scientific reports

Check for updates

OPEN

Prediction of clinical stages of cervical cancer via machine learning integrated with clinical features and ultrasound-based radiomics

Maochun Zhang[1,3,6], Qing Zhang[4,6], Xueying Wang[4,6], Xiaoli Peng[4], Jiao Chen[5] & Hanfeng Yang[1,2]✉

To investigate the prediction of a model constructed by combining machine learning (ML) with clinical features and ultrasound radiomics in the clinical staging of cervical cancer. General clinical and ultrasound data of 227 patients with cervical cancer who received transvaginal ultrasonography were retrospectively analyzed. The region of interest (ROI) radiomics profiles of the original image and derived image were retrieved and profile screening was performed. The chosen profiles were employed in radiomics model and Radscore formula construction. Prediction models were developed utilizing several ML algorithms by Python based on an integrated dataset of clinical features and ultrasound radiomics. Model performances were evaluated via AUC. Plot calibration curves and clinical decision curves were used to assess model efficacy. The model developed by support vector machine (SVM) emerged as the superior model. Integrating clinical characteristics with ultrasound radiomics, it showed notable performance metrics in both the training and validation datasets. Specifically, in the training set, the model obtained an AUC of 0.88 (95% Confidence Interval (CI): 0.83–0.93), alongside a 0.84 accuracy, 0.68 sensitivity, and 0.91 specificity. When validated, the model maintained an AUC of 0.77 (95% CI: 0.63–0.88), with 0.77 accuracy, 0.62 sensitivity, and 0.83 specificity. The calibration curve aligned closely with the perfect calibration line. Additionally, based on the clinical decision curve analysis, the model offers clinical utility over wide-ranging threshold possibilities. The clinical- and radiomics-based SVM model provides a noninvasive tool for predicting cervical cancer stage, integrating ultrasound radiomics and key clinical factors (age, abortion history) to improve risk stratification. This approach could guide personalized treatment (surgery vs. chemoradiation) and optimize staging accuracy, particularly in resource-limited settings where advanced imaging is scarce.

Keywords Cervical cancer, Machine learning, Ultrasound radiomics, Clinical features, Clinical stage

In 2020, cervical cancer is ranked as the second most prevalent cancer among women worldwide, with a mortality rate of 60%[1]. In developed countries and regions, the widespread implementation of cytological screening programs results in a decline in the incidence and mortality rates of cervical cancer[2]. However, in underdeveloped areas, low vaccination rates for human papillomavirus (HPV) and screening coverage are observed due to economic underdevelopment[2]. Recently, the onset age for cervical cancer has trends younger, with some cases diagnosed at advanced stages, which leads to poor prognoses[3]. The clinical stage greatly impacts the treatment options and prognosis for cervical cancer patients. The clinical outcomes for those with advanced stages differ vastly from those with early stages. Based on the guidelines recommended by the International Federation of Gynecology and Obstetrics (FIGO), patients with early-stage cervical cancer (stages I to II A) are

[1]Affiliated Hospital, Jinan University, Guangzhou 510630, China. [2]Department of Radiology, Affiliated Hospital of North Sichuan Medical College, Nanchong 637000, China. [3]Department of Health Management Center, Affiliated Hospital of North Sichuan Medical College, Nanchong 637000, China. [4]Department of Ultrasound, Affiliated Hospital of North Sichuan Medical College, Nanchong 637000, China. [5]Department of Obstetrics and Gynecology, Affiliated Hospital of North Sichuan Medical College, Nanchong 637000, China. [6]Maochun Zhang, Qing Zhang and Xueying Wang contributed equally to this work. ✉email: yhf8929@163.com

often managed with surgical intervention, whereas those with advanced stage (stages II B to IV), concurrent radiation therapy combined with chemotherapy are recommended[4]. Therefore, the early detection and accurate staging of cervical cancer are of significant importance for guiding treatment choices and prognosis assessment. Recent advances in quantitative imaging analysis have positioned radiomics as a transformative tool in oncology, enabling non-invasive characterization of tumor heterogeneity and improved prognostic stratification across cancer types[5]. While initially applied to CT and MRI, radiomics is now being adapted for ultrasound, offering a cost-effective alternative particularly valuable in resource-limited settings[6].

Ultrasound examination is radiation-free, simple, and cost-effective, allowing for multi-angle and multi-plane imaging, which facilitates the detection, measurement, and localization of lesions. It also helps identify lymph node metastases in the pelvic and groin regions. Radiomics provides a means to quantitatively extract high-throughput features from images, enabling a more detailed characterization of tumor heterogeneity with enhanced spatial exploration[7]. Multi-modal approaches integrating clinical data with radiomics features have demonstrated superior performance in cancer staging compared to imaging-only models, as evidenced by recent studies combining MRI radiomics with clinicopathological variables[8]. This synergy justifies our integration of demographic factors (age, abortion history) with ultrasound-derived features. Furthermore, advancements in ultrasound radiomics feature extraction protocols[9] have addressed historical concerns about reproducibility, establishing its reliability for quantitative analysis.

