A machine learning model based on clinical features and ultrasound radiomics features for pancreatic tumor classification

Shunhan Yao[1,2†], Dunwei Yao[3,4†], Yuanxiang Huang[5], Shanyu Qin[3*] and Qingfeng Chen[5*]

[1]Medical College, Guangxi University, Nanning, China, [2]Monash Biomedicine Discovery Institute, Monash University, Melbourne, VIC, Australia, [3]Department of Gastroenterology, The First Affiliated Hospital of Guangxi Medical University, Nanning, China, [4]Department of Gastroenterology, The People's Hospital of Baise, Baise, China, [5]School of Computer, Electronic and Information, Guangxi University, Nanning, China

Objective: This study aimed to construct a machine learning model using clinical variables and ultrasound radiomics features for the prediction of the benign or malignant nature of pancreatic tumors.

Methods: 242 pancreatic tumor patients who were hospitalized at the First Affiliated Hospital of Guangxi Medical University between January 2020 and June 2023 were included in this retrospective study. The patients were randomly divided into a training cohort (n=169) and a test cohort (n=73). We collected 28 clinical features from the patients. Concurrently, 306 radiomics features were extracted from the ultrasound images of the patients' tumors. Initially, a clinical model was constructed using the logistic regression algorithm. Subsequently, radiomics models were built using SVM, random forest, XGBoost, and KNN algorithms. Finally, we combined clinical features with a new feature RAD prob calculated by applying radiomics model to construct a fusion model, and developed a nomogram based on the fusion model.

Results: The performance of the fusion model surpassed that of both the clinical and radiomics models. In the training cohort, the fusion model achieved an AUC of 0.978 (95% CI: 0.96−0.99) during 5-fold cross-validation and an AUC of 0.925 (95% CI: 0.86−0.98) in the test cohort. Calibration curve and decision curve analyses demonstrated that the nomogram constructed from the fusion model has high accuracy and clinical utility.

Conclusion: The fusion model containing clinical and ultrasound radiomics features showed excellent performance in predicting the benign or malignant nature of pancreatic tumors.

KEYWORDS

pancreatic tumors, malignant, clinical features, radiomics features, machine learning, classification nomogram

1 Introduction

Pancreatic cancer is one of the major causes of cancer related deaths in developed countries and one of the most lethal malignant tumors in the world (1). Currently, pancreatic cancer has surpassed breast cancer to become the third leading cause of cancer-related deaths in the United States, and it is projected to become the second leading cause, after only lung cancer, of cancer-related mortality before 2040 (2, 3). The high mortality rate of pancreatic cancer primarily stems from the fact that patients are often diagnosed at a late stage of the disease, missing the optimal window for effective treatment (4). This situation is compounded by the lack of a single attributable risk factor for pancreatic cancer and the pancreas' anatomically difficult-to-access location, which impedes routine disease screening. Additionally, early-stage pancreatic cancer typically presents with no symptoms or with nonspecific symptoms, and the absence of diagnostic biomarkers for early-stage tumors further limits early detection (5). The incidence of pancreatic cancer is rising due to a variety of factors, including the obesity epidemic and increased life expectancy (2, 6). The development of new diagnostic strategies is crucial for enhancing treatment decision-making and assessing patient prognosis.

In recent years, with the continual development of artificial intelligence technologies, an increasing number of researchers have begun to explore their application in the diagnosis of pancreatic cancer. At present, the application of artificial intelligence technologies in the genome and transcriptome analysis of pancreatic cancer is relatively mature. For instance, Biao Zhang and colleagues have developed a pancreatic cancer survival prediction model using random forest algorithms and Cox regression analysis, integrating single-cell and bulk RNA sequencing (7). They have also employed eight different AI algorithms to assess the immunological microenvironment features of pancreatic cancer, providing new biomarkers for its diagnosis and treatment (8). However, there is still a shortage of studies on the AI-assisted diagnosis of pancreatic cancer based on clinical and radiological features. Currently, some researchers have used deep learning techniques to analyze pancreatic CT images and have successfully identified pancreatic cancer lesions (9). Other researchers have also utilized clinical variables to construct mathematical models and scoring systems for assessing the risk of pancreatic cancer (10). These methods have shown varying degrees of effectiveness in clinical trials. However, the use of only clinical features or radiomics alone has been found to be insufficient for reliable differential diagnosis (11–14). The development of mathematical models and scoring systems that combine clinical and radiomics features could improve diagnostic accuracy (15). Yu Hu and colleagues developed a model that integrates clinical risk factors and ultrasound features for predicting the malignancy of soft tissue tumors in the limbs (16). Amogh and colleagues applied deep learning to combine clinical variables with MRI data to construct a classification model for prostate cancer (17). These models have achieved high diagnostic performance, indicating that the synergistic use of clinical and radiomics features in model construction can enhance the diagnostic capabilities of the models.

This study aimed to integrate clinical and ultrasound radiomics features of pancreatic tumor patients to construct a fusion model that can accurately predict the benign or malignant nature of pancreatic tumors.

2 Materials and methods

2.1 Patient enrollment

The ethics review board of The First Affiliated Hospital of Guangxi Medical University approved this retrospective study and waived the requirement for written informed consent. The inclusion criteria for patients were as follows: diagnosed with either benign or malignant pancreatic tumors via pathological examination and treated at the First Affiliated Hospital of Guangxi Medical University from January 2020 to June 2023. The exclusion criteria for patients were as follows: (1) had received any antitumor treatment prior to laboratory tests or ultrasonography. (2) Patients with malignant tumors in other parts of the body. (3) Patients with recurrent pancreatic tumors. (4) Patients with incomplete clinical or ultrasound data. A total of 242 eligible patients were randomly divided into two cohorts at a 7:3 ratio: the training cohort (n=169) and the test cohort (n=73).

