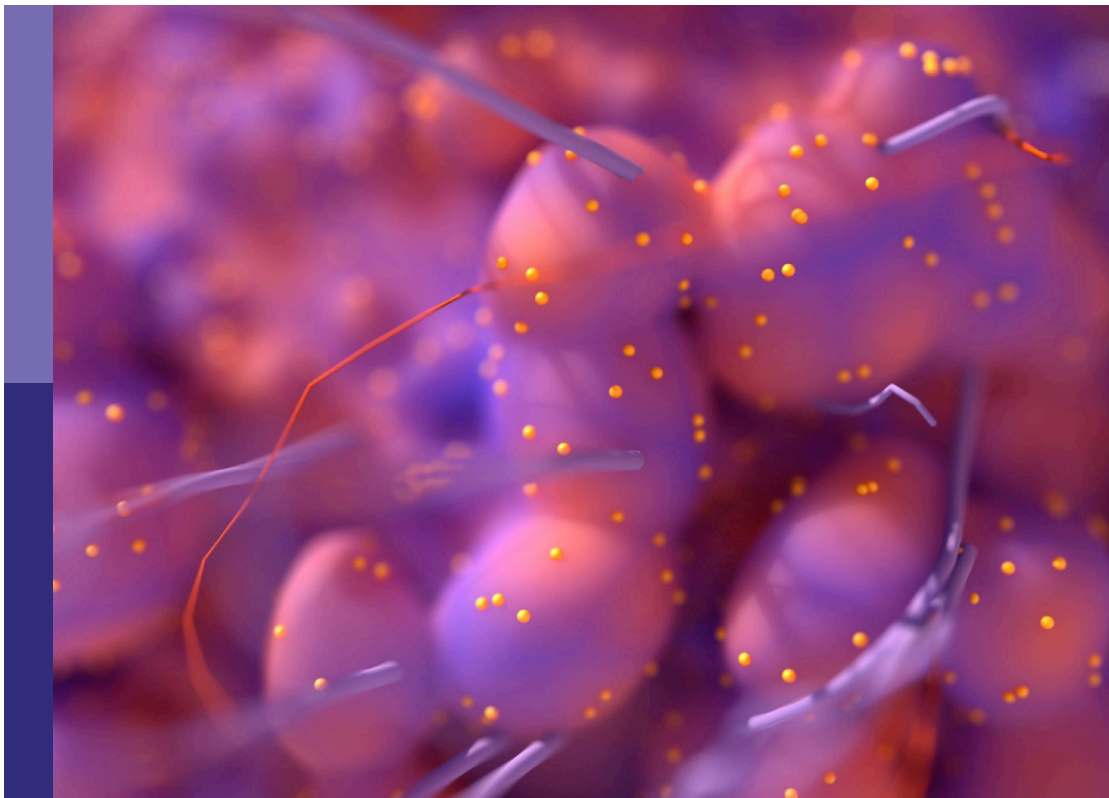
frontiers | Research Topics

Application of radiomics in understanding tumor biological behaviors and treatment response

Edited by
Pei Yang, Zhongxing Liao, Xue Meng, Sweet Ping Ng, Nanna Maria Sijtsema and Lei Xing

Published in
Frontiers in Oncology



ISSN 1664-8714
ISBN 978-2-8325-3397-0
DOI 10.3389/978-2-8325-3397-0

About Frontiers

Frontiers is more than just an open access publisher of scholarly articles: it is a pioneering approach to the world of academia, radically improving the way scholarly research is managed. The grand vision of Frontiers is a world where all people have an equal opportunity to seek, share and generate knowledge. Frontiers provides immediate and permanent online open access to all its publications, but this alone is not enough to realize our grand goals.

Frontiers journal series

The Frontiers journal series is a multi-tier and interdisciplinary set of open-access, online journals, promising a paradigm shift from the current review, selection and dissemination processes in academic publishing. All Frontiers journals are driven by researchers for researchers; therefore, they constitute a service to the scholarly community. At the same time, the *Frontiers journal series* operates on a revolutionary invention, the tiered publishing system, initially addressing specific communities of scholars, and gradually climbing up to broader public understanding, thus serving the interests of the lay society, too.

Dedication to quality

Each Frontiers article is a landmark of the highest quality, thanks to genuinely collaborative interactions between authors and review editors, who include some of the world's best academicians. Research must be certified by peers before entering a stream of knowledge that may eventually reach the public - and shape society; therefore, Frontiers only applies the most rigorous and unbiased reviews. Frontiers revolutionizes research publishing by freely delivering the most outstanding research, evaluated with no bias from both the academic and social point of view. By applying the most advanced information technologies, Frontiers is catapulting scholarly publishing into a new generation.

What are Frontiers Research Topics?

Frontiers Research Topics are very popular trademarks of the *Frontiers journals series*: they are collections of at least ten articles, all centered on a particular subject. With their unique mix of varied contributions from Original Research to Review Articles, Frontiers Research Topics unify the most influential researchers, the latest key findings and historical advances in a hot research area.

Find out more on how to host your own Frontiers Research Topic or contribute to one as an author by contacting the Frontiers editorial office: frontiersin.org/about/contact

Application of radiomics in understanding tumor biological behaviors and treatment response

Topic editors

Pei Yang — Central South University, China
Zhongxing Liao — University of Texas MD Anderson Cancer Center, United States
Xue Meng — Department of Radiation Oncology, Shandong Cancer Hospital, China
Sweet Ping Ng — University of Melbourne, Australia
Nanna Maria Sijtsema — University of Groningen, University Medical Center Groningen, Groningen, Netherlands
Lei Xing — Stanford University, United States

Citation

Yang, P., Liao, Z., Meng, X., Ng, S. P., Sijtsema, N. M., Xing, L., eds. (2023). *Application of radiomics in understanding tumor biological behaviors and treatment response*. Lausanne: Frontiers Media SA. doi: 10.3389/978-2-8325-3397-0

Table of contents

frontiers | Frontiers in Oncology

TYPE Editorial
PUBLISHED 18 August 2023
DOI 10.3389/fonc.2023.1257447

OPEN ACCESS

EDITED AND REVIEWED BY
Zaver Bhujwalla,
Johns Hopkins University, United States

*CORRESPONDENCE
Nanna Maria Sijtsema
n.m.sijtsema@umcg.nl
Pei Yang
proyangpei@163.com

RECEIVED 12 July 2023
ACCEPTED 28 July 2023
PUBLISHED 18 August 2023

CITATION
Xiao N, Pei Z, Lu W, Fang R, Jin Y, Zhou G, Meng X, Ng SP, Xing L, Liao Z, Sijtsema NM and Yang P (2023) Editorial: Application of radiomics in understanding tumor biological behaviors and treatment response. *Front. Oncol.* 13:1257447. doi: 10.3389/fonc.2023.1257447

Editorial: Application of radiomics in understanding tumor biological behaviors and treatment response

Ningping Xiao[1], Zhengda Pei[1,2], Wenhui Lu[2], Rongyao Fang[2], Yi Jin[1], Guanzhi Zhou[3], Xue Meng[4], Sweet Ping Ng[5], Lei Xing[6], Zhongxing Liao[7], Nanna Maria Sijtsema[3]* and Pei Yang[1]*

[1]Hunan Cancer Hospital, Xiangya School of Medicine, Central South University, Changsha, Hunan, China, [2]Graduate Collaborative Training Base of Hunan Cancer Hospital, Hengyang Medical School, University of South China, Hengyang, Hunan, China, [3]Department of Radiotherapy, University Medical Center Groningen, Groningen, Netherlands, [4]Department of Radiation Oncology, Shandong Cancer Hospital, Jinan, Shandong, China, [5]Department of Radiation Oncology, Olivia Newton-John Cancer Centre, Austin Health, Melbourne, VIC, Australia, [6]School of Medicine, Stanford University, Stanford, CA, United States, [7]Department of Radiation Oncology, The University of Texas MD Anderson Cancer Center, Houston, TX, United States

KEYWORDS

radiomics, machine learning, cancer, biological behaviors, artificial intelligence, individual treatment

Editorial on the Research Topic
Application of radiomics in understanding tumor biological behaviors and treatment response

Introduction

Cancer is a serious threat to human health worldwide, with a high mortality rate and increasing morbidity over the years (1). In general, patients with the same clinical stage have significant differences in survival and prognosis due to the high heterogeneity of tumor biological behavior. It is challenging to individualize treatment according to a uniform tumor stage treatment model. Radiomics is one of the indispensable tools for screening, diagnosis, treatment, and follow-up of multiple tumors. Recent data suggest that advanced post-treatment anatomic imaging with post-processing and registration capabilities can be used to characterize the likelihood and the location of potential failures to optimize treatment strategies and improve quality of life

With the wide application of artificial intelligence, the emerging radiomics technology, as a quantitative and high-throughput radiology method, has shown the ability to obtain quantitative texture information from existing medical image data from anatomical structures non-invasively, and has become the gold standard for pre-treatment staging and post-treatment tumor control evaluation (2). In addition, radiomics can further leverage existing "big data" analysis of images to provide hitherto unimaginable predictive power. It leverages powerful big data/machine learning techniques to refine its approach to massive data processing to identify clinically applicable, non-invasive methods to extract oncology outcomes and toxicity prevention data from large-scale data. This

approach offers superior scalability, clinical applicability, cost-effectiveness, ease of implementation and an unmatched value proposition

Papers included in this Research Topic

Radiomics is playing an increasing role in cancer diagnosis and treatment as well as biological research. We introduced this Research Topic to develop a comprehensive predictive model that provides standard recommendations for the multidisciplinary management of multiple cancer causes.The long-term goal is to reduce the mortality rate of cancer patients and improve their quality of life, thereby increasing cost efficiency. We are encouraged by the strong support of the research community for our Research Topic. We were very pleased to see many excellent works submitted to our research project, and we finally published 16 papers, including 13 original studies and 3 reviews, most of which were retrospective studies. The thesis covers brain glioma, nasopharyngeal carcinoma, breast cancer, small cell lung cancer, renal cell carcinoma, hepatocellular carcinoma, pancreatic cancer, rectal cancer and other tumors. The objectives of the study are diverse, including tumor status assessment, differential diagnosis, survival and recurrence assessment, and genomic feature prediction.

Accurate identification and evaluation using radiomics is helpful to develop appropriate treatment plans for patients and avoid unnecessary treatment measures such as surgery, postoperative radiotherapy and chemotherapy. For example, Wang et al. analyzing mammography (DM) images using radiomics, a radiomic line model was established to distinguish benign and malignant circular masses. Gao et al. used radiomics features based on enhanced CT images from the corticomedullary stage (CMP) and nephrography stage (NP) in combination with important clinical factors to distinguish between papillary renal cell carcinoma type 1 (pRCC) and pRCC type 2 tumors by multivariate logistic regression analysis before surgery. In addition, Lu et al. built a radiomic nomogram model based on CT images for patients with focal autoimmune pancreatitis and pancreatic ductal adenocarcinoma in accurate areas. The AUC of the training group and the test group were 0.87 and 0.83, respectively. As a non-invasive predictive tool, the model can improve the accuracy of diagnosis while reducing patient trauma and achieving optimal compliance. The classification and type of tumors are different, and the treatment methods are not consistent. For example, according to the multitask learning model developed by Huang, Y. et al. in combination with support vector machines to distinguish glioblastoma from isolated brain metastases, the mean AUC of the model in the training set and validation set was as high as 0.993 and 0.987. If the preoperative prognosis is primary glioblastoma, aggressive triple therapy, such as postoperative concurrent chemoradiotherapy, is required.

Image-based radiomics models can assist clinicians in treatment evaluation, including predicting the response of individual cancer patients to chemotherapy or immunotherapy, as well as monitoring recurrence and metastasis. These aspects are clearly reflected in the research papers accepted for this radiomics Research Topic: Jiang et al., Yang et al., Wang, Y. et al., Lin et al., and Huang, Y.-M. et al. used several radiomics based ML models and columns to predict patient outcomes, such as survival, mortality, efficacy, postoperative metastasis, and recurrence. Wang, Y. et al.'s radiomics model based on GIST morphological features plays an important role in tumor risk stratification, and the AUC value of the model is 0.933.It can provide reference for clinical diagnosis and treatment plan, formulate the best treatment strategy for individuals according to the predicted results of the model, create customized treatment plan for patients, and improve the treatment effect and later quality of life.

Conclusion

Radiomics has shown promising results in some areas of oncology, including tumor screening, detection, diagnosis, treatment, and prognosis prediction. The 16 studies collected under this study theme apply radiomics to construct comprehensive predictive models to provide optimal recommendations for multidisciplinary management of multiple tumors to address treatment options in clinical practice. Imaging has demonstrated the potential to improve the foresight and accuracy of the diagnosis and treatment of cancer patients. It is promising to applicate radiomics in clinical practice to improve the efficiency of clinicians, reduce the possibility of clinical decision-making errors, and reduce unnecessary procedures, interventions, and medical costs.

All the studies in the subject of this study are retrospective and have certain limitations. The data levels included in each study are not sufficient, and there is no unified standard reference for algorithms such as image source, lesion delineation and feature extraction. These studies are still complex for clinicians and difficult to be thoroughly accepted. More research may need to focus on image and data standardization between different institutions, data sharing, and prospective studies to increase generalization of results.

This Research Topic involves a number of studies and presents the application of radiomics in understanding tumor biological behavior and treatment response. We thank all reviewers and authors for their contributions to this Research Topic. We hope that this Research Topic will attract more attention in related fields.

Author contributions

NX: Writing – original draft. ZP: Writing – original draft. WL: Writing – original draft. RF: Writing – original draft. YJ: Writing – review & editing. GZ: Writing – review & editing. XM: Writing – review & editing. SN: Writing – review & editing. LX: Writing – review & editing. ZL: Supervision, Writing – review & editing. NS: Supervision, Writing – review & editing. PY: Supervision, Writing – review & editing.

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. Torre LA, Siegel RL, Ward EM, Jemal A. Global cancer incidence and mortality rates and trends–an update. *Cancer Epidemiol Biomarkers Prev* (2016) 25(1):16–27. doi: 10.1158/1055-9965.EPI-15-0578

2. Gillies RJ, Kinahan PE, Hricak H. Radiomics: images are more than pictures, they are data. *Radiology* (2016) 278(2):563–77. doi: 10.1148/radiol.2015151169

frontiers | Frontiers in Oncology

ORIGINAL RESEARCH
published: 26 April 2022
doi: 10.3389/fonc.2022.677803

A Radiomics Nomogram for Distinguishing Benign From Malignant Round-Like Breast Tumors

Lanyun Wang[1†], Yi Ding[1†], Wenjun Yang[1], Hao Wang[1], Jinjiang Shen[1], Weiyan Liu[2], Jingjing Xu[3], Ran Wei[1], Wenjuan Hu[1], Yaqiong Ge[4], Bin Song[1] and Bei Zhang[5*]*

[1] Department of Radiology, Minhang Hospital, Fudan University, Shanghai, China, [2] Department of General Surgery, Minhang Hospital, Fudan University, Shanghai, China, [3] Department of Medical Examination Center, Minhang Hospital, Fudan University, Shanghai, China, [4] General Electric (GE) Healthcare, Shanghai, China, [5] Department of Radiology, Ruijin Hospital, School of Medicine, Shanghai Jiao Tong University, Shanghai, China

OPEN ACCESS

Edited by:
Sweet Ping Ng,
University of Melbourne, Australia

Reviewed by:
Chenbin Liu,
Chinese Academy of Medical Sciences and Peking Union Medical College, China
Weiwei Zong,
Henry Ford Health System, United States

***Correspondence:**
Bin Song
songbin@fudan.edu.cn
Bei Zhang
zhangbei650213@126.com

[†]These authors have contributed equally to this work and share first authorship

Specialty section:
This article was submitted to Breast Cancer, a section of the journal Frontiers in Oncology

Received: 08 March 2021
Accepted: 21 March 2022
Published: 26 April 2022

Citation:
Wang L, Ding Y, Yang W, Wang H, Shen J, Liu W, Xu J, Wei R, Hu W, Ge Y, Song B and Zhang B (2022) A Radiomics Nomogram for Distinguishing Benign From Malignant Round-Like Breast Tumors. Front. Oncol. 12:677803. doi: 10.3389/fonc.2022.677803

Objective: The objective of this study is to develop a radiomics nomogram for the presurgical distinction of benign and malignant round-like solid tumors.

Methods: This retrospective trial enrolled patients with round-like tumors who had received preoperative digital mammography (DM) no sooner than 20 days prior to surgery. Breast tumors were segmented manually on DM images in order to extract radiomic features. Four machine learning classification models were constructed, and their corresponding areas under the receiver operating characteristic (ROC) curves (AUCs) for differential tumor diagnosis were calculated. The optimal classifier was then selected for the validation set. After this, predictive machine learning models that employed radiomic features and/or patient features were applied for tumor assessment. The models' AUC, accuracy, negative (NPV) and positive (PPV) predictive values, sensitivity, and specificity were then derived.

Results: In total 129 cases with benign and malignant tumors confirmed by pathological analysis were enrolled in the study, including 91 and 38 in the training and test sets, respectively. The DM images yielded 1,370 features per patient. For the machine learning models, the Least Absolute Shrinkage and Selection Operator for Gradient Boosting Classifier turned out to be the optimal classifier (AUC=0.87, 95% CI 0.76-0.99), and ROC curves for the radiomics nomogram and the DM-only model were statistically different (P<0.001). The radiomics nomogram achieved an AUC of 0.90 (95% CI 0.80-1.00) in the test cohort and was statistically higher than the DM-based model (AUC=0.67, 95% CI 0.51-0.84). The radiomics nomogram was highly efficient in detecting malignancy, with accuracy, sensitivity, specificity, PPV, and NPV in the validation set of 0.868, 0.950, 0.778, 0.826, and 0.933, respectively.

Conclusions: This radiomics nomogram that combines radiomics signatures and clinical characteristics represents a noninvasive, cost-efficient presurgical prediction technique.

Keywords: machine learning, radiomics nomogram, round-like tumors, digital mammography, breast

BACKGROUND

Breast cancer (BC) represents the most common malignant disease in females globally (1–3). However, BC mortality has declined remarkably since the 1970s (4) due in large part to mammography screening and enhanced systemic therapy (5, 6). Digital mammography (DM) constitutes a widely accepted means of breast tumor detection in clinic (6, 7) and has come to play a critical role in the evaluation of breast tumors, taking precedence over other techniques in detecting calcification in breast tumors. Normally, breast cancers display an irregular shape and a spiculated margin in DM images, with or without microcalcification. However, some malignant tumors appear as round-like masses, including mucinous carcinoma, intraductal papillary carcinoma, medullary carcinoma, infiltrating ductal carcinoma, intramammary metastases, metaplastic carcinoma, lymphoma, and phyllodes tumors (8, 9). The margins of these tumors can be circumscribed, microlobulated, and indistinct. Clinicians therefore face a challenge in distinguishing the above tumors from benign lesions by DM alone, especially for dense breast tissues. In addition, the prognosis and clinical treatment of these tumors are necessarily different. Preoperative assessment of round-like tumors can thus help to optimize treatment.

The signals of round-like masses without suspicious malignant or benign macrocalcification in DM are comparable, although their internal structures and densities show substantial differences (9). Previous findings (10) recommend that the classification according to the Breast Imaging Reporting and Data System (BI-RADS) of breast masses found on DM images should be determined in combination with ultrasound (US) or magnetic resonance imaging (MRI), except for completely calcified or fatty masses. Routine imaging techniques such as mammography and US show overt limitations in the differential diagnosis of round-like masses. Although the multimodal technique of breast dynamic contrast-enhanced MRI (DCE-MRI) is highly accurate in distinguishing benign from malignant tumors (11) as well as in differentiating well-circumscribed breast malignant lesions from benign ones (8), it requires contrast media injection and is very expensive. In addition, traditional multimodal diagnosis greatly relies on the radiologist's experience. Though DM is the most applied technique in assessing breast tumors, no quantitative parameters have yet been derived from DM images (12). Therefore, the identification of a measurable DM marker may greatly increase the diagnostic value of this technique for breast tumors. Radiomics could be used to convert digital images into high-dimensional data by extracting a variety of quantitative indices and could thus help to quantitatively evaluate tumor heterogeneity and improve clinical decision making (13, 14).

To this end, this work aims to develop a radiomics nomogram for distinguishing benign from malignant round-like masses without spiculated margins and suspicious malignant calcification or benign macrocalcification, in order to help optimize treatment plans.

Abbreviations: DM, Digital mammography; BI-RADS, Breast Imaging Reporting and Data System; US, ultrasound; MRI, magnetic resonance imaging; DCE-MRI, dynamic contrast-enhanced MRI; PACS, Picture Archiving and Communication System; CC, cranial caudal; MLO, mediolateral oblique; DICOM, Digital Imaging and Communications in Medicine; ROI, Region of interest; ICC, intraclass correlation coefficient; CI, confidence interval; mRMR, maximum correlation minimum redundancy; LASSO, least absolute shrinkage and selection operator; SVM, support vector machine; k-NN, k-Nearest Neighbor; ROC, receiver operating characteristic; AUC, area under the receiver operating characteristic curve; DCA, Decision curve analysis; PPV, positive predictive value; NPV, negative predictive value.

METHODS

Data Cohort

Our institutional review board approved this retrospective study, with no requirement for informed consent. Individuals who received DM screening with benign (excluding fibroadenoma) or malignant tumors confirmed by pathological analysis were continually enrolled between January, 2017 and December, 2019. Due to a high prevalence of fibroadenoma, patients who received DM screening with fibroadenoma confirmed by pathological analysis were continually enrolled between January and December, 2019. Patient data were obtained from the Picture Archiving and Communication System (PACS) of the Affiliated Minhang Hospital of Fudan University, Shanghai, China.

The inclusion criteria were: (1) the presence of an oval or round tumor; (2) a DM exam carried out within 20 preoperative days, with image quality meeting post-processing requirements; (3) nonmalignant or cancerous breast tumor confirmed by histopathology. The exclusion criteria were: (1) receiving treatments, (chemotherapy, surgery, radiotherapy and/or anti-HER2 therapy) before DM screening; (2) the tumor being incompletely displayed in the cranial caudal (CC) or mediolateral oblique (MLO) views; (3) the tumor being architecturally distorted (except for scarring caused by a previous injury or surgery); (4) the tumor showing calcification of BI-RADS 2/4b/4c/5; (5) the tumor having spiculated margins; and (6) the tumor not being displayed due to extremely dense breasts. Ultimately, 129 masses (51 nonmalignant and 78 cancerous) were included, and their histopathologic diagnoses are presented in **Table 1**. The 129 study cases, age 54.6 ± 13.7 years (range, 23–86 years) old, were randomly assigned to the training (n = 91) and test (n = 38) sets.

DM and Image Processing

A GE Senographe Essential DM system (GE Healthcare, Milwaukee, WI) was utilized for data acquisition. In every case, optimal MLO and CC view images were converted into Digital Imaging and Communications in Medicine (DICOM) files. ITK-SNAP software (http://www.itk-snap.org) was utilized for breast tumor segmentation, and regions of interest (ROIs) were manually segmented on MLO and CC views independently by two radiologists (WY and LW) with 10 and 14 years of experience in DM image evaluation, respectively. In cases of obscured tumor margins, both radiology experts reached a consensus by performing an additional image analysis.

Feature Extraction and Selection

Radiomic features were obtained with AK v3.2.2 software (GE healthcare). In total 1,370 features were obtained, including

TABLE 1 | Features of 129 breast tumors confirmed by histology.

Histopathologic type		No. of masses	Proportion (%)	No. of masses with calcifications	BI-RADS category of accompanying calcifications
Benign		51	39.5	0	
	Fibroadenoma	44	34.1	0	
	Intraductal papilloma	2	1.6	0	
	Benign phyllodes tumor	4	3.1	0	
	Tubular gland lymphoma	1	0.7	0	
Malignant		78	60.5	5	
	Invasive ductal carcinoma	53	41.1	4	4a (3) 3 (1)
	Intraductal papillary carcinoma	8	6.2	0	
	Ductal carcinoma *in situ*	1	0.8	0	
	Neuroendocrine carcinoma	1	0.8	0	
	Malignant phyllodes tumor	3	2.3	0	
	Mucinous carcinoma	11	8.5	1	4a
	Sarcomatoid carcinoma	1	0.8	0	

No., number; BI-RADS, Breast Imaging Reporting and Data System.

histogram, shape, gray-level co-occurrence matrix (GLCM), gray-level run-length matrix (GLRLM) and gray-level size zone matrix (GLSZM) features. For interobserver agreement evaluation, CC views were randomly chosen in 30 cases, and another radiologist delineated ROIs independently. After this, intraclass correlation coefficients (ICCs) of these features were calculated. Based on the ICC's 95% confidence interval (CI) (15), values >0.90, from 0.75 to 0.9, from 0.5 to 0.75, and <0.5 were considered to reflect excellent, good, moderate, and poor reliability, respectively. Only features with ICC ≥0.75 were included in subsequent analysis.

The patients were randomized into the training and test sets (ratio of 7:3, respectively). Initially, the maximum correlation minimum redundancy (mRMR) algorithm was used for eliminating redundant and irrelevant parameters in the training set, of which 30 features that showed high correlations with labels were retained. Next, least absolute shrinkage and selection operator (LASSO) analysis with 10-fold cross-validation was performed to further select features *via* λ optimization. The coefficients of select features then underwent compression to zero at the optimal λ value, and only parameters that showed a nonzero coefficient were further retained.

Patient, DM, and US Features

The following clinical information was obtained from the patients' medical records: age, sex, family history of breast cancer, life habits (drinking/smoking), and childbearing information. Next, DM data were analyzed by two radiologists, as stated above, who recorded the following parameters: (1) tumor size (maximum diameter); (2) margin (circumscribed, obscured, microlobulated, or indistinct); (3) density (low, equal, or high); and (4) location (depth) (anterior, middle, or posterior). Additionally, US data were recorded as described in the US report. The imaging features of US were: (1) echo pattern (anechoic, hypoechoic, isoechoic, complex cystic, and solid, heterogeneous, or hyperechoic); (2) edge (clear, partially clear, or unclear); (3) shape (regular, partially regular, or irregular); and (4) blood flow (presence or none).

Radiomics Signature, Clinical Model, and Radiomics Nomogram

Four machine learning models, support vector machine (SVM), k-Nearest Neighbor (k-NN), C-Tree, and logistic regression, were constructed based on the previously obtained optimal feature subset described above. All classifiers underwent training with 10-fold cross-validation with 10 repeats in the training cohort. Their predictive performances were then assessed with cross-validation data and validated in the validation cohort, and the optimal classifier in the validation set was selected. Next, radiomics scores (rad-scores) for various patients were determined. The radiomics signature was evaluated for predictive accuracy by the area under the receiver operating characteristic (ROC) curve (AUC) in both the training and test cohorts.

Clinical data, including age, MG, and US characteristics (continuous data) were analyzed by independent samples *t*-test or the Wilcoxon test (for example, age and tumor size (DM)), and the Chi square test or Fisher's exact test were carried out for analyzing categorical variables such as tumor size (DM), margin (DM), density (DM), location (depth) (DM), echo pattern (US), edge (US), shape (US), and blood flow (US). Univariate logistic analysis was applied to select risk factors for cancerous tumors ($P<0.05$), and this was followed by backward stepwise multivariate logistic regression and likelihood ratio tests in order to build a clinical prediction model. In order to satisfy the collinearity condition, features with both the largest calculated VIF and VIF >10 were eliminated. The model's performance was then determined by ROC curve evaluation.

After this, a radiomics nomogram was built as described above for the clinical model, including the obtained radiomics signature, and its performance was also examined by ROC analysis. Finally, the Hosmer-Lemeshow test was performed to assess consistency between actual and predicted values. The radiomic framework is shown in **Figure 1**.

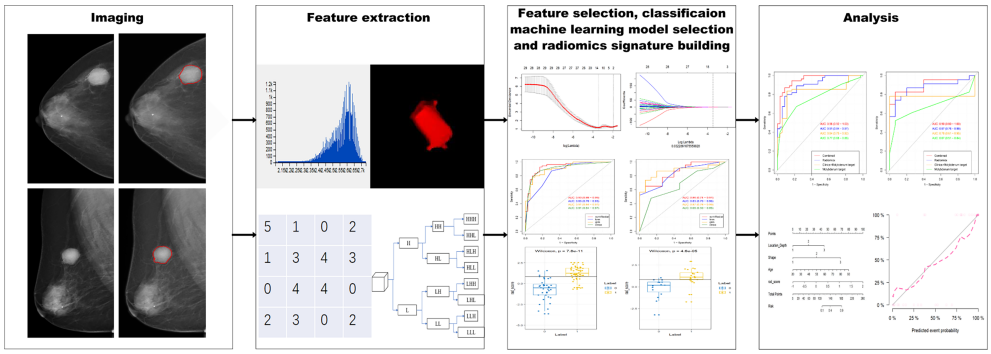


FIGURE 1 | Flow chart of radiomic analysis of round-like masses on DM images.

Radiomics Nomogram Validation and Evaluation

The radiomics nomogram was examined in the training (n = 91) and test (n = 38) cohorts, respectively, with regard to differentiation, calibration and clinical values, and the AUC was determined in order to evaluate the nomogram's performance in distinguishing malignant and benign tumors. The Hosmer–Lemeshow test and the calibration curve were utilized as well to determine the goodness-of-fit. Additionally, internal validation was carried out in the test cohort. A rad-score was derived in the test cohort based on the algorithm built in the training cohort, and decision curve analysis (DCA) was performed to estimate the nomogram's robustness in a clinical setting.

Statistical Analysis

The software packages *R* v. 3.5.1 (https://www.Rproject.org) and SPSS were utilized for all statistical analysis.

RESULTS

Patient and DM/US Features

All patients in this study were female. Regarding some clinical factors, of the 129 patients, three patients (malignant, 2; benign, 1) had a family history of breast cancer; two patients (malignant, 1; benign, 1) had a smoking habit; four patients (malignant, 2; benign, 2) had a drinking habit; and two patients with benign tumors had never given birth. **Table 2** summarizes patient and DM/US features. Age, margin (DM), density (DM), location depth (DM), edge (US), shape (US), and blood flow (US) differed significantly between malignant and benign tumors.

In our univariate logistic regression analysis, factors including age, margin (DM), density (DM), location depth (DM), edge (US), shape (US), and blood flow (US) showed significant associations with malignant masses (all $P<0.05$; **Table 3**), and multivariate logistic regression analysis suggested that age, location depth (DM), shape (US), and rad_score were independent predictors of malignant masses (**Table 4**). These clinical variables were then employed to construct a clinical model that had an AUC value of 0.78 (0.61-0.95) in the testing cohort, which was higher than that of DM 0.67 (0.51-0.84).

Radiomics for Predictive Modeling

In the training set, 13 top-performing features (histogram, shape, and texture features), including 5 and 8 from the CC and MLO views, respectively, were finally selected by the LASSO logistic regression model (**Figures 2A, B**). **Figure 2C** shows the selected radiomics features, and the MLO view had more features than the CC view (8 and 5, respectively). Four classification machine learning models were constructed using the above selected 13 top-performing features, and the performances of the four classification machine learning models are shown in **Figure 3**. The logistic regression model had high AUC values of 0.91 for the training set (**Figure 3A**) and 0.87 for the test set (**Figure 3B**), but the LASSO-based machine learning model showed the best detection performance. The boxplot in **Figure 4** shows the accuracies, AUCs, NPV, PPV, and sensitivities and specificities of the four models after a 100-time cross-validation.

Figure 2D shows the rad-score of each patient determined by logistic regression. Individuals with cancerous tumors generally had higher rad-scores compared to the benign group, and rad-scores were statistically different between individuals with benign and malignant masses in the training and test cohorts (both $P<0.001$). The generated radiomics signature had good predictive accuracy, with AUCs of 0.91 (95% CI 0.86–0.98) and 0.87 (95%CI 0.76–0.99) in the training and test sets, respectively (**Figure 2E**).

Nomogram

According to multivariate logistic regression, location (depth), shape (US), age, and the radiomics signature all independently predicted malignancy in round-like tumors and were therefore included in a radiomics nomogram (**Figure 5A**). **Figures 5B, C** depict the nomogram's calibration curves. In both the training and test cohorts, the curves reflected good calibration, and the Hosmer-Lemeshow test showed non-significance ($P=0.375$), (**Figures 5B, C**).

TABLE 2 | Patient and DM/US characteristics.

Characteristic		Pathological type		*P* value
		Benign	Malignant	
Margin (DM)	circumscribed	17	18	0.021*
	obscured	31	41	
	microlobulated	0	0	
	indistinct	3	19	
Density (DM)	low-density	1	1	0.000*
	equal-density	41	34	
	high- density	9	43	
Location/Depth (DM)	anterior	7	15	0.050*
	middle	35	38	
	posterior	9	25	
Echo pattern (US)	anechoic	0	1	0.104
	hypoechoic	49	64	
	isoechoic	1	1	
	complex cystic and solid	1	3	
	heterogeneous	0	9	
	hyperechoic	0	0	
Edge (US)	clear	7	3	0.002*
	partially clear	43	58	
	unclear	1	17	
Shape (US)	regular	6	3	<0.001*
	partially regular	44	41	
	irregular	1	34	
Blood flow (US)	none	22	13	0.001*
	presence	29	65	
Age #		45 (41~52)	60.5 (50.5~70)	<0.001*
Size #		1.9 (1.6~2.8)	2.3 (1.6~3.225)	0.158

**means P<0.05; # means nonnormal distribution obtained after SK normality test; DM, digital mammography; US, ultrasound.*

TABLE 3 | Positive results of univariate analysis for the differential diagnosis of round-like breast tumors.

Variable	2.5%CI	97.5%CI	OR value	*P* value
Age	1.031	1.110	1.068	0.001*
Margin (DM)	1.236	3.347	1.951	0.008*
Density (DM)	2.010	13.221	4.917	0.001*
Location_Depth (DM)	1.022	4.422	2.064	0.050
Edge (US)	2.335	51.850	8.197	0.005*
Shape (US)	4.334	95.747	15.082	0.000*
Blood_flow (US)	1.228	9.486	3.321	0.020*

**means P<0.05; DM, digital mammography; US, ultrasound; CI, confidence interval; OR, odds ratio.*

Figures 6A, B both present four ROC curves that compare digital mammography, the clinical model, the radiomics signature, and the radiomics nomogram for efficiency in differentiating round-like masses. The DeLong's test showed that the ROC curves of the radiomics nomogram and digital mammography were statistically different (P<0.001), and the radiomics nomogram had an AUC of 0.90 (95% CI 0.80-1.00) in the test set, suggesting a significantly higher performance versus the prediction model constructed only with digital mammography features, which had an AUC of 0.67 (95% CI 0.51-0.84) in the test cohort.

The radiomics nomogram had high efficacy in detecting malignancy, with accuracy, sensitivity, specificity, PPV, and NPV of 0.890, 0.941, 0.825, 0.873,and 0.917, respectively, in the training cohort, and 0.868, 0.950, 0.778, 0.826, and 0.933 in the test cohort, respectively (**Table 5**). Machine learning-based mammography radiomics had an AUC of 0.87 (95% CI 0.76-0.99), indicating a better performance than the clinical model (AUC=0.78, 95% CI 0.61-0.95 in the test cohort) (**Figure 6B**). In the test set, the radiomics signature had higher specificity and PPV compared to the radiomics nomogram, and the radiomics nomogram had improved predictive ability (accuracy, sensitivity, and NPV) compared to the radiomics signature in distinguishing benign and malignant round-like tumors (**Table 5**).

DISCUSSION

This study developed a radiomics signature for predicting malignancy in round-like masses that had good accuracy in

TABLE 4 | Positive results of multivariate logistic regression analysis for the differential diagnosis of round-like breast tumors.

Variable	2.50%CI	97.50%CI	OR	P value
Location_Depth (DM)	1.197	16.582	3.978	0.036*
Shape (US)	1.900	57.442	7.969	0.013*
Age	1.024	1.134	1.072	0.006*
Rad_score	2.821	33.017	8.060	<0.001*
Intercept	<0.001	<0.001	<0.001	<0.001*

**means P<0.05; DM, digital mammography; US, ultrasound; CI, confidence interval; OR, odds ratio.*

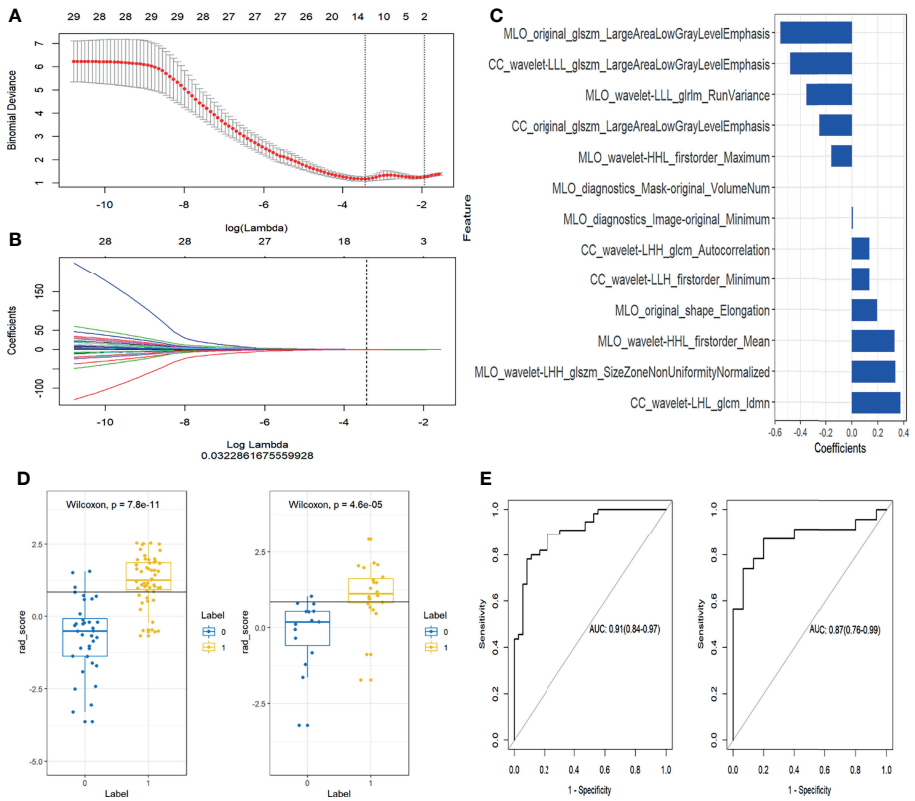


FIGURE 2 | Selection of radiomics features and evaluation of the prediction performance of the radiomics signature. **(A)** Selection of the hyperparameter (λ) in the least absolute shrinkage and selection operator (LASSO) model *via* ten-fold cross-validation based on minimum error; vertical black dotted line, optimal value of λ (best fit). **(B)** Coefficients and log(λ) values; features with nonzero coefficients are shown. **(C)** The 13 features showing nonzero coefficients are displayed. The features utilized for constructing the radiomics signature are shown on the y-axis with the corresponding coefficients in LASSO Cox analysis on the x-axis. **(D)** Rad-scores of benign and malignant masses in the training and test groups. Yellow and blue represent the actual classification: the greater the separation of yellow and blue, the better the rad-score's predictive accuracy. **(E)** Receiver operating characteristic (ROC) curves of the radiomics signature in the training and test set.

identifying the type of lesions (AUC=0.87 in the entire population). Further, a novel radiomics nomogram, built by utilizing multivariate logistic regression data, showed good calibration and was able to distinguish benign from malignant tumors in both the training and test data sets. The AUC of this signature was 0.90, suggesting a higher predictive value of the nomogram compared to mammography alone (AUC=0.67) as well as to the clinical model (AUC=0.78) that was established on the basis of age, DM, and US. In clinic DM and US characteristics are relied upon routinely for differential diagnosis. However, their values are dependent upon the radiologist's experience. In addition, radiomics features are purely objective and quantitative.

This study compared four commonly used classification machine learning methods (SVM, C-Tree, k-NN, and logistic analysis) and found that LASSO had the best performance. As shown above, LASSO had higher AUC compared with the remaining classification machine learning methods. Indeed, LASSO can perform both feature selection and normalization

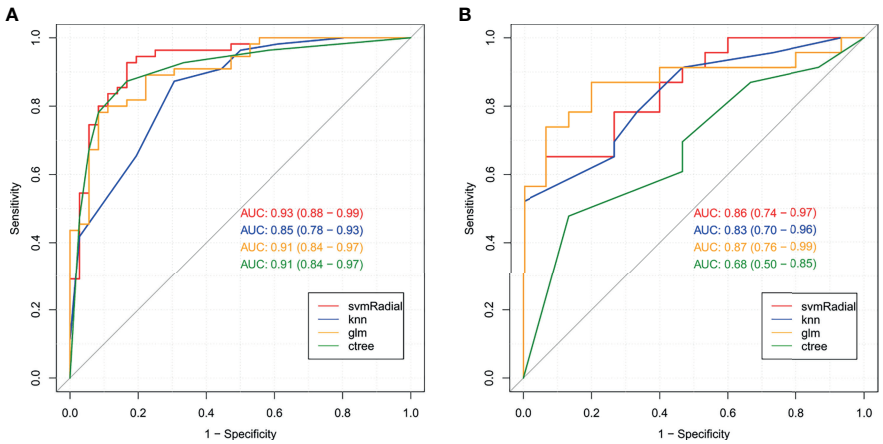


FIGURE 3 | Receiver operating characteristic (ROC) curves for the four classification machine learning models in the training set **(A)** and test set **(B)**.

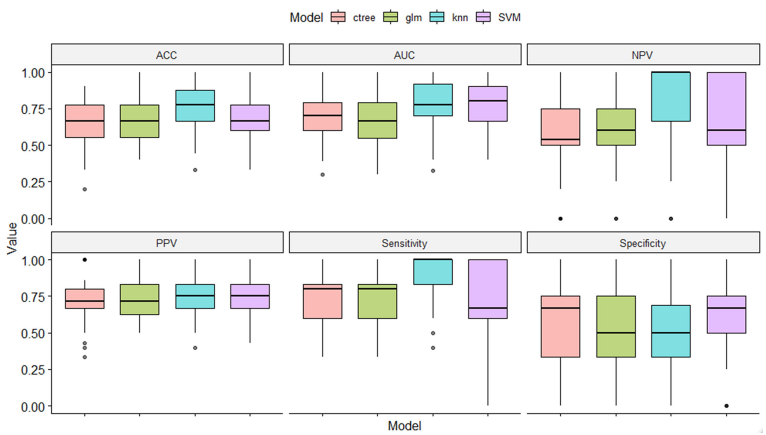


FIGURE 4 | Performance comparison of the four classification machine learning models in distinguishing benign from malignant masses.

for improving prediction accuracy (16) and is able to combine selected radiomic parameters for generating a radiomic signature (17, 18).

Additionally, the radiomics model had a higher ratio of features based on the MLO view compared with the CC view in this study, consistent with one previous study (19), suggesting that the MLO view might be more informative than its CC counterpart. Of course, combining both views provides more data as compared to each individual view (19). The radiomics features selected for modeling in this study included first-order, shape, and texture (including GLCM, GLSZM, and GLRLM) features. Texture features accounted for the largest proportion (7/13), and their correlation coefficients were relatively larger than other features as well. This also indicates that radiomics can reveal deep internal features.

Large Area Low Gray Level Emphasis (LALGLE) assesses the joint distribution of larger size zones showing lower gray-level values in a tumor image. In this study, both the original extraction technique and wavelet analysis were able to extract features from the oblique MLO and CC views. The feature weight was large, and two of three features had the highest magnitude of correlation coefficients (-0.556 and -0.48) in the feature set. Therefore, these features were negatively correlated with malignant status, which may be explained by the fact that most malignant masses have relatively dense cells and elevated density.

Size Zone Nonuniformity Normalized (SZNN) assesses size zone volume variability on a whole image, with reduced values suggesting elevated homogeneity among zone size volumes. The correlation coefficient here (0.334) was relatively large in the feature set, and positively correlated with malignant status, indicating high heterogeneity of malignant lesions. Furthermore, autocorrelation reflects the magnitude of texture fineness and coarseness, and in this study, this latter feature was positively correlated with malignant status, with a correlation coefficient of

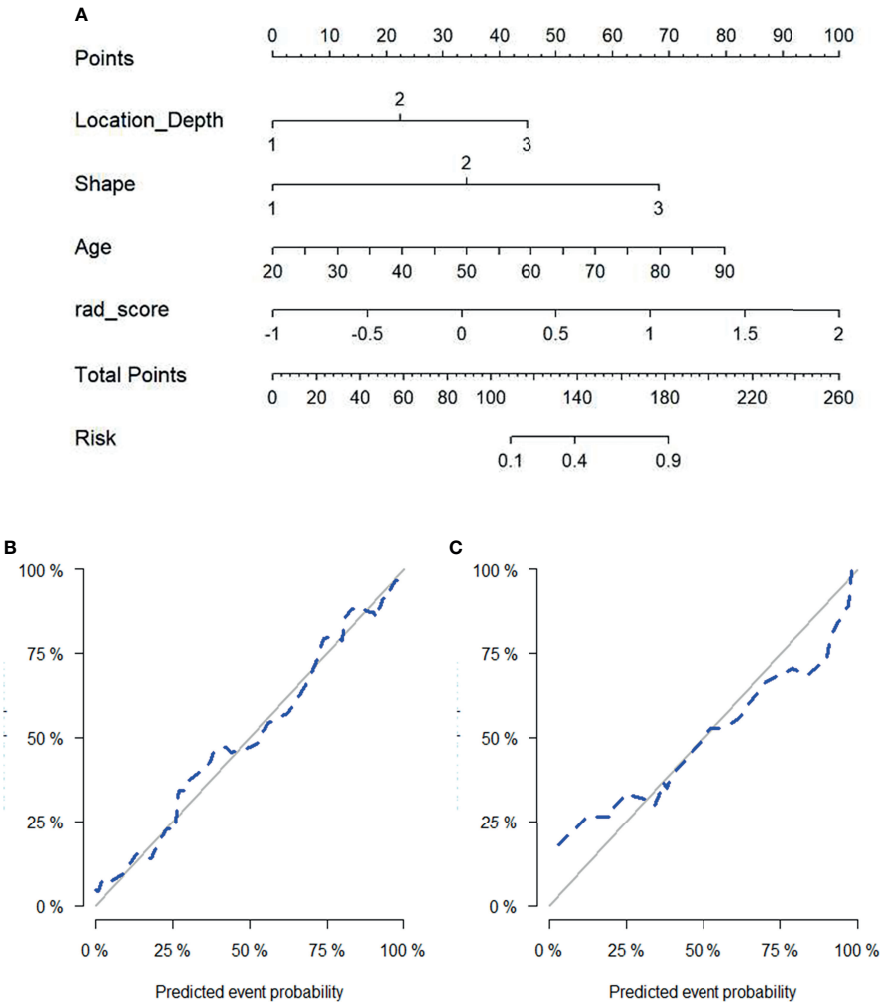


FIGURE 5 | Radiomics nomogram for predicting malignant status of round-like tumors **(A)**. Calibration curves of the radiomics nomogram in the training set **(B)** and test set **(C)**.

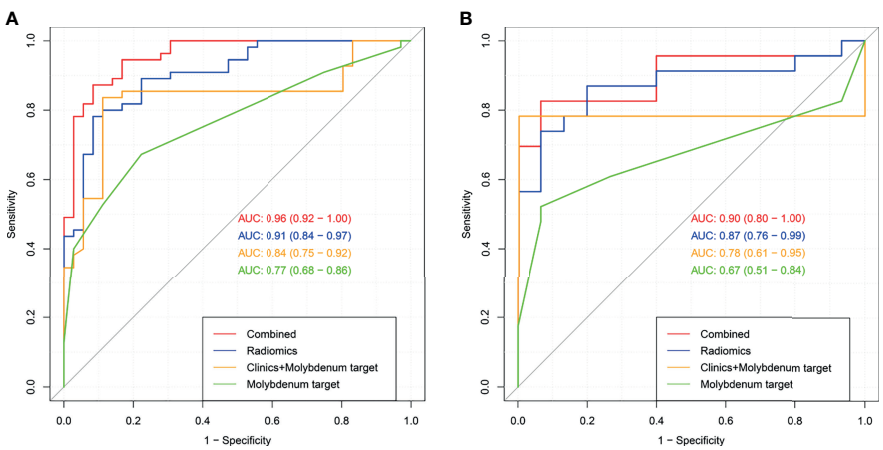


FIGURE 6 | Receiver operating characteristic (ROC) curves for digital mammography, the clinical model, the radiomics signature and the radiomics nomogram in the training set **(A)** and test set **(B)**.

TABLE 5 | Performances of the predictive models in distinguishing benign from malignant tumors.

	Model	Accuracy	Sensitivity	Specificity	PPV	NPV
Training	DM	0.670 (0.564-0.765)	0.879	0.552	0.527	0.889
	Clinics	0.791 (0.693-0.869)	0.909	0.681	0.727	0.888
	Radiomics	0.835 (0.743-0.905)	0.782	0.917	0.935	0.733
	Combined	0.890 (0.807-0.946)	0.941	0.825	0.873	0.917
Test	DM	0.684 (0.513-0.825)	0. 923	0.56	0.522	0.933
	Clinics	0.711 (0.541-0.846)	1	0.577	0.522	1
	Radiomics	0.789 (0.627-0.904)	0.696	0.933	0.941	0.667
	Combined	0.868 (0.719-0.956)	0.95	0.778	0.826	0.933

DM, digital mammography; PPV, positive predictive value; NPV, negative predictive value.

0.133. The texture of malignant masses is generally coarser than that of benign counterparts. Compared with most irregular malignant masses, round-like masses show relatively more uniform growth rates and finer texture, which may explain the lower correlation coefficient.

Five first-order features were also selected in this study, and most of them had low correlation coefficients. Only MLO_wavelet-HHL_firstorder_Mean had a high correlation coefficient (0.329), indicating that the feature was positively correlated with malignant status, which can be explained by the elevated density of malignant tumor cells. Comparing the tumor and contralateral breast gland density by univariate analysis of clinical characteristics, significant differences were found between benign and malignant masses as well.

Multiple studies have shown that radiomics can provide valuable information for clinical diagnostic and prognostic assessments (20–26), and previous researchers have already evaluated DM-derived radiomics for categorizing microcalcification (27), tumors (28), and breast cancer by molecular properties (19, 29). This study focused on the masses that are most difficult to assign to the malignant and benign groups by DM. Consistent with the literature (12), jointly applying DM and radiomics was able to increase overall diagnostic performance remarkably. Such a combination can be used to examine tumor heterogeneity more comprehensively and quantitatively when compared to morphological visual assessment alone.

However, this study is not without its limitations. First, diseased and normal tissues show no overt boundaries in DM, and ROIs were not automatically generated. Therefore, irregularities resulting from manual selection were inevitable. Second, since the sample sizes of cases with specific histopathological subtypes of breast cancer were small, their differential diagnoses by radiomics could not be performed. Further research is therefore needed to address this issue. Finally, our results require multicenter verification with large trials in order to generate more evidence for clinical application.

CONCLUSIONS

This study revealed that DM-based radiomics has good performance in distinguishing benign from malignant round-like masses, and the first-of-its-kind radiomics nomogram was developed and validated for such discrimination, achieving good accuracy. Indeed, DM-derived radiomics has an important clinical value in providing quantitative data to help clinicians read and interpret mammograms.

DATA AVAILABILITY STATEMENT

The datasets examined in this study are available from the corresponding authors upon reasonable request.

ETHICS STATEMENT

The Institutional Review Board of Minhang Hospital, Shanghai, China, approved the current retrospective trial and did not require signed informed consent.

AUTHOR CONTRIBUTIONS

LW, YD, BZ, and BS conceived and designed the study. WY, HW, JS, WL, JX, RW, WH, and YG performed the experiments and collected the data. LW and YD wrote the first draft of the manuscript. All authors read and approved the final manuscript.

FUNDING

The current research was financially supported by the Science and Technology Commission of Minhang District, Shanghai (grant number: 2019MHZ018).

ACKNOWLEDGMENTS

We are grateful to the staff of the Department of Radiology, Pathology, General Surgery (Minhang Hospital, Fudan University) and the Department of Radiology (Ruijin Hospital, Shanghai Jiao Tong University School of Medicine) for sound discussions and contributions to manuscript preparation. The authors thank AiMi Academic Services (www.aimieditor.com) for the English language editing and review services.

REFERENCES

1. Waks AG, Winer EP. Breast Cancer Treatment. *JAMA* (2019) 321(3):288. doi: 10.1001/jama.2018.19323
2. Gartlehner G, Thaler K, Chapman A, Kaminski-Hartenthaler A, Berzaczy D, Van Noord MG, et al. Mammography in Combination With Breast Ultrasonography Versus Mammography for Breast Cancer Screening in Women at Average Risk. *Cochrane Database Syst Rev* (2013) 2013(4): CD009632. doi: 10.1002/14651858.CD009632.pub2
3. Gøtzsche PC, Jørgensen KJ. Screening for Breast Cancer With Mammography (Review). *Cochrane Libr* (2013) 2013(6):81. doi: 10.1002/14651858.CD001877.pub5
4. Berry DA, Cronin KA, Plevritis SK, Fryback DG, Clarke L, Zelen M, et al. Effect of Screening and Adjuvant Therapy on Mortality From Breast Cancer. *N Engl J Med* (2005) 353(17):9. doi: 10.1056/NEJMoa050518
5. Welch HG, Prorok PC, O'Malley AJ, Kramer BS. Breast-Cancer Tumor Size, Overdiagnosis, and Mammography Screening Effectiveness. *N Engl J Med* (2016) 375(15):1438–47. doi: 10.1056/NEJMoa1600249
6. Tabár L, Vitak B, Chen TH, Yen AM, Cohen A, Tot T, et al. Swedish Two-County Trial: Impact of Mammographic Screening on Breast Cancer Mortality During 3 Decades. *Radiology* (2011) 260(3):658–63. doi: 10.1148/radiol.11110469
7. Oeffinger KC, Fontham ETH, Etzioni R, Herzig A, Michaelson JS, Shih YC, et al. Breast Cancer Screening for Women at Average Risk. *JAMA* (2015) 314 (15):1599. doi: 10.1001/jama.2015.12783
8. Yoo JL, Woo OH, Kim YK, Cho KR, Yong HS, Seo BK, et al. Can Mr Imaging Contribute in Characterizing Well-Circumscribed Breast Carcinomas? *Radiographics* (2010) 30(6):1689–702. doi: 10.1148/rg.306105511
9. Larribe M, Thomassin-Piana J, Jalaguier-Coudray A. Breast Cancers With Round Lumps: Correlations Between Imaging and Anatomopathology. *Diagn Interv Imaging* (2014) 95(1):37–46. doi: 10.1016/j.diii.2013.04.003
10. Thomassin-Naggara I, Tardivon A, Chopier J. Standardized Diagnosis and Reporting of Breast Cancer. *Diagn Interv Imaging* (2014) 95(7-8):759–66. doi: 10.1016/j.diii.2014.06.006
11. Turnbull LW. Dynamic Contrast-Enhanced MRI in the Diagnosis and Management of Breast Cancer. *NMR Biomed* (2009) 22(1):28–39. doi: 10.1002/nbm.1273
12. Li Z, Yu L, Wang X, Yu H, Gao Y, Ren Y, et al. Diagnostic Performance of Mammographic Texture Analysis in the Differential Diagnosis of Benign and Malignant Breast Tumors. *Clin Breast Cancer* (2018) 18(4):e621–7. doi: 10.1016/j.clbc.2017.11.004
13. Parekh V, Jacobs MA. Radiomics: A New Application From Established Techniques. *Expert Rev Precis Med Drug Dev* (2016) 1(2):207–26. doi: 10.1080/23808993.2016.1164013
14. Gillies RJ, Kinahan PE, Hricak H. Radiomics: Images Are More Than Pictures, They Are Data. *Radiology* (2016) 278(2):15. doi: 10.1148/radiol.2015151169
15. Koo TK, Li MY. A Guideline of Selecting and Reporting Intraclass Correlation Coefficients for Reliability Research. *J Chiropr Med* (2016) 15(2):155–63. doi: 10.1016/j.jcm.2016.02.012
16. Haris A, W D, Simon N. A Lasso for Hierarchical Interactions. *Ann Stat* (2013) 41(3):31. doi: 10.1214/13-AOS1096
17. Wang H, Song B, Ye N, Ren J, Sun X, Dai Z, et al. Machine Learning-Based Multiparametric MRI Radiomics for Predicting the Aggressiveness of Papillary Thyroid Carcinoma. *Eur J Radiol* (2020) 122:108755. doi: 10.1016/j.ejrad.2019.108755
18. Yin P, Mao N, Zhao C, Wu J, Sun C, Chen L, et al. Comparison of Radiomics Machine-Learning Classifiers and Feature Selection for Differentiation of Sacral Chordoma and Sacral Giant Cell Tumour Based on 3D Computed Tomography Features. *Eur Radiol* (2018) 29(4):1841–7. doi: 10.1007/s00330-018-5730-6
19. Ma W, Zhao Y, Ji Y, Guo X, Jian X, Liu P, et al. Breast Cancer Molecular Subtype Prediction by Mammographic Radiomic Features. *Acad Radiol* (2019) 26(2):196–201. doi: 10.1016/j.acra.2018.01.023
20. Caruso D, Zerunian M, Ciolina M, Santis DD, Rengo M, Soomro MH, et al. Haralick's Texture Features for the Prediction of Response to Therapy in Colorectal Cancer: A Preliminary Study. *La Radiol Med* (2017) 123(3):161–7. doi: 10.1007/s11547-017-0833-8
21. Ganeshan B, Goh V, Mandeville HC, Ng QS, Hoskin PJ, Miles KA. Non-Small Cell Lung Cancer: Histopathologic Correlates for Texture Parameters at CT. *Radiology* (2013) 266(1):11. doi: 10.1148/radiol.12112428
22. Lambin P, Rios-Velazquez E, Leijenaar R, Carvalho S, van Stiphout RG, Granton P, et al. Radiomics: Extracting More Information From Medical Images Using Advanced Feature Analysis. *Eur J Cancer* (2012) 48(4):441–6. doi: 10.1016/j.ejca.2011.11.036
23. Mannil M, Burgstaller JM, Held U, Farshad M, Roman G. Correlation of Texture Analysis of Paraspinal Musculature on MRI With Different Clinical Endpoints: Lumbar Stenosis Outcome Study (LSOS). *Eur Radiol* (2018) 29 (1):22–30. doi: 10.1007/s00330-018-5552-6
24. Li H, Zhu Y, Burnside ES, Drukker K, Hoadley KA, Fan C, et al. MR Imaging Radiomics Signatures for Predicting the Risk of Breast Cancer Recurrence as Given by Research Versions of MammaPrint, Oncotype DX, and PAM50 Gene Assays. *Radiology* (2016) 281(2):10. doi: 10.1148/radiol.2016152110
25. Ji GW, Zhang YD, Zhang H, Laghi A, Voena C. CT-Based Radiomics for Biliary Tract Cancer: A Possible Solution for Predicting Lymph Node Metastases. *Radiology* (2019) 290(1):2. doi: 10.1148/radiol.2018181408
26. Kniep HC, Madesta F, Schneider T, Hanning U, Schönfeld MH, Schön G, et al. Radiomics of Brain MRI: Utility in Prediction of Metastatic Tumor Type. *Radiology* (2019) 290(2):479–87. doi: 10.1148/radiol.2018180946
27. Lei C, Wei W, Liu Z, Xiong Q, Yang C, Yang M, et al. Mammography-Based Radiomic Analysis for Predicting Benign BI-RADS Category 4 Calcifications. *Eur J Radiol* (2019) 121:108711. doi: 10.1016/j.ejrad.2019.108711
28. Drukker K, Giger ML, Joe BN, Kerlikowske K, Greenwood H, Drukteinis JS, et al. Combined Benefit of Quantitative Three-Compartment Breast Image Analysis and Mammography Radiomics in the Classification of Breast Masses in a Clinical Data Set. *Radiology* (2019) 290(3):621–8. doi: 10.1148/radiol.2018180608
29. Zhou J, Tan H, Bai Y, Li J, Lu Q, Chen R, et al. Evaluating the HER-2 Status of Breast Cancer Using Mammography Radiomics Features. *Eur J Radiol* (2019) 121:108718. doi: 10.1016/j.ejrad.2019.108718

Conflict of Interest: Author YG was employed by General Electric (GE) Healthcare.

The remaining 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.

frontiers | Frontiers in Oncology

SYSTEMATIC REVIEW
published: 04 May 2022
doi: 10.3389/fonc.2022.893103

Radiomics for Predicting Response of Neoadjuvant Chemotherapy in Nasopharyngeal Carcinoma: A Systematic Review and Meta-Analysis

Chao Yang[1], Zekun Jiang[1], Tingting Cheng[2,3], Rongrong Zhou[3,4], Guangcan Wang[1], Di Jing[3,4], Linlin Bo[1], Pu Huang[1], Jianbo Wang[5], Daizhou Zhang[6], Jianwei Jiang[7], Xing Wang[8], Hua Lu[1], Zijian Zhang[3,4]* and Dengwang Li[1]**

[1] Shandong Key Laboratory of Medical Physics and Image Processing, Shandong Institute of Industrial Technology for Health Sciences and Precision Medicine, School of Physics and Electronics, Shandong Normal University, Jinan, China, [2] Department of General Practice, Xiangya Hospital, Central South University, Changsha, China, [3] National Clinical Research Center for Geriatric Disorders, Xiangya Hospital, Central South University, Changsha, China, [4] Department of Oncology, Xiangya Hospital, Central South University, Changsha, China, [5] Department of Radiation Oncology, Qilu Hospital, Cheeloo College of Medicine, Shandong University, Jinan, China, [6] Shandong Provincial Key Laboratory of Mucosal and Transdermal Drug Delivery Technologies, Shandong Academy of Pharmaceutical Sciences, Jinan, China, [7] Optical and Digital Image Processing Division, Qingdao NovelBeam Technology Co., Ltd., Qingdao, China, [8] Software Research and Development Center, Shangdong AccurDx Diagnosis of Biotech Co., Ltd., Jinan, China

Purpose: This study examined the methodological quality of radiomics to predict the effectiveness of neoadjuvant chemotherapy in nasopharyngeal carcinoma (NPC). We performed a meta-analysis of radiomics studies evaluating the bias risk and treatment response estimation.

Methods: Our study was conducted through a literature review as per the Preferred Reporting Items for Systematic Reviews and Meta-Analyses guidelines. We included radiomics-related papers, published prior to January 31, 2022, in our analysis to examine the effectiveness of neoadjuvant chemotherapy in NPC. The methodological quality was assessed using the radiomics quality score. The intra-class correlation coefficient (ICC) was employed to evaluate inter-reader reproducibility. The pooled area under the curve (AUC), pooled sensitivity, and pooled specificity were used to assess the ability of radiomics to predict response to neoadjuvant chemotherapy in NPC. Lastly, the Quality Assessment of Diagnostic Accuracy Studies technique was used to analyze the bias risk.

Results: A total of 12 studies were eligible for our systematic review, and 6 papers were included in our meta-analysis. The radiomics quality score was set from 7 to 21 (maximum score: 36). There was satisfactory ICC (ICC = 0.987, 95% CI: 0.957–0.996). The pooled sensitivity and specificity were 0.88 (95% CI: 0.71–0.95) and 0.82 (95% CI: 0.68–0.91), respectively. The overall AUC was 0.91 (95% CI: 0.88–0.93).

OPEN ACCESS

Edited by:
Xue Meng,
Shandong Cancer Hospital, China

Reviewed by:
Yue Qiu,
Sichuan University, China
Likun Wang,
Tianjin Medical University, China

****Correspondence:***
Dengwang Li
dengwang@sdnu.edu.cn
Zijian Zhang
wanzzj@csu.edu.cn
Hua Lu
hualu@sdnu.edu.cn

Specialty section:
This article was submitted to Cancer Imaging and Image-directed Interventions, a section of the journal Frontiers in Oncology

Received: *10 March 2022*
Accepted: *31 March 2022*
Published: *04 May 2022*

Citation:
Yang C, Jiang Z, Cheng T, Zhou R, Wang G, Jing D, Bo L, Huang P, Wang J, Zhang D, Jiang J, Wang X, Lu H, Zhang Z and Li D (2022) Radiomics for Predicting Response of Neoadjuvant Chemotherapy in Nasopharyngeal Carcinoma: A Systematic Review and Meta-Analysis. Front. Oncol. 12:893103. doi: 10.3389/fonc.2022.893103

Conclusion: Prediction response of neoadjuvant chemotherapy in NPC using machine learning and radiomics is beneficial in improving standardization and methodological quality before applying it to clinical practice.

Keywords: nasopharyngeal carcinoma, neoadjuvant chemotherapy, systematic review, meta-analysis, machine learning

INTRODUCTION

Nasopharyngeal carcinoma (NPC) is a malignant head and neck cancer that occurs in the nasopharyngeal space and can spread to the base of the skull and other organs (1–3). Its anatomical location is relatively hidden, causing nearly 70% of NPC patients to be diagnosed at a locally advanced stage (4–6). The pathological subtypes of nasopharyngeal tumors mostly include poorly differentiated or undifferentiated squamous cell carcinomas, which are more sensitive to chemoradiotherapy (7–9). Therefore, definitive concurrent chemoradiotherapy has become the standard of care for NPC patients with locally advanced diseases (10, 11). However, the efficacy of neoadjuvant chemotherapy has been shown to vary greatly in clinical practice, and approximately 30% of patients will develop chemoradiotherapy-related adverse events (12–15).

Radiomics is a highly efficient extraction feature that obtains massive amounts of data from medical images. It transforms imaging data into a high-resolution mineable data space using automated or semiautomated analysis methods (16–18). Given its precise and systematic nature, radiomics can retrieve data that enable the detection of minimal lesions and the prediction of treatment outcomes (19–24). As a result, radiomics is widely used in the study of NPC, and there is huge interest in employing radiomics to predict neoadjuvant chemotherapy efficacy in NPC patients. This information can assist physicians in selecting an optimal scheme and in achieving the maximal anticancer effect. Nevertheless, radiologic data analysis is highly reliant on the subjective interpretation of skilled radiologists. The quantitative data and autonomous imaging markers can serve as an adjunct to expert clinical opinion, thus increasing the prognostic precision (25, 26).

The purpose of this research was to evaluate the methodological quality and analyze the effectiveness of neoadjuvant chemotherapy in NPC among the published radiomics papers. We also performed a meta-analysis of relevant studies to predict the treatment response of neoadjuvant chemotherapy, using the radiomics method, in NPC.

Abbreviations: AUC, area under curve; CR, complete response; ICC, intra-class correlation coefficient; NPC, nasopharyngeal carcinoma; OR, odds ratio; PD, progressive disease; PR, partial response; PRISMA, Preferred Reporting Items for Systematic Reviews and Meta-Analyses; PRISMA-DTA, Preferred Reporting Items for Systematic Reviews and Meta-Analyses for Diagnostic Test Accuracy; QUADAS-2, Quality Assessment of Diagnostic Accuracy Studies; RECIST, Response Evaluation Criteria in Solid Tumors; REML, restricted maximum likelihood; RQS, radiomics quality score; SD, stable disease; SROC, summary receiver operating characteristic.

MATERIALS AND METHODS

Study Protocol and Literature Search

This study was conducted as per the Preferred Reporting Items for Systematic Reviews and Meta-Analyses for Diagnostic Test Accuracy (PRISMA-DTA) guidelines (27). Four databases (Web of Science, PubMed, Embase, and Cochrane Library) were screened to select relevant articles published prior to January 31, 2022. The search terms included were as follows: (Nasopharyngeal Carcinoma OR Carcinomas, Nasopharyngeal) AND (Machine Learning OR Artificial Intelligence OR radiomics) AND (CT OR MRI OR Magnetic Resonance Imaging). Please refer to the **Supplementary Material** for more details on the medical subject heading (MESH terms).

Two independent researchers screened article titles and abstracts to determine inclusion in this study. Case reports, non-original publications, and research on topics of interest other than the effectiveness of neoadjuvant chemotherapy in NPC were excluded. To further evaluate relevant articles, the full texts of articles were retrieved and read to determine eligibility for analysis. The reference list of included papers was also reviewed for potential eligible inclusion. The types of images included in our study included MRI, CT, and PET.

Data Collection

The main endpoints were extracted and adjusted to the largest area under the curve (AUC) in the verification dataset and also prioritized external validation datasets. Among the articles with no external verification dataset, the internal verification dataset (i.e., the test set) was employed. In the absence of an internal verification dataset, the validation set from the training dataset (e.g., leave-one-cross-validation, fivefold cross-validation, and tenfold cross-validation) was employed. The collected models contained radiologically relevant characteristics and sometimes contained characteristics, such as clinical information, pathological types, radiotherapy dose, region of interest (ROI), and imaging features extracted.

Study Evaluation

The radiomics quality score (RQS) assessed the methodological quality of eligible publications, and the Quality Assessment of Diagnostic Accuracy Studies (QUADAS-2) determined the bias risk (28–30).

RQS assesses an investigation's methodological quality by examining protocols, images and segmentation reproducibility, feature reduction and verification, biological verification, clinical application, and model performance, with enhanced evidence and open science (28). The detailed RQS report is provided in the **Supplementary Material**. Overall, 16 items were included in the

RQS, with scores ranging from −8 to 36. The RQS scores were then converted to percentages, whereby −8 to 0 was 0% and 36 was 100% (28). Two experienced physicians independently scored the RQS of eligible articles.

QUADAS-2 evaluates the bias risk in varying domains ("Patient Selection", "Index Test", "Reference Standard", and "Flow and Timing") and can be customized to a particular study question. The bias risk for each included study was determined by the QUADAS item of Review Manager 5.4 in order to determine the quality of diagnostic articles (31).

Meta-Analysis

A meta-analysis of investigations related to the prediction of the treatment response of neoadjuvant chemotherapy in NPC patients was further performed. The data were retrieved by 2 independent reviewers. The internal validity was assessed by a third reviewer. Only studies that provided a two-by-two contingency table or enough data to reconstruct such a table were eligible for analysis. In cases where multiple models were presented, only models with the largest AUC were chosen in our analysis.

Statistical Analysis

Random-effects meta-analyses were conducted with the restricted maximum likelihood (REML) and presented as a log odds ratio (OR). The threshold effect was determined by calculating the sensitivity and specificity of Spearman's correlation coefficients. Forest plots and summary receiver operating characteristic (SROC) curve were generated. The pooled AUC, sensitivity, and specificity were used to assess the ability of radiomics to predict the treatment response of neoadjuvant chemotherapy in NPC patients. A funnel plot assessed publication bias. Cochran's Q test and I^2 score evaluated heterogeneity among eligible studies. An I^2 value of 0%–25% meant unremarkable heterogeneity, 25%–50% meant reduced heterogeneity, 50%–75% meant moderate heterogeneity, and >75% meant high heterogeneity (32).

R (version 4.1.2, https://cran.r-project.org/), IBM SPSS Statistics (version 24; IBM Corporation, Armonk, NY, USA), Stata (version 16.0, https://www.stata.com/), and Review Manager (version 5.4) were employed for statistical analyses.

RESULTS

Literature Search

We initially identified 613 relevant articles; 317 articles were considered for careful evaluation after the elimination of duplicate publications. Upon screening of the titles and abstracts, 18 relevant articles were extracted for further analysis. Four articles that did not contain a radiomics-based model and two conference abstracts were excluded from the analysis. A total of twelve articles that used radiomics-based prediction models were selected for the final systematic review (33–44). Five of the articles examined survival analysis, and seven articles examined the prediction of treatment response. One article that predicted treatment response did not provide enough information to reconstruct a contingency table and calculate the overall outcome (43). Therefore, six articles were included in our meta-analysis. Our PRISMA flowchart is presented in **Figure 1**. The detailed information on all eligible publications is provided in **Table 1**. We summarized detailed information about the selected articles, such as institution, study duration, and type of radiomics features used. The detailed summary table is available in **Tables S4** and **S5**.

Evaluation Criteria for Neoadjuvant Chemotherapy

The response evaluation of neoadjuvant chemotherapy in all included studies was based on the Response Evaluation Criteria in Solid Tumors 1.1 (RECIST 1.1) (45). Complete response (CR) and partial response (PR) were defined as response to treatment, while stable disease (SD) and progressive disease (PD) were defined as no response to treatment.

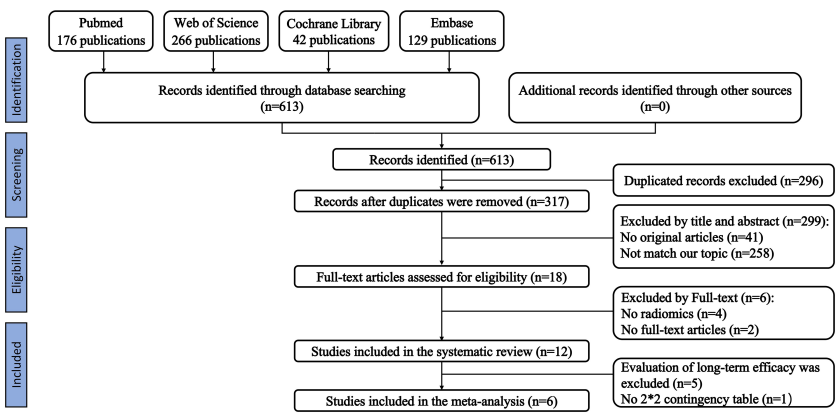


FIGURE 1 | A schematic of the publication selection process.

TABLE 1 | Details of eligible studies.

