0.407 (0.246-0.671)	<0.001
Lymph_node_metastasis	431				
Yes	202	Reference		Reference	
No	229	0.516 (0.347-0.766)	0.001	0.485 (0.313-0.753)	0.001
Postoperative_lymphedema	431				
Yes	144	Reference			
No	287	1.083 (0.715-1.641)	0.705		
Postoperative_pain	431				
Yes	78	Reference			
No	353	1.427 (0.843-2.416)	0.185		
ECOG_PS	431				
0-1	312	Reference			
2	119	1.377 (0.895-2.119)	0.145		
FIGO_staging	431				
IIA	218	Reference		Reference	
IB	213	0.605 (0.408-0.898)	0.013	0.520 (0.333-0.813)	0.004
Pathological_type	431				
Squamous cell carcinoma	356	Reference			
Adenocarcinoma or other	75	1.400 (0.845-2.318)	0.192		
Differentiation_grade	431				
Poor	166	Reference			
Moderate/High	265	1.235 (0.824-1.851)	0.306		
Surgical_margin	431				
Positive	31	Reference			
Negative	400	1.091 (0.509-2.340)	0.823		
Deep_stromal_invasion	431				
≥1/2	284	Reference			
<1/2	147	1.363 (0.906-2.050)	0.137		
Lymphovascular_invasion	431				
Yes	128	Reference			
No	303	0.902 (0.590-1.380)	0.634		

Surgical_approach	431				
Laparoscopic	184	Reference			
Open	247	1.316 (0.884-1.959)	0.176		
HPV	431				
Positive	307	Reference			
Negative	124	1.084 (0.706-1.666)	0.712		
Place_of_residence	431				
Rural	151	Reference			
Urban	280	1.063 (0.706-1.602)	0.769		
Education_level	431				
≥ High school	300	Reference			
< High school	131	0.871 (0.568-1.335)	0.525		
ALB	431				
≥40	100	Reference		Reference	
<40	331	2.567 (1.524-4.322)	<0.001	3.323 (1.856-5.949)	<0.001

Note: BMI, body mass index; ECOG-PS, Eastern Cooperative Oncology Group Performance Status; FIGO, International Federation of Gynecology and Obstetrics; HPV, human papillomavirus; ALB, albumin.

external validation sets. The AUC values were 0.756, 0.742, and 0.807, respectively, indicating good model performance in each group. Training set: accuracy 73.09%, sensitivity 66.46%, specificity 77.04%, positive predictive value (PPV) 63.31%, negative predictive value (NPV) 79.39%, F1 score 64.85%, Youden index 43.50%. Validation set: accuracy 68.89%, sensitivity 57.65%, specificity 78.95%, PPV 71.01%, NPV 67.57%, F1 score 63.64%, Youden index 36.59%. External validation set: Accuracy 73.41%, sensitivity 65.63%, specificity 77.98%, PPV 63.64%, NPV 79.44%, F1 score 64.62%, Youden index 43.61% (see **Table 5**). The risk model formula is: Risk = -0.1410 + 0.8009 × Age + 0.7939 × BMI - 0.7837 × Marital_status - 1.0941 × Caregiving_situation + 0.9001 × Parenteral_nutrition_support + 0.7228 × Lymph_node_metastasis + 0.6542 × FIGO_staging - 1.2008 × ALB.

Model calibration and clinical utility assessment

We validated the model in three aspects. First, it demonstrated strong discriminative power: the C-index was 0.792 for the training set, 0.928 for the validation set, and 0.918 for the external test set. Likelihood ratio, Wald, and Score tests all showed significant differences (all P<0.001), further supporting this finding. Second, the calibration was effective. The Brier scores for the three datasets were 0.1859, 0.2143, and 0.1788, all below 0.22, with two below 0.19. Also, the Hosmer-Lemeshow goodness-of-fit test was significant in the training set (χ^2=485.33, df=8, P<0.001), the internal validation set (χ^2=238.57, df=8, P<0.001), and the external validation set (χ^2=261.69, df=8, P<0.001). Due to the large sample size, the smaller *p*-values likely reflect slight calibration bias rather than clinically significant bias. The calibration plot (**Figure 4**) shows that the predicted risk is highly consistent with the observed risk at each decimal place. Third, the decision curves showed the clinical benefit of the model. As shown in **Figure 5**, the model outperformed both the "treat all patients" and "do not treat any patients" strategies at multiple thresholds. At a threshold of 0.99, the model avoided unnecessary interventions in 62.27 (training set), 52.30 (internal validation set), and 62.63 (external validation set) out of every 100 patients. In short, the model effectively differentiates risks, demonstrates good calibration, and contributes to the decision-making process.

Importance of model features based on SHAP values and their impact on predictions

We used SHAP to rank and interpret the inputs (**Figure 6**). ALB ranked first, followed by age, lymph node status, PN, and FIGO stage. The direction of influence is clearly visible in the summary plot. In the SHAP summary plot, an increased predicted risk of malnutrition was associated with older age, lower BMI, being