Machine learning (ML) is an innovative research tool that employs computers to construct probabilistic statistical models based on data, which are then used to analyze data. ML has emerged as a cornerstone of radiomics analysis, with algorithms like SVM successfully applied to diverse medical imaging tasks—from Alzheimer's disease prediction using PET radiomics[10] to tumor staging in oncology. These studies underscore ML's capacity to uncover complex patterns in high-dimensional data, supporting our multi-model comparison approach. Currently, the integration of radiomics and ML is extensively applied in the medical field, which proves more reliable than traditional methods such as logistic regression models. ML models based on radiomic features can provide valuable information for devising superior treatment plans in clinics[11]. ML algorithms are particularly suited for cancer staging tasks due to their ability to model complex, non-linear relationships between imaging features, clinical variables, and disease outcomes. Traditional statistical methods often assume linear relationships, which may oversimplify the biological complexity of tumor heterogeneity and progression[12]. Advanced ML techniques like SVM can capture these intricate patterns through kernel-based transformations, making them ideal for integrating multi-dimensional data from radiomics and clinical features[13]. However, few studies use ML models to guide treatment for patients with cervical cancer. Therefore, we aimed to explore the prediction of a model constructed by ML based on clinical features and ultrasound radiomics in the clinical staging of cervical cancer. We used various ML methodologies to develop and validate models by integrating ultrasound radiomics features with the clinical characteristics of patients with cervical cancer.

Subjects and methods

Subjects

This retrospective study was approved by the Ethics Committee of the Affiliated Hospital of North Sichuan Medical College (Approval number: 2022ER451-1), and the informed consent requirement was negated owing to the retrospective design of the investigation. All investigational protocols abided by the institutional and national medical ethics committees-recommended guidelines, as well as to the 1964 Declaration of Helsinki and comparable ethical values. A total of 430 patients with cervical squamous cell carcinoma (CSCC) who were admitted to the Affiliated Hospital of North Sichuan Medical College from December 2020 to November 2022 were included in this study. Inclusion criteria: (1) cervical cancer was confirmed by postoperative pathology; (2) transvaginal ultrasonography was performed within one month before surgery and the maximum sectional view of lesions was saved; (3) patients did not undergo relevant intervention before ultrasound evaluation and before treatment. Exclusion criteria: (1) poor image quality, unable to identify lesions; (2) cervical tumors measuring < 1.0 cm in maximal diameter on ultrasound, as lesions below this size threshold (a) may not provide sufficient voxel information for reliable radiomics analysis, (b) are challenging to delineate reproducibly given ultrasound's resolution limits, and (c) represent early-stage cases where staging accuracy is less clinically consequential; (3) patients with other malignancies. Finally, 227 CSCC patients were recruited for analysis. Based on the FIGO 2018 guidelines, stage I-IIA was classified as the early stage, and stage IIB-IV was classified as the advanced stage. There were 159 cases in the early-stage group and 68 cases in the advanced-stage group.

Data collection

The ultrasound images and general clinical data of all patients were collected, including age, number of vaginal births, number of cesarean sections, and number of abortions.

Radiomics feature analysis

Two-dimensional (2D) ultrasound images of 227 patients were selected according to the ataxation criteria. The 2D ultrasound images of the largest slice of the lesion in the vaginal ultrasound one month before surgery were selected. The lesion was confirmed by postoperative pathology. The images were stored in Digital Imaging and Communications in Medicine (DICOM) format within Picture Archiving and Communication Systems (PACS). Original two-dimensional ultrasound images were transferred to the three-dimensional (3D)-Slicer software (https://www.slicer.org). Two highly experienced physicians in the ultrasonic diagnosis of obstetrics and gynecology (Physician A and B, with 15 and 12 years of experience respectively) manually outlined the region of interest (ROI)[14] (Supplementary Fig. 1). The radiomics component in 3D-Slicer software was employed to retrieve ROI radiomics profiles. Finally, a total of 837 radiomics profiles were retrieved and normalized.

To assess the reproducibility of ROI delineation and feature extraction, 30 cases (13.2%) were randomly selected after ten days for re-evaluation. The same two physicians independently re-delineated the ROIs using the same protocol without reference to their initial annotations. This process served two purposes: (1) to evaluate intra-observer variability by comparing each physician's repeat measurements with their initial delineations, and (2) to assess inter-observer consistency by comparing the two physicians' independent annotations. The random selection ensured unbiased representation across all clinical stages.

Prediction model construction

A total of 227 patients were arbitrarily separated into a training set ($n=158$) and a validation set ($n=69$) in a 7:3 ratio using R software (Version 4.2.3, https://www.R-project.org) and R studio software. R software was also employed for redundancy analysis. The Pearson test was conducted for normally distributed data and the Spearman test for nonnormal data. A correlation coefficient threshold of 0.9 was selected to remove redundant characteristic parameters, as this represents the commonly accepted threshold for identifying highly correlated features in radiomics studies[15]. This conservative threshold was chosen to: (1) minimize information loss while reducing multicollinearity, (2) maintain sufficient features for robust model development, and (3) align with previous radiomics research standards[16]. Features with correlation ≥ 0.9 were considered to convey substantially overlapping information, where retaining both could bias machine learning algorithms without adding meaningful independent information.