Author nation, year	Study Type	Cancer	ROI	Imaging	Training set	Test set	External Validation
Piao China, 2021 (33)	Retrospective observational	NPC	GTVnx	MRI	108	0	0
Wang China, 2018 (34)	Retrospective observational	NPC	GTVnx	MRI	120	0	0
Zhang China, 2020 (35)	Retrospective observational	NPC	GTVnx	MRI	81	34	0
Zhang China, 2020 (36)	Retrospective observational	NPC	GTVnx	MRI	169	19	45
Chen China, 2021 (37)	Retrospective observational	NPC	GTVnx GTVnd	MRI	847	400	396
Zhao China, 2020 (38)	Retrospective observational	NPC	GTVnx	MRI	100	23	0
Peng China, 2019 (39)	Retrospective observational	NPC	GTVnx GTVnd	PET/CT	470	237	0
Zhong China, 2020 (40)	Retrospective observational	NPC	GTVnx	MRI	447	191	0
Dong China, 2019 (41)	Retrospective observational	NPC	GTVnx	MRI	254	248	0
Yang China, 2022 (42)	Retrospective observational	NPC	GTVnx	CT	208	89	0
Hu China, 2021 (43)	Retrospective observational	NPC	GTVnx GTVnd CTV PTV	MRI	200	84	0
Liao China, 2021 (44)	Retrospective observational	NPC	GTVnx	MRI	200	86	0

NPC, nasopharyngeal carcinoma; GTVnx, nasopharynx gross tumor volume; GTVnd, lymph node gross tumor volume; CTV, clinical target volume; PTV, planning target volume; MRI, Magnetic Resonance Imaging; CT, Computed Tomography; PET, Positron Emission Tomography.

TABLE 2 | RQS elements, as reported by Lambin et al. (28), and the mean rating of our eligible studies.

RQS scoring item	Interpretation	Average
Image protocol quality	+1 for well-documented protocols, +1 for publicly available protocols	1.25
Multiple segmentations	+1 if segmented multiple times (different physicians, algorithms, or perturbation of regions of interest)	0.92
Phantom study on all scanners	+1 if texture phantoms were used for feature robustness assessment	0
Imaging at multiple time points	+1 if multiple time points for feature robustness assessment	0
Feature reduction or adjustment for multiple testing	−3 if nothing, +3 if either feature reduction or correction for multiple testing	3
Multivariable analysis with non-radiomics feature	+1 if multivariable analysis with non-radiomics features	0.67
Detect and discuss biological correlates	+1 if present	0.33
Cutoff analyses	+1 if cutoff either predefined or at median or continuous risk variable reported	0.71
Discrimination statistics	+1 for discrimination statistic and statistical significance, +1 if resampling applied	1.75
Calibration statistic	+1 for calibration statistic and statistical significance, +1 if resampling applied	1.17
Prospective study registered in a trial database	+7 for prospective validation within a registered study	0
Validation	−5 if validation is missing, +2 if validation is based on a dataset from the same institute, +3 if validation is based on a dataset from another institute, +4 if validation is based on two datasets from two distinct institutes, +4 if the study validates a previously published signature, +5 if validation is based on three or more datasets from distinct institutes	1.83
Comparison to "gold standard"	+2 for comparison to gold standard	1.83
Potential clinical utility	+2 for reporting potential clinical utility	1.5
Cost-effectiveness analysis	+1 for cost-effectiveness analysis	0
Open science and data	+1 if scans are open source, +1 if region of interest segmentations are open source, +1 if code is open source, +1 if radiomics features are calculated on a set of representative ROIs and the calculated features and representative ROIs are open sources	2.04
Total score (maximum score: 36 points)		17

Study Evaluation

The RQS scores, ranging from 7 to 21 (maximum score: 36), are summarized in **Table 2**. The publication with the highest RQS percentage was 58.3%. The intra-class correlation coefficient (ICC) between independent reviewers who assessed the publications was 0.987 (95% CI: 0.957–0.996, $p < 0.001$), which showed excellent reproducibility among readers. The RQS scores examined by the two readers are presented in the **Supplementary Material**. Elevated intra-class association represented the high reliability of quality assessment. Lastly, reviewers reassessed any disagreements until a consensus was reached.

The bias risk, as assessed by QUADAS-2, is presented in **Figure 2**. The publications with high, unclear, or low bias risk in the four domains of patient selection, index test, reference standard, and flow and timing were 0, 4, and 2, respectively. Particularly, three publications failed to present a clear report of the patient selection process. Therefore, they received an unclear bias risk in the patient selection domain (34, 38, 42). One study received an unclear bias risk in the index test domain (33). Three studies received an unclear bias risk in the flow and timing domain (33, 34, 38). All studies in the meta-analysis displayed relatively reduced concerns regarding applicability in the three domains (patient selection, index test, and reference standard).

Meta-Analysis

Seven, out of twelve, selected systematic studies discussed the use of radiomics in predicting the treatment response of neoadjuvant chemotherapy. Only six studies provided sufficient data to allow the reconstruction of a contingency table to compute an overall outcome. Hence, only six studies were included in the meta-analysis.

Spearman's correlation analysis revealed no threshold effect ($\rho = 0.486$, $p = 0.3556$). The SROC curve, pooled AUC, pooled sensitivity, and pooled specificity were used to assess the ability of radiomics to predict the response of neoadjuvant chemotherapy in NPC patients. Based on our data analysis, the pooled sensitivity and specificity were 0.88 (95% CI: 0.71–0.95) and 0.82 (95% CI: 0.68–0.91), respectively, as evidenced by the corresponding forest plots in **Figure 3**. The pooled AUC was 0.91 (95% CI: 0.88–0.93). Cochran's Q was 29.16 ($p < 0.01$), and the I^2 score was 85.8%, which represented a high level of heterogeneity within eligible studies with statistically significant heterogeneity. **Figure 4** depicts the forest plot of the treatment outcome, computed as log OR. The log OR of the radiomics model predicting the neoadjuvant chemotherapy treatment response in NPC patients was 0.31 (95% CI: -1.58–2.21). The SROC curve is provided in **Figure 5**. The funnel plot correlating the outcome to standard error is presented in **Figure 6**. Given that we had less than 10 eligible articles in our meta-analysis, Egger's test was not applicable, as suggested by the Cochrane guidelines (46).

DISCUSSION

Radiomics has excellent prospects in multiple applications and can potentially aid in retrieving more quantitative data from standard medical images (47). In recent years, radiomics has developed rapidly in NPC research. However, despite ongoing

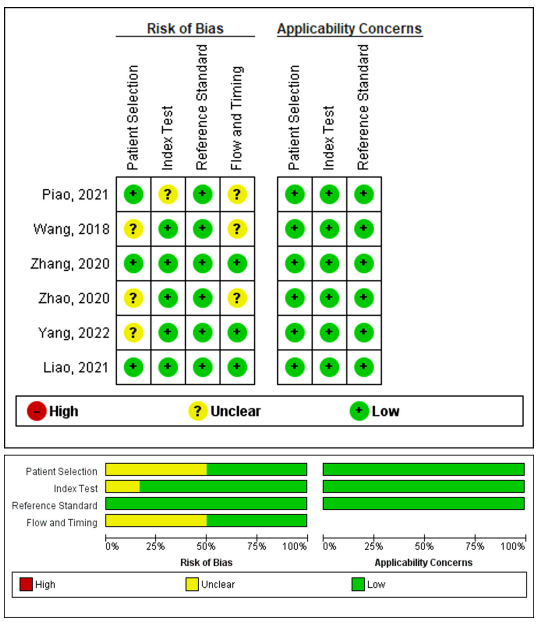


FIGURE 2 | Assessment of the methodological quality of publications included in the meta-analysis, based on the bias risk and applicability using the QUADAS-2 tool. Green, yellow, and red circles denote low, unclear, and high bias risks, respectively.

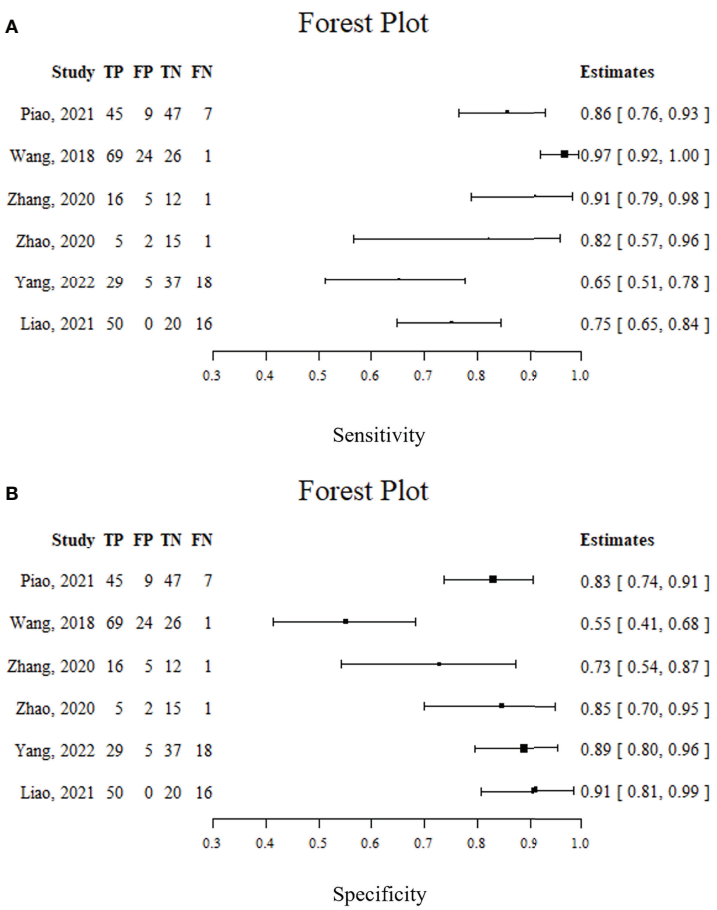


FIGURE 3 | Forest plots. **(A)** sensitivity; **(B)** specificity.

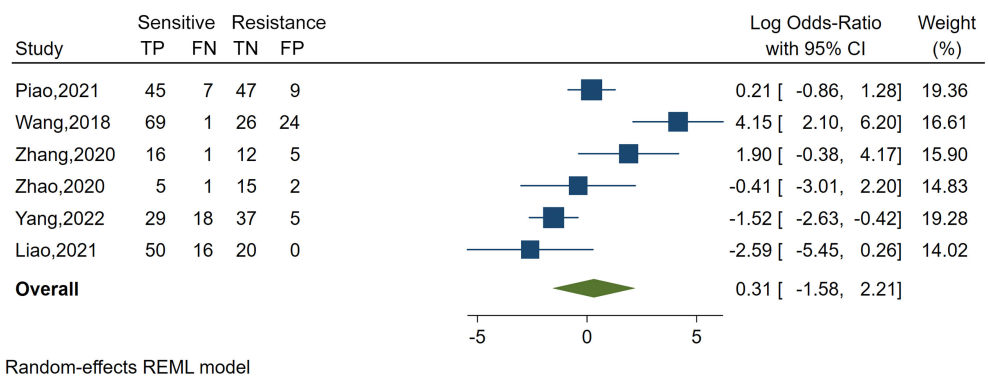


FIGURE 4 | Forest plot of the study outcome, as evidenced by the log odds ratio of six included meta-analysis studies examining the radiomics accuracy in predicting the treatment response to neoadjuvant chemotherapy in treating nasopharyngeal carcinoma. TP, number of patients correctly predicted in the sensitive group; FN, number of patients incorrectly predicted in the resistance group; FP, number of patients incorrectly predicted in the sensitive group; TN, number of patients correctly predicted in the resistance group; x-axis, log-transformed odds ratios; REML, restricted maximum likelihood.

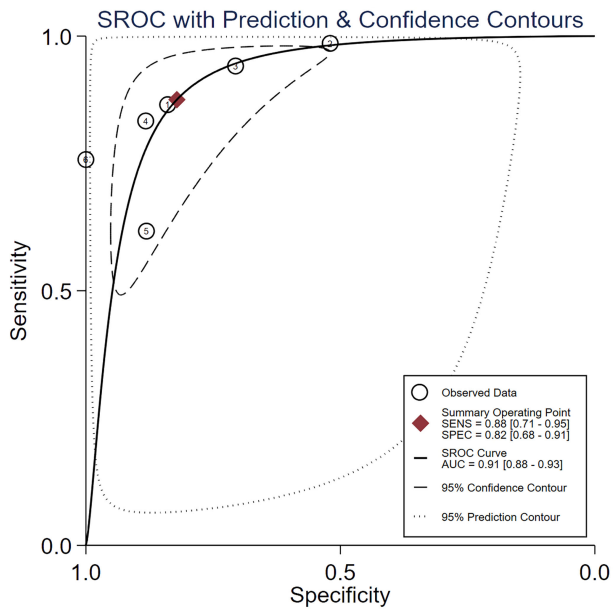


FIGURE 5 | The summary receiver operating characteristic (SROC) curve.

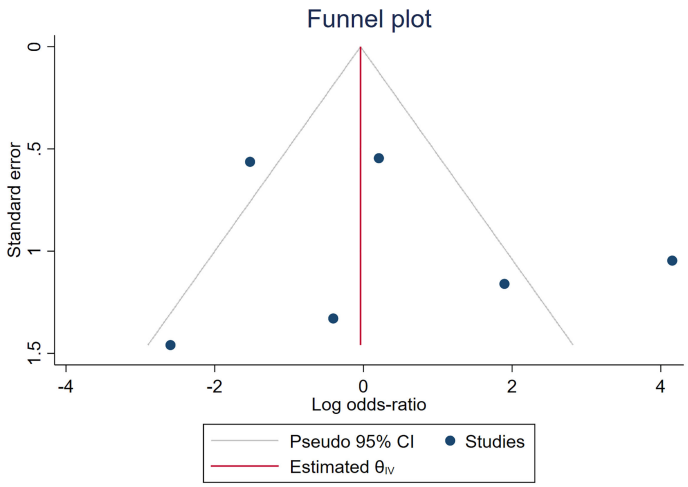


FIGURE 6 | A funnel plot of meta-analyzed studies.

efforts to standardize radiomics extraction features and analysis, their usage outside research is not yet justified (48).

We found several articles based on radiomics to predict response to neoadjuvant chemotherapy in NPC patients, mostly over the last 3 years. This suggested that the use of radiomics in neoadjuvant chemotherapy for NPC patients is novel and remains groundbreaking. Based on our analysis, the characteristics of radiomics investigations were similar among all eligible publications. First, the ROI was manually segmented by two radiologists. Second, the radiomics features were extracted, and relevant features were selected. Third, a model predicting neoadjuvant chemotherapy effectiveness in treating NPC was constructed and evaluated. The texture features were deemed as the most frequent type of radiomics features in the twelve selected articles, and the detailed information is presented in **Table S4**.

Ten of the twelve articles employed texture features in their highest AUC models. The wavelet features were deemed as a frequent occurrence, and others included first-order features and shape features. During the prediction of neoadjuvant

chemotherapy efficacy in NPC, radiomics features including texture, wavelet, first order, and other features extracted from images by artificial intelligence algorithms were able to show a lot of hidden information. With an increasing number of radiomics studies, several studies also revealed that textural features could provide additional predictive information (49–53). This systematic analysis found that the Gray Level Run Length Matrix features, the Gray Level Size Zone Matrix features, and the Gray Level Co-occurrence Matrix features are more frequently used. The textural features were shown to provide good results in predicting the efficacy of neoadjuvant chemotherapy treatment for NPC. One possible reason is that texture features contain information related to the efficacy of neoadjuvant chemotherapy treatment.

The advent of radiomics has made great contributions to overcoming limitations of user-dependent interpretation, thus assisting physicians in solving clinical problems. However, it was undeniable that the quality of our current research on radiomics is uneven. RQS is a common method for assessing the quality of radiomics studies and has been shown to accurately evaluate the methodological quality of radiomics studies. This is essential for the critical appraisal of a massive amount of research articles and prioritizing the verification of high-quality data. Since the first RQS application produced certain variations in inter-rater agreements (54), our independent RQS scoring was conducted by two independent readers experienced in radiomics. This way, we achieved a good level of agreement in terms of overall rating (ICC=0.987) and all scoring elements. The ICC, corresponding to each score category, was greater than 0.75. The RQS of our eligible studies were between 7 and 21 points, with a maximum of 36. But all eligible studies were retrospective in nature; therefore, 7 points was lost. We recommend future prospective studies to obtain higher-quality evidence. Moreover, none of the studies we analyzed conducted a cost–benefit analysis, and no phantom investigations were performed in terms of scan images. These deficiencies in research should be resolved in future radiomics research.

Our meta-analysis examined the prediction accuracy of neoadjuvant chemotherapy efficacy in NPC patients, based on radiomics. The SROC curve, obtained from the meta-analysis, is a ROC curve drawn from the OR of different radiomics studies. We demonstrated an enhanced prediction with a pooled AUC of 0.91. Our pooled sensitivity and pooled specificity reached 0.88 and 0.82, respectively. In terms of the publications that were eligible for meta-analysis, our QUADAS-2 assessment revealed a reduced bias risk while highlighting some critical matters. Particularly, three articles exhibited incomplete reporting of the inclusion–exclusion criteria, which can inadvertently introduce bias in the patient selection process (34, 38, 42). Moreover, one study received an unclear bias risk in the index test domain (33), due to the low number of features analyzed to the point of potential bias. In addition, three studies received an unclear bias risk in the flow and timing domain. Among them, one study failed to report the neoadjuvant chemotherapy duration (38). The remaining two studies showed less standardized processes (33, 34), and neither study employed a test set to validate the radiomics model. One study (33) employed a leave-one-out cross-validation method to evaluate the model, and another (34) used the bootstrap-validated model. Although internal validation in the training set can evaluate the performance of the radiomics model, this validation method may have introduced bias. All these concerns are sources of possible bias and should be clearly stated to eliminate bias.

The limitations of our work include the following. First, all studies were retrospective, and no prospective radiomics studies were found. Second, the radiomics features may have been affected by imaging technology. In the future, multicenter prospective investigations should be conducted to fully examine the predictability of radiomics studies (55). Third, RQS is a purely methodological scoring system that does not account for alterations in the study aim. Fourth, our sample size was relatively low, and the included studies were all from China. Fifth, although the QUADAS-2 assessment provided some unclear bias risks, no high bias sources were found. Moreover, being a qualitative score, the QUADAS-2 interpretation is not easily interpretable. Given our small sample size, our publication bias assessment is open to question. Sixth, we noted a high study heterogeneity, but this is typically common among machine learning meta-analyses and diagnostic meta-analyses (56–59).

CONCLUSION

Radiomics studies investigating the efficacy of neoadjuvant chemotherapy in NPC patients demonstrated promising results. We, therefore, recommend properly designed prospective trials in the future, including the validation and standardization of methodological data analysis.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/**Supplementary Material**. Further inquiries can be directed to the corresponding authors.

AUTHOR CONTRIBUTIONS

Conceptualization: CY, ZJ, TC, RZ, JW, DZ; Data collection: CY, GW, ZJ, PH; Data analysis: CY, ZJ, ZZ, LB; Data curation: GW, CY, JJ, XW; Writing-original draft preparation: CY; Writing-review and editing: ZJ, DJ, ZZ, HL; Supervision: DL, ZZ and HL. All authors have read and agreed to the published version of the manuscript.

FUNDING

This work was funded by the National Natural Science Foundation of China (61971271), the Taishan Scholars Project of Shandong Province (Tsqn20161023), the Jinan City-School Integration Development Strategy Project

(JNSX2021023), the Natural Science Foundation of Shandong Province (ZR2019PF011), the Natural Science Foundation of Hunan Province, China (S2021JJKWLH0218), and 2020 Hunan Provincial clinical medical technology innovation guidance project (S2020SFTLJS0217).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fonc.2022.893103/full#supplementary-material

REFERENCES

1. Bray F, Ferlay J, Soerjomataram I, Siegel R, Torre L, Jemal AJCCJC. Erratum: Global Cancer Statistics 2018: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. *CA Cancer J Clin* (2020) 70(4):313. doi: 10.3322/caac.21609
2. Chen W, Zheng R, Baade PD, Zhang S, Zeng H, Bray F, et al. Cancer Statistics in China, 2015. *CA Cancer J Clin* (2016) 66(2):115–32. doi: 10.3322/caac.21338
3. Wei KR, Zheng RS, Zhang SW, Liang ZH, Li ZM, Chen WQ. Nasopharyngeal Carcinoma Incidence and Mortality in China, 2013. *Chin J Cancer* (2017) 36 (1):90. doi: 10.1186/s40880-017-0257-9
4. Pan JJ, Ng WT, Zong JF, Chan LL, O'Sullivan B, Lin SJ, et al. Proposal for the 8th Edition of the AJCC/UICC Staging System for Nasopharyngeal Cancer in the Era of Intensity-Modulated Radiotherapy. *Cancer* (2016) 122(4):546–58. doi: 10.1002/cncr.29795
5. Mao YP, Xie FY, Liu LZ, Sun Y, Li L, Tang LL, et al. Re-Evaluation of 6th Edition of Ajcc Staging System for Nasopharyngeal Carcinoma and Proposed Improvement Based on Magnetic Resonance Imaging. *Int J Radiat Oncol* (2009) 73(5):1326–34. doi: 10.1016/j.ijrobp.2008.07.062
6. Tang LL, Chen WQ, Xue WQ, He YQ, Zheng RS, Zeng YX, et al. Global Trends in Incidence and Mortality of Nasopharyngeal Carcinoma. *Cancer Lett* (2016) 374(1):22–30. doi: 10.1016/j.canlet.2016.01.040
7. Wei WI, Sham JST. Nasopharyngeal Carcinoma. *Lancet* (2005) 365 (9476):2041–54. doi: 10.1016/s0140-6736(05)66698-6
8. Micheau C, Rilke F, Pilotti S. Proposal for a New Histopathological Classification of the Carcinomas of the Nasopharynx. *Tumori* (1978) 64 (5):513–8. doi: 10.1177/030089167806400509
9. Marks JE, Phillips JL, Menck HR. The National Cancer Data Base Report on the Relationship of Race and National Origin to the Histology of Nasopharyngeal Carcinoma. *Cancer* (1998) 83(3):582–8. doi: 10.1002/(sici)1097-0142(19980801)83:3<582::aid-cncr29>3.0.co;2-r
10. Tang LL, Chen YP, Mao YP, Wang ZX, Guo R, Chen L, et al. Validation of the 8th Edition of the UICC/AJCC Staging System for Nasopharyngeal Carcinoma From Endemic Areas in the Intensity-Modulated Radiotherapy Era. *J Natl Compr Canc Ne* (2017) 15(7):913–9. doi: 10.6004/jnccn.2017.0121
11. Qiu WZ, Huang PY, Shi JL, Xia HQ, Zhao C, Cao KJ. Neoadjuvant Chemotherapy Plus Intensity-Modulated Radiotherapy Versus Concurrent Chemoradiotherapy Plus Adjuvant Chemotherapy for the Treatment of Locoregionally Advanced Nasopharyngeal Carcinoma: A Retrospective Controlled Study. *Chin J Cancer* (2016) 35:2. doi: 10.1186/s40880-015-0076-9
12. Liu SL, Sun XS, Yan JJ, Chen QY, Lin HX, Wena YF, et al. Optimal Cumulative Cisplatin Dose in Nasopharyngeal Carcinoma Patients Based on Induction Chemotherapy Response. *Radiother Oncol* (2019) 137:83–94. doi: 10.1016/j.radonc.2019.04.020
13. Lan XW, Xiao Y, Zou XB, Zhang XM, OuYang PY, Xie FY. Outcomes of Adding Induction Chemotherapy to Concurrent Chemoradiotherapy for Stage T3N0-1 Nasopharyngeal Carcinoma: A Propensity-Matched Study. *Onco Targets Ther* (2017) 10:3853–60. doi: 10.2147/OTT.S133917
14. Yen RF, Chen THH, Ting LL, Tzen KY, Pan MH, Hong RL. Early Restaging Whole-Body(18)F-FDG PET During Induction Chemotherapy Predicts Clinical Outcome in Patients With Locoregionally Advanced Nasopharyngeal Carcinoma. *Eur J Nucl Med Mol I* (2005) 32(10):1152–9. doi: 10.1007/s00259-005-1837-5
15. Peng H, Chen L, Zhang Y, Li WF, Mao YP, Liu X, et al. The Tumour Response to Induction Chemotherapy has Prognostic Value for Long-Term Survival Outcomes After Intensity-Modulated Radiation Therapy in Nasopharyngeal Carcinoma. *Sci Rep-Uk* (2016) 6:24835. doi: 10.1038/srep24835
16. Lambin P, Rios-Velazquez E, Leijenaar R, Carvalho S, van Stiphout RG, Granton P, et al. Radiomics: Extracting More Information From Medical Images Using Advanced Feature Analysis. *Eur J Cancer* (2012) 48(4):441–6. doi: 10.1016/j.ejca.2011.11.036
17. Kumar V, Gu YH, Basu S, Berglund A, Eschrich SA, Schabath MB, et al. Radiomics: The Process and the Challenges. *Magn Reson Imaging* (2012) 30 (9):1234–48. doi: 10.1016/j.mri.2012.06.010
18. Mayerhoefer ME, Materka A, Langs G, Haggstrom I, Szczypinski P, Gibbs P, et al. Introduction to Radiomics. *J Nucl Med* (2020) 61(4):488–95. doi: 10.2967/jnumed.118.222893
19. Castellano G, Bonilha L, Li LM, Cendes F. Texture Analysis of Medical Images. *Clin Radiol* (2004) 59(12):1061–9. doi: 10.1016/j.crad.2004.07.008
20. Gillies RJ, Kinahan PE, Hricak H. Radiomics: Images Are More Than Pictures, They Are Data. *Radiology* (2016) 278(2):563–77. doi: 10.1148/radiol.2015151169
21. Jiang ZK, Dong YJ, Yang LK, Lv YH, Dong S, Yuan SH, et al. CT-Based Hand-Crafted Radiomic Signatures Can Predict PD-L1 Expression Levels in Non-Small Cell Lung Cancer: A Two-Center Study. *J Digit Imaging* (2021) 34 (5):1073–85. doi: 10.1007/s10278-021-00484-9
22. Yang FC, Zhang JY, Zhou L, Xia W, Zhang R, Wei HF, et al. CT-Based Radiomics Signatures can Predict the Tumor Response of Non-Small Cell Lung Cancer Patients Treated With First-Line Chemotherapy and Targeted Therapy. *Eur Radiol* (2022) 32(3):1538–47. doi: 10.1007/s00330-021-08277-y
23. Giannini V, Mazzetti S, Bertotto I, Chiarenza C, Cauda S, Delmastro E, et al. Predicting Locally Advanced Rectal Cancer Response to Neoadjuvant Therapy With (18)F-FDG PET and MRI Radiomics Features. *Eur J Nucl Med Mol Imaging* (2019) 46(4):878–88. doi: 10.1007/s00259-018-4250-6
24. Zhuang Z, Liu Z, Li J, Wang X, Xie P, Xiong F, et al. Radiomic Signature of the FOWARC Trial Predicts Pathological Response to Neoadjuvant Treatment in Rectal Cancer. *J Transl Med* (2021) 19(1):256. doi: 10.1186/s12967-021-02919-x
25. Sullivan DC, Obuchowski NA, Kessler LG, Raunig DL, Gatsonis C, Huang EP, et al. Metrology Standards for Quantitative Imaging Biomarkers. *Radiology* (2015) 277(3):813–25. doi: 10.1148/radiol.2015142202
26. Raunig DL, McShane LM, Pennello G, Gatsonis C, Carson PL, Voyvodic JT, et al. Quantitative Imaging Biomarkers: A Review of Statistical Methods for Technical Performance Assessment. *Stat Methods Med Res* (2015) 24(1):27–67. doi: 10.1177/0962280214537344
27. McInnes MDF, Moher D, Thombs BD. Preferred Reporting Items for a Systematic Review and Meta-Analysis of Diagnostic Test Accuracy Studies: The PRISMA-DTA Statement (Vol 319, Pg 388, 2018). *Jama-J Am Med Assoc* (2019) 322(20):2026–. doi: 10.1001/jama.2019.18307
28. Lambin P, Leijenaar RTH, Deist TM, Peerlings J, de Jong EEC, van Timmeren J, et al. Radiomics: The Bridge Between Medical Imaging and Personalized Medicine. *Nat Rev Clin Oncol* (2017) 14(12):749–62. doi: 10.1038/nrclinonc.2017.141
29. Whiting PF, Rutjes AWS, Westwood ME, Mallett S, Deeks JJ, Reitsma JB, et al. QUADAS-2: A Revised Tool for the Quality Assessment of Diagnostic Accuracy Studies. *Ann Intern Med* (2011) 155(8):529–U104. doi: 10.7326/0003-4819-155-8-201110180-00009
30. Park JE, Kim D, Kim HS, Park SY, Kim JY, Cho SJ, 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* (2020) 30(1):523–36. doi: 10.1007/s00330-019-06360-z
31. Whiting P, Rutjes AW, Reitsma JB, Bossuyt PM, Kleijnen J. The Development of QUADAS: A Tool for the Quality Assessment of Studies of Diagnostic Accuracy Included in Systematic Reviews. *BMC Med Res Methodol* (2003) 3:25. doi: 10.1186/1471-2288-3-25
32. Higgins JPT, Thompson SG, Deeks JJ, Altman DG. Measuring Inconsistency in Meta-Analyses. *Brit Med J* (2003) 327(7414):557–60. doi: 10.1136/bmj.327.7414.557
33. Piao Y, Jiang C, Wang L, Yan F, Ye Z, Fu Z, et al. The Usefulness of Pretreatment MR-Based Radiomics on Early Response of Neoadjuvant

Chemotherapy in Patients With Locally Advanced Nasopharyngeal Carcinoma. *Oncol Res* (2021) 28(6):605–13. doi: 10.3727/096504020X16022401878096
34. Wang G, He L, Yuan C, Huang Y, Liu Z, Liang C. Pretreatment MR Imaging Radiomics Signatures for Response Prediction to Induction Chemotherapy in Patients With Nasopharyngeal Carcinoma. *Eur J Radiol* (2018) 98:100–6. doi: 10.1016/j.ejrad.2017.11.007
35. Zhang L, Ye Z, Ruan L, Jiang M. Pretreatment MRI-Derived Radiomics May Evaluate the Response of Different Induction Chemotherapy Regimens in Locally Advanced Nasopharyngeal Carcinoma. *Acad Radiol* (2020) 27 (12):1655–64. doi: 10.1016/j.acra.2020.09.002
36. Zhang L, Wu X, Liu J, Zhang B, Mo X, Chen Q, et al. MRI-Based Deep-Learning Model for Distant Metastasis-Free Survival in Locoregionally Advanced Nasopharyngeal Carcinoma. *J Magn Reson Imaging* (2021) 53 (1):167–78. doi: 10.1002/jmri.27308
37. Chen X, Li Y, Li X, Cao X, Xiang Y, Xia W, et al. An Interpretable Machine Learning Prognostic System for Locoregionally Advanced Nasopharyngeal Carcinoma Based on Tumor Burden Features. *Oral Oncol* (2021) 118:105335. doi: 10.1016/j.oraloncology.2021.105335
38. Zhao L, Gong J, Xi Y, Xu M, Li C, Kang X, et al. MRI-Based Radiomics Nomogram may Predict the Response to Induction Chemotherapy and Survival in Locally Advanced Nasopharyngeal Carcinoma. *Eur Radiol* (2020) 30(1):537–46. doi: 10.1007/s00330-019-06211-x
39. Peng H, Dong D, Fang MJ, Li L, Tang LL, Chen L, et al. Prognostic Value of Deep Learning PET/CT-Based Radiomics: Potential Role for Future Individual Induction Chemotherapy in Advanced Nasopharyngeal Carcinoma. *Clin Cancer Res* (2019) 25(14):4271–9. doi: 10.1158/1078-0432.CCR-18-3065
40. Zhong LZ, Fang XL, Dong D, Peng H, Fang MJ, Huang CL, et al. A Deep Learning MR-Based Radiomic Nomogram may Predict Survival for Nasopharyngeal Carcinoma Patients With Stage T3N1M0. *Radiother Oncol* (2020) 151:1–9. doi: 10.1016/j.radonc.2020.06.050
41. Dong D, Zhang F, Zhong LZ, Fang MJ, Huang CL, Yao JJ, et al. Development and Validation of a Novel MR Imaging Predictor of Response to Induction Chemotherapy in Locoregionally Advanced Nasopharyngeal Cancer: A Randomized Controlled Trial Substudy (NCT01245959). *BMC Med* (2019) 17(1):190. doi: 10.1186/s12916-019-1422-6
42. Yang Y, Wang M, Qiu K, Wang Y, Ma X. Computed Tomography-Based Deep-Learning Prediction of Induction Chemotherapy Treatment Response in Locally Advanced Nasopharyngeal Carcinoma. *Strahlenther Onkol.* (2022) 198(2):183–93. doi: 10.1007/s00066-021-01874-2
43. Hu CM, Zheng DC, Cao XS, Pang PP, Fang YH, Lu T, et al. Application Value of Magnetic Resonance Radiomics and Clinical Nomograms in Evaluating the Sensitivity of Neoadjuvant Chemotherapy for Nasopharyngeal Carcinoma. *Front Oncol* (2021) 11:740776. doi: 10.3389/fonc.2021.740776
44. Liao H, Chen XB, Lu SL, Jin GQ, Pei W, Li Y, et al. MRI-Based Back Propagation Neural Network Model as a Powerful Tool for Predicting the Response to Induction Chemotherapy in Locoregionally Advanced Nasopharyngeal Carcinoma. *J Magn Reson Imaging* (2021). doi: 10.1002/jmri.28047
45. Eisenhauer EA, Therasse P, Bogaerts J, Schwartz LH, Sargent D, Ford R, et al. New Response Evaluation Criteria in Solid Tumours: Revised RECIST Guideline (Version 1.1). *Eur J Cancer* (2009) 45(2):228–47. doi: 10.1016/j.ejca.2008.10.026
46. Nasser M. Cochrane Handbook for Systematic Reviews of Interventions. *Am J Public Health* (2020) 110(6):753–4. doi: 10.2105/Ajph.2020.305609
47. Gul M, Bonjoc KC, Gorlin D, Wong CW, Salem A, La V, et al. Diagnostic Utility of Radiomics in Thyroid and Head and Neck Cancers. *Front Oncol* (2021) 11:639326. doi: 10.3389/fonc.2021.639326
48. Zwanenburg A, Vallieres M, Abdalah MA, Aerts HJWL, Andrearczyk V, Apte A, et al. The Image Biomarker Standardization Initiative: Standardized Quantitative Radiomics for High-Throughput Image-Based Phenotyping. *Radiology* (2020) 295(2):328–38. doi: 10.1148/radiol.2020191145
49. Jiang Z, Wang B, Han X, Zhao P, Gao M, Zhang Y, et al. Multimodality MRI-Based Radiomics Approach to Predict the Posttreatment Response of Lung Cancer Brain Metastases to Gamma Knife Radiosurgery. *Eur Radiol* (2022) 32 (4):2266–76. doi: 10.1007/s00330-021-08368-w
50. Zhang X, Xu X, Tian Q, Li B, Wu Y, Yang Z, et al. Radiomics Assessment of Bladder Cancer Grade Using Texture Features From Diffusion-Weighted Imaging. *J Magn Reson Imaging* (2017) 46(5):1281–8. doi: 10.1002/jmri.25669
51. Cunliffe A, Armato SG3rd, Castillo R, Pham N, Guerrero T, Al-Hallaq HA. Lung Texture in Serial Thoracic Computed Tomography Scans: Correlation of Radiomics-Based Features With Radiation Therapy Dose and Radiation Pneumonitis Development. *Int J Radiat Oncol Biol Phys* (2015) 91(5):1048–56. doi: 10.1016/j.ijrobp.2014.11.030
52. Tian Q, Yan L-F, Zhang X, Zhang X, Hu Y-C, Han Y, et al. Radiomics Strategy for Glioma Grading Using Texture Features From Multiparametric MRI. *J Magn Reson Imaging* (2018) 48(6):1518–28. doi: 10.1002/jmri.26010
53. Bo L, Zhang Z, Jiang Z, Yang C, Huang P, Chen T, et al. Differentiation of Brain Abscess From Cystic Glioma Using Conventional MRI Based on Deep Transfer Learning Features and Hand-Crafted Radiomics Features. *Front Med (Lausanne)* (2021) 8:748144. doi: 10.3389/fmed.2021.748144
54. Sanduleanu S, Woodruff HC, de Jong EEC, van Timmeren JE, Jochems A, Dubois L, et al. Tracking Tumor Biology With Radiomics: A Systematic Review Utilizing a Radiomics Quality Score. *Radiother Oncol* (2018) 127 (3):349–60. doi: 10.1016/j.radonc.2018.03.033
55. Park CM. Can Artificial Intelligence Fix the Reproducibility Problem of Radiomics? *Radiology* (2019) 292(2):374–5. doi: 10.1148/radiol.2019191154
56. Cuocolo R, Cipullo MB, Stanzione A, Romeo V, Green R, Cantoni V, et al. Machine Learning for the Identification of Clinically Significant Prostate Cancer on MRI: A Meta-Analysis. *Eur Radiol* (2020) 30(12):6877–87. doi: 10.1007/s00330-020-07027-w
57. Cronin P, Kelly AM, Altaee D, Foerster B, Petrou M, Dwamena BA. How to Perform a Systematic Review and Meta-Analysis of Diagnostic Imaging Studies. *Acad Radio* (2018) 25(5):573–93. doi: 10.1016/j.acra.2017.12.007
58. Kao YS, Hsu Y. A Meta-Analysis for Using Radiomics to Predict Complete Pathological Response in Esophageal Cancer Patients Receiving Neoadjuvant Chemoradiation. *In Vivo* (2021) 35(3):1857–63. doi: 10.21873/invivo.12448
59. Ugga L, Perillo T, Cuocolo R, Stanzione A, Romeo V, Green R, et al. Meningioma MRI Radiomics and Machine Learning: Systematic Review, Quality Score Assessment, and Meta-Analysis. *Neuroradiology* (2021) 63 (8):1293–304. doi: 10.1007/s00234-021-02668-0

Conflict of Interest: Author DZ is employed by Shandong Provincial Key Laboratory of Transmucosal and Transdermal Drug Delivery, Shandong Freda Pharmaceutical Group Co., Ltd., Shandong Academy of Pharmaceutical Sciences. Author JJ is employed by Qingdao NovelBeam Technology Co., Ltd. Author XW is employed by Shangdong AccurDx Diagnosis of Biotech Co., Ltd.

The remaining 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.

The handling editor XM declared a shared parent affiliation with the author JW at the time of review.

frontiers | Frontiers in Oncology

ORIGINAL RESEARCH
published: 03 June 2022
doi: 10.3389/fonc.2022.854979

Differential Diagnosis of Type 1 and Type 2 Papillary Renal Cell Carcinoma Based on Enhanced CT Radiomics Nomogram

Yankun Gao[1], Xingwei Wang[1], Shihui Wang[2], Yingying Miao[1], Chao Zhu[1], Cuiping Li[1], Guoquan Huang[3], Yan Jiang[4], Jianying Li[5], Xiaoying Zhao[1] and Xingwang Wu[1]**

[1] Department of Radiology, The First Affiliated Hospital of Anhui Medical University, Hefei, China, [2] Department of Radiology, The First Affiliated Hospital of Wannan Medical college, Wuhu, China, [3] Department of Imaging, Wuhu Second People's Hospital, Wuhu, China, [4] Department of Pathology, The First Affiliated Hospital of Anhui Medical University, Hefei, China, [5] CT Research Center, GE Healthcare China, Shanghai, China

OPEN ACCESS

Edited by:
Pei Yang,
Central South University, China

Reviewed by:
Jun Shen,
Sun Yat-sen University, China
Bassem Youssef,
American University of Beirut, Lebanon
Stefania Volpe,
University of Milan, Italy

***Correspondence:**
Xiaoying Zhao
jay2491112@163.com
Xingwang Wu
duobi2004@126.com

Specialty section:
This article was submitted to Genitourinary Oncology, a section of the journal Frontiers in Oncology

Received: 14 January 2022
Accepted: 11 May 2022
Published: 03 June 2022

Citation:
Gao Y, Wang X, Wang S, Miao Y, Zhu C, Li C, Huang G, Jiang Y, Li J, Zhao X and Wu X (2022) Differential Diagnosis of Type 1 and Type 2 Papillary Renal Cell Carcinoma Based on Enhanced CT Radiomics Nomogram. Front. Oncol. 12:854979. doi: 10.3389/fonc.2022.854979

Objectives: To construct a contrast-enhanced CT-based radiomics nomogram that combines clinical factors and a radiomics signature to distinguish papillary renal cell carcinoma (pRCC) type 1 from pRCC type 2 tumours.

Methods: A total of 131 patients with 60 in pRCC type 1 and 71 in pRCC type 2 were enrolled and divided into training set (n=91) and testing set (n=40). Patient demographics and enhanced CT imaging characteristics were evaluated to set up a clinical factors model. A radiomics signature was constructed and radiomics score (Rad-score) was calculated by extracting radiomics features from contrast-enhanced CT images in corticomedullary phase (CMP) and nephrographic phase (NP). A radiomics nomogram was then built by incorporating the Rad-score and significant clinical factors according to multivariate logistic regression analysis. The diagnostic performance of the clinical factors model, radiomics signature and radiomics nomogram was evaluated on both the training and testing sets.

Results: Three validated features were extracted from the CT images and used to construct the radiomics signature. Boundary blurring as an independent risk factor for tumours was used to build clinical factors model. The AUC value of the radiomics nomogram, which was based on the selected clinical factors and Rad-score, were 0.855 and 0.831 in the training and testing sets, respectively. The decision curves of the radiomics nomogram and radiomics signature in the training set indicated an overall net benefit over the clinical factors model.

Conclusion: Radiomics nomogram combining clinical factors and radiomics signature is a non-invasive prediction method with a good prediction for pRCC type 1 tumours and type 2 tumours preoperatively and has some significance in guiding clinicians selecting subsequent treatment plans.

Keywords: radiomics nomogram, papillary renal cell carcinoma, differential diagnosis, computed tomography, tumour subtypes

INTRODUCTION

Renal cell carcinoma (RCC) is the most common malignancy of the kidney in adults, accounting for approximately 85% of renal tumours (1). Clear cell RCC (ccRCC), papillary RCC (pRCC) and chromophobe RCC (chRCC), accounting for 70-80%, 10-20% and 3-7% of RCCs, respectively (2). PRCC is the second most common subtype after ccRCC. Among the subtypes of RCC, pRCC has a higher 5-year survival rate and a better prognosis. In 1997, Delahunt et al. initially subdivided pRCC into type 1 and type 2 according to morphological and immunohistochemical characteristics (3). Typically, type 1 exhibits papillae covered by a single layer of monolayer cuboidal epithelium with a lack of cytoplasm, whereas type 2 is characterized by the presence of nuclear pseudostratification (4). Previous studies have shown that type 2 tumours tend to have a higher pathological stage, a higher nuclear grade, as patients with type 2 tumours have a worse prognosis (5–7). As type 2 tumours are more aggressive, an early and accurate diagnosis is essential. Due to the low malignancy of type 1 tumours, relatively conservative treatment options such as follow-up, ablation and partial nephrectomy are usually available in clinical practice. According to the National Comprehensive Cancer Network (NCCN) RCC guidelines, less aggressive RCC can be treated by active surveillance or partial nephrectomy. In contrast, most highly aggressive RCC patients usually undergo radical nephrectomy with consideration of adjuvant therapy. The precise preoperative differentiation between these two types of the tumours will determine different treatment options and different prognoses.

Pathological biopsy by percutaneous puncture biopsy or surgical excision is the most accurate method of identifying the pRCC subtype, but it is after all an invasive test, and we would like to be able to make a non-invasive diagnosis preoperatively. Although some studies have shown a higher heterogeneity of type 2 tumours compared to type 1 tumours on conventional computed tomography (CT) and magnetic resonance imaging (MRI) images, typically type 2 tumours are large, have blurred margins and tend to invade blood vessels and metastasize to surrounding lymph nodes (8–10). However, these two types of tumours have many overlapping imaging features on conventional CT or MRI images, and it is often difficult to distinguish subtypes of pRCC based on imaging features alone.

Radiomics is a recent emerging research approach that uses high-throughput data feature extraction algorithms to translate medical images into high-dimensional, useable quantitative image features, and it uses various algorithms for deeper analysis of the features. This method can be used not only for preoperative pathological classification and grading of the tumour, but also for the prediction of prognosis and survival rate of tumour patients (11–13). Currently, radiomics studies for RCC have been focused on the identification of the three most common subtypes of RCC (ccRCC, pRCC, chRCC) and on the nuclear grading of RCC (10, 14–18). For example, Deng et al. (14)showed that CT-based texture analysis was not only able to identify ccRCC and pRCC, but also to predict the Fuhrman grade of the tumour. Some studies have shown that CT and MRI-based texture analysis techniques can differentiate between pRCC subtypes (10, 17, 18). However, to the best of our knowledge, apart from some studies that have identified pRCC subtypes based on textural features alone, there is no study that combines radiomics features with clinical factors to make a differential diagnosis of pRCC subtypes. In our study, we quantified radiomics signature by calculating the rad-score value form contrast-enhanced CT images of each patient and attempted to build a contrast enhanced CT-based radiomics nomogram that included both rad-score and clinical factors to better discriminate between the two subtypes of pRCC.

MATERIALS AND METHODS

Patients

The ethics review board at our hospital approved this retrospective study and patient informed consent was waived. Patients who underwent non-enhanced and contrast-enhanced CT scans from January 2013 to October 2021 at our hospital for diagnosing kidney disease were considered. Percutaneous puncture or surgical excision specimens diagnosed as pRCC type 1 or 2 were selected by searching the hospital's picture archiving and communication system (PACS). The inclusion criteria were as follows: (1) Patients who had a definitive pathologic diagnosis of pRCC. (2) Patients with available preoperative plain and enhanced CT scans, and the image quality was satisfactory for analysis (clear image with no artifacts). (3) Patients with complete clinic-pathological data. The exclusion criteria were as follows: (1) The subtype of pRCC patients could not be determined as type 1 or type 2. (2) Patients who had a history of abdominal surgery. (3) Patients received abdominal radiotherapy or chemotherapy prior to CT scan.

CT Image Acquisition

CT scan protocols are shown in **Table 1**. A power injector administered a 90-100-ml of nonionic contrast medium (Omnipaque, GE Healthcare or Ultravist, Bayer, Schering Pharma) into the antecubital vein at a rate of 3 mL/s. Pre-contrast CT of the abdomen was first acquired, followed by three

TABLE 1 | CT scan protocols.

Manufacturer	Siemens	General Electric	Philips
Scanner model	Sensation 64	Discovery 750	Brilliance
Sequence	Axial	Axial	Axial
Gantry rotation time (s)	0.5	0.5	0.5
Tube voltage (kV)	120	120	120
Tube current (mA)	200	250-400	180-450
Detector collimation (mm)	64×0.6	64×0.625	64×0.625
Matrix	512×512	512×512	512×512
Pitch	1.0	1.375	1.0
Slice thickness (mm)	5	5	5
Corticomedullary phase (s)	30	30	30
Nephrographic phase (s)	80	80	80
Excretory phase (s)	180	180	180

s, second; kV, kilovolt; mA, milliampere; mm, millimetre.

post-contrast CT scans obtained in the corticomedullary phase (CMP, acquired 30 s after contrast injection), nephrographic phase (NP, acquired 80 s after contrast injection) and excretory phase (EP, acquired 180 s after contrast injection).

CT Characteristic Evaluation

The CT image were scrutinized by two radiologists with 4 years (reader 1, Y.G.) and 7 years (reader 2, X.W.) of diagnostic abdominal imaging experience. If there was disagreement, two radiologists needed to reach a consensus. Without knowledge of the clinicopathologic data, the two readers interpreted the following CT characteristics together: the maximum diameter of the tumour on axial CT images; shape (round or not round); location (left or right); boundary (clear or blurred boundary); calcification (present or not, "calcification was considered as high density seen during pre-enhancement CT"); necrosis (present or not, "necrosis was considered as the non-enhanced liquid area of tumour accounting for more than 50% of the tumour"); renal vein invasion (present or not, "renal vein invasion was considered as the tumour tissue in the renal vein and inferior vena cava was observed on the imaging"); lymph node metastasis (present or not, "lymph node metastasis was considered as the short-axis diameter of the perirenal and retroperitoneal lymph nodes were greater than 10mm") (9, 19).

To standardize the measurement of tumour enhancement, it is generally necessary to select the appropriate region of interests (ROIs) within the tumour and characterize the tumour enhancement according to the changes in CT values of the ROIs on different scan phases. Since the tumour had been enhanced to some extent on the CMP images and the various heterogenous components of the tumour could be better displayed at the stage, all ROIs in this study were selected based on the CMP images. To accurately assess the extent of tumour enhancement, the ROIs avoided components such as necrosis, calcification, and vascularity that are clearly visible on the images and include only the substantial components of the tumour. The reader 1 select 2 non-overlapping ROIs, made separate measurements and averaged the two numbers to obtain the final measurement. Due to individual patient factors and factors that are difficult to control when performing CT scanning operations, the iodine contrast load during the scan was not identical in each case, and this variation could constitute a systematic error in the measurement of tumour enhancement. In this study, the cortical area of the kidney on the side of the tumour was selected as the reference area for iodine contrast loading normalization during the scan to correct for such systematic errors. **Figure 1** shows an example of this approach.

The ROIs selected in CMP were copied and pasted into the non-enhanced and NP images to obtain the average tumour attenuation value (TAV) in each scan phase. The average CT value of the reference area in each corresponding scan phase was used as the cortex attenuation value (CAV). The tumour enhancement value (TEV) and the cortex enhancement value (CEV) were calculated by subtracting the values of the same ROI in the non-enhanced phase: $TEV_x = TAV_x - TAV_0$ and $CEV_x = CAV_x - CAV_0$, where x represents the phase (0, non-enhanced; 1, CMP; 2, NP). The relative enhancement value (REV) was defined as the ratio of TEV to CEV: $REV_x = TEV_x/CEV_x$, representing the degree of enhancement within the tumour relative to the renal cortex (20).

Construction of the Clinical Factors Model

Univariate analysis was used to compare the differences in clinical factors (including clinical data and CT characteristcs) between the type 1 and type 2 tumours. The significant variables acquired in the univariate analysis were used as inputs, and a multivariate logistic regression analysis was applied to establish a clinical factors model. Odds ratios (OR) was calculated for each independent factor as a relative risk estimate with a 95% confidence interval (CI).

Three-Dimensional Segmentation of Tumour Images and Radiomics Feature Extraction

The basic steps of a radiomics model for renal tumours are detailed in **Figure 2**. Three-dimensional (3-D) segmentation of

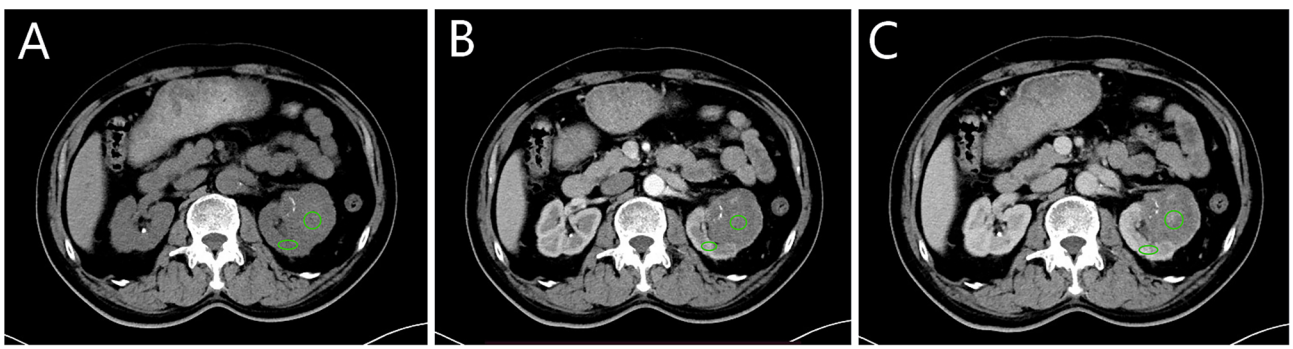


FIGURE 1 | Selection of region of interests (ROIs) and reference region. **(A–C)** correspond to the non-enhanced, corticomedullary phase (CMP) and nephrographic phase (NP). The green circle is one of two tumour ROIs, selected from the parenchymal portion of the tumour where enhancement is evident. The green oval is the reference region located in the cortical portion of the kidney. The zones of ROIs and reference region are in the same position in each scan phase.

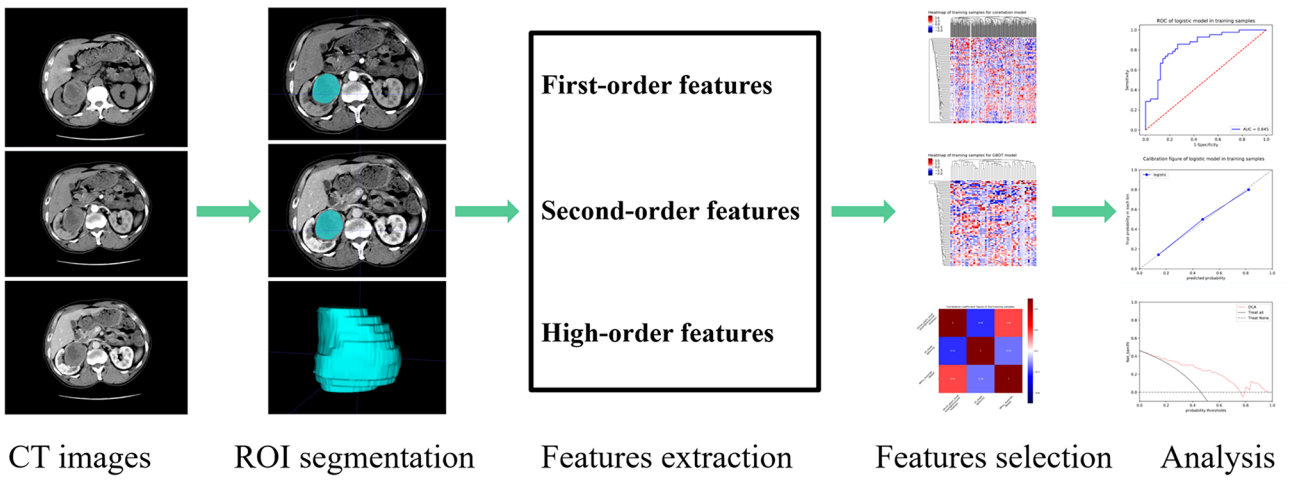


FIGURE 2 | Papillary renal cell carcinoma (pRCC) study radiomics flow chart.

tumours was performed using the ITK-SNAP software (version 3.8, www.itksnap.org). Contouring was drawn using the ROIs within the tumour borders on CMP and NP images, 1-2mm from the tumour boundary. An example of the use of manual segmentation in a renal tumour is shown in **Figure 3**.

Features extraction was executed using the PHIgo Workstation (General Electric Company). As the images were derived from three CT scanners with different parameters, normalization and image resampling had to be performed before features could be extracted from the ROIs of the CMP and NP images. The image data is normalized using a z-score in the following form:

$$z = \frac{x - \mu}{\sigma},$$

Where μ is the mean of the whole data, σ is the standard deviation of the whole data. In addition, all CT images were resampled to 1.0×1.0×1.0 mm^3 voxels to standardize the slice thickness use B-spline interpolation sampling technology. 1316 radiomics features were extracted from the ROIs of the CMP and NP images, respectively.

Inter-observer reliability and intra-observer repeatability of radiomics feature extraction were usually assessed using inter- and intra- class correlation coefficients (ICC). We randomly chose 20 cases of CT images (8 pRCC type 1 and 12 pRCC type 2); ROI segmentation was performed by reader 1 and reader 2. After two weeks, reader 1 repeated the same steps to evaluate the degree of matching of feature extraction. When the ICC value is more than 0.75, it indicates that the extracted features have a good consistency. Then the remaining image segmentation will be carried out by reader 1 alone.

Construction of the Radiomics Signature

To prevent overfitting of the radiomics features, features were further selected before the construction of the radiomics signature. First, features with ICC >0.75 within the training set were retained. Second, statistically significant features were screened out using the univariate logistic analysis. Third, the most valuable features were selected using Gradient Boosting Decision Tree and multivariate logistic analysis. Finally, a radiomics score (Rad-score) was calculated by using a formula based on the radiomics features.

Rad – score was used to establish aradiomics signature multivariate logistic regression.

Construction of Radiomics Nomogram and Performance Evaluation of Different Models

A radiomics nomogram was constructed by combining the significant variables of clinical factors and the Rad-score. Calibration curves were used to evaluate the calibration of the nomogram. The Hosmer-Lemeshow test was conducted to assess the goodness-of-fit of the nomogram. The diagnostic performance of the clinical factors model, the radiomics signature model and the radiomics nomogram for differentiating pRCC type 1 from pRCC type 2 was evaluated based on the area under the receiver operator characteristic (ROC) curve (AUC) in both the training and testing sets. To evaluate the clinical effectiveness of the radiomics nomogram, a decision curve analysis (DCA) was performed by calculating the net benefit of a threshold probability range across the training and testing sets.

Statistical Analysis

Statistical tests were performed using SPSS (version 25.0, IBM) and IPM statistical (version 2.4.0, General Electric Company). Univariate analysis was used to compare the differences in clinical factors between type 1 and type 2 tumours. Chi-square test or Fisher exact test was used for categorical variables, and Mann-Whitney U test was used for continuous variables. A two-side $p < 0.05$ was considered significant.

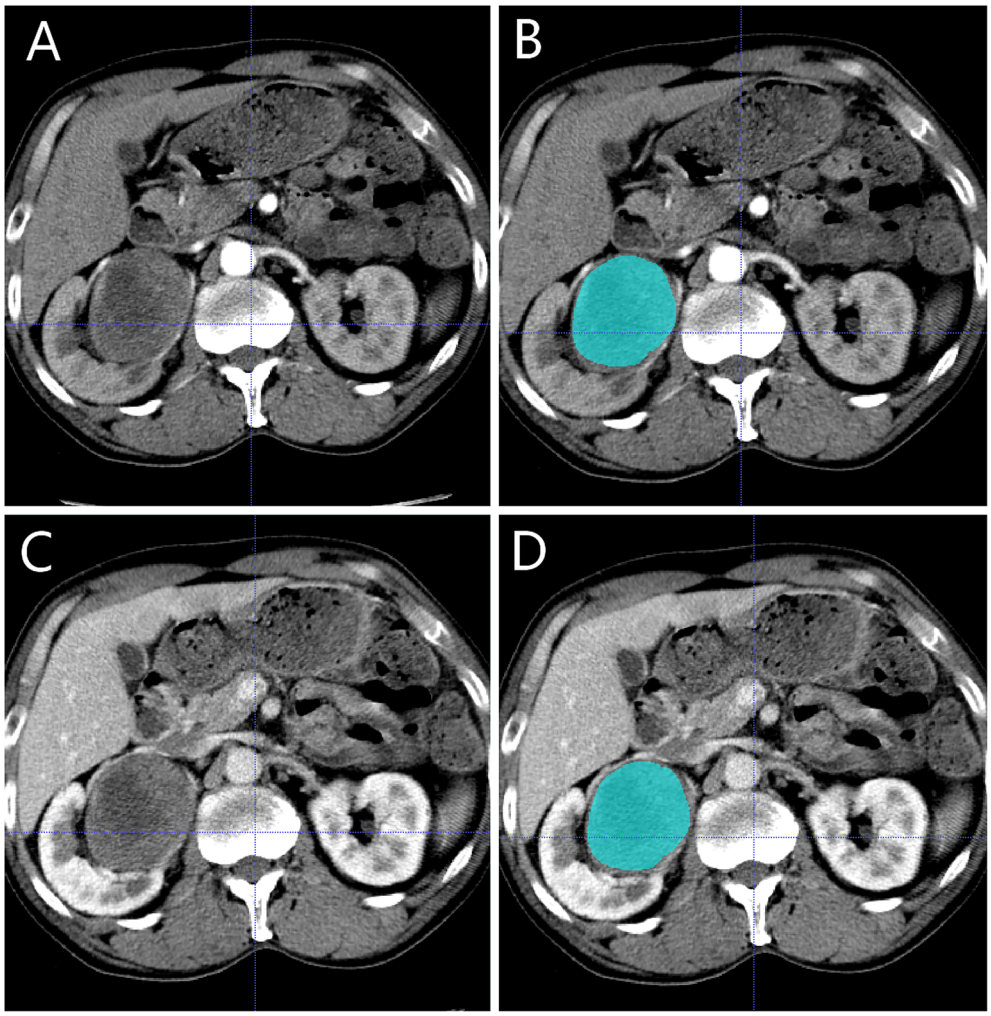


FIGURE 3 | Manual segmentation of the tumour on the center axial slice of the pRCC type 2. **(A, B)** is the corticomedullary phase (CMP); **(C, D)** is the nephrographic phase (NP).

RESULTS

Clinical Factors of the Patients and Construction of the Clinical Factor Model

A total of 131 patients were finally enrolled in this study according to the inclusion and exclusion criteria, including 60 type 1 patients (51 men and 9 women; mean age, 57.17 ± 12.17 years old) and 71 type 2 patients (55 men and 16 women; mean age, 58.56 ± 13.09 years old). The entire cohort of patients conforming to the inclusion criteria was divided randomly into the training set (n=91) and testing set (n=40) in a ratio of 7:3. The clinical factor data in the training and testing sets are shown in **Table 2**. Maximum diameter, shape, boundary, calcification, necrosis, renal vein invasion, lymph node metastasis and REV2 were statistically significant in differentiating pRCC type 1 and type 2 tumours after univariate analysis in the training set (both $p<0.05$). Multivariate logistic regression analyses were performed on the eight statistically significant clinical factors listed above. The p-value were 0.111, 0.770, 0.026, 0.342, 0.945, 1.000, 0.999 and 0.971 respectively. If the tumour boundary is blurred (OR, 2.352; 95%CI, 1.743-3.174), it is more likely to be a pRCC type 2 tumour.

Feature Extraction, Selection, and Radiomics Signature Construction

A total of 2632 radiomics features were extracted from the CMP and NP CT images, of which 1876 features had an ICCs greater than 0.75, indicating good inter-and intra- observer agreement for these features. By univariate correlation analysis, 282 radiomics features showed significant differences between type 1 and type 2 tumours. These features were sequentially imported into Gradient Boosting Decision Tree (21) and multivariate logistic analyses to obtain the most valuable features, resulting in three useful features (**Figure 4**). Finally, the radiomics signature was established by using three features. The AUC were 0.845 (95%CI 0.775-0.913) in the training set and 0.821

TABLE 2 | Clinical factors.

Clinical factors	Training cohort (n=91)		*p*	Testing cohort (n=40)		*p*
	Type1 (n=42)	Type2 (n=49)		Type1 (n=18)	Type2 (n=22)	
Gender			0.121			0.579
Male	35 (83%)	34 (69%)		16 (89%)	21 (95%)	
Female	7 (17%)	15 (31%)		2 (11%)	1 (5%)	
Age (years)	57.4 ± 12.2	58.8 ± 12.6	0.571	56.7 ± 12.4	58.1 ± 14.3	0.693
Maximum diameter (cm)	3.5 ± 2.0	6.0 ± 3.5	<0.001	3.6 ± 2.2	6.6 ± 2.7	0.001
Shape			0.001			<0.001
Round	38 (90%)	29 (59%)		17 (94%)	9 (41%)	
Not round	4(10%)	20 (41%)		1 (6%)	13 (59%)	
Location			0.174			0.356
Left	18 (43%)	28 (57%)		8 (44%)	13 (59%)	
Right	24 (57%)	21 (43%)		10 (56%)	9 (41%)	
Boundary			<0.001			0.004
Clear	41 (98%)	28 (57%)		16 (89%)	10 (45%)	
Blurred	1 (2%)	21 (43%)		2 (11%)	12 (55%)	
Calcification			0.004			0.427
Present	4 (10%)	17 (35%)		2 (11%)	5 (23%)	
Absent	38 (90%)	32 (65%)		16 (89%)	17 (77%)	
Necrosis			<0.001			<0.001
Present	6 (14%)	26 (53%)		2 (11%)	16 (73%)	
Absent	36 (86%)	23 (47%)		16 (89%)	6 (27%)	
Renal vein invasion			0.003			0.011
Present	0 (0)	9 (18%)		0 (0)	7 (32%)	
Absent	42 (100%)	40 (82%)		18 (100%)	15 (68%)	
Lymph node metastasis			0.001			0.005
Present	0 (0)	12 (24%)		0 (0)	8 (36%)	
Absent	42 (100%)	37 (76%)		18 (100%)	14 (64%)	
TEV1 (HU)	18.5 ± 17.4	27.4 ± 32.5	0.063	13.4 ± 7.3	29.8 ± 31.5	0.121
TEV2 (HU)	32.0 ± 21.3	38.6 ± 23.3	0.068	28.2 ± 11.4	43.1 ± 30.1	0.178
REV1	0.19 ± 0.16	0.28 ± 0.24	0.058	0.19 ± 0.20	0.30 ± 0.28	0.103
REV2	0.28 ± 0.28	0.33 ± 0.17	0.014	0.23 ± 0.12	0.42 ± 0.49	0.092

TEV, tumour enhancement value; REV, relative enhancement value; 1, corticomedullary phase; 2, nephrographic phase.

(95%CI 0.702-0.922) in the testing set. The Rad-score was calculated using the following formula:

$$\text{Rad} - \text{score} = -0.2964 - 1.1110 \times \text{CMP} - \text{wavelet}-\text{LHH}\_\text{glszm}\_\text{SmallAreaHighGrayLevelEmpha} + 1.1637 \times \text{CMP} - \text{original}\_\text{shape}\_\text{Sphericity} + 0.9529 \times \text{NP} - \text{wavelet} - \text{HLL}\_\text{firstorder}\_\text{Median}$$

The distribution of the Rad-score in the training and testing sets is shown in **Figure 5**.

Establishment of Radiomics Nomogram and Evaluation of Performance Between Different Models

Using the data in the training set, a radiomics nomogram was established by combining important clinical factors which was the boundary information and Rad-score (**Figure 6**), and the radiomics nomogram score (Nomo-score) was calculated based on multivariate logistic regression analysis. The formula for calculating the Nomo-score for this study is shown below: Nomo-score = -2.1459 + B×2.3959 + R×0.8423 (B = Boundary; R = Rad-score). The calibration curves of the radiomics nomogram in **Figure 7** showed good calibration in both the training and testing cohorts. The discriminatory efficacies of the three diagnostic models (clinical factors model, radiomics signature and radiomics nomogram) are shown in **Table 3**. **Figure 8** plots the clinical factors model, radiomics signature and radiomics nomogram ROC curves based on the training cohort and testing cohort comparing the accuracies of these three models in identifying pRCC type 1 and type 2 tumours. The decision curves showed that in most training cohorts within reasonable threshold probabilities, the radiomics nomogram added greater overall net benefit in differentiating between pRCC type 1 and type 2 tumours compared to the clinical factors and radiomics signature. The DCA value for the three models in the training cohort are shown in **Figure 9**.

DISCUSSION

PRCC is the second most common subtype of RCC, second only to ccRCC. PRCC can be divided into two different subtypes, type 1 and type 2 (2). The systematic review and meta-analysis by Xiong et al. collected a total of 4494 pRCC patients from 22 studies and showed that overall survival and cancer specific survival was worse in type 2 pRCC patients than in type 1 pRCC patients (22). Because the two tumours have many

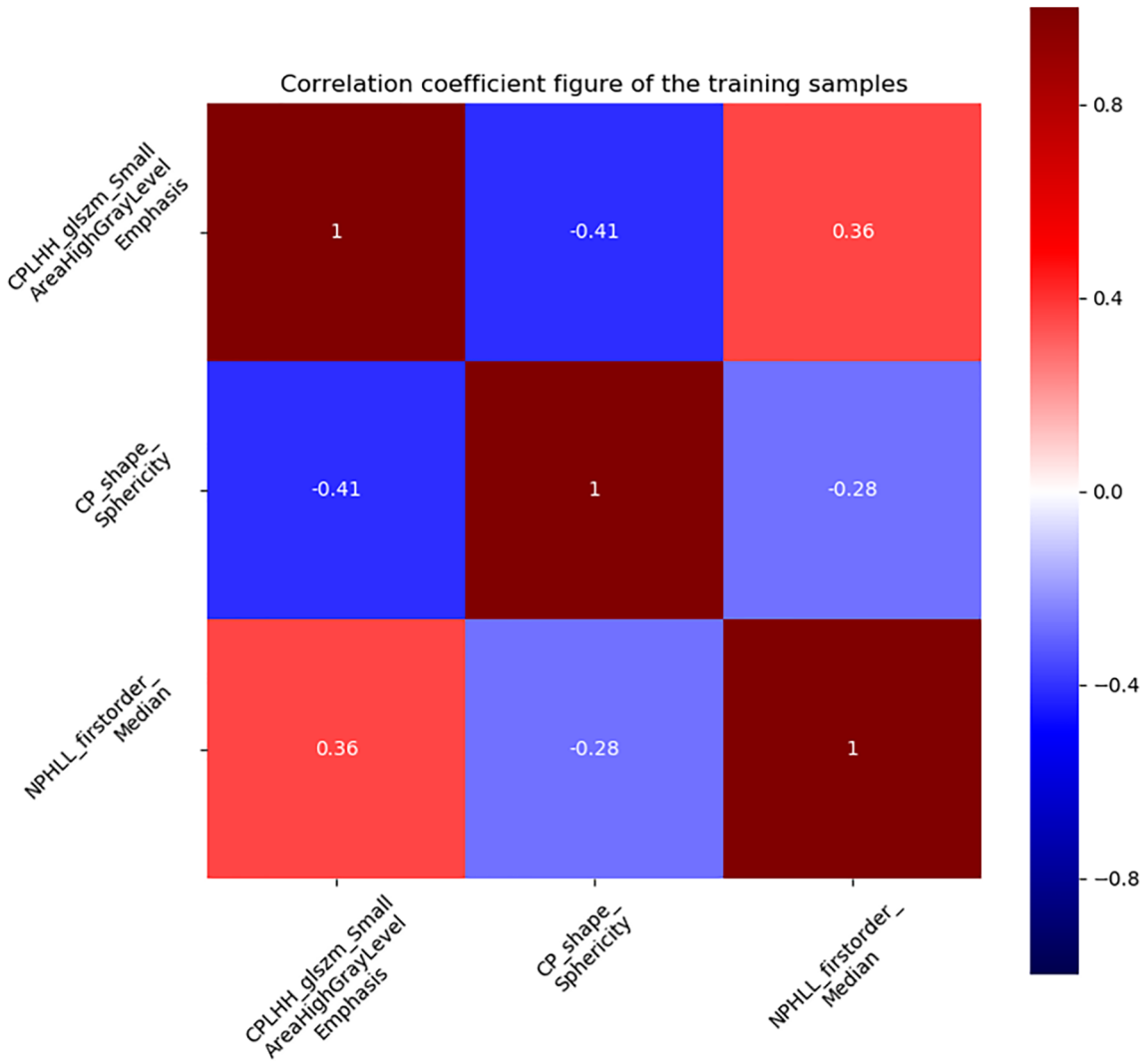


FIGURE 4 | The correlation diagram of the three effective features screened out.

differences in biology, treatment options and prognosis, it is clinically important to distinguish accurately between type 1 and type 2 tumours preoperatively. In the present study, the radiomics nomogram was constructed by combining clinical factors with Rad-score and was found to be highly accurate in distinguishing pRCC subtypes, with an AUC value of 0.855 in the training cohort.

Previous studies have shown that clinical and conventional CT and MRI images can help distinguish between pRCC type 1 and type 2 tumours (2, 8, 9, 23). Fourteen clinical factors were used for analysis in our study, mainly including gender, age, maximum diameter, shape, location, boundary, calcification, necrosis, renal vein invasion, lymph node metastasis, tumour enhancement value (TEV1 and TEV2) and relative enhancement value (REV1 and REV2). After multivariate logistic regression analysis, blurred tumour boundaries could be used as an independent predictor of type 2 tumours, which is in line with the results of previous studies (8–10). We believe that the most likely reason for this result is that type 2 tumours are highly malignant and aggressive, more likely to invade the fatty layer surrounding the kidney, resulting in poorly defined borders on CT images. In this study, type 2 tumours were significantly larger in diameter than type 1 tumours, and the difference was statistically significant (p<0.001), which is consistent with the findings of Egbert et al. (8) and Yamada et al. (9). In contrast, some of the findings showed that the difference in diameter between type 2 and type 1 tumours was not statistically significant (24–26). We found that type 2 tumours had more necrosis and calcification compared to type 1 tumours and that this difference was statistically significant, which is consistent with previous findings (8, 18, 27). Mydlo et al. (28) found that type 1 tumours were less enhanced than type 2 tumours on CT-

FIGURE 5 | The radiomics score (Rad-score) for each patient in the training **(A)** and testing **(B)** sets.

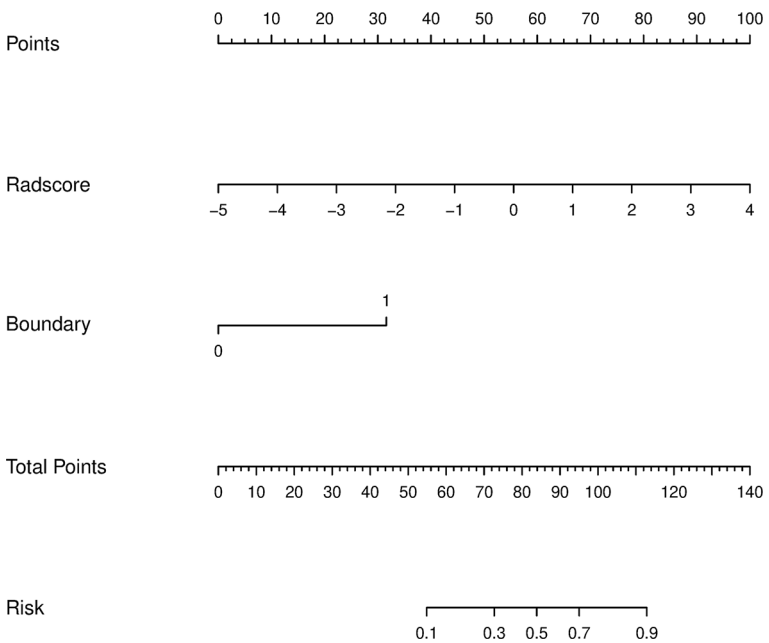


FIGURE 6 | A radiomics nomogram distinguishing between type 1 and type 2 tumours. The nomogram was constructed by combining boundary and radiomics score (Rad-score) on the basis of a training cohort. The corresponding points are estimated from the boundary and Rad-score values, and these are added together to obtain total points. The likelihood of type 2 pRCC was estimated from the total points, the greater the total points, the greater the probability of type 2 pRCC.

enhanced scans, and we found no statistically significant difference between TEV1, REV1 in CMP and TEV2 in NP, while REV2 in NP type 2 tumours were significantly greater than type 1 tumours and the difference was statistically significant (p=0.014). We believe the reason for this outcome is the higher malignancy of type 2 tumours and the abundance of tumour neovascularization.