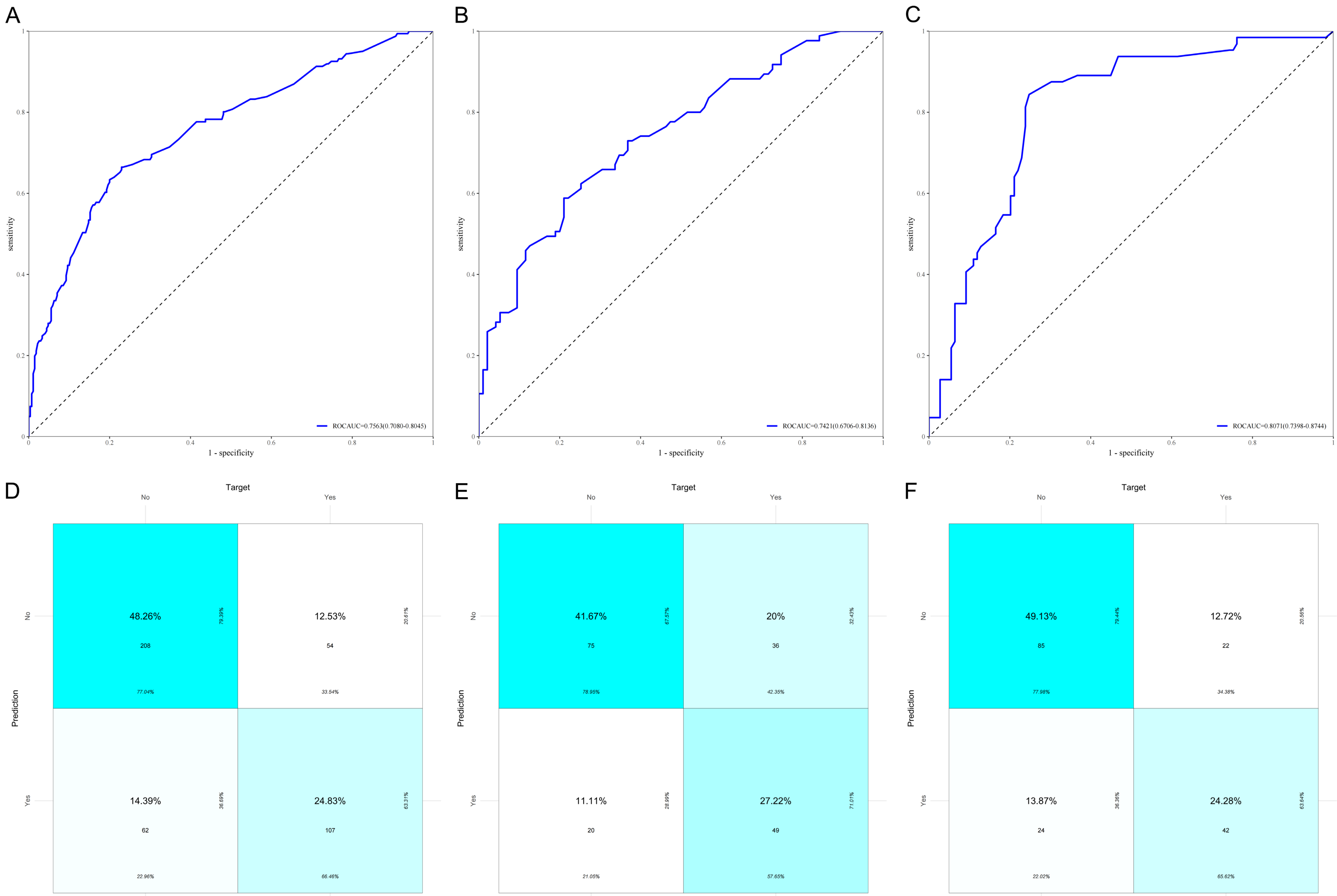

Figure 3. Model ROC curves for training, validation, and external validation groups. A: ROC curve for the training group. B: ROC curve for the validation group. C: ROC curve for the external validation group. D: Confusion matrix for the training group. E: Confusion matrix for the validation group. F: Confusion matrix for the external validation group. Note: ROC, Receiver Operating Characteristic curve; AUC, Area Under the Curve; PPV, Positive Predictive Value; NPV, Negative Predictive Value.

Table 5. Model performance metrics across training, validation, and external validation groups

Data Group	Positive Samples	Negative Samples	Accuracy	Sensitivity	Specificity	PPV	NPV	F1 Score	Youden Index
Training Group	161	270	73.09%	66.46%	77.04%	63.31%	79.39%	64.85%	43.50%
Validation Group	85	95	68.89%	57.65%	78.95%	71.01%	67.57%	63.64%	36.59%
External Validation Group	64	109	73.41%	65.63%	77.98%	63.64%	79.44%	64.62%	43.61%

Note: PPV, Positive Predictive Value; NPV, Negative Predictive Value.

unmarried, not receiving family care, receiving PN support, having lymph node metastasis, and a higher FIGO stage. In contrast, a decreased predicted risk of malnutrition was associated with higher albumin levels and family care. Among all factors, PN support was the only one with a positive SHAP value, indicating a higher estimated risk. This relationship between PN and malnutrition is unlikely to be direct, but rather an inverse causal relationship. This is because PN is only given to patients with severe clinical conditions and/or those considered to be at high nutritional risk, rather than being the cause of malnutrition. These patterns are consistent with clinical expectations, making the outputs more trustworthy in clinical practice. Quantitative SHAP importance confirms these patterns. The mean absolute SAP value for ALB was the highest (0.0728), followed by age (0.0696), lymph node metastasis (0.0676), PN support (0.0651), FIGO stage (0.0639), marital status (0.0562), care status (0.0544), and BMI (0.0427). This numerical contribution validates the ranking of feature importance seen in the summary plot and dependency plot.

SHAP value decomposition analysis for typical samples: comparison of normonourished and malnourished cases

Figure 7 details SHAP decomposition of two typical samples, explaining the prediction results for each sample. **Figure 7A** (Sample 1, malnourished; label = 0, prediction probability = 0.24, explained R^2=0.87) shows high interpretability. Blue bars for care status, ALB, age, and BMI (left: reduced risk) support normal nutrition; red bars for marital status, lymph node metastasis, FIGO stage, and preoperative support (right: increased risk) have a slight opposite effect, but the final classification is correct. **Figure 7B** (sample 274, malnutrition; label = 1, probability = 0.80, R^2=0.87) shows the opposite: red bars for care status, ALB, marital status, and metastasis increase risk; blue bars for support, BMI, age, and stage moderate the risk, but red bars dominate, accurately confirming malnutrition.

ROC curve comparison of RISK model and single ALB indicator in pooled samples

Figure 8 compares the ROC curves of the RISK model and ALB in the pooled sample (training + test + new data). **Figure 8A** (RISK): AUC=0.761 (95% CI: 0.727-0.795); the optimal cutoff has a sensitivity of 70.65%, a specificity of 71.94%, a Youden index of 42.59%, an accuracy of 71.43%, a precision of 70.65%, and an F1 score of 66.16%. **Figure 8B** (ALB): AUC=0.590 (95% CI: 0.550-0.630); the sensitivity is 81.29%, the specificity is 35.65%, the Youden index is 16.94%, the accuracy is 46.30%, the precision is 18.71%, and the F1 score is 21.60%. DeLong's test confirmed the superiority of the RISK model (difference = 0.171, 95% CI: 0.119-0.223; Z=6.423, P<0.001; **Table 6**), validating that the multivariate model is superior to the single indicator in assessing the risk of malnutrition.