2.2 Clinical data collection

All patients' clinical data, ultrasound images, and pathological examination results were sourced from the hospital's Health Information System (HIS). The clinical data included general patient information, clinical signs, and laboratory test results. General information included sex, age, and body mass index (BMI). Clinical signs included abdominal pain, jaundice, tumor location, tumor diameter, and tumor subtype. The laboratory test results included blood type, blood sugar, total bilirubin, low-density lipoprotein cholesterol, apolipoprotein B, CA125, CA199, and other parameters. During data preprocessing, features with more than 10% missing values were discarded. For continuous numerical variables, missing values were imputed using the median. For categorical variables, the mode was employed to fill in missing values. Continuous variables such as BMI, CA125 and CA199 were converted into categorical variables according to the criteria of the WHO classification. Ultimately, a complete set of 28 clinical features was obtained.

2.3 Radiomics feature collection

All the procedures were performed according to the Image Biomarker Standardization Initiative (IBSI) standards. Ultrasound images were acquired by Olympus or Fuji ultrasonic equipment with linear probes of 5.0–7.5 MHz and saved in DICOM format for further analysis. Echo texture analysis of the ultrasound images was conducted using a computer program specifically designed for ultrasound image analysis (MaZda v4.6; Institute of Electronics, Technical University of Lodz, Poland) (18).

Endoscopic ultrasonography was performed on patients by gastroenterologists with 10 years of experience, and 1 image containing the tumor was collected for each patient for feature extraction. The ultrasound grayscale images were imported into MaZda software, where a physician with ten years of experience delineated the region of interest (ROI) encompassing the tumor on the images. Subsequently, MaZda software was used to extract the features of the ROIs. Then, to measure interobserver reproducibility, a total of 100 images were randomly selected to be re-segmented by a senior radiologist with 20 years of experience. An interclass correlation coefficient (ICC)>0.75 indicates high feature stability. Discrepancies were resolved through consultation. The radiologists were unaware of the specific histopathological type when delineating the ROI.

A total of 306 features were extracted and categorized into seven major types: (1) first-order statistics; (2) shape-based; (3) gray level co-occurrence matrix (GLCM); (4) gray level run length matrix (GLRLM); (5) gray level size zone matrix (GLSZM); (6) neighboring gray tone difference matrix (NGTDM); (7) gray level dependence matrix (GLDM). Feature preprocessing included handling outliers through log transformation and normalization using min–max scaling. Following these steps, we obtained a radiomics feature dataset for subsequent analysis from 242 patients. The process of radiomics feature extraction and modeling is illustrated in Figure 1.

2.4 Model establishment

Clinical Model: We utilized 28 candidate clinical features to construct the clinical model. In the training cohort, both univariate and multivariate logistic regression analyses were employed to compare the differences in clinical features between patients with benign and malignant pancreatic tumors. Features with significant differences ($p < 0.05$) were selected to establish the clinical model.

Radiomics Model: For the radiomics model, 306 candidate radiomics features were used. In the training cohort, four machine learning algorithms—support vector machine (SVM),

random forest, XGBoost, and K-nearest neighbors (KNN) —were utilized to construct classification models, and the best-performing model was chosen as the radiomics model.

Fusion Model: We calculated the probability of each patient developing a malignant pancreatic tumor using the radiomics model, and named this probability the RDA-prob. A fusion model was constructed through multivariate logistic regression analysis by combining RDA-prob data with clinical data. Based on the fusion model, a nomogram was developed. The process of model construction is illustrated in Figure 2.

2.5 Statistical analysis

Statistical analysis was performed using Python software (version 3.12), R software (version 4.0.2), and SPSS software (version 26.0). For binary categorical variables, the chi-square test was utilized; for continuous variables, comparisons between the two groups were performed using the independent samples t-test or the Wilcoxon rank sum test, depending on the data distribution. A two-tailed P-value < 0.05 was considered to indicate statistical significance. The performance metrics for the models included accuracy, precision, recall, F1 score, and area under the curve (AUC) value, along with the generation of ROC curves. The model's performance was evaluated on the training cohort using 5-fold cross-validation and was also calculated on the test cohort. The performance of the nomogram was assessed through calibration curves and decision curves.

3 Results

3.1 Study population and baseline characteristics

A total of 242 patients with pancreatic tumors were randomly divided at a 7:3 ratio into a training cohort (n = 169) and a test
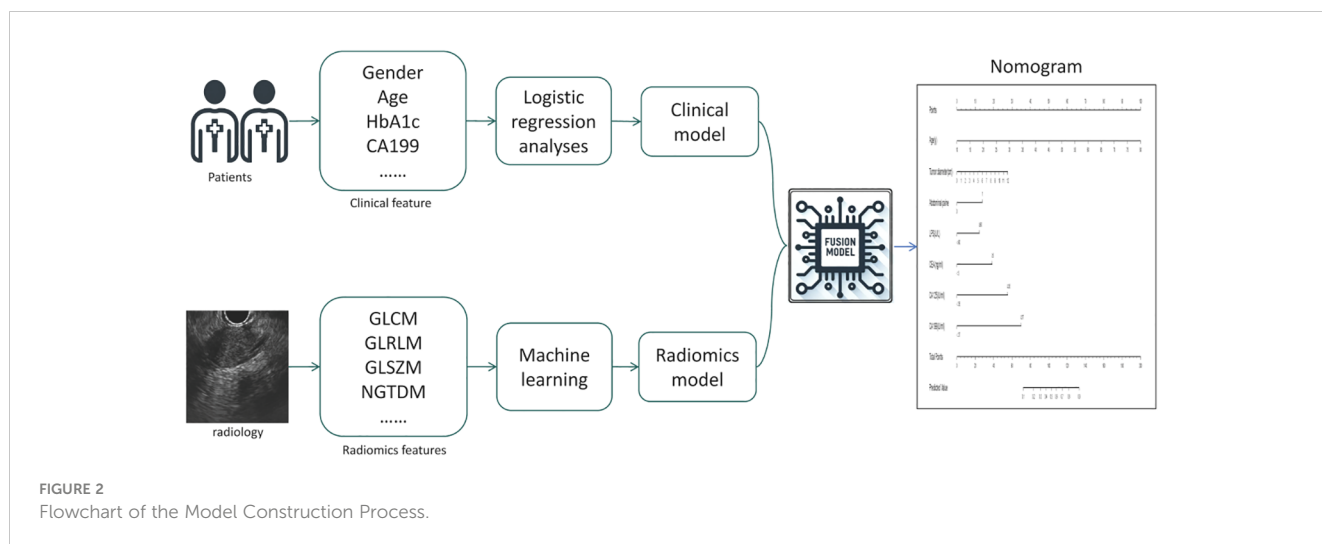