Radiomics is a newly emerging research method that has been widely used in the diagnosis and differential diagnosis of kidney tumours (17, 29–31). It aids clinical decision-making by extracting high-throughput quantitative data from images, thus enabling non-invasion analysis of tumour heterogeneity. Previous findings show that CT and MRI-based radiomics can be used to differentiate between pRCC type 1 and type 2 tumours. Wang et al. (32) collected 77 patients with RCC, including 32 ccRCC, 23 pRCC and 22 cRCC. The patients all underwent routine MRI (T2WI, EN-T1WI CMP, EN-T1WI NP) preoperatively, and a total of 39 radiomics features were

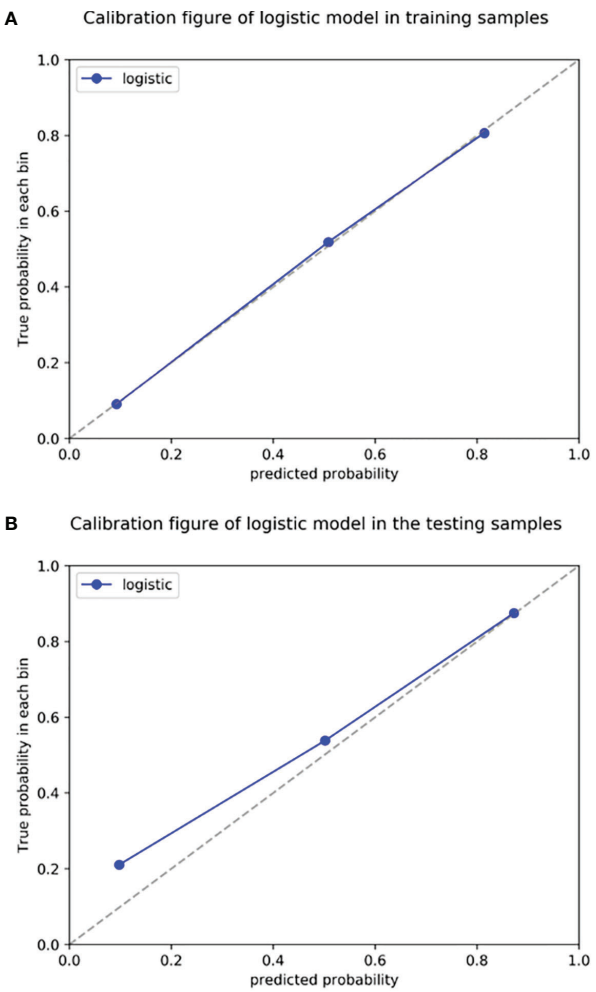


FIGURE 7 | Radiomics nomogram calibration curves for the training **(A)** and testing **(B)** sets. The calibration curves show a good fit of the nomogram. The 45° straight lines indicate a perfect match between the true (Y-axis) and predicted (X-axis) probabilities. The closer the distance between the two curves, the better the accuracy.

extracted from the three sequences mentioned above. The final ROC curves were constructed and showed AUC values of 0.631-0.951 for differentiating ccRCC and cRCC; AUC values of 0.688-0.955 for differentiating pRCC and cRCC, and AUC values of 0.747-0.890 for differentiating ccRCC and pRCC. Yap et al. (33) extracted a total of 33 shape and 760 texture features from preoperative CT images of 735 renal masses (539 malignant and 196 benign) and used these features to build a radiomics model based on multiple machine learning classifiers for identifying benign and malignant renal masses. The AUC values were 0.64-0.68 for the shape features, 0.67-0.75 for the texture features, and 0.68-0.75 for the combination of shape and texture features. Nie et al. (30) collected a total of 99 patients who underwent preoperative CT examination and divided into a training set (n=80) and testing set (n=19) in order to construct a radiomics nomogram that could distinguish AML.wolf from hm-ccRCC preoperatively. A total of 14 valid features were selected from CMP and NP to build radiomics nomogram, which showed good discriminatory efficacy in both the training set (AUC, 0.896; 95% CI, 0.810-0.983) and the testing set (AUC, 0.949; 95%CI, 0.856-1.000). Its discriminatory power was higher than that of the clinical factors model and the radiomics signature. Doshi et al. (10) assessed whether qualitative features (signal intensity, heterogeneity, and margin) and quantitative textural features (ADC, HASTE, and contrast-enhanced entropy) from preoperative MRI images of 21 pRCC type 1 tumours and 17 type 2 tumours could be for preoperative differentiation between type 1 and type 2 tumours. The results showed that the AUC values were 0.822 for the qualitative feature model, 0.682-0.716 for the quantitative feature model, and 0.859 for the combined qualitative and quantitative feature model. Duan et al. (17) extracted textures features based on 62 preoperative three-phase enhanced CT images of pRCC (30 type 1 tumours and 32 type 2 tumours) and built a model based on an SVM classifier. The AUC values were 0.772-0.753 for the CMP-based model, 0.832-0.841 for the NP-based model, 0.849-0.858 for the EP-based model, and 0.922 for the combined three-phase model. The results showed that CT-based texture analysis could be used to preoperatively differentiate between type 1 and type 2 tumours.

The nomogram is a practical and straightforward statistical prediction tool that has been widely used to combine multiple risk factors to predict medical prognosis and outcomes (34). Huang et al. (35) combined clinical factors with radiomics signature to construct a nomogram for predicting disease-free survival in non-small lung cancer. The nomogram's diagnostic efficacy was higher than clinical factors alone. Our study builds a nomogram based on boundary and Rad-score to predict the probability of type 1 tumours with AUC values of 0.855, 0.831 in the training and testing sets, respectively. The AUC values for the model constructed on clinical factors alone were 0.702,0.717 in

TABLE 3 | Diagnostic performance of the clinical factors model, the radiomics signature and the radiomics nomogram.

Model	Training cohort		Testing cohort	
	AUC (95%CI)	Accuracy %	AUC (95%CI)	Accuracy %
Clinical factors model	0.702 (0.643,0.764)	53.8	0.717 (0.611,0.826)	55.0
Radiomics signature	0.845 (0.775,0.913)	78.0	0.821 (0.702,0.922)	75.0
Radiomics nomogram	0.855 (0.787,0.918)	78.0	0.831 (0.716,0.930)	75.0

AUC, area under the curve.

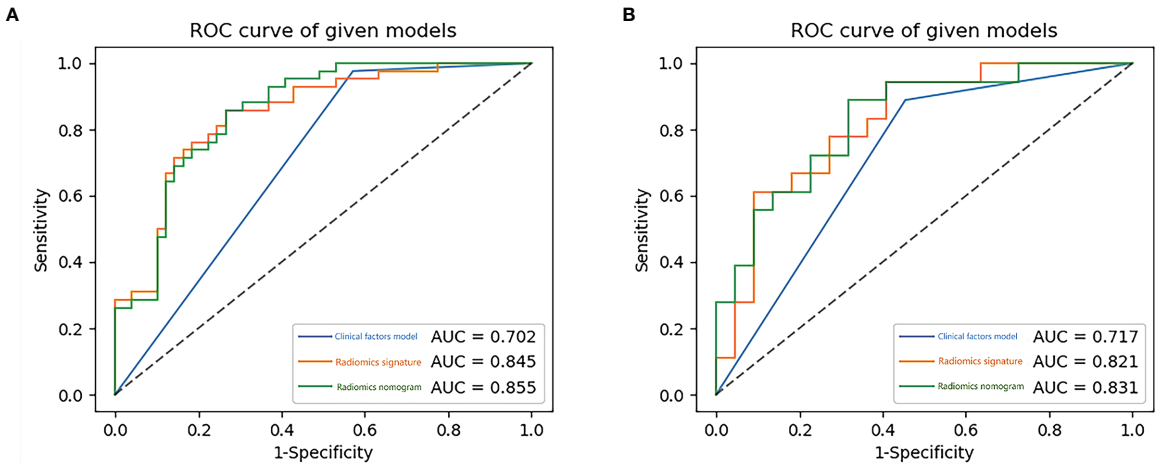


FIGURE 8 | The receiver operating characteristic (ROC) curves of the clinical factors model, the radiomics signature and the radiomics nomogram for training **(A)** and testing **(B)** sets.

the training and testing sets, respectively. The high diagnostic efficacy of the nomogram over clinical factors alone suggests that the Rad-score is of high value in differentiating between type 1 and type 2 tumours.

Compared to the above radiomics studies, our study had some differences and provided some improvements: First, our study focused on distinguishing between type 1 pRCC and type 2 pRCC, mainly because type 1 and type 2 tumours often have many overlapping imaging presentations in CT images. Second, a total of three radiomics features were extracted from the CMP and NP images, two of which were derived from the CMP images, indicating that the CMP images have higher diagnostic efficacy in differentiating type 1 and type 2 tumours. Third, this study combined clinical factors with radiomics features for the construction of the model, enabling a more comprehensive assessment of tumour characteristics and allowing more reliable results to be obtained. Fourth, most previous studies tend to base their texture analysis on one dimension of the tumour, whereas we mainly used all dimensions of the tumour to analyse the tumour and obtain more features. While previous studies mainly extracted a few dozen features, we extracted over 1000 features. Finally, although pRCC is a relatively rare type of RCC, a total of 131 cases of pRCC were collected in our study. To our knowledge, this is the largest sample size to date to study a radiomics-based subtype of pRCC, and our sample was derived from multiple centres.

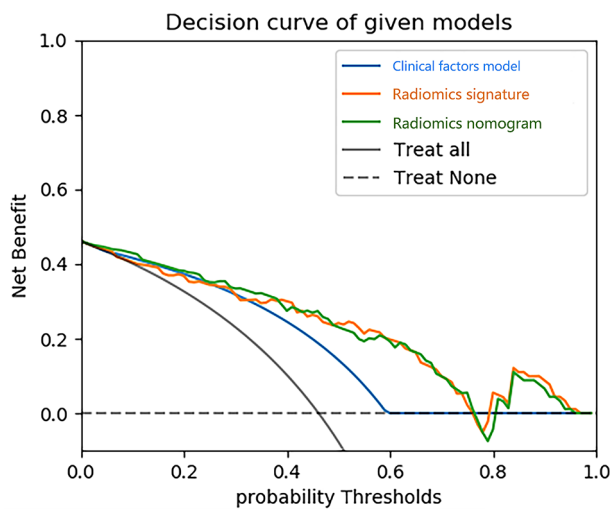


FIGURE 9 | Decision curve analysis (DCA) for three models. The y-axis indicates the net benefit; x-axis indicates probability thresholds. The blue line, yellow line and green line represent net benefit of the clinical factors model, the radiomics signature and the radiomics nomogram, respectively. Both the radiomics nomogram and the radiomics signature showed a higher overall net benefit in differentiating type 1 from type 2 than the clinical factors model.

There are several limitations to our current study. First, this study was a retrospective study, which may introduce bias in the selection of the sample and overestimation of diagnostic accuracy, so external validation may be included in subsequent studies. Second, our study only extracted radiomics features from the CMP and NP images for tumour analysis, and in the future more features may be extracted from the four-phase images of CT to obtain more radiomics information of the tumour. Third, in this study, we used a variety of CT scanners from different suppliers, and although we have normalized the images before extracting the features, there is still the potential for error in the experiment. Fourth, manual segmentation of 3D ROI is both time-consuming and complicated, especially for tumours with unclear borders. Further research should focus on developing an automatic segmentation method for renal tumours with better reliability and reproducibility. Final, the primary target of this study was pRCC and did not include other types of renal tumours. In subsequent studies, we will collect more cases to

build a complete model for differentiating subtypes of renal tumours.

In conclusion, our study demonstrates the importance of combining clinical factors with radiomics features to construct a CT-based radiomics nomogram of CMP and NP images. Our radiomics nomogram can distinguish between pRCC type 1 and type 2 tumours preoperatively and has good diagnostic performance. As a new non-invasion, quantitative diagnostic method, the use of radiomics nomogram needs further validation before it can be used in the clinic.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/supplementary material. Further inquiries can be directed to the corresponding author.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by The First Affiliated Hospital of Anhui Medical University. The ethics committee waived the requirement of written informed consent for participation.

AUTHOR CONTRIBUTIONS

YG: The acquisition of data, analysis of data, and drafting the article. XWWa, SW, YM, CZ, CL and GH: The acquisition of data. YJ: The data analysis. JL: The analysis and interpretation of data. XZ: Drafting and revision of manuscripts. XWWu: Final approval of the version to be submitted. All authors contributed to the article and approved the submitted version.

FUNDING

This work was supported by 2021 Medical Empowerment- Pilot Elite Research Project Special Fund (NO. XM_HR_YXFN_2021_05_19).

ACKNOWLEDGEMENTS

The authors thank General Electric for its technical support.

REFERENCES

1. Siegel RL, Miller K D, Jemal A. Cancer Statistics, 2020. *CA Cancer J Clin* (2020) 70:7–30. doi: 10.3322/caac.21590
2. Motzer RJ, Jonasch E, Michaelson MD, Nandagopal L, Gore JL, George S, et al. NCCN Guidelines Insights: Kidney Cancer, Version 2.2020. *J Natl Compr Canc Netw* (2019) 17:1278–85. doi: 10.6004/jnccn.2019.0054
3. Delahunt B, Eble J. Papillary Renal Cell Carcinoma: A Clinicopathologic and Immunohistochemical Study of 105 Tumors. *Modern Pathol* (1997) 10:537–44.
4. Moch H, Cubilla A, Humphrey P, Reuter V, Ulbright T. The 2016 WHO Classification of Tumours of the Urinary System and Male Genital Organs-Part A: Renal, Penile, and Testicular Tumours. *Eur Urol* (2016) 70:93–105. doi: 10.1016/j.eururo.2016.02.029
5. Pan H, Ye L, Zhu Q, Yang Z, Hu M. The Effect of the Papillary Renal Cell Carcinoma Subtype on Oncological Outcomes. *Sci Rep* (2020) 10:21073. doi: 10.1038/s41598-020-78174-9
6. Wong ECL, Di Lena R, Breau RH, Pouliot F, Finelli A, Lavallee LT, et al. Morphologic Subtyping as a Prognostic Predictor for Survival in Papillary Renal Cell Carcinoma: Type 1 vs. Type 2. *Urol Oncol* (2019) 37:721–6. doi: 10.1016/j.urolonc.2019.05.009
7. Klatte T, Pantuck AJ, Said JW, Seligson DB, Rao NP, LaRochelle JC, et al. Cytogenetic and Molecular Tumor Profiling for Type 1 and Type 2 Papillary Renal Cell Carcinoma. *Clin Cancer Res* (2009) 15:1162–9. doi: 10.1158/1078-0432.CCR-08-1229
8. Egbert ND, Caoili EM, Cohan RH, Davenport MS, Francis IR, Kunju LP, et al. Differentiation of Papillary Renal Cell Carcinoma Subtypes on CT and MRI. *AJR Am J Roentgenol* (2013) 201:347–55. doi: 10.2214/AJR.12.9451
9. Yamada T, Endo M, Tsuboi M, Matsuhashi T, Takase K, Higano S, et al. Differentiation of Pathologic Subtypes of Papillary Renal Cell Carcinoma on CT. *AJR Am J Roentgenol* (2008) 191:1559–63. doi: 10.2214/AJR.07.3181
10. Doshi AM, Ream JM, Kierans AS, Bilbily M, Rusinek H, Huang WC, et al. Use of MRI in Differentiation of Papillary Renal Cell Carcinoma Subtypes: Qualitative and Quantitative Analysis. *AJR Am J Roentgenol* (2016) 206:566–72. doi: 10.2214/AJR.15.15004
11. Ng F, Kozarski R, Ganeshan B, Goh V. Assessment of Tumor Heterogeneity by CT Texture Analysis: Can the Largest Cross-Sectional Area be Used as an Alternative to Whole Tumor Analysis? *Eur J Radiol* (2013) 82:342–8. doi: 10.1016/j.ejrad.2012.10.023
12. Raman SP, Chen Y, Schroeder JL, Huang P, Fishman EK. CT Texture Analysis of Renal Masses: Pilot Study Using Random Forest Classification for Prediction of Pathology. *Acad Radiol* (2014) 21:1587–96. doi: 10.1016/j.acra.2014.07.023
13. Davnall F, Yip C, Ljungqvist G, Selmi M, Ng F, Sanghera B, et al. Assessment of Tumor Heterogeneity: An Emerging Imaging Tool for Clinical Practice? *Insights Imaging* (2012) 3:573–89. doi: 10.1007/s13244-012-0196-6
14. Deng Y, Soule E, Samuel A, Shah S, Cui E, Asare-Sawiri M, et al. CT Texture Analysis in the Differentiation of Major Renal Cell Carcinoma Subtypes and Correlation With Fuhrman Grade. *Eur Radiol* (2019) 29:6922–9. doi: 10.1007/s00330-019-06260-2
15. Wang P, Pei X, Yin XP, Ren JL, Wang Y, Ma LY, et al. Radiomics Models Based on Enhanced Computed Tomography to Distinguish Clear Cell From non-Clear Cell Renal Cell Carcinomas. *Sci Rep* (2021) 11:13729. doi: 10.1038/s41598-021-93069-z
16. Li Q, Liu YJ, Dong D, Bai X, Huang QB, Guo AT, et al. Multiparametric MRI Radiomic Model for Preoperative Predicting WHO/ISUP Nuclear Grade of Clear Cell Renal Cell Carcinoma. *J Magn Reson Imaging* (2020) 52:1557–66. doi: 10.1002/jmri.27182
17. Duan C, Li N, Niu L, Wang G, Zhao J, Liu F, et al. CT Texture Analysis for the Differentiation of Papillary Renal Cell Carcinoma Subtypes. *Abdom Radiol (NY)* (2020) 45:3860–8. doi: 10.1007/s00261-020-02588-2
18. Vendrami CL, Velichko YS, Miller FH, Chatterjee A, Villavicencio CP, Yaghmai V, et al. Differentiation of Papillary Renal Cell Carcinoma Subtypes on MRI: Qualitative and Texture Analysis. *AJR Am J Roentgenol* (2018) 211:1234–45. doi: 10.2214/AJR.17.19213
19. Couvidat C, Eiss D, Verkarre V, Merran S, Correas JM, Mejean A, et al. Renal Papillary Carcinoma: CT and MRI Features. *Diagn Interv Imaging* (2014) 95:1055–63. doi: 10.1016/j.diii.2014.03.013
20. Zhu YH, Wang X, Zhang J, Chen YH, Kong W, Huang YR. Low Enhancement on Multiphase Contrast-Enhanced CT Images: An Independent Predictor of the Presence of High Tumor Grade of Clear Cell Renal Cell Carcinoma. *AJR Am J Roentgenol* (2014) 203:W295–300. doi: 10.2214/AJR.13.12297
21. Li T, Yang K, Stein J D, Nallasamy N. Gradient Boosting Decision Tree Algorithm for the Prediction of Postoperative Intraocular Lens Position in

Cataract Surgery. *Transl Vis Sci Technol* (2020) 9:38. doi: 10.1167/tvst.9.13.38
22. Xiong S, Zhu W, Li X, Yu Y, Yang K, Zhang L, et al. Whether Histologic Subtyping Affect the Oncological Outcomes of Patients With Papillary Renal Cell Carcinoma: Evidence From a Systematic Review and Meta-Analysis. *Transl Androl Urol* (2021) 10:3255–66. doi: 10.21037/tau-21-329
23. Vargas H, Chaim J, Lefkowitz R, Lakhman Y, Zheng J, Moskowitz C, et al. Renal Cortical Tumors: Use of Multiphasic Contrast-Enhanced MR Imaging to Differentiate Benign and Malignant Histologic Subtypes. *Radiology* (2012) 264:779–88. doi: 10.1148/radiol.12110746
24. Rosenkrantz AB, Sekhar A, Genega EM, Melamed J, Babb JS, Patel AD, et al. Prognostic Implications of the Magnetic Resonance Imaging Appearance in Papillary Renal Cell Carcinoma. *Eur Radiol* (2013) 23:579–87. doi: 10.1007/s00330-012-2631-y
25. Waldert M, Haitel A, Marberger M, Katzenbeisser D, Ozsoy M, Stadler E, et al. Comparison of Type I and II Papillary Renal Cell Carcinoma (RCC) and Clear Cell RCC. *BJU Int* (2008) 102:1381–4. doi: 10.1111/j.1464-410X.2008.07999.x
26. Young JR, Coy H, Douek M, Lo P, Sayre J, Pantuck AJ, et al. Type 1 Papillary Renal Cell Carcinoma: Differentiation From Type 2 Papillary RCC on Multiphasic MDCT. *Abdom Radiol (NY)* (2017) 42:1911–8. doi: 10.1007/s00261-017-1091-x
27. Dilauro M, Quon M, McInnes MD, Vakili M, Chung A, Flood TA, et al. Comparison of Contrast-Enhanced Multiphase Renal Protocol CT Versus MRI for Diagnosis of Papillary Renal Cell Carcinoma. *AJR Am J Roentgenol* (2016) 206:319–25. doi: 10.2214/AJR.15.14932
28. Mydlo J, Weinstein R, Misseri R, Axiotis C, Thelmo W. Radiologic, Pathologic and Molecular Attributes of Two Types of Papillary Renal Adenocarcinomas. *Scandinav J Urol Nephrol* (2001) 35:262–9. doi: 10.1080/003655901750425819
29. Li X, Ma Q, Tao C, Liu J, Nie P, Dong C. A CT-Based Radiomics Nomogram for Differentiation of Small Masses (< 4 Cm) of Renal Oncocytoma From Clear Cell Renal Cell Carcinoma. *Abdom Radiol (NY)* (2021) 46:5240–9. doi: 10.1007/s00261-021-03213-6
30. Nie P, Yang G, Wang Z, Yan L, Miao W, Hao D, et al. A CT-Based Radiomics Nomogram for Differentiation of Renal Angiomyolipoma Without Visible Fat From Homogeneous Clear Cell Renal Cell Carcinoma. *Eur Radiol* (2020) 30:1274–84. doi: 10.1007/s00330-019-06427-x
31. Li X, Ma Q, Nie P, Zheng Y, Dong C, Xu W. A CT-Based Radiomics Nomogram for Differentiation of Renal Oncocytoma and Chromophobe Renal Cell Carcinoma With a Central Scar-Matched Study. *Br J Radiol* (2021) 95:20210534. doi: 10.1259/bjr.20210534
32. Wang W, Cao K, Jin S, Zhu X, Ding J, Peng W. Differentiation of Renal Cell Carcinoma Subtypes Through MRI-Based Radiomics Analysis. *Eur Radiol* (2020) 30:5738–47. doi: 10.1007/s00330-020-06896-5
33. Yap FY, Varghese BA, Cen SY, Hwang DH, Lei X, Desai B, et al. Shape and Texture-Based Radiomics Signature on CT Effectively Discriminates Benign From Malignant Renal Masses. *Eur Radiol* (2021) 31:1011–21. doi: 10.1007/s00330-020-07158-0
34. Chen S, Wan Q, Zhou D, Wang T, Hu J, He Y, et al. A Simple-To-Use Nomogram for Predicting the Survival of Early Hepatocellular Carcinoma Patients. *Front Oncol* (2019) 9:584. doi: 10.3389/fonc.2019.00584
35. Huang Y, Liu Z, He L, Chen X, Pan D, Ma Z, et al. Radiomics Signature: A Potential Biomarker for the Prediction of Disease-Free Survival in Early-Stage (I or II) Non-Small Cell Lung Cancer. *Radiology* (2016) 281:947–57. doi: 10.1148/radiol.2016152234

Conflict of Interest: Author JL was employed by GE Healthcare China.

The remaining 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: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

frontiers | Frontiers in Oncology

TYPE Original Research
PUBLISHED 15 July 2022
DOI 10.3389/fonc.2022.918830

OPEN ACCESS

EDITED BY
Nanna Maria Sijtsema,
University of Groningen, Netherlands

REVIEWED BY
Zhenyu Shu,
Zhejiang Provincial People's Hospital, China
Rodrigo Delgadillo,
University of Miami Health System, United States

*CORRESPONDENCE
Yong Lu
paullu2323@vip.sina.com
Jianping Lu
cjr.lujianping@vip.163.com
Fu Shen
ssff_53@163.com

†These authors have contributed equally to this work

SPECIALTY SECTION
This article was submitted to Cancer Imaging and Image-directed Interventions, a section of the journal Frontiers in Oncology

RECEIVED 12 April 2022
ACCEPTED 27 June 2022
PUBLISHED 15 July 2022

CITATION
Jing G, Xing P, Li Z, Ma X, Lu H, Shao C, Lu Y, Lu J and Shen F (2022) Prediction of clinically significant prostate cancer with a multimodal MRI-based radiomics nomogram. *Front. Oncol.* 12:918830. doi: 10.3389/fonc.2022.918830

Prediction of clinically significant prostate cancer with a multimodal MRI-based radiomics nomogram

Guodong Jing[1†], Pengyi Xing[2†], Zhihui Li[3†], Xiaolu Ma[1], Haidi Lu[1], Chengwei Shao[1], Yong Lu[4*], Jianping Lu[1*] and Fu Shen[1*]

[1]Department of Radiology, Changhai Hospital, Shanghai, China, [2]Department of Radiology, 989th Hospital of the joint logistic support force of the Chinese People's Liberation Army, Luoyang, China, [3]Department of Radiology, Ruijin Hospital Luwan Branch, Shanghai Jiaotong University School of Medicine, Shanghai, China, [4]Department of Radiology, Ruijin Hospital, Shanghai Jiaotong University School of Medicine, Shanghai, China

Objective: To develop and validate a multimodal MRI-based radiomics nomogram for predicting clinically significant prostate cancer (CS-PCa).

Methods: Patients who underwent radical prostatectomy with pre-biopsy prostate MRI in three different centers were assessed retrospectively. Totally 141 and 60 cases were included in the training and test sets in cohort 1, respectively. Then, 66 and 122 cases were enrolled in cohorts 2 and 3, as external validation sets 1 and 2, respectively. Two different manual segmentation methods were established, including lesion segmentation and whole prostate segmentation on T2WI and DWI scans, respectively. Radiomics features were obtained from the different segmentation methods and selected to construct a radiomics signature. The final nomogram was employed for assessing CS-PCa, combining radiomics signature and PI-RADS. Diagnostic performance was determined by receiver operating characteristic (ROC) curve analysis, net reclassification improvement (NRI) and decision curve analysis (DCA).

Results: Ten features associated with CS-PCa were selected from the model integrating whole prostate (T2WI) + lesion (DWI) for radiomics signature development. The nomogram that combined the radiomics signature with PI-RADS outperformed the subjective evaluation alone according to ROC analysis in all datasets (all $p<0.05$). NRI and DCA confirmed that the developed nomogram had an improved performance in predicting CS-PCa.

Conclusions: The established nomogram combining a biparametric MRI-based radiomics signature and PI-RADS could be utilized for noninvasive and accurate prediction of CS-PCa.

KEYWORDS

magnetic resonance imaging, nomogram, radiomics, prostate cancer, clinically significant

Introduction

Prostate cancer (PCa) was the second commonest male malignancy in 2020 around the world, causing great harm to the male genitourinary system (1, 2). The descriptive phrase "clinically significant" is broadly utilized for differentiating PCa that might result in morbidity and/or death from harmless PCa subtypes. Such differentiation is critical because "insignificant" PCa not causing harm is commonly encountered (2, 3). Overtreatment of insignificant PCa is considered an important limitation of prostate-specific antigen (PSA) testing.

The European Association of Urology (EAU)-European Association of Nuclear Medicine (EANM)-European Society for Radiotherapy and Oncology (ESTRO)-European Society of Urogenital Radiology (ESUR)-International Society of Geriatric Oncology (SIOG) guidelines (2020 version) for PCa summarized the newest data and advised active surveillance (AS) or watchful waiting (WW) in PCa cases showing a Gleason score (GS) < 7, while clinically significant prostate cancer (CS-PCa) patients with GS ≥ 7 should undergo timely treatment and intervention because of increased risk of progression and short overall survival in clinical practice (2). Therefore, accurately evaluating CS-PCa preoperatively is critical for predicting long-term prognosis and selecting therapeutic options, which would result in more personalized and effective treatments. However, clearly defining CS-PCa is difficult.

The currently applied standard practice of MRI-targeted and template biopsy shows low diagnostic inaccuracy (4, 5). The IP1-PROSTAGRAM trial showed higher detection of CS-PCa with MRI Prostate Imaging–Reporting and Data System (PI-RADS) > 2 in comparison with transrectal ultrasound-guided prostate (TRUS) biopsy (6). However, cancer detection rates (CDRs) are only 6% and 9% for PI-RADS 1 and PI-RADS 2, respectively (4); high-grade cancer may still be missed especially with previous MRI showing suspicious lesions. Patients and clinicians should recognize the considerable uncertainty about prediction (2).

Currently early individualized detection attracts increasing attention. With recent progress in high-throughput analytical tools, radiomics models integrating clinical parameters show overt advantages in generating critical data regarding tissue properties otherwise not detectable by the naked eye (7–13). Indeed, increasing evidence suggests that radiomics could be superior in GS prediction in PCa over routine imaging strategies (14–17). However, which sequence and segmentation method could yield higher clinical benefit have not been evaluated. Thus, a comparison of the predictive capacity of combinations of sequences and various segmentation approaches is urgently required to establish the best radiomics methodology. Therefore, this study aimed to develop a radiomics model considering multimodal MRI and evaluate its predictive potential in CS-PCa with external validation.

Materials and methods

Participants

The current retrospective trial had approval from the local Institutional Review Board (Committee on Ethics of Biomedicine, Changhai Hospital; Committee on Ethics of Biomedicine, Ruijin Hospital Luwan Branch; Committee on Ethics of Biomedicine, 989th Hospital of the joint logistic support force of the Chinese People's Liberation Army).

Individuals who underwent radical prostatectomy with pre-biopsy prostate MRI were searched in the hospitals' databases. Exclusion criteria were: (1) no histological confirmation of PCa with baseline MRI in our institutions (2) no PSA test within 8 weeks prior to baseline MRI; (3) a history of previous therapy for prostate cancer; (4) poor quality of MR images (such as susceptibility artifact); (4) time from baseline MRI to surgical procedure exceeding 12 weeks.

Eventually, 201 cases were identified and enrolled in Changhai hospital from January 2016 to December 2019 as cohort 1. The primary cohort was randomized into the training set (n = 141) and test set (n = 60) at a ratio of 7:3. Next, 66 and 122 cases were enrolled from January 2019 to December 2021 in Ruijin Hospital Luwan Branch and 989th Hospital of the joint logistic support force of the Chinese People's Liberation Army, respectively, as cohorts 2 and 3 (external validation sets), respectively. The study flowchart is shown in **Figure 1A**.

Clinicopathologic data

Clinicopathology factors, including age, BMI, PSA levels, location of each tumor and GS post-prostatectomy, were retrieved from patient records. Radical prostatectomy samples underwent sectioning from apex to base at 3- to 5-mm intervals, and the PCa borders were delineated. All pathological GSs obtained from surgical samples were categorized as follows: GS < 7 [International Society of Urological Pathology (ISUP)] grade 1 PCa considered clinically insignificant; GS ≥ 7 (ISUP grade 2 and above) defined as clinically significant PCa (2, 3).

Imaging and image analysis

Prostate MRI was carried out on a 3.0T MR scanner with an abdominal phase array coil without endorectal coil, following a 4-h fasting period and enema treatment with glycerin (20 ml).

Routine sequences, including sagittal T2WI, axial high-resolution T2WI, axial DWI, axial T1WI and gadolinium contrast-enhanced T1WI, were applied. **Supplementary Table 1** shows axial T2WI and DWI parameters utilized for PI-RADS and radiomics model development.

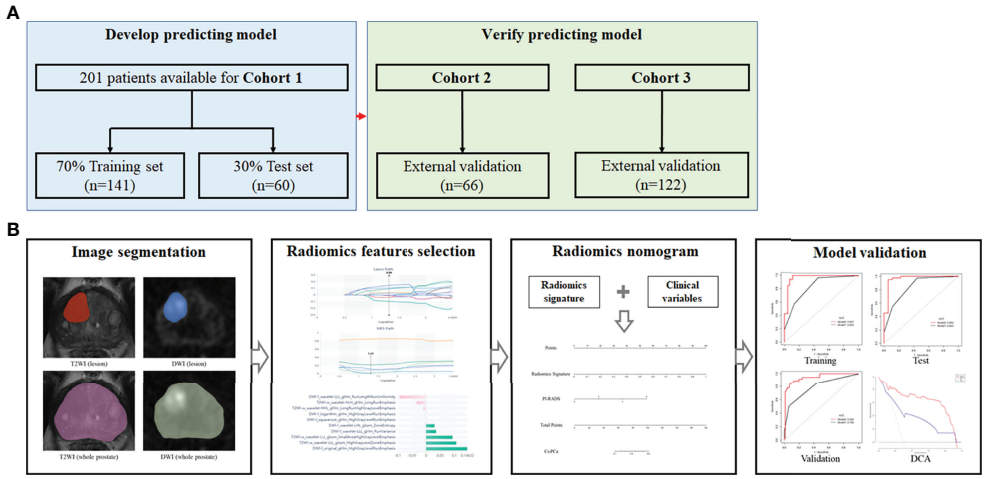


FIGURE 1
Study flowchart and nomogram workflow. **(A)** Study flowchart. Cohort 1, Changhai Hospital; Cohort 2, Ruijin Hospital Luwan Branch; Cohort 3, 989th Hospital of the joint logistic support force of the Chinese People's Liberation Army. **(B)** Workflow for nomogram analysis.

The PI-RADS (version 2.1) score for each case was assessed by three radiologists, including ZH.L., GD.J. and PY.X. with 8, 9 and 12 years of experience in MRI diagnosis, respectively, blinded to pathological data with the exception of tumor location. Any discrepancy among the three observers was resolved by discussion until at least two of them agreed.

Image segmentation

The T2WI and DWI DICOM data acquired pre-biopsy were imported into the Radcloud radiomics platform (Huiying Medical Technology, China. http://radcloud.cn/). Since the original images were obtained from distinct cohorts, their normalization was critical to minimize signal variations for subsequent radiomics analysis (PyRadiomics package, class radiomics.imageoperations.normalizeImage; using the following formula: $f(x)=s(x-\mu_x)/\sigma_x$, where $f(x)$ indicates the normalized intensity; x indicates the original intensity; μ refers to the mean value; σ indicates the variance; s is an optional scaling, by default, it is set to 1. While reserving the diagnostic intensity discrepancy, the signal discrepancy in MR parameters was decreased). In addition, the resampling used (the radiomics.imageoperations.resampleImage function (the default interpolator is Bspline).

Two different segmentation methods were employed: (i) lesion segmentation, which only delineates the border that best fits the lesion area; (ii) whole prostate segmentation, which delineates the whole prostate region. Regions of interest (ROIs) were obtained by manual delineation in individual slices for each MR image (T2WI and DWI with b = 1500 s/mm^2) by the above two segmentation methods in all specimens.

The first radiologist (GD.J.), who was blinded to clinical data, independently carried out the segmentation process for every case on the platform, comprising lesion segmentation and whole prostate segmentation, respectively. Then, ROIs were utilized to obtain volumes of interest (VOIs) in all cohorts. Next, two radiologists (ZH.L. and GD.J.) repeated segmentations for 30 random cases one week later for observer's agreement analysis. Additionally, segmentations were performed under the supervision of a senior radiologist (F.S.), with 14 years of related work experience, for avoiding overt lesion misidentification.

Radiomics feature extraction and reduction

Based on the derived VOIs, four groups of features were obtained: (1) **first-order** features, quantifying voxel intensity distribution on MR scans; (2) shape features, reflecting the 3D features of VOIs; (3) texture features, quantification of region heterogeneity differences, including gray-level co-occurrence, run length, size zone and neighborhood gray-tone difference matrices; (4) higher-order features, encompassing transformed first-order statistics and texture features, including logarithm, exponential, gradient, square, square root, local binary pattern [LBP] and wavelet transformations. In all, 1409 radiomics features were respectively obtained with the above platform from each VOI, based on the Python software package "pyradiomics" (version 6.1). Features complied with the image biomarker standardisation initiative (IBSI) standard (18).

For each cohort, inter- and intra-observer correlation coefficients (ICCs) were determined to assess feature robustness.

Features with ICCs above 0.9 were subsequently utilized for model building, with excellent feature reproducibility.

Based on the two different segmentation methods, ten types of models were obtained: Model 1, DWI (lesion + whole prostate); Model 2, DWI (lesion); Model 3, DWI (whole prostate); Model 4, T2WI (lesion + whole prostate); Model 5, T2WI (lesion); Model 6, T2WI (whole prostate); Model 7, lesion (DWI + T2WI); Model 8, whole prostate (DWI + T2WI); Model 9, whole prostate (DWI) + lesion (T2WI); Model 10, whole prostate (T2WI) + lesion (DWI). For selecting optimal features related to CS-PCa in each model, the variance threshold algorithm, Select-K-best and the least absolute shrinkage and selection operator (LASSO) algorithm were employed.

Radiomics signature building

The selected features (non-zero coefficients in the LASSO algorithm) were employed to develop a radiomics signature for scoring patients in the 10 models, respectively. The predictive value of the radiomics signature was assessed by determining the area under the receiver operator characteristic (ROC) curve (AUC) and Delong test in the training set.

Nomogram model establishment

The predictive abilities of clinical variables and the radiomics signature were assessed by univariate logistic regression analysis. Parameters with $p<0.05$ were subsequently combined to build the nomogram model by multivariable logistic regression analysis ($p<0.05$). Next, the nomogram was examined for performance in each cohort. **Figure 1B shows** the nomogram's workflow.

Statistical analysis

The distribution of continuous data was evaluated by the Kolmogorov-Smirnov test, and the t-test or Wilcoxon test was utilized for comparing these data. The Chi-square or Fisher's exact test was performed for qualitative data analysis. In the variance threshold approach, a threshold of 0.8 was applied, so that the eigenvalues of the variance smaller than 0.8 were removed. The select-K-best approach, which belongs to a single variable feature selection method, retained all features showing $p<0.05$. In the LASSO model, the L1 regularizer constituted the cost function, applying 5 as the cross-validation error and 1000 iterations at most (11–13). Sensitivity, specificity, accuracy, positive predictive value (PPV), negative predictive value (NPV), positive likelihood ratio (PLR) and negative likelihood ratio (NLR) were determined. The goodness of fit for the monogram was assessed by the Hosmer-Lemeshow test. AUC calculation, NRI, and the DeLong test were carried out for comparing the nomogram and PI-RADS V2.1. DCA was carried out for determining the nomogram's clinical usefulness by assessing net benefits at distinct threshold probabilities. The nomogram was examined with R 3.6.3. The remaining data were assessed with SPSS (version 22.0, Inc., Chicago, IL, USA) and MedCalc v19.6.1. $P<0.05$ was deemed statistically significant.

Results

Patient features

Table 1 lists the features of all patients. Clinicopathological parameters were similar in the three cohorts ($p>0.05$). According to pathological GS based on final surgical specimens, 139/201 (69.2%), 45/66 (68.2%) and 86/122 (70.5%) cases were defined as CS-PCa (GS ≥7) in the three cohorts, respectively.

Radiomics features

Feature repeatability based on ICCs in distinct cohorts is shown in Supplementary Figure 1. After inter/intraobserver agreement analysis, 1239/1409 T2WI (lesion) (87.9%), 1243/1409 T2WI (whole) (88.2%), 1096/1409 DWI (lesion) (77.8%) and 1199/1409 DWI (whole) (85.1%) features had excellent robustness and were subsequently utilized in radiomics analysis (inter- and intra-observer ICCs ≥0.9). There was excellent reproducibility for VOI size of lesion segmentation (ICC of T2WI, 0.931; ICC of DWI, 0.910) and whole prostate segmentation (ICC of T2WI, 0.942; ICC of DWI, 0.913). Eventually, optimal features were obtained with the LASSO algorithm for each model and presented in Supplementary Table 2.

ROC analyses of the radiomics signature

The selected features were utilized for the radiomics signature (RS) in each model, respectively. The detailed ROC curve analyses for the 10 models, PSA and PI-RADS are listed in Table 2. ROC curves and their comparisons (Delong test) are shown in Supplementary Figure 2. Among the 10 models, PSA and PI-RADS, whole prostate (T2WI) + lesion (DWI) was determined to have the best performance by ROC curve analysis in the training set (AUC=0.967, specificity=90.9%, sensitivity=92.8% and accuracy=92.2%). The ten optimal features of whole prostate (T2WI) + lesion (DWI) are shown in Figure 2. The correlation analysis of selected features is shown in Supplementary Figure 3.

TABLE 1 Clinical characteristics of patients with prostate cancer in all cohorts.

Characteristic		Cohort 1 (n=201)	Cohort 2 (n = 66)	Cohort 3 (n = 122)	*P* value
Age (year, mean ± SD)		58.547 ± 10.351	59.167 ± 10.181	58.492 ± 10.811	0.902
BMI (kg/m^2, mean ± SD)		23.977 ± 2.706	23.664 ± 2.734	24.442 ± 2.971	0.153
Tumor location (%)	Peripheral zone	99 (49.3)	30 (45.5)	61 (50.0)	0.975
	Transitional zone	63 (31.3)	23 (34.8)	39 (32.0)	
	Peripheral + Transitional zone	39 (19.4)	13 (19.7)	22 (18.0)	
PI-RADS (%)	1	0 (0)	0 (0)	0 (0)	0.957
	2	60 (29.9)	16 (24.2)	36 (29.5)	
	3	34 (16.9)	14 (21.2)	24 (19.7)	
	4	75 (37.3)	24 (36.4)	42 (34.4)	
	5	32 (15.9)	12 (18.2)	20 (16.4)	
Gleason score (%)	<7	62 (30.8)	21 (31.8)	36 (29.5)	0.826
	7 (3 + 4)	48 (23.9)	15 (22.7)	24 (19.7)	
	7 (4 + 3)	42 (20.9)	12 (18.2)	25 (20.5)	
	8 (4 + 4 or 3 + 5 or 5 + 3)	38 (18.9)	11 (16.7)	24 (19.7)	
	9, 10	11 (5.5)	7 (10.6)	13 (10.6)	
Pathological T stage #	T2	136 (67.7)	36 (54.5)	66 (54.1)	0.070
	T3a	34 (16.9)	17 (25.8)	35 (28.7)	
	T3b	31 (15.4)	13 (19.7)	21 (17.2)	
PSA (ng/ml, median IQR) *		12.600 (7.782, 23.280)	12.525 (7.730, 20.578)	13.485 (9.479, 26.995)	0.493

Cohort 1: Training and test sets; Cohort 2: Validation set 1; Cohort 3: Validation set 2.
BMI: Body mass index; PI-RADS: Prostate imaging reporting and data system; PSA: Prostate-specific antigen; IQR: interquartile range.
#The current Union for International Cancer Control (UICC) no longer recognizes pT2 substages.
*Postoperative blood samples.

Logistic regression analysis and nomogram model establishment

Univariate analysis showed the RS, PSA and PI-RADS had significant associations with CS-PCa. Then, predictive model development employed multivariate logistic regression analysis of the selected risk factors (PI-RADS, OR=7.688, p=0.011; RS, OR=7.650×10^5, p=0.002) in the training set (**Table 3**). The radiomics signature also showed a high predictive value for CS-PCa in the test and validation sets (**Table 4**). The regression formula was as follows: prediction probability=−10.943+9.527*RS+1.742*PI-RADS. **Figure 3** shows the monogram.

AUCs for the nomogram were 0.967, 0.964, 0.945 and 0.942 in the training set, test set, validation set 1 and validation set 2, respectively. The Hosmer-Lemeshow test revealed the nomogram model had favorable calibration in all cohorts (p>0.05); details are listed in **Supplementary Table 3**. In all data sets, the nomogram showed elevated AUCs in comparison with the PI-RADS utilized alone. The DeLong test demonstrated significant differences (all p<0.05). NRIs were 0.326 to 0.372, showing the nomogram had an improved clinical utility compared with the PI-RADS for CS-PCa (**Table 5** and **Figure 4**). DCA of validation cohorts confirmed the nomogram's superiority over the PI-RADS at large probability thresholds (**Figure 5**).

Discussion

This work showed that whole prostate (T2WI) + lesion (DWI) was the best segmentation for radiomics model building. According to the AUC, NRI, and DCA results, a radiomics nomogram was developed, which seems to have higher predictive ability than the PI-RADS for CS-PCa in three hospital databases. Clinicians can use this model to more accurately screen patients with CS-PCa before surgery and conduct individuated treatments.

The European Association of Urology' Guidelines on Prostate Cancer recommend active surveillance and follow-up observation for PCa patients with a Gleason score (GS) < 7, whereas clinically significant prostate cancer (CS-PCa) patients with GS ≥ 7 should undergo timely treatment and intervention because of increased risk of disease progression and short overall survival (2). Therefore, accurate clinical assessment is vital for patients to choose the best treatment.

In recent years, multiparametric MRI has been increasingly utilized for PCa's qualitative evaluation (19, 20). The Prostate Imaging Reporting and Data System (PI-RADS) was proposed for better standardization of prostate MRI performance and image interpretation. PI-RADS guidelines v2.1 in 2019 introduced the concept of biparametric magnetic resonance imaging (including T2WI and DWI only) to simplify prostate

TABLE 2 ROC curve analysis in the training set.

	AUC	95% CI	Specificity	Sensitivity	Accuracy	PLR	NLR	PPV	NPV
Model 10	0.967	0.939-0.995	0.909	0.928	0.922	10.206	0.079	0.957	0.851
Model 6	0.929	0.883-0.976	0.841	0.948	0.915	5.962	0.061	0.929	0.881
Model 8	0.920	0.876-0.963	1.000	0.845	0.894	infinity	0.155	1.000	0.746
Model 4	0.911	0.862-0.960	0.704	1.000	0.908	3.385	0.000	0.882	1.000
Model 3	0.909	0.864-0.954	1.000	0.742	0.823	infinity	0.258	1.000	0.638
Model 7	0.903	0.854-0.952	1.000	0.722	0.808	infinity	0.278	1.000	0.620
Model 9	0.899	0.822-0.976	0.864	1.000	0.957	7.333	0.000	0.942	1.000
Model 2	0.888	0.836-0.941	0.886	0.784	0.816	6.895	0.244	0.938	0.650
Model 1	0.837	0.750-0.923	0.727	0.928	0.865	3.402	0.099	0.882	0.820
PI-RADS	0.835	0.766-0.904	0.545	0.969	0.837	2.132	0.057	0.825	0.889
Model 5	0.800	0.707-0.892	0.704	0.866	0.816	2.931	0.190	0.866	0.704
PSA	0.776	0.702-0.851	1.000	0.557	0.695	infinity	0.443	1.000	0.506

Model 1: DWI (lesion + whole prostate).
Model 2: DWI (lesion).
Model 3: DWI (whole prostate).
Model 4: T2WI (lesion + whole prostate).
Model 5: T2WI (lesion).
Model 6: T2WI (whole prostate).
Model 7: lesion (DWI + T2WI).
Model 8: whole prostate (DWI + T2WI).
Model 9: whole prostate (DWI) + lesion (T2WI).
Model 10: whole prostate (T2WI) + lesion (DWI).
AUC, area under the curve; PLR, positive likelihood ratio; NLR, negative likelihood ratio; NPV, negative predictive value; PPV, positive predictive value.

MRI (21). Prostate MRI categorizes suspected PCa into low- and high risk types, considering risk scores from 1 to 5. PI-RADS grades of 3-5 are recommended to undergo MRI-directed biopsy (22), which could decrease the amounts of avoidable biopsies. However, such approach may miss a small portion of CS-PCas (23), due to low cancer detection rates, i.e., only 6% (0-20%) and 9% (5-13%) for PI-RADS 1 and PI-RADS 2, respectively, in patient level analysis (4). In addition, the commonly used clinical application of the PSA shows limitations, including overdiagnosis and resulting overtreatment (24, 25). Therefore, novel methods for timely and accurate PCa risk stratification are urgently required for improving patient prognosis.

Radiomics is a novel approach that converts traditional medical imaging findings into data mining and high-throughput quantitative analysis. The analysis approach of radiomics provides a non-invasive tool for evaluating the biological characteristics and heterogeneity of prostate cancer more comprehensively and quantitatively than morphological

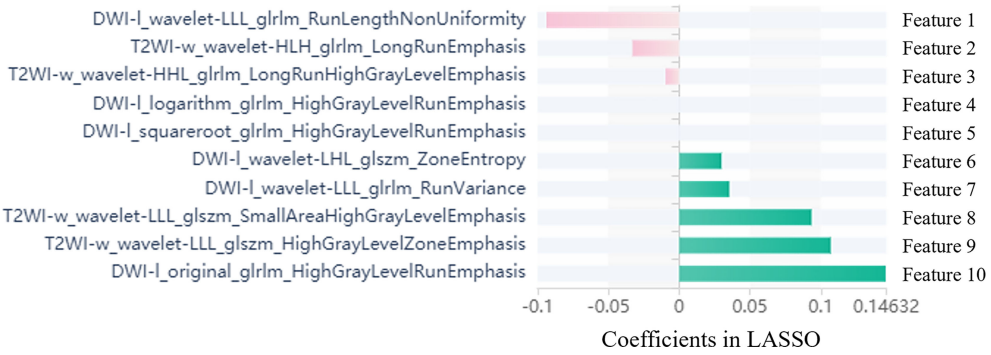


FIGURE 2
Selected radiomics features with associated coefficients in the LASSO model. DWI-l: lesion segmentation of DWI; T2WI-w: whole prostate segmentation of T2WI. GLSZM: Gray level size zone matrix; GLDM: Gray Level dependence; GLRLM: Gray level run length matrix; NGTDM: Neighborhood gray tone difference matrix; Wavelet: The wavelet transform decomposes the tumor area image into low-frequency components (L) or high-frequency components (H) in the x, y, and z axes.

TABLE 3 Univariate and multivariate logistic regression analyses in the training set.

	Univariable analyses		Multivariable analyses	
	OR (95% CI)	*P* value	OR (95% CI)	*P* value
Age (year)	0.967 (0.933, 1.003)	0.068	/	/
BMI (kg/m^2)	0.891 (0.778, 1.021)	0.097	/	/
PSA	1.172 (1.080, 1.271)	**<0.001**	1.391 (0.991, 1.952)	0.056
Location	1.616 (0.950, 2.751)	0.077	/	/
PI-RADS	7.120 (3.569, 14.202)	**<0.001**	7.688 (1.594, 37.085)	**0.011**
Radiomics signature	4.517×10^4 (899.309, 2268910.875)	**<0.001**	7.650×10^5 (128.450, 4.560×10^9)	**0.002**

OR, odds ratio.
Bold values mean p<0.05.

TABLE 4 Multivariate logistic regression analysis in the test and validation sets.

	Test set (n=60)		Validation set 1 (n=66)		Validation set 2 (n=122)	
	OR (95% CI)	*P* value	OR (95% CI)	*P* value	OR (95% CI)	*P* value
PSA	1.155 (0.904, 1.475)	0.250	1.185 (0.941, 1.492)	0.150	1.001 (0.996, 1.007)	0.627
PI-RADS	14.204 (1.150, 175.495)	**0.039**	4.751 (0.916, 24.655)	0.064	4.065 (1.833, 9.017)	**0.001**
Radiomics signature	9.420×10^6 (1.206, 7.351×10^{13})	**0.047**	11624.241 (6.780, 1.993×10^7)	**0.014**	1.021 (1.011, 1.031)	**<0.001**

OR, odds ratio.
Bold values mean p<0.05.

visual representation. Several studies have demonstrated that the current MRI-related radiomics application could be widely used for GS assessment in PCa (14–17). Although they found that multiparametric radiomics models show great potential in predicting GS, there is currently no comparative assessment of different combinations of sequences and patterns of segmentation for model building, which can yield higher clinical benefit for CS-PCa with external validation.

The most valuable aspect of the present study is the multi-pattern approach that enhances MRI-based radiomics by mining complementary information provided by multi-pattern MRI and considering the heterogeneity of tumors for predicting differential

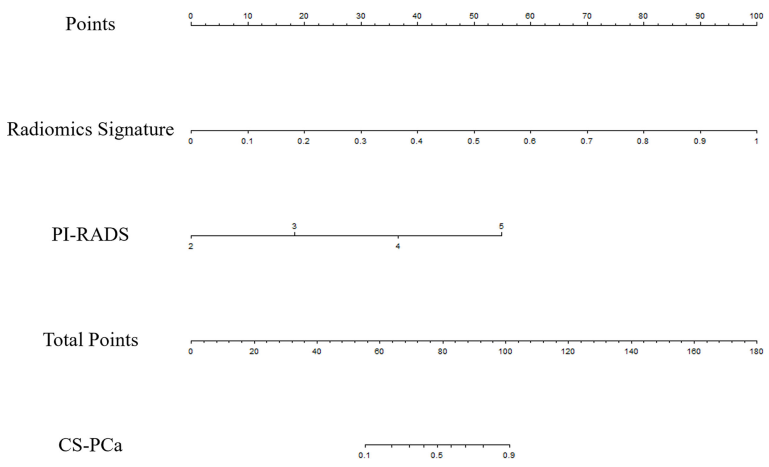


FIGURE 3
The nomogram developed using the training set for predicting CS-PCa, based on the radiomics signature and PI-RADS.

TABLE 5 ROC curve analysis and comparison of prediction models in all data sets.

		AUC	95% CI	Specificity	Sensitivity	Accuracy	*P* value	NRI
Training set (n=141)	PI-RADS	0.835	0.766-0.904	0.545	0.969	0.837	<0.001	0.372
	Nomogram	0.967	0.930-1.000	0.886	1.000	0.964		
Test set (n=60)	PI-RADS	0.843	0.737-0.948	0.556	0.976	0.850	0.01	0.365
	Nomogram	0.964	0.904-1.000	0.944	0.952	0.950		
Validation set 1 (n=66)	PI-RADS	0.824	0.719-0.929	0.524	0.978	0.833	0.01	0.333
	Nomogram	0.945	0.869-1.000	0.857	0.978	0.939		
Validation set 2 (n=122)	PI-RADS	0.796	0.710-0.882	0.942	0.500	0.812	<0.001	0.326
	Nomogram	0.942	0.896-0.987	0.907	0.861	0.893		

AUC, area under the curve; NRI, net reclassification index.

features involved in CS-PCa (26). Among the factors that affecting radiomics assessment, segmentation represents the first critical step of imaging processing. Manual ROI drawing represents the most conventionally utilized segmentation method nowadays (27). Most prior studies assessed lesion-derived radiomics models with AUCs from 0.648 to 0.910 (14–16). Gong et al. (17) investigated the potential of prostate gland radiomic features in identifying GS, with an AUC of 0.794 in the validation cohort. However, the various patterns of segmentation for model building have been less discussed and requires further quantitative assessment. Therefore, in this study, we established multi-pattern segmentations, including prostate lesions (T2WI or DWI), whole prostate (T2WI or DWI), and the combination of different methods, which were applied for radiomics analysis to detect clinically significant prostate cancer. Following feature selection, 10 optimal features based on the whole prostate (T2WI) + lesion (DWI) model were selected to develop a

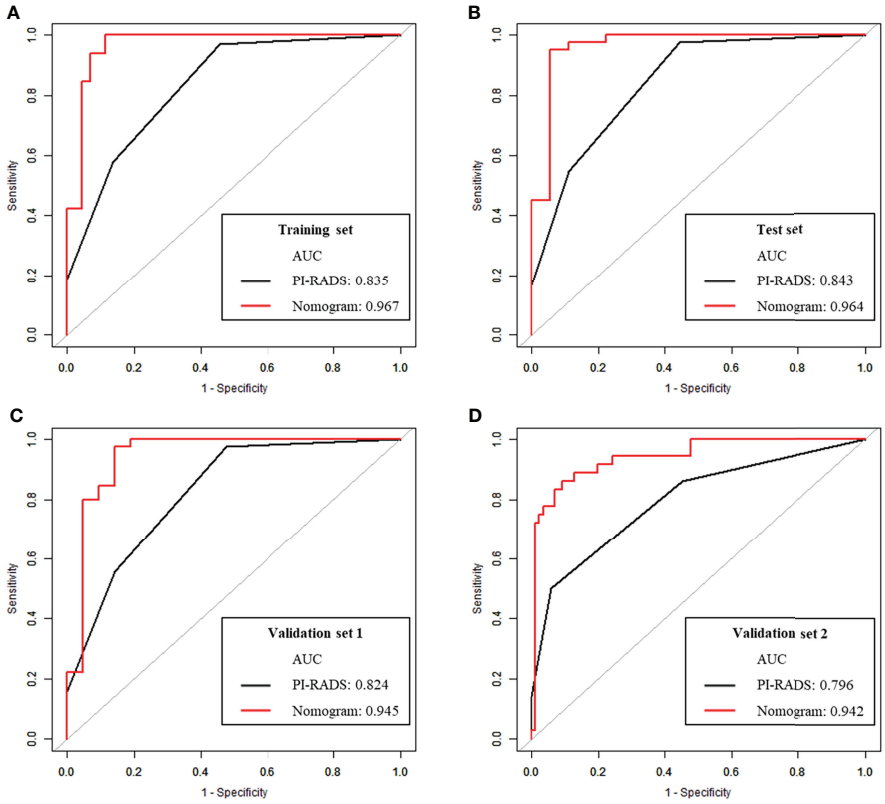


FIGURE 4
ROC curve analysis of the nomogram and PI-RADS for CS-PCa prediction. **(A)** In the training set. **(B)** In the test set. **(C)** In validation set 1. **(D)** In validation set 2.

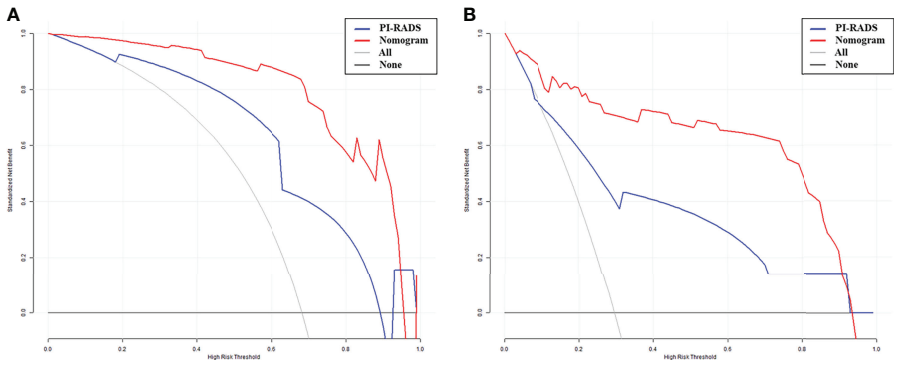


FIGURE 5
Decision curve analysis (DCA) of the nomogram and PI-RADS models. X-axis, risk threshold of CS-PCa; Y-axis, net benefit. Black line, all cases assumed to be clinically insignificant; gray line, all cases considered clinically significant. The nomogram model had enhanced net benefit compared with the PI-RADS at large probability thresholds (0.0-0.9). **(A)** In validation set 1. **(B)** In validation set 2.

radiomics signature for preoperative prediction of CS-PCa, with favorable discriminatory potential (Table 2). A possible explanation is that the whole prostate (T2WI) model contained phenotypic features for the entire prostate, while the lesion (DWI) model involved heterogeneous data describing microcirculation for the focal lesion.

Since the PI-RADS v2.1 introduced the biparametric prostate MRI, which was widely recognized by radiologists and urologists, several prior studies extracted radiomic features from T2W and DWI (14, 16, 28, 29). Thus, combining biparametric MRI and deep mining of correlations among distinct radiomics features could allow a comprehensive assessment of tumor heterogeneity, which might increase the predictive efficiency and potentially guide in distinguishing cases requiring individualized treatments (30–32).

The second noteworthy aspect of the current study is that the radiomics signature and PI-RADS were combined to develop a radiomics nomogram with improved discriminatory ability, which constitutes a visualization tool to predict CS-PCa. Zhang et al. reported a radiomics nomogram model, which did not incorporate the PI-RADS v2 score, showed an AUC of 0.910 (15). Montoya et al. reported that the use of radiomics model failed to outperform PI-RADS v2.1 scales and their combination did not lead to further performance gains (AUC=0.830, p>0.05) (28). However, our results showed that the nomogram model incorporated subjective evaluation exhibited a higher AUC compared with the PI-RADS alone (p<0.05) in all cohorts. NRI analysis determined the predictive value was improved by using the nomogram in lieu of the traditional PI-RADS v2.1, and good clinical usefulness was demonstrated by DCA. These data suggest the developed nomogram could be utilized to guide clinical practice.

The third vital aspect of this study is that we had two actual external validation datasets, adding value to our previous reports. Using external cohorts is very helpful for overcoming the weakness that the developed model has no exposure to a validation cohort in the training phase.

However, the current study still had some limitations. First, an important drawback of the current retrospective trial was its relatively small sample size. This implies selection bias and low generalizability of the obtained results, although external validation cohorts were analyzed. Therefore, larger multicenter studies are warranted for reducing the effects of selection bias on model accuracy. Secondly, the imaging segmentation approach was manual rather than semi-automatic/automatic delineation, favoring subjective errors, with no suitability for large data processing (33). Thirdly, the current work failed to develop and validate deep learning tools for the prediction of CS-PCa, which may show more advantages and deserve further investigation (34).

Conclusion

Overall, based on preoperative biparametric MRI [whole prostate (T2WI) + lesion (DWI)], a quantitative radiomics signature was built. The nomogram model combined with the radiomics signature and PI-RADS had improved clinical benefit in comparison with the subjective evaluation only in predicting clinically significant prostate cancer.

Data availability statement

The original contributions presented in the study are included in the article/**Supplementary Material**. Further inquiries can be directed to the corresponding authors.

Ethics statement

Written informed consent was not obtained from the individual(s) for the publication of any potentially identifiable images or data included in this article.

Author contributions

JL, CS, YL, and FS conceived the project. XM acquired the data. PX and GJ analyzed and interpreted the patient data regarding radiomics features. ZL and HL performed statistical analyses and feature extraction. ZL, PX, and GJ was a major contributor in writing the manuscript. All authors contributed to the article and approved the submitted version.

Funding

The present study was supported by the Project of the Action Plan of Major Diseases Prevention and Treatment (2017ZX01001-S12), the Special Project of Integrated Traditional Chinese and Western Medicine in General Hospitals of Shanghai (ZHYY-ZXYJHZX-201901), and the Changhai hospital discipline construction project (2020YXK034). The funders developed the main idea and designed the study.

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

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fonc.2022.918830/full#supplementary-material.

SUPPLEMENTARY TABLE 1
Main MRI sequences and parameters.

SUPPLEMENTARY TABLE 2
Description of the selected radiomics features with associated feature groups and filters.

SUPPLEMENTARY TABLE 3
Prediction performances of the nomogram model in all data sets.

SUPPLEMENTARY FIGURE 1
Repeatability of radiomics features based on ICCs in different cohorts.

SUPPLEMENTARY FIGURE 2
ROC curves for 10 models, PSA and PI-RADS, and comparisons (Delong test).

SUPPLEMENTARY FIGURE 3
Correlation analysis of the selected ten optimal features of the whole prostate (T2WI) + lesion (DWI) model.

References

1. Culp MB, Soerjomataram I, Efstathiou JA, Bray F, Jemal A. Recent global patterns in prostate cancer incidence and mortality rates. *Eur Urol* (2020) 77(1):38–52. doi: 10.1016/j.eururo.2019.08.005

2. Mottet N, van den Bergh RCN, Briers E, Van den Broeck T, Cumberbatch MG, De Santis M, et al. EAU-EANM-ESTRO-ESUR-SIOG guidelines on prostate cancer-2020 update. part 1: Screening, diagnosis, and local treatment with curative intent. *Eur Urol* (2021) 79(2):243–62.doi: 10.1016/j.eururo.2020.09.042

3. Goel S, Shoag JE, Gross MD, Al Hussein Al Awamlh B, Robinson B. Concordance between biopsy and radical prostatectomy pathology in the era of targeted biopsy: a systematic review and meta-analysis. *Eur Urol Oncol* (2020) 3 (1):10–20. doi: 10.1016/j.euo.2019.08.001.

4. Goel S, Shoag JE, Gross MD, et al. Concordance between biopsy and radical prostatectomy pathology in the era of targeted biopsy: a systematic review and meta-analysis. *Eur Urol Oncol* (2020) 3(1):10–20. doi: 10.1016/j.euo.2019.08.001

5. Oerther B, Engel H, Bamberg F, Sigle A, Gratzke C, Benndorf M. Cancer detection rates of the PI-RADSv2.1 assessment categories: systematic review and meta-analysis on lesion level and patient level. *Prostate Cancer Prostatic Dis* (2021). doi: 10.1038/s41391-021-00417-1

6. Eldred-Evans D, Burak P, Connor MJ, Day E, Evans M, Fiorentino F, et al. Population-based prostate cancer screening with magnetic resonance imaging or ultrasonography: the IP1-PROSTAGRAM study. *JAMA Oncol* (2021) 7(3):395–402. doi: 10.1001/jamaoncol.2020.7456

7. Lambin P, Rios-Velazquez E, Leijenaar R, Carvalho S, van Stiphout RG, Granton P, et al. Radiomics: extracting more information from medical images using advanced feature analysis. *Eur J Cancer* (2012) 48(4):441–6. doi: 10.1016/j.ejca.2011.11.036

8. Kumar V, Gu Y, Basu S, Berglund A, Eschrich SA, Schabath MB, et al. Radiomics: the process and the challenges. *Magn Reson Imaging* (2012) 30 (9):1234–48. doi: 10.1016/j.mri.2012.06.010

9. Aerts HJ, Velazquez ER, Leijenaar RT, Parmar C, Grossmann P, Carvalho S, et al. Decoding tumour phenotype by noninvasive imaging using a quantitative radiomics approach. *Nat Commun* (2014) 5:4006. doi: 10.1038/ncomms5006

10. Gillies RJ, Kinahan PE, Hricak H. Radiomics: images are more than pictures, they are data. *Radiology* (2016) 278:563–77. doi: 10.1148/radiol.2015151169

11. Li Z, Li S, Zang S, Ma X, Chen F, Xia Y, et al. Predicting treatment response to neoadjuvant chemoradiotherapy in rectal mucinous adenocarcinoma using an MRI-based radiomics nomogram. *Front Oncol* (2021) 11:671636. doi: 10.3389/fonc.2021.671636

12. Ma X, Shen F, Jia Y, Xia Y, Li Q, Lu J. MRI-Based radiomics of rectal cancer: preoperative assessment of the pathological features. *BMC Med Imaging* (2019) 19 (1):86. doi: 10.1186/s12880-019-0392-7

13. Liu M, Ma X, Shen F, Xia Y, Jia Y, Lu J. MRI-Based radiomics nomogram to predict synchronous liver metastasis in primary rectal cancer patients. *Cancer Med* (2020) 9(14):5155–63. doi: 10.1002/cam4.3185

14. Toivonen J, Montoya Perez I, Movahedi P, Merisaari H, Pesola M, Taimen P, et al. Radiomics and machine learning of multisequence multiparametric prostate MRI: Towards improved non-invasive prostate cancer characterization. *PloS One* (2019) 14(7):e0217702. doi: 10.1371/journal.pone.0217702

15. Zhang GM, Han YQ, Wei JW, Qi YF, Gu DS, Lei J, et al. Radiomics based on MRI as a biomarker to guide therapy by predicting upgrading of prostate cancer from biopsy to radical prostatectomy. *J Magn Reson Imaging* (2020) 52(4):1239–48. doi: 10.1002/jmri.27138

16. Chaddad A, Kucharczyk MJ, Niazi T. Multimodal radiomic features for the predicting gleason score of prostate cancer. *Cancers (Basel)* (2018) 10(8):249. doi: 10.3390/cancers10080249

17. Gong L, Xu M, Fang M, He B, Li H, Fang X, et al. The potential of prostate gland radiomic features in identifying the gleason score. *Comput Biol Med* (2022) 144:105318. doi: 10.1016/j.compbiomed.2022.105318

18. Zwanenburg A, Vallières M, Abdalah MA, Aerts HJWL, Andrearczyk V, Apte A, et al. The image biomarker standardization initiative: standardized quantitative radiomics for high-throughput image-based phenotyping. *Radiology* (2020) 295(2):328–38. doi: 10.1148/radiol.2020191145

19. Ueno Y, Tamada T, Bist V, Reinhold C, Miyake H, Tanaka U, et al. Multiparametric magnetic resonance imaging: current role in prostate cancer management. *Int J Urol* (2016) 23(7):550–7. doi: 10.1111/iju.13119

20. Aydın H, Kızılgöz V, Tekin BO. Overview of current multiparametric magnetic resonance imaging approach in the diagnosis and staging of prostate cancer. *Kaohsiung J Med Sci* (2015) 31(4):167–78. doi: 10.1016/j.kjms.2015.01.002

21. Turkbey B, Rosenkrantz AB, Haider MA, Padhani AR, Villeirs G, Macura KJ, et al. Prostate imaging reporting and data system version 2.1: 2019 update of prostate imaging reporting and data system version 2. *Eur Urol* (2019) 76(3):340–51. doi: 10.1016/j.eururo.2019.02.033

22. Padhani AR, Barentsz J, Villeirs G, Rosenkrantz AB, Margolis DJ, Turkbey B, et al. PI-RADS steering committee: the PI-RADS multiparametric MRI and MRI-directed biopsy pathway. *Radiology* (2019) 292(2):464–74. doi: 10.1148/radiol.2019182946

23. Schoots IG, Padhani AR, Rouvière O, Barentsz JO, Richenberg J. Analysis of magnetic resonance imaging-directed biopsy strategies for changing the paradigm of prostate cancer diagnosis. *Eur Urol Oncol* (2020) 3(1):32–41. doi: 10.1016/j.euo.2019.10.001

24. Salami SS, Vira MA, Turkbey B, khoury M, Yaskiv O, Villani R, et al. Multiparametric magnetic resonance imaging outperforms the prostate cancer prevention trial risk calculator in predicting clinically significant prostate cancer. *Cancer* (2014) 120(18):2876–82. doi: 10.1002/cncr.28790

25. Bhat NR, Vetter JM, Andriole GL, Shetty AS, Ippolito JE, Kim EH. Magnetic resonance imaging-defined prostate-specific antigen density significantly improves the risk prediction for clinically significant prostate cancer on biopsy. *Urology* (2019) 126:152–7. doi: 10.1016/j.urology.2018.12.010

26. Li Z, Dai H, Liu Y, Pan F, Yang Y, Zhang M. Radiomics analysis of multi-sequence MR images for predicting microsatellite instability status preoperatively in rectal cancer. *Front Oncol* (2021) 11:697497. doi: 10.3389/fonc.2021.697497

27. Rizzo S, Botta F, Raimondi S, Origgi D, Fanciullo C, Morganti AG, et al. Radiomics: the facts and the challenges of image analysis. *Eur Radiol Exp* (2018) 2 (1):36. doi: 10.1186/s41747-018-0068-z

28. Montoya Perez I, Merisaari H, Jambor I, Ettala O, Taimen P, Knaapila J, et al. Detection of prostate cancer using biparametric prostate MRI, radiomics, and kallikreins: a retrospective multicenter study of men with a clinical suspicion of prostate cancer. *J Magn Reson Imaging* (2022) 55(2):465–77. doi: 10.1002/jmri.27811

29. Zhang L, Jiang D, Chen C, Yang X, Lei H, Kang Z, et al. Development and validation of a multiparametric MRI-based radiomics signature for distinguishing between indolent and aggressive prostate cancer. *Br J Radiol* (2022) 95 (1131):20210191. doi: 10.1259/bjr.20210191

30. Scialpi M. Simplified PI-RADS-based biparametric MRI: a rationale for detecting and managing prostate cancer. *Clin Imaging* (2021) 80:290–1. doi: 10.1016/j.clinimag.2021.07.024

31. Zhang L, Zhe X, Tang M, Zhang J, Ren J, Zhang X, et al. Predicting the grade of prostate cancer based on a biparametric MRI radiomics signature. *Contrast Media Mol Imaging* (2021) 2021:7830909. doi: 10.1155/2021/7830909

32. Xu L, Zhang G, Shi B, Liu Y, Zou T, Yan W, et al. Comparison of biparametric and multiparametric MRI in the diagnosis of prostate cancer. *Cancer Imaging* (2019) 19(1):90. doi: 10.1186/s40644-019-0274-9

33. Price WN2nd, Cohen IG. Privacy in the age of medical big data. *Nat Med* (2019) 25:37–43. doi: 10.1038/s41591-018-0272-7

34. Hosny A, Parmar C, Quackenbush J, Schwartz LH, Aerts H. Artificial intelligence in radiology. *Nat Rev Cancer* (2018) 18:500–10. doi: 10.1038/s41568-018-0016-5

frontiers | Frontiers in Oncology

TYPE Original Research
PUBLISHED 02 August 2022
DOI 10.3389/fonc.2022.967360

OPEN ACCESS

EDITED BY
Xue Meng,
Shandong Cancer Hospital, China

REVIEWED BY
Xin Jiang,
The First Hospital of Jilin University, China
Wencheng Zhang,
Tianjin Medical University Cancer Institute and Hospital, China
Guang Han,
Hubei Cancer Hospital, China

*CORRESPONDENCE
Haitao Jiang
jianght@zjcc.org.cn
Yujin Xu
xuyj@zjcc.org.cn

†These authors have contributed equally to this work and share first authorship

SPECIALTY SECTION
This article was submitted to Cancer Imaging and Image-directed Interventions, a section of the journal Frontiers in Oncology

RECEIVED 12 June 2022
ACCEPTED 05 July 2022
PUBLISHED 02 August 2022

CITATION
Yang H, Wang L, Shao G, Dong B, Wang F, Wei Y, Li P, Chen H, Chen W, Zheng Y, He Y, Zhao Y, Du X, Sun X, Wang Z, Wang Y, Zhou X, Lai X, Feng W, Shen L, Qiu G, Ji Y, Chen J, Jiang Y, Liu J, Zeng J, Wang C, Zhao Q, Yang X, Hu X, Ma H, Chen Q, Chen M, Jiang H and Xu Y (2022) A combined predictive model based on radiomics features and clinical factors for disease progression in early-stage non-small cell lung cancer treated with stereotactic ablative radiotherapy. *Front. Oncol.* 12:967360. doi: 10.3389/fonc.2022.967360

A combined predictive model based on radiomics features and clinical factors for disease progression in early-stage non-small cell lung cancer treated with stereotactic ablative radiotherapy

Hong Yang[1†], Lin Wang[2†], Guoliang Shao[1], Baiqiang Dong[3], Fang Wang[1], Yuguo Wei[4], Pu Li[5], Haiyan Chen[1], Wujie Chen[1], Yao Zheng[1], Yiwei He[1], Yankun Zhao[1], Xianghui Du[6], Xiaojiang Sun[6], Zhun Wang[6], Yuezhen Wang[6], Xia Zhou[6], Xiaojing Lai[6], Wei Feng[6], Liming Shen[6], Guoqing Qiu[6], Yongling Ji[6], Jianxiang Chen[6], Youhua Jiang[7], Jinshi Liu[7], Jian Zeng[7], Changchun Wang[7], Qiang Zhao[7], Xun Yang[7], Xiao Hu[6], Honglian Ma[6], Qixun Chen[7], Ming Chen[3], Haitao Jiang[1*] and Yujin Xu[6*]

[1]Department of Radiology, Cancer Hospital of the University of Chinese Academy of Sciences (Zhejiang Cancer Hospital), Institute of Basic Medicine and Cancer (IBMC), Chinese Academy of Sciences, Hangzhou, China, [2]Shaoxing University School of Medicine, Shaoxing, China, [3]Department of Radiation Oncology, Sun Yat-sen University Cancer Center, State Key Laboratory of Oncology in South China, Collaborative Innovation Center for Cancer Medicine, Sun Yat-sen University, Guangzhou, China, [4]Precision Health Institution, General Electric (GE) Healthcare, Hangzhou, China, [5]Department of Radiation Physics, Cancer Hospital of the University of Chinese Academy of Sciences (Zhejiang Cancer Hospital), Institute of Basic Medicine and Cancer (IBMC), Chinese Academy of Sciences, Hangzhou, China, [6]Department of Radiation Oncology, Cancer Hospital of the University of Chinese Academy of Sciences (Zhejiang Cancer Hospital), Institute of Basic Medicine and Cancer (IBMC), Chinese Academy of Sciences, Hangzhou, China, [7]Department of Thoracic Surgery, Cancer Hospital of the University of Chinese Academy of Sciences (Zhejiang Cancer Hospital), Institute of Basic Medicine and Cancer (IBMC), Chinese Academy of Sciences, Hangzhou, China

Purpose: To accurately assess disease progression after Stereotactic Ablative Radiotherapy (SABR) of early-stage Non-Small Cell Lung Cancer (NSCLC), a combined predictive model based on pre-treatment CT radiomics features and clinical factors was established.