Univariable and multivariable Cox regression analysis of factors influencing OS in cervical cancer patients

Univariate Cox regression analysis showed that marital status, lymph node metastasis, FIGO stage, differentiation degree, HPV status, ALB level, and comprehensive risk score were associated with OS (P<0.05). The multivariate model preserved the following independent influencing factors: married vs. unmarried (hazard ratio [HR]=2.210, 95% CI: 1.159-4.212, P=0.016); presence vs. absence of metastasis (HR=4.737, 95% CI: 2.689-8.346, P<0.001); FIGO stage; and overall risk score. Stage IB-IIA (higher stage) vs. lower stage (HR=5.150, 95% CI: 2.817-9.414, P<0.001); poorly differentiated vs. well-differentiated (HR=2.123, 95% CI:

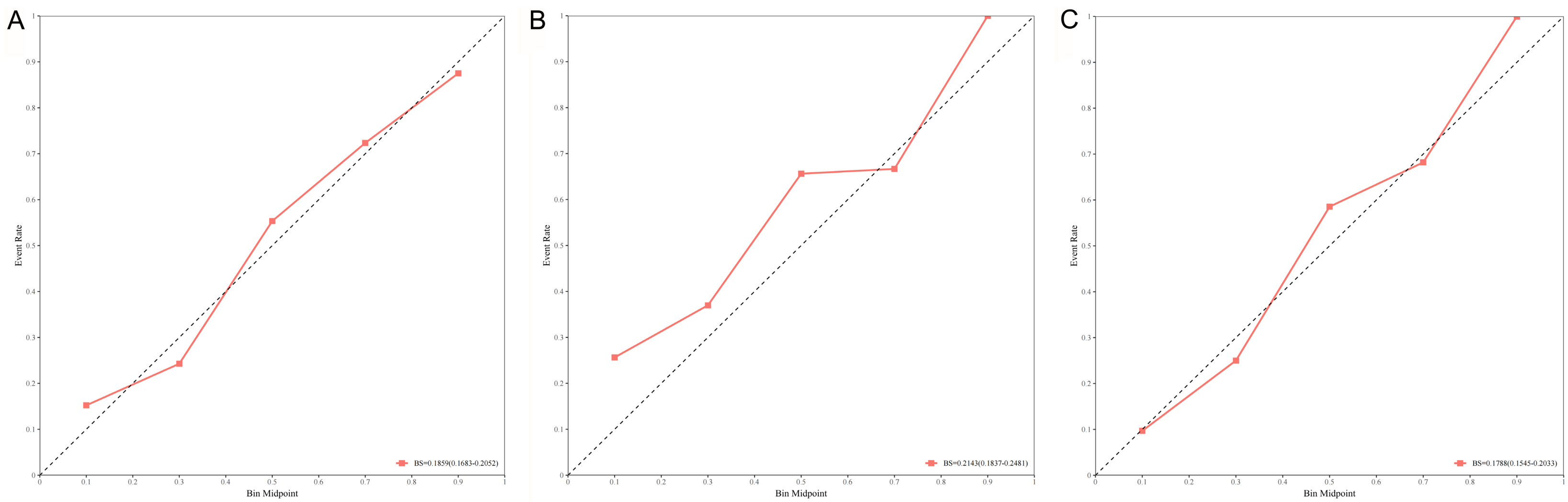

Figure 4. Calibration curves for training set, test set, and new samples. A: Calibration curve for the training set. B: Calibration curve for the test set. C: Calibration curve for new samples. Note: The calibration curve illustrates the concordance between predicted probabilities and observed outcomes.

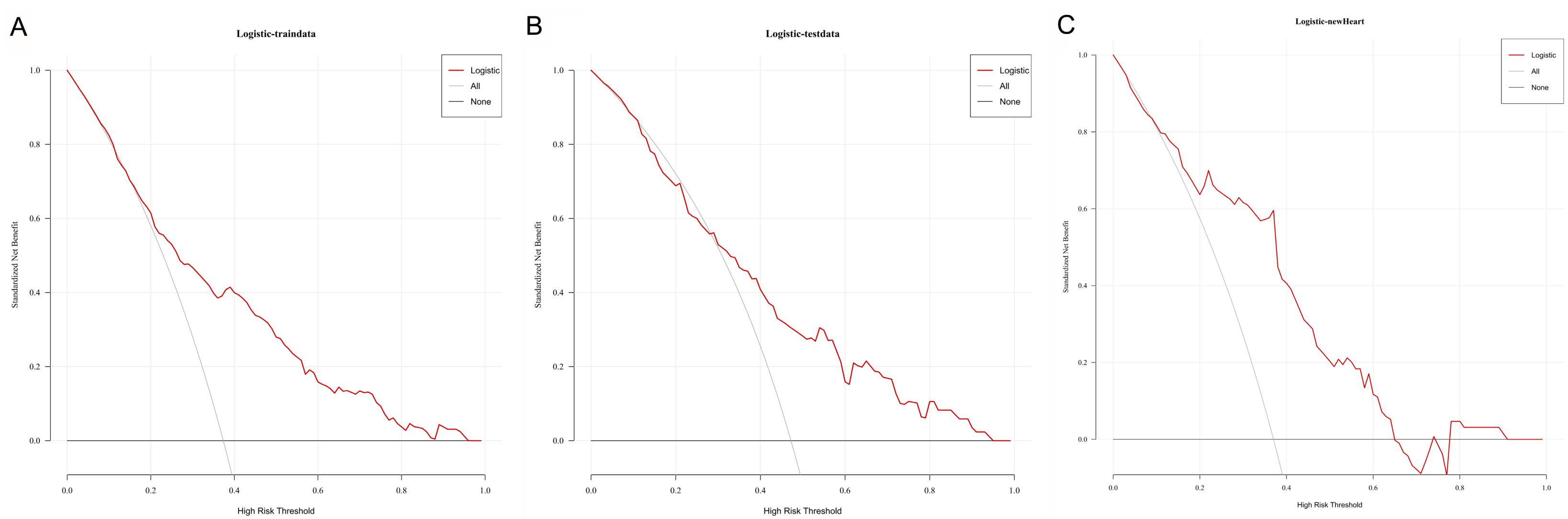

Figure 5. Decision Curve Analysis (DCA) for training set, test set, and new samples. A: DCA curve for the training set. B: DCA curve for the test set. C: DCA curve for new samples. Note: DCA, Decision Curve Analysis; The red curve represents the predictive model; The gray curve represents the treat-all strategy; The black curve represents the treat-none strategy; The horizontal axis represents threshold probability; The vertical axis represents net benefit.