FIGURE 1
Extraction and modeling of radiomics features.
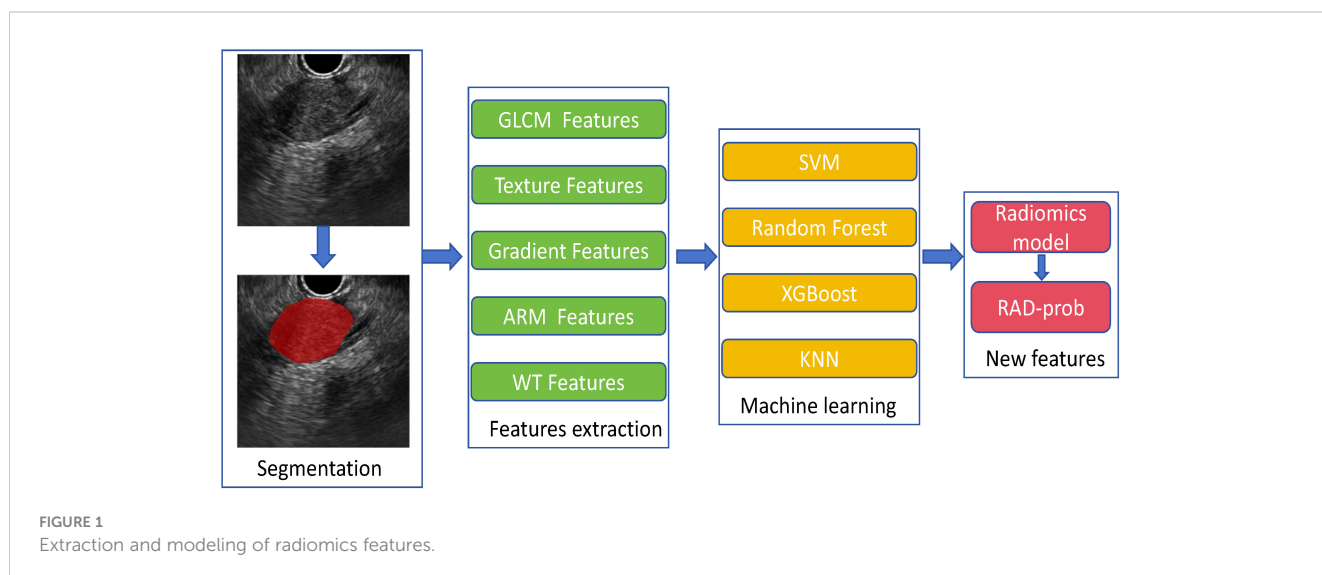
FIGURE 2
Flowchart of the Model Construction Process.

cohort (n = 73). There were no statistically significant differences in basic clinical characteristics, such as sex, age, BMI, abdominal pain, jaundice, tumor diameter, or tumor classification (P > 0.05), between the training and test cohorts. The basic clinical characteristics of the patients are shown in Table 1.

3.2 Clinical model construction and evaluation

In the training cohort, univariate and multivariate logistic regression analyses were conducted on 28 clinical features, revealing that age, abdominal pain, CEA, CA125, CA199, and HbA1c were significant independent predictors. Therefore, these factors were used to construct the classification model. Although the P-value for tumor diameter was slightly above 0.05 (0.058), it was still close to the level of significance. Given its clinical

TABLE 1   Baseline clinical information of patients.

Variable	Training cohort (n=169)	Test cohort (n=73)	P-Value
Male/female	91/78	42/31	0.709
Age	56.84 ± 13.40	58.12 ± 11.67	0.454
BMI	21.90 ± 3.57	21.97 ± 3.06	0.879
abdominal pain	126	53	0.751
Jaundice	35	14	0.785
Pancreatic head	104	47	0.675
Pancreatic body	40	15	0.595
Pancreas tail	34	14	0.866
Tumor diameter	4.04 ± 2.04	3.59 ± 1.72	0.081
Malignant tumor	117	52	0.755
Benign tumor	52	21	0.755

importance, tumor diameter was also included as a feature in the clinical model. The clinical model achieved an AUC of 0.892 (95% CI: 0.85–0.94) in 5-fold cross-validation. In the test cohort, the AUC was 0.882 (95% CI: 0.74–0.95). The results of univariate and multivariate logistic regression analyses in the training cohort are shown in Table 2. The detailed results of the clinical model's performance are presented in Table 3. The ROC curve is illustrated in Figure 3A.

3.3 Radiomics model construction and evaluation

In the training cohort, classification models were constructed using 306 radiomics features with four different algorithms: SVM, random forest, XGBoost, and KNN. The performance of the model constructed using these four algorithms is shown in Table 4. Among them, the model built using the KNN algorithm performed the best and is identified as the radiomics model. The radiomics model achieved an AUC of 0.854 (95% CI: 0.78–0.92) in 5-fold cross-validation (5-CV), and an AUC of 0.739 (95% CI: 0.61–0.87) in the test cohort. The ROC curve for the radiomics model is illustrated in Figure 3B. The top 20 primary features of the radiomics model are shown in Figure 3D.

3.4 Fusion model construction and evaluation

The radiomics model was utilized to calculate the probability of each patient developing a malignant pancreatic tumor, and these probabilities were assigned to a new feature named the RDA-prob. By combining the RDA-prob with clinical data in a multivariate logistic regression analysis, four features were selected for constructing the fusion model: age, RDA-prob, CA125, and CA199. The fusion model achieved an AUC of 0.978 (95% CI: 0.96–0.99) in 5-fold cross-validation and an AUC of 0.925 (95% CI: 0.86–0.98) in the test cohort. The fusion model consistently outperformed the other two models in both the five-fold cross-

TABLE 2 Results of the logistic regression analysis in the Training cohort.