After the redundancy analysis, the least absolute shrinkage and selection operator (LASSO) method was adopted for feature or profile screening. Tenfold cross-validation was used to minimize overfitting[17]. A Radscore formula was constructed according to the obtained features. The radiomics model was generated with the help of the RadScore. In the training set, several models were established utilizing Logistic Regression (LR)[18], Random Forest (RF)[19], Support Vector Machines (SVM)[20], Gaussian Naive Bayes (GNB)[21], and Extreme Gradient Boosting (XGBoost)[22] as previously described. Further verification was conducted in the validation set. The optimal model identified from the aforementioned methods was then selected to generate a model integrating both clinical profiles and the RadScore. For comparative purposes, a separate clinical model encompassing only clinical features was also established. All model development and evaluation were implemented using Python (Version: 3.11.4, https://www.python.org/downloads/).

Statistical analysis

The interclass correlation coefficients (ICC) were employed to assess profiles retrieved by the two physicians and doctors with the same names. ICC > 0.75 indicated good consistency, and $P<0.05$ indicated the statistical significance threshold. All general clinical demographic information were numerical variables. The normally distributed data were presented as means ± standard deviation (SD) and compared via *t*-test according to whether the variance was homogeneous. Non-normally distributed data were described by median (Min, P25, P75, Max) and compared by the Wilcoxon Mann-Whitney test. The same test was also employed for the early-stage Radscore difference and the advanced-stage comparison between the training and validation sets. The performance of the five models was compared by plotting the receiver operating characteristic (ROC) curve and calculating the area under the curve (AUC), accuracy, sensitivity, and specificity, thereby identifying the optimal radiomics model. A calibration curve was employed to assess the model calibration efficiency, and Decision Curve Analysis (DCA) was employed for the model's clinical application value prediction.

Results

General clinical data

Based on the FIGO staging system and pathological findings, there were 13 cases of stage IA, 90 cases of stage IB, 56 cases of stage IIA, 20 cases of stage IIB, 44 cases of stage III, and 4 cases of stage IV. Stages I to IIA were included in the early ($n=159$) and stages IIB to IV in the advanced stage ($n=68$). We observed substantial differences in patient age, number of vaginal births, and number of abortions between the early stage (I-IIA) and advanced stage (IIB-IV) groups ($P<0.05$), but no statistically significant differences in the number of cesarean section ($P>0.05$, Table 1).

Radiomics profile and model evaluation

The ICC values of the two retrieved radiomics profiles by the same and those extracted by two physicians were all more than 0.9, with satisfactory consistency (all $P<0.05$). We retrieved 837 profiles. After normalization processing, redundancy analysis, tenfold cross-validation and LASSO dimension reduction screening (Supplementary Fig. 2), 18 radiomics features were finally obtained, including 1 first-order feature and 17 high-

General clinical data	Early stage ($n=159$)	Advanced stage ($n=68$)	*P* (Shapiro–wilk test)	t/*P* (t-test)	Z/*P* (WM-W test)	*P* (logistic)
Age (years)	53.7 ± 9.8	57.7 ± 10.8	0.107	− 2.741/0.007		0.034
Number of natural births	2(0,1,2,7)	2(0,2,3,6)	< 0.001		− 3.270/0.001	0.546
Number of cesarean sections	0(0,0,0,2)	0(0,0,0,2)	< 0.001		− 1.402/0.161	
Number of abortions	2(0,1,3,7)	1(0,0,2,6)	< 0.001		− 2.528/0.011	0.003
Radscore						< 0.001

Table 1. General clinical data.

order features (Supplementary Table 1). The obtained feature and corresponding non-zero coefficient were used to construct a formula to obtain the Radscore:

Radscore = − 2.142.
+ Original_firstorder_Minimum*(− 0.410).
+ wavelet.LLH_firstorder_Minimum*(− 0.059).
+ wavelet.LLH_glcm_MCC*(− 1.960).
+ wavelet.LLH_glrlm_ShortRunHighGrayLevelEmphasis*1.843.
+ wavelet.LLH_glszm_ZoneEntropy*(− 0.366).
+ wavelet.LHL_glszm_SmallAreaEmphasis*(− 0.112).
+ wavelet.LHL_glszm_SmallAreaLowGrayLevelEmphasis*(− 0.609).
+ wavelet.LHL_glszm_ZoneEntropy*0.427.
+ wavelet.LHH_gldm_DependenceNonUniformityNormalized*0.935.
+ wavelet.LHH_gldm_GrayLevelVariance*0.263.
+ wavelet.LHH_glszm_LowGrayLevelZoneEmphasis*(− 0.113).
+ wavelet.LHH_glszm_SmallAreaEmphasis*1.559.
+ wavelet.HLL_glcm_Imc2*(− 1.173).
+ wavelet.HHH_firstorder_Mean*(− 0.184).
+ wavelet.HHH_glszm_SmallAreaHighGrayLevelEmphasis *0.571.
+ wavelet.LLL_firstorder_Kurtosis*(− 0.621).
+ wavelet.LLL_glcm_InverseVariance*(− 1.769).
+ wavelet.LLL_glrlm_RunEntropy*0.261.