Methods: This study retrospectively analyzed the data of 96 patients with early-stage NSCLC treated with SABR. Clinical factors included general information (e.g. gender, age, KPS, Charlson score, lung function, smoking status), pre-treatment lesion status (e.g. diameter, location, pathological type, T stage), radiation parameters (biological effective dose, BED), the type of peritumoral

radiation-induced lung injury (RILI). Independent risk factors were screened by logistic regression analysis. Radiomics features were extracted from pre-treatment CT. The minimum Redundancy Maximum Relevance (mRMR) and the Least Absolute Shrinkage and Selection Operator (LASSO) were adopted for the dimensionality reduction and feature selection. According to the weight coefficient of the features, the Radscore was calculated, and the radiomics model was constructed. Multiple logistic regression analysis was applied to establish the combined model based on radiomics features and clinical factors. Receiver Operating Characteristic (ROC) curve, DeLong test, Hosmer-Lemeshow test, and Decision Curve Analysis (DCA) were used to evaluate the model's diagnostic efficiency and clinical practicability.

Results: With the median follow-up of 59.1 months, 29 patients developed progression and 67 remained good controlled within two years. Among the clinical factors, the type of peritumoral RILI was the only independent risk factor for progression ($P< 0.05$). Eleven features were selected from 1781 features to construct a radiomics model. For predicting disease progression after SABR, the Area Under the Curve (AUC) of training and validation cohorts in the radiomics model was 0.88 (95%CI 0.80-0.96) and 0.80 (95%CI 0.62-0.98), and AUC of training and validation cohorts in the combined model were 0.88 (95%CI 0.81-0.96) and 0.81 (95%CI 0.62-0.99). Both the radiomics and the combined models have good prediction efficiency in the training and validation cohorts. Still, DeLong test shows that there is no difference between them.

Conclusions: Compared with the clinical model, the radiomics model and the combined model can better predict the disease progression of early-stage NSCLC after SABR, which might contribute to individualized follow-up plans and treatment strategies.

KEYWORDS

non-small cell lung cancer, stereotactic ablative radiotherapy, progression, radiomics, predictive model

Introduction

Lung cancer is the second incidence of diagnosed tumor and is the primary leading cause of cancer-related deaths worldwide (1). Non-small cell lung cancer (NSCLC) accounts for 80-85% of lung cancer. Currently, surgery remains the standard of care for NSCLC (2). For patients who are medically inoperable due to their existing severe chronic disease or their rejection of surgery, the treatment of stereotactic ablative radiotherapy (SABR) has been established as the standard alternative therapy (3–5).

SABR is a non-invasive external beam radiation modality which could facilitate the delivery of ablative doses to the tumor, sparing the surrounding normal tissues over a limited number of fractions. Previous studies had shown that the local control rate could reach 85%~98%, and the 3-year overall survival (OS) can get 48%~65% after SABR in early-stage NSCLC (6–8). However, the patients still had the risk of locoregional recurrence (4% ~14%) and distant metastasis (13%~23%) after SABR (9, 10), which is a great challenge for clinicians. Chemotherapy is not ideal because most patients who receive SABR cannot take the risk of surgery due to poor cardiopulmonary function and aged physical condition. In the era of immunotherapy, Immune Checkpoint Inhibitors (ICI) have represented a revolution in treating various stages of NSCLC. The addition of ICI to SABR seems promising, and several multicenter, prospective, randomized controlled clinical trials are underway. A systematic literature review indicated that the ICI-SABR combination has a good safety profile and achieves high rates of local control and greater chances of obtaining abscopal responses than SABR alone, with a relevant impact on progression-free survival (PFS) (11). However, most patients with early-stage NSCLC could be cured after SABR alone, and

they shall be waived from the suffering of the harm of systematic therapy. For this reason, finding out those patients will have a high risk of disease progression is becoming essential. Therefore, establishing an effective predictive model to assess the risk of progression and survival probability of early-stage NSCLC patients is of great significance for treatment plan selection or the individual design of follow-up.

Resulting from the heterogeneity of tumors, the growth rate, invasive ability, drug sensitivity, and prognosis of tumors can be different, and the divergence can limit the usefulness of molecular testing-based tissue biopsies (12). Radiomics extracts quantitative features from Computed Tomography (CT), Magnetic Resonance Imaging (MRI), Positron Emission Tomography (PET), and other medical images with high throughput by utilizing computer software (13). Through statistical or computer learning methods, the characteristics most related to clinical results are selected to establish models, which can provide valuable predictive information for the diagnosis and treatment of diseases, and can provide information on tumor cells more comprehensively, systematically, and deeply (12–17). In this study, the radiomics method was used to deeply mine the pre-treatment CT radiomics features, combined with clinical factors, to construct and validate a predictive model for the disease progression of early-stage NSCLC after SABR, providing a feasible and practical reference for clinical guidance of individualized treatment of patients.

Materials and methods

This retrospective study was approved by the ethics committees of Cancer Hospital of the University of Chinese Academy of Sciences (Zhejiang Cancer Hospital). The requirement for informed consent was waived.

Patient data

The clinical and imaging data of patients with early-stage NSCLC treated with SABR in the Department of Thoracic Radiation Oncology, Cancer Hospital of the University of Chinese Academy of Sciences from 2012 to 2018 were collected. General information (e.g. gender, age, KPS, Charlson score, lung function, smoking status), pre-treatment lesion status (e.g. diameter, location, pathological type, T stage), radiation parameters (biological effective dose, BED), and the type of peritumoral radiation-induced lung injury (RILI) was classified.

Inclusion criteria: 1) Pathologically confirmed primary NSCLC by bronchoscopy or percutaneous CT-guided biopsy; 2) The TNM clinical stage I~II according to the American Joint Committee on Cancer (AJCC) (8th edition); 3) Have not received other prior antitumor therapy; 4) Thorax CT examination performed before treatment and every 3-6 months follow-up after SABR. Exclusion criteria: 1) coexisting with other primary malignant tumors; 2) incomplete clinical and imaging data; 3) lesions cannot be accurately segmented (e.g. the lesion and peripheral atelectasis cannot be accurately segmented.); 4) lost to follow-up.

SABR treatment

All patients performed four-dimensional CT simulations with free breathing. The Internal Gross Target Volume (IGTV) was derived from the Maximum Intensity Projection (MIP) of 4DCT and the Planning Target Volume (PTV) was expanded by a 5-mm margin in all directions around the IGTV. The total radiation dose and fraction dose were determined by the radiation oncologists based on the lesion location, volume, and peripheral organs at risk. Target delineation, conformity, and dose limitations in normal tissues were referred to the American Radiation Therapy Oncology Group (RTOG) 0236 study (18). The prescription dose was 5-15Gy per fraction, once a day, with a total dose of 40-70Gy. The BED was calculated using the formula, $BED_{\alpha/\beta} = nd\ (1+ d/\alpha/\beta)$, where n=number of fractions, d=dose per fraction, and α/β=10 Gy for the lung cancer.

Follow-up

All patients underwent enhanced thorax CT examination one month after the end of treatment and every three months thereafter, and every six months after two years. If progression is suspected, PET-CT or pathological biopsy is performed. Enhanced thorax CT was performed with GE 64-slice CT or Siemens 64-slice CT, tube current 100~300mAs, tube voltage 120 kV, pitch 5.0 mm, slice thickness 5.0 mm. The contrast agent was selected from Opitiray (Ioversol) or Ultravist (Iopromide), and the high-pressure syringe was injected rapidly through the dorsal vein of the hand, the injection rate was 2.5ml/s, and the dose was 80-95ml. Enhanced thorax CT examination was performed 38s after contrast agent injection. Disease progression within two years of follow-up was defined as a high-risk group, and progression or no progression for more than two years was defined as a low-risk group.

Radiomic analysis

The workflow of the study was shown in Figure 1.

Medical imaging segmentation

The lung-window CT images (window width of 1600 Hounsfield units (HU) and window level of -450 HU; DICM format) of early-stage NSCLC patients before SABR treatment were imported into ITK-SNAP software (Version 3.4.0,

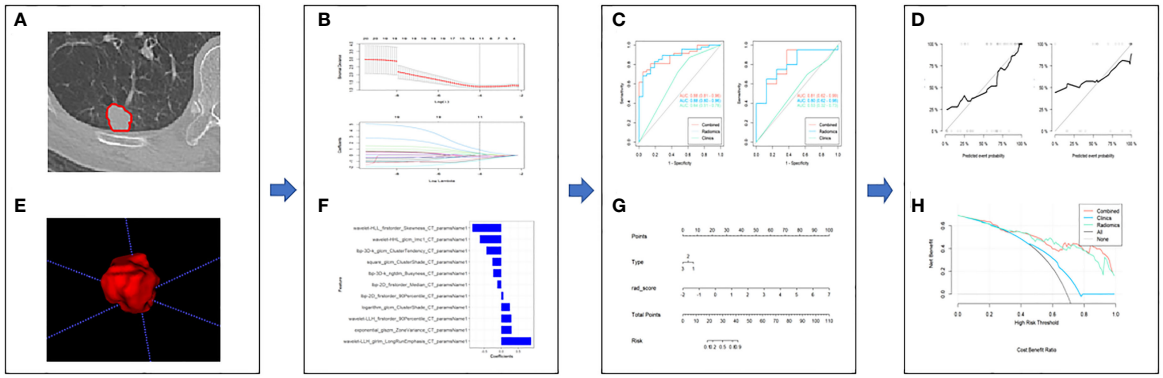


FIGURE 1
|The framework for the radiomics workflow. **(A, B)** Medical imaging segmentation; **(C, D)** Feature extraction and selection; **(E, F)** The ROC curves and nomogram; **(G, H)** Hosmer-Lemeshow Test and the decision curve.

http://www.itksnap.org/). A region of interest (ROI) was manually delineated layer by layer by an attending radiologist (who had 10 years of experience with thorax CT images), and a Volume of Interest (VOI) was synthesized. Adjacent aorta, ribs, and pulmonary bullae were excluded. At the same time, a senior radiologist (who had 15 years of experience with thorax CT images) randomly selected 30 patients and repeated the delineation process. The Intraclass Correlation Coefficient (ICC) was used to evaluate consistency between observers.

Feature extraction and selection

Image preprocessing and radiomics feature extraction were performed using python pyradiomics (version 3.0.1), which complies with IBSI (19). Image preprocessing includes resampling, denoising, and intensity standardization. Feature parameters include morphological features, first-order features, texture features, and transformation-based features. Before feature selection, radiomics features of different dimensions were normalized using a Z-score, which was used to remove the mean and variance normalization. The minimum Redundancy Maximum Relevance (mRMR) and the Least Absolute Shrinkage and Selection Operator (LASSO) were used for dimensionality reduction and feature selection.

Model construction and evaluation

Univariate logistic regression analysis was used to screen independent clinical risk factors. According to the ratio of 7: 3, the patients were randomly divided into the training cohort and the validation cohort. The data of the training cohort were used to construct the model, and the data of the validation cohort were used to test. According to the radiomics labels and their weight coefficients, the radiomics score (Radscore) of every patient was calculated, and a radiomics model was established. Multivariate logistic regression analysis was used to establish a combined model based on radiomics features and clinical factors, and a nomogram was constructed. The area under curve (AUC) was calculated by receiver operating characteristic (ROC) curve analysis, and the performance of the training cohort and the validation cohort models was evaluated. The accuracy, sensitivity, specificity, Positive Predictive Value (PPV), and Negative Predictive Value (NPV) of the models were obtained. Delong test, Hosmer-Lemeshow test, and Decision Curve Analysis (DCA) were used to evaluate the diagnostic efficiency and clinical utility of the model.

Statistical analysis

All data analysis was performed by using IBM SPSS version 24.0 (IBM Corp., Armonk, NY, USA). The continuous variables that conformed to be normally distributed were analyzed by the independent samples t-test. Otherwise, the continuous variables were analyzed by the Wilcoxon Rank-Sum test. The categorical variables were used the chi-square test or Fisher's exact test. $P< 0.05$ was considered statistically significant.

Results

Patient characteristics

A total of 96 patients were included in this study. With the median follow-up of 59.1 months, 29 patients developed progression and 67 remained good controlled within two years. All patients were randomly assigned to the training cohort (n=68) and the validation cohort (n=28). There were no statistically significant differences in clinical factors between the training and validation cohorts ($P > 0.05$). Statistical

characteristics were summarized in Table 1. Univariate logistic regression analysis showed that the type of peritumoral RILI was significantly different between the high-risk group and the low-risk group for disease progression (OR, 0.48; 95% CI: 0.25-0.90; *P*=0.022). Thus, a clinical model is established through this independent risk factor.

Analysis based on CT radiomics features

Feature selection and model construction

A total of 1781 radiomics features (including morphological features, first-order features, texture features, and transformation-based features) were extracted from the

TABLE 1 Characteristics of patients in the training and validation cohorts.

	Training Cohort			Validation Cohort		
	High-risk (n=21)	**Low-risk (n=47)**	***p***	**High-risk (n=8)**	**Low-risk (n=20)**	***p***
Gender (%)						
Female	4 (19.0)	12 (25.5)		2 (25.0)	9 (45.0)	
Male	17 (81.0)	35 (74.5)	0.7849	6 (75.0)	11 (55.0)	0.58188
Age, years (mean ± SD)	74.5 (6.7)	73.6 (7.9)	0.6588	72.5 (6.6)	72 (9.2)	0.88141
KPS	88.8 (8.6)	89.8 (6.4)	0.6034	90 (7.6)	89.5 (7.6)	0.87475
Charlson	0.9 (1.2)	0.7 (1)	0.5613	1.6 (1.5)	0.6 (0.9)	0.02219*
Diameter, cm (mean ± SD)	2.3 (0.8)	2.5 (0.8)	0.2269	2.1 (0.7)	2.4 (0.9)	0.47013
Histology						
Adenocarcinoma	9 (42.9)	25 (53.2)		2 (25.0)	11 (55.0)	
Squamous cell carcinoma	10 (47.6)	13 (27.7)		3 (37.5)	4 (20.0)	
Not otherwise Specified	2 (9.5)	9 (19.1)	0.2404	3 (37.5)	5 (25.0)	0.34642
T stage						
1	16 (76.2)	34 (72.3)		4 (50.0)	17 (85.0)	
2	5 (23.8)	12 (25.5)		4 (50.0)	3 (15.0)	
3	0 (0.0)	1 (2.1)	0.7814	0 (0.0)	0 (0.0)	NA
Tumor location						
Central	2 (9.5)	3 (6.4)		1 (12.5)	0 (0.0)	
Peripheral	19 (90.5)	44 (93.6)	1.0000	7 (87.5)	20 (100.0)	0.62906
Involved lobe						
RLL/RML	8 (38.1)	15 (31.9)		2 (25.0)	8 (40.0)	
LLL	8 (38.1)	11 (23.4)		3 (37.5)	2 (10.0)	
LUL	2 (9.5)	7 (14.9)		2 (25.0)	4 (20.0)	
RUL	3 (14.3)	14 (29.8)	0.3922	1 (12.5)	6 (30.0)	0.31476
Pulmonary function						
Normal	3 (14.3)	4 (8.5)		1 (12.5)	1 (5.0)	
Mild	1 (4.8)	8 (17.0)		1 (12.5)	4 (20.0)	
Moderate	11 (52.4)	18 (38.3)		3 (37.5)	5 (25.0)	
Severe	6 (28.6)	17 (36.2)	0.3853	3 (37.5)	10 (50.0)	0.76868
Smoker						
No	6 (28.6)	21 (44.7)		4 (50.0)	10 (50.0)	
Yes	15 (71.4)	26 (55.3)	0.3241	4 (50.0)	10 (50.0)	1.00000
BED	98.1 (14.5)	98.5 (12.5)	0.9062	86.4 (14.8)	93.7 (19.3)	0.33499
BED≥100						
No	7 (33.3)	18 (38.3)		4 (50.0)	10 (50.0)	
Yes	14 (66.7)	29 (61.7)	0.9044	4 (50.0)	10 (50.0)	1.00000
Type						
1	10 (47.6)	34 (72.3)		5 (62.5)	14 (70.0)	
2	3 (14.3)	7 (14.9)		2 (25.0)	3 (15.0)	
3	8 (38.1)	6 (12.8)	0.0524	1 (12.5)	3 (15.0)	0.82186

KPS, karnofsky performance status; RLL, right lower lobe; RML, right middle lobe; LLL, left lower lobe; LUL, left upper lobe; RUL, right upper lobe; BED, biologically effective dose; Type, the type of peritumoral radiation-induced lung injury. *p< 0.05, expressive significance.

pre-treatment CT images of early-stage NSCLC treated with SABR by using the python pyradiomics (version 3.0.1), ICC=0.82>0.75, indicating good inter-group consistency. After dimensionality reduction and feature selection by mRMR and LASSO, the 11 most valuable features and their corresponding coefficients were retained, as shown in Figure 2. The values of 11 features were input into the formula to obtain Radscore, and the radiomics model reflecting the disease progression was established. The box plot showed the Radscore distribution of high- and low-risk group for disease progression in training and validation cohorts, as shown in Figure 3. The resulting formula was as follows:

Combined with the Radscore and the type of peritumoral RILI, a combined model was constructed, and a visual nomogram was formed, as shown in Figure 4.

Model performance evaluation

Figure 4 showed that the AUC with its 95% confidence interval (CI) of the radiomics model, clinical model, and combined model was 0.88 (95%CI 0.80-0.96), 0.64 (95%CI 0.51-0.78), and 0.88 (95% CI 0.81-0.96) in the training cohort and 0.80 (95%CI 0.62-0.98),0.53 (95%CI 0.32-0.73) and 0.81 (95%CI 0.62-0.99) in the validation cohort, respectively. Table 2 showed that the accuracy values of the radiomics model, clinical model, and combined

$$\begin{aligned}\text{Radscore} = &-0.12*\text{lbp} - 2\text{D}\_\text{firstorder}\_\text{Median} + 0.879*\text{wavelet} - \text{LLH}\_\text{glrlm}\_\text{LongRunEmphasis} + -0.237*\text{lbp} - 3\text{D} - \text{k\_ngtdm}\_\text{Busyness} + 0.254*\text{logarithm}\_\text{glcm}\_\text{ClusterShade} + -0.266*\text{square\_glcm\_ClusterShade} + -0.852*\text{wavelet-HLL\_firstorder\_Skewness} + -0.635*\text{wavelet-HHL\_glcm\_Imc1} + -0.442*\text{lbp-3D-k\_glcm\_ClusterTendency} + 0.3*\text{exponential}\\ &\text{l\_glszm\_ZoneVariance} + 0.297*\text{wavelet-LLH\_firstorder\_90Percentile} + 0.059*\text{lbp-2D}\_\\ &\text{firstorder\_90Percentile} + 1.212\end{aligned}$$

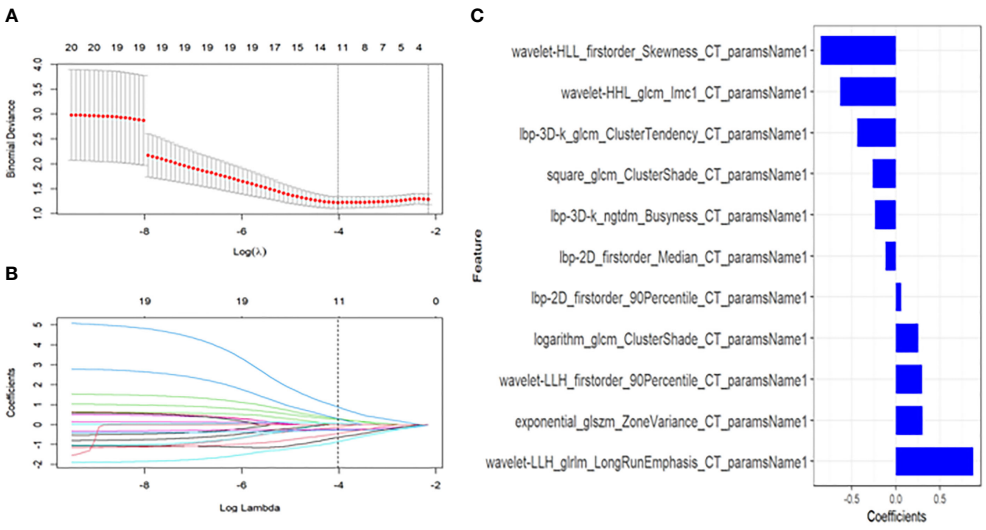


FIGURE 2
Textural feature selection using the Least Absolute Shrinkage and Selection Operator (LASSO) binary logistic regression. **(A)** Tuning parameters(λ) for the LASSO model were selected by 10-fold cross-validation using the minimum criteria. Partial likelihood deviance was plotted against log (λ). The dotted vertical lines correspond to the optimal values according to the minimum criteria and 1-SE criterion. The 11 features with the smallest binomial deviance were selected. **(B)** A feature coefficient convergence graph for filtering features using 10-fold cross-validation in the LASSO regression model. **(C)** LASSO coefficient profiles of texture features. Vertical lines correspond to the values selected by 10-fold cross-validation of the log(λ) sequence; the 11 nonzero coefficients are indicated.

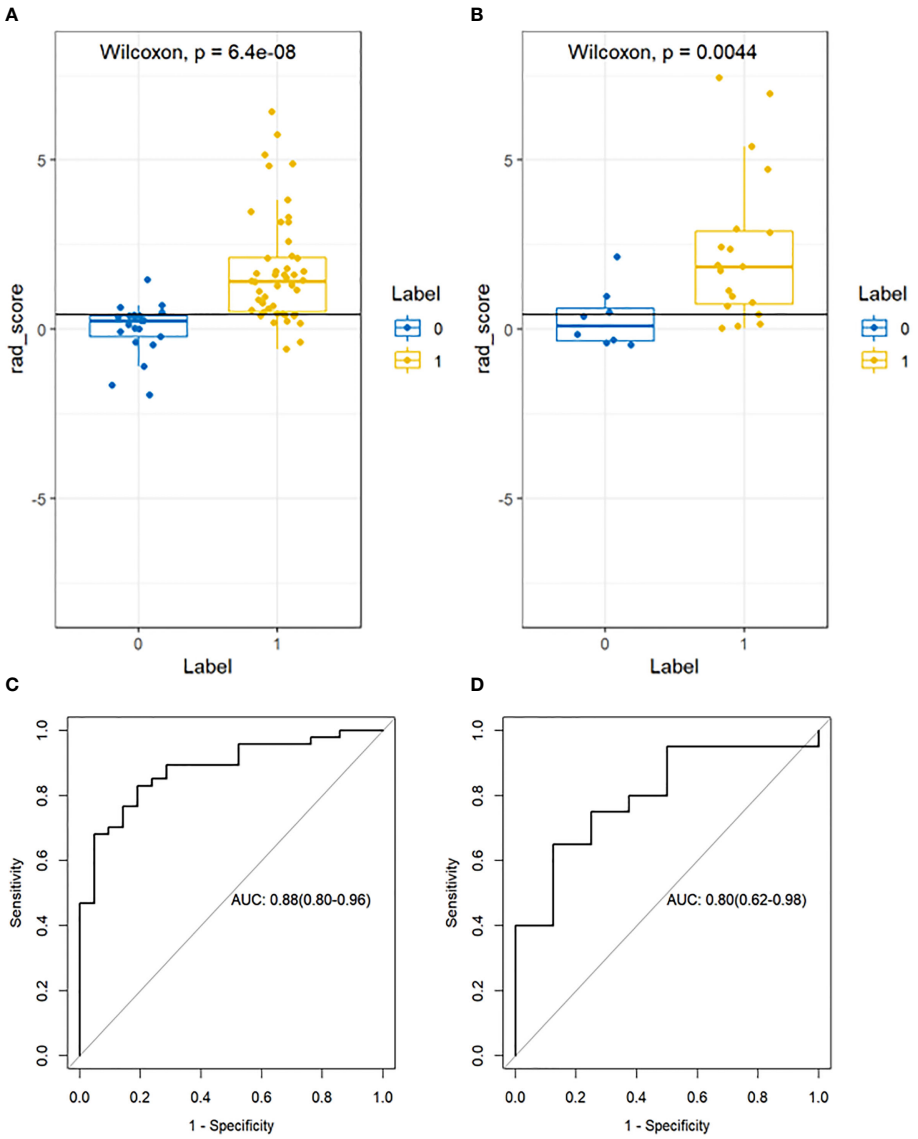


FIGURE 3
Box plot showing the Radscore distribution of high and low risk group for disease progression on training and validation cohorts. p-value from Wilcoxon Rank-Sum test **(A, B)**. Receiver Operator Characteristic (ROC) curves (training and validation cohorts) **(C, D)**. The prediction performance of the ROC curves for radiomics signature for training and validation cohorts.

model were 82.3%, 72.1%, and 79.4% in the training cohort, and 71.4%, 64.3%, and 85.6% in the validation cohort, respectively. The results showed that both the radiomics model and the combined model have good prediction efficiency in the training cohort and the validation cohort.

According to the DeLong test, the performance of the radiomics model and combined model in the training and validation cohort was significantly better than that of the clinical model ($P<0.05$), but there was no statistically significant difference between the radiomics model and combined model ($P > 0.05$), as shown in Table 3. Hosmer-Lemeshow Test of the nomograms of the training and validation cohorts were shown in Figure 5, in which the results showed that the prediction of disease progression in the training cohort was well-calibrated ($P>0.05$). DCA results for the three discrimination models were shown in Figure 6. The results showed that the radiomics and combined models have high clinical benefits.

Discussion

Radiomics can extract many disease features that cannot be observed with the naked eye from medical images and non-invasively capture information inside tumors that may be related to tumor recurrence, thereby realizing the goal of personalized

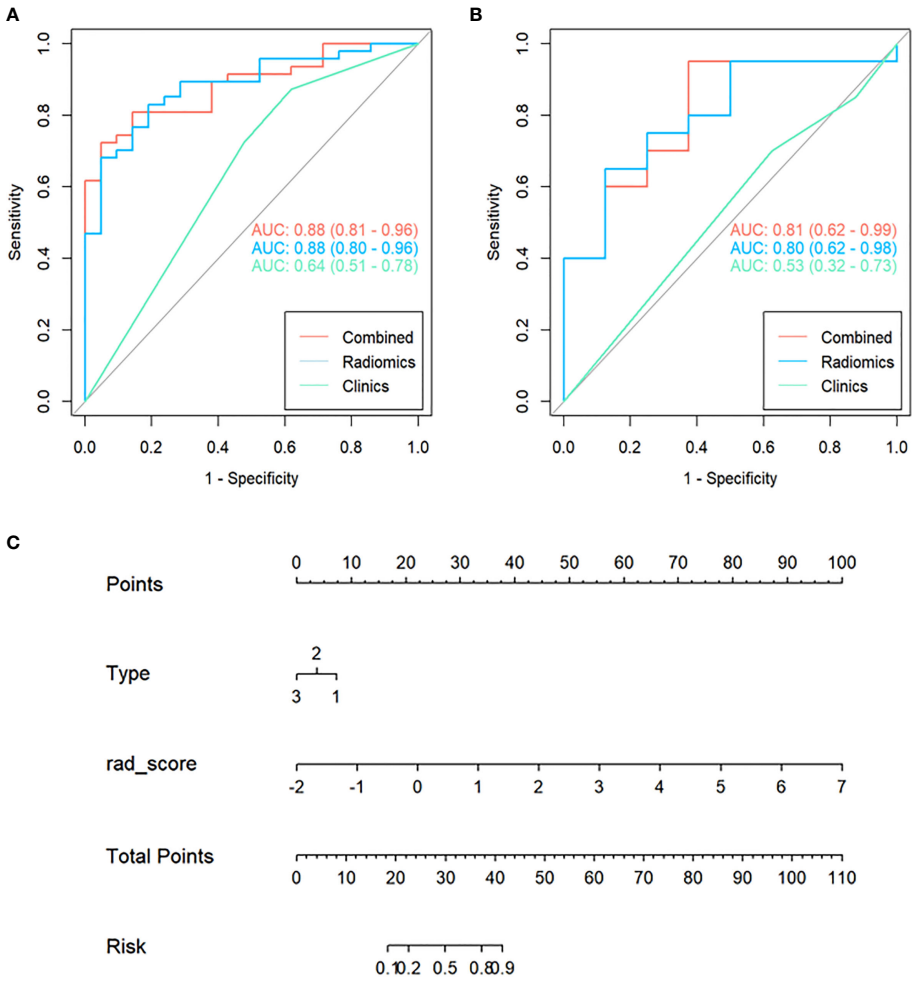


FIGURE 4
Receiver Operating Characteristic (ROC) curves of the clinical, radiomics, and combined model used to discriminate between the high and low risk of disease progression of lung cancer treated with SABR in the training and validation cohorts **(A, B)**. Radiomics nomogram **(C)** was used to discriminate the high and low risk of disease progression in lung cancer patients treated with SABR. The nomogram was based on the training cohort; the Radscore was shown. Initially, vertical lines were drawn at the Radscore values to determine the values of the points. The final point value was the sum of those of the two points. Finally, a vertical line was drawn at the total point value to determine the risk of disease progression of lung cancer treated with SABR.

TABLE 2 **Predictive performance of three prediction models for training and validation cohort.**

Training cohort	AUC	95%CI	Sensitivity	Specificity	Accuracy	PPV	NPV
Clinical model	0.64	0.51-0.78	0.872	0.381	0.721	0.759	0.571
Radiomics model	0.88	0.80-0.96	0.830	0.810	0.824	0.907	0.680
Combined model	0.88	0.81-0.96	0.971	0.606	0.794	0.723	0.952
Validation cohort	AUC	95%CI	Sensitivity	Specificity	Accuracy	PPV	NPV
Clinical model	0.53	0.32-0.73	0.850	0.125	0.643	0.708	0.250
Radiomics model	0.80	0.62-0.98	0.750	0.625	0.714	0.833	0.500
Combined model	0.81	0.62-0.99	0.864	0.833	0.857	0.950	0.625

AUC, the area under the curve; CI, confidence interval; PPV, positive predictive value; NPV, negative predictive value.

TABLE 3 Comparison of ROC curves with DeLong test in the training and validation cohort.

	Clinical vs Radiomics		Clinical vs Combined		Radiomics vs Combined	
	Z	P	Z	P	Z	P
Training Cohort	2.87	0.004*	3.48	<0.001*	0.093	0.926
Validation Cohort	2.08	0.038*	2.35	0.019*	0.24	0.812

*p< 0.05, expressive significance.

medicine. In our study, a combined model based on pre-treatment thorax CT radiomics features and clinical factors was developed and validated to predict the likelihood of disease progression after SABR in early-stage NSCLC.

Distant metastasis was one of the main reasons for SABR treated in early-stage NSCLC. The RTOG 0236 study showed that the 5-year distant metastasis rate was 31% (18). In addition, the metastasis usually develops soon after the treatment of the primary lesion, and the survival time is significantly reduced once it occurs. For these patients with a high risk of early distant metastasis, systemic therapy combined with SABR may reduce the risk of metastasis and improve the OS. Therefore, it is of great significance to establish an accurate and effective predictive model to assess the risk of disease progression in patients of early-stage NSCLC.

Several studies have focused on the relationship between SABR prognosis and clinicopathological factors. Onishi et al. showed that BED ≥100Gy had significantly better local control rates and OS than those receiving BED< 100Gy (20). The predictive survival model showed that BED_{10}< 113Gy was an independent risk factor for OS and PFS and was significantly associated with both local and distant progression (21). The prescription of BED ≥ 100Gy was currently recommended by international guidelines, including the National Comprehensive Cancer Network (NCCN) and the European Society of Medical Oncology (ESMO) guidelines. Kang et al. constructed a survival prediction model for stage I NSCLC treated with SABR, showing that tumor diameter >2.45 cm was an independent predictor of OS and PFS, which had a significant correlation with both local and distant progression (21). It is unclear whether there is any difference in the prognosis of different pathological types after SABR. Abel et al. analyzed 15,110 patients with early-stage (I ~ IIA) NSCLC who received SABR, and the 5-year OS of patients with adenocarcinoma and squamous cell carcinoma were 36% and 24% (P<0.0001), respectively. Squamous cell carcinoma was an independent poor prognostic factor (22). In our study, BED, tumor diameter, and pathological type did not correlate with disease progression, which may be related to the relatively concentrated BED dose (95% concentrated between 93.4Gy-99.4Gy), relatively uniform clinical factors, and a small number of cases and so on. Therefore, it is difficult to construct predictive models solely on the clinicopathological characteristics.

The type of peritumoral RILI was the only independent risk factor for tumor progression among clinical factors (P< 0.05). The pattern of changes in lung parenchyma on CT post-SABR can generally be categorized as acute (within six months, corresponding to pneumonitis) or late (after six months, corresponding to fibrosis) (23). Several papers have classified

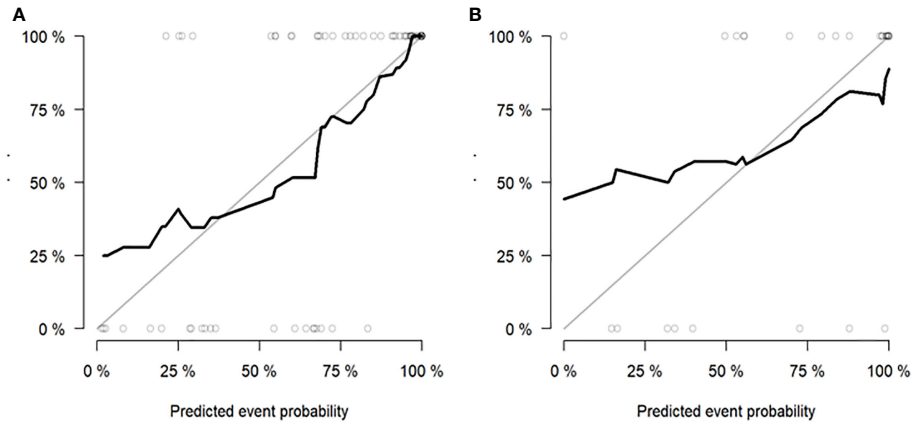


FIGURE 5
Hosmer-Lemeshow Test of the nomogram of the training **(A)** and validation **(B)** cohorts. The diagonal dotted lines represent the ideal predictions; the solid lines represent nomogram performance. A closer fit to the diagonal line indicates that the model matches better.

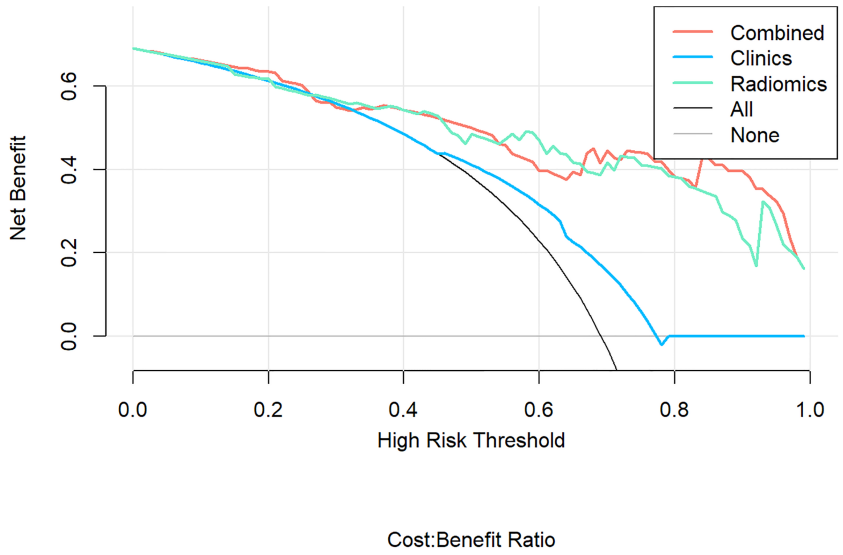


FIGURE 6
Decision Curve Analysis (DCA) results for the three discrimination models. The Y-axis represents the net benefit, calculated by summing the benefits (true positives) and subtracting the weighted harm (i.e., deleting false positives). The optimal method for feature selection is that with the highest net benefit.

acute changes into one of five general patterns: diffuse consolidation, patchy consolidation, diffuse ground-glass opacities (GGO), patchy GGO, and no change (24–26). In the past, the vast majority of literature discussed the identification of RILI and tumor recurrence (27–29), and there were few studies on the correlation between them. Based on the above considerations, we redefined the peritumoral RILI and divided them into three types (Figure 7). Type I is diffuse consolidation around the tumor, also called severe RILI. Type II is diffuse GGO around the tumor, which is distributed over 180 degrees around the tumor; we also call it moderate RILI. Type III is patchy GGO within a range of fewer than 180 degrees around the tumor, or there is no change; we call it mild RILI.

In this study, a total of 1781 radiomics features were extracted. After screening of radiomics features by mRMR and LASSO, 11 features were finally retained, including four first-order features and seven texture features, all based on transformation. First-order features describe the gray value distribution of tumor image ROIs. In this study, Skewness reflects the asymmetry of gray value distribution relative to the mean. The more low-signal gray distribution in the lesion is, the higher the tumor heterogeneity.

In the texture feature, the Gray Level Co-occurrence Matrix (GLCM) studies the spatial correlation characteristics between the gray levels of two points in a certain distance and direction in the image so as to reflect the texture information of the image in direction, interval, change amplitude and speed. In this study, Informational Measure of Correlation (IMC) 1 assesses the correlation between the probability distributions of i and j (quantifying the complexity of the texture); Cluster Tendency is a measure of groupings of voxels with similar gray-level values; Cluster Shade is a measure of the skewness and uniformity of the GLCM, a higher cluster shade implies more significant asymmetry about the mean. The Gray Level Run Length Matrix (GLRLM) mainly reflects texture roughness and directionality. It is used to describe the length of the same pixel gray level that appears continuously in a specified direction. In this study, Long Run Emphasis (LRE) measures the distribution of long-run lengths, with a more excellent value indicative of long run lengths and more coarse structural textures. The Gray Level Size Zone Matrix (GLSZM) provides information about the spatial distribution of corresponding adjacent pixels or voxels at the same gray level. In this study, Zone Variance (ZV) measures the variance in zone size volumes for the zones, and the more significant the value, the greater the heterogeneity. The Neighbouring Gray Tone Difference Matrix (NGTDM) represents the difference between the gray value of a point and the average gray value in the neighborhood at a certain distance, thereby capturing the spatial rate of gray intensity changes. In this study, busyness is a measure of the shift from a pixel to its neighbor; a high value for busyness indicates a 'busy' image, with rapid changes of intensity between pixels and their neighborhood. The gray information of these images can quantitatively analyze tumor heterogeneity so as to conduct quantitative studies at the microscopic level, which can effectively predict the disease progression of patients (30, 31). The features selected in this study were all processed by filters, which may be because filters can extract and reconstruct the parts of the original images, thus mining deeper image information.

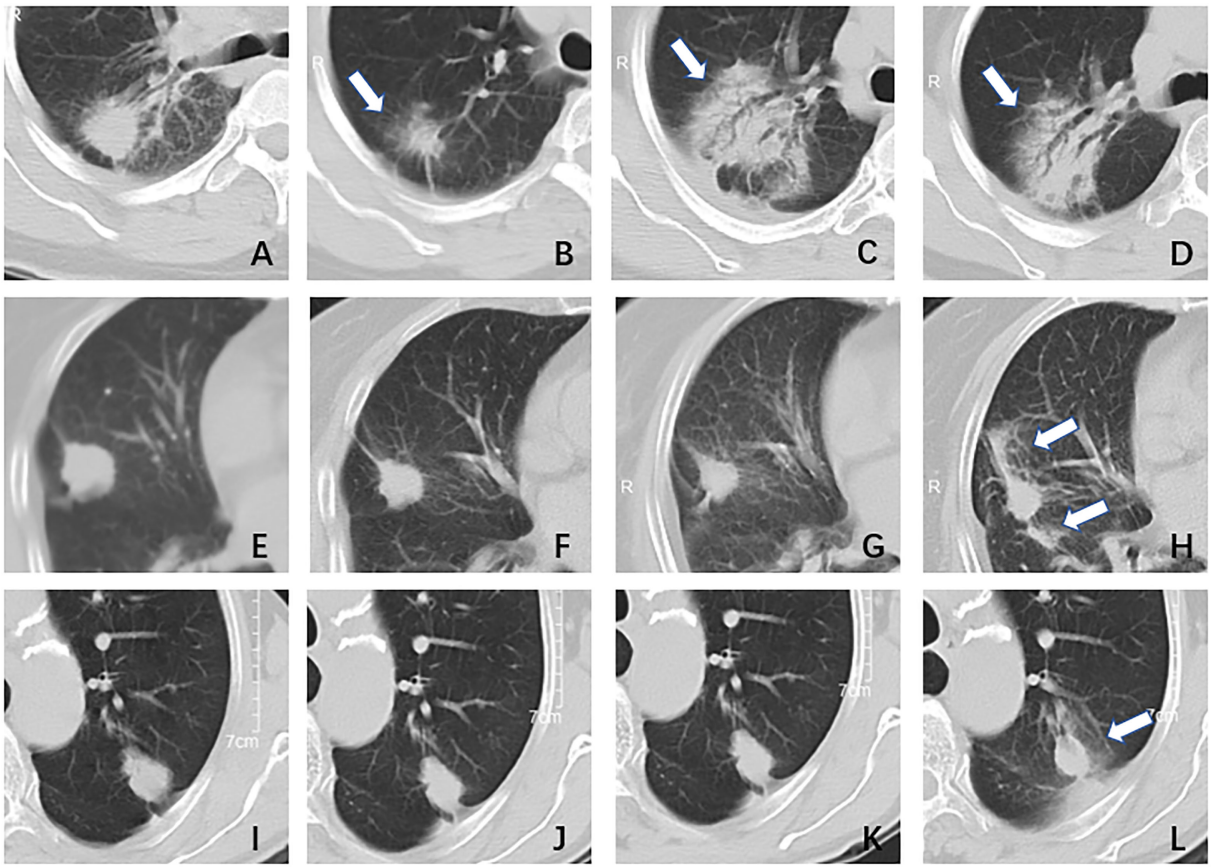


FIGURE 7
The type of peritumoral radiation-induced lung injury. Type I, female, 51 years, adenocarcinoma in the right lung, DT40GY/5F; **(A)** pre-treatment: a nodule with blurred boundary and spicule sign; **(B)** one month after treatment: the tumor shrunk and there was a surrounding ground-glass opacity; **(C)** three months after treatment: the tumor area showed diffuse consolidation and was indistinguishable from the tumor; **(D)** six months after treatment: the imaging findings were similar to **(C)**. Type II, female, 79 years, adenocarcinoma in right lung, DT55GY/5F; **(E)** pre-treatment: a nodule with a clear boundary and shallow lobed; **(F)**one month after treatment: the tumor has shrunk a little, no ground glass opacity surrounding it; **(G)** four months after treatment: there was no significant change; **(H)** six months after treatment: the tumor was surrounded by ground-glass opacity, more than 1/2. Type III, male,70 years, adenocarcinoma in left lung, DT50GY/5F; **(I)** pre-treatment: a nodule with a clear boundary and shallow lobed; **(J)** two months after treatment: there was no significant change; **(K)** four months after treatment: there was no significant change; **(L)** six months after treatment: the tumor was surrounded by ground-glass opacity, less than 1/2.

In this study, the AUC of the radiomics model and combined model in the training and validation cohorts were all above 0.80, and the accuracy rates were above 0.7. The model we developed showed a good predictive efficiency of disease progression after SABR, which provided important information for subsequent clinical therapy and follow-up. Rainer et al. also had similar findings. This study predicted tumor progression six months after Stereotactic Body Radiation Therapy (SBRT) for early-stage lung cancer, enrolled 399 patients from 13 different units, and finally retained seven radiomics features to establish a Support Vector Machine (SVM) model, using 10-fold cross-validation and AUC to evaluate the performance of the classifier. The results showed that the AUC was 0.789, sensitivity was 67.0%, and specificity was 78.7%, which was a good prediction (32). Lafata et al. also proposed the potential relationship between radiomics features extracted from pre-treatment CT images and clinical outcomes following SBRT for NSCLC; the results showed that two features demonstrated a statistically significant association with local failure: Homogeneity2 (p=0.022) and Long-Run-High-Gray-Level-Emphasis (p=0.048) multivariable logistic regression models produced AUC values of 0.83 (33).

In the DeLong test, we found no statistically significant difference between the radiomics model based solely on the CT images and the combined model, which indirectly confirmed the dominant role of CT images in the prediction model. Even so, compared with the ROC, the AUC value of the combined model is higher than that of the pure CT radiomics model. Therefore, clinical variables (the type of peritumoral RILI) still have a specific positive effect on the comprehensive judgment of

the model. Luo et al. and Li et al. also proposed that clinical variables were significantly correlated with the clinical outcomes of patients receiving SBRT for lung cancer and proved that the combined model based on clinical factors and radiomics features could effectively improve model prediction efficiency (34, 35).

Limitations of this study: Firstly, this study was a retrospective study, which can only be analyzed based on existing data, and prospective studies can be carried out in the future to incorporate some new variables. Secondly, the number of cases in this study was limited, and the sample size needed to be further expanded to improve the stability of the model. Thirdly, the data in this study came from the same hospital, and only internal validation was performed. Data from other hospitals should be added for external validation to improve model repeatability.

Conclusions

In conclusion, the radiomics model established based on pre-treatment thorax CT images of early-stage NSCLC can predict the disease progression after SABR treatment. At the same time, the nomogram we developed has a better predictive ability for the disease progression and provides a feasible and practical reference value for clinical guidance of individualized treatment, follow-up, and evaluation strategies for patients undergoing SABR.

Data availability statement

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

Ethics statement

The studies involving human participants were reviewed and approved by Ethics committees of Cancer Hospital of the University of Chinese Academy of Sciences. Written informed consent for participation was not required for this study in accordance with the national legislation and the institutional requirements.

Author contributions

HY had full access to all of the data in the study and took responsibility for the integrity of the data and the accuracy of the data analysis. HY and LW are co-first authors of this article. Collection and assembly of data: BD, XD, and HM. Resources data curation: XS, ZW, XZ, XL, WF, LS, GQ, JL, JZ, CW, QZ, XY, QC, MC, PL, YW, YJ, JC,YJ and XH. Data analysis and interpretation: GS, FW, HC, WC, YZ,YH and YZ. Writing-original draft preparation: HY and LW. Statistical analysis: HY, YGW. Study concept, supervision, funding acquisition, project administration, and writing-review: HJ and YX. All authors contributed to the article and approved the submitted version.

Funding

This study was supported by grants from Medical and Health Research Project of Zhejiang Province (Grant Number: 2020KY486 ; 2020KY079); Beijing Xisike Clinical Oncology Research Foundation (Y-2019AZMS-0061); and Beijing Science and Technology Innovation Medical Development Foundation (KC2021-JX-0186-63).

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. Sung H, Ferlay J, Siegel RL, Laversanne M, Soerjomataram I, Jemal A, et al. Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin* (2021) 71:209–49. doi: 10.3322/caac.21660

2. Scott WJ, Howington J, Feigenberg S, Movsas B, Pisters K. Treatment of non-small cell lung cancer stage I and stage II: ACCP evidence-based clinical practice guidelines (2nd edition). *Chest* (2007) 132:234–42. doi: 10.1378/chest.07-1378

3. Stanic S, Paulus R, Timmerman RD, Michalski JM, Barriger RB, Bezjak A, et al. No clinically significant changes in pulmonary function following stereotactic body radiation therapy for early-stage peripheral non-small cell lung cancer: an analysis of RTOG 0236. *Int J Radiat Oncol Biol Phys* (2014) 88:1092–9. doi: 10.1016/j.ijrobp.2013.12.050

4. Zheng X, Schipper M, Kidwell K, Lin J, Reddy R, Ren Y, et al. Survival outcome after stereotactic body radiation therapy and surgery for stage I non-small cell lung cancer: a meta-analysis. *Int J Radiat Oncol Biol Phys* (2014) 90:603–11. doi: 10.1016/j.ijrobp.2014.05.055

5. Ma L, Xiang J. Clinical outcomes of video-assisted thoracic surgery and stereotactic body radiation therapy for early-stage non-small cell lung cancer: A meta-analysis. *Thorac Cancer* (2016) 7:442–51. doi: 10.1111/1759-7714.12352

6. Senthi S, Lagerwaard FJ, Haasbeek CJ, Slotman BJ, Senan S. Patterns of disease recurrence after stereotactic ablative radiotherapy for early stage non-small cell lung cancer: a retrospective analysis. *Lancet Oncol* (2012) 13:802–9. doi: 10.1016/S1470-2045(12)70242-5

7. Chang JY, Liu H, Balter P, Komaki R, Liao Z, Welsh J, et al. Clinical outcome and predictors of survival and pneumonitis after stereotactic ablative radiotherapy for stage I non-small cell lung cancer. *Radiat Oncol* (2012) 7:152. doi: 10.1186/1748-717X-7-152

8. Grills IS, Hope AJ, Guckenberger M, Kestin LL, Werner-Wasik M, Yan D, et al. A collaborative analysis of stereotactic lung radiotherapy outcomes for early-stage non-small-cell lung cancer using daily online cone-beam computed tomography image-guided radiotherapy. *J Thorac Oncol* (2012) 7:1382–93. doi: 10.1097/JTO.0b013e318260e00d

9. Baumann P, Nyman J, Hoyer M, Wennberg B, Gagliardi G, Lax I, et al. Outcome in a prospective phase II trial of medically inoperable stageInon-small-cell lung cancer patients treated with stereotactic body radiotherapy. *Clin Oncol* (2009) 27:3290–6. doi: 10.1200/JCO.2008.21.5681

10. Grills IS, Mangona VS, Welsh R, Chmielewski G, Mcinerney E, Martin S, et al. Outcomes after stereotactic lung radiotherapy or wedge resection for stage Inon-small-cell lung cancer. *J Clin Oncol* (2010) 28:928–35. doi: 10.1016/j.ijrobp.2009.07.155

11. Chicas-Sett R, Morales-Orue I, Castilla-Martinez J, Zafra-Martin J, Kannemann A, Blanco J, et al. Stereotactic ablative radiotherapy combined with immune checkpoint inhibitors reboots the immune response assisted by immunotherapy in metastatic lung cancer: A systematic review. *Int J Mol Sci* (2019) 20:2173–90. doi: 10.3390/ijms20092173

12. Limkin EJ, Sun R, Dercle L, Zacharaki EI, Robert C, Reuzé S, et al. Promises and challenges for the implementation of computational medical imaging (radiomics) in oncology. *Ann Oncol* (2017) 28:1191–206. doi: 10.1093/annonc/mdx034

13. Lambin P, Leijenaar RTH, Deist TM, Peerlings J, de Jong EEC, van Timmeren J, et al. Radiomics: the bridge between medical imaging and personalized medicine. *Nat Rev Clin Oncol* (2017) 14:749–62. doi: 10.1038/nrclinonc.2017.141

14. Liu J, Cui J, Liu F, Yuan Y, Guo F, Zhang G, et al. Multi-subtype classification model for non-small cell lung cancer based on radiomics: SLS model. *Med Phys* (2019) 46:3091–100. doi: 10.1002/mp.13551

15. Rossi G, Barabino E, Fedeli A, Ficarra G, Coco S, Russo A, et al. Radiomic detection of EGFR mutations in NSCLC. *Cancer Res* (2021) 81:724–31. doi: 10.1158/0008-5472.CAN-20-0999

16. van Timmeren JE, van Elmpt W, Leijenaar RTH, Reymen B, Monshouwer R, Bussink J, et al. Longitudinal radiomics of cone-beam CT images from non-small cell lung cancer patients: Evaluation of the added prognostic value for overall survival and locoregional recurrence. *Radiother Oncol* (2019) 136:78–85. doi: 10.1016/j.radonc.2019.03.032

17. Kadoya N, Tanaka S, Kajikawa T, Tanabe S, Abe K, Nakajima Y, et al. Homology-based radiomic features for prediction of the prognosis of lung cancer based on CT-based radiomics. *Med Phys* (2020) 47:2197–205. doi: 10.1002/mp.14104

18. Timmerman RD, Hu C, Michalski J, Straube W, Galvin J, Johnstone D, et al. Long-term results of RTOG 0236: a phase II trial of stereotactic body radiation therapy (SBRT) in the treatment of patients with medically inoperable stage I non-small cell lung cancer. *Int J RadiatOncol* (2014) 90:S30. doi: 10.1016/j.ijrobp.2014.05.135

19. van Griethuysen JJM, Fedorov A, Parmar C, Hosny A, Aucoin N, Narayan V, et al. Computational radiomics system to decode the radiographic phenotype. *Cancer Res* (2017) 77:e104–7. doi: 10.1158/0008-5472.CAN-17-0339

20. Onishi H, Araki T, Shirato H, Nagata Y, Hiraoka M, Gomi K, et al. Stereotactic hypofractionated high-dose irradiation for stage I non-small cell lung carcinoma: clinical outcomes in 245 subjects in a Japanese multi-institutional study. *Cancer* (2004) 101:1623-1631. doi: 10.1002/cncr.20539

21. Kang J, Ning MS, Feng H, Li H, Bahig H, Brooks ED, et al. Predicting 5–year progression and survival outcomes for early stage non-small cell lung cancer treated with stereotactic ablative radiation therapy: development and validation of robust prognostic nomograms. *Int J Radiat Oncol Biol Phys* (2020) 106:90–9. doi: 10.1016/j.ijrobp.2019.09.037

22. Abel S, Hasan S, White R, Schumacher L, Finley G, Colonias A, et al. Stereotactic ablative radiotherapy (SABR) in early stage non-small cell lung cancer: comparing survival outcomes in adenocarcinoma and squamous cell carcinoma. *Lung Cancer* (2019) 128:127–33. doi: 10.1016/j.lungcan.2018.12.022

23. Palma DA, Senan S, Haasbeek CJA, Verbakel WFAR, Vincent A, Lagerwaard F, et al. Radiological and clinical pneumonitis after stereotactic lung radiotherapy: a matched analysis of three-dimensional conformal and volumetric-modulated arc therapy techniques. *Int J Radiat Oncol Biol Phys* (2011) 80:506–513. doi: 10.1016/j.ijrobp.2010.02.032

24. Dahele M, Palma D, Lagerwaard F, Slotman B, Senan S. Radiological changes after stereotactic radiotherapy for stage I lung cancer. *J Thorac Oncol* (2011) 6:1221–8. doi: 10.1097/JTO.0b013e318219aac5

25. Bhatt AD, El-Ghamry MN, Dunlap NE, Bhatt G, Harkenrider MM, Schuler JC, et al. Tumor volume change with stereotactic body radiotherapy (SBRT) for early-stage lung cancer: Evaluating the potential for adaptive SBRT. *Am J Clin Oncol* (2015) 38:41–6. doi: 10.1097/COC.0b013e318287bd7f

26. Yang Y, Li G, Li S, Wang Y, Zhao Y, Dong B, et al. CT appearance pattern after stereotactic body radiation therapy predicts outcomes in early-stage non-Small-Cell lung cancer. *Front Oncol* (2021) 11:746785. doi: 10.3389/fonc.2021.746785

27. Huang K, Dahele M, Senan S, Guckenberger M, Rodrigues GB, Ward A, et al. Radiographic changes after lung stereotactic ablative radiotherapy (SABR) – can we distinguish recurrence from fibrosis? a systematic review of the literature. *Radiother Oncol* (2012) 102:335–42. doi: 10.1016/j.radonc.2011.12.018

28. Kato S, Nambu A, Onishi H, Saito A, Kuriyama K, Komiyama T, et al. Computed tomography appearances of local recurrence after stereotactic body radiation therapy for stage I non-small-cell lung carcinoma. *Jpn J Radiol* (2010) 28:259–65. doi: 10.1007/s11604-009-0415-3

29. Halpenny D, Ridge CA, Hayes S, Zheng J, Moskowitz CS, Rimner A, et al. Computed tomographic features predictive of local recurrence in patients with early stage lung cancer treated with stereotactic body radiation therapy. *Clin Imaging* (2015) 39:254–8. doi: 10.1016/j.clinimag.2014.12.005

30. Avanzo M, Stancanello J, El Naqa I. Beyond imaging: The promise of radiomics. *Phys Med* (2017) 38:122–39. doi: 10.1016/j.ejmp.2017.05.071

31. Abbasian Ardakani A, Bureau NJ, Ciaccio EJ, Acharya UR. Interpretation of radiomics features–a pictorial review. *Comput Methods Prog BioMed* (2022) 215:106609. doi: 10.1016/j.cmpb.2021.106609

32. Klement RJ, Allgäuer M, Appold S, Dieckmann K, Ernst I, Ganswindt U, et al. Support vector machine-based prediction of local tumor control after stereotactic body radiation therapy for early-stage non-small cell lung cancer. *Int J Radiat Oncol Biol Phys* (2014) 88:732–8. doi: 10.1016/j.ijrobp.2013.11.216

33. Lafata KJ, Hong JC, Geng R, Ackerson BG, Liu JG, Zhou Z, et al. Association of pre-treatment radiomics features with lung cancer recurrence following stereotactic body radiation therapy. *Phys Med Biol* (2019) 64:025007. doi: 10.1088/1361-6560/aaf5a5

34. Luo LM, Huang BT, Chen CZ, Wang Y, Su CH, Peng GB, et al. A combined model to improve the prediction of local control for lung cancer patients undergoing stereotactic body radiotherapy based on radiomic signature plus clinical and dosimetric parameters. *Front Oncol* (2021) 11:819047. doi: 10.3389/fonc.2021.819047

35. Li Q, Kim J, Balagurunathan Y, Liu Y, Latifi K, Stringfield O, et al. Imaging features from pretreatment CT scans are associated with clinical outcomes in non-small-cell lung cancer patients treated with stereotactic body radiotherapy. *Med Phys* (2017) 44:4341–9. doi: 10.1002/mp.12309

frontiers | Frontiers in Oncology

TYPE Original Research
PUBLISHED 09 August 2022
DOI 10.3389/fonc.2022.975881

OPEN ACCESS

EDITED BY
Xue Meng,
Shandong Cancer Hospital, China

REVIEWED BY
Zijian Zhang,
Central South University, China
Jianbo Wang,
Shandong University, China

*CORRESPONDENCE
Haitao Jiang
jianght@zjcc.org.cn

SPECIALTY SECTION
This article was submitted to Cancer Imaging and Image-directed Interventions, a section of the journal Frontiers in Oncology

RECEIVED 22 June 2022
ACCEPTED 18 July 2022
PUBLISHED 09 August 2022

CITATION
Jiang T, Tan Y, Nan S, Wang F, Chen W, Wei Y, Liu T, Qin W, Lu F, Jiang F and Jiang H (2022) Radiomics based on pretreatment MRI for predicting distant metastasis of nasopharyngeal carcinoma: A preliminary study. *Front. Oncol.* 12:975881. doi: 10.3389/fonc.2022.975881

Radiomics based on pretreatment MRI for predicting distant metastasis of nasopharyngeal carcinoma: A preliminary study

Tingting Jiang[1,2], Yalan Tan[1,2], Shuaimin Nan[1,2], Fang Wang[1,2], Wujie Chen[1,2], Yuguo Wei[3], Tongxin Liu[2,4], Weifeng Qin[2,4], Fangxiao Lu[1,2], Feng Jiang[2,4] and Haitao Jiang[1,2]*

[1]Department of Radiology, Cancer Hospital of the University of Chinese Academy of Sciences (Zhejiang Cancer Hospital), Hangzhou, China, [2]Institute of Basic Medicine and Cancer (IBMC), Chinese Academy of Sciences, Hangzhou, China, [3]Precision Health Institution, General Electric (GE) Healthcare, Hangzhou, China, [4]Department of Radiation Oncology, Cancer Hospital of the University of Chinese Academy of Sciences (Zhejiang Cancer Hospital), Hangzhou, China

Objective: To explore the feasibility of predicting distant metastasis (DM) of nasopharyngeal carcinoma (NPC) patients based on MRI radiomics model.

Methods: A total of 146 patients with NPC pathologically confirmed, who did not exhibit DM before treatment, were retrospectively reviewed and followed up for at least one year to analyze the DM risk of the disease. The MRI images of these patients including T2WI and CE-T1WI sequences were extracted. The cases were randomly divided into training group (n=116) and validation group (n=30). The images were filtered before radiomics feature extraction. The least absolute shrinkage and selection operator (LASSO) regression was used to develop the dimension of texture parameters and the logistic regression was used to construct the prediction model. The ROC curve and calibration curve were used to evaluate the predictive performance of the model, and the area under curve (AUC), accuracy, sensitivity, and specificity were calculated.

Results: 72 patients had DM and 74 patients had no DM. The AUC, accuracy, sensitivity and specificity of the model were 0. 80 (95% CI: 0.72~0. 88), 75.0%, 76.8%, 73.3%. and0.70 (95% CI: 0.51~0.90), 66.7%, 72.7%, 63.2% in training group and validation group, respectively.

Conclusion: The radiomics model based on logistic regression algorithm has application potential for evaluating the DM risk of patients with NPC.

KEYWORDS

nasopharyngeal carcinoma, distant metastasis, radiomics, prediction model, magnetic resonance imaging

1 Introduction

Nasopharyngeal carcinoma (NPC) is one of the most common head and neck cancers with a high incidence in South China, Southeast Asia, and North Africa (1). According to the latest data from the International Agency for Research on Cancer, the number of NPC patients from China in 2020 was 62444, of which 34,810 patients were died for the disease (2). More than 75% of patients were diagnosed with Locally advanced NPC (TNM stage III or IVA) at the first visit (3). Notably, due to highly sensitivity to radiotherapy, the prognosis of NPC has been greatly improved with advancements in radiotherapy and optimizations in chemotherapy regimens (4). At present, distant metastasis is still the major cause of treatment failure in NPC (5). As such, early diagnosis and accurate identification of DM is indispensable for timely implementation of reasonable treatment.

Currently, the anatomical tumor-node-metastasis (TNM) staging system is the main indicator for prognostic prediction, but this system has limitations in predicting DM and stratification for treatment decisions (6). Recent studies have shown that although patients within the same TNM stage received equivalent standard treatments, more than 20% of patients eventually developing DM showed poor efficacy and prognosis (7). The possible explanation is that TNM staging is mainly based on the anatomical information and cannot reflect the presence of heterogeneity in tumors. Hence, exploring an effective strategy to accurately identify patients at a high risk of DM is essential.

Radiomics, an emerging field of medical research, involves the transformation of traditional medical images into analyzable quantitative imaging features for model construction, and has shown great advantages in early diagnosis, efficacy evaluation, and prognosis prediction of tumors (8). This study aims to explore the feasibility of predicting distant metastasis risk (DM) of nasopharyngeal carcinoma (NPC) based on MRI radiomics model.

2 Materials and methods

2.1 Patients and datasets

In this retrospective study, the medical records and imaging data of NPC patients were obtained at Zhejiang Cancer Hospital from January 2010 to December 2016. The clinical features including gender, age, T stage, and histological types were collected. Inclusion criteria were: (1) two sequences (axial T2-weighted [T2WI] and T1-weighted contrast [CE-T1WI]) of head and neck MRI were all collected; (2) all patients were diagnosed with NPC by pathology and did not exhibit DM before treatment; (3) all patients were followed up for more than one year, with a maximum of 6 years; (4) no prior malignancy. Exclusion criteria were: (1) received antitumor therapy before MRI examination; (2) artifacts in MRI images; (3) the tumor is too small (there is volume effect when sketching the target ROI). Finally, a total of 146 NPC patients were included, all patients were randomly divided into training group (n = 116) and validation group (n = 30) as a ratio of 4:1 by computer. This study was approved by the Ethics Committees of Zhejiang Cancer Hospital.

2.2 Image acquisition

MRI was performed with Siemens 3.0T MR scan equipment and a 16-channel head and neck joint coil. Scanning sequence and parameters: (1) axial T1-weighted imaging (T1WI): TR: 498ms, TE: 8ms, slice thickness: 5mm, FOV: 260×260mm, matrix size: 288×229. (2) axial T2-weighted imaging (T2WI): TR: 3020ms, TE: 100ms, slice thickness: 4mm, FOV: 260×260mm, matrix size: 372×363. The contrast medium was Gd-DTPA, dose 15mL (0.1mmol/kg), injection rate 2.0mL/s. Contrast enhanced T1WI (CE-T1WI scan was performed after 1min after intravenous injection of the elbow.

2.3 Research methods of radiomics

Tumor segmentation: T2WI and CE-T1WI images were introduced into ITK-SNAP software (version 3.8.0, http://www.itksnap.org/) for tumor segmentation. Two doctors (with more than 8 years of experience in neck diagnosis, respectively) manually sketched the target ROI layer by layer, and selected axial images of each sequence to avoid enlarged lymph nodes in parapharyngeal space as far as possible.

Radiomics feature extraction and radiomics model building: all the segmented ROI data were imported into the Darwin research platform for feature extraction. The definition and calculation formula of features are in line with the PyRadiomics standard (9). In order to avoid reducing the speed of calculation, the extracted features were standardized by the minimum and maximum scaling algorithm. The optimal feature selection was used for removing the low performance features, the K was set to 15% and the f calssif function was selected which refers to the first 15% features sorted by F value were selected by the analysis variance of F test statistics. The minimum absolute contraction and selection operator (LASSO) regression was used to further reduce the dimension of the feature parameters. Finally, 15 combinatorial features were obtained. Logistic regression algorithm was used to establish a model includes above selected feature parameters. The diagnostic efficacy evaluation of the training group and validation group model were obtained from the receiver operating characteristic (ROC) and area under curve (AUC),

accuracy, sensitivity, and specificity. The workflow of the radiomics procession is presented in Figure 1.

2.4 Statistical analysis

The general characteristics of patients were statistically analyzed by SPSS 26.0. The classified data was compared by chi-square test or Fisher's exact test. The independent sample t-test was applied in the analysis of the quantitative data which according to normal distribution were expressed as mean ± standard deviation (). The Mann-Whitney U test was used for the comparison of quantitative data which did accord to normal distribution. $P< 0.05$ indicates that the difference is statistically significant.

3 Results

3.1 Clinical characteristics of the patients

A total of 146 patients with NPC were collected and followed up for at least one year, of which 72 patients had DM, and other 74 patients had no DM. Histological subtypes were divided into three types: type I differentiated keratinizing carcinoma (n=15), type II differentiated nonkeratinizing carcinoma (n=69), and type III undifferentiated nonkeratinizing carcinoma (n=62). There was a significant difference in T-stage between the two groups ($P<$ 0.001). The T stage in DM group was higher as a whole. There was no significant difference in sex, age and histological types between the two groups ($P>$ 0.05), as shown in Table 1.

3.2 Radiomics feature selection results

1781 radiomics features were extracted from magnetic resonance images: (1) first-order features, (2) shape features, (3) texture features: gray level co-occurrence matrix (GLCM), gray level run length matrix (GLRLM), gray level size zone matrix (GLSZM), gray level dependence matrix (GLDM), neighbouring gray tone difference matrix (NGTDM). There were 15 most valuable imaging features after dimensionality reduction with LASSO, including first-order features (n=4) and texture features (n=11). The first-order features including: inter quartile range, skewness (n=2), kurtosis. GLCM features include: cluster tendency, difference average (DA), sum squares (SS), correlation, cluster shade (CS), informational measure of correlation (Imc1). GLRLM features include: run variance (RV), long run low gray level emphasis (LRLGLE) (n=2). GLSZM features including: small area low gray level emphasis (SALGLE); NGTDM features including: strength. The Workflow of radiomics model was shown in Figure 2.

3.3 Prediction model results

In the training group, AUC was 0.80 (95% CI: 0.72- 0.88), the sensitivity was 76.8%, the specificity was 73.3% and the accuracy was 75.0%. In the validation group, the AUC value of the model was 0.70 (95% CI: 0.51- 0. 90), the sensitivity was 72.7%, the specificity was 63.2% and the accuracy was 66.7%. The corresponding characteristic coefficients and the comparison of characteristic parameters between the two groups were shown in Table 2, Figures 3, 4.

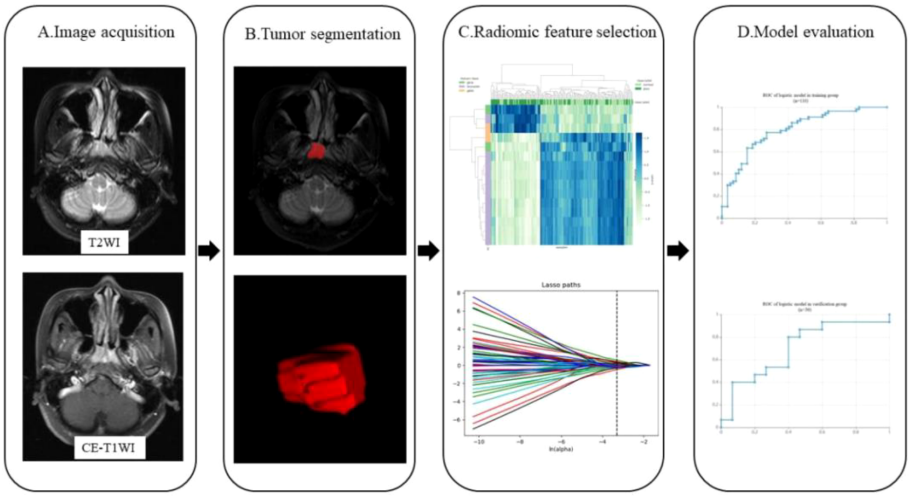


FIGURE 1
Workflow showing the establishment of a radiomics model based on MRI for predicting DM of NPC. The steps include **(A)** MR image acquisition, **(B)** tumor segmentation, **(C)** radiomics features selection, and **(D)** model evaluation.

10.3389/fonc.2022.975881

TABLE 1 Comparison of general characteristics of patients with NPC.

	DM group (n=72)	Non-DM group (n=74)	Statistical value	*P*
Gender (male/female)	58/16	54/20	0.744	0.388[b]
age	47.9 ± 12.5	48.6 ± 12.9	0.362	0.718[a]
Tumor stage			5.943	<0.001[c]
T1	3	15		
T2	13	35		
T3	25	18		
T4	31	6		
Histological type			0.658	0.510[c]
I type	6	9		
II type	34	35		
III type	32	30		

a: t value; b: x^2 value; c: Z value.

4 Discussion

MRI is one of the most commonly used in the early diagnosis and evaluation of NPC, and its sensitivity and resolution of lesions are better than CT images. Although radiotherapy and simultaneous radiotherapy and chemotherapy significantly reduced the local recurrence rate of NPC, the treatment response of patients with DM was poor, and the 5-year survival rate was less than 5%, which was the main cause of treatment failure (10). Therefore, it is necessary to evaluate the risk of metastasis in patients with NPC before treatment in order to adopt more aggressive therapy strategies for high-risk patients. MRI only simply reflected the anatomical structure of tumor invasion, but ignored the heterogeneity of tumor. It was difficult to monitor and evaluate the risk of tumor patients with DM during treatment, which has many limitations. By extracting and analyzing medical images and obtaining quantitative feature data that cannot be recognized by the naked eye, the imaging model can more comprehensively and carefully show the microscopic characteristics and heterogeneity of the tumor. Up to now, few studies have used imaging analysis to predict the occurrence of DM in NPC (11, 12). Most of them used imaging models to predict tumor stage, recurrence, curative effect evaluation and prognosis evaluation (13–18). Zhang (12) reported that combined MRI imaging features with clinical features to evaluate DM risk in patients with NPC before the first treatment, and found that the combined model had good diagnostic efficacy in both training group and validation group. Peng (11) found that the combination of sequence floating forward selection (SFFS) and support vector machine (SVM) classifier can further improve the accuracy of imaging prediction model by analyzing the characteristics of preprocessed PET/CT images for prediction the recurrence and DM in patients with locally advanced NPC.