Feature	Univariate analysis OR (95%CI)	p-value	Multivariate analysis OR (95%CI)	p-value
Sex	0.89 [1.47, 3.89]	0.738	–	–
Age	1.11 [0.00, 0.05]	0.000	1.12 [1.08, 1.17]	0.000
BMI	0.91 [0.83, 1.00]	0.046	0.95 [0.84, 1.06]	0.353
Abdominal pain	4.42 [2.12, 9.21]	0.000	4.32 [1.61, 11.59]	0.004
Jaundice	2.02 [1.38, 2.83]	0.126	–	–
Pancreatic head	1.12 [1.25, 3.52]	0.732	–	–
Pancreatic body	1.45 [1.43, 2.99]	0.367	–	–
Pancreas tail	0.77 [1.64, 3.44]	0.523	–	–
Tumor diameter	1.26 [0.43, 1.98]	0.016	1.24 [0.99, 1.55]	0.058
Blood type A	1.01 [0.49, 2.08]	0.977	–	–
Blood type B	1.20 [0.53, 2.71]	0.661	–	–
Blood type AB	–	1.000	–	–
Blood type O	0.63 [0.32, 1.23]	0.176	–	–
D-II	1.96 [0.43, 4.45]	0.109	–	–
HbA1c	3.07 [0.86, 6.24]	0.002	3.44 [1.22, 9.68]	0.02
BS	1.28 [0.61, 2.69]	0.514	–	–
T-BIL	1.89 0.92, 3.86]	0.081	–	–
D-BIL	2.32 [1.12, 4.79]	0.023	2.51 [0.52, 12.14]	0.25
ALP	2.17 [1.06, 4.43]	0.033	0.36 [0.07, 1.76]	0.21
T-CHO	1.08 [0.51, 2.29]	0.838	–	–
TG	0.93 [0.45, 1.92]	0.841	–	–
LDL-C	0.80 [0.40, 1.59]	0.531	–	–
apoB	0.82 [0.38, 1.76]	0.602	–	–

(Continued)

TABLE 2 Continued

Feature	Univariate analysis OR (95%CI)	p-value	Multivariate analysis OR (95%CI)	p-value
LPS	1.34 [0.61, 2.94]	0.459	–	–
AMS	0.64 [0.21, 1.90]	0.420	–	–
CEA	11.36 [3.35, 38.58]	0.000	4.99 [1.17, 21.28]	0.03
CA125	12.63 [4.28, 37.29]	0.000	10.31 [2.73, 38.91]	0.00
CA199	10.81 [4.93, 23.68]	0.000	12.91 [4.84, 34.41]	0.00

BMI, body mass index; D-II, D-Dimer; HbA1c, Glycated hemoglobin A1c; BS, Blood Sugar; T-BIL, Total Bilirubin; D-BIL, Direct Bilirubin; ALP, Alkaline Phosphatase; T-CHO, Total Cholesterol; TG, Triglycerides; LDL-C, Low-density lipoprotein cholesterol; apoB, Apolipoprotein B; LPS, Lipopolysaccharide; AMS, Amylase; CEA, Carcinoembryonic Antigen; CA125, Carbohydrate Antigen 125; CA199, Carbohydrate Antigen 19–9.

validation and the test cohort, demonstrating its robustness and generalizability. The detailed results of the performance of the fusion model are provided in Table 3. The ROC curve for the fusion model is shown in Figure 3C.

3.5 Nomogram

A nomogram was constructed based on the variables selected by the fusion model (Figure 4A). This nomogram provides a visual method for measuring the results of the fusion model, allowing for the rapid calculation of predictive probabilities for individual patients. To use the nomogram, one must first determine the position of each variable on its respective axis and then draw a line to the points axis to obtain the score for that variable. The scores for all variables are summed to obtain a total score, which is then located on the total points axis. The probability corresponding to this total score is the likelihood of the pancreatic tumor being malignant. Calibration curves (Figures 4B, C) demonstrate that the nomogram has good accuracy in both the training and test cohorts. The decision curve (Figure 4D) indicates that the nomogram has high clinical utility.

4 Discussion

In our study, we discovered that age, abdominal pain, tumor diameter, CA199, CEA, CA125, LPS, and HbA1c are independent predictors of malignant pancreatic tumors. These findings align with those of other studies (19–25). Several studies have shown a positive correlation between age, abdominal pain, tumor diameter and the probability of predicting malignant pancreatic tumors (19, 20). CA199 is the only serum biomarker for pancreatic cancer recommended in the European Oncology Guidelines, yet its sensitivity and specificity are approximately 80%, indicating limited diagnostic performance (21). Like CA199, CA125 is a high-molecular-weight glycoprotein that is

TABLE 3 Performance of the clinical model, radiomics model, and fusion model.

	Accuracy	Precision	Recall	F1 Score	AUC
5-CV					
clinical model	0.834	0.868	0.897	0.882	0.892
Radiomics model	0.811	0.824	0.923	0.871	0.854
Fusions model	0.917	0.940	0.940	0.940	0.978
Test cohort					
clinical model	0.849	0.918	0.865	0.891	0.882
Radiomics model	0.699	0.759	0.846	0.800	0.739
Fusions model	0.863	0.904	0.904	0.904	0.925

currently primarily used as an auxiliary diagnostic tool for ovarian cancer (22). Several studies have shown a positive correlation between age, abdominal pain, tumor diameter, and the probability of predicting malignant pancreatic tumors (23). CEA, LPS, and HbA1c have also been confirmed by multiple studies to be associated with pancreatic cancer (24, 25). The results of our study support these findings, and based on the identified independent predictors, we constructed a clinical classification model. This model demonstrated good classification performance and exhibited robustness in both the training cohort and the test cohort.