After integrating general clinical data with statistical significance and Radscore for multivariate logistic regression, there was no discernible difference in the number of vaginal births ($P > 0.05$), and it was removed. Finally, age, number of abortions, and Radscore were used to construct combined models. Figure 1A illustrated the ROC curves of the five radiomics models on the training set, while Fig. 1B presented the ROC curves of the five radiomics models on the validation set. Comparisons of the performances of the five models were presented in Table 2. Although the XGBoost model exhibited good predictive power on the training set, it performed inadequately on the validation set, indicating a tendency towards overfitting. The RF model demonstrated a strong AUC value, however, it was characterized by low sensitivity on the validation set. Similarly, both the LR and GNB models showed low sensitivity in both training and validation sets. Consequently, to confirm the stability and consistency of the five models, the SVM model was ultimately selected as the optimal model. Furthermore, a heatmap illustrated a strong correlation between the Radscore and the two clinical variables (age and the number of abortions) (Supplementary Fig. 3). Consequently, an integrated model was developed using SVM, which combined age, the number of abortions, and the RadScore.

Performance of integrated model

Figure 2A displayed the ROC curves of integrated model for both the training and validation sets. Figure 2B presented the calibration plots of the integrated model for both the training and validation sets, demonstrating the consistency between the predicted probability and actual outcomes. Meanwhile, Fig. 2C showed the clinical decision curve of the integrated model, illustrating its usage in guiding therapeutic decisions based on predicted outcomes. Additionally, a baseline clinical model comprising solely these three fundamental clinical variables was constructed for comparison. The outcomes of model evaluations were summarized in Table 3.

Discussion

In this study, we established five models by using ML with clinical features and ultrasound-based radiomics to predict the clinical outcomes of patients with cervical cancer. We found that the SVM model combining age and the number of abortions showed good performance. In the training set, this model had an AUC of 0.88, an accuracy of 0.83, a sensitivity of 0.68, and a specificity of 0.89. In the validation set, this model had an AUC of

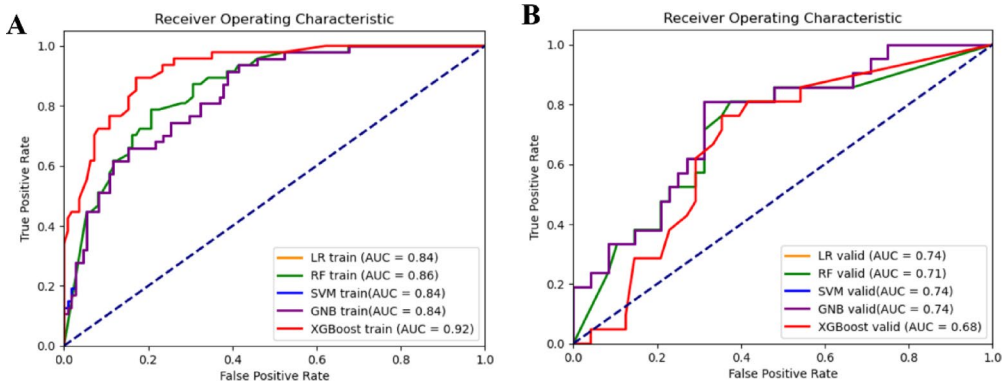

Fig. 1. The receiver operating characteristic (ROC) curves of the five models. (**A**) ROC curves for the training set; (**B**) ROC curves for the validation set.

Items		AUC (95%CI)	Accuracy	Sensitivity	Specificity
Training set	LR	0.84(0.77–0.90)	0.79	0.51	0.91
	RF	0.86(0.80–0.91)	0.80	0.62	0.88
	SVM	0.84(0.77–0.90)	0.71	0.79	0.68
	GNB	0.84(0.77–0.90)	0.80	0.57	0.89
	XGBoost	0.92(0.88–0.96)	0.84	0.64	0.93
Validation set	LR	0.74(0.61–0.86)	0.71	0.38	0.85
	RF	0.71(0.56–0.83)	0.70	0.48	0.79
	SVM	0.74(0.61–0.86)	0.70	0.71	0.69
	GNB	0.74(0.61–0.86)	0.67	0.38	0.79
	XGBoost	0.68(0.54–0.81)	0.67	0.29	0.83

Table 2. Diagnostic performance of the five models in predicting the stage of cervical cancer in the training and validation sets. CI represents confdence interval. AUC represents area under the receiver operating characteristic curve. LR represents logistic regression model. RF represents random forest model. SVM represents support vector machine model. GNB represents Gaussian Naive Bayes model. XGBoost represents extreme gradient boosting model.