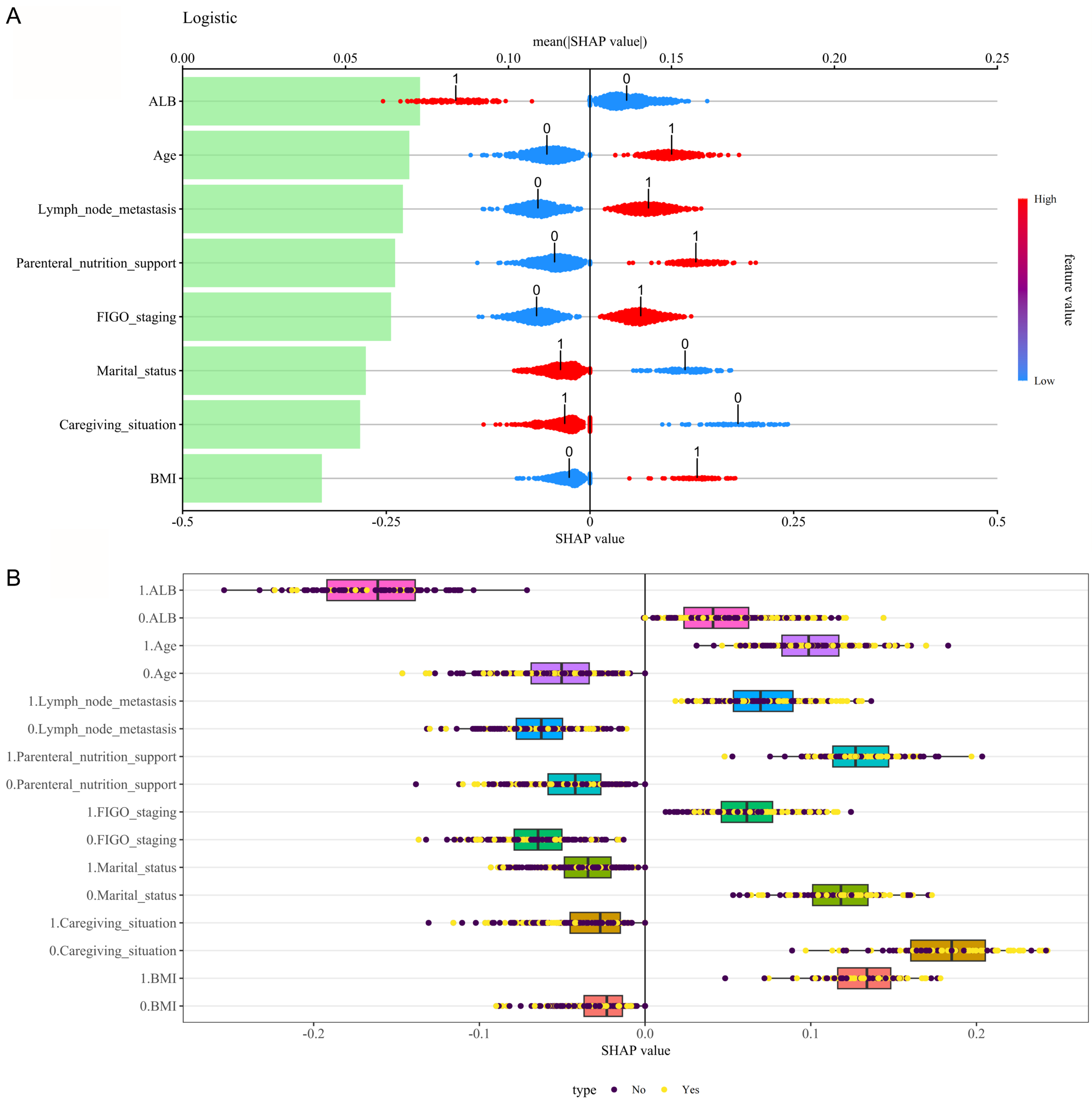

Figure 6. Feature importance and impact analysis based on SHAP values. A: Feature importance ranking. B: SHAP summary plot. Note: SHAP, SHapley Additive exPlanations; ALD, Preoperative Albumin; FIGO, International Federation of Gynecology and Obstetrics; BMI, Body Mass Index; Red indicates high feature values, blue indicates low feature values; Positive SHAP values indicate increased predictive risk, negative SHAP values indicate decreased predictive risk.

1.350-3.338, P=0.001); HPV positive vs. negative (HR=0.631, 95% CI: 0.399-0.996, P=0.048); risk score (HR=1.883, 95% CI: 1.180-3.007, P=0.008). Metastasis and advanced FIGO stage significantly increased mortality; HPV positivity showed a protective trend (higher ALB levels also showed a protective trend, P=0.074). Other factors, including age (P=0.071), BMI (P=0.633), parenteral nutrition support (P=0.780), caregiver status (P=0.349), ECOG-PS (P=0.948), surgical approach (P=0.561), place of residence (P=0.266), and education level (P=0.556), were not statistically significant in univariable analysis (**Table 7**).

Discussion

We developed a machine learning model that integrates factors such as age, BMI, marital status, caregiver type, PN support, lymph node metastasis, FIGO stage, and albumin to predict the risk of malnutrition in patients with cervi-