In this study, we extracted the imaging features of T2WI and CE-T1WI sequences in MR images and constructed a radiomics model based on logistic regression to predict the risk of DM in NPC patients before initiating treatment. The model has higher

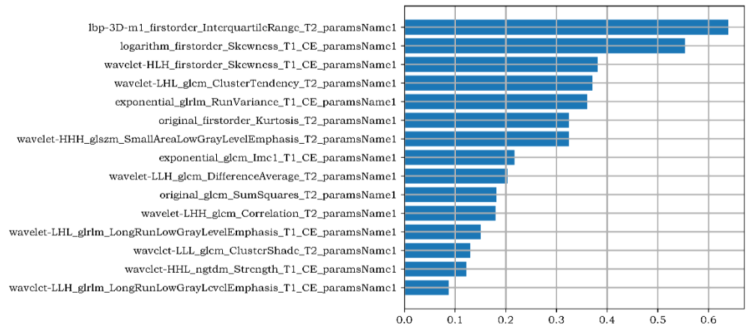


FIGURE 2
The Gini Coefficient importance analysis of radiomics features. The three radiomics features with the highest contribution are the inter quartile range of the first-order feature, the skewness of the first-order feature after logarithm, and the skewness of the first-order feature filtered by high-low- high wavelet filters in XYZ direction,.

10.3389/fonc.2022.975881

TABLE 2 Predictive effectiveness of radiomics model in the training group and validation group.

	training group (n=116)	validation group(n=30)
AUC (95% CI)	0. 80(0.72~0. 88)	0.70(0.51~0. 90)
sensitivity	76.8%	72.7%
specificity	73.3%	63.2%
accuracy	75.0%	66.7%

AUC value, sensitivity, specificity, and accuracy in the training group and validation group, suggesting that it has a good diagnostic performance. Similarly, previous studies have found that the ability of image group feature prediction based on T2WI and CE-T1WI image extraction was better than that of independently T2WI sequence or independently CE-T1WI sequence image extraction (19). T2WI images mainly provide anatomical information, and CE-T1WI images mainly evaluate the blood supply of tumor. The combination of two sequence images is a key factor in judging the prognosis of NPC.

In this study, 1781 radiomics features were extracted. After dimensionality reduction with LASSO, 15 radiomics features were remained, including first-order features (n=4) and second-order GLCM (n=6), GLRLM (n=3), GLSZM (n=1) and NGTDM (n=1). The first-order feature mainly describes the distribution of voxel gray values in ROI. The second-order histogram feature or texture feature is a feature describing the spatial distribution intensity level of voxels. GLCM describes the joint distribution of two gray pixels with a certain spatial position relationship, in which the correlation is the feature extracted from the T2WI image, which reflects the consistency of the image texture. The greater the value difference, the higher the heterogeneity in the tumor. Cluster Tendency is a measure of groupings of voxels with similar gray-level values. DA refers to the relationship between similar intensity values and different intensity values. SS quantifies the distribution of neighbouring intensity level pairs about the mean intensity level. Correlation represents the linear dependency of gray level values to their respective voxels. CS is a measure of the skewness and uniformity of the GLCM: a higher CS implies greater asymmetry about the mean. GLRLM is a quantitative index of the smoothness of image texture, in which the larger the LRLGLE value, the higher the gray value and the smoother of the texture in image. RV is a measure of the variance in runs for the run lengths. Similar to GLRLM, GLSZM, mainly describes the quantitative index of image texture uniformity: SALGLE refers to the proportion in the image of the joint distribution of smaller size zones with lower gray-level values, and this feature is positively correlated with tumor heterogeneity. NGTDM refers to the sum of the difference between a gray value and the average gray value of its neighbours: texture intensity is related to contrast and coarseness, with small values for coarseness textures and high values for busyness or fine textures. Strength is a measure of the primitives in an image. Its value is high when the primitives are easily defined and visible. These features are objective and quantitative information that cannot be observed by the human eye, and usually reflect the pathophysiological information inside the tumor. The results of this study showed that the differences in the characteristic parameter values between the DM group and the non-DM group were in line with the above rules, and the overall texture distribution of the tumors in the DM group was uneven, that is, the inherent

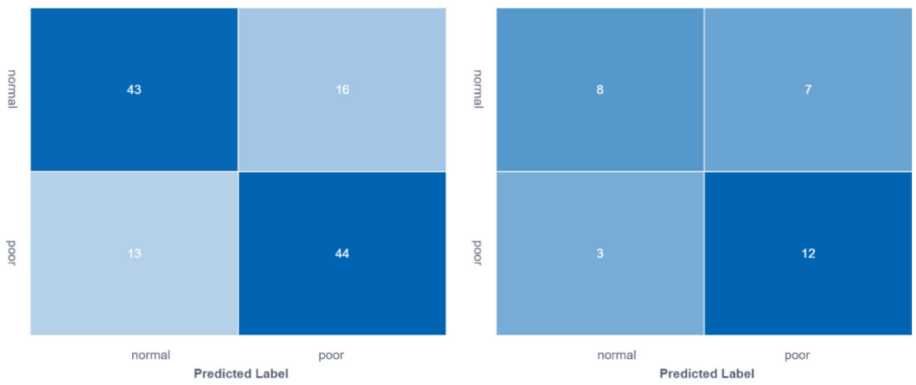


FIGURE 3
Comparison of feature model cross-validation performance between training group and validation group.

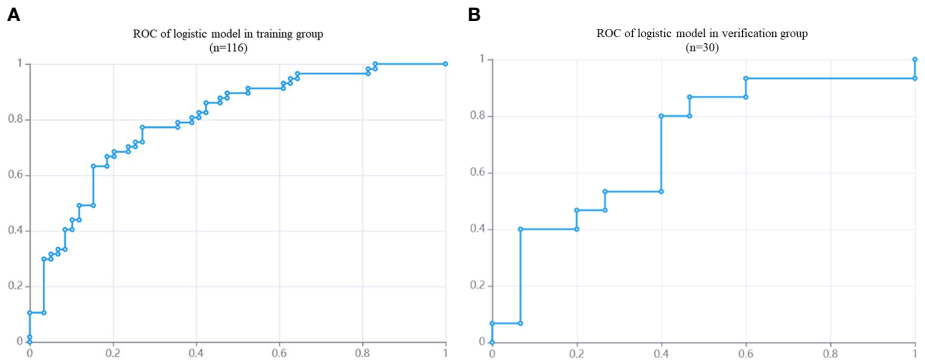


FIGURE 4
The ROC of DM in NPC patients based on MRI radiomics model. **(A)**: the AUC of training group (n = 116) is 0.80. **(B)**: the AUC of validation group (n = 30) is 0.70.

heterogeneity of tumor may provide additional information about pathophysiological features.

There were several limitations in the present study: (1) this study is a retrospective analysis of single-center with a small sample size, which needs to be verified by multicenter, prospective studies with a large sample size, (2) the thick slice thickness (5mm) of the images used in this study may have effects on ROI segmentation, (3) different types of MRI equipment are used for inspection, the scanning parameters are not unified.

5 Conclusions

Radiomics analysis can objectively quantify the morphological and internal heterogeneity changes of NPC, and the radiomics model can effectively evaluate the risk of DM in patients with NPC, which may assist physicians in screening patients with DM and accordingly formulating individualized treatment plans for patients.

Data availability statement

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

Ethics statement

The studies involving human participants were reviewed and approved by Ethics Committees of Zhejiang Cancer Hospital. Written informed consent for participation was not required for this study in accordance with the national legislation and the institutional requirements.

Author contributions

TJ had full access to all of the data in the study and took responsibility for the integrity of the data and the accuracy of the data analysis. Collection and assembly of data: TJ, YT, SN, WC, FW, TL, FL. Resources and data curation: TL, WQ, FJ. Data analysis and interpretation: TJ, YW, HJ, FW. Writing-original draft preparation: TJ. Statistical analysis: TJ, YW, HJ. Study concept, supervision, funding acquisition, project administration, writing-review: HJ and FJ. All authors contributed to the article and approved the submitted version.

Funding

This study was supported by grants from Medical and Health Research Project of Zhejiang Province (Grant Number: 2020KY486), and Natural Science Foundation of Zhejiang Province (2022C03072).

Conflict of interest

Author YW was employed by General Electric (GE) Healthcare.

The remaining 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. Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin* (2018) 68:394–424. doi: 10.3322/caac.21492

2. Sung H, Ferlay J, Siegel RL, Laversanne M, Soerjomataram I, Jemal A, et al. Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin* (2021) 71:209–49. doi: 10.3322/caac.21660

3. Pan JJ, Ng WT, Zong JF, Lee SW, Choi HC, Chan LL, et al. Prognostic nomogram for refining the prognostication of the proposed 8th edition of the AJCC/UICC staging system for nasopharyngeal cancer in the era of intensity-modulated radiotherapy. *Cancer* (2016) 122:3307–15. doi: 10.1002/cncr.30198

4. Setton J, Han J, Kannarunimit D, Wuu YR, Rosenberg SA, Deselm C, et al. Long-term patterns of relapse and survival following definitive intensity-modulated radiotherapy for non-endemic nasopharyngeal carcinoma. *Oral Oncol* (2016) 53:67–73. doi: 10.1016/j.oraloncology.2015.11.015

5. Li Y, Yang X, Du X, Lei Y, He Q, Hong X, et al. RAB37 hypermethylation regulates metastasis and resistance to docetaxel-based induction chemotherapy in nasopharyngeal carcinoma. *Clin Cancer Res* (2018) 24:6495–508. doi: 10.1158/1078-0432.Ccr-18-0532

6. Liu SL, Sun XS, Chen QY, Liu ZX, Bian LJ, Yuan L, et al. Development and validation of a transcriptomics-based gene signature to predict distant metastasis and guide induction chemotherapy in locoregionally advanced nasopharyngeal carcinoma. *Eur J Cancer* (2022) 163:26–34. doi: 10.1016/j.ejca.2021.12.017

7. Zhang L, Huang Y, Hong S, Yang Y, Yu G, Jia J, et al. Gemcitabine plus cisplatin versus fluorouracil plus cisplatin in recurrent or metastatic nasopharyngeal carcinoma: A multicentre, randomised, open-label, phase 3 trial. *Lancet* (2016) 388:1883–92. doi: 10.1016/s0140-6736(16)31388-5

8. Lambin P, Leijenaar RTH, Deist TM, Peerlings J, De Jong EEC, Van Timmeren J, et al. Radiomics: the bridge between medical imaging and personalized medicine. *Nat Rev Clin Oncol* (2017) 14:749–62. doi: 10.1038/nrclinonc.2017.141

9. Van Griethuysen J, Fedorov A, Parmar C, Hosny A, Aucoin N, Narayan V, et al. Computational radiomics system to decode the radiographic phenotype. *Cancer Res* (2017) 77:e104–7. doi: 10.1158/0008-5472.Can-17-0339

10. Zhang L, Wu X, Liu J, Zhang B, Mo X, Chen Q, et al. MRI-Based deep-learning model for distant metastasis-free survival in locoregionally advanced nasopharyngeal carcinoma. *J Magn Reson Imaging* (2021) 53:167–78. doi: 10.1002/jmri.27308

11. Peng L, Hong X, Yuan Q, Lu L, Wang Q, Chen W. Prediction of local recurrence and distant metastasis using radiomics analysis of pretreatment nasopharyngeal [18F]FDG PET/CT images. *Ann Nucl Med* (2021) 35:458–68. doi: 10.1007/s12149-021-01585-9

12. Zhang L, Dong D, Li H, Tian J, Ouyang F, Mo X, et al. Development and validation of a magnetic resonance imaging-based model for the prediction of distant metastasis before initial treatment of nasopharyngeal carcinoma: A retrospective cohort study. *EBioMedicine* (2019) 40:327–35. doi: 10.1016/j.ebiom.2019.01.013

13. Zhuo EH, Zhang WJ, Li HJ, Zhang GY, Jing BZ, Zhou J, et al. Radiomics on multi-modalities MR sequences can subtype patients with non-metastatic nasopharyngeal carcinoma (NPC) into distinct survival subgroups. *Eur Radiol* (2019) 29:5590–9. doi: 10.1007/s00330-019-06075-1

14. Liao H, Chen X, Lu S, Jin G, Pei W, Li Y, et al. MRI-Based back propagation neural network model as a powerful tool for predicting the response to induction chemotherapy in locoregionally advanced nasopharyngeal carcinoma. *J Magn Reson Imaging* (2022) 56:547–59. doi: 10.1002/jmri.28047

15. Xie C, Du R, Ho JW, Pang HH, Chiu KW, Lee EY, et al. Effect of machine learning re-sampling techniques for imbalanced datasets in (18)F-FDG PET-based radiomics model on prognostication performance in cohorts of head and neck cancer patients. *Eur J Nucl Med Mol Imaging* (2020) 47:2826–35. doi: 10.1007/s00259-020-04756-4

16. Du D, Feng H, Lv W, Ashrafinia S, Yuan Q, Wang Q, et al. Machine learning methods for optimal radiomics-based differentiation between recurrence and inflammation: Application to nasopharyngeal carcinoma post-therapy PET/CT images. *Mol Imaging Biol* (2020) 22:730–8. doi: 10.1007/s11307-019-01411-9

17. Bologna M, Corino V, Calareso G, Tenconi C, Alfieri S, Iacovelli NA, et al. Baseline MRI-radiomics can predict overall survival in non-endemic EBV-related nasopharyngeal carcinoma patients. *Cancers (Basel)* (2020) 12:2958. doi: 10.3390/cancers12102958

18. Peng H, Dong D, Fang MJ, Li L, Tang LL, Chen L, et al. Prognostic value of deep learning PET/CT-based radiomics: Potential role for future individual induction chemotherapy in advanced nasopharyngeal carcinoma. *Clin Cancer Res* (2019) 25:4271–9. doi: 10.1158/1078-0432.Ccr-18-3065

19. Zhang B, Tian J, Dong D, Gu D, Dong Y, Zhang L, et al. Radiomics features of multiparametric MRI as novel prognostic factors in advanced nasopharyngeal carcinoma. *Clin Cancer Res* (2017) 23:4259–69. doi: 10.1158/1078-0432.Ccr-16-2910

frontiers | Frontiers in Oncology

TYPE Original Research
PUBLISHED 16 August 2022
DOI 10.3389/fonc.2022.966743

OPEN ACCESS

EDITED BY
Pei Yang,
Central South University, China

REVIEWED BY
Tian-wu Chen,
Affiliated Hospital of North Sichuan Medical College, China
Jianbo Wang,
Shandong University, China

*CORRESPONDENCE
Xiaoping Yin
yinxiaoping78@sina.com
Fei Yang
hiyangfei@126.com

†These authors share first authorship

SPECIALTY SECTION
This article was submitted to Cancer Imaging and Image-directed Interventions, a section of the journal Frontiers in Oncology

RECEIVED 11 June 2022
ACCEPTED 25 July 2022
PUBLISHED 16 August 2022

CITATION
Wang Y, Wang Y-R, Ren J-L, Jia L-Y, Ma L-Y, Yin X-P, Yang F and Gao B-L (2022) Malignancy risk of gastrointestinal stromal tumors evaluated with noninvasive radiomics: A multi-center study. *Front. Oncol.* 12:966743. doi: 10.3389/fonc.2022.966743

Malignancy risk of gastrointestinal stromal tumors evaluated with noninvasive radiomics: A multi-center study

Yun Wang[1†], Yurui Wang[2†], Jialiang Ren[3], Linyi Jia[4], Luyao Ma[1], Xiaoping Yin[1*], Fei Yang[5*] and Bu-Lang Gao[1]

[1]Affiliated Hospital of Hebei University/Hebei University (Clinical Medical College), Baoding, China, [2]Tangshan Gongren Hospital, Tangshan, China, [3]General Electric Pharmaceutical Co., Ltd, Shanghai, China, [4]Xingtai People's Hospital, Xingtai, China, [5]Medical Imaging Department, The First Affiliated Hospital of Hebei North University, Zhangjiakou, China

Purpose: This study was to investigate the diagnostic efficacy of radiomics models based on the enhanced CT images in differentiating the malignant risk of gastrointestinal stromal tumors (GIST) in comparison with the clinical indicators model and traditional CT diagnostic criteria.

Materials and methods: A total of 342 patients with GISTs confirmed histopathologically were enrolled from five medical centers. Data of patients wrom two centers comprised the training group (n=196), and data from the remaining three centers constituted the validation group (n=146). After CT image segmentation and feature extraction and selection, the arterial phase model and venous phase model were established. The maximum diameter of the tumor and internal necrosis were used to establish a clinical indicators model. The traditional CT diagnostic criteria were established for the classification of malignant potential of tumor. The performance of the four models was assessed using the receiver operating characteristics curve.

Reuslts: In the training group, the area under the curves(AUCs) of the arterial phase model, venous phase model, clinical indicators model, and traditional CT diagnostic criteria were 0.930 [95% confidence interval (CI): 0.895-0.965), 0.933 (95%CI 0.898-0.967), 0.917 (95%CI 0.872-0.961) and 0.782 (95%CI 0.717-0.848), respectively. In the validation group, the AUCs of the models were 0.960 (95%CI 0.930-0.990), 0.961 (95% CI 0.930-0.992), 0.922 (95%CI 0.884-0.960) and 0.768 (95%CI 0.692-0.844), respectively. No significant difference was detected in the AUC between the arterial phase model, venous phase model, and clinical indicators model by the DeLong test, whereas a significant difference was observed between the traditional CT diagnostic criteria and the other three models.

Conclusion: The radiomics model using the morphological features of GISTs play a significant role in tumor risk stratification and can provide a reference for clinical diagnosis and treatment plan.

KEYWORDS

gastrointestinal stromal tumors, traditional CT diagnosis, enhance different periods, radiomics, multiple centers

Introduction

Gastrointestinal stromal tumors (GISTs) are the most common mesenchymal tumors in the gastrointestinal tract of middle-aged and elderly (60-70 years old) patients. The common sites of GIST are stomach (50%-60%), small intestine (20%-30%), colorectal (5%-10%), and esophagus (< 5%) (1, 2). GISTs exhibit a specific malignant potential as well as early liver and abdominal metastasis. According to the National Institute of Health (NIH) 2008 standard (3), the risk of GIST can be divided into very low risk, low risk, medium risk, and high risk. Typically, GISTs with a very low or low risk are classified as potential malignant, whereas those with a medium or high risk are classified as malignant. Because of the heterogeneity, different individuals considered different malignant potentials with varied treatment approaches in the same GIST lesion. Clinically, the potentially malignant GISTs are treated as a benign tumor, whereas malignant GISTs are treated with imatinib mesylate and other drugs before or after the operation to prevent recurrence or metastasis (4). The gold standard for malignant diagnosis of GIST is based on the pathological results, including tumor size, mitotic count and tumor site (Table 1) (3, 5). In order to obtain pathological samples of tumor for risk grading and evaluation of the tumor, a puncture biopsy is essential. However, this is an invasive method and might lead to tumor cell metastasis and tumor bleeding. Therefore, risk classification of the tumor should be obtained at the earliest time possible for selection of an appropriate clinical treatment plan. Although computed tomography (CT) is of a great value in detecting GISTs (6), it is still difficult to judge the malignant potential of tumors due to lack of understanding of the images or unclear tumor signs.

In recent years, rapid development in medical imaging analysis and imaging pattern recognition tools has promoted the development of a high-throughput quantitative feature extraction process, the radiomics, which converts images into exploitable data for analysis (7). This technique can be used to diagnose noninvasively the nature of lesions and ultimately assist the radiologist in making an accurate diagnosis. In the evaluation of the malignancy of GISTs, radiomics has been applied using data of ultrasound, magnetic resonance imaging, and CT (8–13). However, no studies have been performed wtih CT data in the arterial and venous phase to extract the radiomics features for evaluation of the malignancy of GISTs. The present study aimed to explore the radiomics diagnostic models of GIST with different degrees of risk based on the CT image data in the arterial phase and venous phase from five medical centers, with four models being estalibshed, including the arterial phase model, venous phase model, clinical indicators model, and traditional CT diagnostic criteria. The data in two centers were set up as the training group to reduce the sampling bias and to establish a more ubiquitous radiomics model than those in one center only, with the slice thickness of images as 5 mm (14). The diagnostic efficiency was also evaluated to find the best model to guide the correct clinical decision-making process.

Materials and methods

Patients

This retrospective study was approved by the Institutional Review Board of the Affiliated Hospital of Hebei University, and all patients had given their signed informed consent to

TABLE 1 NIH 2008 criteria for risk stratification of GIST recurrence after surgery.

Risk category	Tumor size (cm)	Mitotic index (per 50 HPF)	Location
Very low risk	≤ 2.0	≤ 5.0	Any
Low risk	2.1-5.0	≤ 5.0	Any
Intermediate risk	≤ 5.0	6-10	Gastric
	5.1-10.0	≤ 5.0	Gastric
High risk	>10.0	Any	Any
	Any	>10	Any
	>5.0	>5	Any
	≤ 5.0	>5	Non-gastric
	5.1-10.0	≤ 5	Non-gastric

GIST, gastrointestinal stromal tumor; HPF, high-power field.

participate. All methods were performed in accordance with the relevant guidelines and regulations. The data of CT images of 342 patients were collected from five medical centers from January 2015 to August 2021. Two centers were randomly selected and assigned to the training group, and the data of the other three centers were set up as the validation group.

The inclusion criteria were patients with GISTs confirmed by pathology, complete clinical and pathological data (lesion size, origin location, and risk classification), and standard dynamic enhanced CT scan at least 15 days before the operation. The exclusion criteria were patients with a previous history of other coexisting malignant tumors, neoadjuvant chemoradiotherapy before CT scan, and poor image quality precluding quantitative analyses. The selection process of patient cohorts is shown in Figure 1.

CT scanning instruments and methods

The Discovery CT750 HD scanner (GE Medical Systems, Milwaukee, WI, USA), Toshiba Aquilion 64-slice spiral CT scanner (Tokyo harbor area, Japan), Philips 256-slice ICT scanner

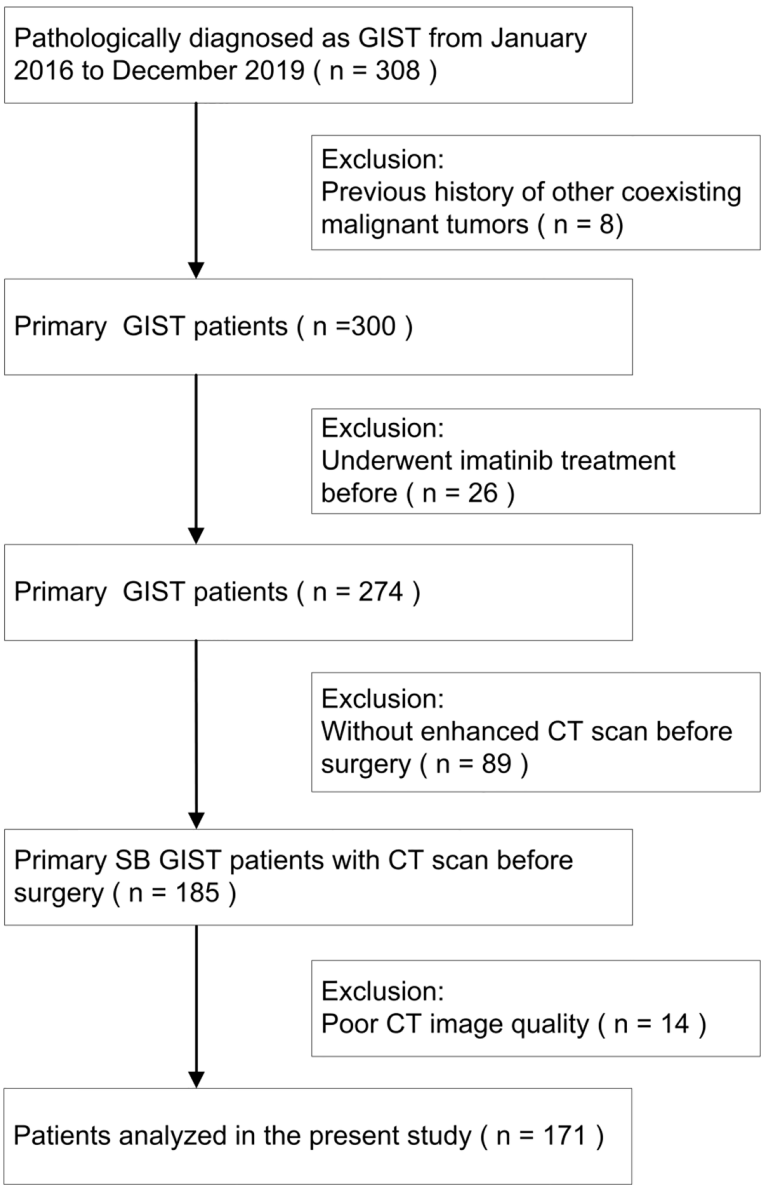


FIGURE 1
A flowchart shows selection of study population and exclusion criteria.

(Amsterdam, The Netherlands), and Philips brilliance 64-slice CT scanner (Amsterdam, The Netherlands) were used for CT scanning. After fasting for 6-8 h, the patient had warm water (500-1000 mL) 10 min before the examination with plain and enhanced abdominal scanning in the supine position. The scanning parameters were as follows: slice thickness 5 mm, pitch 0.9-1.0, scanning field 350 mm×350 mm, matrix 512×512, tube voltage 100–120 kV, tube current 160–300 mA, and X-ray tube rotation time 0.5–0.8 s. The contrast agent was injected through the elbow vein at a flow rate of 3.0–3.5 mL/s and a dose of 1.0–1.2 mL/kg body weight. The scanning time of the arterial phase, venous phase, and delayed phase was 30-35 s, 50-60 s, and 180 s, respectively, after injection of contrast agent. The CT images at the arterial and venous phase were selected for imaging analysis.

Clinical data

The clinical data including age, gender of patients, and tumor location were collected based on pathological results. The imaging data including tumor maximal diameter and necrosis within the tumor lesion were collected based on CT imaging. In the malignant potential classification using the traditional CT diagnosis method, the CT images were assessed by five radiologists (with 19, 15, 10, 8 and 4 years of working experience, respectively) who were blinded to the pathological diagnoses in all cases. The tumor is divided into potentially malignant and malignant according to the CT image characteristics, including tumor size, location, shape, boundary, enhancement mode and degree, infiltration of peripheral organs, and lymph node enlargement (15–17). In disagreement, a consensus was reached after discussion.

CT image segmentation

Two radiologists (physicians 1 and 2) with 10 years of experience in the abdominal imaging diagnosis applied the ITK-SNAP software (version 3.8.0, https://www.itksnap.org) to delineate the CT-enhanced images at the arterial and venous phases. The delineated areas included the tumor lesion as much as possible without inclusion of the surrounding normal tissues or other tissues in order to generate a two-dimensional (2D) region of interest (ROI) (Figure 2). The 2D ROI was then recombined to generate a 3D volume of interest (VOI) for subsequent image feature extraction and analysis.

Radiomic feature extraction and selection

CT images with different scanning parameters were preprocessed. The linear interpolation method was used to resample the image to 1×1×1 mm^3, attempting to alleviate the influence of different layer thicknesses. The image gray was

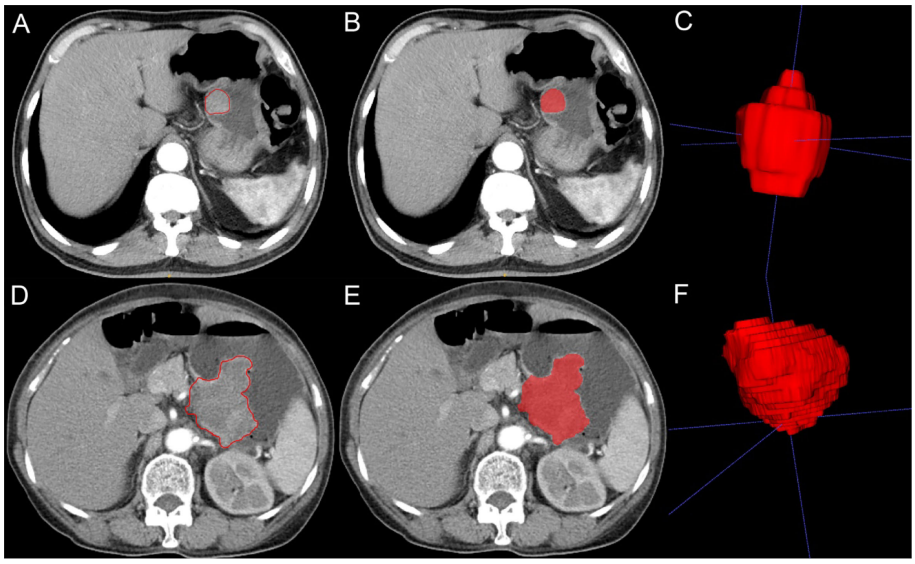


FIGURE 2
Imaging segmentation of gastrointestinal stromal tumors (GIST) on computed tomography (CT) imaging. **(A)** Two-dimensional (2D) CT arterial phase image of potential malignant GIST. The tumor is quasicircular and uniformly enhanced. The red outline is the boundary drawn by radiologists to show the tumor. **(B)** 2D segmentation of the tumor. **(C)** 3D segmentation of the tumor lesion. **(D)** 2D CT arterial phase image of malignant GIST with irregular shape and uneven internal enhancement of the tumor. The red outline is the boundary drawn by radiologists to show the tumor lesion. **(E)** 2D segmentation of the tumor. **(F)** 3D segmentation of the tumor.

discretized, the bin width was set to 25, and the image filtering process was used to highlight different bandwidth signals and prevent the noise in ROI from interfering with texture information. Parameters σ 3 and 5 of Laplace of Gaussian (LOG) filter and wavelet were used. After wavelet decomposition, eight categories of information of the filtered features were obtained from the original set of feature information. The preprocessed image and the outlined ROI files were imported into the platform of "pyradiomics" for radiomics feature extraction, and two sets of image filtering were utilized. A total of 1037 features were generated from the histogram (18 features), morphological feature (14 features), texture feature of gray level co-occurrence matrix (GLCM, 24 features), gray-level run-length matrix (GLRLM, 16 features), gray-level size zone matrix (GLSZM, 16 features), gray-level dependence matrix (GLDM, 14 features), and neighborhood gray-tone difference matrix (NGTDM, 5 features).

In order to prevent overfitting risk, it was necessary to reduce the dimension of data features and select those with the best efficiency and most research significance. A total of 20 important features were selected using the minimum redundancy maximum relevance feature selection (mRMR) based on relevant references in the literature to prevent distortion of model (18–20). Subsequently, the least absolute contraction selection operator (LASSO) was used to further eliminate collinear features, and 5 features in the arterial phase and 11 features in the venous phase were kept.

Radiomics models building

The arterial phase model and venous phase model of radiomics: After feature selection, 5 features of arterial phase and 11 features of venous phase were used to retain the minimum Akaike information criterion (AIC) feature set by the multifactor stepwise regression.

Clinical indicators model of radiomics: The clinical indicators of the tumor diameter (>5 cm) and internal necrosis were used as clinical features. The model was established by multiple regression analysis with the maximal diameter of the tumor and presence of necrosis as the feature of the model and benign or malignant nature as the goal.

Traditional CT diagnostic criteria of radiommics: According to the consolidated GIST CT diagnosis results of the five evaluating radiologists, the traditional CT diagnostic criteria were used for classification of the malignant potential of the tumors.

Sample size estimation

In the training group, 196 consecutive patients were enrolled in two centers between January 2016 and December 2019. The training cohort contained 58 low-risk GIST patients and 138 high-risk patients. There were in total 2 predictors in our model (internal tumor necrosis and tumor diameter), making an event-per-predictor ratio of large than 10, which fell in the range of 5-9 in the rule of thumb for event-per-predictor in logistic regression models (21). In the validation group, the validation sample size was determined according to the method of sample size estimation for clinical research by Chow and colleagues (22), with the sample size being calculated to test whether the means of two groups were significantly different. Based on this method, the mimimal number of validation sameples were 14 (low-risk) and 36 (high-risk) in the group with the desired two-sided significance level of α=0.05 and power of 1-β=95%.

Statistical analysis

All statistical analyses were performed using the R software (version 4.1.0, www.rporject.org). Measurement data were presented as median [Q1-Q3] if in non-normal distribution and tested with the Mann-Whitney U test, and enumeration data were expressed as numbers of cases (n) or percentage (%) and tested with the Chi square test. The non-normal distribution data of measurement were presented as median and interquartile range and tested with the Chi square test. Interclass and intraclass correlation coefficients (ICC) were used to evaluate the consistency of imaging features within and between observers. A total of 30 cases of CT images were randomly selected for ROI segmentation by physicians 1 and 2. One week later, physician 1 repeated the same steps, with an ICC >0.75 indicating good consistency in feature extraction. The segmentation of the remaining image was also completed by physician 1. The receiver operator characteristic (ROC) curve was used to evaluate the predictive efficacy of the malignant potential of GISTs in the models. The larger the area under the ROC curve (AUC), the higher the diagnostic efficiency. The AUC, accuracy, sensitivity, and specificity were calculated, and the ROC curves were assessed by the Delong test. All indexes were evaluated separately in the training and validation groups. Two-side P<0.05 was set as statistic significant.

Results

Clinical characteristics

According to the inclusion and exclusion criteria, data from 342subjects with GISTs, including 156 (45.6%) males and 186 (54.4%) females with an age range 33-82 (62.00 [54.00-69.00]) years, were collected. The GIST lesion was in the stomach in 226 (66.1%) cases and of a non-stomach location in 106 (33.9%), including 104 (30.4%) cases with potential malignancy (26 cases with an extremely low risk and 78 cases with a low risk) and 238 (69.6%) cases with malignancy (86 cases with a moderate risk and 152 cases with a high risk). According to the traditional CT

TABLE 2 Clinical data of the training and validation groups.

Variables	Training (n=196)	Validation (n=146)	P
Gender			0.181[1]
Female	112	74	
Male	84	72	
Age [median, Q1-Q3]	62.000 [56.000-69.000]	63.000 [52.000-69.000]	0.508[2]
Real malignant potential			0.793[1]
Potential malignancy	58(29.6%)	46(31.5%)	
Malignant	138(70.4%)	100(68.5%)	
Traditional CT classification			0.990[1]
Potential malignancy	64	48	
Malignant	132	98	
Maximal diameter ≥5 cm			0.847[1]
No	72	56	
Yes	124	90	
Internal necrosis			0.006[1]
No	82	84	
Yes	114	62	

Q1,First quarter; Q3, Three quarter; [1] Chi square test; [2]Mann-Whitney U test.

diagnosis criteria for GISTs, 130 (38.0%) cases exhibited potential malignancy, whereas 212 (62.0%) cases were malignant. The diameter of the tumors was 1-24 (mean 6.9 ± 4.1) cm, with the tumor maximal diameter ≥5 cm in 214 (62.6%) cases and <5 cm in 128 (37.4%). Internal necrosis was presented in 176 (51.5%) cases. In the radiomics model, the patients were divided into the training (n=196) and validation (n=146) group (Table 2), with no significant (P>0.05) difference in the age, gender, malignancy potential, CT diagnostic grade, tumor maximal diameter >5 cm, and internal necrosis between the training and validation groups.

Univariable and multivariable analysis

In univariate analysis of GIST parameters, the tumor maximal diameter and internal necrosis were statistically significant (P<0.001) between potentially malignant and malignant GISTs (Table 3). Using the significant variables from the univariate analysis as inputs, multivariate logistic regression analysis showed that lesion diameter ≥ 5cm (coeficient 3.264, OR 26.17 (7.832-109.083), P<0.001) and lesion internal necrosis (coefficient 2.014, OR 7.491 (1.969-31.461), P=0.003) were independent factors for predicting malignant GIST.

ICC of radiomic features

A total of 1037 radiomics features with good consistency (mena ICC 0.95, range 0.75-1.0) were selected, whereas 95 features with bad consistency (ICC <0.75) were removed.

Predictive performance of radiomics models

After feature selection, the radiomics features of the arterial and venous phases only preserved the morphological features. ROC curve analyses were performed for the arterial and venous phase models, clinical indicators model, and traditional CT diagnostic criteria (Figure 3), with a good calibration demonstrated in the arterial and venous phase models (Figure 4). The Radscore distribution of the arterial and venous phase models in the training and validation group were shown in Figure 5.

In the ROC curve analyses for the training group, the AUC, accuracy, sensitivity, and specificity for grading tumor malignancy were 0.930 (95%CI: 0.895-0.965), 0.888, 0.928, and

TABLE 3 Univariable analysis of potentially malignant and malignant GISTs.

Variables	Training group			Validation group		
	Potentially malignant (n=104)	Malignant (n=238)	*P*	Potentially malignant (n=104)	Malignant (n=238)	*P*
Sex			0.910[1]			0.316[1]
Female	34(58.621%)	78(56.522%)		20(43.478%)	54(54.000%)	
Male	24(41.379%)	60(43.478%)		26(56.522%)	46(46.000%)	
Age[meadian, Q1-Q3]	61.000 [58.000-68.000]	63.000 [55.000-69.000]	0.683[2]	64.000 [54.250-71.500]	62.500 [51.000-68.000]	0.172[2]
Diameter ≥5 cm			<0.001[1]			<0.001[1]
No	54(93.103%)	18(13.043%)		42(91.304%)	14(14.000%)	
Yes	4(6.897%)	120(86.957%)		4(8.696%)	86(86.000%)	
Internal necrosis			<0.001[1]			<0.001[1]
No	54(93.103%)	28(20.290%)		46(100.000%)	38(38.000%)	
Yes	4(6.897%)	110(79.710%)		0(0.000%)	62(62.000%)	

GIST, gastrointestinal stromal tumors; Q1, First quarter; Q3, Three quarter; [1] Chi square test; [2] Mann-Whitney U test.

0.793, respectively, for the arterial phase model; 0.933 (95%CI: 0.898-0.967), 0.857, 0.855, and 0.862, respectively, for the venous phase model; 0.917 (95%CI: 0.872-0.961), 0.918, 0.913, and 0.931, respectively, for the clinical indicators model;, and 0.782 (95%CI: 0.717-0.848), 0.806, 0.841, and 0.724, respectively, for the traditional CT diagnostic criteria. In the validation group, the AUC, accuracy, sensitivity, and specificity were 0.960(95% CI: 0.930-0.990), 0.932, 0.920, and 0.957, respectively, for the arterial phase model; 0.961 (95%CI: 0.930-0.992), 0.932, 0.920, and 0.957, respectively, for the venous phase model; 0.922 (95% CI: 0.884-0.960), 0.890, 0.880, and 0.913, respectively, for the clinical indicators model; and 0.768 (95%CI: 0.692-0.844), 0.795, 0.840 and 0.696, respectively, for the traditional CT diagnostic criteria (Table 4).

Comparison of the AUC values in grading tumor malignancy between different models using the Delong test was performed (Table 5). No significant (P>0.05) difference was detected in the AUC between the arterial and venous phase models, and clinical indicators model, whereas significant (P<0.01) differences were detected between the traditional CT diagnostic criteria (CT) and any of the other three models. The AUC value was significantly (P<0.01) better in the arterial phase model, venous phase model, and clinical indicators model than that in the traditional CT diagnostic criteria.

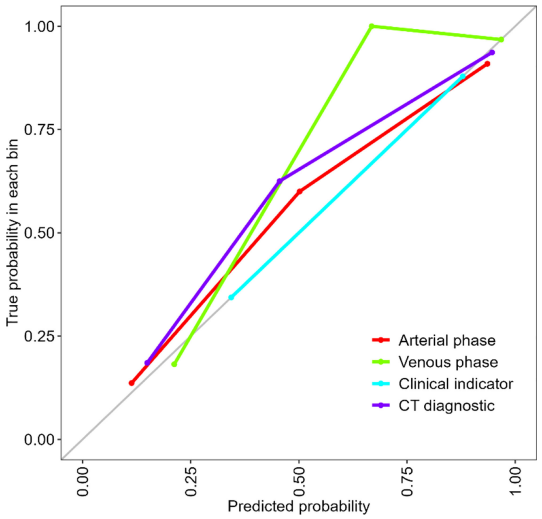


FIGURE 4
Calibration curve for the arterial phase model, venous phase model, clinical indicators model and the traditional CT diagnostic criteria. The calibration of the four models was depicted by the calibration curve in terms of the agrement between the predicted risks of gastrointestinal stromal tumors (GISTs) and the actual results based on the modified criteria. The grey line represents an ideal prediction, and the other lines represent the predictive performance of the models. The closer the fit of the purpole line to the ideal line, the better the prediction.

Discussion

This study investigated the value of radiomics models in grading tumor malignancy of GISTs using enhanced CT imaging data from five medical centers, and four radiomics models were established based on the morphological features of the arterial and venous phase, clinical indicators, and traditional CT diagnostic criteria for GISTs. The models of the arterial phase, venous phase, and clinical indicators were significantly better than the traditional CT diagnostic criteria in grading the tumor malignancy of GISTs.

After studying CT venous phase images and radiomics features of GISTs in 222 cases including one training group (n=130) and one validation group (n=92) in the raiomics, Chen

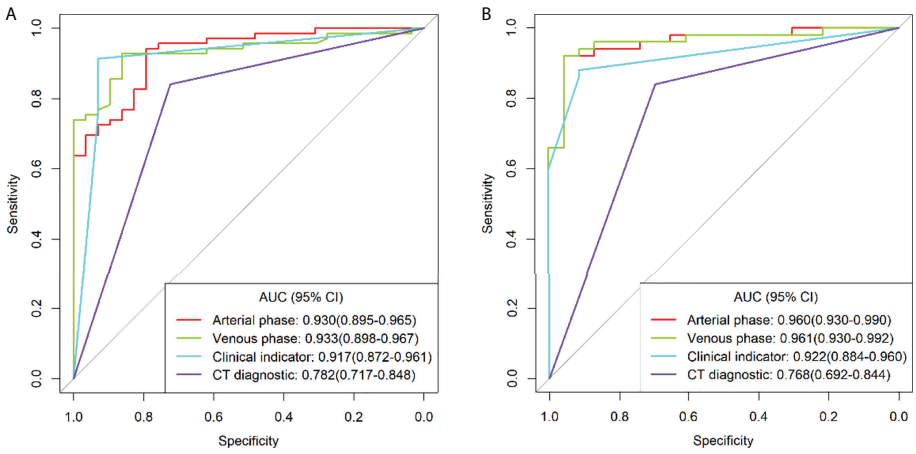


FIGURE 3
Receiver operating characteristics (ROC) curve analysis for different models in the training group **(A)** and validation group **(B)**.

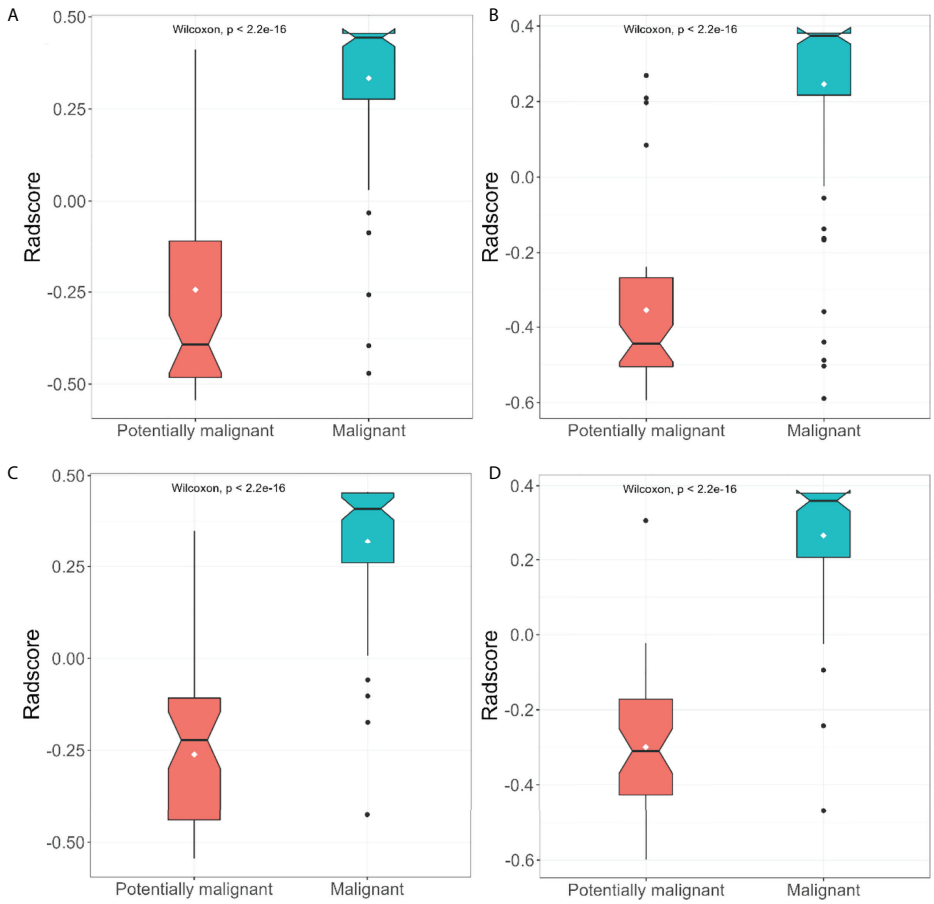


FIGURE 5
The distribution of arterial and venous phase model radscore between patients suffered from malignant and potiential malignant tumors in the training group **(A, B)** and the validation group **(C, D)**.

TABLE 4 Effectiveness of radiomics models in the grading of GIST malignancy.

Model	Training group(n=196)				Validation group(n=146)			
	AUC	Accuracy	Sensitivity	Specificity	AUC	Accuracy	Sensitivity	Specificity
A	0.930	0.888	0.928	0.793	0.960	0.932	0.920	0.957
V	0.933	0.857	0.855	0.862	0.961	0.932	0.920	0.957
Clinical	0.917	0.918	0.913	0.931	0.922	0.890	0.880	0.913
CT	0.782	0.806	0.841	0.724	0.768	0.795	0.840	0.696

GIST, gastrointestinal stromal tumor; AUC, area under the receiver operator characteristic curve; A, arterial phase model; V, venous phase model; Clinical, clinical indicators model; CT, traditional CT diagnostic criteria.

et al. (9) found that the radiomics features combined with clinical indicators and traditional CT characteristics were more effective in judging the malignant potential of GISTs as compared to the clinical indicators or traditional CT characteristic models. Through investigating 339 cases of GISTs from four centers including the training group (n=148), internal verification group (n=41), and external validation group (n=150), Zhang et al. (23) found that the radiomics features of enhanced CT were significantly correlated with the expression of Ki-67 in GISTs and that the tumor size had the highest prediction accuracy of Ki-67 expression. Wang et al. (24) established a radiomics model to predict the malignant potential and mitotic count of GISTs by analyzing the portal venous-phase images of 333 GIST cases, and it was also found

TABLE 5 Comparison of AUC results among models by the Delong test.

Model	Training group *P*	Validation group *P*
A-V	0.879	0.897
A-Clinical	0.590	0.013
V-Clinical	0.439	0.013
A-CT	<0.001	<0.001
V-CT	<0.001	<0.001
Clinical-CT	<0.001	<0.001

AUC, area under the receiver operator characteristic curve; A, arterial phase model; V, venous phase model; Clinical, clinical indicators model; CT, traditional CT diagnostic criteria.

that the combination of radiomics features, subjective CT examination results, and clinical indicators could be used to realize individualized risk prediction and improve the diagnostic level. However, these studies only selected the venous phase of GIST images with enhanced scanning as the research object, and no studies have investigated the difference in the radiomics characteristics of GISTs between the arterial phase and venous phase. Moreover, the data of the training group were from one single center, lacking multicenter data and consequently efficiency for generalization.

The texture performance of enhanced CT images at different periods varies, and to set up an appropriate radiomics model, it is crucial to select the texture features at different enhancement phases such as those of the arterial phase and venous phase. Several investigators have studied the CT enhancement degree of GIST, albeit different in the conclusions (1, 25–30). With the increase of GIST risk stratification, some researchers had found a declining trend in the CT value at each phase of enhanced scanning (28), whereas others had revealed that the degree of GIST enhancement was not related to risk classification (16). In addition, some investigators (17) had demonstrated that the GIST of the small intestine was highly malignant, with the tumor enhancement degree equal to that of adjacent intestinal wall. In case of an unclear correlation between tumor risk and CT enhancement degree, the radiomics features of CT images at different enhancement periods were used to stratify the GIST risk. Liu et al. (1) evaluated 78 patients with GISTs and found significant differences in the CT texture parameters with different GIST risks between the arterial phase and venous phase. Feng et al. (25) found that the entropy value at the venous phase was more accurate in distinguishing low-risk small bowel GIST from medium- and high-risk small bowel GIST as compared to that at the arterial phase. Also, some studies established radiomics models based on the CT arterial phase images (30) or venous phase images (29) so as to provide a noninvasive detection method for prediction of potential malignancy and malignancy of the GIST. Our study was based on GIST data from multicenters, and after extracting and selecting the radiomic features of the arterial and venous phases, only one morphological feature remained: the maximal diameter of the tumor. With only one morpholigcal feature left, good consistency could be easily obtained in the tumor delineation process, with similar efficiency in the arterial phase and venous phase features. The fact that there were no other radiomics features left could be attributed to the GIST data from multiple centers. Strikingly, the imaging parameter settings and scanning parameters of different CT scanners manufactured by different companies varied greatly, which may cause inconsistency in the data of radiomics model. When the image was analyzed and extracted in the digital form, the differences between extracted texture features might lead to some potential changes in the acquired images (26), which need further investigation for confirmation.

Although the AUC of the clinical indicators model was lower than that of the arterial and venous phase models, its specificity was improved as compared to the latter two models. Tumors with a large volume or a large diameter was more likely to have internal necrosis than those with a small volume. The internal necrosis of tumors exhibited uneven enhancement on enhanced CT imaging. In one study (27) investigating tumor location, size, shape, tumor growth, imaging enhancement mode and degree, tumor necrosis percentage, and distant metastasis on CT imaging in 42 patients with GISTs, it was found that the malignant degree of GIST can be predicted from the location, size, and necrosis rate of the tumor. Another study (31) evaluating 1303 patients with GISTs showed that tumor size >5 cm was significantly correlated with the increased rate of tumor recurrence. Tumor size had also been found to be of important diagnostic value in the risk classification of GISTs, irrespective of the NIH standard, AFIP standard, or AJCC staging system (32). It can be seen that the maximal diameter and internal necrosis of GISTs are significant in clinical diagnosis of potential malignancy and malignant tumors, as our study had confirmed the significant role of tumor morphology at the arterial and venous phase.

The traditional CT diagnostic criteria of radiommics showed low efficiency in the diagnosis of GISTs, with a significantly low AUC value compared with the other three models. Accurate diagnosis of the GIST tumor is closely related to the experience of the radiologists and appropriate understanding of tumor signs, especially atypical CT signs which may make differential diagnosis even more difficult.

Currently, some radiomics studies on grading the GIST malignant degree have been performed using ultrasound and magnetic resonance imaging besides CT imaging data (12, 33–39). Liu et al. (35) applied multicenter endoscopic ultrasound imaging data of 914 patients to set up a triple normalization-based deep learning framework with ultrasound-specific pretraining and meta attention (TN-USMA model) to automatically grading high- and low-risk GISTs. In comparing the diagnostic performance of one radiomics-based method and two state-of-the-art deep learning approaches, the TN-USMA model which was composed of intensity normalization, size normalization, and spatial resolution normalization achieved an overall accuracy of 0.834 (95% CI 0.772-0.885), an AUC of 0.881 (95% CI 0.825, 0.924), a sensitivity of 0.844, and a specificity of 0.832. Although the AUC

of the TN-USMA model significantly outperformed the other two deep learning approaches ($P < 0.05$), it was less superior to our models of radiomics. Yang et al. (12) employed the magnetic resonance diffusion-weighted imaging (DWI) data of 91 patients with pathologically-confirmed GIST for radiomic model establishment and risk stratification, and the nomogram incorporating the texture signature features, maximal tumor diameter and location demonstrated a good discriminating effect of GIST malignancy with an AUC of 0.878 in the training and 0.903 in the validation group, suggesting that the texture-based model could be used to predict the mitotic index and risk potential of GISTs before surgery. Other radiomics models based on magnetic resonance imaging data of T1WI, T2WI, and ADC (apparent diffusion coefficient) had also be investigated in grading the malignant risk of GISTs (36), although with good effects in differentiating high-, intermediate- and low-risk GISTs, the AUC value was below 0.85 for T1WI, T2WI, and ADC. In radiomics models based on CT imaging data without the use of internal tumor necrosis and tumor size for evaluating the malignant risk of GISTs (33, 34, 37–39), good effects had been achieved on distinguishing high- and low-risk malignancy, but the AUC values were all below 0.90. In our study, the radiomics models using the internal tumor necrosis and tumor diameter imgaging data at the arterial and venous phases achieved an AUC value over 0.93 in both the training and validation group, suggesting a greater value of these radiomics models in differentiating the malignant risk of GISTs.

Some limitations existed in our study including the retrospective nature, Chinese patients enrolled only, a small cohort of patients at each center, and differences in the CT scanners and scanning parameters. All these issues may affect the publication bias, and the results should be explained in caution. Future studies will have to resolve these issues for better performances.

In conclusion, the morphological radiomic features of GISTs play a significant role in tumor risk stratification and can provide a reference for clinical diagnosis and treatment plan.

Data availability statement

The original contributions presented in the study are included in the article/Supplementary Material. Further inquiries can be directed to the corresponding authors.

Ethics statement

The studies involving human participants were reviewed and approved by Ethics committee of Affiliated Hospital of Hebei University. The patients/participants provided their written informed consent to participate in this study.

Author contributions

JR and XY designed the study. WY, YW, LJ, and LM collected the data. WY, JR, and B-LG analyzed the data. FY supervised the study. WY wrote the original version of article. B-LG Revised the original version. All authors approved the article.

Funding

This study was supported by the Hebei University, Health Commission of Hebei Province and Hebei Natural Fund Project (2020B05, G2019041 and 2021201017).

Conflict of interest

Author JR was employed by General Electric Pharmaceutical (Shanghai) Co., Ltd., Shanghai, China.

The remaining 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. Liu S, Pan X, Liu R, Zheng H, Chen L, Guan W, et al. Texture analysis of ct images in predicting malignancy risk of gastrointestinal stromal tumours. *Clin Radiol* (2018) 73:266–74. doi: 10.1016/j.crad.2017.09.003

2. Søreide K, Sandvik OM, Søreide JA, Giljaca V, Jureckova A, Bulusu VR. Global epidemiology of gastrointestinal stromal tumours (gist): A systematic review of population-based cohort studies. *Cancer Epidemiol* (2016) 40:39–46. doi: 10.1016/j.canep.2015.10.031

3. Joensuu H. Risk stratification of patients diagnosed with gastrointestinal stromal tumor. *Hum Pathology* (2008) 39:1411–9. doi: 10.1016/j.humpath.2008.06.025

4. Mantese G. Gastrointestinal stromal tumor: Epidemiology, diagnosis, and treatment. *Curr Opin Gastroenterol* (2019) 35:555–9. doi: 10.1097/MOG.0000000000000584

5. Cannella R, Tabone E, Porrello G, Cappello G, Gozzo C, Incorvaia L, et al. Assessment of morphological ct imaging features for the prediction of risk stratification, mutations, and prognosis of gastrointestinal stromal tumors. *Eur Radiol* (2021) 3111:8554–64. doi: 10.1007/s00330-021-07961-3

6. Wang JK. Predictive value and modeling analysis of msct signs in gastrointestinal stromal tumors (gists) to pathological risk degree. *Eur Rev Med Pharmacol Sci* (2017) 21:999–1005.

7. Gillies RJ, Kinahan PE, Hricak H. Radiomics: Images are more than pictures, they are data. *Radiology* (2016) 278:563–77. doi: 10.1148/radiol.2015151169

8. Cannella R, La Grutta L, Midiri M, Bartolotta TV. New advances in radiomics of gastrointestinal stromal tumors. *World J Gastroenterol* (2020) 26:4729–38. doi: 10.3748/wjg.v26.i32.4729

9. Chen T, Ning Z, Xu L, Feng X, Han S, Roth HR, et al. Radiomics nomogram for predicting the malignant potential of gastrointestinal stromal tumours preoperatively. *Eur Radiol* (2019) 29:1074–82. doi: 10.1007/s00330-018-5629-2

10. Li X, Jiang F, Guo Y, Jin Z, Wang Y. Computer-aided diagnosis of gastrointestinal stromal tumors: A radiomics method on endoscopic ultrasound image. *Int J Comput Assist Radiol Surg* (2019) 14:1635–45. doi: 10.1007/s11548-019-01993-3

11. Ning Z, Luo J, Li Y, Han S, Feng Q, Xu Y, et al. Pattern classification for gastrointestinal stromal tumors by integration of radiomics and deep convolutional features. *IEEE J BioMed Health Inform* (2019) 23:1181–91. doi: 10.1109/JBHI.2018.2841992

12. Yang L, Zheng T, Dong Y, Wang Z, Liu D, Du J, et al. Mri texture-based models for predicting mitotic index and risk classification of gastrointestinal stromal tumors. *J Magn Reson Imaging* (2021) 53:1054–65. doi: 10.1002/jmri.27390

13. Zhang QW, Zhou XX, Zhang RY, Chen SL, Liu Q, Wang J, et al. Comparison of malignancy-prediction efficiency between contrast and non-contract ct-based radiomics features in gastrointestinal stromal tumors: A multicenter study. *Clin Transl Med* (2020) 10:e291. doi: 10.1002/ctm2.91

14. Park S, Lee SM, Do K-H, Lee J-G, Bae W, Park H, et al. Deep learning algorithm for reducing ct slice thickness: Effect on reproducibility of radiomic features in lung cancer. *Korean J radiology* (2019) 20:1431–40. doi: 10.3348/kjr.2019.0212

15. Chen T, Xu L, Dong X, Li Y, Yu J, Xiong W, et al. The roles of ct and eus in the preoperative evaluation of gastric gastrointestinal stromal tumors larger than 2 cm. *Eur Radiol* (2019) 29:2481–9. doi: 10.1007/s00330-018-5945-6

16. Iannicelli E, Carbonetti F, Federici GF, Martini I, Caterino S, Pilozzi E, et al. Evaluation of the relationships between computed tomography features, pathological findings, and prognostic risk assessment in gastrointestinal stromal tumors. *J Comput Assist Tomogr* (2017) 41:271–8. doi: 10.1097/RCT.0000000000000499

17. Maldonado FJ, Sheedy SP, Iyer VR, Hansel SL, Bruining DH, McCollough CH, et al. Reproducible imaging features of biologically aggressive gastrointestinal stromal tumors of the small bowel. *Abdom Radiol (NY)* (2018) 43:1567–74. doi: 10.1007/s00261-017-1370-6

18. Linning E, Lu L, Li L, Yang H, Schwartz LH, Zhao B. Radiomics for classification of lung cancer histological subtypes based on nonenhanced computed tomography. *Acad Radiol* (2019) 26:1245–52. doi: 10.1016/j.acra.2018.10.013

19. Parmar C, Grossmann P, Bussink J, Lambin P, Aerts H. Machine learning methods for quantitative radiomic biomarkers. *Sci Rep* (2015) 5:13087. doi: 10.1038/srep13087

20. Rios Velazquez E, Parmar C, Liu Y, Coroller TP, Cruz G, Stringfield O, et al. Somatic mutations drive distinct imaging phenotypes in lung cancer. *Cancer Res* (2017) 77:3922–30. doi: 10.1158/0008-5472.CAN-17-0122

21. Vittinghoff E, McCulloch CE. Relaxing the rule of ten events per variable in logistic and cox regression. *Am J Epidemiol* (2007) 165:710–8. doi: 10.1093/aje/kwk052

22. Chow SC, Wang H, Shao J. *Sample Size Calculations in Clinical Research*, *Second* Edition. Taylor & Francis, (2007).

23. Zhang QW, Gao YJ, Zhang RY, Zhou XX, Chen SL, Zhang Y, et al. Personalized ct-based radiomics nomogram preoperative predicting ki-67 expression in gastrointestinal stromal tumors: A multicenter development and validation cohort. *Clin Transl Med* (2020) 9:12. doi: 10.1186/s40169-020-0263-4

24. Wang C, Li H, Jiaerken Y, Huang P, Sun L, Dong F, et al. Building ct radiomics-based models for preoperatively predicting malignant potential and mitotic count of gastrointestinal stromal tumors. *Transl Oncol* (2019) 12:1229–36. doi: 10.1016/j.tranon.2019.06.005

25. Feng C, Lu F, Shen Y, Li A, Yu H, Tang H, et al. Tumor heterogeneity in gastrointestinal stromal tumors of the small bowel: Volumetric ct texture analysis as a potential biomarker for risk stratification. *Cancer Imaging* (2018) 18:46. doi: 10.1186/s40644-018-0182-4

26. Lu L, Liang Y, Schwartz LH, Zhao B. Reliability of radiomic features across multiple abdominal ct image acquisition settings: A pilot study using acr ct phantom. *Tomography* (2019) 5:226–31. doi: 10.18383/j.tom.2019.00005

27. Mazzei MA, Cioffi Squitieri N, Vindigni C, Guerrini S, Gentili F, Sadotti G, et al. Gastrointestinal stromal tumors (gist): A proposal of a "ct-based predictive model of miettinen index" in predicting the risk of malignancy. *Abdom Radiol (NY)* (2020) 45:2989–96. doi: 10.1007/s00261-019-02209-7

28. Su Q, Wang Q, Zhang H, Yu D, Wang Y, Liu Z, et al. Computed tomography findings of small bowel gastrointestinal stromal tumors with different histologic risks of progression. *Abdom Radiol (NY)* (2018) 43:2651–8. doi: 10.1007/s00261-018-1511-6

29. Yan J, Zhao X, Han S, Wang T, Miao F. Evaluation of clinical plus imaging features and multidetector computed tomography texture analysis in preoperative risk grade prediction of small bowel gastrointestinal stromal tumors. *J Comput Assist Tomogr* (2018) 42:714–20. doi: 10.1097/RCT.0000000000000756

30. Zhang L, Kang L, Li G, Zhang X, Ren J, Shi Z, et al. Computed tomography-based radiomics model for discriminating the risk stratification of gastrointestinal stromal tumors. *Radiol Med* (2020) 125:465–73. doi: 10.1007/s11547-020-01138-6

31. Chen T, Ye LY, Feng XY, Qiu HB, Zhang P, Luo YX, et al. Performance of risk stratification systems for gastrointestinal stromal tumors: A multicenter study. *World J Gastroenterol* (2019) 25:1238–47. doi: 10.3748/wjg.v25.i10.1238

32. Zhao B, Zhang J, Mei D, Zhang J, Luo R, Xu H, et al. The assessment of different risk classification systems for gastrointestinal stromal tumors (gists): The analytic results from the seer database. *Scand J Gastroenterol* (2018) 53:1319–27. doi: 10.1080/00365521.2018.1515319

33. Chen Z, Xu L, Zhang C, Huang C, Wang M, Feng Z, et al. Ct radiomics model for discriminating the risk stratification of gastrointestinal stromal tumors: A multi-class classification and multi-center study. *Front Oncol* (2021) 11:654114. doi: 10.3389/fonc.2021.654114

34. Chu H, Pang P, He J, Zhang D, Zhang M, Qiu Y, et al. Value of radiomics model based on enhanced computed tomography in risk grade prediction of gastrointestinal stromal tumors. *Sci Rep* (2021) 11:12009. doi: 10.1038/s41598-021-91508-5

35. Liu C, Qiao M, Jiang F, Guo Y, Jin Z, Wang Y. Tn-usma net: Triple normalization-based gastrointestinal stromal tumors classification on multicenter eus images with ultrasound-specific pretraining and meta attention. *Med Phys* (2021) 48(11):7199–214. doi: 10.1002/mp.15172

36. Mao H, Zhang B, Zou M, Huang Y, Yang L, Wang C, et al. Mri-based radiomics models for predicting risk classification of gastrointestinal stromal tumors. *Front Oncol* (2021) 11:631927. doi: 10.3389/fonc.2021.631927

37. Song Y, Li J, Wang H, Liu B, Yuan C, Liu H, et al. Radiomics nomogram based on contrast-enhanced ct to predict the malignant potential of gastrointestinal stromal tumor: A two-center study. *Acad Radiol* (2021) 29(6):806–16. doi: 10.1016/j.acra.2021.05.005

38. Wang M, Feng Z, Zhou L, Zhang L, Hao X, Zhai J. Computed-tomography-based radiomics model for predicting the malignant potential of gastrointestinal stromal tumors preoperatively: A multi-classifier and multicenter study. *Front Oncol* (2021) 11:582847. doi: 10.3389/fonc.2021.582847

39. Wu X, Dong D, Zhang L, Fang M, Zhu Y, He B, et al. Exploring the predictive value of additional peritumoral regions based on deep learning and radiomics: A multicenter study. *Med Phys* (2021) 48:2374–85. doi: 10.1002/mp.14767

frontiers | Frontiers in Oncology

TYPE Systematic Review
PUBLISHED 06 September 2022
DOI 10.3389/fonc.2022.951557

OPEN ACCESS

EDITED BY
Pei Yang,
Central South University, China

REVIEWED BY
Hiroshi Kagamu,
Saitama Medical University International Medical Center, Japan
Angelo Castello,
IRCCS Ca 'Granda Foundation Maggiore Policlinico Hospital, Italy

*CORRESPONDENCE
Xiaosong Ben
benxs@163.com
Chunxia Jing
jcxphd@gmail.com

SPECIALTY SECTION
This article was submitted to Cancer Imaging and Image-directed Interventions, a section of the journal Frontiers in Oncology

RECEIVED 24 May 2022
ACCEPTED 08 August 2022
PUBLISHED 06 September 2022

CITATION
Zhu K, Su D, Wang J, Cheng Z, Chin Y, Chen L, Chan C, Zhang R, Gao T, Ben X and Jing C (2022) Predictive value of baseline metabolic tumor volume for non-small-cell lung cancer patients treated with immune checkpoint inhibitors: A meta-analysis. *Front. Oncol.* 12:951557. doi: 10.3389/fonc.2022.951557

Predictive value of baseline metabolic tumor volume for non-small-cell lung cancer patients treated with immune checkpoint inhibitors: A meta-analysis

Ke Zhu[1,2], Danqian Su[1,2], Jianing Wang[1,2], Zhouen Cheng[1,2], Yiqiao Chin[1,2], Luyin Chen[1,2], Chingtin Chan[1,2], Rongcai Zhang[1,2], Tianyu Gao[1], Xiaosong Ben[3]* and Chunxia Jing[1,4]*

[1]Department of Public Health and Preventive Medicine, School of Medicine, Jinan University, Guangzhou, China, [2]International School, Jinan University, Guangzhou, China, [3]Department of Thoracic Surgery, Guangdong Provincial People's Hospital, Guangdong Academy of Medical Sciences, Guangzhou, China, [4]Guangdong Key Laboratory of Environmental Pollution and Health, Jinan University, Guangzhou, China

Background: Immune checkpoint inhibitors (ICIs) have emerged as a promising treatment option for advanced non-small-cell lung cancer (NSCLC) patients, highlighting the need for biomarkers to identify responders and predict the outcome of ICIs. The purpose of this study was to evaluate the predictive value of baseline standardized uptake value (SUV), metabolic tumor volume (MTV) and total lesion glycolysis (TLG) derived from 18F-FDG-PET/CT in advanced NSCLC patients receiving ICIs.

Methods: PubMed and Web of Science databases were searched from January 1st, 2011 to July 18th, 2022, utilizing the search terms "non-small-cell lung cancer", "PET/CT", "standardized uptake value", "metabolic tumor volume", " total lesion glycolysis", and "immune checkpoint inhibitors". Studies that analyzed the association between PET/CT parameters and objective response, immune-related adverse events (irAEs) and prognosis of NSCLC patients treated with ICIs were included. We extracted the hazard ratio (HR) with a 95% confidence interval (CI) for progression-free survival (PFS) and overall survival (OS). We performed a meta-analysis of HR using Review Manager v.5.4.1.

Results: Sixteen studies were included for review and thirteen for meta-analysis covering 770 patients. As for objective response and irAEs after ICIs, more studies with consistent assessment methods are needed to determine their relationship with MTV. In the meta-analysis, low SUVmax corresponded to poor PFS with a pooled HR of 0.74 (95% CI, 0.57-0.96, P=0.02). And a high level of baseline MTV level was related to shorter PFS (HR=1.45, 95% CI, 1.11-1.89,

P<0.01) and OS (HR, 2.72; 95% CI, 1.97-3.73, P<0.01) especially when the cut-off value was set between 50-100 cm^3. SUVmean and TLG were not associated with the prognosis of NSCLC patients receiving ICIs.

Conclusions: High level of baseline MTV corresponded to shorter PFS and OS, especially when the cut-off value was set between 50-100 cm^3. MTV is a potential predictive value for the outcome of ICIs in NSCLC patients.

KEYWORDS

PET/CT (18)F-FDG, standardized uptake value, metabolic tumor volume, non-small-cell lung cancer, immune checkpoint inhibitor

1 Introduction

Lung cancer is the most common cause of cancer-related deaths worldwide in 2020, accounting for 1.80 million deaths (1). Non-small-cell lung cancer (NSCLC), compromising 80-85% of the lung cancer cases (2), has raised significant public health concerns. NSCLC is mainly composed of squamous cell carcinoma and adenocarcinoma (3), and the 5-year survival rate is 25% (4).Clinically, more than 60% of NSCLC patients had locally progressed or metastatic diseases (stage III or IV) at the time of diagnosis, when the tumor can not be effectively treated by surgical treatment alone (5), and the median overall survival varies between 7.0 and 12.2 months (6).

For the treatment of advanced NSCLC, chemotherapy remains the primary conventional therapy. But the response rate of NSCLC patients to chemotherapy was only about 20% (7), and the adverse events such as vomiting and diarrhea had a significant impact on patients' daily lives. The advent of immune checkpoint inhibitors (ICIs) targeting programmed cell death 1 (PD-1) or its ligand (PD-L1) has brought about a promising treatment option for the management of advanced NSCLC (8). A meta-analysis of 13 randomized controlled trials (RCTs) has proved that ICIs show better efficacy and result in fewer adverse events than chemotherapy as the treatment for advanced NSCLC (9). However, the benefits of ICIs remain limited to only 20% of advanced NSCLC patients (10). Thus it's necessary to identify potential biomarkers to identify NSCLC patients who would benefit from ICIs treatment.

^{18}F-fluorodeoxyglucose positron emission tomography/computed tomography (^{18}F-FDG PET/CT) monitors the uptake of ^{18}F-FDG of tumor cells. It is a convenient imaging modality for the staging, treatment guidance, and response predicting in NSCLC patients and is more practical and noninvasive than abdominal ultrasound and mediastinoscopy (11).

As metabolic parameters on PET/CT, SUV is associated with 18F-FDG uptake of the tumor; MTV combines the information of 18F-FDG uptake and tumor volume; TLG is the product of MTV and SUVmean, and is related to both tumor volume and tumor glycolytic activity. They reflect both tumor burden and aggressiveness (12). Takada et al. found that the accumulation of ^{18}F-FDG as SUVmax and SUVmean in tumor cells was significantly associated with PD-L1 expression in NSCLC patients (13). In addition, MTV and TLG have been potential prognostic factors in NSCLC patients treated with surgery (14) and chemotherapy (15). Thus SUV, MTV and TLG are expected to evaluate the efficacy of ICIs in advanced NSCLC patients. However, relevant studies showed inconclusive results. Monaco et al. have demonstrated that NSCLC patients with MTV and TLG values lower than the median values had improved outcomes of ICIs compared to those with higher values (16). No significant relationship was found between MTV, TLG, and ICIs response in studies conducted by Yamaguchi et al. (17) and Castello et al. (18).

Thus, we conducted this meta-analysis to assess the predictive value of SUV, MTV and TLG for advanced NSCLC patients receiving ICIs.

2 Material and methods

2.1 Data search and study selection

From January 1st, 2011, to July 18th, 2022, We searched comprehensively English language publications from PubMed and Web of Science using the terms "non-small-cell lung cancer", "PET/CT", "Standardized uptake value", "metabolic tumor volume", " total lesion glycolysis", and "immune checkpoint inhibitors". We extracted data from the full-text

articles that met the following inclusion criteria: studies limited to NSCLC; ICIs administered alone for the patients; ^{18}F-FDG PET/CT completed before ICIs initiation; studies reported objective response, immune-related adverse events (irAEs), survival data, including progression-free survival (PFS) or overall survival (OS); hazard ratio (HR) with 95% CI was provided for PFS or OS. Reviews, meeting abstracts, and editorial material were excluded. Two authors conducted the searches and screening independently. A consensus resolved any discrepancies.

2.2 Data extraction

Data were extracted from the publications independently by two reviewers (YC and CC), and the following information was recorded: first author's name, year of the paper published, country, types of ICIs, median follow-up, number of patients, median age of patients, median values of MTV and TLG, HR and p-value for PFS and OS. The data were collected and organized in a standardized data extraction table for analysis. We also formed a table including median values of MTV or numbers of patients in different objective response groups and the related p-value to demonstrate the relationship between MTV and objective response. When there was uncertainty in the inclusion of data, a third researcher assisted with confirming the data.

2.3 Quality assessment

We used ROBINS-I (Risk Of Bias In Non-randomised Studies - of Interventions) to assess the quality of included articles from seven bias domains, including confounding bias, selection bias, bias due to classification of interventions, bias from intended interventions, bias due to missing data, bias in outcomes measure, and bias due to selection reporting result. We classified each article as low, moderate, or high risk according to detailed guidance from ROBINS-I (19).