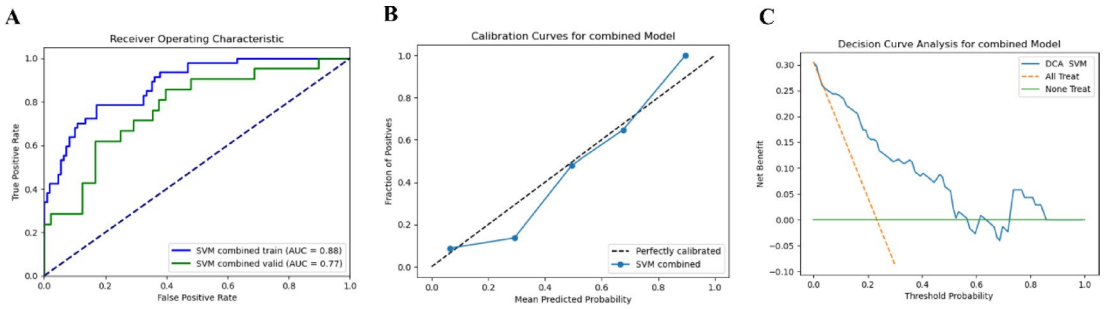

Fig. 2. The performance of the integrated model. (**A**) the ROC curves of an integrated model for both the training and validation sets. (**B**) the calibration plots of the integrated model for both the training and validation sets. (**C**) the clinical decision curve of an integrated model.

Items		AUC (95%CI)	Accuracy	Sensitivity	Specificity
Training set	RadScore	0.84(0.77–0.90)	0.71	0.79	0.68
	Clinical	0.65(0.56–0.74)	0.59	0.55	0.60
	Combined	0.88(0.82–0.93)	0.83	0.68	0.89
Validation set	RadScore	0.74(0.61–0.86)	0.70	0.71	0.69
	Clinical	0.60(0.46–0.75)	0.55	0.62	0.52
	Combined	0.77(0.64–0.89)	0.75	0.62	0.81

Table 3. Diagnostic performance of each SVM model in predicting the stage of cervical squamous cell carcinoma in the training and validation sets.

0.77, an accuracy of 0.75, a sensitivity of 0.62, and a specificity of 0.81. Our results showed that the SVM model can guide the therapeutic decisions of patients with cervical cancer based on predicted outcomes.

The FIGO staging system for cervical cancer dominates in clinical practice as well as in cancer database reporting[23]. Many scholars have conducted in-depth research involving the radiomics significance in predicting cervical cancer. Wang et al. revealed that cervical adenocarcinoma exhibits higher textural heterogeneity compared to CSCC (14). A combined model constructed using multi-sequence MRI images yielded an AUC of 0.89 in differentiating CSCC from cervical adenocarcinoma, which is higher than that of individual sequence models[24]. Shi et al. showed that the combined model of T2WI and voxel incoherent motion diffusion-weighted imaging had superior performance in evaluating tissue differentiation compared to individual sequences[25]. There has been extensive prior research on the application of radiomics in predicting pathological staging, lymph node and vascular space involvement, treatment response, and prognosis assessment of cervical cancer[26–28]. This presents new approaches for diagnosing and evaluating the treatment responses in cervical cancer. However, these studies have almost exclusively relied on magnetic resonance imaging (MRI), Computerized tomography (CT), or positron emission tomography (PET)-CT images. As far as we know, there have been scant investigations

on radiomics based on ultrasound images for the clinical staging of cervical cancer. In addition, previous studies often employed a single ML modeling approach, whereas our research has established five different ML models.

ML offers more accurate, objective, and reliable models that help in clinical decision-making[29]. Zhang et al. discovered that the RF model could accurately predict early mortality in metastatic colorectal cancer, outperforming models based on LR, CatBoost, XGBoost, and LightGBM, and the RF model demonstrated greater clinical utility[30]. Radiomics models are essential in disease diagnosis and intervention planning strategies, with their predictive accuracy being of paramount importance, which largely hinges on the appropriateness of the selected ML algorithms[31–33].

In this study, among the five ML models established by Radscore, the SVM model performed the best, with an AUC (95% Confidence Interval), accuracy, sensitivity, and specificity of 0.84 (0.77–0.90), 0.71, 0.79, and 0.68 in the training set, and 0.74 (0.61–0.86), 0.70, 0.71, and 0.69 in the validation set. Furthermore, this study also found that combining clinical parameters (age and number of abortions) with radiomics features in a combined model substantially increased the model's diagnostic performance. The combined model showed an AUC (95% CI), accuracy, sensitivity, and specificity of 0.88 (0.82–0.93), 0.83, 0.68, and 0.89 in the training set, and 0.77 (0.64–0.89), 0.75, 0.62, and 0.81 in the validation set. The performance of the integrated model surpassed that of both standalone clinical models and standalone radiomics models. The probable reason may be that standalone radiomics models fail to adequately account for individual patient variations and the impact of other clinical factors on tumor progression. Consequently, it can furnish more comprehensive and precise diagnostic information by integrating clinical characteristics with tumor radiomics models. This combination approach acknowledges the complexity of cancer diagnosis, where both biological characteristics visible through imaging and patient-specific clinical data contribute to a more holistic understanding of the disease state.