Subsequently, our study analyzed endoscopic ultrasound images of patients with pancreatic tumors. Radiographic images can provide quantitative and reproducible texture information (26), and by extracting and quantifying texture features within images, significant reference information can be provided for the diagnosis, treatment, and prognosis, with the potential to enhance the diagnostic performance of models. Recently, several papers have been published by scholars on the application of image texture features in patients with pancreatic lesions. Zixing Huang and colleagues compared CT image texture features of patients with
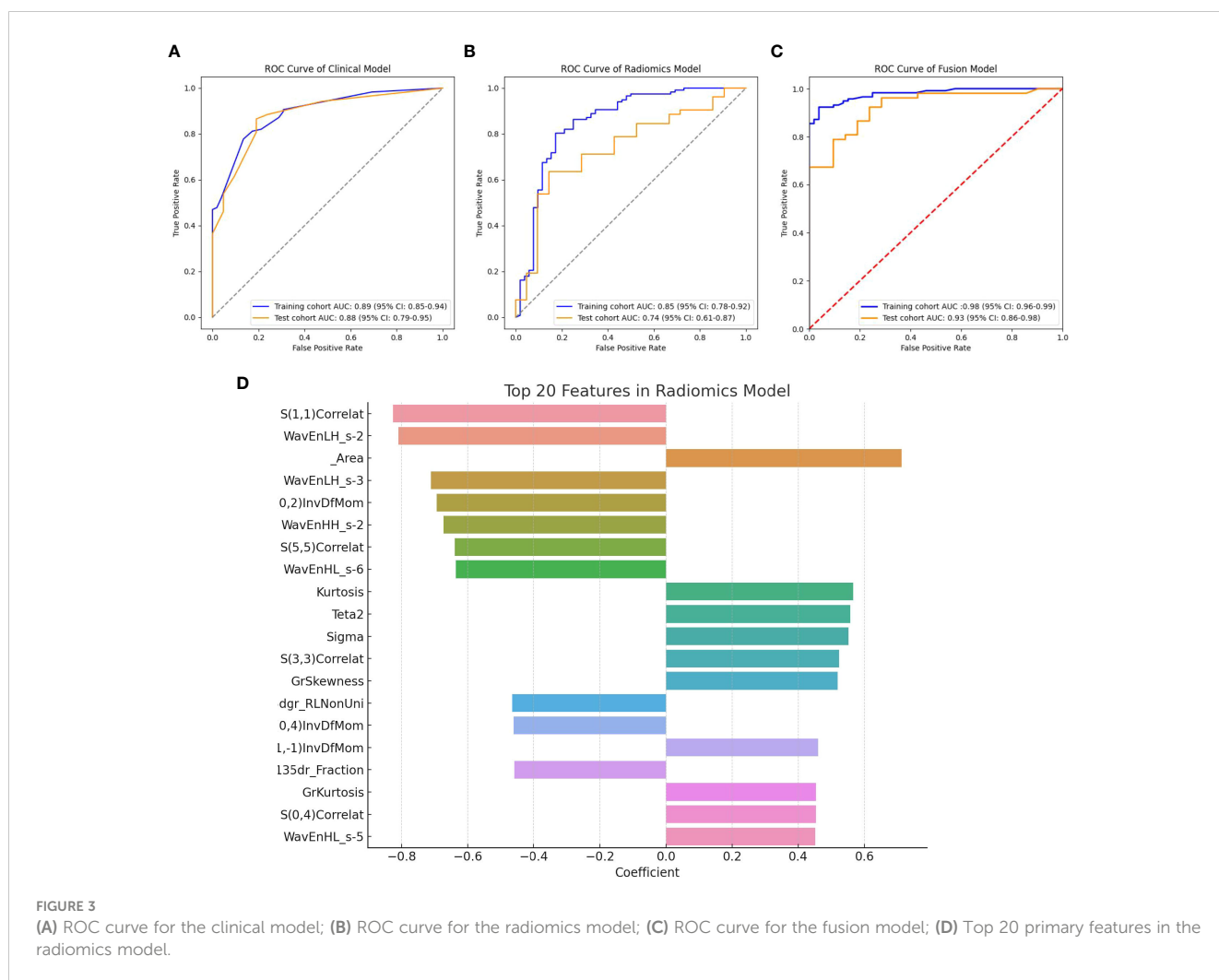

FIGURE 3
(A) ROC curve for the clinical model; (B) ROC curve for the radiomics model; (C) ROC curve for the fusion model; (D) Top 20 primary features in the radiomics model.

TABLE 4 Performance of radiation models constructed by different algorithms.

	Accuracy	Precision	Recall	F1 Score	AUC
5-CV					
SVM	0.740	0.759	0.923	0.831	0.780
Random Forest	0.734	0.776	0.873	0.819	0.780
XGBoost	0.752	0.801	0.855	0.827	0.798
KNN	0.811	0.824	0.923	0.871	0.854
Test cohort					
SVM	0.753	0.793	0.885	0.836	0.639
Random Forest	0.671	0.769	0.769	0.769	0.620
XGBoost	0.644	0.750	0.750	0.750	0.601
KNN	0.699	0.759	0.846	0.800	0.739

pancreatic lymphoma and pancreatic adenocarcinoma and identified 60 texture features with significant differences (27). Xudong Li and others analyzed MRI images of 119 pancreatic tumor patients and found four texture features that could provide information for tumor differentiation (28). These research results offer promise for the development of new radiological biomarkers.

In our study, we extracted 306 radiomics features from endoscopic ultrasound images of pancreatic tumors and constructed radiomic classification models using four different machine learning algorithms to compare their performance. The results showed that the model based on the KNN algorithm performed optimally, hence it was chosen as the primary radiomic model. Given the difficulty in interpreting the association between individual ultrasound radiomics features and potential pathological characteristics, we adopted the research methodologies of Fei Yao (29) and Weichen Zhang (30), integrating multiple ultrasonic radiomics features into a multi-element parameter, RAD-prob, to simplify the complexity of multi-feature analysis. The radiomics features used to construct RAD-prob range from basic morphological characteristics to advanced texture features, revealing the variability, heterogeneity, and inconsistency of the lesions. For instance, Gray Level Co-occurrence Matrix (GLCM) features such as S(1,1) Correlat and S (5,5) Correlat demonstrate the correlation between pixels in different directions; wavelet transform energy features like WavEnLH_s-2 and WavEnHH_s-2 reveal high-frequency textures and edge details of the image; first-order statistical features such as Area and Kurtosis describe the size of the image region and the peak characteristics of the gray level distribution; and inverse difference moment features like 0,2) InvDfMom indicate the uniformity and local similarity of the image texture. These comprehensive feature analyses not only exhibit the physical and geometrical attributes of pancreatic cancer but also highlight key biological markers in the ultrasound images, providing valuable insights into the structure and tissue characteristics of the lesion area, thereby supporting clinical diagnosis and treatment decisions.