2.4 Statistics analysis

We performed all statistical analyses using Review Manager v.5.4.1 and pooled the hazard ratio (HR) and its 95% confidence index (CI) of PFS and OS using the inverse variance method. An HR greater than 1 indicated worse survival for patients with high SUV, MTV or TLG, while an HR less than 1 indicated a better survival for patients with a high SUV, MTV or TLG. Chi-square test and I^2 statistics were used to detect heterogeneity between studies. I^2 values of more than 50% were considered high heterogeneity. If high heterogeneity was found between primary studies, a random effect model would be used for meta-analysis. Otherwise, a fixed effect model would be applied. P values less than 0.05 were considered statistically significant.

3 Results

3.1 Literature search

Eight hundred and sixteen studies were retrieved from the systematic search of PubMed and Web of Science from January 1st, 2011, to July 18th, 2022. We excluded 134 duplicate studies and further screened the remaining 682 using titles and abstracts. 641 studies did not meet the inclusion criteria and thus were excluded. The full texts of the 41 potentially eligible studies were evaluated. Then 25 studies were excluded for the following reasons: not single ICIs as treatment (n=3), no available data (n = 19), and overlapped data (n=3). Ultimately, sixteen studies were included for review and thirteen studies assessing the predictive value of SUV, MTV and TLG in NSCLC patients receiving ICIs were included in this meta-analysis. Figure 1 shows the flowchart diagram.

3.2 Characteristics of included studies

The thirteen articles, including 770 patients, were analyzed in this meta-analysis. Characteristics of the included studies are summarized in Table 1. Four studies were conducted in France (20, 22, 26, 27), followed by three in Italy (16, 18, 21) and three in Japan (16, 18, 21). We also identified a single study in the United States (19), Israel (24) and Belgium (28). Two studies were of a prospective design (18, 20). SUV, MTV and TLG were measured in four studies (18, 20, 25, 28) and MTV alone was measured in five studies (17, 19, 21, 22, 24).

Regarding types of ICIs, nine studies (17, 20–25, 27, 28) reported using PD-1 inhibitors, while three used PD-1 and PD-L1 inhibitors (16, 18, 26). Patients were divided into high or low SUV/MTV/TLG groups in each study based on the cut-off values, and their PFS/OS were analyzed. And eleven of the thirteen studies used median MTV/TLG as cut-off values (16, 18–22, 24–28). The left two used log-rank test (23) and receiver operating characteristic (ROC) curve analysis (17) to determine cut-off values, respectively.

3.3 Quality assessment

We used the Cochrane collaboration tool to assess the risk of bias in included studies. The risks of the selected studies are shown in Figure 2. As shown, the overall risk of bias was relatively low, and the overall quality met the requirements of the meta-analysis.

10.3389/fonc.2022.951557

TABLE 1 Characteristics and results of included studies.

Studies	Year	Country	Study design	Types of ICIs	Median follow-up	No. of patients	Median Age	Median values as cut-offs	Outcome			
									HR (95% CI) for PFS	p value	HR (95% CI) for OS	p value
Andraos et al. (19)	2022	USA	R	–	17.0 months	124	67	MTV: 87.8 (cut-offs: 88.0)	1.36 (0.91-2.01)	0.131	2.23 (1.35-3.69)	0.002
Castello et al. (18)	2021	Italy	P	Nivolumab/ pembrolizumab/ atezolizumab	12.4 months	50	73	SUVmax: 13.6	0.9 (0.5-1.8)	0.75	0.9 (0.4-2.0)	0.75
								SUVmean: 5.9	0.9 (1.0-1.7)	0.75	0.8 (0.4-1.9)	0.71
								MTV: 63.7	2.5 (1.2-4.8)	0.01	2.3 (1.0-5.3)	0.04
								TLG: 330.1	1.8 (0.9-3.6)	0.08	1.5 (0.7-3.6)	0.27
Chardin et al. (20)	2020	France	P	Nivolumab/ pembrolizumab	12.3 months	79	64	SUVmax: 13.4	–	–	1.31 (0.63-2.75)	0.5
								SUVpeak: 9.7	–	–	1.15 (0.55-2.40)	0.7
								MTV:36.5	–	–	5.37 (2.17-13.3)	<0.0001
								TLG: 267.0	–	–	5.05 (2.05-12.5)	0.0001
Dall'Olio et al. (21)	2021	Italy	R	Pembrolizumab	20.3 months	34	66.6	MTV: 75.0	–	–	5.37 (1.72-16.77)	0.004
Eude et al. (22)	2022	France	R	Pembrolizumab	–	65	64.1	MTV: 188.3	–	–	1.314	0.012
Hashimoto et al. (23)	2020	Japan	R	Nivolumab/ pembrolizumab	–	85	–	MTV: 17.8 (cut-offs: 5.0)	1.28 (0.97-1.73)	0.07	1.59 (1.09-2.45)	0.001
								TLG: 75.4 (cut-offs: 20.0)	1.21 (0.92-1.63)	0.16	1.47 (1.03-2.21)	0.03
Icht et al. (24)	2020	Israel	R	Nivolumab/ pembrolizumab	–	58	65	MTV:12.95	1.1 (0.87-1.4)	0.4	1.2 (0.86-1.73)	0.26
Kitajima et al. (25)	2021	Japan	R	Nivolumab/ pembrolizumab	36.8 months	40	69.1	SUVmax: 8.57	1.04 (0.49-2.18)	0.92	1.56 (0.67-3.69)	0.3
								MTV: 15.5	2.15 (1.03-4.73)	0.042	2.15 (1.03-4.73)	0.042
								TLG: 87.7	1.15 (0.55-2.42)	0.7	1.35 (0.59-3.13)	0.47
Monaco et al. (16)	2021	Italy	R	Nivolumab/ pembrolizumab/ atezolizumab	–	92	70	SUVmean: 4.9	0.365 (0.150-0.890)	0.027	0.261 (0.084-0.808)	0.02
								MTV: 94.9	1.139 (0.989-1.311)	0.07	1.221 (1.063-1.402)	0.005
Seban et al. (26)	2019	France	R	Nivolumab/ pembrolizumab/ atezolizumab	11.6 months	80	61.9	SUVmax: 12.8	0.8 (0.5-1.3)	0.35	0.9 (0.5-1.5)	–
								MTV: 75.0	1.0 (0.9-1.1)	0.25	3.1 (1.7-5.7)	0.0001
Seban et al. (27)	2020	France	R	Pembrolizumab	13.4 months	63	65	SUVmax: 18	0.6 (0.3-1.1)	0.11	0.6 (0.2-1.6)	0.31
								SUVmean: 10.1	0.5 (0.3-1.1)	0.04	0.8 (0.3-1.9)	0.56
								MTV: 84.0	2.1 (1.1-4.3)	0.02	3.1 (1.1-8.3)	0.03
Vekens et al. (28)	2021	Belgium	R	Pembrolizumab	20 months	30	67	SUVmax: 15.7	0.62 (0.39-0.98)	0.04	0.54 (0.29-1.01)	0.06

(Continued)

TABLE 1 Continued

Studies	Year	Country	Study design	Types of ICIs	Median follow-up	No. of patients	Median Age	Median values as cut-offs	Outcome			
									HR (95% CI) for PFS	p value	HR (95% CI) for OS	p value
								SUVpeak: 10.2	1.43 (0.97-2.11)	0.07	1.71 (0.97-3.03)	0.06
								SUVmean: 6.06	1.76 (0.54-5.79)	0.35	1.51 (0.46-4.93)	0.5
								MTV: 123.9	1.01 (0.99-1.03)	0.25	1.01 (0.99-1.02)	0.29
								TLG: 802.6	0.99 (0.99-1.00)	0.29	0.99 (0.99-1.00)	0.42
Yamaguchi et al. (17)	2020	Japan	R	Pembrolizumab	346 days	48	69	MTV: 112.0 (cut-offs: 268.0)	1.49 (0.77-3.24)	0.32	1.57 (0.98-2.41)	0.04

3.4 Outcomes of included studies

3.4.1 PET/CT parameters and response assessment

Eight studies discussed whether PET/CT parameters including MTV, TLG, SUVmax, SUVmean, and SUVpeak, can predict the response of ICIs in different patients. All of them classified responses to ICIs as complete remission (CR), partial response (PR), stable disease (SD), and progression of disease (PD) based on the Response Evaluation Criteria In Solid Tumors (RECIST) 1.1. Four articles demonstrated that none of the PET/CT parameters significantly correlated with ICIs response (17, 23, 28, 29). However, the other four studies showed that NSCLC patients who achieved CR, PR, or SD after ICIs treatment had significantly lower median MTV values than those with PD (16, 26, 27, 30). The detailed data were shown in Table 2. In addition, Seban et al. found that SUVmean was

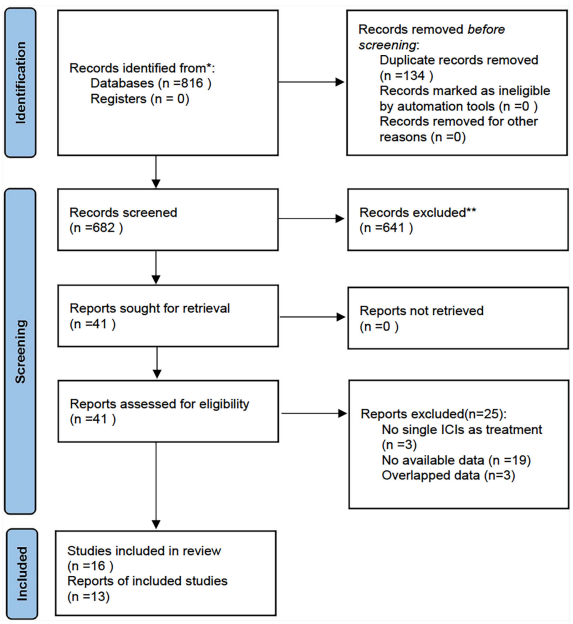


FIGURE 1
Flowchart diagram for the literature search.

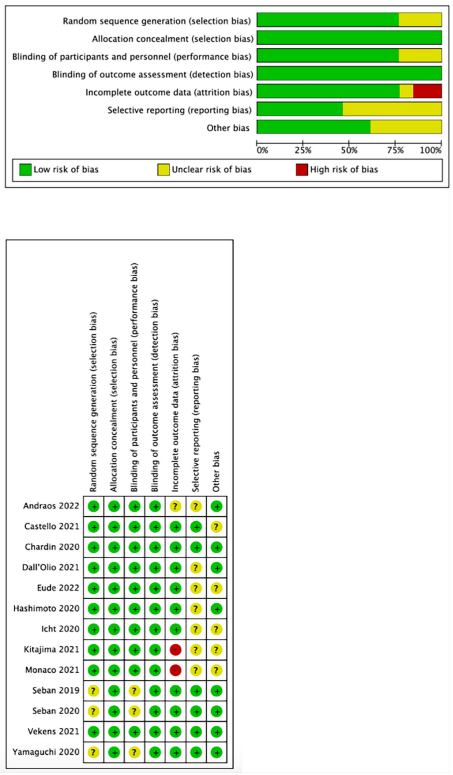


FIGURE 2
Results of quality assessment.

significantly higher in patients who achieved long-term benefit (LTB, defined as CR, PR or SD maintained ±12 months) compared to those without LTB (27), while Polveri et al. concluded that TLG was significantly associated with progressive vs non-progressive disease status (30).

3.4.2 PET/CT parameters and immune-related adverse events (irAEs)

Two studies discussed the relationship between PET/CT parameters and irAEs. In the analysis of Mu et al. (31), SUVmax and MTV were not correlated with irAEs, with the

TABLE 2 MTV values and objective response (RECIST 1.1).

Authors	Published year	CR+PR+SD group		PD group		P value
		value	Number of patients	value	Number of patients	
Median value of MTV						
Ferrari et al. (29)	2021	203.0	15	–	13	0.387
Monaco et al. (16)	2021	77	61	160.2	31	**0.039**
Polvari et al. (30)	2020	57.4	27	124.4	30	**0.028**
Seban et al. (26)	2019	55.4	32	83.4	48	**0.04**
Seban et al. (27)	2020	59.4	17	90.5	46	0.05
Vekens et al. (28)	2021	192.8	23	119.8	7	0.17
Number of patients in high/low MTV group						
Hashimoto et al. (23)	2020	High MTV: 36 Low MTV: 17		High MTV: 18 Low MTV: 9		>0.99
Yamaguchi et al. (17)	2020	High MTV: 3 Low MTV: 20		High MTV: 7 Low MTV: 15		0.16

CR, complete response; PR, partial response; SD, stable disease; PD, progressive disease; MTV, metabolic tumor volume; Bold means statistically significant.

odds ratio of 0.95 (95%CI, 0.87-1.05, P=0.34) and 0.99 (95%CI, 0.98-1.00, P=0.27), respectively. However, Hashimoto et al. (23) reported that the frequency of irAE was significantly higher in patients with low values of SUVmax, MTV, and TLG than in those with high values, inconsistent with the result of Mu et al.

3.4.3 PET/CT parameters and NSCLC survival

3.4.3.1 SUVmax and NSCLC survival

Six studies (18, 20, 25–28) analyzed the relationship between SUVmax and PFS/OS, as shown in Figure 3. The cut-off values of SUVmax ranged from 8.57 to 18 cm^3. Five studies analyzing PFS showed a pooled HR of 0.74 (95% CI, 0.57-0.96, P=0.02). However, SUVmax was not significantly associated with OS (HR, 0.89; 95% CI, 0.64-1.23, P=0.48). There was no significant heterogeneity between studies in both PFS (I^2 = 0%, P=0.72) and OS group (I^2 = 13%, P=0.33).

3.4.3.2 SUVmean and NSCLC survival

We performed SUVmean and survival analysis based on four studies (16, 18, 27, 28) with cut-off values between 4.9 and 10.1cm^3 (Figure 4). SUVmean was not associated with either PFS (HR, 0.67; 95% CI, 0.39-1.16, P=0.15) or OS (HR, 1.11; 95% CI, 0.65-1.19, P=0.69). The heterogeneity test didn't show significant heterogeneity in PFS (I^2 = 53%, P=0.1) and OS group (I^2 = 19%, P=0.29).

3.4.3.3 MTV and NSCLC survival

Thirteen studies (16–28) analyzed the relationship between MTV and PFS/OS, as shown in Figure 5. The cut-off values of MTV ranged from 5.0 to 268.0cm^3, so we performed a subgroup analysis based on the cut-off values, dividing them into three groups: MTV < 50cm^3, MTV between 50-100cm^3, and MTV >100cm^3.

In eleven studies analyzing PFS, a pooled HR of 1.21 (95% CI, 1.06-1.36, P<0.01) was shown. There was statistically significant heterogeneity between studies, with an I^2 of 79.4% (P<0.01). It is also demonstrated that patients with higher MTV would have shorter PFS (HR=1.45, 95% CI, 1.11-1.89, P<0.01) when the cut-off values was set at 50-100cm^3. There was no evidence of a significant association between MTV and PFS in the other two subgroups.

OS was analyzed in thirteen MTV studies. The pooled HR was 1.67 (95% CI, 1.36-2.06, P<0.01) with statistically significant heterogeneity between studies (I^2 = 84%, P<0.01). High MTV was significantly associated with poor OS, with an HR of 1.90 (95% CI, 1.15-3.15, P=0.01) and 2.35 (95% CI, 1.43-3.87, P<0.01) when the cut-off value was set below 50cm^3 and 50-100cm^3, respectively. The left subgroup showed no evidence of significant association.

3.4.3.4 TLG and NSCLC survival

TLG and survival analysis was performed based on five studies (18, 20, 23, 25, 28) with cut-off values between 20 and 802.6 (Figure 6). TLG was not associated with either PFS (HR, 1.10; 95% CI, 0.91-1.33, P=0.34) or OS (HR, 1.52; 95% CI, 0.98-2.34, P=0.06). The heterogeneity test showed high heterogeneity in OS (I^2 = 79%, P<0.01) and no significant results in PFS (I^2 = 41%, P=0.17).

4 Discussion

This study evaluated the predictive values of PET/CT parameters including SUVmax, SUVmean, MTV and TLG in

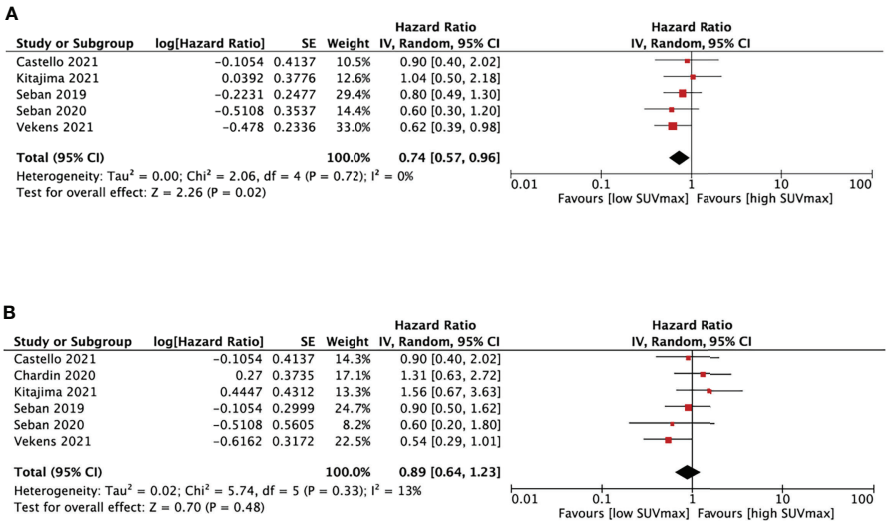


FIGURE 3
Forest plots of hazard ratios comparing progression free survival **(A)** or overall survival **(B)** of patients with high level versus low level max standardized uptake value treating with immune checkpoint inhibitors.

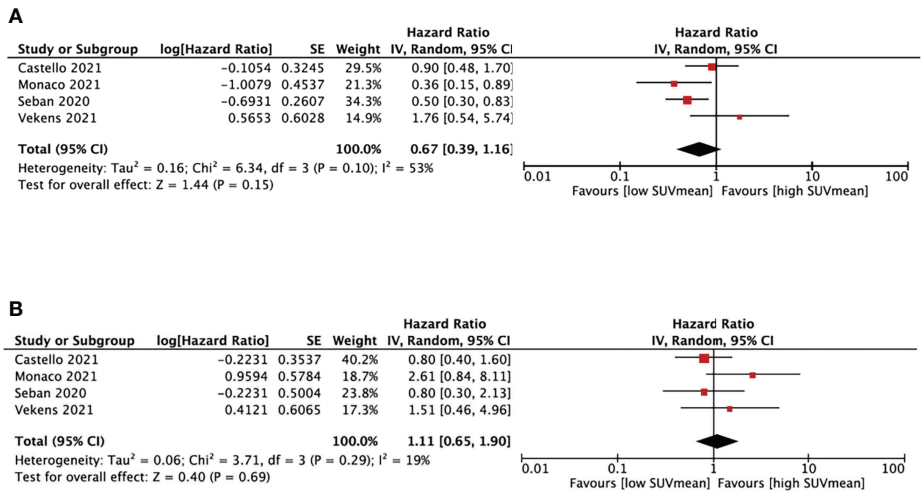


FIGURE 4
Forest plots of hazard ratios comparing progression free survival **(A)** or overall survival **(B)** of patients with high level versus low level mean standardized uptake value treating with immune checkpoint inhibitors.

NSCLC patients receiving ICIs. The cut-off values categorized patients into high or low-level parameter groups in the included studies.

Firstly, we analyzed the relationship between PET/CT parameters and the objective response of ICIs. Eight studies assessed the objective response based on RECIST 1.1. Four studies showed NSCLC patients who achieved CR, PR, or SD after ICIs treatment had significantly lower median MTV values than those with PD (16, 26, 27, 30), while four demonstrated no significant correlation (17, 23, 28, 29). More studies with consistent response assessments are needed to determine whether MTV is associated with the objective response of ICIs. SUVmean (27) and TLG (30) were also said to have a significant relationship with disease status in a single study, respectively.

ICIs may alter the physiological homeostasis of the immune response, thus leading to the development of irAEs. Two studies discussed the relationship between PET/CT parameters and irAEs (23, 31). However, no consistent results could be yet concluded.

We also discussed whether PET/CT parameters could predict NSCLC survival by PFS and OS after ICIs. We found that lower SUVmax corresponded to shorter PFS. Lopci et al. found a positive association between SUVmax and CD8-tumor infiltrating lymphocytes and PD-1 expression (32). SUVmax were also independent predictors of PD-L1 positivity by Takada et al. (13). However, the predictive role of baseline SUVmax is still under discussion since only one of the five included studies about SUVmax showed significant results.

In terms of MTV, we found that a high baseline MTV level was significantly associated with shorter PFS and OS than a low MTV level for patients treated with ICIs.

MTV refers to the metabolically active volume of tumors segmented using FDG PET (33), reflecting tumor burden and the metabolic status. Regarding tumor burden, Kim et al. concluded that larger-size tumors are more immunosuppressive than smaller-size tumors, which negatively affects the immune responses induced by immunotherapy (34).The experiments in mice also verified that PD-L1 blocker is less effective in mice bearing larger lung squamous cell tumors (35). On the cell level, Wang et al. analyzed one hundred twenty-two NSCLC tumor specimens by immunohistochemistry and found a significantly positive correlation between MTV and CD163-TAM, Foxp3-Tregs (36). CD163-TAMs were tumor-promoting M2 macrophages (37), and Foxp3-Tregs were a kind of immune regulatory cells (35), both of which are immunosuppressive cells. Therefore, we hypothesize that patients with a higher MTV would have a worse prognosis when treated with ICIs than those with a lower MTV, since a higher MTV would result in a more immunosuppressive tumor microenvironment.

In respective of tumor glycolysis, a higher MTV indicates a larger metabolically active volume of glucose uptake by the tumor (38). Different from normal cells, tumor cells can uptake a large amount of glucose at a rapid rate, consuming most of nutrients from the surrounding environment, and metabolizing glucose into lactic acid (Warburg effect) (39). Tumors with higher MTV would have worse response to ICIs by affecting T cells responsiveness by the following possible ways.

Firstly, in tumor microenvironment (TME), tumor cells and T cells compete for glucose as their primary energy source (40). Tumors with higher MTV would consume more glucose and

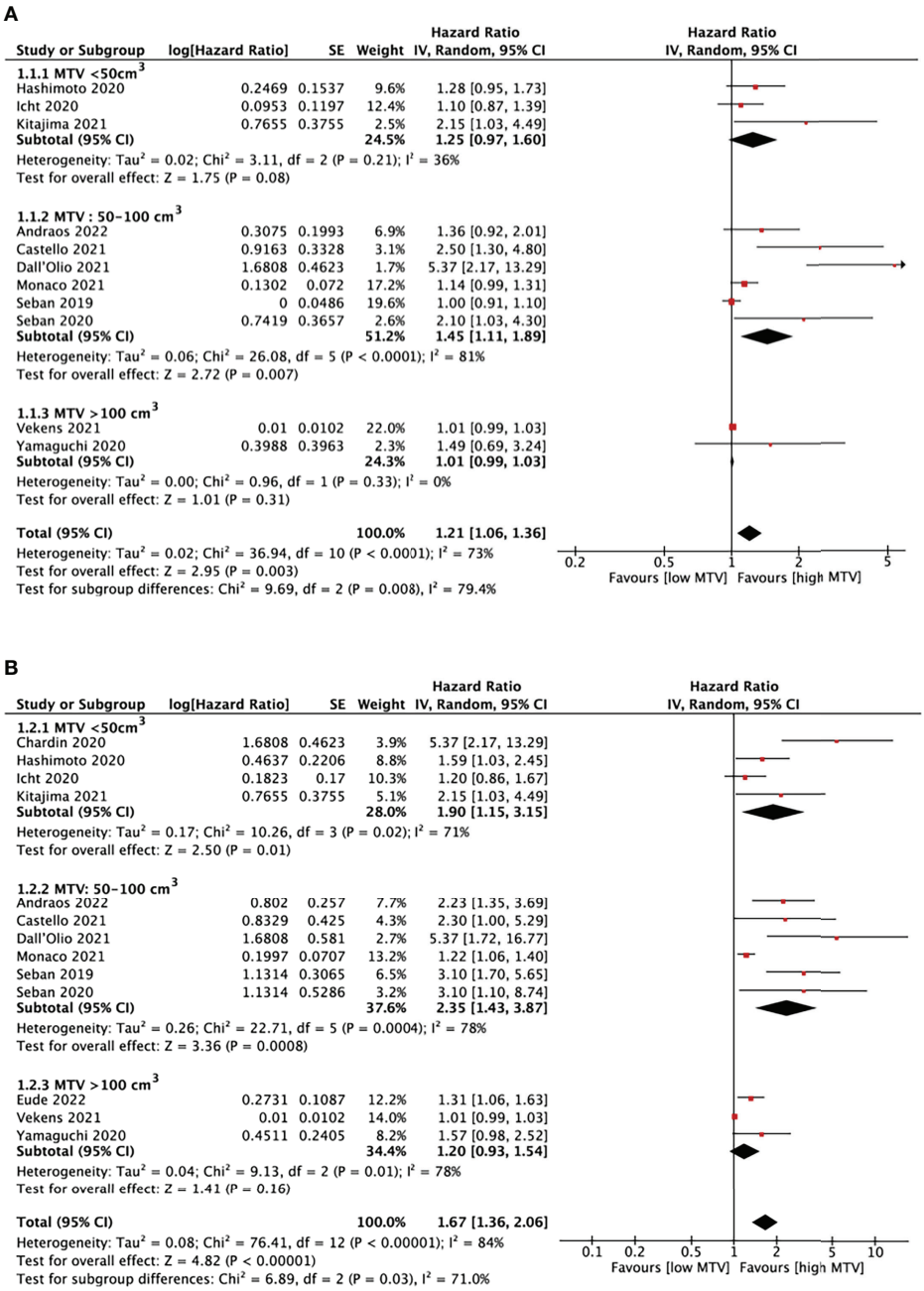


FIGURE 5
Forest plots of hazard ratios comparing progression free survival **(A)** or overall survival **(B)** of patients with high level versus low level metabolic tumor volume treating with immune checkpoint inhibitors.

lead to glucose deprivation of T cells, decreasing T cells' ability to produce effector cytokines like interferon gamma (IFN-γ), which has impact on the function of tumor infiltrating CD8+ T cells (41).

In contrast, Harley et al. found that melanoma tumors with less glycolysis would provide more glucose for infiltrating T cells and are associated with increased antigen presentation and better response to anti-PD-1 ICIs (42). Secondly, the accumulation of lactate in the TME will inhibit CD8+T cell proliferation and activation by preventing lactic acid export from CD8+T cell (43) or inhibiting CD25 expression, a T cells activation marker (44). More studies are still needed to explain why MTV could predict the outcome of immunotherapy in patients with NSCLC.

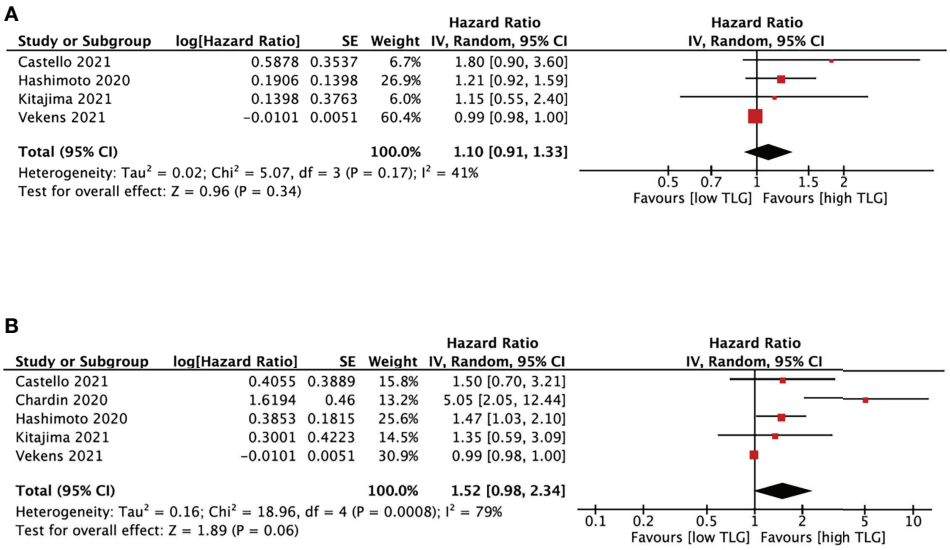


FIGURE 6
Forest plots of hazard ratios comparing progression free survival **(A)** or overall survival **(B)** of patients with high level versus low level total lesion glycolysis treating with immune checkpoint inhibitors.

Since the cut-off values of MTV ranged from 5.0 to 268.0 cm^3 in different studies, we also did a subgroup analysis to determine the impact of cut-off values on outcome assessment. Our result showed that a high baseline MTV level was significantly associated with shorter PFS when setting the cut-off values of MTV below 50cm^3 and shorter OS in the groups with cut-off values lower than 50 cm^3 or between 50cm^3 and 100 cm^3.

The baseline MTV level didn't show any predictive value when the cut-off values were more than 100 cm^3. Thus further studies with a larger sample size should focus on cut-off values of MTV between 50 and 100 cm^3 and try to figure out a more precise cut-off value to improve the efficacy of MTV prediction on response assessment to ICIs in NSCLC patients.

Although SUVmean and TLG were potential prognostic markers of NSCLC (45), our pooled results showed that they were not significantly associated with PFS and OS in NSCLC patients receiving ICIs.

In addition to NSCLC, PET/CT parameters also played potential predictive roles in other cancers treated with ICIs, supporting our findings. Zhang et al. reported that total SUVmax ≥12.5 was associated with worse PFS in head and neck squamous cell carcinoma (46). And according to a systematic review and meta-analysis of metastatic melenoma (47), MTV and TLG were promising predictors of OS for metastatic melanoma patients who received ICIs.

^{18}F-FDG PET/CT is a convenient and noninvasive imaging modality, and SUVmax and MTV are easily obtained. Since our study proved that SUVmax and MTV have the potential predictive value for ICIs in NSCLC patients, further studies are needed to define the role of SUVmax and MTV in providing individualized treatments for advanced NSCLC patients. Early identification of NSCLC patients for ICIs can improve the efficacy of ICIs in responders and avoid the side effects and high costs of ICIs in non-responders, allowing them to initiate other treatments timely.

Our study also has several limitations. Firstly, majority of the included studies are retrospective studies. Potential selection bias may exist and impact the reliability of this meta-analysis. Secondly, the methods of PET/CT were not consistent between different studies. A golden method should be defined to ensure the homogeneity of studies. Thirdly, cut-off values of SUV, MTV and TLG ranged widely and were determined by different methods, including median values, log-rank test and ROC curve analysis. Thus the pooled results may show some risk of bias.

In conclusion, our study showed that high baseline MTV levels correspond to shorter PFS and OS compared with low baseline MTV levels especially when the cut-off value was set between 50-100 cm^3. MTV is a potential predictor of ICI outcomes in NSCLC patients.

Data availability statement

The original contributions presented in the study are included in the article/supplementary material. Further inquiries can be directed to the corresponding author.

Author contributions

CJ, TG and KZ conceived and designed research. Data collection was performed by DS, JW and ZC. Data extraction was performed by YC and CC and verified by KZ. Statistical analysis was performed by LC, RZ and KZ. KZ, DS, JW, ZC, YC, LC, CC, XB and CJ participated in drafting article. All authors gave final approval to the version submitted.

Funding

This work was funded by the Guangdong Pharmaceutical Association Scientific Research Fund (No: 2021ZX03).

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. World Health Organization. *Cancer* (2020). Available at: https://www.who.int/news-room/fact-sheets/detail/cancer (Accessed December 13, 2021).

2. Salehi-Rad R, Li R, Paul MK, Dubinett SM, Liu B. The biology of lung cancer: Development of more effective methods for prevention, diagnosis, and treatment. *Clin Chest* (2020) 41(1):25–38. doi: 10.1016/j.ccm.2019.10.003

3. Wang BY, Huang JY, Chen HC, Lin CH, Lin SH, Hung WH, et al. The comparison between adenocarcinoma and squamous cell carcinoma in lung cancer patients. *J Cancer Res Clin Oncol* (2020) 146(1):43–52. doi: 10.1007/s00432-019-03079-8

4. *Cancer. lung cancer - non-small cell: Statistics* (2021). Available at: https://www.cancer.net/cancer-types/lung-cancer-non-small-cell/statistics (Accessed December 13, 2021).

5. Osmani L, Askin F, Gabrielson E, Li QK. Current WHO guidelines and the critical role of immunohistochemical markers in the subclassification of non-small cell lung carcinoma (NSCLC): Moving from targeted therapy to immunotherapy. *Semin Cancer Biol* (2018) 52:103–9. doi: 10.1016/j.semcancer.2017.11.019

6. Schad F, Thronicke A, Steele ML, Merkle A, Matthes B, Grah C, et al. Overall survival of stage IV non-small cell lung cancer patients treated with viscum album l. @ in addition to chemotherapy, a real-world observational multicenter analysis. *PloS One* (2018) 13(8):e0203058. doi: 10.1371/journal.pone.0203058

7. Scheff RJ, Schneider BJ. Non-small-cell lung cancer: Treatment of late stage disease: Chemotherapeutics and new frontiers. *Semin Intervent Radiol* (2013) 30 (2):191–8. doi: 10.1055/s-0033-1342961

8. Kim JH, Kim HS, Kim BJ. Prognostic value of KRAS mutation in advanced non-small-cell lung cancer treated with immune checkpoint inhibitors: A meta-analysis and review. *Oncotarget* (2017) 8(29):48248–52. doi: 10.18632/oncotarget.17594

9. Yu DP, Cheng X, Liu ZD, Xu SF. Comparative beneficiary effects of immunotherapy against chemotherapy in patients with advanced NSCLC: Meta-analysis and systematic review. *Oncol Lett* (2017) 14(2):1568–80. doi: 10.3892/ol.2017.6274

10. Song P, Yang D, Wang H, Cui X, Si X, Zhang X, et al. Relationship between the efficacy of immunotherapy and characteristics of specific tumor mutation genes in non-small cell lung cancer patients. *Thorac Cancer* (2020) 11(6):1647–54. doi: 10.1111/1759-7714.13447

11. Hochhegger B, Alves GR, Irion KL, Fritscher CC, Fritscher LG, Concatto NH, et al. PET/CT imaging in lung cancer: Indications and findings. *J Bras Pneumol* (2015) 41(3):264–74. doi: 10.1590/S1806-37132015000004479

12. Lim R, Eaton A, Lee NY, Setton J, Ohri N, Rao S, et al. 18F-FDG PET/CT metabolic tumor volume and total lesion glycolysis predict outcome in oropharyngeal squamous cell carcinoma. *J Nucl Med* (2012) 53(10):1506–13. doi: 10.2967/jnumed.111.101402

13. Takada K, Toyokawa G, Okamoto T, Baba S, Kozuma Y, Matsubara T, et al. Metabolic characteristics of programmed cell death-ligand 1-expressing lung cancer on 18 f-fluorodeoxyglucose positron emission tomography/computed tomography. *Cancer Med* (2017) 6(11):2552–61. doi: 10.1002/cam4.1215

14. Liu J, Dong M, Sun X, Li W, Xing L, Yu J. Prognostic value of 18F-FDG PET/CT in surgical non-small cell lung cancer: A meta-analysis. *PloS One* (2016) 11(1):e0146195. doi: 10.1371/journal.pone.0146195

15. Li X, Wang D, Yu L. Prognostic and predictive values of metabolic parameters of 18F-FDG PET/CT in patients with non-small cell lung cancer treated with chemotherapy. *Mol Imaging* (2019) 18:1536012119846025. doi: 10.1177/1536012119846025

16. Monaco L, Gemelli M, Gotuzzo I, Bauckneht M, Crivellaro C, Genova C, et al. Metabolic parameters as biomarkers of response to immunotherapy and prognosis in non-small cell lung cancer (NSCLC): A real world experience. *Cancers (Basel)* (2021) 13(7):1634. doi: 10.3390/cancers13071634

17. Yamaguchi O, Kaira K, Hashimoto K, Mouri A, Shiono A, Miura Y, et al. Tumor metabolic volume by 18F-FDG-PET as a prognostic predictor of first-line pembrolizumab for NSCLC patients with PD-L1 ≥ 50. *Sci Rep* (2020) 10(1):14990. doi: 10.1038/s41598-020-71735-y

18. Castello A, Rossi S, Toschi L, Lopci E. Impact of antibiotic therapy and metabolic parameters in non-small cell lung cancer patients receiving checkpoint inhibitors. *J Clin Med* (2021) 10(6):1251. doi: 10.3390/jcm10061251

19. Andraos TY, Halmos B, Cheng H, Huntzinger C, Shirvani SM, Ohri N. Disease Burden on PET Predicts Outcomes for Advanced NSCLC Patients Treated with First-Line Immunotherapy. *Clin Lung Cancer* (2022) 23(4):291–9. doi: 10.1016/j.cllc.2022.02.003

20. Chardin D, Paquet M, Schiappa R, Darcourt J, Bailleux C, Poudenx M, et al. Baseline metabolic tumor volume as a strong predictive and prognostic biomarker in patients with non-small cell lung cancer treated with PD1 inhibitors: A prospective study. *J Immunother Cancer* (2020) 8(2):e000645. doi: 10.1136/jitc-2020-000645

21. Dall'Olio FG, Calabrò D, Conci N, Argalia G, Marchese PV, Fabbri F, et al. Baseline total metabolic tumour volume on 2-deoxy-2-[18F]fluoro-d-glucose positron emission tomography-computed tomography as a promising biomarker in patients with advanced non-small cell lung cancer treated with first-line pembrolizumab. *Eur J Cancer* (2021) 150:99–107. doi: 10.1016/j.ejca.2021.03.020

22. Eude F, Guisier F, Salaün M, Thiberville L, Pressat-Laffouilhere T, Vera P, et al. Prognostic value of total tumour volume, adding necrosis to metabolic tumour volume, in advanced or metastatic non-small cell lung cancer treated with first-line pembrolizumab. *Ann Nucl Med* (2022) 36(3):224–34. doi: 10.1007/s12149-021-01694-5

23. Hashimoto K, Kaira K, Yamaguchi O, Mouri A, Shiono A, Miura Y, et al. Potential of FDG-PET as prognostic significance after anti-PD-1 antibody against patients with previously treated non-small cell lung cancer. *J Clin Med* (2020) 9 (3):725. doi: 10.3390/jcm9030725

24. Icht O, Domachevsky L, Groshar D, Dudnik E, Rotem O, Allen AM, et al. Lower tumor volume is associated with increased benefit from immune checkpoint inhibitors in patients with advanced non-small-cell lung cancer. *Asia Pac J Clin Oncol* (2021) 17(2):e125–31. doi: 10.1111/ajco.13360

25. Kitajima K, Kawanaka Y, Komoto H, Minami T, Yokoi T, Kuribayashi K, et al. The utility of 68F-FDG PET/CT for evaluation of tumor response to immune

checkpoint inhibitor therapy and prognosis prediction in patients with non-small-cell lung cancer. *Hell J Nucl Med* (2021) 24(3):186–98. doi: 10.1967/s002449912402

26. Seban RD, Mezquita L, Berenbaum A, Dercle L, Botticella A, Le Pechoux C, et al. Baseline metabolic tumor burden on FDG PET/CT scans predicts outcome in advanced NSCLC patients treated with immune checkpoint inhibitors. *Eur J Nucl Med Mol Imaging* (2020) 47(5):1147–57. doi: 10.1007/s00259-019-04615-x

27. Seban RD, Assie JB, Giroux-Leprieur E, Massiani MA, Soussan M, Bonardel G, et al. FDG-PET biomarkers associated with long-term benefit from first-line immunotherapy in patients with advanced non-small cell lung cancer. *Ann Nucl Med* (2020) 34(12):968–74. doi: 10.1007/s12149-020-01539-7

28. Vekens K, Everaert H, Neyns B, Ilsen B, Decoster L. The value of 18F-FDG PET/CT in predicting the response to PD-1 blocking immunotherapy in advanced NSCLC patients with high-level PD-L1 expression. *Clin Lung Cancer* (2021) 22 (5):432–40. doi: 10.1016/j.cllc.2021.03.001

29. Ferrari C, Santo G, Merenda N, Branca A, Mammucci P, Pizzutilo P, et al. Immune checkpoint inhibitors in advanced NSCLC: [18F]FDG PET/CT as a troubleshooter in treatment response. *Diagnostics (Basel)* (2021) 11(9):1681. doi: 10.3390/diagnostics11091681

30. Polverari G, Ceci F, Bertaglia V, Reale ML, Rampado O, Gallio E, et al. 18F-FDG pet parameters and radiomics features analysis in advanced nsclc treated with immunotherapy as predictors of therapy response and survival. *Cancers (Basel)* (2020) 12(5):1163. doi: 10.3390/cancers12051163

31. Mu W, Tunali I, Qi J, Schabath MB, Gillies RJ. Radiomics of 18F fluorodeoxyglucose PET/CT images predicts severe immune-related adverse events in patients with NSCLC. *Radiol Artif Intell* (2020) 2(1):e190063. doi: 10.1148/ryai.2019190063

32. Lopci E, Toschi L, Grizzi F, Rahal D, Olivari L, Castino GF, et al. Correlation of metabolic information on FDG-PET with tissue expression of immune markers in patients with non-small cell lung cancer (NSCLC) who are candidates for upfront surgery. *Eur J Nucl Med Mol Imaging* (2016) 43(11):1954–61. doi: 10.1007/s00259-016-3425-2

33. Im HJ, Bradshaw T, Solaiyappan M, Cho SY. Current methods to define metabolic tumor volume in positron emission tomography: Which one is better? *Nucl Med Mol Imaging* (2018) 52(1):5–15. doi: 10.1007/s13139-017-0493-6

34. Kim SI, Cassella CR, Byrne KT. Tumor burden and immunotherapy: Impact on immune infiltration and therapeutic outcomes. *Front Immunol* (2021) 11:629722. doi: 10.3389/fimmu.2020.629722

35. Guisier F, Cousse S, Jeanvoine M, Thiberville L, Salaun M. A rationale for surgical debulking to improve anti-PD1 therapy outcome in non small cell lung cancer. *Sci Rep* (2019) 9(1):16902. doi: 10.1038/s41598-019-52913-z

36. Wang Y, Zhao N, Wu Z, Pan N, Shen X, Liu T, et al. New insight on the correlation of metabolic status on 18F-FDG PET/CT with immune marker expression in patients with non-small cell lung cancer. *Eur J Nucl Med Mol Imaging* (2020) 47(5):1127–36. doi: 10.1007/s00259-019-04500-7

37. Sumitomo R, Hirai T, Fujita M, Murakami H, Otake Y, Huang CL. M2 tumor-associated macrophages promote tumor progression in non-small-cell lung cancer. *Exp Ther Med* (2019) 18(6):4490–8. doi: 10.3892/etm.2019.8068

38. Moon SH, Hyun SH, Choi JY. Prognostic significance of volume-based PET parameters in cancer patients. *Korean J Radiol* (2013) 14(1):1–12. doi: 10.3348/kjr.2013.14.1.1

39. Vaupel P, Schmidberger H, Mayer A. The warburg effect: Essential part of metabolic reprogramming and central contributor to cancer progression. *Int J Radiat Biol* (2019) 95(7):912–9. doi: 10.1080/09553002.2019.1589653

40. Chang CH, Qiu J, O'Sullivan D, Buck MD, Noguchi T, Curtis JD, et al. Metabolic competition in the tumor microenvironment is a driver of cancer progression. *Cell* (2015) 162(6):1229–41. doi: 10.1016/j.cell.2015.08.016

41. Chang CH, Curtis JD, Maggi LBJr, Faubert B, Villarino AV, O'Sullivan D, et al. Posttranscriptional control of T cell effector function by aerobic glycolysis. *Cell* (2013) 153(6):1239–51. doi: 10.1016/j.cell.2013.05.016

42. Harel M, Ortenberg R, Varanasi SK, Mangalhara KC, Mardamshina M, Markovits E, et al. Proteomics of melanoma response to immunotherapy reveals mitochondrial dependence. *Cell* (2019) 179(1):236–250.e18. doi: 10.1016/j.cell.2019.08.012

43. Fischer K, Hoffmann P, Voelkl S, Meidenbauer N, Ammer J, Edinger M, et al. Inhibitory effect of tumor cell-derived lactic acid on human T cells. *Blood* (2007) 109(9):3812–9. doi: 10.1182/blood-2006-07-035972

44. Bosticardo M, Ariotti S, Losana G, Bernabei P, Forni G, Novelli F. Biased activation of human T lymphocytes due to low extracellular pH is antagonized by B7/CD28 costimulation. *Eur J Immunol* (2001) 31(9):2829–38. doi: 10.1002/1521-4141(200109)31:9<2829::aid-immu2829>3.0.co;2-u

45. Yıldırım F, Yurdakul AS, Özkaya S, Akdemir ÜÖ, Öztürk C. Total lesion glycolysis by 18F-FDG PET/CT is independent prognostic factor in patients with advanced non-small cell lung cancer. *Clin Respir J* (2017) 11(5):602–11. doi: 10.1111/crj.12391

46. Zhang S, Zhang R, Gong W, Wang C, Zeng C, Zhai Y, et al. Positron emission tomography-computed tomography parameters predict efficacy of immunotherapy in head and neck squamous cell carcinomas. *Front Oncol* (2021) 11:728040. doi: 10.3389/fonc.2021.728040

47. Ayati N, Sadeghi R, Kiamanesh Z, Lee ST, Zakavi SR, Scott AM. The value of 18F-FDG PET/CT for predicting or monitoring immunotherapy response in patients with metastatic melanoma: A systematic review and meta-analysis. *Eur J Nucl Med Mol Imaging* (2021) 48(2):428–48. doi: 10.1007/s00259-020-04967-9

frontiers | Frontiers in Oncology

TYPE Original Research
PUBLISHED 07 September 2022
DOI 10.3389/fonc.2022.922185

OPEN ACCESS

EDITED BY
Nanna Maria Sijtsema,
University Medical Center Groningen,
Netherlands

REVIEWED BY
Anandh Kilpattu Ramaniharan,
Anna University, India
Lan He,
Guangdong Academy of Medical Sciences, China

*CORRESPONDENCE
Qingshi Zeng
zengqingshi@sina.com

SPECIALTY SECTION
This article was submitted to Cancer Imaging and Image-directed Interventions, a section of the journal Frontiers in Oncology

RECEIVED 17 April 2022
ACCEPTED 08 August 2022
PUBLISHED 07 September 2022

CITATION
Fang C, Zhang J, Li J, Shang H, Li K, Jiao T, Yin D, Li F, Cui Y and Zeng Q (2022) Clinical-radiomics nomogram for identifying HER2 status in patients with breast cancer: A multicenter study.
Front. Oncol. 12:922185.
doi: 10.3389/fonc.2022.922185

Clinical-radiomics nomogram for identifying HER2 status in patients with breast cancer: A multicenter study

Caiyun Fang[1,2], Juntao Zhang[3], Jizhen Li[4], Hui Shang[1,2], Kejian Li[1,2], Tianyu Jiao[1,2], Di Yin[1], Fuyan Li[5], Yi Cui[6] and Qingshi Zeng[1*]

[1]Department of Radiology, Shandong Provincial Qianfoshan Hospital, The First Hospital Affiliated Hospital of Shandong First Medical University, Jinan, China, [2]Postgraduate Department, Shandong First Medical University and Shandong Academy of Medical Sciences, Jinan, China, [3]GE Healthcare Precision Health Institution, Shanghai, China, [4]Department of Radiology, Shandong Mental Health Center, Jinan, China, [5]Department of Radiology, Shandong Provincial Hospital Affiliated to Shandong First Medical University, Jinan, China, [6]Department of Radiology, Qilu Hospital of Shandong University, Jinan, China

Purpose: To develop and validate a clinical-radiomics nomogram based on radiomics features and clinical risk factors for identification of human epidermal growth factor receptor 2 (HER2) status in patients with breast cancer (BC).

Methods: Two hundred and thirty-five female patients with BC were enrolled from July 2018 to February 2022 and divided into a training group (from center I, 115 patients), internal validation group (from center I, 49 patients), and external validation group (from centers II and III, 71 patients). The preoperative MRI of all patients was obtained, and radiomics features were extracted by a free open-source software called 3D Slicer. The Least Absolute Shrinkage and Selection Operator regression model was used to identify the most useful features. The radiomics score (Rad-score) was calculated by using the radiomics signature-based formula. A clinical-radiomics nomogram combining clinical factors and Rad-score was developed through multivariate logistic regression analysis. The performance of the nomogram was evaluated using receiver operating characteristic (ROC) curve and decision curve analysis (DCA).

Results: A total of 2,553 radiomics features were extracted, and 21 radiomics features were selected as the most useful radiomics features. Multivariate logistic regression analysis indicated that Rad-score, progesterone receptor (PR), and Ki-67 were independent parameters to distinguish HER2 status. The clinical-radiomics nomogram, which comprised Rad-score, PR, and Ki-67, showed a favorable classification capability, with AUC of 0.87 [95% confidence internal (CI), 0.80 to 0.93] in the training group, 0.81 (95% CI, 0.69 to 0.94) in the internal validation group, and 0.84 (95% CI, 0.75 to 0.93) in

the external validation group. DCA illustrated that the nomogram was useful in clinical practice.

Conclusions: The nomogram combined with Rad-score, PR, and Ki-67 can identify the HER2 status of BC.

KEYWORDS

breast cancer, human epidermal growth factor receptor 2, radiomics, nomogram, magnetic resonance imaging

Introduction

Breast cancer (BC) is the most common malignancy worldwide and the main cause of cancer-related death in women (1, 2). The prognosis of BC has improved since the appearance of targeted therapies, especially for patients with a human epidermal growth factor receptor 2 (HER2)-positive subtype (3). HER2-positive BC is characterized by high invasiveness, high degree of malignancy, recurrence, and metastasis, and poor prognosis (4, 5). Therefore, accurate assessment of the HER2 status is very important for the prognosis prediction and treatment decision-making for BC patients.

At present, the HER2 status is mainly detected by immunohistochemistry (IHC) or fluorescence *in situ* hybridization (FLSH), both of which are invasive methods involving tissue samples (6). However, the consistency of the HER2 status between core needle biopsy and subsequent resection biopsy of the same BC is 81%–96% (7, 8). Therefore, the development of a non-invasive and reliable method is essential for the assessment of the HER2 status in BC patients. Magnetic resonance imaging (MRI), an essential tool in breast imaging, is considered to be one of the most sensitive imaging methods for detecting BC and monitoring neoadjuvant chemotherapy (9, 10). T2WI can be used to detect bleeding, edema, and cyst in breast lesions (11). Diffusion-weighted imaging (DWI), a common method to evaluate the micro-architecture of the tumors based on the measurement of the Brownian motion of water molecules, improves the accuracy of breast tumor diagnosis (12). Dynamic contrast-enhanced MRI (DCE-MRI), another common method to evaluate BC, can provide information on blood perfusion and microvessel distribution (13). The so-called imaging features, such as blurred boundary, irregular shape, and lobulated or burr mass, are useful for the diagnosis of BC, whereas the features have limited performance in predicting the HER2 status (14). Radiomics is a new machine learning method that aims to extract a large number of quantitative features from medical images using data characterization algorithms (15). These quantitative features have been applied to identify benign and malignant breast lesions and predict neoadjuvant chemotherapy response and lymph node metastasis (16–18).

In the present study, to identify the HER2 status of BC patients, we hypothesized that the combination of radiomics signatures and clinical factors could evaluate the HER2 status in BC patients. To verify the feasibility of our hypothesis, radiomics features were selected using the Least Absolute Shrinkage and Selection Operator (LASSO) logistic model based on the radiomics features extracted from fat suppression T2WI (FS-T2WI), DWI, and DCE-MRI. A clinical-radiomics nomogram model integrating radiomics signatures and clinical risk factors was constructed by multivariate logistic regression analysis and verified by the multicenter dataset.

Materials and methods

Patients

The retrospective study was approved by the local institutional review board, and the requirement for informed consent was waived. From July 2018 to February 2022, the MR images and pathological data of BC patients were collected from three clinical centers (center I, the First Affiliated Hospital of Shandong First Medical University; center II, Provincial Hospital Affiliated to Shandong First Medical University; and center III, Qilu Hospital of Shandong University). The inclusion criteria were as follows: (1) postoperative pathology confirmed that BC was an invasive ductal carcinoma of no special type; (2) breast MRI was performed within 2 weeks before surgery; (3) no preoperative radiotherapy or neoadjuvant chemotherapy. The exclusion criteria were as follows: (1) incomplete clinical data or insufficient MRI quality; (2) the HER2 status was not tested by IHC or FLSH after surgery, or the IHC intensity score of patient specimens was 2 +, and FLSH was not further tested. The flowchart is shown in Figure 1.

In addition, the following clinical information was obtained through the patient's electronic medical record

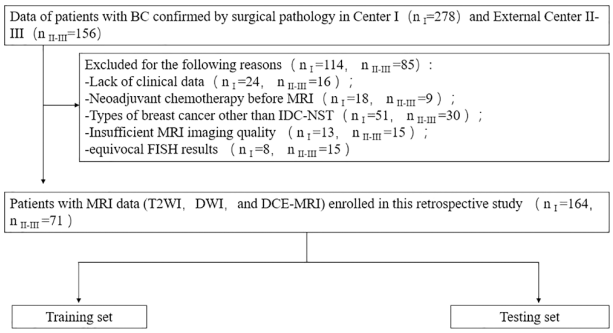


FIGURE 1
Patient recruitment routes in center I and external centers II-III. n_{I}, number of patients in center I; n_{II-III}, total number of patients in external centers II-III.

system: age, tumor diameter, tumor grade, estrogen receptor (ER), progesterone receptor (PR), Ki-67 proliferation index, HER2, and pathological axillary lymph node (ALN) metastasis status.

Postoperative pathological assessment

The status of HER2 was detected by IHC or FLSH after operation. According to the guidelines of the American Society of Clinical Oncology/College of American Pathologists (ASCO/CAP) (6), if the IHC result was 0 or 1+, HER2 was defined as negative; if the result was 3+, it is positive; for tumors with an IHC result of 2+, further FLSH detection was required. If gene amplification occurred, it was defined as positive. For the ER/PR test, the nuclear staining of ≥1% of tumors was defined as ER/PR positive. The critical threshold of Ki-67 to 14% was set, and tumors ≥14% were defined as high expression.

MRI acquisition and image segmentation

Breast MRI examinations were performed using a 3.0-T MRI scanner, equipped with a special breast phased-array surface coil. Patients were placed in the prone position, and the bilateral mammary glands naturally hung in the coil to fully extend the mammary glands. FS-T2WI, DWI, and DCE-MRI were sequentially obtained, and the detailed parameters of MRI acquisition are summarized in Supplementary Table 1.

A free open-source software called 3D Slicer (www.slicer.org) was used to perform image segmentation. On FS-T2WI, DWI, and DCE-MRI (the peak enhancement phase of multiphase-enhanced MRI selected according to the time intensity curve), the region of interest (ROI) of each tumor was manually outlined layer by layer along the tumor contour by excluding the areas of necrosis and calcification. Figure 2 shows an example of manual ROI drawing.

Radiomics feature extraction and selection

Due to the difference between MRI scanning parameters and devices, we preprocessed the images before the extraction of radiomics features. We resampled the voxels of all images to 1 mm × 1 mm × 1 mm using three-line interpolation and standardized its intensity range to 0 to 255. The 3D Slicer software was also used for feature exaction according to guidelines defined by the image biomarker standardization initiative (19). Four groups of features were extracted from the FS-T2WI, DWI, DCE-MRI, and their combination (FS-T2WI+DWI+DCE-MRI).

To evaluate the intra- and interobserver agreement of feature exaction, the MR images of 30 patients were randomly selected. Two experienced radiologists (reader 1 and reader 2) blinded to clinical information completed the process manually and independently with the same criteria. Reader 1 repeated the process after 3 weeks to assess intraobserver reproducibility. The reliability of measurements was assessed by intra- and interclass correlation coefficients (ICCs). ICC values above 0.75 were considered to have good consistency, and the remaining MRI feature exaction was completed by reader 1.

Two feature selection methods, minimum-redundancy maximum-relevance (mRMR) and LASSO, were used to obtain the most significant characteristics for evaluation of the HER2 status. At first, mRMR was carried out to narrow the range of redundant and irrelevant features; 30 features were retained. Then, the retained features were filtered with LASSO to obtain the best features, and 10-fold cross validations were utilized to determine the optimal values of λ. The radiomics score (Rad-score) of each patient was calculated by selecting the

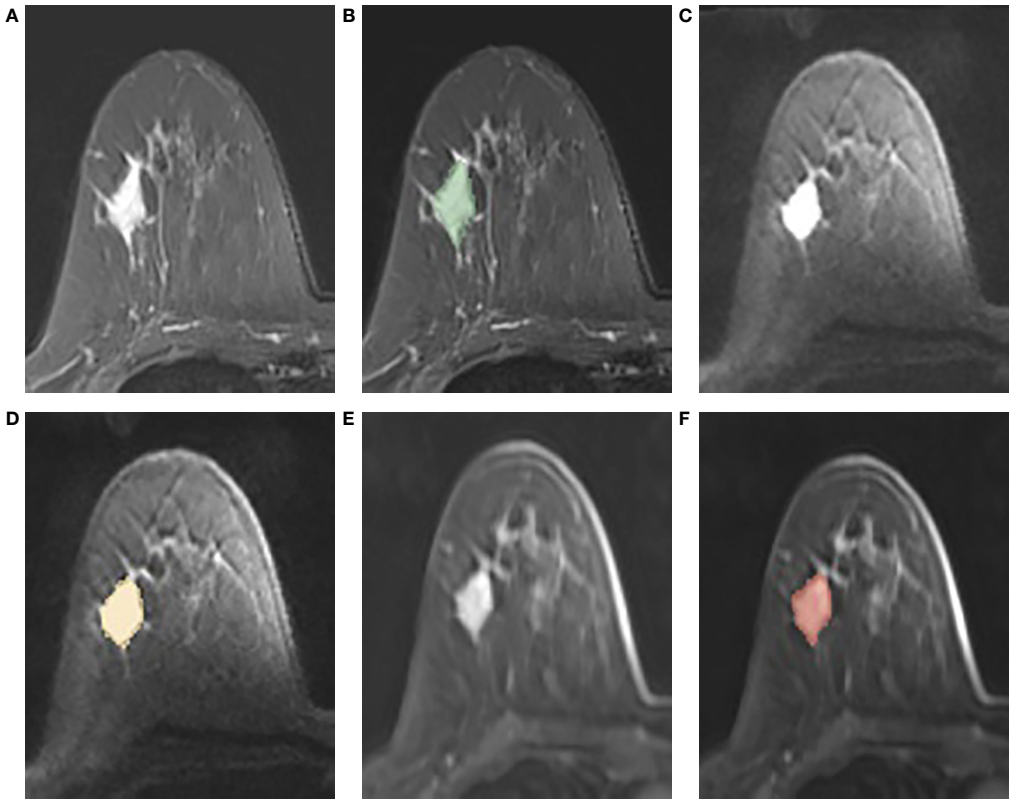


FIGURE 2
(A–F): An example of manual segmentation in breast cancer. **(A, B)**: tumor area (green in fat suppression T2WI image); **(C, D)**: tumor area (orange in DWI image, b = 1,000 s/mm^2); **(E, F)**: tumor area (red in DCE-MR image).

linear combination of features and the product of their respective coefficients

Model construction and validation

The seven clinical parameters (age, tumor diameter, tumor grade, ER, PR, Ki-67, and ALN metastasis status) were first analyzed by univariate logistic regression to screen out the clinical features of $P < 0.05$. To obtain the clinical risk factors identifying the HER2 status and building the clinical model, the significant variables in univariate analysis were input for stepwise multivariate logistic regression analysis. Moreover, we used multivariate logistic regression analysis to develop a clinical-radiomics model based on Rad-score and clinical risk factors, which is displayed by a nomogram.

Therefore, a total of three models were constructed to identify the HER2 status of BC: clinical model, radiomics model, and clinical-radiomics model. The area under the receiver operating characteristic (ROC) curve (AUC) was used to evaluate the discrimination performance of the three models in the training and validation cohorts. Finally, to explore the clinical utility of nomogram, decision curve analysis (DCA) was carried out based on three models to determine the utility of nomogram in a series of threshold probabilities.

Statistical analysis

R programming language (version 4.1.0, www.programmingr.com) was applied for statistical analysis and data processing. The differences of continuous variables (age, tumor diameter) between the HER2-negative group and HER2-positive group were compared by the independent sample *t*-test or Mann–Whitney U-test and described as mean ± standard deviation (SD). The differences of categorical variables (ER, PR, Ki-67, ALN metastasis status, and tumor grade) between the two groups were compared using chi-square test or Fisher's exact test and expressed as absolute numbers (n) and proportions (%). Univariate and multivariate logistic regression analyses were used to evaluate the relationship between HER2 overexpression status and clinical risk factors. All statistical tests were two-sided, and *P*-values of <0.05 were regarded as significant.

Results

Clinical characteristics

A total of 235 breast cancer patients was consecutively enrolled (met the inclusion criteria, but not the exclusion criteria). All the patients were women with the mean age of 50.76 ± 10.82 years (range: 26-83 years). The patients were divided into three independent groups: a training group (from center I, 115 patients), internal validation group (from center I, 49 patients), and external validation group (from centers II and III, 71 patients).

The clinical characteristics of the three groups were compared as shown in Table 1. The HER2-positive proportions in the training, internal validation, and external validation sets were 27.1%, 26.5%, and 32.4%, respectively.

Intraobserver and interobserver agreement for radiomics features extraction

The intraobserver ICC was 0.761 to 0.990, and the interobserver ICC ranged from 0.759 to 0.989 for evaluation of the radiomics features extraction. The results showed good consistency of feature extraction within and between observers.

Feature selection and development of the radiomics model

A total of 851 quantitative radiomics features were extracted from each sequence, which could be summarized into the following four groups: 14 volume and shape features (2D and 3D), 18 first-order features, 75 texture features, and 744 ([18 + 75] * 8) wavelet transform features.

By mRMR and LASSO, 6, 10, and 3 optimal radiomics features were selected from FS-T2WI, DWI, and DCE-MRI, respectively. Then, combining these three sequences, two radiomics features (one from DWI and one from DCE-MRI) were selected and executed from 2,553 (851×3) features to construct a radiomics model (Figure 3). The Rad-score of each patient was calculated using the formula presented in Supplementary Materials 2.

Development of the clinical and clinical-radiomics models

Based on the univariate and stepwise multivariate logistic regression analyses, two clinical risk factors (PR and Ki-67) were obtained for identification of the HER2 status and were used for construction of the clinical model. In addition, logistic

TABLE 1 Patient characteristics in the training and validation cohorts (mean ± standard deviation).

Clinicopathological features	Training group (N = 115)		*P*	Internal validation group (N = 49)		*P*	External validation group (N = 71)		*P*
	HER2- (n = 84)	HER2+ (n = 31)		HER2- (n = 36)	HER2+ (n = 13)		HER2- (n = 48)	HER2+ (n = 23)	
Age (years, mean± SD)	49.5 ± 10.5	52.5 ± 8.9	0.163	52.6 ± 11.4	50.3 ± 7.5	0.494	49.8 ± 12.4	52.4 ± 11.3	0.396
Diameter (cm, mean± SD)	2.1 ± 0.9	2.2 ± 0.7	0.658	1.8 ± 0.7	2.8 ± 1.1	0.000	1.9 ± 0.8	2.4 ± 0.8	0.034
ER			0.003			1.000			0.034
Positive	72 (85.7%)	18 (58.1%)		34 (94.4%)	12 (92.3%)		37 (77.1%)	12 (52.2%)	
Negative	12 (14.3%)	13 (41.9%)		2 (5.6%)	1 (7.7%)		11 (22.9%)	11 (47.8%)	
PR			0.001			0.352			0.000
Positive	68 (81.0%)	15 (48.4%)		31 (86.1%)	9 (69.2%)		38 (79.2%)	8 (34.8%)	
Negative	16 (19.0%)	16 (51.6%)		5 (13.9%)	4 (30.8%)		10 (20.8%)	15 (65.2%)	
Ki-67			0.001			0.040			0.038
≥14%	48 (57.1%)	28 (90.3%)		20 (55.6%)	12 (92.3%)		40 (83.3%)	23 (100%)	
<14%	36 (42.9%)	3 (9.7%)		16 (44.4%)	1 (7.7%)		8 (16.7%)	0	
Pathological ALN metastasis			0.751			0.176			0.399
Positive	28 (33.3%)	12 (38.7%)		10 (27.8%)	7 (53.8%)		22 (45.8%)	13 (56.5%)	
Negative	56 (66.7%)	19 (61.3%)		26 (72.2%)	6 (46.2%)		26 (54.2%)	10 (43.5%)	
Histological grade			0.210			0.289			0.017
I	17 (20.2%)	2 (6.5%)		9 (25.0%)	0		2 (4.2%)	0	
II	51 (60.7%)	22 (71.0%)		22 (61.1%)	8 (61.5%)		38 (79.2%)	12 (52.2%)	
III	16 (19.0%)	7 (22.6%)		5 (13.9%)	5 (38.5%)		8 (16.7%)	11 (47.8%)	
Rad-score(median)	-1.3[-0.9, -0.1]	-0.7[-0.9, -0.1]	<1e-04	-1.1[-1.7, -0.8]	-0.6[-0.9, -0.3]	0.002	-1.1[-1.7, -0.8]	-0.9[-1.2, -0.8]	0.010

ER, estrogen receptor; PR, progesterone receptor; HER2, human epidermal growth factor receptor-2; ALN, axillary lymph node.

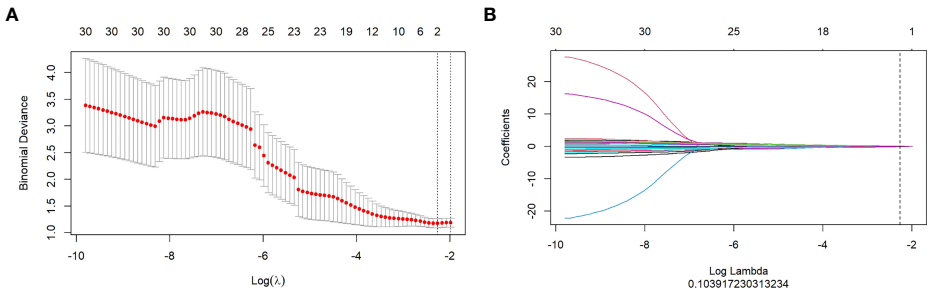


FIGURE 3
(A, B): Texture feature selection using the Least Absolute Shrinkage and Selection Operator (LASSO) regression. **(A)**: Optimal tuning parameters (λ) in the LASSO model binomial deviation diagram. **(B)**: LASSO coefficient profile of the features.

TABLE 2 Univariate and multivariate analyses of risk factors for HER2.

Variable	Univariate logistic analysis		Multivariate logistic analysis	
	OR (95% CI)	*P*	OR (95% CI)	*P*
ER	0.23 [0.09, 0.59]	0.002	NA	NA
PR	0.22 [0.09, 0.53]	0.000	0.37 [0.13, 1.07]	0.067
Ki-67	7.00 [1.97, 24.84]	0.002	4.12 [0.98, 17.37]	0.053
Rad-score	11.85 [4.25, 33.02]	<1e-04	9.88 [3.43, 28.43]	<1e-04

OR, odds ratio; CI, confidence interval; NA, not available; ER, estrogen receptor; PR, progesterone receptor; HER2, human epidermal growth factor receptor-2.

regression analysis showed that the Rad-score was an independent variable to identify HER2 status (Table 2). Therefore, a clinical-radiomics model was constructed by combining Rad-score and clinical risk factors.

Comparison of models and establishment of clinical-radiomics nomogram

To compare the performance of the clinical-radiomics model, the radiomics model, and the clinical model in identifying the HER2 status, we plotted the ROC curves of the three models (Figure 4). In the training cohort, the clinical-radiomics model showed the highest discrimination between HER2-negative and positive cases, with an AUC of 0.87 (95% CI, 0.80 to 0.93). The AUC value of the clinical-radiomics model was significantly higher than that of the radiomics model (AUC = 0.84, 95% CI, 0.76 to 0.92) and clinical model (AUC = 0.73, 95% CI, 0.64 to 0.82). In the internal validation and external validation cohorts, the AUC of the clinical-radiomics model was 0.81 (95% CI, 0.69 to 0.94) and 0.84 (95% CI, 0.75 to 0.93), respectively, which was superior to the single radiomics model and clinical model. The clinical-radiomics model showed the best ability to identify the HER2 status. Therefore, a clinical-radiomics nomogram was developed based on the clinical-radiomics model (Figure 5).

Clinical application

Figure 6 shows the DCA curves of the clinical-radiomics model, the radiomics model, and the clinical model. According to the DCA, the clinical-radiomics nomogram showed good clinical practicability in all threshold probabilities and obtained the greatest benefit. It indicated that the nomogram was a reliable clinical tool and could be used to identify the HER2 status of BC.

Discussion

BC is one of the most common death causes of cancer among women in the world. However, the way of BC treatment has changed drastically since HER2 is a target of the monoclonal antibody trastuzumab as well as of other anti-HER2 compounds. In this study, to identify the HER2 status in BC patients, we developed and validated a clinical-radiomics nomogram based on radiomics features and clinical risk factors. It successfully

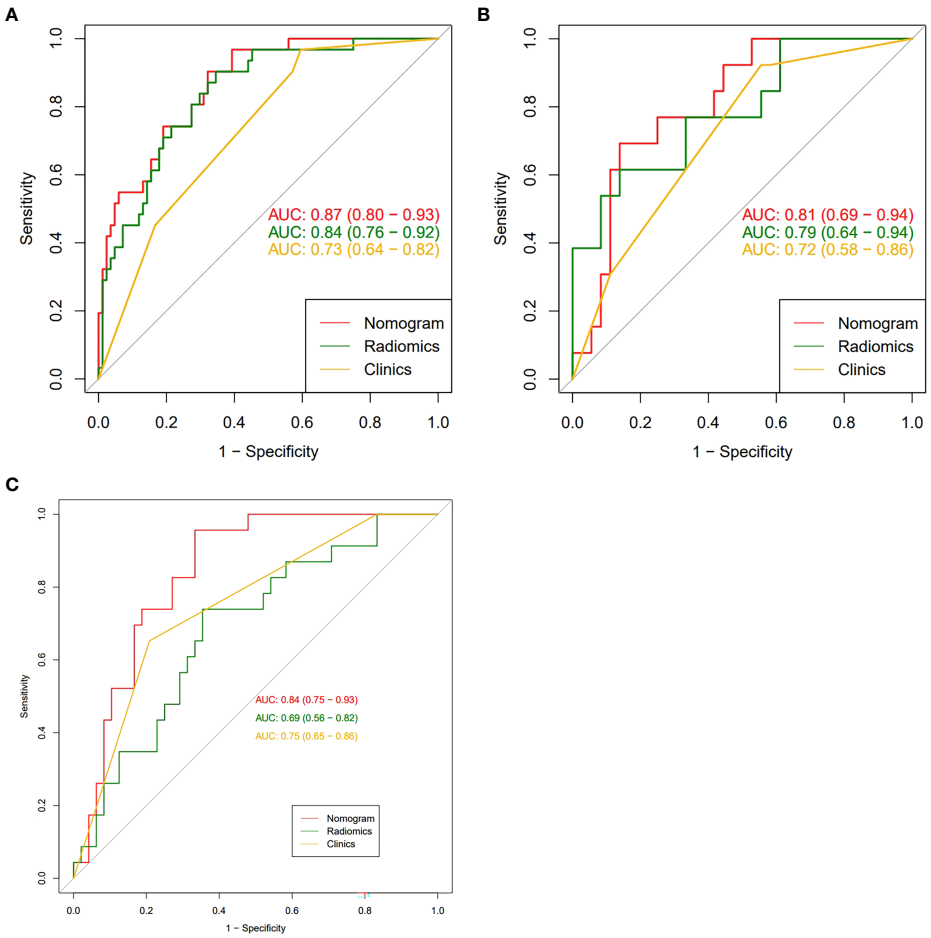


FIGURE 4
(**A–C**): The receiver operating characteristic curves of nomogram, radiomic signatures, and clinical risk factors for identifying the HER2 status of breast cancer were presented in the training group **(A)**, the internal validation group **(B)** and the external validation group **(C)**, respectively. The nomogram obtained the highest area under the curve (AUC).

stratified BC patients according to HER2 status and performed well in the training, internal, and external validation groups.

In the present study, PR and Ki-67 were identified as clinical risk factors for distinguishing the HER2 status by multivariate logistic regression analysis. PR promotes cell growth through nuclear pathways and non-nuclear pathways. There is a negative correlation between HER2 overexpression and PR expression, which is due to the loss of the PR protein caused by HER2 overexpression through the PI3K/Akt signaling pathway (20). A previous study has shown that the expression level of PR in BC with overexpression or high amplification level of HER2 is lower than that of low-level tumor (21). Ki-67 is a nuclear protein, which is usually used to detect and quantify tumor-proliferating cells. Its increased expression is related to cell growth (22). The Ki-67 index is positively correlated with HER2 status, which indicates that HER2 overexpression may upregulate the expression of Ki-67 (23). This was consistent with our results. Based on these clinical risk factors, we further obtained the clinical model to identify the HER2 status through multivariable logistic regression analysis. The AUC in the training, internal, and external validation groups were 0.73, 0.72, and 0.75 respectively, indicating that the discrimination efficiency of the model is good.