SVM was recognized as an effective and robust method for addressing nonlinear problems. It was designed to capture the complex and nonlinear relationships between radiative phenotypic and genotypic information while ensuring stability and efficiency[34]. RF is an algorithm based on an ensemble of multiple decision trees, which often achieves high accuracy across many datasets[35]. GNB is a probabilistic classification algorithm that can enhance specificity[36]. The XGBoost algorithm excels in both computational speed and model precision[37]. Garavand et al. established a diagnostic model for coronary artery disease using SVM, with performance metrics including Precision, Recall, F-measure, Matthews Correlation Coefficient (MCC), AUC, and Precision-Recall Curve (PRC) at 0.88, 0.88, 0.88, 0.70, 0.88, and 0.83, respectively. This model outperforms other methods such as Multilayer Perceptrons (MLP), LR, J48 (a decision tree algorithm), k-nearest Neighbors (KNN), and Naive Bayes (NB)[38]. This study indicated that the SVM model exhibited superior diagnostic performance, suggesting that cervical cancer may be nonlinear or not linearly separable, and thus is more suitably addressed by employing the SVM algorithm. The superior performance of SVM in our study likely stems from its ability to handle non-linear decision boundaries that characterize the relationship between ultrasound radiomics features and cervical cancer stage. Unlike linear models that assume simple additive effects, SVM's kernel trick allows it to project features into higher-dimensional spaces where complex patterns become separable[39]. This capability is particularly valuable in cancer staging where: (1) tumor biology follows non-linear progression patterns, (2) imaging features interact in complex ways, and (3) clinical variables may have threshold effects rather than linear relationships with outcomes[40]. Our findings align with recent studies demonstrating SVM's effectiveness in modeling similar complex medical decision boundaries[41].

The model's performance characteristics - with higher specificity (0.81–0.91) than sensitivity (0.62–0.68) - carry distinct clinical implications for early versus advanced stage management. For early-stage patients (I-IIA), the high specificity is particularly valuable as it minimizes false positives that might inappropriately escalate treatment from surgery to chemoradiation. This is crucial for preserving fertility options in younger patients where unnecessary chemoradiation could be devastating. However, the moderate sensitivity requires caution in advanced-stage cases (IIB-IV), as approximately 32–38% of truly advanced cases might be missed. In practice, this suggests the model should: (1) strongly support treatment intensification when positive (given high specificity), but (2) not rule out advanced disease when negative, particularly in high-risk patients (older age, multiple abortions) who may benefit from supplemental MRI even with negative predictions. This performance profile makes the model particularly suitable as a 'rule-in' tool for advanced disease, while maintaining conventional staging for uncertain cases - a balanced approach that could optimize resource utilization without compromising patient outcomes. The trade-off mirrors clinical decision thresholds where overtreatment of early-stage disease is often considered more consequential than under-treatment of advanced cases, given the irreversible consequences of unnecessary chemoradiation.

However, this study still has certain limitations. First, it is a single-center retrospective study without external verification, and the sample size of advanced-stage cervical cancer cases is relatively small, which may lead to overfitting. Secondly, the manual contouring of ROIs in this study is laborious and labor-intensive, with a certain degree of subjectivity and error in determining the lesion margins. Third, while our model incorporated key clinical variables (age and abortion history), additional clinically relevant features such as body mass index (BMI), HPV genotype, and genetic markers could potentially improve predictive performance. Recent studies in breast cancer have demonstrated the value of combining radiomics with comprehensive clinical profiles. Future multicenter studies should consider collecting these variables to build more robust models. ourth, despite our rigorous protocol employing dual-physician delineation and achieving excellent inter-observer concordance (ICC > 0.9), manual ROI definition remains susceptible to variability, particularly for heterogeneous or poorly-marginated tumors. Emerging strategies to address this limitation include: (1) AI-assisted segmentation using deep learning models pre-trained on similar datasets to generate initial contours for clinician refinement, (2) standardized consensus protocols for ultrasound tumor boundary definition (e.g., systematic inclusion/exclusion criteria for necrotic regions), and (3) multi-modal image fusion techniques that co-register ultrasound with MRI/CT to improve anatomical reference. Additionally, this study only explores the value of ultrasound

radiomics based on 2D ultrasound images, failing to incorporate color Doppler ultrasound and ultrasound contrast agents into the research. Lastly, the clinical parameters included in the study were limited to age and number of abortions, which may contribute to the inferior performance of the clinical model. Therefore, future studies should be performed by increasing sample sizes, and incorporating more clinical variables into the models. Moreover, multi-center and multi-modal studies are needed to further validate our results.

Conclusion

This study leverages ML to construct and validate an integrated model involving clinical profiles and ultrasound radiomics, which effectively predicts the clinical stage of cervical cancer. The integrated model employing the SVM algorithm demonstrates high predictive value in assessing the clinical stages of cervical cancer, thereby holding significant promise for enhancing clinical diagnosis.

Data availability

The simulation experiment data used to support the findings of this study are available from the corresponding author upon request.