To date, no single radiological or laboratory test has been able to reliably distinguish between malignant and nonmalignant pancreatic tumors (31). The preliminary diagnosis of pancreatic
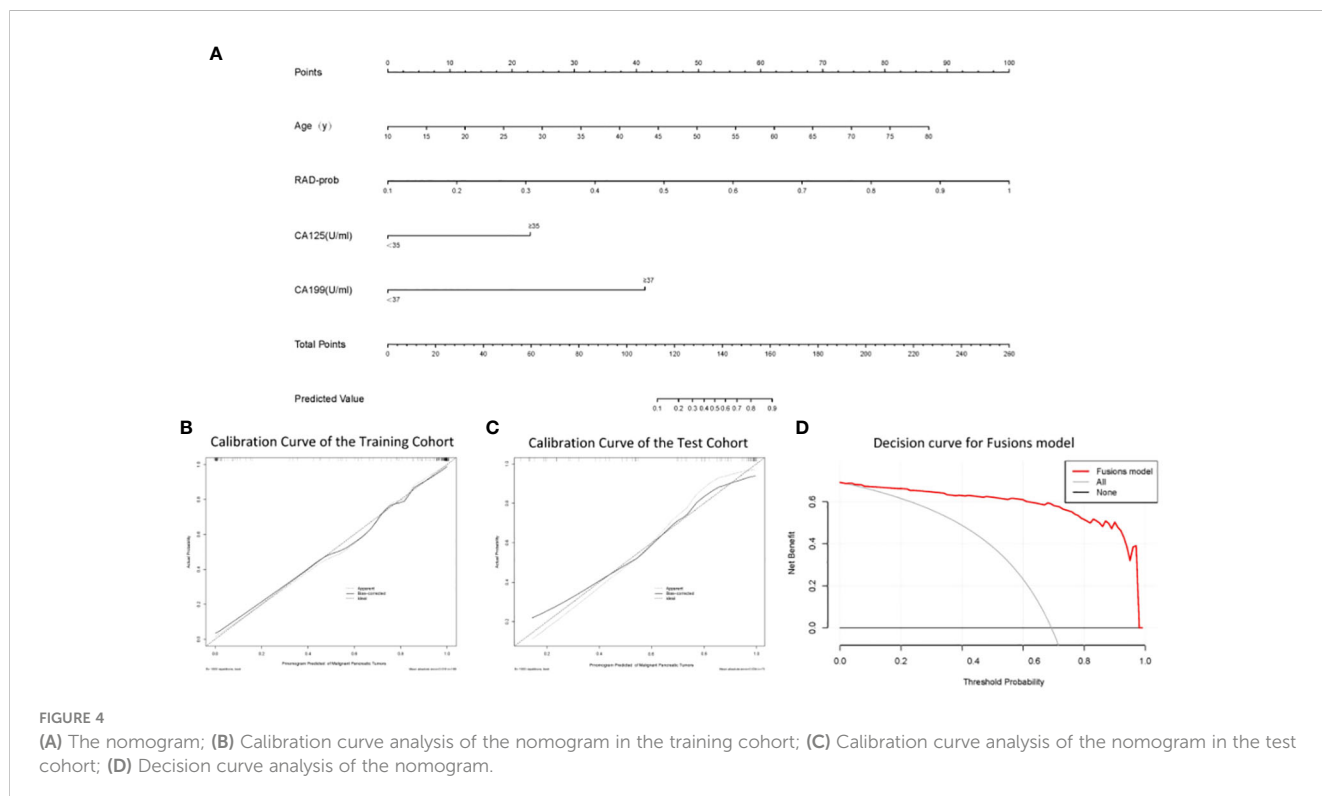

FIGURE 4
(A) The nomogram; (B) Calibration curve analysis of the nomogram in the training cohort; (C) Calibration curve analysis of the nomogram in the test cohort; (D) Decision curve analysis of the nomogram.

tumors depends on a comprehensive analysis of clinical features, laboratory tests, and radiological examinations rather than a single parameter (32). Traditional diagnostic methods, which rely on doctors interpreting test reports, are limited by the individual experiences of the doctors, potentially leading to inconsistent diagnoses (33). Artificial intelligence-based diagnostic models can objectively assist in diagnosis by utilizing a variety of different parameters (33, 34). Qian Lu and colleagues developed a multimodal model that integrated CT features and clinical characteristics, which performed better in predicting lymph node metastasis in pancreatic cancer than single-modality models (35). Wenjie Liang and colleagues developed and validated a nomogram model that combines radiomics features and clinical characteristics and showed excellent performance in differentiating the pathological grade of pancreatic neuroendocrine tumors (36). In our study, we also constructed a classification model that integrates clinical and radiomics features. We first used the radiomics model to calculate a new radiomics feature named RAD-prob and then combined the RAD-prob with clinical features to construct a fusion model through logistic regression. The performance of the fusion model surpassed that of both the clinical and radiomics models. In the training cohort, the combined model achieved an AUC of 0.978 (95% CI: 0.96–0.999) during 5-fold cross-validation. In the test cohort, the AUC was 0.925 (95% CI: 0.86–0.98).