Radiomics, a research hotspot in the field of medical imaging analysis recently, is gaining importance in the evaluation of cancer by improving tumor diagnostic, prognostic, and predictive accuracy. The advantage of radiomics is the application of a large number of automatic data feature extraction algorithms to transform image data into quantitative features. In the present study, a radiomics

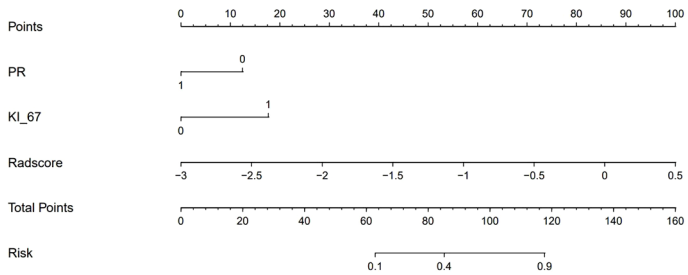


FIGURE 5
A clinical-radiomics nomogram. The nomogram was composed of Rad-score, PR, and Ki-67. PR: 0 = negative, 1 = positive; Ki-67: 0 = low expression, 1 = high expression.

model for identification of the HER2 status of BC patients was constructed on the basis of the extracted features (one from DWI and one from DCE-MRI). The AUC of the constructed radiomics model was 0.79 (internal validation group), which was similar to the previous studies (24, 25). Zhou et al. (24) reported a development of radiomic features based on mammography, including mediolateral oblique and cranial caudal views, to evaluate the BC HER2 status. The best combination of the two views was achieved, and the AUC of the test set was 0.787. In another study, the features extracted from T2WI in combination with DCE-MRI showed that the ability of predicting the HER2 status of BC patients was better than that of single-parameter MRI, and the AUC of the validation set was 0.81 (25).

Accurate identification of the HER2 status plays an essential role in the evaluation of treatment options for BC patients. The use of HER2 expression as a predictive biomarker of target drug response to trastuzumab is becoming a standard recommendation for the treatment of invasive breast cancer (26). To accurately identify the HER2 status of BC patients, we further established the clinical-radiomics nomogram based on radiomics features and clinical risk factors. The performance of the nomogram in identifying HER2 status was further improved, with AUC of 0.87 in the training group and 0.81 in the internal validation group. In the present study, we used the external validation set to verify the clinical-radiomics nomogram. The results showed that it had good prediction efficiency (AUC = 0.84), and its identification ability was significantly superior than that of single radiomics features and clinical features. Our multicenter data provided additional radiomics evidence for predicting the HER2 status of BC. It can be used as a non-invasive identification tool for the HER2 status. Doctors can add the scores of each prediction index to get the total score according to the individual differences of patients, so as to make a more accurate prediction, help clinical decision-

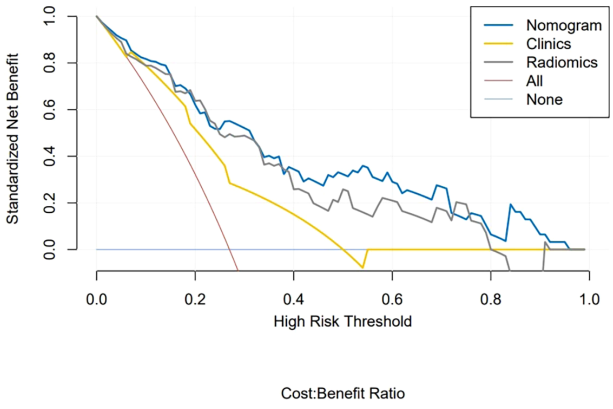


FIGURE 6
Decision curve analysis of clinical application evaluation of the nomogram. The vertical axis displays standardized net benefit. The two horizontal axes show the corresponding relationship between risk threshold and cost-benefit ratio. Compared with the radiomics signature (gray line) and clinical characteristics (yellow line), the nomogram (blue line) achieved the highest net benefit.

making more intuitively, and make personalized treatment plans. In addition, this study uses DCA to evaluate the clinical application of the nomogram. The DCA results show that the net benefit of the nomogram is higher than that of the radiomics model and clinical model, which increases the reliability of the model.

However, this study still has some limitations. Firstly, we only included the invasive ductal carcinoma of no special type in this study, because this pathological type accounts for 80% of all BC. This choice can avoid confounding factors associated with pathological types. Secondly, in DCE-MR images, we only selected the most obvious enhancing phase according to the time intensity curve and did not analyze the pre-contrasts and other enhanced images. Finally, this study is retrospective, and the sample size is relatively small, so some bias is inevitable. In future studies, large sample size prospective randomized studies are needed to verify the results of this study.

Conclusions

In conclusion, combined with radiomics features and clinical risk factors, a clinical-radiomics nomogram was constructed to evaluate the HER2 status of BC patients. It can be used for identifying the HER2 status in BC patients, helping clinical decision-making, and providing supplementary information for precise medical treatment.

Data availability statement

The original contributions presented in the study are included in the article/Supplementary Material. Further inquiries can be directed to the corresponding author.

Ethics statement

The studies involving human participants were reviewed and approved by The Ethical Review Committee of The First Affiliated of Shandong First Medical University, Provincial Hospital Affiliated to Shandong First Medical University and Qilu Hospital of Shandong university. Written informed consent for participation was not required for this study in accordance with the national legislation and the institutional requirements.

Author contributions

CF wrote the original draft preparation and design. KL, TJ, DY, FL, and YC collected the data. CF, JZ, JL, and HS analyzed the data and built the prediction models. QZ revised the manuscript. All authors contributed to the article and approved the submitted version.

Conflict of interest

Author JZ was employed by GE Healthcare, Shanghai, China.

The remaining 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

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fonc.2022.922185/full#supplementary-material

References

1. Sung H, Ferlay J, Siegel RL, Laversanne M, Soerjomataram I, Jemal A, et al. Global cancer statistics 2020: Globocan estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA: Cancer J Clin* (2021) 71(3):209–49. doi: 10.3322/caac.21660

2. Siegel RL, Miller KD, Fuchs HE, Jemal A. Cancer statistics, 2022. *CA: Cancer J Clin* (2022) 72(1):7–33. doi: 10.3322/caac.21708

3. Starczynski J, Atkey N, Connelly Y, O'Grady T, Campbell FM, di Palma S, et al. Her2 gene amplification in breast cancer: A rogues' gallery of challenging diagnostic cases: Ukneqas interpretation guidelines and research recommendations. *Am J Clin Pathol* (2012) 137(4):595–605. doi: 10.1309/ajcpatbz2jfn1qqc

4. Slamon DJ, Clark GM, Wong SG, Levin WJ, Ullrich A, McGuire WL. Human breast cancer: Correlation of relapse and survival with amplification of the her-2/Neu oncogene. *Sci (New York NY)* (1987) 235(4785):177–82. doi: 10.1126/science.3798106

5. Slamon DJ, Godolphin W, Jones LA, Holt JA, Wong SG, Keith DE, et al. Studies of the her-2/Neu proto-oncogene in human breast and ovarian cancer. *Sci (New York NY)* (1989) 244(4905):707–12. doi: 10.1126/science.2470152

6. Wolff AC, Hammond MEH, Allison KH, Harvey BE, Mangu PB, Bartlett JMS, et al. Human epidermal growth factor receptor 2 testing in breast cancer: American society of clinical Oncology/College of American pathologists clinical

practice guideline focused update. *J Clin Oncol* (2018) 36(20):2105–22. doi: 10.1200/jco.2018.77.8738

7. Tamaki K, Sasano H, Ishida T, Miyashita M, Takeda M, Amari M, et al. Comparison of core needle biopsy (Cnb) and surgical specimens for accurate preoperative evaluation of er, pgr and Her2 status of breast cancer patients. *Cancer Sci* (2010) 101(9):2074–9. doi: 10.1111/j.1349-7006.2010.01630.x

8. D'Alfonso T, Liu YF, Monni S, Rosen PP, Shin SJ. Accurately assessing her-2/Neu status in needle core biopsies of breast cancer patients in the era of neoadjuvant therapy: Emerging questions and considerations addressed. *Am J Surg Pathol* (2010) 34(4):575–81. doi: 10.1097/PAS.0b013e3181d65639

9. Houssami N, Ciatto S, Macaskill P, Lord SJ, Warren RM, Dixon JM, et al. Accuracy and surgical impact of magnetic resonance imaging in breast cancer staging: Systematic review and meta-analysis in detection of multifocal and multicentric cancer. *J Clin Oncol* (2008) 26(19):3248–58. doi: 10.1200/jco.2007.15.2108

10. Alaref A, Hassan A, Sharma Kandel R, Mishra R, Gautam J, Jahan N. Magnetic resonance imaging features in different types of invasive breast cancer: A systematic review of the literature. *Cureus* (2021) 13(3):e13854. doi: 10.7759/cureus.13854

11. Harada TL, Uematsu T, Nakashima K, Kawabata T, Nishimura S, Takahashi K, et al. Evaluation of breast edema findings at T2-weighted breast mri is useful for diagnosing occult inflammatory breast cancer and can predict prognosis after neoadjuvant chemotherapy. *Radiology* (2021) 299(1):53–62. doi: 10.1148/radiol.2021202604

12. Deike-Hofmann K, Kuder T, König F, Paech D, Dreher C, Delorme S, et al. Diffusion-weighted breast imaging. *Der Radiologe* (2018) 58(Suppl 1):14–9. doi: 10.1007/s00117-018-0423-3

13. Berg WA, Zhang Z, Lehrer D, Jong RA, Pisano ED, Barr RG, et al. Detection of breast cancer with addition of annual screening ultrasound or a single screening mri to mammography in women with elevated breast cancer risk. *Jama* (2012) 307(13):1394–404. doi: 10.1001/jama.2012.388

14. Montemurro F, Martincich L, Sarotto I, Bertotto I, Ponzone R, Cellini L, et al. Relationship between dce-mri morphological and functional features and histopathological characteristics of breast cancer. *Eur Radiol* (2007) 17(6):1490–7. doi: 10.1007/s00330-006-0505-x

15. Gillies RJ, Kinahan PE, Hricak H. Radiomics: Images are more than pictures, they are data. *Radiology* (2016) 278(2):563–77. doi: 10.1148/radiol.2015151169

16. Zhang Q, Peng Y, Liu W, Bai J, Zheng J, Yang X, et al. Radiomics based on multimodal mri for the differential diagnosis of benign and malignant breast lesions. *J Magn Reson Imaging* (2020) 52(2):596–607. doi: 10.1002/jmri.27098

17. Zhuang X, Chen C, Liu Z, Zhang L, Zhou X, Cheng M, et al. Multiparametric mri-based radiomics analysis for the prediction of breast tumor regression patterns after neoadjuvant chemotherapy. *Trans Oncol* (2020) 13(11):100831. doi: 10.1016/j.tranon.2020.100831

18. Tan H, Gan F, Wu Y, Zhou J, Tian J, Lin Y, et al. Preoperative prediction of axillary lymph node metastasis in breast carcinoma using radiomics features based on the fat-suppressed T2 sequence. *Acad Radiol* (2020) 27(9):1217–25. doi: 10.1016/j.acra.2019.11.004

19. Zwanenburg A, Vallières M, Abdalah MA, Aerts H, Andrearczyk V, Apte A, et al. The image biomarker standardization initiative: Standardized quantitative radiomics for high-throughput image-based phenotyping. *Radiology* (2020) 295(2):328–38. doi: 10.1148/radiol.2020191145

20. Kim HJ, Cui X, Hilsenbeck SG, Lee AV. Progesterone receptor loss correlates with human epidermal growth factor receptor 2 overexpression in estrogen receptor-positive breast cancer. *Clin Cancer Res* (2006) 12(3 Pt 2):1013s–8s. doi: 10.1158/1078-0432.Ccr-05-2128

21. Konecny G, Pauletti G, Pegram M, Untch M, Dandekar S, Aguilar Z, et al. Quantitative association between her-2/Neu and steroid hormone receptors in hormone receptor-positive primary breast cancer. *J Natl Cancer Inst* (2003) 95(2):142–53. doi: 10.1093/jnci/95.2.142

22. Milde-Langosch K, Karn T, Müller V, Witzel I, Rody A, Schmidt M, et al. Validity of the proliferation markers Ki67, Top2a, and Racgap1 in molecular subgroups of breast cancer. *Breast Cancer Res Treat* (2013) 137(1):57–67. doi: 10.1007/s10549-012-2296-x

23. Dorić M, Kuskunović-Vlahovljak S, Lazović Salčin E, Radović S, Čamdžić N, Babić M, et al. Correlation between numerical and categorical immunohistochemical score of ki-67 and Her2 with clinicopathological parameters of breast cancer. *Medicinski glasnik* (2021) 18(1):107–13. doi: 10.17392/1203-21

24. Zhou J, Tan H, Bai Y, Li J, Lu Q, Chen R, et al. Evaluating the her-2 status of breast cancer using mammography radiomics features. *Eur J Radiol* (2019) 121:108718. doi: 10.1016/j.ejrad.2019.108718

25. Zhou J, Tan H, Li W, Liu Z, Wu Y, Bai Y, et al. Radiomics signatures based on multiparametric mri for the preoperative prediction of the Her2 status of patients with breast cancer. *Acad Radiol* (2021) 28(10):1352–60. doi: 10.1016/j.acra.2020.05.040

26. Mendoza G, Portillo A, Olmos-Soto J. Accurate breast cancer diagnosis through real-time pcr her-2 gene quantification using immunohistochemically-identified biopsies. *Oncol Lett* (2013) 5(1):295–8. doi: 10.3892/ol.2012.984

frontiers | Frontiers in Oncology

TYPE Original Research
PUBLISHED 16 September 2022
DOI 10.3389/fonc.2022.916126

OPEN ACCESS

EDITED BY
Sweet Ping Ng,
University of Melbourne, Australia

REVIEWED BY
Fabrizio Urraro,
University of Campania Luigi Vanvitelli, Italy
Yongyi Zeng,
First Affiliated Hospital of Fujian Medical University, China

*CORRESPONDENCE
Ling Zhang
41733348@qq.com
Genggeng Qin
zealotq@smu.edu.cn

†These authors have contributed equally to this work

SPECIALTY SECTION
This article was submitted to Cancer Imaging and Image-directed Interventions, a section of the journal Frontiers in Oncology

RECEIVED 08 April 2022
ACCEPTED 05 September 2022
PUBLISHED 16 September 2022

CITATION
Zeng F, Dai H, Li X, Guo L, Jia N, Yang J, Huang D, Zeng H, Chen W, Zhang L and Qin G (2022) Preoperative radiomics model using gadobenate dimeglumine-enhanced magnetic resonance imaging for predicting β-catenin mutation in patients with hepatocellular carcinoma: A retrospective study. *Front. Oncol.* 12:916126. doi: 10.3389/fonc.2022.916126

Preoperative radiomics model using gadobenate dimeglumine-enhanced magnetic resonance imaging for predicting β-catenin mutation in patients with hepatocellular carcinoma: A retrospective study

Fengxia Zeng[1†], Hui Dai[2,3†], Xu Li[4], Le Guo[1], Ningyang Jia[5], Jun Yang[1], Danping Huang[1], Hui Zeng[1], Weiguo Chen[1], Ling Zhang[1*] and Genggeng Qin[1,6*]

[1]Department of Radiology, Nanfang Hospital, Southern Medical University, Guangzhou, China, [2]Hospital Office, Ganzhou People's Hospital, Ganzhou, China, [3]Hospital Office, Ganzhou Hospital-Nanfang Hospital, Southern Medical University, Ganzhou, China, [4]School of Biomedical Engineering, Southern Medical University, Guangzhou, China, [5]Department of Radiology, Eastern Hepatobiliary Surgery Hospital, Second Military Medical University, Shanghai, China, [6]Department of Radiology, Ganzhou Hospital-Nanfang Hospital, Southern Medical University, Ganzhou, China

Objective: To compare and evaluate radiomics models to preoperatively predict β-catenin mutation in patients with hepatocellular carcinoma (HCC).

Methods: Ninety-eight patients who underwent preoperative gadobenate dimeglumine (Gd-BOPTA)-enhanced MRI were retrospectively included. Volumes of interest were manually delineated on arterial phase, portal venous phase, delay phase, and hepatobiliary phase (HBP) images. Radiomics features extracted from different combinations of imaging phases were analyzed and validated. A linear support vector classifier was applied to develop different models.

Results: Among all 15 types of radiomics models, the model with the best performance was seen in the R^{HBP} radiomics model. The area under the receiver operating characteristic curve (AUC), accuracy, sensitivity, specificity of the R^{HBP} radiomics model in the training and validation cohorts were 0.86 (95% confidence interval [CI], 0.75–0.93), 0.75, 1.0, and 0.65 and 0.82 (95% CI, 0.63–0.93), 0.73, 0.67, and 0.76, respectively. The combined model integrated radiomics features in the R^{HBP} radiomics model, and signatures in the clinical model did not improve further compared to the single HBP radiomics model with AUCs of 0.86 and 0.76. Good calibration for the best R^{HBP} radiomics model was displayed in both cohorts; the decision curve showed that the net benefit could achieve 0.15. The most important radiomics features were low

10.3389/fonc.2022.916126

and high gray-level zone emphases based on gray-level size zone matrix with the same Shapley additive explanation values of 0.424.

Conclusion: The R^{HBP} radiomics model may be used as an effective model indicative of HCCs with β-catenin mutation preoperatively and thus could guide personalized medicine.

KEYWORDS

hepatocellular carcinoma, β-catenin mutation, magnetic resonance imaging, Gd-BOPTA, radiomics

Highlights

1. β-catenin can be preoperatively estimated using radiomics model based on Gd-BOPTA-enhanced MRI.
2. Among all 15 types of radiomics models, the model with the best performance was seen in the R^{HBP} radiomics model.
3. The R^{HBP} radiomics model may assist in the selection of appropriate decision-making for personalized medicine in patients with hepatocellular carcinoma.

Introduction

Hepatocellular carcinoma (HCC) has become the third most common cause of cancer-related deaths globally in 2020, making it a health problem worldwide (1). Although there has been recent progress in the treatment of HCC, the 5-year tumor recurrence occurs in approximately 35% of cases after liver transplantation, and 70% after hepatectomy indicates an unsatisfactory overall survival for patients with HCC (2–4). Additionally, patients with advanced HCC are not eligible for curative therapies. Thus, immunotherapy plays a critical role in HCC (5, 6).

Recently, programmed cell death 1 (PD-1) immune checkpoint inhibitors have shown efficacy in patients with HCC at advanced stages (7, 8). Anti-PD-1 therapy has revealed unprecedented response and disease control rates in clinical trials and has become the second-line therapy for HCC treatment granted by the Food and Drug Administration (9–11). Unfortunately, some patients failed to respond effectively, and not all patients showed positive results on prognosis, although excellent antitumor responses were observed with anti-PD-1 therapy (12, 13). Hence, selecting the subgroup that would receive the best benefit from anti-PD-1 antibody is important in the management of patients with HCC (13).

β-catenin activation can induce immune escape and resistance to anti-PD-1 therapy in HCC cases (14–17). Morita et al. (16) also showed that β-catenin without mutation was significantly correlated with longer survival in both progression-free survival and overall survival with anti-PD-1 therapy. The β-catenin mutation, which contributes to the activation of the Wnt/β-catenin signaling pathway, can be observed in approximately 30%–40% of patients with HCC (18). β-catenin, which is an intracellular signal transducer in the WNT signaling pathway, is encoded by CTNNB1 and closely related to the occurrence and development of liver tumors. Most liver tumors have mutations in genes encoding key components of the WNT/β-catenin signaling pathway (14, 18, 19). Studies have reported that the Wnt/β-catenin pathway is correlated with carcinogenesis, especially in hepatocellular adenoma (20, 21). However, in patients with HCC, presence of the β-catenin mutation may suggest better cell differentiation and a more favorable prognosis (22, 23). Currently, the diagnosis of β-catenin mutation depends on polymerase chain reaction or immunohistochemical analysis. Nuclear expression of β-catenin in immunohistochemical analysis can hint at the β-catenin mutation and activation of the β-catenin pathway. However, the sensitivity and specificity of nuclear β-catenin expression are limited. Its transcriptional product, glutamine synthetase (GS) expression, is a reliable biomarker of the β-catenin mutation, while GS expression in human HCC is not always associated with β-catenin mutation (24–26). Therefore, the diagnosis of β-catenin mutation should be confirmed by the expression of β-catenin and GS by immunohistochemical analysis.

No stable serological or genomic biomarkers of the β-catenin mutation have been found to date due to the high heterogeneity of HCC. Moreover, the diagnosis of HCC depends not only on postoperative histologic examination, but also on imaging methods (3, 27). Thus, an accurate and noninvasive β-catenin mutation before surgery plays an important role in the

prognostic assessment and selection of patients for anti-PD-1 therapy.

Signal intensity in liver-specific contrast-enhanced magnetic resonance imaging (MRI) is correlated with genetic alterations and molecular expression of HCC (28–32). Kitao et al. (30) suggested that the β-catenin mutation in HCC might show distinctive imaging findings in hepatobiliary phase (HBP) images. Nevertheless, these immunohistochemical predictors are difficult to detect using preoperative imaging methods. Radiomics is a rapidly growing methodology that permits digital decoding of medical images into multidimensional radiological features for noninvasive profiling of tumors. Several studies on HCC have demonstrated that unviewable radiomic features are closely associated with histopathologic features, especially in hepatobiliary-specific contrast-enhanced MRI. Feng et al. (33) constructed a radiomic feature-based nomogram using gadoxetic acid-enhanced MRI to achieve satisfactory preoperative prediction of microvascular invasion in patients with HCC. Wang et al. (34) showed that preoperative arterial and HBP imaging radiomic features could be a reliable biomarker to evaluate the CK19 status of HCC.

To our knowledge, no study has investigated the radiomics model of β-catenin mutation prediction. In this study, we aimed to determine the performance of a radiomic feature-based model using gadobenate dimeglumine (Gd-BOPTA)-enhanced MRI to predict β-catenin mutation and assist in the selection of optimal therapeutic strategies for patients with HCC.

Materials and methods

Ethics statements

Our study was approved by the Ethics Review Board of Eastern Hepatobiliary Surgery Hospital. The requirement for informed consent was waived because no protected health information was needed.

Study design and patient cohort

Between September 2016 and June 2017, 463 patients who underwent preoperative MRI at the Eastern Hepatobiliary Surgery Hospital were retrospectively enrolled and analyzed. Inclusion and exclusion criteria are shown in Figure 1.

Magnetic resonance imaging

Gd-BOPTA-enhanced MRI was performed using GE Optima MR360 1.5 T. After injecting Gd-BOPTA (0.1 mmoL/kg; MultiHance, Braccom) into patients' cubital veins at a flow rate of 2.0 mL/s, enhanced scanning of the arterial phase (AP, 22–25 seconds), portal venous phases (PVP, 50–60 seconds), and delayed phases (DP, 90–120 seconds) was performed, respectively. The scanning of HBP imaging was completed 60 minutes after injecting patients with normal liver function and 120 minutes after injecting patients with impaired function. The diagnosis of impaired liver function was made by clinical physicians. The scanning parameters are presented in Table 1.

Pathological examination

All tumor sections were reviewed by two experienced pathologists with 5–10 years of experience in HCC. Each tumor sample was first sectioned and then stained with hematoxylin-eosin. To determine the β-catenin mutation, the expressions of β-catenin and GS were examined by immunohistochemical analysis. In our study, the positive expressions of β-catenin and GS were attributed to the activation of β-catenin. However, no expression of β-catenin or GS was considered as HCC without a β-catenin mutation. HCC with a β-catenin mutation was shown in Supplementary Figure 1 and without a β-catenin mutation was shown in Supplementary Figure 2.

Magnetic resonance imaging signatures and clinical factor acquisition

Clinical factors (CFs), which included sex, age, alpha-fetoprotein, hepatitis B surface antigen (HBsAg), hepatitis B e antigen, and cirrhosis, were collected from patients' electronic medical records. Based on the Liver Imaging Reporting and Data System (2017) (LI-RADS-2017), two abdominal radiologists with >10 years of experience assessed the radiological features and worked in consensus. MRI signatures included the tumor size, tumor morphology, capsule, margins, rim, and peritumoral enhancement in the AP, hypointensity, and peritumoral hypointensity in the HBP.

Region of interest segmentation and feature extraction

Region of interest (ROI) segmentation was performed by a radiologist with 5 years of work experience and validated by a radiologist with >10 years of work experience using ITK-SNAP software (www.itk-snap.org). ROIs in AP, PVP, DP, and HBP images were delineated on each slice of the lesions and three-dimensional ROIs were generated accordingly (Figure 2). Overall, 1674 radiomic features were extracted from each MRI phase, which included 324 first-order features, 432 gray-level co-occurrence matrix features, 288 gray-level size zone matrix (GLSZM) features, 288 gray-level run length matrix features,

10.3389/fonc.2022.916126

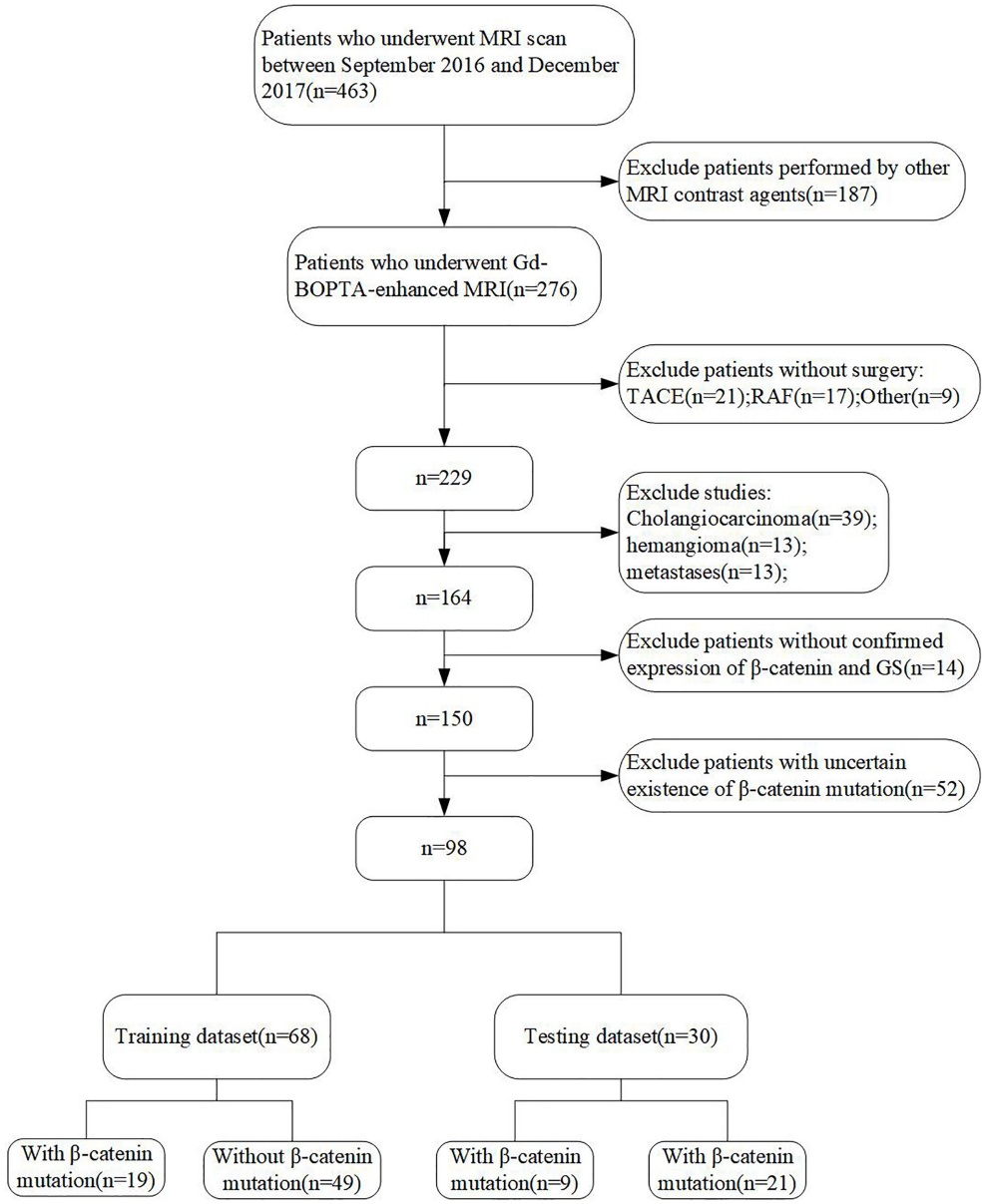


FIGURE 1
Flow chart of the study population. HCC, hepatocellular carcinoma; RAF, radiofrequency ablation; TACE, transcatheter arterial chemoembolization.

TABLE 1 Sequences and parameters of Gd-BOPTA dynamic-enhanced MRI.

Sequences	TR/TE (msec)	FOV (mm)	Thickness (mm)	Flip angle	Matrix
T1WI	190/4.3 (2)	420×420	6	80	256×160
T2WI	6667/85	420×420	6	160	320×224
DCE	3.7/1.7	420×420	2.5	15	256×192
HBP	3.7/1.7	420×420	2.5	15	256×192

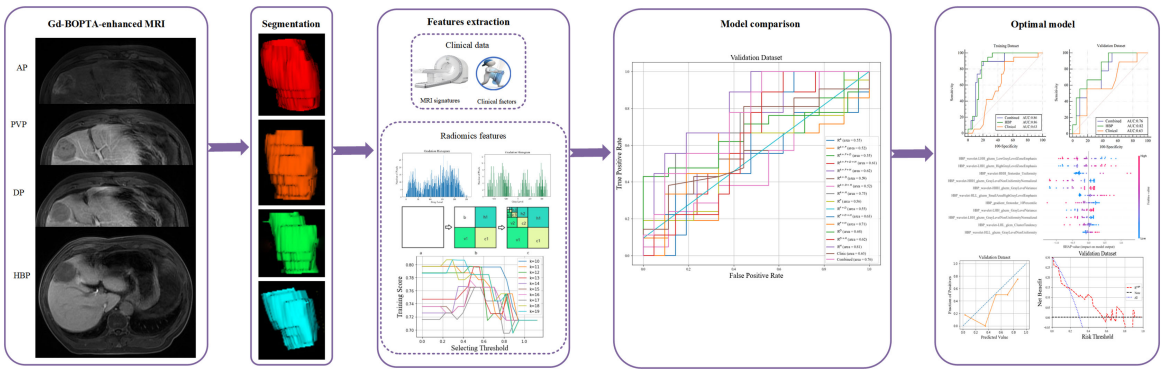


FIGURE 2
The workflow of the study.

90 neighboring gray-tone difference matrix features, and 252 gray-level dependence matrix features. All radiomic features were extracted using the open-source software Pyradiomics (https://pyradiomics.readthedocs.io/en/latest/index.html). Features in each of the MRI sequences were combined and categorized into different groups.

Feature selection and model construction

Patients were randomly allocated into a training/validation set (n=68) and a validation set (n=30) at a ratio of 7:3. The corresponding features were combined and categorized into four groups:

1. group 1 (n=4): features from one phase, named R^{AP}, R^{PVP}, R^{DP}, and R^{HBP} (1674 features for each type);
2. group 2 (n=6): features from any two phases, named R^{AP+PVP}, R^{AP+DP}, R^{AP+HBP}, R^{PVP+DP}, $R^{PVP+HBP}$, and R^{DP+HBP} (3348 features for each type);
3. group 3 (n=4): features from any three phases, named $R^{AP+PVP+DP}$, $R^{AP+PVP+HBP}$, and $R^{PVP+DP+HBP}$,and $R^{AP+DP+HBP}$ (5022 features for each type); and
4. group 4 (n=1): features from any four phases, named $R^{AP+PVP+DP+HBP}$ (6696 features for this type).

All the features in different groups were first analyzed using the F-test. The top 30 features in each phase and ROI were selected based on the results of the F-test. For the clinical model, all 14 CF and MRI signatures were analyzed by the F-test, and the top five features were chosen. Next, features in 15 types of combinations and clinical models were selected by linear support vector classifier (LSVC), and were then cross-validated by a 10-fold cross-validation using the training set to obtain the optimal parameter.

We chose the LSVC as the only classifier in all radiomic and clinical models. The predictive power of these models was evaluated by the area under the receiver operating characteristic curve (AUC), accuracy, sensitivity, and specificity. The best radiomics model for predicting β-catenin mutation was determined using the validation dataset by comparing the AUC values. A combined model that integrated the radiomic signature and clinical characteristics was also developed using the LSVC. Calibration curves were applied to analyze the performance of the best model, and a decision curve analysis was used to determine the clinical usefulness. Additionally, a local interpretability technique called Shapley additive explanation (SHAP) was used to break down predictions and show the impact of each feature. This assessment was based on SHAP values (Shap0.32.1), which were equal to the prediction from the original forecasted value minus the deletion of the feature. Positive results signified that the feature supported the prediction, whereas negative values signified that the prediction was not supported.

Statistical analysis

For baseline characteristics, continuous variables were analyzed using the t-test, and categorical variables were analyzed using the chi-square test or Fisher exact test. The AUC was compared using the DeLong test. Statistical analysis was performed using SPSS (version 22.0; IBM Corp.) and MedCalc (version 19.4.1). Statistical significance was set at $p<0.05$.

Results

Patient characteristics

The final cohort included 98 patients, who were randomly divided into training and validation datasets at a ratio of 7:3 (Figure 1). The longest interval between the MRI examination and

10.3389/fonc.2022.916126

surgery was approximately 2 weeks. The clinical characteristics and MRI features of all patients are shown in Table 2. There were no significant differences between the training and validation cohorts. In both datasets, most of the patients with HCC were male (85.3% and 83.3%, respectively), and the median ages were 55.63 and 54.13 years, respectively. The positive incidences of a β-catenin mutation were similar in the training and validation datasets (27.9% and 30%, respectively).

TABLE 2 Baseline clinical characteristics of the training and validation cohort.

Characteristic	Training dataset (N=68)	Validation dataset (N=30)	*P* value
Sex, N (%)			0.770*
Female	10 (14.7%)	5 (16.7%)	
Male	58 (85.3%)	25 (83.3%)	
Age, median ± SD (years)	55.63 ± 10.77	54.13 ± 11.90	0.540
Maximum diameter of tumor, median ± SD (mm)	41.60 ± 20.26	42.27 ± 18.76	0.879
Alpha fetoprotein, N (%)			0.816*
<20ng/ml	34 (50%)	17 (56.7%)	
20ng/ml-400ng/ml	27 (39.7%)	10 (33.3%)	
>400ng/ml	7 (10.3%)	3 (10%)	
HBsAg, N (%)			0.770*
Negative	10 (14.7%)	5 (16.7%)	
Positive	58 (85.3%)	25 (83.3%)	
HBeAg, N (%)			0.590
Negative	49 (72.1%)	20 (66.7%)	
Positive	19 (27.9%)	10 (33.3%)	
Cirrhosis, N (%)			0.743
Negative	18 (26.5%)	7 (23.3%)	
Positive	50 (73.5%)	23 (76.7%)	
Capsule, N (%)			0.506
Negative	54 (79.4%)	22 (73.3%)	
Positive	14 (20.6%)	8 (26.7%)	
Arterial rim enhancement, N (%)			0.219
Negative	43 (63.2%)	15 (50%)	
Positive	25 (36.8%)	15 (50%)	
Arterial peritumoral enhancement, N (%)			0.697*
Negative	63 (92.6%)	27 (90%)	
Positive	5 (7.4%)	3 (10%)	
Tumor margin, N (%)			0.810
Smooth	21 (30.9%)	10 (33.3%)	
Nonsmooth	47 (69.1%)	20 (66.7%)	
Tumor hypointensity on HBP, N (%)			\
Yes	68 (100%)	30 (100%)	
No	0	0	
Peritumoral hypointensity on HBP, N (%)			1.00*
Absent	59 (86.8%)	26 (86.7%)	
Present	9 (13.2%)	4 (13.3%)	
Shape, N (%)			0.095*
Round	15 (22.1%)	14 (46.7%)	
Oval	11 (16.2%)	4 (13.3%)	
Lobular	32 (47.1%)	8 (26.7%)	
Irregular	10 (14.7%)	4 (13.3%)	
β-catenin mutation, N (%)			0.835
Absent	49 (72.1%)	21 (70%)	
Present	19 (27.9%)	9 (30%)	

* Calculated by Fisher's exact test.

Performance of the clinical model

In total, 14 clinical characteristics and MRI features were used to construct the clinical model. After feature selection, four features including sex, HBsAg, age, and peritumoral hypointensity in the HBP were included in the clinical model. The AUCs of the clinical model were 0.63 (95% confidence interval [CI]: 0.50–0.74) in the training dataset and 0.63 (95% CI: 0.54–0.80) in the validation dataset.

Radiomics signature calculation

We performed LSVC modeling on AP, PVP, DP, and HBP features to explore the value of β-catenin mutation discrimination. For group 1, 12, four, four, and 11 features were selected for R^{AP}, R^{PVP}, R^{DP}, and R^{HBP} model construction, respectively. For the biphasic and triphasic MRI images, the number of features selected as putatively effective features ranged from 2 to 16. Twelve features were selected for group 4. The details are shown in Table 3.

Performance of the proposed models

All 15 types of radiomics models were compared to determine the best phases or combinations. The AUCs for all combinations of radiomics models are displayed in Table 4. The model with the best performance of was seen in the R^{HBP} radiomics model in the validation datasets, with AUCs of 0.82 among the R^{AP} (AUC=0.55), R^{PVP} (AUC=0.56), and R^{DP} radiomics models (AUC=0.68). For the biphasic MRI image, R^{AP+HBP}, $R^{PVP+HBP}$, and R^{AP+PVP} achieved excellent performance in the training datasets, but only the R^{AP+HBP} and $R^{PVP+HBP}$ radiomics models retained a moderate AUC (AUC=0.75 and 0.71, respectively) in the validation datasets. Additionally, the radiomics models did not perform better in the tri-phasic or quad-phasic MRI images in the validation datasets (Figure 3).

After combining clinical characteristics in the clinical model with radiomics features in the R^{HBP} radiomics model, the combined model did not perform better than the single HBP radiomics model in the differentiation of the β-catenin mutation, with AUCs of 0.86 in the training dataset and 0.76 in the validation dataset. Comparisons were made between the best-performing R^{HBP} radiomics model and the clinical model. The Delong test resulted in a *p*-value of 0.0043 in the training dataset (AUC: 0.86 versus [vs.] 0.63). Although there was no significant difference in the validation dataset (p=0.151), there was a trend in that the single HBP radiomics model showed better performance than the clinical model with a higher AUC in the validation dataset (Table 4 and Figure 4).

Calibration curves for the R^{HBP} radiomics model showed no significant difference between the predicted probabilities of the model and the ideal β-catenin mutation estimates in the training and validation datasets (p=0.081 and 0.454, respectively). The decision curve for the R^{HBP} radiomics model is shown in Figure 5. In our study, the net benefit could achieve 0.15, and the corresponding threshold probability of the curve was 0.18.

TABLE 3 Selected radiomics features of the proposed models.

Different model	Firstorder	GLCM	GLDM	GLRLM	GLSZM	NGTDM	Total
R^{AP}	1	7	0	2	2	0	12
R^{PVP}	0	0	0	1	3	0	4
R^{DP}	3	0	0	0	1	0	4
R^{HBP}	2	1	0	0	8	0	11
R^{AP+PVP}	2	2	1	1	6	0	12
R^{AP+DP}	2	0	0	1	2	0	5
R^{AP+HBP}	1	2	0	2	10	0	15
R^{PVP+DP}	2	0	0	3	4	0	9
$R^{PVP+HBP}$	1	1	0	0	10	0	12
R^{DP+HBP}	1	1	0	1	5	0	8
$R^{AP+PVP+DP}$	1	0	0	2	4	0	7
$R^{AP+PVP+HBP}$	2	2	0	3	9	0	16
$R^{AP+DP+HBP}$	1	0	0	0	1	0	2
$R^{PVP+DP+HBP}$	0	0	0	4	5	0	9
$R^{AP+PVP+DP+HBP}$	1	0	0	4	6	0	11

TABLE 4 Predictive performance of the proposed models.

Different model	Training dataset				Validation dataset			
	AUC (95%CI)	ACC	SEN	SPE	AUC(95%CI)	ACC	SEN	SPE
R^{AP}	0.74 (0.62-0.84)	0.69	0.68	0.69	0.55 (0.35-0.73)	0.70	0.44	0.81
R^{PVP}	0.82 (0.71-0.91)	0.71	0.95	0.61	0.56 (0.36-0.74)	0.46	0.33	0.52
R^{DP}	0.79 (0.67-0.88)	0.69	0.79	0.65	0.68 (0.48-0.83)	0.36	0.44	0.33
R^{HBP}	0.86 (0.75-0.93)	0.75	1	0.65	0.82 (0.63-0.93)	0.73	0.67	0.76
R^{AP+PVP}	0.90 (0.80-0.96)	0.80	0.89	0.76	0.52 (0.48-0.73)	0.60	0.33	0.71
R^{AP+DP}	0.86 (0.75-0.93)	0.69	0.95	0.59	0.59 (0.60-0.82)	0.40	0.66	0.28
R^{AP+HBP}	0.87 (0.77-0.94)	0.76	0.89	0.71	0.75 (0.61-0.86)	0.63	0.67	0.62
R^{PVP+DP}	0.47 (0.43-0.68)	0.75	0.84	0.71	0.55 (0.40-0.65)	0.60	0.33	0.71
$R^{PVP+HBP}$	0.93 (0.84-0.97)	0.81	0.95	0.76	0.71 (0.52-0.84)	0.60	0.56	0.61
R^{DP+HBP}	0.85 (0.75-0.92)	0.69	0.84	0.63	0.62 (0.56-0.79)	0.57	0.67	0.52
$R^{AP+PVP+DP}$	0.87 (0.77-0.94)	0.82	0.89	0.79	0.55 (0.52-0.76)	0.70	0.44	0.81
$R^{AP+PVP+HBP}$	0.94 (0.85-0.98)	0.81	0.89	0.78	0.62 (0.54-0.78)	0.60	0.44	0.67
$R^{AP+DP+HBP}$	0.85 (0.74-0.92)	0.76	0.89	0.71	0.52 (0.47-0.72)	0.50	0.67	0.43
$R^{PVP+DP+HBP}$	0.89 (0.79-0.95)	0.83	0.74	0.88	0.63 (0.57-0.81)	0.66	0.44	0.76
$R^{AP+PVP+DP+HBP}$	0.90 (0.81-0.96)	0.85	0.79	0.88	0.60 (0.51-0.82)	0.63	0.33	0.76
CF	0.63 (0.50-0.74)	0.57	0.89	0.45	0.63 (0.54-0.80)	0.3	0.44	0.24
Combined	0.86 (0.75-0.93)	0.69	0.89	0.61	0.76 (0.57-0.89)	0.63	0.56	0.67

CF, clinical factor; AP, arterial phase; PVP, portal venous phase; DP, delay phase; HBP, hepatobiliary phase; AUC, area under the curve; 95%CI, 95% confidence index; ACC, accuracy; SEN, sensitivity; SPE, specificity.

Importance of the features in the R^{HBP} radiomics model

We further explored the interpretability of the R^{HBP} radiomics model. The SHAP values for the features were obtained to explore their contributions to the model (Figure 6). The most important radiomics features were LGLZE and HGLZE based on GLSZM with the same SHAP values of 0.424. However, the difference between LGLZE and HGLZE was that the former supported the absence of the β-catenin mutation, whereas the latter supported the presence of the β-catenin mutation.

Discussion

To our knowledge, this is the first study to develop and validate Gd-BOPTA-enhanced MRI radiomics models for predicting the β-catenin mutation in HCC. Among the single-phase radiomics models in our study, the R^{HBP} radiomics model outperformed the other models in both the training and validation datasets. Some models, including bi-phase radiomics or others, could achieve better performance in the training dataset, but gain reverse results in the validation dataset. Furthermore, we established a combined model including radiomics from the best R^{HBP} radiomics model and features from a clinical model. However, the results showed that the combined model had the same performance as the single HBP radiomics model in the training datasets, but much worse performance in the validation datasets. This implies that radiomics based only on R^{HBP} images may aid in determining the activation of the β-catenin pathway effectively. The Hosmer-Lemeshow test of the training and validation datasets showed that the predicted curve was aligned with the ideal curve. Decision curve analysis in our study showed that the therapy strategy based on the R^{HBP} radiomics model was clinically useful.

We developed a clinical model combining clinical baseline factors and Gd-BOPTA-enhanced MRI features that demonstrated poor performance for preoperative prediction of the β-catenin mutation. Thus, it is difficult to distinguish the status of β-catenin by the macroscopical MRI features. Some scholars (30, 31, 35) have analyzed imaging findings, including a high enhancement ratio on R^{HBP} images, and a high apparent diffusion coefficient on diffusion-weighted imaging may be useful for identifying HCCs with the β-catenin mutation. While, these semi-quantitative parameter features cannot accurately assess the properties of the whole tumor. Radiomics can provide quantitative analysis of tumors. In our study, all the patients enrolled were showed hypointensity on R^{HBP} images, including patients with β-catenin mutation. However, the R^{HBP}

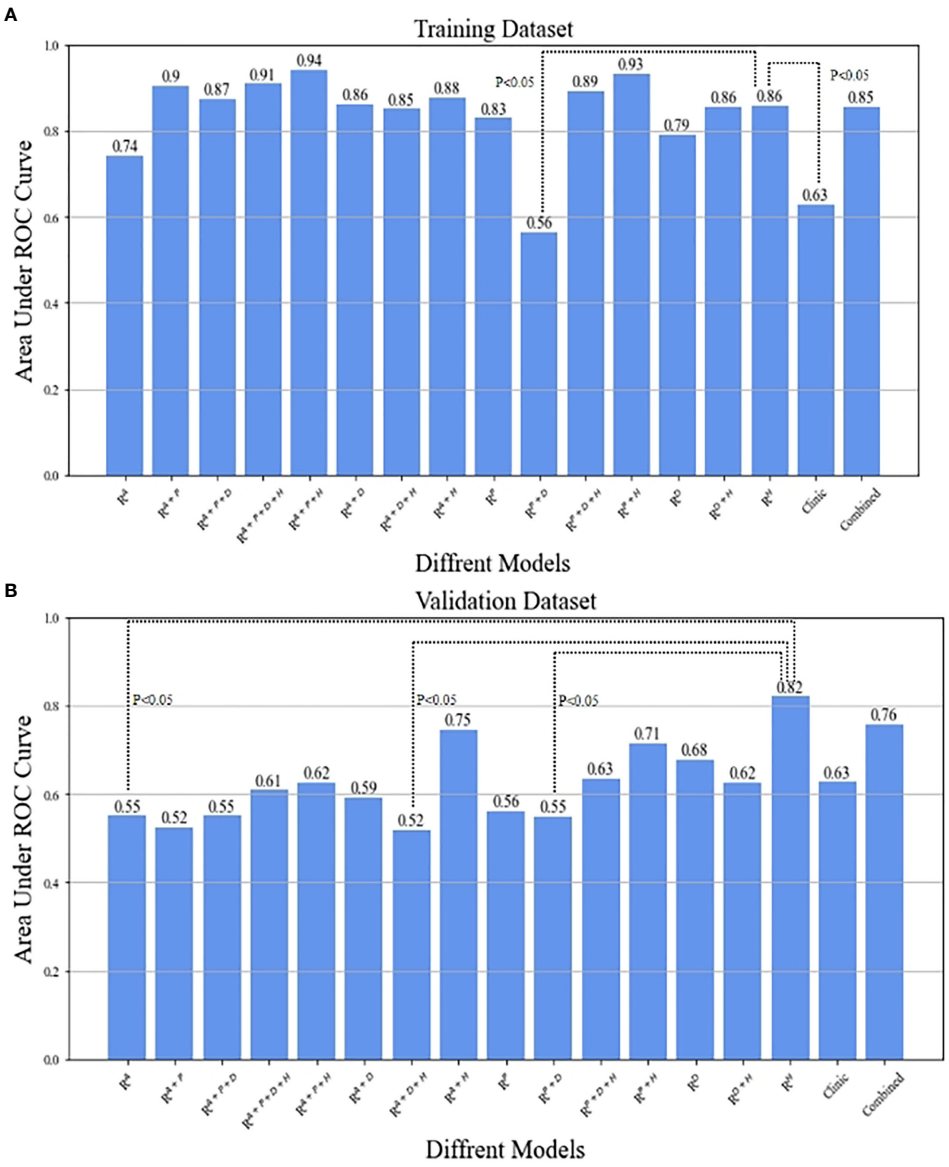


FIGURE 3
The comparison between R^{HBP} radiomics model in predicting β-catenin mutation and the other models by means of Delong test in the training **(A)** and validation **(B)** datasets.

radiomics model has achieved the best performance among all 15 types of radiomics models. We believe that more information related to the existence of the β- catenin mutation has been contained in the R^{HBP} images. Moreover, compared with visual and subjective imaging characteristics, radiomics can deeply excavate the information and provide a better prediction of the β-catenin mutation.

Many previous studies (31, 32) have demonstrated an intense correlation between signal intensity on the HBP images after liver-specific MR contrast administration and the expression of membranous uptake transporter organic anion-transporting polypeptide (OATP) 1B3. Additionally, the expression of OATP1B3 is due to the activation of β-catenin and/or hepatocyte nuclear factor 4α (36). Kitao et al. (30) reported a positive correlation between the expression of β-catenin and OATP1B3. Their results also showed that HCCs with the β-catenin mutation showed significantly higher enhancement ratios on HBP images than HCCs without. Reizine et al.'s study (35) of hepatocellular adenoma reported that liver-specific contrast uptake was strongly associated with

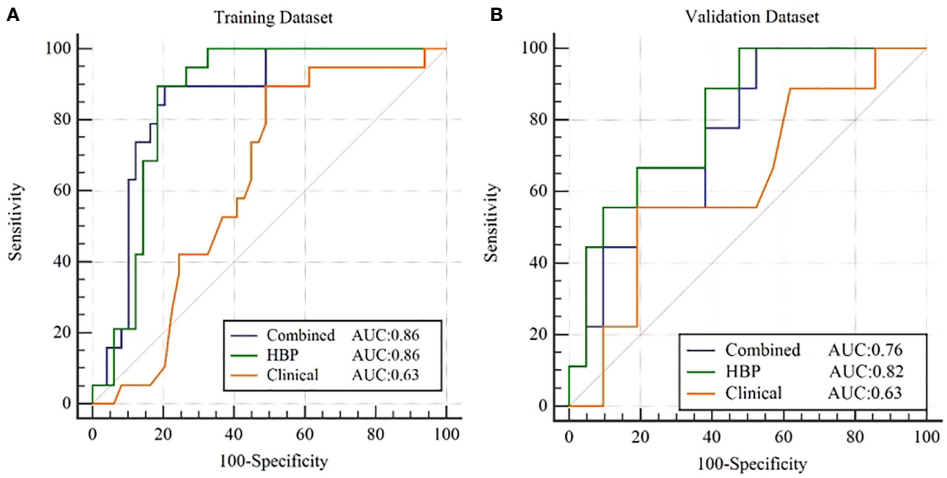


FIGURE 4
ROC curves for β-catenin mutation prediction of the clinical model, R^{HBP} radiomics model and combined model in the training **(A)** and validation **(B)** dataset.

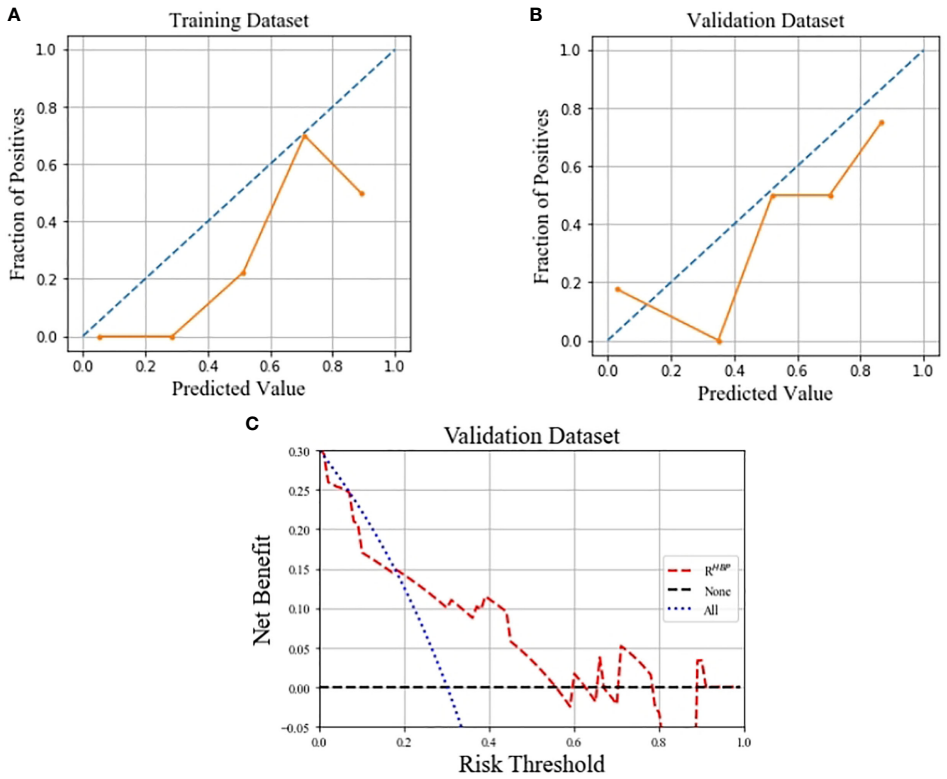


FIGURE 5
Calibration curves of the R^{HBP} radiomics model in predicting β-catenin mutation on the training **(A)** and validation **(B)** dataset, which demonstrated good agreement with the ideal curve. Decision curve analysis for the R^{HBP} radiomics model in the validation dataset **(C)**.

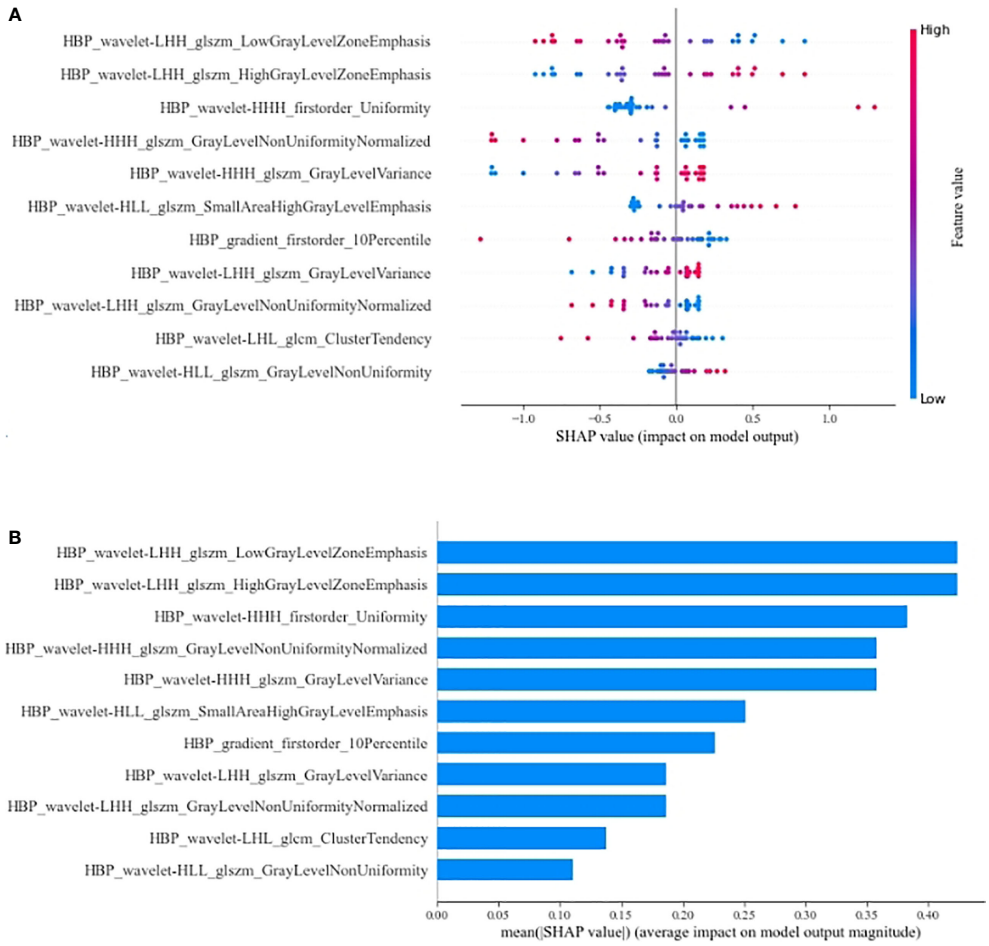


FIGURE 6
Summary plot **(A)** and bar plot **(B)** for the SHAP value of radiomics features on R^{HBP} radiomics model. LGLZE, low gray level zone emphasis; HGLZE, high gray level zone emphasis.

activation of the β-catenin pathway. These findings were consistent with our results. Our results showed that the R^{HBP} radiomics model can achieve better performance than other single-phase and combined-phase radiomics models in the test datasets. By assigning a corresponding SHAP value to each feature in the best R^{HBP} radiomics model, we found that the most important features were LGLZE and HGLZE with the same SHAP value. LGLZE and HGLZE belong to GLSZM features, which describe the darkness and brightness of a lesion. LGLZE measures the distribution of lower gray-level zones, with a higher value indicating a greater proportion of lower gray-level values and size zones in the image, and HGLZE measures the distribution of the higher gray-level values, with a higher value indicating a greater proportion of higher gray-level values and size zones in the image. A larger value of LGLZE indicates a darker lesion on HBP images, which represents a low probability of HCC with the β-catenin mutation. Similarly, a larger value of HGLZE indicates a higher probability for the activation of the β-catenin pathway.

The current study has several advantages. First, unlike previous studies, we systematically evaluated and compared different MRI phases and their combinations. Additionally, we found an easy-to-use model that included only the proposed R^{HBP} radiomics signature to simplify prediction of the β-catenin mutation. Finally, the similar performance in the training and validation datasets made our R^{HBP} radiomics model more objective than others.

Our study also has limitations. First, since numerous potential participants did not undergo immunohistochemical analysis or preoperative liver-specific contrast-enhanced MRI, they were excluded, which may have created a selection bias and limited the validity of our results. Larger datasets should be used to

validate the performance of our model in future studies. Second, the diagnosis of β-catenin activation was based on immunohistochemical analysis. The β-catenin mutation is supposed to be confirmed by polymerase chain reaction, but this was difficult to perform because of the retrospective nature of this study. However, previous studies (15, 16, 30) have shown that the expression of β-catenin and GS could represent the activation of β-catenin. We believe that the data in our study were sufficient to conclude that radiomics on HBP images was useful in preoperative individual prediction of HCCs with the β-catenin mutation. Our validation and training datasets were from the same center. Data from multiple centers should be used to assess the stability and generalizability of our findings.

In conclusion, the R^{HBP} radiomics model showed an excellent AUC with moderate sensitivity and specificity in both the training and validation cohorts. It can be used as an effective model to predicts HCCs with the β-catenin mutation preoperatively and thus may assist in the selection of appropriate decision-making for personalized medicine in patients with HCC.

Methodology

- retrospective
- diagnostic or prognostic study
- performed at one institution

Data availability statement

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

Ethics statement

The studies involving human participants were reviewed and approved by the Ethics Review Board of Eastern Hepatobiliary Surgery Hospital. Written informed consent for participation was not required for this study in accordance with the national legislation and the institutional requirements.

Author contributions

FZ, HD, and GQ designed the research. FZ, HD, and LZ performed the experiment. FZ, XL, LG, NJ, YJ, DH, and HZ collected and analyzed the data; FZ and HD prepared the original draft. LZ, GQ, and WC reviewed and edited the paper. All authors contributed to the article and approved the submitted version.

Funding

This study was funded by the National Natural Science Foundation of China (82171929), National Key Research and Development Program of China [2019YFC0121903] and [2019YFC0117301], the Foundation of President of Nanfang Hospital [2021C007], Natural Science Funding of Guangdong Province [2018A0303130215], [2018A030313951] and [2019A1515011168].

Acknowledgments

We would like to thank the whole study team for continuous support.

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

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fonc.2022.916126/full#supplementary-material

SUPPLEMENTARY FIGURE 1
Photomicrographs show pathologic findings of HCCs with β-catenin mutation. **(A)** Hematoxylineosin staining (magnification, ×100) shows HCC with polygonal cells and round nucleus. At immunohistochemical analysis (magnification, ×200), tumor shows intense expression of, **(B)** nuclear β-catenin, **(C)** cytoplasmic GS.

SUPPLEMENTARY FIGURE 2
Photomicrographs show pathologic findings of HCCs without β-catenin mutation. **(A)** Hematoxylineosin staining (magnification, ×100) shows HCC with polygonal cells and round nucleus. At immunohistochemical analysis (magnification, ×200), tumor shows no definite expression of, **(B)** nuclear β-catenin, **(C)** cytoplasmic GS.

References

1. Sung H, Ferlay J, Siegel RL, Laversanne M, Soerjomataram I, Jemal A, et al. Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin* (2021) 71(3):209–49. doi: 10.3322/caac.21660

2. Raoul JL, Edeline J. Systemic treatment of hepatocellular carcinoma: standard of care in China and elsewhere. *Lancet Oncol* (2020) 21(4):479–81. doi: 10.1016/S1470-2045(20)30082-6

3. Marrero JA, Kulik LM, Sirlin CB, Zhu AX, Finn RS, Abecassis MM, et al. Diagnosis, staging, and management of hepatocellular carcinoma: 2018 practice guidance by the American association for the study of liver diseases. *Hepatology* (2018) 68(2):723–50. doi: 10.1002/hep.29913

4. Dimitroulis D, Damaskos C, Valsami S, Davakis S, Garmpis N, Spartalis E, et al. From diagnosis to treatment of hepatocellular carcinoma: An epidemic problem for both developed and developing world. *World J Gastroenterol* (2017) 23 (29):5282–94. doi: 10.3748/wjg.v23.i29.5282

5. Llovet JM, Montal R, Sia D, Finn RS. Molecular therapies and precision medicine for hepatocellular carcinoma. *Nat Rev Clin Oncol* (2018) 15(10):599–616. doi: 10.1038/s41571-018-0073-4

6. Sonbol MB, Riaz IB, Naqvi SAA, Almquist DR, Mina S, Almasri J, et al. Systemic therapy and sequencing options in advanced hepatocellular carcinoma: A systematic review and network meta-analysis. *JAMA Oncol* (2020) 6(12):e204930. doi: 10.1001/jamaoncol.2020.4930

7. Fessas P, Lee H, Ikemizu S, Janowitz T. A molecular and preclinical comparison of the PD-1-targeted T-cell checkpoint inhibitors nivolumab and pembrolizumab. *Semin Oncol* (2017) 44(2):136–40. doi: 10.1053/j.seminoncol.2017.06.002

8. Yau T, Kang YK, Kim TY, El-Khoueiry AB, Santoro A, Sangro B, et al. Efficacy and safety of nivolumab plus ipilimumab in patients with advanced hepatocellular carcinoma previously treated with sorafenib: The CheckMate 040 randomized clinical trial. *JAMA Oncol* (2020) 6(11):e204564. doi: 10.1001/jamaoncol.2020.4564

9. El-Khoueiry AB, Sangro B, Yau T, Crocenzi TS, Kudo M, Hsu C, et al. Nivolumab in patients with advanced hepatocellular carcinoma (CheckMate 040): an open-label, non-comparative, phase 1/2 dose escalation and expansion trial. *Lancet* (2017) 389(10088):2492–502. doi: 10.1016/S0140-6736(17)31046-2

10. Ren Z, Xu J, Bai Y, Xu A, Cang S, Du C, et al. Sintilimab plus a bevacizumab biosimilar (IBI305) versus sorafenib in unresectable hepatocellular carcinoma (ORIENT-32): a randomised, open-label, phase 2-3 study. *Lancet Oncol* (2021) 22(7):977–90. doi: 10.1016/S1470-2045(21)00252-7

11. Saung MT, Pelosof L, Casak S, Donoghue M, Lemery S, Yuan M, et al. FDA Approval summary: Nivolumab plus ipilimumab for the treatment of patients with hepatocellular carcinoma previously treated with sorafenib. *Oncologist* (2021) 26 (9):797–806. doi: 10.1002/onco.13819

12. Ott PA, Hodi FS, Kaufman HL, Wigginton JM, Wolchok JD. Combination immunotherapy: a road map. *J Immunother Cancer* (2017) 5:16. doi: 10.1186/s40425-017-0218-5

13. Ernstoff MS, Gandhi S, Pandey M, Puzanov I, Grivas P, Montero A, et al. Challenges faced when identifying patients for combination immunotherapy. *Future Oncol* (2017) 13(18):1607–18. doi: 10.2217/fon-2017-0218

14. He S, Tang S. WNT/β-catenin signaling in the development of liver cancers. *BioMed Pharmacother* (2020) 132:110851. doi: 10.1016/j.biopha.2020.110851

15. Ruiz de Galarreta M, Bresnahan E, Molina-Sánchez P, Lindblad KE, Maier B, Sia D, et al. β-catenin activation promotes immune escape and resistance to anti-PD-1 therapy in hepatocellular carcinoma. *Cancer Discovery* (2019) 9(8):1124–41. doi: 10.1158/2159-8290.CD-19-0074

16. Morita M, Nishida N, Sakai K, Aoki T, Chishina H, Takita M, et al. Immunological microenvironment predicts the survival of the patients with hepatocellular carcinoma treated with anti-PD-1 antibody. *Liver Cancer* (2021) 10(4):380–93. doi: 10.1159/000516899

17. Deldar Abad Paskeh M, Mirzaei S, Ashrafizadeh M, Ashrafizadeh M, Zarrabi A, Sethi G, et al. Wnt/β-catenin signaling as a driver of hepatocellular carcinoma progression: An emphasis on molecular pathways. *J Hepatocell Carcinoma* (2021) 8:1415–44. doi: 10.2147/JHC.S336858

18. Miyoshi Y, Iwao K, Nagasawa Y, Aihara T, Sasaki Y, Imaoka S, et al. Activation of the beta-catenin gene in primary hepatocellular carcinomas by somatic alterations involving exon 3. *Cancer Res* (1998) 58(12):2524–7.

19. Rogacki K, Kasprzak A, Stępiński A. Alterations of wnt/β-catenin signaling pathway in hepatocellular carcinomas associated with hepatitis c virus. *Pol J Pathol* (2015) 66(1):9–21. doi: 10.5114/pjp.2015.51148

20. Nault JC, Couchy G, Balabaud C, Morcrette G, Caruso S, Blanc JF, et al. Molecular classification of hepatocellular adenoma associates with risk factors, bleeding, and malignant transformation. *Gastroenterology* (2017) 152(4):880–894.e6. doi: 10.1053/j.gastro.2016.11.042

21. Nault JC, Paradis V, Cherqui D, Vilgrain V, Zucman-Rossi J. Molecular classification of hepatocellular adenoma in clinical practice. *J Hepatol* (2017) 67 (5):1074–83. doi: 10.1016/j.jhep.2017.07.009

22. Zucman-Rossi J, Benhamouche S, Godard C, Boyault S, Grimber G, Balabaud C, et al. Differential effects of inactivated Axin1 and activated beta-catenin mutations in human hepatocellular carcinomas. *Oncogene* (2007) 26 (5):774–80. doi: 10.1038/sj.onc.1209824

23. Dal Bello B, Rosa L, Campanini N, Tinelli C, Torello Viera F, D'Ambrosio G, et al. Glutamine synthetase immunostaining correlates with pathologic features of hepatocellular carcinoma and better survival after radiofrequency thermal ablation. *Clin Cancer Res* (2010) 16(7):2157–66. doi: 10.1158/1078-0432.CCR-09-1978

24. Austinat M, Dunsch R, Wittekind C, Tannapfel A, Gebhardt R, Gaunitz F, et al. Correlation between beta-catenin mutations and expression of wnt-signaling target genes in hepatocellular carcinoma. *Mol Cancer* (2008) 7:21. doi: 10.1186/1476-4598-7-21

25. Adebayo Michael AO, Ko S, Tao J, Moghe A, Yang H, Xu M, et al. Inhibiting glutamine-dependent mTORC1 activation ameliorates liver cancers driven by β-catenin mutations. *Cell Metab* (2019) 29(5):1135–1150.e6. doi: 10.1016/j.cmet.2019.01.002

26. Cadoret A, Ovejero C, Terris B, Souil E, Lévy L, Lamers WH, et al. New targets of beta-catenin signaling in the liver are involved in the glutamine metabolism. *Oncogene* (2002) 21(54):8293–301. doi: 10.1038/sj.onc.1206118

27. Vogel A, Cervantes A, Chau I, Daniele B, Llovet JM, Meyer T, et al. Hepatocellular carcinoma: ESMO clinical practice guidelines for diagnosis, treatment and follow-up. *Ann Oncol* (2018) 29(Suppl 4):iv238–55. doi: 10.1093/annonc/mdy308

28. Chen J, Wu Z, Xia C, Jiang H, Liu X, Duan T, et al. Noninvasive prediction of HCC with progenitor phenotype based on gadoxetic acid-enhanced MRI. *Eur Radiol* (2020) 30(2):1232–42. doi: 10.1007/s00330-019-06414-2

29. Ye Z, Cao L, Wei Y, Chen J, Zhang Z, Yao S, et al. Preoperative prediction of hepatocellular carcinoma with highly aggressive characteristics using quantitative parameters derived from hepatobiliary phase MR images. *Ann Transl Med* (2020) 8 (4):85. doi: 10.21037/atm.2020.01.04

30. Kitao A, Matsui O, Yoneda N, Kozaka K, Kobayashi S, Sanada J, et al. Hepatocellular carcinoma with β-catenin mutation: Imaging and pathologic characteristics. *Radiology* (2015) 275(3):708–17. doi: 10.1148/radiol.14141315

31. Kitao A, Zen Y, Matsui O, Gabata T, Kobayashi S, Koda W, et al. Hepatocellular carcinoma: signal intensity at gadoxetic acid-enhanced MR imaging–correlation with molecular transporters and histopathologic features. *Radiology* (2010) 256(3):817–26. doi: 10.1148/radiol.10092214

32. Narita M, Hatano E, Arizono S, Hayashino A, Isoda H, Kitamura K, et al. Expression of OATP1B3 determines uptake of gd-EOB-DTPA in hepatocellular carcinoma. *J Gastroenterol* (2009) 44(7):793–8. doi: 10.1007/s00535-009-0056-4

33. Feng ST, Jia Y, Liao B, Huang B, Zhou Q, Li X, et al. Preoperative prediction of microvascular invasion in hepatocellular cancer: a radiomics model using gd-EOB-DTPA-enhanced MRI. *Eur Radiol* (2019) 29(9):4648–59. doi: 10.1007/s00330-018-5935-8

34. Wang W, Gu D, Wei J, Ding Y, Yang L, Zhu K, et al. A radiomics-based biomarker for cytokeratin 19 status of hepatocellular carcinoma with gadoxetic acid-enhanced MRI. *Eur Radiol* (2020) 30(5):3004–14. doi: 10.1007/s00330-019-06585-y

35. Reizine E, Ronot M, Ghosn M, Calderaro J, Frulio N, Bioulac-Sage, et al. Hepatospecific MR contrast agent uptake on hepatobiliary phase can be used as a biomarker of marked β-catenin activation in hepatocellular adenoma. *Eur Radiol* (2021) 31(5):3417–26. doi: 10.1007/s00330-020-07434-z

36. Sekine S, Ogawa R, Ojima H, Kanai Y. Expression of SLCO1B3 is associated with intratumoral cholestasis and CTNNB1 mutations in hepatocellular carcinoma. *Cancer Sci* (2011) 102(9):1742–7. doi: 10.1111/j.1349-7006.2011.01990.x

frontiers | Frontiers in Oncology

TYPE Original Research
PUBLISHED 20 September 2022
DOI 10.3389/fonc.2022.906498

OPEN ACCESS

EDITED BY
Nanna Maria Sijtsema,
University Medical Center Groningen, Netherlands

REVIEWED BY
Wen-Yen Huang,
Tri-Service General Hospital, Taiwan
Shixiong Liang,
Guangxi Medical University Cancer Hospital, China

*CORRESPONDENCE
Yu-Jen Chen
chenmdphd@gmail.com
Shih-Ming Hsu
smhsu@ym.edu.tw

SPECIALTY SECTION
This article was submitted to Cancer Imaging and Image-directed Interventions, a section of the journal Frontiers in Oncology

RECEIVED 28 March 2022
ACCEPTED 26 August 2022
PUBLISHED 20 September 2022

CITATION
Huang Y-M, Wang T-E, Chen M-J, Lin C-C, Chang C-W, Tai H-C, Hsu S-M and Chen Y-J (2022) Radiomics-based nomogram as predictive model for prognosis of hepatocellular carcinoma with portal vein tumor thrombosis receiving radiotherapy. *Front. Oncol.* 12:906498. doi: 10.3389/fonc.2022.906498

Radiomics-based nomogram as predictive model for prognosis of hepatocellular carcinoma with portal vein tumor thrombosis receiving radiotherapy

Yu-Ming Huang[1,2,3], Tsang-En Wang[2,4,5], Ming-Jen Chen[2,4,5], Ching-Chung Lin[2,4,5], Ching-Wei Chang[2,4,5], Hung-Chi Tai[3,6], Shih-Ming Hsu[3]* and Yu-Jen Chen[2,5,6,7,8]*

[1]Department of Radiation Oncology, Taipei Hospital, Ministry of Health and Welfare, New Taipei City, Taiwan, [2]Department of Medicine, MacKay Medical College, New Taipei City, Taiwan, [3]Department of Biomedical Imaging and Radiological Sciences, National Yang Ming Chiao Tung University, Taipei, Taiwan, [4]Division of Gastroenterology, Department of Internal Medicine, MacKay Memorial Hospital, Taipei, Taiwan, [5]Department of Artificial Intelligence and Medical Application, MacKay Junior College of Medicine, Nursing, and Management, New Taipei City, Taiwan, [6]Department of Radiation Oncology, MacKay Memorial Hospital, Taipei, Taiwan, [7]Department of Medical Research, MacKay Memorial Hospital, Taipei, Taiwan, [8]Department of Medical Research, China Medical University Hospital, Taichung, Taiwan

Background: This study aims to establish and validate a predictive model based on radiomics features, clinical features, and radiation therapy (RT) dosimetric parameters for overall survival (OS) in hepatocellular carcinoma (HCC) patients treated with RT for portal vein tumor thrombosis (PVTT).

Methods: We retrospectively reviewed 131 patients. Patients were randomly divided into the training (n = 105) and validation (n = 26) cohorts. The clinical target volume was contoured on pre-RT computed tomography images and 48 textural features were extracted. The least absolute shrinkage and selection operator regression was used to determine the radiomics score (rad-score). A nomogram based on rad-score, clinical features, and dosimetric parameters was developed using the results of multivariate regression analysis. The predictive nomogram was evaluated using Harrell's concordance index (C-index), area under the curve (AUC), and calibration curve.