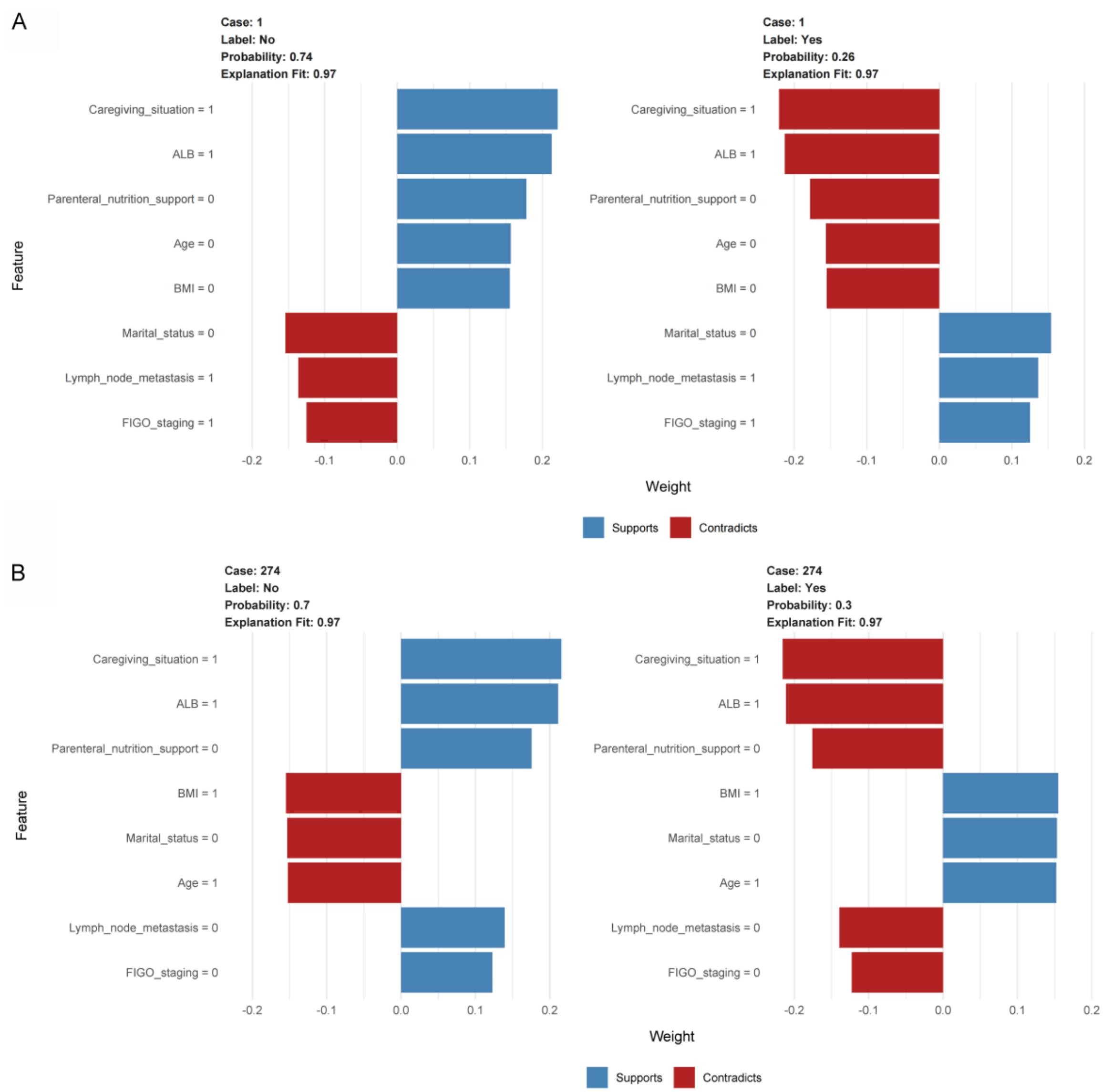

Figure 7. SHAP value decomposition analysis for representative samples. A: SHAP decomposition diagram for Sample 1 (normal nutrition case). B: SHAP decomposition diagram for Sample 274 (malnutrition case). Note: SHAP, SHapley Additive exPlanations; Blue bars (Supports) indicate features that decrease malnutrition risk; Red bars (Contradicts) indicate features that increase malnutrition risk; Weight represents the contribution of features to predictive outcomes; Label displays the actual label (0 for normal, 1 for malnutrition); Pred represents the model-predicted probability; Explanation R^2 indicates the degree of model interpretability for the sample prediction.

cal cancer undergoing surgical treatment. The model demonstrated stable performance in the training cohort, internal validation cohort, and external validation cohort, with good discrimination, reliable calibration, and meaningful clinical applicability.

Malnutrition affects many women with cervical cancer, and this condition often worsens as treatment progresses. In our cohort, the prevalence of malnutrition was 17.7% at baseline, rising to 47.1% during follow-up. This increase alone suggests the need for earlier and more aggressive screening. Previous studies have reported linking nutritional status to survival in this population; one analysis estimated the mortality hazard ratio for malnourished patients to be 3.12 [17]. While screening tools exist, their limitations have been exposed in routine practice. The NRS-2002 is widely used because it is easy to implement in various wards, but it still requires subjective judgment

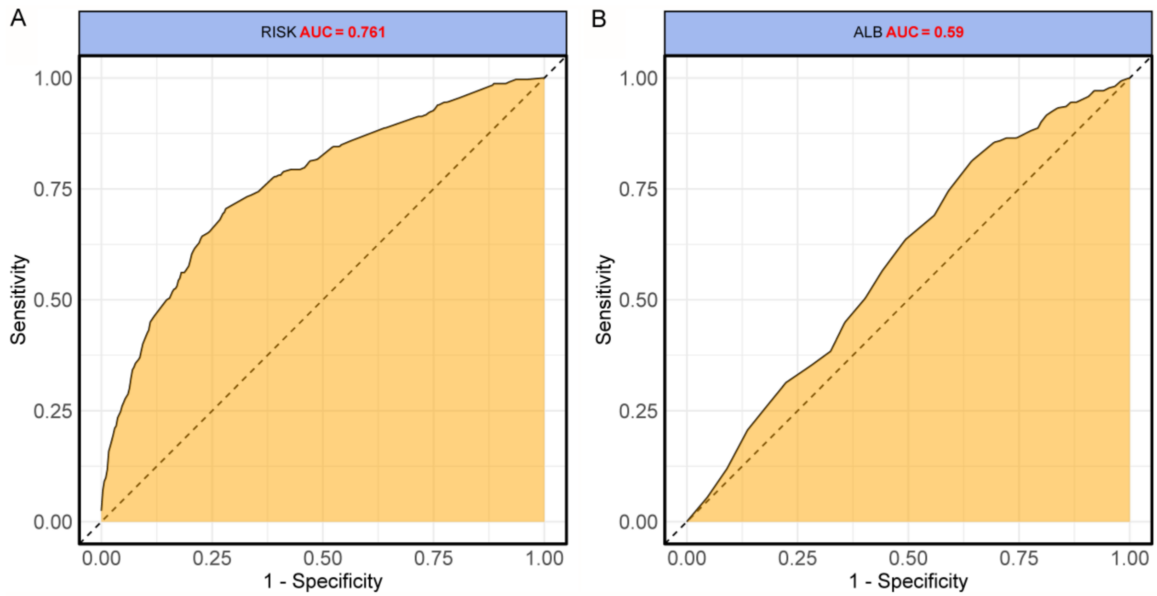

Figure 8. ROC curve comparison between RISK model and single indicator ALB in pooled samples. A: ROC curve for the RISK predictive model. B: ROC curve for the ALB single indicator. Note: ROC, Receiver Operating Characteristic; AUC, Area Under the Curve; RISK represents the multi-feature combined predictive model; ALB represents the preoperative albumin single indicator; The orange shaded area represents the area under the ROC curve; The dashed line represents the reference line (AUC=0.5).

Table 6. Performance metrics and DeLong test results for RISK model vs. single indicator ALB in pooled samples

Variable	95% CI	Specificity	Sensitivity	Youden Index	Accuracy	Precision	F1 Score
RISK	0.727-0.795	71.94%	70.65%	42.59%	71.43%	70.65%	66.16%
ALB	0.550-0.630	35.65%	81.29%	16.94%	46.30%	18.71%	21.60%

Note: AUC, Area Under the Curve; CI, Confidence Interval; ALB, Albumin; Youden Index = Sensitivity + Specificity - 1; F1 Score represents the harmonic mean of precision and recall.