Our fusion model offers a promising tool to assist doctors in making diagnostic decisions. To facilitate the clinical application of the model, we developed a nomogram based on the model parameters. This nomogram can quickly calculate predictive probabilities for individual patients, which may lead to its widespread use. Although the model developed in this study achieved good predictive results, there is still room for further optimization. In the future, methods such as ensemble learning and deep learning could be employed to further enhance the model's predictive performance. Additionally, validation with multicenter data is needed to assess the model's performance in real-world settings.

5 Conclusion

In summary, this study constructed three pancreatic tumor classification models using clinical and ultrasound radiomics features, each capable of differentiating between benign and malignant pancreatic tumors. The fusion model, which includes both clinical and ultrasound radiomics features, demonstrated exceptional performance in predicting the benign or malignant nature of pancreatic tumors. The nomogram of the fusion model, as a visual and personalized tool, can assist doctors in accurately identifying the type of pancreatic tumor.

Data availability statement

The data supporting this study's findings are available from the corresponding author upon reasonable request.

Ethics statement

The studies involving humans were approved by The Ethics Review Board of The First Affiliated Hospital of Guangxi Medical University. The studies were conducted in accordance with the local legislation and institutional requirements. Written informed consent for participation was not required from the participants or the participants' legal guardians/next of kin in accordance with the national legislation and institutional requirements.

Author contributions

SY: Data curation, Formal analysis, Methodology, Writing – original draft. DY: Data curation, Methodology, Writing – original draft. YH: Methodology, Writing – original draft. SQ: Conceptualization, Data curation, Writing – review & editing. QC: Conceptualization, Writing – review & editing.

Funding

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

References

1. Huang J, Lok V, Ngai CH, Zhang L, Yuan J, Lao XQ, et al. Worldwide Burden of, Risk Factors for, and Trends in Pancreatic Cancer. *Gastroenterology*. (2021) 160:744–54. doi: 10.1053/j.gastro.2020.10.007

2. Kleeff J, Korc M, Apte M, La Vecchia C, Johnson CD, Biankin AV, et al. Pancreatic cancer. *Nat Rev Dis Primers*. (2016) 2:16022. doi: 10.1038/nrdp.2016.22

3. Halbrook CJ, Lyssiotis CA, Pasca di Magliano M, Maitra A. Pancreatic cancer: Advances and challenges. *Cell*. (2023) 186:1729–54. doi: 10.1016/j.cell.2023.02.014

4. Stoffel EM, Brand RE, Goggins M. Pancreatic cancer: changing epidemiology and new approaches to risk assessment, early detection, and prevention. *Gastroenterology*. (2023) 164:752–65. doi: 10.1053/j.gastro.2023.02.012

5. Klein AP. Pancreatic cancer epidemiology: understanding the role of lifestyle and inherited risk factors. *Nat Rev Gastroenterol Hepatol*. (2021) 18:493–502. doi: 10.1038/s41575-021-00457-x

6. McWilliams RR, Petersen GM. Overweight, obesity, and pancreatic cancer: beyond risk alone. *JAMA*. (2009) 301:2592–3. doi: 10.1001/jama.2009.922

7. Zhang B, Sun J, Guan H, Guo H, Huang B, Chen X, et al. Integrated single-cell and bulk RNA sequencing revealed the molecular characteristics and prognostic roles of neutrophils in pancreatic cancer. *Aging (Albany NY)*. (2023) 15:9718–42. doi: 10.18632/aging.205044

8. Zhang B, Liu J, Li H, Huang B, Zhang B, Song B, et al. Integrated multi-omics identified the novel intratumor microbiome-derived subtypes and signature to predict the outcome, tumor microenvironment heterogeneity, and immunotherapy response for pancreatic cancer patients. *Front Pharmacol*. (2023) 14:1244752. doi: 10.3389/fphar.2023.1244752

9. Gandikota HP, S A, M SK. CT scan pancreatic cancer segmentation and classification using deep learning and the tunicate swarm algorithm. *PloS One*. (2023) 18:e0292785. doi: 10.1371/journal.pone.0292785

10. Peng L, Chen H. A novel nomogram and risk classification system based on inflammatory and immune indicators for predicting prognosis of pancreatic cancer patients with liver metastases. *Cancer Med*. (2023) 12:18622–32. doi: 10.1002/cam4.6471

11. Singhi AD, Wood LD. Early detection of pancreatic cancer using DNA-based molecular approaches. *Nat Rev Gastroenterol Hepatol*. (2021) 18:457–68. doi: 10.1038/s41575-021-00470-0

12. Young Bang J, Krall K, Jhala N, Jhala N, Singh C, Tejani M, Arnoletti JP, et al. Comparing needles and methods of endoscopic ultrasound-guided fine-needle biopsy to optimize specimen quality and diagnostic accuracy for patients with pancreatic masses in a randomized trial. *Clin Gastroenterol Hepatol*. (2021) 19:825–835.e7. doi: 10.1016/j.cgh.2020.06.042

13. Wang J, Zheng T, Liao Y, Geng S, Li J, Zhang Z, et al. Machine learning prediction model for post- hepatectomy liver failure in hepatocellular carcinoma: A multicenter study. *Front Oncol*. (2022) 12:986867. doi: 10.3389/fonc.2022.986867

14. Zhao G, Chen X, Zhu M, Liu Y, Wang Y. Exploring the application and future outlook of Artificial intelligence in pancreatic cancer. *Front Oncol*. (2024) 14:1345810. doi: 10.3389/fonc.2024.1345810

15. Thomasian NM, Kamel IR, Bai HX. Machine intelligence in non-invasive endocrine cancer diagnostics. *Nat Rev Endocrinol*. (2022) 18:81–95. doi: 10.1038/s41574-021-00543-9

16. Hu Y, Li A, Zhao CK, Ye XH, Peng XJ, Wang PP, et al. A multiparametric clinic-ultrasomics nomogram for predicting extremity soft-tissue tumor Malignancy: a combined retrospective and prospective bicentric study. *Radiol Med*. (2023) 128:784–97. doi: 10.1007/s11547-023-01639-0