Received: 25 October 2024; Accepted: 19 May 2025
Published online: 29 May 2025

References

1. Sung, H. et al. Global Cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *Cancer J. Clin.* **71**(3), 209–249 (2021).
2. Bray, F. et al. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *Ca-a Cancer J. Clin.* **68**, 394–424 (2018).
3. Buskwofie, A., David-West, G. & Clare, C. A. A review of cervical cancer: Incidence and disparities. *J. Natl Med. Assoc.* **112**, 229–232 (2020).
4. Bhatla, N., Aoki, D., Sharma, D. N. & Sankaranarayanan, R. Cancer of the cervix Uteri: 2021 update. *Int. J. Gynecol. Obstet.* **155**(Suppl 1), 28–44 (2021).
5. Ai, Y., Zhu, H., Xie, C. & Jin, X. Radiomics in cervical cancer: Current applications and future potential. *Crit. Rev. Oncol. Hematol.* **152**, 102985. https://doi.org/10.1016/j.critrevonc.2020.102985 (2020).
6. Cui, L. et al. Ultrasound radiomics in oncology: Current advances and future trends. *J. Imaging* **7**(2), 59. https://doi.org/10.3390/jimaging11020059 (2021).
7. Fiset, S. et al. Repeatability and reproducibility of MRI-based radiomic features in cervical cancer. *Radiother. Oncol.* **135**, 107–114 (2019).
8. Halle, M. K. et al. Radiomic profiles improve prognostication and reveal targets for therapy in cervical cancer. *Sci. Rep.* **14**(1), 11339. https://doi.org/10.1038/s41598-024-61271-4 (2024).
9. Ali, M. et al. Applications of artificial intelligence, deep learning, and machine learning to support the analysis of microscopic images of cells and tissues. *J. Imaging* **11**(2), 59. https://doi.org/10.3390/jimaging11020059 (2025). Published 2025 Feb 15.
10. Conti, A. et al. Radiomics analysis on SPM-based segmentation for Alzheimer's disease prediction: A machine-learning approach. *Diagnostics* **12**(4), 933. https://doi.org/10.3390/diagnostics12040933 (2022).
11. Zheng, Y., Zhou, D., Liu, H. & Wen, M. CT-based radiomics analysis of different machine learning models for differentiating benign and malignant parotid tumors. *Eur. Radiol.* **32**(10), 6953–6964 (2022).
12. Kumar, V. et al. Radiomics: The process and the challenges. *Magn. Reson. Imaging* **30**(9), 1234–1248. https://doi.org/10.1016/j.mri.2012.06.010 (2012).
13. Parmar, C. et al. Machine learning methods for quantitative radiomic biomarkers. *Sci. Rep.* **5**, 13087. https://doi.org/10.1038/srep13087 (2015).
14. Bove, S. et al. A ultrasound-based radiomic approach to predict the nodal status in clinically negative breast cancer patients. *Sci. Rep.* **12**(1), 7914. https://doi.org/10.1038/s41598-022-11876-4 (2022).
15. Park, J. E. et al. Quality of science and reporting of radiomics in oncologic studies: Room for improvement according to radiomics quality score and TRIPOD statement. *Eur. Radiol.* **30**, 523–536 (2020).
16. Zwanenburg, A. et al. The image biomarker standardization initiative: standardized quantitative radiomics for high-throughput image-based phenotyping. *Radiology* **295**, 328–338 (2020).
17. Manganaro, L. et al. Radiomics in cervical and endometrial cancer. *Br. J. Radiol.* **94**, 20201314 (2021).
18. Brusco, M. J., Steinley, D. & Watts, A. L. A comparison of logistic regression methods for ising model estimation. *Behav. Res. Methods* **55**(7), 3566–3584 (2023).
19. Feng, H. et al. A random forest model for peptide classification based on virtual docking data. *Int. J. Mol. Sci.* **24**(14), 11409 (2023).
20. Wu, F., Zhang, X., Fang, Z. & Yu, X. Support vector machine-based global classification model of the toxicity of organic compounds to Vibrio fischeri. *Molecules (Basel Switzerland)* **28**(6), 2703 (2023).
21. Zhang, J., Hao, L., Xu, Q. & Gao, F. Radiomics and clinical characters based Gaussian Naive Bayes (GNB) model for preoperative differentiation of pulmonary pure invasive mucinous adenocarcinoma from mixed mucinous adenocarcinoma. *Technol. Cancer Res. Treat.* **23**, 15330338241258415 (2024).
22. Hou, N. et al. Predicting 30-days mortality for MIMIC-III patients with sepsis-3: a machine learning approach using XGboost. *J. Translational Med.* **18**(1), 462 (2020).
23. Lee, S. I. & Atri, M. FIGO staging system for uterine cervical cancer: Enter cross-sectional imaging. *Radiology* **292**, 15–24 (2019).
24. Wang, W. et al. Multiparametric MRI-based radiomics analysis: Differentiation of subtypes of cervical cancer in the early stage. *Acta Radiol.* **63**, 847–856 (2022).
25. Shi, B. et al. A combination analysis of IVIM-DWI biomarkers and T2WI-based texture features for tumor differentiation grade of cervical squamous cell carcinoma. *Contrast Med. Mol. Imaging* **2022**, 2837905 (2022).
26. Cui, L. et al. Multi-parametric MRI-based peritumoral radiomics on prediction of lymph-vascular space invasion in early-stage cervical cancer. *Diagn. Interven. Radiol.* **28**, 312–321 (2022).
27. Jha, A. K. et al. Systematic review and meta-analysis of prediction models used in cervical cancer. *Artif. Intell. Med.* **139**, 102549 (2023).
28. Wagner-Larsen, K. S. et al. MRI-based radiomic signatures for pretreatment prognostication in cervical cancer. *Cancer Med.* **12**, 20251–20265 (2023).
29. Zhu, Y., Yang, L. & Shen, H. Value of the application of CE-MRI radiomics and machine learning in preoperative prediction of sentinel lymph node metastasis in breast cancer. *Front. Oncol.* **11**, 757111 (2021).