Table 7. Univariable and multivariable Cox regression analysis of factors influencing OS in cervical cancer patients

Characteristics	Total (N)	Univariate Analysis		Multivariable Analysis	
		HR (95% CI)	*P* Value	HR (95% CI)	*P* Value
Age					
≥60 years	285				
<60 years	499	0.630 (0.382-1.040)	0.071		
BMI					
<18.5	125				
≥18.5	659	0.856 (0.452-1.621)	0.633		
Marital_status					
Married	606				
Other	178	1.927 (1.018-3.649)	0.044	2.21 (1.159-4.212)	0.016
Caregiving_situation					
Family caregiver	656				
Other caregiver	128	1.373 (0.707-2.669)	0.349		
Parenteral_nutrition_support					
Yes	188				
No	596	0.928 (0.547-1.573)	0.780		

Lymph_node_metastasis					
Yes	383				
No	401	4.647 (2.644-8.169)	<0.001	4.737 (2.689-8.346)	<0.001
Postoperative_lymphedema					
Yes	246				
No	538	1.501 (0.952-2.367)	0.081		
Postoperative_pain					
Yes	124				
No	660	0.784 (0.403-1.523)	0.473		
ECOG_PS					
0-1	575				
2	209	1.017 (0.611-1.693)	0.948		
FIGO_staging					
IIA	400				
IB	384	5.091 (2.804-9.242)	<0.001	5.15 (2.817-9.414)	<0.001
Pathological_type					
Squamous cell carcinoma	644				
Adenocarcinoma or other	140	0.978 (0.548-1.746)	0.941		
Differentiation_grade					
Poor	285				
Moderate/High	499	2.364 (1.507-3.706)	<0.001	2.123 (1.350-3.338)	0.001
Surgical_margin					
Positive	53				
Negative	731	0.761 (0.278-2.082)	0.595		
Deep_stromal_invasion					
≥1/2	510				
<1/2	274	1.624 (0.976-2.703)	0.062		
Lymphovascular_invasion					
Yes	234				
No	550	0.947 (0.577-1.552)	0.828		
Surgical_approach					
Laparoscopic	336				
Open	448	0.873 (0.552-1.38)	0.561		
HPV					
Positive	546				
Negative	238	0.476 (0.304-0.745)	0.001	0.631 (0.399-0.996)	0.048
Place_of_residence					
Rural	280				
Urban	504	1.293 (0.822-2.035)	0.266		
Education_level					
≥ High school	542				
< High school	242	1.160 (0.708-1.902)	0.556		
ALB					
≥40	190				
<40	594	1.960 (1.232-3.118)	0.005	1.540 (0.959-2.474)	0.074
Risk					
Low risk (<-0.581)	398				
High risk (≥-0.581)	386	1.718 (1.083-2.724)	0.021	1.883 (1.180-3.007)	0.008

Note: Risk was categorized by median value -0.581 into high-expression group (≥-0.581) and low-expression group (<-0.581). BMI, body mass index; ECOG-PS, Eastern Cooperative Oncology Group Performance Status; FIGO, International Federation of Gynecology and Obstetrics; HPV, human papillomavirus; ALB, albumin.

from clinicians and is time-consuming. Its component scores, such as nutritional status and disease severity items, are associated with length of hospital stay and mortality in general inpatient cohorts [18]. Even so, the performance of NRS-2002 varies when compared with other short assessment tools, such as subjective global assessment for hospitalized adults [19]. No single checklist can cover all at-risk patients, and its performance changes with context. In a cervical cancer surgery cohort, Tian et al. showed that patient-generated subjective global assessment (PG-SGA) outperformed NRS-2002: Youden index was 0.795 and 0.550, respectively [20]. The message is simple: using only one tool will miss some patients.

We constructed and tested a multivariate model to bridge this gap. The model achieved standard discrimination across all three datasets: AUC of 0.756, 0.742, and 0.807, respectively. The external dataset performed best, indicating that the model generalizes rather than simply memorizing the training data. The difference compared to albumin alone was significant. Albumin had an AUC of 0.590; our model outperformed it by 0.171, with a De Lon test *p*-value <0.001. This result is consistent with findings in oncology: models incorporating multiple signals outperform any single biomarker because they capture interactions and clinical context [21, 22]. What does this mean for practical applications? The risk of malnutrition increases during treatment. Routine screening misses a significant portion of patients, and a single laboratory test cannot compensate for this deficiency. A model that integrates common variables can identify risks early, allowing sufficient time to adjust treatment plans. This is precisely the goal of this study: earlier initiation of screening, easy replication of screening, and provision of usable results for healthcare teams.

The calibration results are reliable. The Brill score is 0.1859 on the training set, 0.2143 on the internal validation set, and 0.1788 on the external validation set. A score below 0.22 is considered acceptable, with scores below 0.19 on two datasets, indicating a high degree of consistency between predicted and observed risks. This is crucial in clinical practice, as clinicians require reliable data applicable to individual patients, not just average data. DCA showed net benefits across a threshold range of 0.04 to 0.99. Within this range, the model outperforms both the "all-treatment" and "no-treatment" strategies. At a very high cutoff (0.99), there are 62.27, 52.30, and 62.63 unnecessary interventions per 100 patients, respectively, on the training, internal, and external validation sets. In practice, this translates to clearer triage: focusing nutritional support on those most in need while avoiding additional interventions for low-risk patients.