17. Hiremath A, Shiradkar R, Fu P, Mahran A, Rastinehad AR, Tewari A, et al. An integrated nomogram combining deep learning, Prostate Imaging-Reporting and Data System (PI-RADS) scoring, and clinical variables for identification of clinically significant prostate cancer on biparametric MRI: a retrospective multicentre study. *Lancet Digit Health*. (2021) 3:e445–54. doi: 10.1016/S2589-7500(21)00082-0

18. Strzelecki M, Szczypinski P, Materka A, Klepaczko A. A software tool for automatic classification and segmentation of 2D/3D medical images. *Nucl Inst Methods Phys Research A*. (2013) 702:137–40. doi: 10.1016/j.nima.2012.09.006

19. Pan M, Yang Y, Teng T, Lu F, Chen Y, Huang H. Development and validation of a simple-to-use nomogram to predict liver metastasis in patients with pancreatic neuroendocrine neoplasms: a large cohort study. *BMC Gastroenterol*. (2021) 21:101. doi: 10.1186/s12876-021-01685-w

20. Vernerey D, Huguet F, Vienot A, Goldstein D, Paget-Bailly S, Van Laethem JL, et al. Prognostic nomogram and score to predict overall survival in locally advanced untreated pancreatic cancer (PROLAP). *Br J Cancer*. (2016) 115:281–9. doi: 10.1038/bjc.2016.212

21. Ducreux M, Cuhna AS, Caramella C, Hollebecque A, Burtin P, Goéré D, et al. Cancer of the pancreas: ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up. *Ann Oncol*. (2017) 28:iv167–8. doi: 10.1093/annonc/mdv295

22. Bottoni P, Scatena R. The role of CA 125 as tumor marker: biochemical and clinical aspects. *Adv Exp Med Biol*. (2015) 867:229–44. doi: 10.1007/978-94-017-7215-0_14

23. Skulimowski A, Durczyński A, Strzelczyk J, Hogendorf P. Comparison of clinical usefulness of serum Ca125 and CA19-9 in pancreatic adenocarcinoma diagnosis: meta-analysis and systematic review of literature. *Biomarkers*. (2021) 26:287–95. doi: 10.1080/1354750X.2021.1876770

24. Terakawa N, Satoi S, Takai S, Yanagimoto H, Takahashi K, Komiyama Y, et al. Clinical monitoring of innate cellular immunity of monocytes/macrophages by tumor necrosis factor alpha productivity in whole blood stimulated by lipopolysaccharide in patients with pancreatic cancer. *Pancreas*. (2006) 33:31–7. doi: 10.1097/01.mpa.0000226880.66130.79

25. Haas M, Heinemann V, Kullmann F, Laubender RP, Klose C, Bruns CJ, et al. Prognostic value of CA 19-9, CEA, CRP, LDH and bilirubin levels in locally advanced and metastatic pancreatic cancer: results from a multicenter, pooled analysis of patients receiving palliative chemotherapy. *J Cancer Res Clin Oncol*. (2013) 139:681–9. doi: 10.1007/s00432-012-1371-3

26. Acharya UR, Hagiwara Y, Sudarshan VK, Chan WY, Ng KH. Towards precision medicine: from quantitative imaging to radiomics. *J Zhejiang Univ Sci B*. (2018) 19:6–24. doi: 10.1631/jzus.B1700260

27. Huang Z, Li M, He D, Wei Y, Yu H, Wang Y, et al. Two-dimensional texture analysis based on CT images to differentiate pancreatic lymphoma and pancreatic adenocarcinoma: A preliminary study. *Acad Radiol*. (2019) 26:e189–95. doi: 10.1016/j.acra.2018.07.021

28. Li X, Zhu H, Qian X, Chen N, Lin X. MRI texture analysis for differentiating nonfunctional pancreatic neuroendocrine neoplasms from solid pseudopapillary neoplasms of the pancreas. *Acad Radiol*. (2020) 27:815–23. doi: 10.1016/j.acra.2019.07.012

29. Yao F, Ding J, Lin F, Xu X, Jiang Q, Zhang L, et al. Nomogram based on ultrasound radiomics score and clinical variables for predicting histologic subtypes of epithelial ovarian cancer. *Br J Radiol*. (2022) 95:20211332. doi: 10.1259/bjr.20211332

30. Zhang W, Wang Q, Liang K, Lin H, Wu D, Han Y, et al. Deep learning nomogram for preoperative distinction between Xanthogranulomatous cholecystitis and gallbladder carcinoma: A novel approach for surgical decision. *Comput Biol Med*. (2024) 168:107786. doi: 10.1016/j.compbiomed.2023.107786

31. Wang K, Wang X, Pan Q, Zhao B. Liquid biopsy techniques and pancreatic cancer: diagnosis, monitoring, and evaluation. *Mol Cancer*. (2023) 22:167. doi: 10.1186/s12943-023-01870-3

32. Wood LD, Canto MI, Jaffee EM, Simeone DM. Pancreatic cancer: pathogenesis, screening, diagnosis, and treatment. *Gastroenterology*. (2022) 163:386–402.e1. doi: 10.1053/j.gastro.2022.03.056

33. Gore JC. Artificial intelligence in medical imaging. *Magn Reson Imaging*. (2020) 68:A1–4. doi: 10.1016/j.mri.2019.12.006

34. Chen M, Copley SJ, Viola P, Lu H, Aboagye EO. Radiomics and artificial intelligence for precision medicine in lung cancer treatment. *Semin Cancer Biol*. (2023) 93:97–113. doi: 10.1016/j.semcancer.2023.05.004

35. Lu Q, Zhou C, Zhang H, Liang L, Zhang Q, Chen X, et al. A multimodal model fusing multiphase contrast-enhanced CT and clinical characteristics for predicting lymph node metastases of pancreatic cancer. *Phys Med Biol*. (2022) 67:10.1088/1361-6560/ac858e. doi: 10.1088/1361-6560/ac858e

36. Liang W, Yang P, Huang R, Xu L, Wang J, Liu W, et al. A combined nomogram model to preoperatively predict histologic grade in pancreatic neuroendocrine tumors. *Clin Cancer Res*. (2019) 25:584–94. doi: 10.1158/1078-0432.CCR-18-1305