30. Liu, S., Zhou, Y., Wang, C., Shen, J. & Zheng, Y. Prediction of lymph node status in patients with early-stage cervical cancer based on radiomic features of magnetic resonance imaging (MRI) images. *BMC Med. Imaging* **23**(1), 101. https://doi.org/10.1186/s12880-023-01059-6 (2023).
31. Tomita, H. et al. Nodal-based radiomics analysis for identifying cervical lymph node metastasis at levels I and II in patients with oral squamous cell carcinoma using contrast-enhanced computed tomography. *Eur. Radiol.* **31**(10), 7440–7449. https://doi.org/10.1007/s00330-021-07758-4 (2021).
32. Moro, F. et al. Developing and validating ultrasound-based radiomics models for predicting high-risk endometrial cancer. *Ultrasound Obstet. Gynecol.* **60**(2), 256–268. https://doi.org/10.1002/uog.24805 (2022).
33. Zhang, Y., Zhang, Z., Wei, L. & Wei, S. Construction and validation of nomograms combined with novel machine learning algorithms to predict early death of patients with metastatic colorectal cancer. *Front. Public. Health* **10**, 1008137 (2022).
34. Erickson, B. J., Korfiatis, P., Akkus, Z. & Kline, T. L. Machine learning for medical imaging. *Radiographics* **37**, 505–515 (2017).
35. Sarica, A., Cerasa, A. & Quattrone, A. Random forest algorithm for the classification of neuroimaging data in Alzheimer's disease: A systematic review. *Front. Aging Neurosci.* **9**, 329 (2017).
36. Uddin, S., Khan, A., Hossain, M. E. & Moni, M. A. Comparing different supervised machine learning algorithms for disease prediction. *BMC Med. Inf. Decis. Mak.* **19**, 281 (2019).
37. Wang, R., Zhang, J., Shan, B., He, M. & Xu, J. XGBoost machine learning algorithm for prediction of outcome in aneurysmal subarachnoid hemorrhage. *Neuropsychiatric Dis. Treat.* **18**, 659–667 (2022).
38. Garavand, A. et al. Efficient model for coronary artery disease diagnosis: A comparative study of several machine learning algorithms. *J. Healthc. Eng.* **2022**, 5359540 (2022).
39. Noble, W. S. What is a support vector machine? *Nat. Biotechnol.* **24**(12), 1565–1567. https://doi.org/10.1038/nbt1206-1565 (2006).
40. Gillies, R. J., Kinahan, P. E. & Hricak, H. Radiomics: Images are more than pictures, they are data. *Radiology* **278**(2), 563–577. https://doi.org/10.1148/radiol.2015151169 (2016).
41. Huang, Y. Q. et al. Development and validation of a radiomics nomogram for preoperative prediction of lymph node metastasis in colorectal cancer. *J. Clin. Oncol.* **34**(18), 2157–2164. https://doi.org/10.1200/JCO.2015.65.9128 (2016).

Author contributions

Guarantor of integrity of the entire study：X.P., J.C. study concepts：X.P., J.C. study design：X.P., M.Z., Q.Z. definition of intellectual content：M.Z., Q.Z. literature research：M.Z., Q.Z., X.W., J.C. clinical studies：Q.Z., X.W. data acquisition：X.P., J.C. data analysis：X.P., J.C., M.Z., Q.Z. statistical analysis：M.Z., Q.Z. manuscript preparation：M.Z., Q.Z., X.W. manuscript editing：M.Z., Q.Z., X.W., H.Y. manuscript review：H.Y..

Declarations

Competing interests

The authors declare no competing interests.

Ethical approval

All procedures performed in studies involving human participants were in accordance with the 1964 Helsinki declaration and its later amendments or comparable ethical standards. This study is approved by the Medical Ethics Committee of Affiliated Hospital of North Sichuan Medical College, approval number 2022ER451-1 and Written informed consent was obtained.

Additional information

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1038/s41598-025-03170-w.

Correspondence and requests for materials should be addressed to H.Y.