Multivariate logistic regression yielded eight independent predictors. The risk was reduced for those under 60 years of age, with an odds ratio of 0.449. This is consistent with physiological principles, as older adults typically have lower reserves and impaired digestive function, which increases susceptibility to malnutrition; large gynecologic oncology studies have identified advanced age as an independent risk factor [4]. A BMI of 18.5 kg/m^2 or higher also had a protective effect, with an odds ratio of 0.452. Lower BMI means less energy and protein reserves when experiencing surgical stress and disease-related catabolism. Psychosocial factors were also important. Being unmarried increased the risk, with an odds ratio of 2.190, while not having a family caregiver further increased the risk, with an odds ratio of 2.986. Family caregivers tend to provide more attentive dietary assistance and more consistent encouragement, thus promoting eating; a lack of such support usually means poorer meal preparation, less motivation to eat, and greater stress, all of which negatively impact nutritional status [23]. Albumin was the strongest single predictor, with an odds ratio of 3.323. ALB levels below 40 g/L indicate low protein reserves and impaired hepatic synthesis. We still favor a multifactorial approach because albumin levels vary with inflammation, organ function, and humoral status, which are not purely nutritional signals. The advantages of this model over albumin alone support this choice. Disease burden is also associated with risk. Absence of lymph node involvement reduced the risk of malnutrition, with an odds ratio of 0.485, as did earlier FIGO IIA stage, with an odds ratio of 0.520. These findings are consistent with previous studies that have associated tumor burden and systemic inflammation with declining nutritional status; one study found that the AUC for lymph node metastasis reached 0.857 after incorporating nutritional variables, highlighting the bidirectional link between disease and nutrition [24]. Several studies have also shown that low albumin levels are associated with poorer outcomes and survival in patients with cervical cancer [25, 26]. The relationship between increased risk of malnutrition and PN support should be interpreted with caution. Although PN was found to be a positive predictor in regression and SHAP analyses, this appears to be a reverse causal relationship. Clinicians typically initiate PN because patients already have nutritional problems, gastrointestinal intolerances, or postoperative complications, even though PN itself does not cause malnutrition. In such cases, PN is viewed as a marker of potential clinical severity rather than a cause of malnutrition, a limitation that is now clearly recognized.

Cox regression correlated the RISK score with survival. The score remained an independent prognostic factor with a hazard ratio of 1.883 and a *P* value of 0.008. Each unit increase in the score was associated with an 88.3% increase in the risk of death, consistent with published associations between malnutrition and survival [19]. As expected of established prognostic biomarkers, FIGO stage had a hazard ratio of 5.150 and lymph node metastasis had a hazard ratio of 4.737. HPV positivity reduced the risk with a hazard ratio of 0.631 and a *p*-value of 0.048, consistent with earlier reports [3]. Poor tumor differentiation increased mortality (HR 2.123). Being unmarried also increased mortality (HR 2.210). Albumin showed some signal but did not reach significance (P=0.074); this suggests the need for larger prospective cohort studies. To make the model output readable and testable, we used SHAP to score the impact of each feature on risk. Albumin ranked first. Next came age and lymph node status, followed by PN support and FIGO stage. Marital status, caregiver type, and BMI followed closely. These charts clearly showed the direction of risk changes for specific patients: some values increased risk, while others decreased it. Clinicians could see the scores and their causes, which helped build trust and develop personalized nutrition plans at the bedside.

This study had real advantages. We included clinical, demographic, and psychosocial data to reflect the actual occurrence of malnutrition [27]. The sample size was large enough to support the analysis (784 patients). We also tested the model outside the derivation environment: 173 patients from independent institutions, which supported the model's applicability beyond a single center [28]. Few studies do this; many rely solely on internal checks.

Recent research has focused on predicting malnutrition in cancer populations. According to Yu et al. [16], a radiomics-based nomogram was developed for IB1-IIA2 stage cervical cancer, with an AUC of 0.972 on the training set and 0.805 on the validation set. Nevertheless, the CT-derived radiomics features and the lack of external validation limit its reproducible clinical application. In addition, the AUC of the machine learning model for elderly hospitalized cancer patients was 0.945 [13]. However, its important predictors, such as activities of daily living, are specific to elderly patients and therefore cannot be simply applied to gynecologic oncology. Our model uses only easily collected clinical and psychosocial features, has SHAP interpretability, and has been externally validated, making it an attractive and generalizable model for cervical cancer care.

Limitations remain. The retrospective design carries the risk of selection bias and information bias. We did not collect detailed dietary intake data, which could have increased the accuracy of risk estimates [29]. Our results apply to patients undergoing FIGO stage IB-IIA surgery; other groups require testing, including those receiving initial chemoradiotherapy and those with advanced disease [30]. Although there was a numerical difference in the incidence of malnutrition between the internal and external validation cohorts, this difference was not statistically significant (P=0.056), suggesting sampling variation rather than true heterogeneity, and this did not affect the model's stable external performance. Although the NRS-2002 is widely used, using it as a reference standard increases subjectivity. Further prospective and multicenter studies should be conducted to test its durability and clinical impact.

Conclusion

This nomogram incorporates eight clinically available predictors, such as age, sex, marital status, presence of caregivers, PN support, lymph node metastasis, FIGO stage, and albumin. Older age, lack of marital support, non-family care, use of PN, presence of lymph node metastasis, and higher FIGO stage increased the risk of malnutrition; while higher BMI and higher albumin levels had a protective effect. The model has good discriminative power, calibration, and clinical applicability, which supports its use in early postoperative identification of nutritional risks.

Disclosure of conflict of interest

None.

Address correspondence to: Beiyu Tuo, Department of Oncology, Yan'an Traditional Chinese Medicine Hospital, No. 26, Xuanyuan Avenue, New District, Yan'an 716000, Shaanxi, China. E-mail: tby1138437542@163.com

References

[1] Wei F, Georges D, Man I, Baussano I and Clifford GM. Causal attribution of human papillomavirus genotypes to invasive cervical cancer worldwide: a systematic analysis of the global literature. Lancet 2024; 404: 435-444.

[2] Lemp JM, De Neve JW, Bussmann H, Chen S, Manne-Goehler J, Theilmann M, Marcus ME, Ebert C, Probst C, Tsabedze-Sibanyoni L, Sturua L, Kibachio JM, Moghaddam SS, Martins JS, Houinato D, Houehanou C, Gurung MS, Gathecha G, Farzadfar F, Dryden-Peterson S, Davies JI, Atun R, Vollmer S, Bärnighausen T and Geldsetzer P. Lifetime prevalence of cervical cancer screening in 55 low- and middle-income countries. JAMA 2020; 324: 1532-1542.

[3] 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.

[4] Goins EC, Weber JM, Truong T, Moss HA, Previs RA, Davidson BA and Havrilesky LJ. Malnutrition as a risk factor for post-operative morbidity in gynecologic cancer: analysis using a national surgical outcomes database. Gynecol Oncol 2022; 165: 309-316.

[5] Batool R, Butt MS, Sultan MT, Saeed F and Naz R. Protein-energy malnutrition: a risk factor for various ailments. Crit Rev Food Sci Nutr 2015; 55: 242-253.

[6] Yuan J, Zhang Y, Wu S and Zheng L. The intestinal flora and nutritional status and immune function characteristics of obese colon cancer patients. BMC Gastroenterol 2024; 24: 237.

[7] Yang J, Dai E and Yin T. Effects of different nutritional support methods on nutritional status and immune function in patients undergoing radiotherapy for head and neck cancer. Clin Transl Oncol 2025; 27: 1310-1319.

[8] Seebacher V, Rockall A, Nobbenhuis M, Sohaib SA, Knogler T, Alvarez RM, Kolomainen D, Shepherd JH, Shaw C and Barton DP. The impact of nutritional risk factors and sarcopenia on survival in patients treated with pelvic exenteration for recurrent gynaecological malignancy: a retrospective cohort study. Arch Gynecol Obstet 2022; 305: 1343-1352.

[9] Tan F, Xia R, Zeng L, Xie H, Long X and Peng C. Value of preoperative controlling nutritional status score in prognosis of patients with high-risk factors for early-stage cervical cancer. Pak J Med Sci 2024; 40: 120-127.

[10] Chen X, Liu X, Ji W, Zhao Y, He Y, Liu Y, Li Q, Shi H and Cui J. The PG-SGA outperforms the NRS 2002 for nutritional risk screening in cancer patients: a retrospective study from China. Front Nutr 2023; 10: 1272420.

[11] Panchal M, Saseedharan S, Navade J and Gada M. Comparison of mNutric score, NRS-2002 score, and SGA score with the gold standard GLIM criteria in the diagnosis of malnutrition. Nutr Health 2025; 31: 1819-1824.

[12] Chen Y, Xiang Q, Li C, Zeng Y, Dong J, Zhang P, Li Y, Wang Y and Wang K. Nutritional risk and assessment for patients with cancer pain. Nutr Cancer 2022; 74: 168-174.

[13] Duan R, Li Q, Yuan QX, Hu J, Feng T and Ren T. Predictive model for assessing malnutrition in elderly hospitalized cancer patients: a machine learning approach. Geriatr Nurs 2024; 58: 388-398.

[14] Bernstein AM, Janeke P, Riggs RV, Burke E, Meyer J, Moyer MF, Murofushi K, Botha RA and Meyer JEM. Artificial intelligence-based hospital malnutrition screening: validation of a novel machine learning model. Appl Clin Inform 2025; 16: 1646-1657.

[15] Parchure P, Besculides M, Zhan S, Cheng FY, Timsina P, Cheertirala SN, Kersch I, Wilson S, Freeman R, Reich D, Mazumdar M and Kia A. Malnutrition risk assessment using a machine learning-based screening tool: a multicentre retrospective cohort. J Hum Nutr Diet 2024; 37: 622-632.

[16] Yu W, Xu H, Chen F, Shou H, Chen Y, Jia Y, Zhang H, Ding J, Xiong H, Wang Y and Song T. Development and validation of a radiomics-based nomogram for the prediction of postoperative malnutrition in stage IB1-IIA2 cervical carcinoma. Front Nutr 2023; 10: 1113588.

[17] Argefa TG and Roets L. Malnutrition and the survival of cervical cancer patients: a prospective cohort study using the PG-SGA tool. Nutr Cancer 2022; 74: 605-612.
[18] Sahli L, Hagenbuch N, Ballmer PE, Rühlin M and Imoberdorf R. NRS-2002 components, nutritional score and severity of disease score, and their association with hospital length of stay and mortality. Swiss Med Wkly 2021; 151: w20517.
[19] Kroc Ł, Fife E, Piechocka-Wochniak E, Sołtysik B and Kostka T. Comparison of nutrition risk screening 2002 and subjective global assessment form as short nutrition assessment tools in older hospitalized adults. Nutrients 2021; 13: 225.
[20] Tian M, Fu H and Du J. Application value of NRS2002 and PG-SGA in nutritional assessment for patients with cervical cancer surgery. Am J Transl Res 2021; 13: 7186-7192.
[21] Wang S, Zhao M, Gao Z, Yang X, Wang Y, Hua K and Fu J. A survival nomogram involving nutritional-inflammatory indicators for cervical cancer patients receiving adjuvant radiotherapy. J Cancer 2024; 15: 5773-5785.
[22] Medici F, Ferioli M, Zamfir AA, Buwenge M, Macchia G, Deodato F, Castellucci P, Tagliaferri L, Perrone AM, De Iaco P, Strigari L, Bazzocchi A, Rizzo SMR, Donati CM, Arcelli A, Fanti S, Morganti AG and Cilla S. Integrating novel and classical prognostic factors in locally advanced cervical cancer: a machine learning-based predictive model (ESTHER study). J Pers Med 2025; 15: 153.
[23] Ou M, Zhu L, Chen H, Wang G, Chen F and Xiao Z. Perioperative change trajectories and predictors of swallowing function and swallowing-related quality of life in patients with oral cancer: a longitudinal observational study. BMJ Open 2023; 13: e075401.
[24] Li N, Peng E and Liu F. Prediction of lymph node metastasis in cervical cancer patients using AdaBoost machine learning model: analysis of risk factors. Am J Cancer Res 2025; 15: 1158-1173.
[25] Wang X, Xu J, Zhang H and Qu P. The effect of albumin and hemoglobin levels on the prognosis of early-stage cervical cancer: a prospective, single-center-based cohort study. BMC Womens Health 2023; 23: 553.
[26] Yoshikawa N, Yoshihara M, Tamauchi S, Ikeda Y, Yokoi A and Kajiyama H. Hypoalbuminemia for the prediction of survival in patients with stage IVB cervical cancer. PLoS One 2022; 17: e0273876.
[27] Zhang G, Zhang Y, He F, Wu H, Wang C and Fu C. Preoperative controlling nutritional status (CONUT) score is a prognostic factor for early-stage cervical cancer patients with high-risk factors. Gynecol Oncol 2021; 162: 763-769.
[28] Li X, Deng L, Xie H, Li S, Zhao H, Liu T, Liu X, Lin S, Liu C and Shi HP. NCR as a biomarker for nutritional status and inflammation in predicting outcomes in patients with cancer cachexia: a prospective, multicenter study. BMC Cancer 2025; 25: 539.
[29] Mota AP, Aredes MA, De Oliveira Miguel J and Chaves GV. Nutritional status assessed by patient-generated subjective global assessment is associated with toxicity to chemoradiotherapy in women with cervical cancer: a prospective study. Eur J Clin Nutr 2022; 76: 1740-1747.
[30] Jou J, Coulter E, Roberts T, Binder P, Saenz C, McHale M, Plaxe S, Mayadev J and Eskander RN. Assessment of malnutrition by unintentional weight loss and its implications on oncologic outcomes in patient with locally advanced cervical cancer receiving primary chemoradiation. Gynecol Oncol 2021; 160: 721-728.