Results: Two radiomics features were extracted to calculate the rad-score for the prediction of OS. The radiomics-based nomogram had better performance than the clinical nomogram for the prediction of OS, with a C-index of 0.73 (95% CI, 0.67–0.79) and an AUC of 0.71 (95% CI, 0.62–0.79). The predictive accuracy was assessed by a calibration curve.

Conclusion: The radiomics-based predictive model significantly improved OS prediction in HCC patients treated with RT for PVTT.

KEYWORDS

hepatocellular carcinoma, portal vein tumor thrombosis, radiation therapy, radiomics, predictive model

Introduction

Hepatocellular carcinoma (HCC) is the sixth most common cancer and the third leading cause of cancer death worldwide. The prognosis of HCC is poor, with a 5-year survival rate of 5%–18% (1–4). Approximately 70% of newly diagnosed HCC patients are not suitable for curative local treatment (5). The major cause is macrovascular invasion, in which tumor cells invade the portal vein, hepatic vein, or the inferior vena cava in the liver (6). Portal vein tumor thrombosis (PVTT) is a common complication of HCC and is related to poor prognosis and poor response to local treatment. The incidence of PVTT in HCC ranges from 44% to 62% (7). PVTT can interfere with the portal blood supply in the normal liver and deteriorate liver function. It may contribute to intrahepatic or extrahepatic metastasis (8). This locally advanced and mostly unresectable disease is associated with rapid cancer progression and deterioration of liver function. Patients with PVTT have a median survival rate of only 3 months without treatment (9). Current treatments for HCC with PVTT include targeted therapy with sorafenib and lenvatinib and locoregional treatments such as operation (OP), radiation therapy (RT), transarterial chemoembolization (TACE), and transarterial radioembolization (TARE) (10–13). However, there is no consensus on the best forms of treatment for HCC patients with PVTT. Several clinical studies have reported that RT alone or combined with TACE is an effective treatment for HCC with PVTT (14–17). The clinical target volume (CTV) of RT for PVTT usually encompasses the area of PVTT and/or visible tumor with a 5–10–mm margin to cover the involved portal vein region (18). The advantages of RT for HCC with PVTT are local tumor control, portal vein patency, and survival benefit (19). No universal marker or method of clinical utility that can predict the survival of HCC patients treated with RT for PVTT is known. An effective predictive model that may guide precision medicine for these patients with generally poor survival is required.

HCC can be diagnosed on contrast-enhanced computed tomography (CT) or magnetic resonance imaging (MRI) (20). Therefore, HCC is frequently diagnosed on images alone, precluding the requirement for tissue proof. Currently, CT is routinely used by physicians for diagnosis, staging, and RT planning for HCC. Radiomics is an emerging and promising methodology for medical image analysis that converts medical images into high-dimensional quantitative features using machine learning algorithms and statistical analysis software. Thus, it may facilitate the detection of lesions (21, 22), improve diagnostic accuracy (23–25), predict disease risk and prognosis (26–32), evaluate the risk of treatment and treatment-related toxicities (33–37), and guide treatment strategies (38, 39) in different types of diseases, especially malignancies. Several studies have been published on the use of radiomics in HCC (40–43). Wang et al. analyzed the prognostic value of MRI textural features in HCC in 201 patients who underwent OP (44). Meng et al. integrated intratumoral and peritumoral CT radiomics features and clinical features to develop and validate a radiomics-based predictive nomogram to predict overall survival (OS) in HCC patients undergoing TACE (45). Cozzi et al. appraised the ability of a radiomics-based analysis to predict local response and OS in HCC patients who were eligible for curative or palliative RT (46). To the best of our knowledge, relatively limited data and few studies focused on prognosis estimation in HCC patients treated with RT for PVTT with radiomics analysis are available. This study uses radiomics features of CTV, which are derived from the pre-RT CT of HCC patients with PVTT, in combination with clinical features and RT dosimetric parameters to develop a predictive model for HCC with PVTT.

Material and methods

Patients

We retrospectively reviewed HCC patients newly diagnosed with PVTT between December 2007 and December 2019 in one institution. A contrast-enhanced CT or MRI was performed for diagnosis and staging. According to the 7th edition American Joint Committee on Cancer/American Joint Committee on

Cancer staging system, all patients were staged IIIB (patients with a single tumor or multiple tumors of any size involving a major branch of the portal vein or hepatic vein, Vp4 in Liver Cancer Study Group of Japan classification). All patients had an Eastern Cooperative Oncology Group (ECOG) performance status of 0 to 2. In this study, patients were either inoperable or not eligible for TACE or TARE. The primary treatments were RT and/or targeted therapy. Patients with a history of OP, RT, TACE, or TARE were excluded. A total of 131 patients were enrolled and randomly divided into the training cohort (n = 105) and validation cohort (n = 26), with a ratio of 4:1.

RT protocol

Patients underwent CT simulation in the supine position and were immobilized with an alpha cradle. Planning CT images with a slice thickness of 3 mm were acquired through the entire upper abdomen. Contrast-enhanced CT was used to localize the PVTT along with the primary tumor and to assess the enhancement patterns of lesions. The gross tumor volume (GTV) was delineated using the diagnostic and simulation images of the PVTT with or without the primary liver tumor. The CTV was determined by expanding the GTV margin by 5–10 mm to consider areas at significant risk of microscopic disease. The planning target volume (PTV) was generated by adding a 5–10–mm margin to the CTV in all directions for a setup error. RT was delivered using either three-dimensional conformal radiotherapy or intensity-modulated radiation therapy (IMRT) based on physician preference. The treatment plans were designed using 6- or 10-MV photons. All patients were treated with linear accelerators. Dosimetric parameters such as the dose of the CTV and normal organs were extracted from RT planning systems (Eclipse Treatment Planning System; Varian Medical Systems Inc., Palo Alto, CA, USA). The prescribed dose was 45, 50, or 60 Gy delivered in 1.8–2 Gy per fraction (BED10: 53.1–72.0 Gy). The goals were to deliver the prescribed dose to ≥95% of the PTV and 95% of the prescribed dose to ≥99% of the PTV. The dosimetric parameters were recorded for evaluation. After RT, abdominal CT or MRI was performed for response assessment. Most patients underwent abdominal CT or MRI 1 month after RT. The patency status of the portal vein area was evaluated by experienced radiologists.

Acquisition of CT images

Contrast-enhanced CT was performed using Philips MRC 800 (Philips Medical Systems, Amsterdam, Netherlands) with a peak tube voltage of 120 kVp, tube current of 325 mA, rotation time of 0.75 s, matrix of 512 × 512, field of view of 50 cm, and slice thickness of 3 mm for RT planning and radiomics analysis.

Texture analysis

The CTV, the region of interest (ROI), was contoured by experienced radiation oncologists on all axial CT images. Segmentation was performed using the Eclipse system. Three-dimension ROI was visualized using Local Image Features Extraction (LIFEx) version 5.10 (http://www.lifexsoft.org; Orsay, France) (47). The LIFEx software was used to extract the textural features of the ROI. A total of 48 textural features of the images were extracted, including features of a histogram-based matrix, gray-level co-occurrence matrix (GLCM), gray-level run length matrix (GLRLM), neighborhood gray-level dependence matrix (NGLDM), and gray-level zone length matrix (GLZLM) (Figure 1).

Extraction of radiomics features

The study population was divided into the training and validation cohorts in a ratio of 4:1 using the sample function of R (version 3.6.1) software (https://www.r-project.org; Vienna, Austria) to make randomization. The least absolute shrinkage and selection operator (LASSO) Cox regression was performed to determine the radiomics features that can predict OS in the training cohort. We performed the 10-fold cross-validation 20 times. The final value of lambda (penalized parameter) was determined with the minimized mean deviance and the corresponding subset of covariates with non-zero coefficients. Features were selected by the total times of non-zero coefficient in 20 randomized 10-fold cross-validations. The Cox proportional-hazard model was fitted with the selected features, and the radiomics score (rad-score) predicting OS could be calculated linearly.

Clinical feature extraction

The following 17 clinical features were selected: age, gender, etiology of viral hepatitis, drinking history, ECOG performance status, Child-Pugh class, tumor size, anemia status, serum levels of alpha-fetoprotein (AFP), white blood cell, platelet, albumin, alanine aminotransferase, aspartate aminotransferase (AST), total bilirubin, creatinine, and prothrombin time.

RT dosimetric parameters

The prescribed RT doses, RT fields as involved PVTT with or without primary liver tumors, CTV, normal liver volume (NLV), and mean liver doses (MLDs) of all patients were recorded (Figure 2).

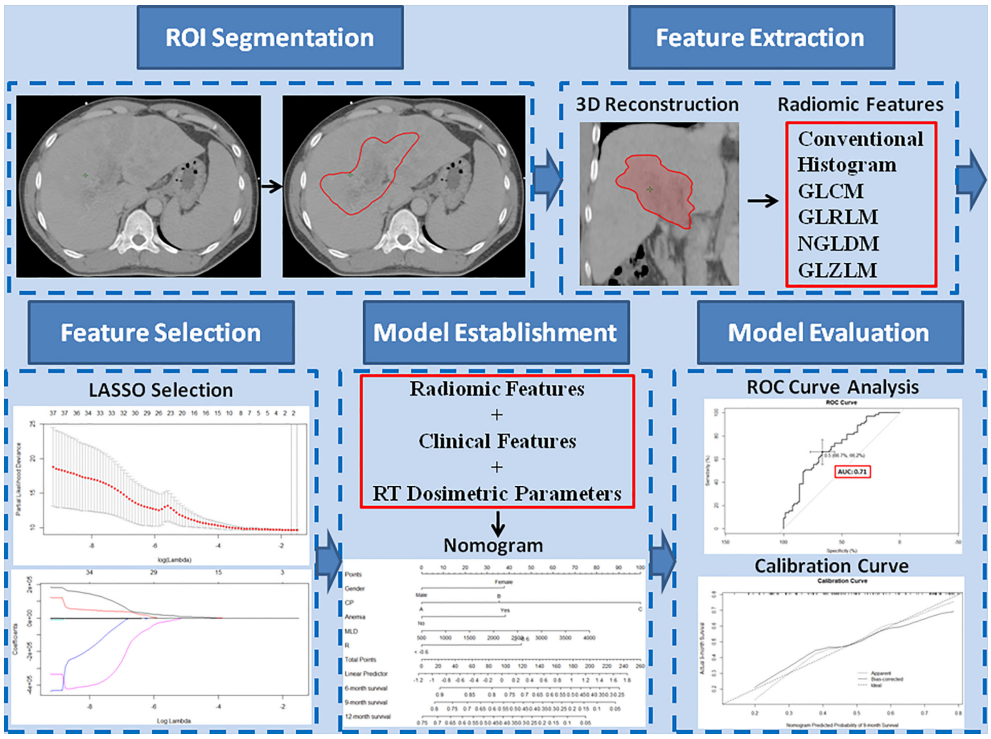


FIGURE 1
Study workflow. The region of interest (ROI) was segmented on all transverse contrast-enhanced computed tomography images by experienced radiation oncologists using the Eclipse system. After a three-dimensional reconstruction of the ROI, 48 textural features, including conventional features, histogram features, gray-level co-occurrence matrix, gray-level run-length matrix, neighborhood gray-level dependence matrix, and gray-level zone length matrix, were extracted. The extracted features were selected by least absolute shrinkage and selection operator regression. Based on the selected radiomics features, clinical features, and radiation therapy dosimetric parameters, a nomogram model was established to predict overall survival. The performance of the predictive model was evaluated with concordance index, area under the curve of the receiver operating characteristic curve, and calibration curve.

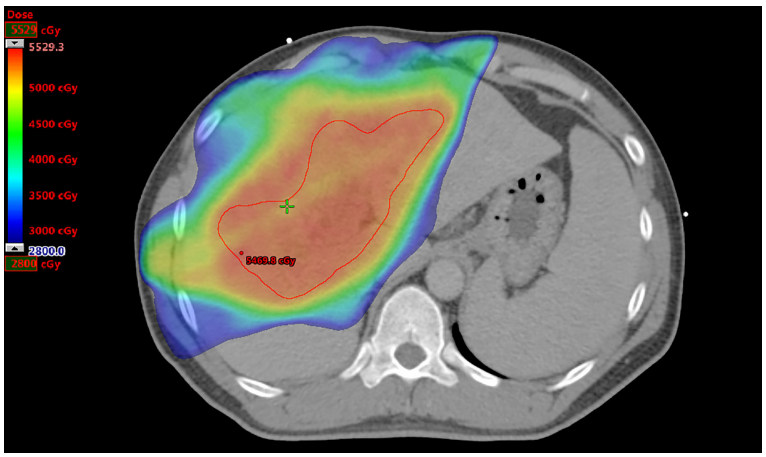


FIGURE 2
Radiation therapy plan for a patient. The prescribed dose to treat portal vein tumor thrombosis only is 50 Gy. The clinical target volume as the region of interest is contoured in red, and the volume is 280.7 cm^3. The normal liver volume is 1,769.0 cm^3, and the mean liver dose is 2,189.4 cGy.

Statistical analysis

Statistical analysis was performed using R (version 3.6.1) and SPSS version 24.0 (IBM Corporation, Armonk, NY, USA). Numerical data are presented as mean ± standard deviation. LASSO regression analysis was performed using the "glmnet" package to select the radiomics features for rad-score to predict OS. The optimal cutoff value of the rad-score was determined using X-tile software (Yale University, New Haven, CT, USA). The survival curves were plotted using the Kaplan–Meier method and assessed using the log-rank test. The Chi-square test was used to assess categorical variables, and the Mann–Whitney *U* test was used to assess continuous variables. Univariate Cox regression analysis was performed to determine the predictors of OS from rad-score, clinical features, and RT dosimetric parameters. Thereafter, multivariate Cox regression analysis was used to select prognostic factors for the establishment of predictive nomogram models. The "survival" and "rms" packages were used for survival analysis, nomogram model construction, Harrell's concordance index (C-index) calculation, and calibration curve. The C-index was a measure of goodness-of-fit for outcomes in a regression model, ranging from 0.5 to 1. A C-index value of 0.5 indicated that the predictive ability of the model was no better than a random chance, whereas C-index values of >0.7 and >0.8 indicated that the model was good and strong, respectively. A value of 1 implied that the model perfectly predicted the outcome. The "survivalROC" package was used for calculation and comparison of the area under the curve (AUC) of the receiver operating characteristic (ROC) curve for evaluation of the nomogram. The AUC ranged from 0.5 to 1. The discrimination potent of the model based on the value of AUC was as follows: 0.5, no discrimination potent; 0.7–0.8, acceptable; 0.8–0.9, excellent; and >0.9, outstanding. Differences were considered significant at $p < 0.05$.

Establishment of predictive models

Based on the results of multivariate Cox regression analysis, the nomogram models with significant clinical features, RT dosimetric parameters, and/or rad-score were constructed to predict OS in HCC patients treated with RT for PVTT. The confirmation of nomograms was subjected to a 1,000 resampling bootstrap analysis for validation. The predictive models were evaluated with C-indexes, AUC of ROC curves, and calibration curves (Figure 1).

Ethical statement

This study was approved by the Institutional Review Board of our institution (IRB number: 20MMHIS215e).

Results

Patient characteristics

A summary of the baseline characteristics of the 131 patients is presented in Table 1. The median age at diagnosis was 61 years (range: 36–87 years), and 108 (82.4%) of the patients were men. A total of 110 (84.0%) patients had hepatitis B/C virus infection, and 93 (71.0%) patients had a drinking history. In this study, 83, 44, and 4 patients had Child–Pugh classes A, B, and C, respectively. The pre-RT tumor size was 9.5 ± 5.2 cm. Before RT, 59 (45.0%) patients had anemia, and the median serum AFP

TABLE 1 Baseline characteristics of all patients.

Characteristics	*N* = 131
Age (median (range), year)	61 (36–87)
Gender (*N* (%))	
Male	108 (82.4)
Female	23 (17.6)
Hepatitis (B/C) (*N* (%))	
Yes	110 (84.0)
No	21 (16.0)
Drinking history (*N* (%))	
Yes	93 (71.0)
No	38 (29.0)
ECOG (*N* (%))	
0	36 (27.5)
1	71 (54.2)
2	24 (18.3)
Child–Pugh class (*N* (%))	
A	83 (63.4)
B	44 (33.6)
C	4 (3.0)
Tumor size (mean (SD), cm)	9.5 (5.2)
Anemia (*N* (%))	
Yes	59 (45.0)
No	72 (55.0)
AFP (median (range), ng/ml)	149.3 (1.2–515,800.0)
WBC (mean (SD), 10^3/μl)	6.3 (2.5)
PLT (mean (SD), 10^3/μl)	172.3 (105.5)
ALB (mean (SD), g/dl)	3.5 (0.6)
ALT (mean (SD), IU/L)	53.2 (42.1)
AST (mean (SD), IU/L)	86.0 (80.8)
TBIL (mean (SD), mg/dl)	1.5 (0.9)
Cr (mean (SD), mg/dl)	0.9 (0.3)
PT (mean (SD), s)	11.8 (1.0)

ECOG, Eastern Cooperative Oncology Group; SD, standard deviation; AFP, alpha-fetoprotein; WBC, white blood cell; PLT, platelet; ALB, albumin; ALT, alanine aminotransferase; AST, aspartate aminotransferase; TBIL, total bilirubin; Cr, creatinine; PT, prothrombin time.
The performance status was graded with the ECOG score, in which grade 0 indicated fully active, grade 1 indicated able to perform light work, and grade 2 indicated capable of all self-care but unable to perform any work activities.

level was 149.3 ng/ml (range: 1.2–515,800.0 ng/ml). The baseline characteristics of the training and validation cohorts are summarized in Table 2. No significant differences were found in the baseline characteristics of the two cohorts.

RT dosimetric parameters

The RT dosimetric parameters for 131 patients are presented in Table 3. A total of 25, 101, and 5 patients were treated with an RT dose of 45, 50, and 60 Gy, respectively. The RT field in 88 (67.2%) patients involved PVTT only and that in 43 (32.8%) patients involved PVTT and primary liver tumors. The median CTV was 164.6 cm^3 (range: 19.5–2,189.0 cm^3). The NLV was 1,140.7 ± 480.8 cm^3, and the MLD was 1891.3 ± 651.5 cGy. The RT dosimetric parameters for the training and validation cohorts are summarized in Table 4. No significant differences were found between the two cohorts for RT dosimetric parameters.

Treatment outcome

The treatment outcomes of the patients are presented in Table 5. The median follow-up time was 9.8 months (range, 1.6–57.9 months), and 101 (77.1%) patients underwent abdominal CT or MRI images for response assessment. Three

TABLE 2 Baseline characteristics of the training and validation cohorts.

Characteristics	Training cohort (*N* = 105)	Validation cohort (*N* = 26)	*p*
Age (median (range), year)	61 (36–87)	62 (45–84)	0.36
Gender (*N* (%))			
Male	88 (83.8)	20 (76.9)	0.41
Female	17 (16.2)	6 (23.1)	
Etiology of viral hepatitis (*N* (%))			
Hepatitis B	64 (60.9)	12 (46.2)	0.11
Hepatitis C	23 (21.9)	4 (15.4)	
Hepatitis B + C	5 (4.8)	2 (7.7)	
None	13 (12.4)	8 (30.7)	
Drinking history (*N* (%))			
Yes	76 (72.4)	17 (65.4)	0.48
No	29 (27.6)	9 (34.6)	
ECOG (*N* (%))			
0	28 (26.7)	8 (30.7)	0.92
1	59 (56.2)	12 (46.2)	
2	18 (17.1)	6 (23.1)	
Child–Pugh class (*N* (%))			
A	68 (64.8)	15 (57.7)	0.79
B	34 (32.4)	10 (38.5)	
C	3 (2.8)	1 (3.8)	
Tumor size (mean (SD), cm)	9.5 (5.1)	9.4 (5.2)	0.93
Anemia (*N* (%))			
Yes	46 (43.8)	13 (50.0)	0.57
No	59 (56.2)	13 (50.0)	
AFP (median (range), ng/ml)	149.3 (1.2–515,800.0)	136.9 (2.0–121,480.0)	0.59
WBC (mean (SD), 10^3/μl)	6.3 (2.4)	6.2 (2.7)	0.84
PLT (mean (SD), 10^3/μl)	177.4 (108.7)	151.5 (89.9)	0.28
ALB (mean (SD), g/dl)	3.5 (0.6)	3.6 (0.6)	0.29
ALT (mean (SD), IU/L)	54.0 (43.9)	50.0 (33.8)	0.51
AST (mean (SD), IU/L)	87.1 (81.2)	81.5 (63.7)	0.64
TBIL (mean (SD), mg/dl)	1.5 (0.9)	1.5 (0.9)	0.85
Cr (mean (SD), mg/dl)	0.9 (0.3)	0.9 (0.2)	0.83
PT (mean (SD), s)	11.7 (1.0)	12.1 (1.2)	0.15

ECOG, Eastern Cooperative Oncology Group; SD, standard deviation; AFP, alpha-fetoprotein; WBC, white blood cell; PLT, platelet; ALB, albumin; ALT, alanine aminotransferase; AST, aspartate aminotransferase; TBIL, total bilirubin; Cr, creatinine; PT, prothrombin time.
The performance status was graded with the ECOG score. A two-sided p-value of < 0.05 was considered statistically significant.

TABLE 3 RT dosimetric parameters of all patients.

Parameters	*N* = 131
RT dose (*N* (%))	
45 Gy	25 (19.1)
50 Gy	101 (77.1)
60 Gy	5 (3.8)
RT field (*N* (%))	
Involved PVTT	88 (67.2)
Involved PVTT + liver tumors	43 (32.8)
CTV (median (range), cm^3)	164.6 (19.5–2,189.0)
NLV (mean (SD), cm^3)	1,140.7 (480.8)
MLD (mean (SD), cGy)	1,891.3 (651.5)

RT, radiation therapy; PVTT, portal vein tumor thrombosis; CTV, clinical target volume; NLV, normal liver volume; SD, standard deviation; MLD, mean liver dose.

cohort (Figure 3). The rad-score formula was GLRLM_HGRE × −2.973897e−05 + GLRLM_SRHGE × −2.504878e−05.

Rad-score and correlation with OS

The optimal cutoff value of the rad-score, as determined by X-tile software, was −0.6. The patients were divided into the high- (≧−0.6) and low-risk (<−0.6) groups based on the cutoff value of the rad-score. The median OS rates in the high- and low-risk groups were 7.4 months (95% CI, 6.5–10.7) and 12.4 months (95% CI, 10.0–16.8), respectively (p = 0.007). Considering the training cohort, the median OS rates in the high- and low-risk groups were 7.5 months (95% CI, 6.5–11.2) and 11.8 months (95% CI, 9.6–16.8), respectively (p = 0.038). In the validation cohort, the median OS rates in the high- and low-risk groups were 6.8 months (95% CI, 4.3–NA) and 12.6 months

TABLE 4 RT dosimetric parameters of the training and validation cohorts.

Parameters	Training cohort (*N* = 105)	Validation cohort (*N* = 26)	*p*
RT dose (*N* (%))			
45 Gy	21 (20.0)	4 (15.4)	0.87
50 Gy	80 (76.2)	21 (80.8)	
60 Gy	4 (3.8)	1 (3.8)	
RT field (*N* (%))			
Involved PVTT	72 (68.6)	16 (61.5)	0.49
PVTT + liver tumors	33 (31.4)	10 (38.5)	
CTV (median (range), cm^3)	175.6 (27.3–2,189.0)	154.2 (19.5–1,958.0)	0.63
NLV (mean (SD), cm^3)	1137.4 (498.2)	1,154.0 (411.2)	0.88
MLD (mean (SD), cGy)	1,895.5 (661.5)	1,874.5 (621.5)	0.88

RT, radiation therapy; PVTT, portal vein tumor thrombosis; CTV, clinical target volume; NLV, normal liver volume; SD, standard deviation; MLD, mean liver dose.
A two-sided p-value of < 0.05 was considered statistically significant.

(2.3%) patients were alive at the time of the current analysis. Sixteen (15.8%) patients had patent portal veins after RT. The median OS was 9.8 months (95% CI, 8.0–11.6 months), and the median progression-free survival (PFS) was 5.6 months (95% CI, 4.8–6.4 months). Distant metastases were found in 22 (16.8%) patients. The treatment outcomes of the training and validation cohorts are summarized in Table 6. No significant differences in treatment outcomes were found between the two cohorts.

Radiomics feature extraction and development of the rad-score

A total of 48 radiomics features were extracted from the imaging data of all patients. Two features were selected by LASSO Cox regression analysis to predict the OS in the training

(95% CI, 10.7–NA), respectively (p = 0.033). The median OS rates were significantly lower in the high-risk groups than in the low-risk groups in both the training and validation cohorts (Figure 4).

Extraction of significant features

Univariate and multivariate Cox regression analyses were performed to determine the predictors of OS from rad-score, clinical features, and RT dosimetric parameters. Univariate analysis revealed seven predictors, namely gender, Child–Pugh class, anemia status, rad-score, MLD, tumor size, and AST, for OS prediction. Gender, Child–Pugh class, anemia status, rad-score, and MLD were found to be independent predictors in multivariate analysis (Table 7).

TABLE 5 Treatment outcomes of all patients.

Outcomes	*N* = 131
Patency (*N* (%))	
Yes	16 (15.8)
No	85 (84.2)
OS (median (95% CI), m)	9.8 (8.0–11.6)
PFS (median (95% CI), m)	5.6 (4.8–6.4)
DM (*N* (%))	
Yes	22 (16.8)
No	109 (83.2)

OS, overall survival; CI, confidence interval; PFS, progression-free survival; DM, distant metastasis.

A clinical nomogram with selected clinical features and RT dosimetric parameters was developed for OS prediction (Figure 5).

Performances of different predictive nomograms and significant features

C-indexes were used to evaluate the discrimination power of significant features, clinical nomogram, and radiomics-based nomogram. The C-index profiles are presented in Table 8. In this study, the radiomics-based nomogram showed the best discrimination power, which was examined by internal validation. ROC analyses for 9-month survival, the AUCs for

TABLE 6 Treatment outcomes of the training and validation cohorts.

Outcomes	Training cohort (*N* = 105)	Validation cohort (*N* = 26)	*p*
Patency (*N* (%))			
Yes	13 (15.5)	3 (17.6)	0.82
No	71 (84.5)	14 (82.4)	
OS (median (95% CI), m)	9.8 (7.9–11.7)	10.1 (6.7–13.7)	0.87
PFS (median (95% CI), m)	5.2 (4.4–6.1)	5.9 (4.1–7.7)	0.47
DM (*N* (%))			
Yes	19 (18.1)	3 (11.5)	0.64
No	86 (81.9)	23 (88.5)	

OS, overall survival; CI, confidence interval; PFS, progression-free survival; DM, distant metastasis.
A two-sided p-value of <0.05 was considered statistically significant.

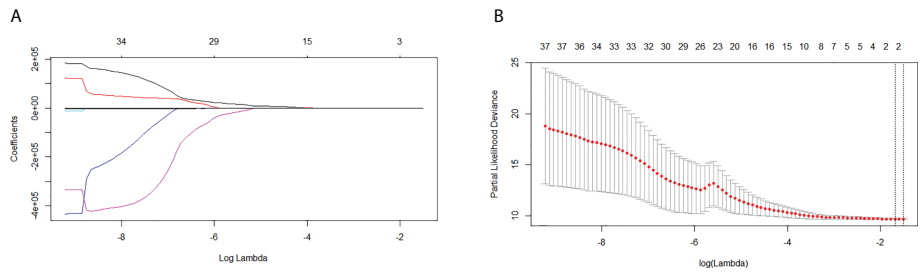


FIGURE 3
Least absolute shrinkage and selection operator (LASSO) regression analysis for the selection of significant radiomics features from the 48 textural features. **(A)** Coefficient profile of the LASSO model. **(B)** Optimal tuning parameter (lambda) selection using 10-fold cross-validation with minimum criteria. Two significant radiomics features were extracted.

Establishment of the predictive model

Based on the result of multivariate Cox regression analysis, a radiomics-based nomogram with significant clinical features, RT dosimetric parameters, and rad-score was developed to predict OS. the radiomics-based nomogram and clinical nomogram were 0.71 (95% CI, 0.63–0.79) and 0.61 (95% CI, 0.51–0.71), respectively (Figure 6). The calibration curves of the radiomics-based nomogram and clinical nomogram are presented in Figure 7. The radiomics-based nomogram exhibited better predictive

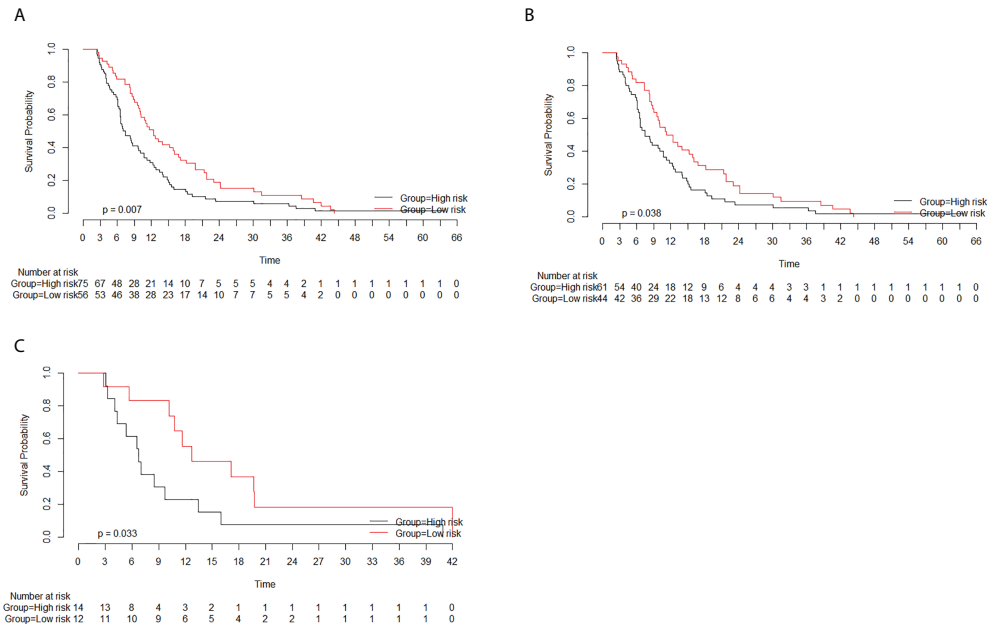


FIGURE 4
Survival curves of the high- and low-risk groups based on the radiomics score (rad-score) classification. The rad-scores in the high- and low-risk groups were more than −0.6 and less than −0.6, respectively. **(A)** Considering all patients, the median overall survival (OS) rates in the high- and low-risk groups were 7.4 months (95% CI, 6.5–10.7) and 12.4 months (95% CI, 10.0–16.8), respectively (p = 0.007). **(B)** Considering the training cohort, the median OS rates in the high- and low-risk groups were 7.5 months (95% CI, 6.5–11.2) and 11.8 months (95% CI, 9.6–16.8), respectively (p = 0.038). **(C)** Considering the validation cohort, the median OS rates in the high- and low-risk groups were 6.8 months (95% CI, 4.3–NA) and 12.6 months (95% CI, 10.7–NA), respectively (p = 0.033).

TABLE 7 Univariate and multivariate analyses for predictors of OS.

Predictors	Univariate analysis			Multivariate analysis		
	HR	95% CI	*p*	HR	95% CI	*p*
Gender						
Male	1			1		
Female	1.704	1.062–2.735	0.025	1.886	1.013–3.512	0.045
Child–Pugh class						
A	1			1		
B and C	1.515	1.042–2.201	0.028	1.672	1.053–2.655	0.029
Anemia						
No	1			1		
Yes	1.617	1.123–2.328	0.009	1.690	1.043–2.739	0.033
Rad-score						
<−0.6	1			1		
≧−0.6	1.635	1.137–2.351	0.008	1.540	1.007–2.355	0.047
MLD (cGy)	1.000	1.000–1.001	0.050	1.001	1.000–1.001	0.002
Tumor size (cm)	1.054	1.016–1.093	0.005	1.040	0.990–1.092	0.119
AST (U/L)	1.003	1.001–1.004	0.001	1.002	0.998–1.006	0.292

OS, overall survival; HR, hazard ratio; CI, confidence interval; rad-score, radiomics score; MLD, mean liver dose; AST, aspartate aminotransferase.
A two-sided p-value of <0.05 was considered statistically significant.

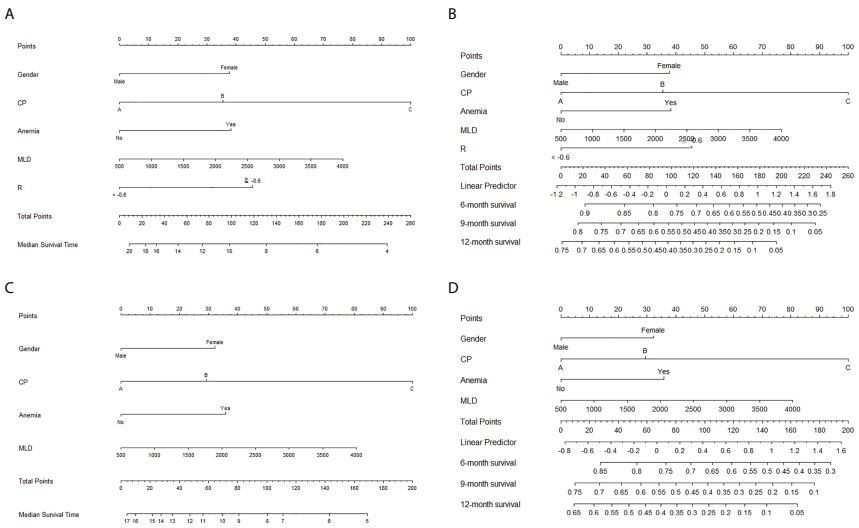


FIGURE 5
Nomograms for the prediction of overall survival. Nomograms with radiomics score, significant clinical features, and radiation therapy (RT) dosimetric parameters for the prediction of **(A)** median survival time and **(B)** 6-, 9-, and 12-month survival rates. Nomograms with selected clinical features and RT dosimetric parameters for the prediction of **(C)** median survival time and **(D)** 6-, 9-, and 12-month survival rates.

TABLE 8 C-indexes of significant features, clinical nomogram, and radiomics-based nomogram.

Variables	Training cohort		Validation cohort		All patients	
	C-index	95% CI	C-index	95% CI	C-index	95% CI
Gender	0.54	0.50–0.58	0.53	0.41–0.65	0.54	0.50–0.58
Child–Pugh class	0.54	0.48–0.60	0.65	0.54–0.76	0.56	0.51–0.61
Anemia	0.56	0.50–0.62	0.58	0.46–0.70	0.56	0.51–0.61
MLD	0.51	0.44–0.58	0.59	0.46–0.72	0.53	0.47–0.59
Rad-score	0.57	0.52–0.62	0.64	0.53–0.75	0.58	0.53–0.63
Clinical nomogram	0.60	0.53–0.67	0.72	0.58–0.86	0.61	0.55–0.67
Radiomics-based nomogram	0.72	0.65–0.79	0.82	0.69–0.95	0.73	0.67–0.79

C-index, concordance index; CI, confidence interval; MLD, mean liver dose; rad-score, radiomics score.

accuracy than the clinical nomogram for the prediction of 9-month survival.

Discussion

In this study, we intended to develop a radiomics-based nomogram using pre-RT CT data. Univariate and multivariate analyses revealed that the rad-score significantly influenced OS. The performance of the radiomics-based nomogram was better than the clinical nomogram, and the predictive accuracy of each significant feature in the C-index and ROC analysis was examined by the calibration curve.

This study was conducted in the Department of Radiation Oncology in a medical center. All HCC patients were treated with the same RT protocol to ensure the standardization of the treatment and CT quality. CT was performed according to the American Association of Physicists in Medicine (AAPM) and American College of Radiology (ACR) guidelines (AAPM report #74 and #96 and ACR CT QC manual) and standard quality assurance measures.

Few studies have reported the application of radiomics and clinical features to predict treatment outcomes and prognosis in different types of cancers treated with RT. Hou et al. established an integrated model that combined posttreatment CT radiomics features and clinical features for response and OS prediction in esophageal cancer patients undergoing neoadjuvant chemoradiotherapy (48). Wu et al. developed a nomogram using radiomics and clinical features to predict OS in HCC patients treated with stereotactic body radiotherapy (SBRT) for

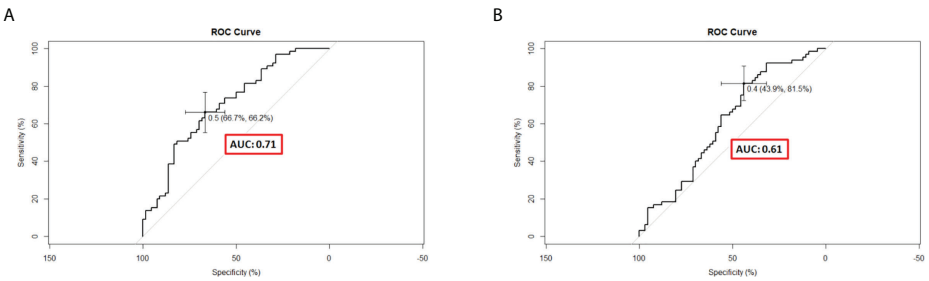


FIGURE 6
Receiver operating characteristic curves of different predictive nomograms for 9-month survival. **(A)** The area under the curve (AUC) was 0.71 (95% CI, 0.62–0.79) in the radiomics-based nomogram. **(B)** The AUC was 0.61 (95% CI, 0.51–0.71) in the clinical nomogram.

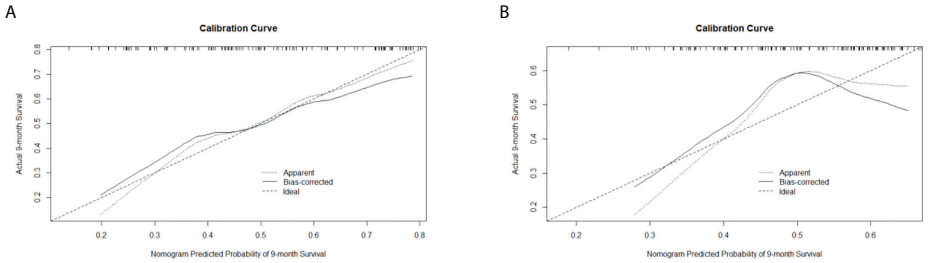


FIGURE 7
Calibration curves of **(A)** the radiomics-based nomogram and **(B)** clinical nomogram for the prediction of 9-month survival.

PVTT (49). Parr et al. indicated that a radiomics-based predictive model combined with clinical features is better than an analysis of clinical features alone for predicting OS in pancreatic cancer patients treated with SBRT (50). Thus, radiomics combined with clinical features may have a better performance than analysis with clinical features alone for treatment response and OS prediction.

This study has several limitations. First, MRI is the preferred imaging modality for the evaluation of liver lesions. The technique of MRI-guided RT with MRI simulation and planning is rapidly developing (51). However, contrast-enhanced CT is still the main imaging methodology for diagnosis, staging, and RT planning for HCC, with acceptable sensitivity and high specificity. It is noninvasive, well-developed in current clinical practice, and not time-consuming or labor-consuming. Second, sorafenib has been used as the standard systemic treatment of advanced HCC during the investigation period of our study population. Currently, different agents such as lenvatinib, checkpoint inhibitors, and antivascular endothelial growth factor receptor antibodies have demonstrated efficacy in the treatment of advanced HCC. The effects of different systemic treatments should be examined in the future. Third, this study included 30 patients without CT or MRI follow-up data. Fourth, the small number of patients from a single institute could not draw a firm conclusion for application in other hospitals. In this retrospective review study, the standardization of CT simulation protocol, RT dose to PTV, and follow-up schedule lasted for 12 years, which might provide an informative database for analyzing radiomics and clinical outcomes. The current results may provide proof-of-concept information and practical procedures for other hospitals trying to apply radiomics in each institute. A prospective large-scale and multicenter study is required. Finally, the data in this study are derived from one hospital. Although internal validation was conducted for verification, further multicenter analysis is required for external validation.

Conclusion

Radiomics features combined with clinical features and dosimetric parameters have better performance than each

significant feature and clinical nomogram. This study recommends the development of a predictive model with significant clinical features, radiomics features, and dosimetric parameters. The multicenter analysis is warranted after the standardization of treatment protocol, radiology imaging, and radiomics data in all hospitals for external validation to ensure the accuracy of the universal predictive model.

Data availability statement

The original contributions presented in the study are included in the article/supplementary material. Further inquiries can be directed to the corresponding authors.

Ethics statement

The studies involving human participants were reviewed and approved by MacKay Memorial Hospital. The patients/participants provided their written informed consent to participate in this study.

Author contributions

Conceptualization: Y-MH and Y-JC. Methodology: Y-MH, H-CT, S-MH, and Y-JC. Software: Y-MH Validation: Y-MH, S-MH, and Y-JC. Formal analysis: Y-MH. Investigation: Y-MH. Resources: Y-MH, T-EW, M-JC, C-CL, C-WC, and Y-JC. Data curation: Y-MH. Writing—original draft preparation: Y-MH. Writing—review and editing: Y-MH, S-MH, and Y-JC. Visualization: Y-MH. Supervision: Y-JC. Project administration: Y-JC. Funding acquisition: Y-JC. All authors have read and agreed to the published version of the manuscript.

Funding

This research was funded by MacKay Memorial Hospital (Grant numbers MMH-E-109-13 and MMH-E-110-13) and the Ministry of Science and Technology of Taiwan (Grant number MOST 109-2314-B-195-003-MY3).

Acknowledgments

We would like to thank Editage for providing their editing services.

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. Sung H, Ferlay J, Siegel RL, Laversanne M, Soerjomataram I, Jemal A, et al. Global cancer statistics 2020: globocan estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin* (2021) 71(3):209–49. doi: 10.3322/caac.21660

2. Villanueva A, Longo DL. Hepatocellular carcinoma. *N Engl J Med* (2019) 380 (15):1450–62. doi: 10.1056/NEJMra1713263

3. Jemal A, Ward EM, Johnson CJ, Cronin KA, Ma J, Ryerson AB, et al. Annual report to the nation on the status of cancer, 1975–2014, featuring survival. *J Natl Cancer Inst* (2017) 109(9):djx030. doi: 10.1093/jnci/djx030

4. Llovet JM, Kelley RK, Villanueva A, Singal AG, Pikarsky E, Roayaie S, et al. Hepatocellular carcinoma. *Nat Rev Dis Primers* (2021) 7(1):6. doi: 10.1038/s41572-020-00240-3

5. Vilarinho S, Calvisi DF. New advances in precision medicine for hepatocellular carcinoma recurrence prediction and treatment. *Hepatology* (2014) 60(6):1812–4. doi: 10.1002/hep.27311

6. Costentin CE, Ferrone CR, Arellano RS, Ganguli S, Hong TS, Zhu AX. Hepatocellular carcinoma with macrovascular invasion: defining the optimal treatment strategy. *Liver Cancer* (2017) 6(4):360–74. doi: 10.1159/000481315

7. Zhang ZM, Lai ECH, Zhang C, Yu HW, Liu Z, Wan BJ, et al. The strategies for treating primary hepatocellular carcinoma with portal vein tumor thrombus. *Int J Surg* (2015) 20:8–16. doi: 10.1016/j.ijsu.2015.05.009

8. Addario L, Tritto G, Cavaglià E, Amodio F, Giannelli E, Di Costanzo GG. Preserved liver function, portal thrombosis and absence of oesophageal varices are risk factors for metastasis of hepatocellular carcinoma. *Dig Liver Dis* (2011) 43 (4):319–24. doi: 10.1016/j.dld.2010.09.003

9. Cabibbo G, Enea M, Attanasio M, Bruix J, Craxì A, Cammà C. A meta-analysis of survival rates of untreated patients in randomized clinical trials of hepatocellular carcinoma. *Hepatology* (2010) 51(4):1274–83. doi: 10.1002/hep.23485

10. Tiao G, Mähringer-Kunz A, Steinle V, Kloeckner R, Schotten S, Hahn F, et al. The impact of portal vein tumor thrombosis on survival in patients with hepatocellular carcinoma treated with different therapies: a cohort study. *PloS One* (2021) 16(5):e0249426. doi: 10.1371/journal.pone.0249426

11. Faivre S, Rimassa L, Finn RS. Molecular therapies for HCC: looking outside the box. *J Hepatol* (2020) 72(2):342–52. doi: 10.1016/j.jhep.2019.09.010

12. Zhang ZY, Dong KS, Zhang EL, Zhang LW, Chen XP, Dong HH. Resection might be a meaningful choice for hepatocellular carcinoma with portal vein

thrombosis. *Med (Baltimore)* (2019) 98(50):e18362. doi: 10.1097/md.0000000000018362

13. Woloschak GE, Somma F, Stoia V, Serra N, D'Angelo R, Gatta G, et al. Yttrium-90 trans-arterial radioembolization in advanced-stage HCC: the impact of portal vein thrombosis on survival. *PloS One* (2019) 14(5):e0216935. doi: 10.1371/journal.pone.0216935

14. Holliday EB, Tao R, Brownlee Z, Das P, Krishnan S, Taniguchi C, et al. Definitive radiation therapy for hepatocellular carcinoma with portal vein tumor thrombus. *Clin Transl Radiat Oncol* (2017) 4:39–45. doi: 10.1016/j.ctro.2017.04.003

15. Rim CH, Kim CY, Yang DS, Yoon WS. Comparison of radiation therapy modalities for hepatocellular carcinoma with portal vein thrombosis: a meta-analysis and systematic review. *Radiother Oncol* (2018) 129(1):112–22. doi: 10.1016/j.radonc.2017.11.013

16. Yu JI, Park HC. Radiotherapy as valid modality for hepatocellular carcinoma with portal vein tumor thrombosis. *World J Gastroenterol* (2016) 22 (30):6851–63. doi: 10.3748/wjg.v22.i30.6851

17. Yoon SM, Ryoo BY, Lee SJ, Kim JH, Shin JH, An JH, et al. Efficacy and safety of transarterial chemoembolization plus external beam radiotherapy vs sorafenib in hepatocellular carcinoma with macroscopic vascular invasion. *JAMA Oncol* (2018) 4(5):661–9. doi: 10.1001/jamaoncol.2017.5847

18. Hong TS, Bosch WR, Krishnan S, Kim TK, Mamon HJ, Shyn P, et al. Interobserver variability in target definition for hepatocellular carcinoma with and without portal vein thrombus: radiation therapy oncology group consensus guidelines. *Int J Radiat Oncol Biol Phys* (2014) 89(4):804–13. doi: 10.1016/j.ijrobp.2014.03.041

19. Lin CS, Jen YM, Chiu SY, Hwang JM, Chao HL, Lin HY, et al. Treatment of portal vein tumor thrombosis of hepatoma patients with either stereotactic radiotherapy or three-dimensional conformal radiotherapy. *Jpn J Clin Oncol* (2006) 36(4):212–7. doi: 10.1093/jjco/hyl006

20. Wang G, Zhu S, Li X. Comparison of values of CT and MRI imaging in the diagnosis of hepatocellular carcinoma and analysis of prognostic factors. *Oncol Lett* (2019) 17(1):1184–8. doi: 10.3892/ol.2018.9690

21. Yu NC, Chaudhari V, Raman SS, Lassman C, Tong MJ, Busuttil RW, et al. CT and MRI improve detection of hepatocellular carcinoma, compared with ultrasound alone, in patients with cirrhosis. *Clin Gastroenterol Hepatol* (2011) 9 (2):161–7. doi: 10.1016/j.cgh.2010.09.017

22. Mayerhoefer ME, Riedl CC, Kumar A, Dogan A, Gibbs P, Weber M, et al. [18F]FDG-PET/CT radiomics for prediction of bone marrow involvement in mantle cell lymphoma: a retrospective study in 97 patients. *Cancers (Basel)* (2020) 12(5):1138. doi: 10.3390/cancers12051138

23. Yang J, Guo X, Ou X, Zhang W, Ma X. Discrimination of pancreatic serous cystadenomas from mucinous cystadenomas with CT textural features: based on machine learning. *Front Oncol* (2019) 9:494. doi: 10.3389/fonc.2019.00494

24. Wu J, Liu A, Cui J, Chen A, Song Q, Xie L. Radiomics-based classification of hepatocellular carcinoma and hepatic haemangioma on precontrast magnetic resonance images. *BMC Med Imaging* (2019) 19(1):23. doi: 10.1186/s12880-019-0321-9

25. Zhang T, Zhang Y, Liu X, Xu H, Chen C, Zhou X, et al. Application of radiomics analysis based on CT combined with machine learning in diagnostic of pancreatic neuroendocrine tumors patient's pathological grades. *Front Oncol* (2021) 10:521831. doi: 10.3389/fonc.2020.521831

26. Liu Z, Zhu G, Jiang X, Zhao Y, Zeng H, Jing J, et al. Survival prediction in gallbladder cancer using CT based machine learning. *Front Oncol* (2020) 10:604288. doi: 10.3389/fonc.2020.604288

27. Peng H, Dong D, Fang MJ, Li L, Tang LL, Chen L, et al. Prognostic value of deep learning PET/CT-based radiomics: potential role for future individual induction chemotherapy in advanced nasopharyngeal carcinoma. *Clin Cancer Res* (2019) 25(14):4271–9. doi: 10.1158/1078-0432.ccr-18-3065

28. Zhao Y, Yang J, Luo M, Yang Y, Guo X, Zhang T, et al. Contrast-enhanced CT-based textural parameters as potential prognostic factors of survival for colorectal cancer patients receiving targeted therapy. *Mol Imaging Biol* (2020) 23 (3):427–35. doi: 10.1007/s11307-020-01552-2

29. Bourbonne V, Fournier G, Vallières M, Lucia F, Doucet L, Tissot V, et al. External validation of an MRI-derived radiomics model to predict biochemical recurrence after surgery for high-risk prostate cancer. *Cancers (Basel)* (2020) 12 (4):814. doi: 10.3390/cancers12040814

30. Botta F, Raimondi S, Rinaldi L, Bellerba F, Corso F, Bagnardi V, et al. Association of a CT-based clinical and radiomics score of non-small cell lung cancer (NSCLC) with lymph node status and overall survival. *Cancers (Basel)* (2020) 12(6):1432. doi: 10.3390/cancers12061432

31. Kocher M, Ruge MI, Galldiks N, Lohmann P. Applications of radiomics and machine learning for radiotherapy of malignant brain tumors. *Strahlenther Onkol* (2020) 196(10):856–67. doi: 10.1007/s00066-020-01626-8

32. Dissaux G, Visvikis D, Da-Ano R, Pradier O, Chajon E, Barillot I, et al. Pretreatment 18F-FDG PET/CT radiomics predict local recurrence in patients treated with stereotactic body radiotherapy for early-stage non-small cell lung cancer: a multicentric study. *J Nucl Med* (2020) 61(6):814–20. doi: 10.2967/jnumed.119.228106

33. Yan M, Wang W. A radiomics model of predicting tumor volume change of patients with stage III non-small cell lung cancer after radiotherapy. *Sci Prog* (2021) 104(1):3685042199729. doi: 10.1177/0036850421997295

34. Sun R, Limkin EJ, Vakalopoulou M, Dercle L, Champiat S, Han SR, et al. A radiomics approach to assess tumour-infiltrating CD8 cells and response to anti-PD-1 or anti-PD-L1 immunotherapy: an imaging biomarker, retrospective multicohort study. *Lancet Oncol* (2018) 19(9):1180–91. doi: 10.1016/s1470-2045(18)30413-3

35. Nardone V, Tini P, Pastina P, Botta C, Reginelli A, Carbone S, et al. Radiomics predicts survival of patients with advanced non–small cell lung cancer undergoing PD–1 blockade using nivolumab. *Oncol Lett* (2020) 19(2):1559–66. doi: 10.3892/ol.2019.11220

36. Nardone V, Tini P, Nioche C, Mazzei MA, Carfagno T, Battaglia G, et al. Texture analysis as a predictor of radiation-induced xerostomia in head and neck patients undergoing IMRT. *Radiol Med* (2018) 123(6):415–23. doi: 10.1007/s11547-017-0850-7

37. Ekert K, Hinterleitner C, Baumgartner K, Fritz J, Horger M. Extended texture analysis of non-enhanced whole-body MRI image data for response assessment in multiple myeloma patients undergoing systemic therapy. *Cancers (Basel)* (2020) 12(3):761. doi: 10.3390/cancers12030761

38. Yang B, Ji H, Zhong J, Ma L, Zhong J, Dong H, et al. Value of 18F-FDG PET/CT-based radiomics nomogram to predict survival outcomes and guide personalized targeted therapy in lung adenocarcinoma with EGFR mutations. *Front Oncol* (2020) 10:567160. doi: 10.3389/fonc.2020.567160

39. Tran WT, Jerzak K, Lu FI, Klein J, Tabbarah S, Lagree A, et al. Personalized breast cancer treatments using artificial intelligence in radiomics and pathomics. *J Med Imaging Radiat Sci* (2019) 50(4):S32–41. doi: 10.1016/j.jmir.2019.07.010

40. He M, Zhang P, Ma X, He B, Fang C, Jia F. Radiomic feature-based predictive model for microvascular invasion in patients with hepatocellular carcinoma. *Front Oncol* (2020) 10:574228. doi: 10.3389/fonc.2020.574228

41. Zheng BH, Liu LZ, Zhang ZZ, Shi JY, Dong LQ, Tian LY, et al. Radiomics score: a potential prognostic imaging feature for postoperative survival of solitary HCC patients. *BMC Cancer* (2018) 18(1):1148. doi: 10.1186/s12885-018-5024-z

42. Zhou Y, Ding J, Qin Z, Wang Y, Zhang J, Jia K, et al. Predicting the survival rate of patients with hepatocellular carcinoma after thermal ablation by nomograms. *Ann Transl Med* (2020) 8(18):1159. doi: 10.21037/atm-20-6116

43. Song W, Yu X, Guo D, Liu H, Tang Z, Liu X, et al. MRI-Based radiomics: associations with the recurrence-free survival of patients with hepatocellular carcinoma treated with conventional transcatheter arterial chemoembolization. *J Magn Reson Imaging* (2019) 52(2):461–73. doi: 10.1002/jmri.26977

44. Wang XH, Long LH, Cui Y, Jia AY, Zhu XG, Wang HZ, et al. MRI-Based radiomics model for preoperative prediction of 5-year survival in patients with hepatocellular carcinoma. *Br J Cancer* (2020) 122(7):978–85. doi: 10.1038/s41416-019-0706-0

45. Meng XP, Wang YC, Ju S, Lu CQ, Zhong BY, Ni CF, et al. Radiomics analysis on multiphase contrast-enhanced CT: a survival prediction tool in patients with hepatocellular carcinoma undergoing transarterial chemoembolization. *Front Oncol* (2020) 10:1196. doi: 10.3389/fonc.2020.01196

46. Cozzi L, Dinapoli N, Fogliata A, Hsu WC, Reggiori G, Lobefalo F, et al. Radiomics based analysis to predict local control and survival in hepatocellular carcinoma patients treated with volumetric modulated arc therapy. *BMC Cancer* (2017) 17(1):829. doi: 10.1186/s12885-017-3847-7

47. Nioche C, Orlhac F, Boughdad S, Reuzé S, Goya-Outi J, Robert C, et al. LIFEx: a freeware for radiomic feature calculation in multimodality imaging to accelerate advances in the characterization of tumor heterogeneity. *Cancer Res* (2018) 78(16):4786–9. doi: 10.1158/0008-5472.can-18-0125

48. Hou TC, Huang WC, Tai HC, Chen YJ. Integrated radiomic model for predicting the prognosis of esophageal squamous cell carcinoma patients undergoing neoadjuvant chemoradiation. *Ther Radiol Oncol* (2019) 3:28. doi: 10.21037/tro.2019.07.03

49. Wu K, Shui Y, Sun W, Lin S, Pang H. Utility of radiomics for predicting patient survival in hepatocellular carcinoma with portal vein tumor thrombosis treated with stereotactic body radiotherapy. *Front Oncol* (2020) 10:569435. doi: 10.3389/fonc.2020.569435

50. Parr E, Du Q, Zhang C, Lin C, Kamal A, McAlister J, et al. Radiomics-based outcome prediction for pancreatic cancer following stereotactic body radiotherapy. *Cancers (Basel)* (2020) 12(4):1051. doi: 10.3390/cancers12041051

51. Witt JS, Rosenberg SA, Bassetti MF. MRI-Guided adaptive radiotherapy for liver tumours: visualising the future. *Lancet Oncol* (2020) 21(2):e74–82. doi: 10.1016/s1470-2045(20)30034-6

frontiers | Frontiers in Oncology

TYPE Original Research
PUBLISHED 21 September 2022
DOI 10.3389/fonc.2022.1000471

OPEN ACCESS

EDITED BY
Xue Meng,
Shandong Cancer Hospital, China

REVIEWED BY
Fada Guan,
Yale University, United States
Qilei Chen,
University of Massachusetts Lowell, United States

*CORRESPONDENCE
Zhiyong Liu
zhiyongliuicu@csu.edu.cn

SPECIALTY SECTION
This article was submitted to Cancer Imaging and Image-directed Interventions, a section of the journal Frontiers in Oncology

RECEIVED 22 July 2022
ACCEPTED 16 August 2022
PUBLISHED 21 September 2022

CITATION
Huang Y, Huang S and Liu Z (2022) Multi-task learning-based feature selection and classification models for glioblastoma and solitary brain metastases. *Front. Oncol.* 12:1000471. doi: 10.3389/fonc.2022.1000471

Multi-task learning-based feature selection and classification models for glioblastoma and solitary brain metastases

Ya Huang[1], Shan Huang[1] and Zhiyong Liu[2,3]*

[1]Key Laboratory of Computing and Stochastic Mathematics, School of Mathematics and Statistics, Hunan Normal University, Changsha, China, [2]Department of Intensive Care, Xiangya Hospital, Central South University, Changsha, China, [3]National Clinical Research Center for Geriatric Disorders, Xiangya Hospital, Central South University, Changsha, China

Purpose: To investigate the diagnostic performance of feature selection *via* a multi-task learning model in distinguishing primary glioblastoma from solitary brain metastases.

Method: The study involved 187 patients diagnosed at Xiangya Hospital, Yunnan Provincial Cancer Hospital, and Southern Cancer Hospital between January 2010 and December 2018. Radiomic features were extracted from conventional magnetic resonance imaging including T1-weighted, T2-weighted, and contrast-enhanced T1-weighted sequences. We proposed a new multi-task learning model using these three sequences as three tasks. Multi-series fusion was performed to complement the information from different dimensions in order to enhance model robustness. Logical loss was used in the model as the data-fitting item, and the feature weights were expressed in the logical loss space as the sum of shared weights and private weights to select the common features of each task and the characteristics having an essential impact on a single task. A diagnostic model was constructed as a feature selection method as well as a classification method. We calculated accuracy, recall, precision, and area under the curve (AUC) and compared the performance of our new multi-task model with traditional diagnostic model performance.

Results: A diagnostic model combining the support vector machine algorithm as a classification algorithm and our model as a feature selection method had an average AUC of 0.993 in the training set, with AUC, accuracy, precision, and recall rates respectively of 0.992, 0.920, 0.969, and 0.871 in the test set. The diagnostic model built on our multi-task model alone, in the training set, had an average AUC of 0.987, and in the test set, the AUC, accuracy, precision, and recall rates were 0.984, 0.895, 0.954, and 0.838.

Conclusion: It is feasible to implement the multi-task learning model developed in our study using logistic regression to differentiate between glioblastoma and solitary brain metastases.

KEYWORDS

solitary brain metastases, glioblastoma, multi-task learning, feature selection, classification, logistic regression

1 Introduction

Brain tumors, also known as intracranial tumors, are a growth or mass of abnormal cells or tissue in the brain (1). Brain tumors are generally subdivided into two main types: primary or secondary (metastatic) (2). Glioblastoma (GBM) is a typical malignant primary brain tumor that affects an average of 3 out of 100,000 people (3). Solitary brain metastases (SBM) are secondary malignant brain tumors, which are more common than GBM. The incidence rate of SBM is approximately 7 to 14 per 100,000 people (3, 4). As the standard treatment course for GBM is aggressive trimodality therapy compared to surgery or radiosurgery for SBM, it is of great clinical importance to accurately and rapidly distinguish between these two types of brain malignancies as rapidly as possible. As the main diagnostic tool for brain tumors (5), magnetic resonance imaging (MRI) methods create clear and detailed three-dimensional images of brain and tumor anatomy. However, for patients with SBM and GBM, their MR images both show ring enhancement, intra-tumor necrosis, and per femoral T2 high signal (6, 7), which poses a challenge for the accurate differentiation between GBM and SBM.

Radiomics (8–12), the application of advanced image feature analysis algorithms, can be used to capture intra-tumoral heterogeneity in a non-invasive manner. Numerous studies have applied radiomics to tumor classification. Austin et al. used a filtered histogram texture analysis-based imaging historic approach to identify high-grade and low-grade gliomas (13), where the AUC on the test set reached 0.90. However, the data of their study were highly unbalance in the number of high-grade and low-grade gliomas. Among the three feature selection methods, packing, filtering, and embedding, the embedding method can obtain a higher computational efficiency and classification performance than the filtering (14, 15) and packing (16) methods (17). Therefore, the embedding method has received increasing attention recently. Qian et al. used the feature selection method of least absolute shrinkage and selection operator (Lasso) combined with a support vector machine(SVM) classifier, to obtain an AUC of 0.90 in their test set (18). Cho et al. used a machine learning approach to classify gliomas based on radiomics (19), which ultimately selected five significant features with an average AUC of 0.903 on their test set. Artzi et al. found that the SVM approach had the best results for classifying between GBM and SBM subtypes (20), with an AUC of 0.96.

Liu et al. combined handcrafted radiomics and deep learning-based radiomics and used a random forest algorithm for feature selection and classification (21), the AUC reached 0.97 for single contrast-enhanced T1-weighted(T1C) MRI sequence. Because tumor sites behave differently under different sequences of MRI, patients usually have multiple series of imaging data acquired to accurately determine the tumor location, size, and additional information during the treatment. Different sequences provide different information, and multiple sequence fusion can complement information from different dimensions, thus enhancing the robustness of a model. As such, we introduced a multi-task learning model (22, 23) to fuse T1, T1C, and T2 sequence information to develop a robust prediction model to aid in clinical diagnosis.

Nowadays, most studies on multi-task-based feature selection focus on different canonical terms. The features are selected by constraints of different paradigms, such as the commonly used $\ell_{1,1}$ (24), $\ell_{2,1}$ (25) paradigms, etc. (17, 26). In the present report, we have improved the data-fitting term in the multi-task learning model so that the model can be used not only for feature selection but also for classification functions to ultimately achieve a higher accuracy than traditional diagnostic models.

The present work proposes a new feature selection and classification model based on the multi-task learning model, which treats the 1106 features extracted from each sample in each sequence (T1, T2, T1C) as a task. It uses the logical loss function as a data-fitting term to ensure the feasibility of classifying GBM and SBM. Taking $\ell_{1,1}$, ℓ_1 as regular terms to ensure the sparseness of feature selection. At the same time, private weights are introduced based on a common weight to make full use of the relevant similarities and differences between each task. The result is a 3-task feature selection classification model. In this study, we mainly utilized the alternating iterative method and the fast-iterative shrinkage threshold algorithm based on the backtracking method to solve equations. The experimental results demonstrate that

our model, whether as a feature selection model combined with SVM classification methods to form a diagnostic model or as a standalone diagnostic model, successfully integrated multiple sequence information to provide a robust predictive model for clinical diagnosis while also the diagnostic model consisting of one model improves the efficiency of our tumor classification.

2 Materials and methods

2.1 Data acquisition

The data used in this study were obtained from 120 patients with SBM and 67 patients with GBM admitted to the Xiangya Hospital, Yunnan Cancer Hospital and Southern Cancer Hospital between January 1, 2010 and December 31, 2018. All patients were histologically diagnosed according to the tumor grading guidelines published by the World Health Organization in 2021. This retrospective analysis of data from MR images was approved by the institutional review board, and the requirement of informed consent was waived.

MRI on all patients was performed by the hospital radiology department using 3.0-T systems. Each patient had T1, T2, and T1C MR image series performed. High-quality MR images were obtained using the following protocols:

- Axial T1: layer thickness =5 mm, layer spacing =1.5 mm, matrix =320×256, and field of view (FOV)=24×24 cm.
- Axial T2: layer thickness =5 mm, layer spacing =1.5 mm, matrix =384×384, and FOV =24×24 cm.
- Axial T1C: layer thickness =5 mm, layer spacing =1.5 mm, matrix =320×256, and FOV =24×24 cm.

MR image data of the patient for the present study can be found in the reference (21).

2.2 Data preprocessing

The overall workflow of the current study is shown in Figure 1 with a description of each involved step.

2.2.1 Delineation of the region of interest (ROI)

We preprocessed each image by noise reduction, offset field correction, and strict intra-target alignment using the public software package FSL v6.0.4. All images were evaluated independently by two neuroradiologists who have between 5-10 years of experience. ROIs of the entire tumor on T1, T2, and T1C images were created manually using the ITK-SNAP software layer by layer around the enhanced tumor layer (27); areas of macroscopic necrosis, cystic degeneration, and edema were avoided. A third senior neuroradiologist with 15 years of experience reexamined the images and made a final diagnosis when there was a conflict between the two original neuroradiologists (21).

2.2.2 Data normalization

Differences in instrumentation and imaging parameters, tumor sites of patients, and other factors can lead to significant differences in MR images. These differences will result in significant issues for imaging histology analysis. Therefore, we performed MIL(modality mismatch, intensity distribution variance, and layer-spacing differences) normalization for all MR images (28). First, we used B-sample interpolation for body mode matching for all patient MR images to obtain a total of 120 SBM samples and 67 GBM samples,

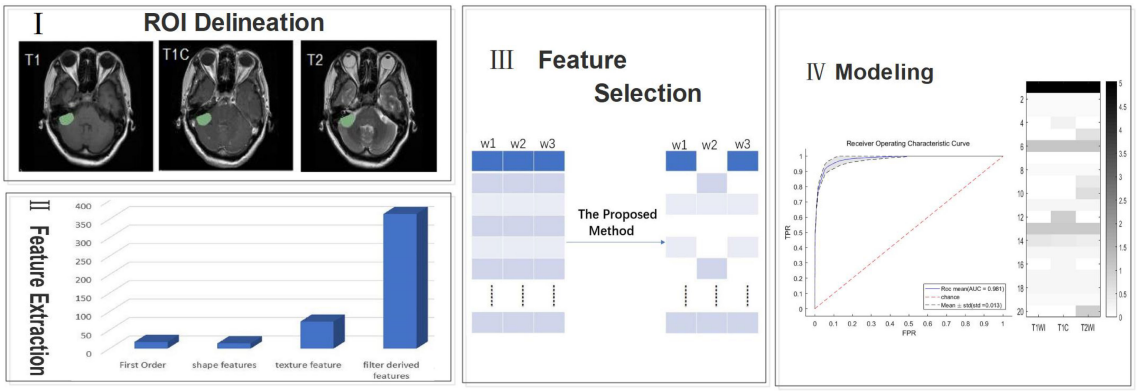


FIGURE 1
A generic framework for creating classification models using radiological features. Steps include ROI(region of interest) delineation, feature extraction, feature selection, and modeling.

Second, we set the interlayer gap of all MR images to 1 mm. Third, we applied MIL data normalization to make the intensity of MR image distribution consistent.

2.2.3 Feature extraction

We used the platform PyRadiomics (http://www.radiomics.io/pyradiomics.html) to perform feature extraction on all MIL normalized data. 1106 features were extracted from each MRI series. With each patient undergoing three different sequences of MRI, we extracted a total of 3318 features per patient. The extracted features are shown in Table 1.

To account for the large difference in the number of samples between the two tumor groups, we used the Synthetic Minority Over-sampling Technique (SMOTE) (29) to oversample the GBM group in order to generate the same number of samples as the SBM group. In total, we obtained 120 SBM samples and 120 GBM samples. Finally, we randomly selected 24 SBM samples and 24 GBM samples as the test set, and the remaining 96 SBM samples and 96 GBM samples as the training set.

2.3 Feature selection and classification

2.3.1 Proposed model

We introduce the logical loss function into the multi-task learning model to obtain the data fitting term as

$$L(X, y, W) = \sum_{m=1}^{M}\sum_{j=1}^{N}(- y_i^m X_i^m w^m + \ln\ (1 + \exp\ (X_i^m w^m)) \quad (1)$$

where the number of samples for each task is N and $y_j^m \in \{+1, -1\}$ denotes the class of samples. When $y = -1$, the patient has GBM.$y = +1$ means the patient has SBM.

To obtain the characteristics unique to a single task, we introduced a personal value of b based on the expected weight of s and finally derived our new model

TABLE 1 Extract the specific content of radiomics features.

Original image		Derived image	
feature	**number**	**name**	**number**
Shape-based	14	LoG	273
First Order	18	Wavelet	364
Gray Level Co-occurrence Matrix	22	Square	91
Gray Level Run Length Matrix	16	Square Root	91
Gray Level Size Zone Matrix	16	Logarithmic	91
Gray Level Dependence Matrix	14	Exponential	91
Neighboring Gray Tone Difference Matrix	5	–	–

$$\min_{s,B}\ \sum_{m=1}^{M}\sum_{i=1}^{N}\frac{1}{N}(-y_i X_i^m(s + b^m) + \ln\ (1 + \exp\ (X_i^m(s + b^m))))$$
$$+ \lambda_s \left\| s \right\|_1 + \lambda_b \left\| B \right\|_{1,1} \quad (2)$$

Where y_i is the label of the i^{th} sample. M is the number of tasks where $M=3$ as described in the text. We assume that the number of samples for all tasks is $N.X_i^m \in \mathbb{R}^{1\times(d+1)}$ is the i^{th} sample of the m^{th} task, i.e., the i^{th} row of the matrix X^m. Depending on the nature of logistic regression, we populate the last column of X^m with an N-dimensional 1 vector. $B = [b^1, b^2, \cdots, b^M] \in \mathbb{R}^{(d+1)\times M}$, then $b^m$is the m^{th} column of matrix $B, ||s||_1 = \sum_{i=1}^{d+1}|s_i|$ and $||B||_{1,1} = \sum_i^{(d+1)}\sum_m^M |b_i^m|$ are the regularization terms where s and bmare the shared and private weights, respectively.λ_s, λ_b are the two regularization parameters, and let $\lambda_s < M\lambda_b$.

2.3.2 Model solving

We rewrote (2) in the following form

$$\min_{s,B}\sum_{m=1}^{M}\sum_{i=1}^{N}\frac{1}{N}(- y_i X_i^m(s + b^m) + \ln\ (1 + \exp\ (X_i^m(s + b^m))))$$
$$+ \lambda_s\sum_{i}^{p+1}|s_i| + \lambda_b\sum_{i}^{d+1}\sum_{m}^{M}|b_i^m| \quad (3)$$

We considered the matrix B as a constant matrix, with a minimization of the variable s

$$\min_{s}\ \sum_{m=1}^{M}\sum_{i=1}^{N}\frac{1}{N}(- y_i X_i^m(s + b^m) + \ln\ (1 + \exp\ (X_i^m(s$$
$$+ b^m)))) + \lambda_s\sum_{i}^{p+1}|s_i| \quad (4)$$

Similarly, fixing the vector s and considering it as a constant vector, then minimizing B is equivalent to minimizing the following problem for any m

$$\min_{b^m}\ \sum_{i=1}^{N}\frac{1}{N}(- y_i X_i^m(s + b^m) + \ln\ (1 + \exp\ (X_i^m(s + b^m))))$$
$$+ \lambda_b\sum_{i}^{d+1}|b_i^m| \quad (5)$$

It is further shown that the computation of b^m is only relevant for a single task.

To select the final feature, equation (4) and (5) must be solved. Because of the non-differentiation of the $\ell 1$ norm, we used a fast-iterative shrinkage threshold algorithm to solve the above two subproblems in the course of our study (30). Both equations were solved in a similar manner, but the calculation in the model (5) involved only the mth task, independent of the other tasks, while solving model (4) required the participation of all tasks.

2.3.3 Model analysis

We represent the weights solved by the model in the form of a sum of shared weights s and private weights b^m. For the computation of s, all tasks need to be involved, and when $s_i \neq 0$, it is assumed that all tasks pick the i^{th} feature. However, the computation of b^m requires data from only the m^{th} task, and $b^m \neq 0$ indicates the i^{th} feature is important for the m^{th} task but may not be important for other tasks. Finally, we denoted the features selected by the mth task by $(s + b^m)$. The $\ell_1$and $\ell_{1,1}$ regularization forced both weights to be sparse, thus satisfying the "feature selection" requirement. The feature selection methods with $\ell_{1,1}$ norm as the regular term or $\ell_{2,1}$ norm as the regular term are two classical methods in the sparse embedding method (31). The multi-task Lasso model based on $\ell_{1,1}$ regularization had an entirely separable form. Each task can separately compute its own weight without being influenced by other tasks, so the sparse multi-task Lasso model is equivalent to the single-task Lasso model. The single difference is that all tasks of the sparse multi-task Lasso method use the same regularization parameter. In contrast, the regularization parameter of the single-task Lasso model can be determined by each task individually. The multi-task Lasso model based on the $\ell_{2,1}$regularization makes the features selected by different tasks almost the same, which reasonably exploits the correlation between different tasks but loses the specificity between different tasks. In contrast, our model not only makes full use of the correlation between different tasks but also highlights the specificity between different tasks and fuses the information of multiple sequences, thus providing a more robust prediction model for clinical diagnosis. Their differences are given visually in Figure 2.

Accurate preoperative diagnosis can be effective in formulating accurate and personalized treatment for patients, especially when MR images of SBM and GBM are extremely similar. In the current study, the proposed multi-task learning based on the logistic loss function can be used not only for feature selection but also for tumor classification tasks.

In contrast, the Lasso model based on mean square loss can only be used as a feature selection algorithm and cannot perform the subsequent classification task independently. In contrast, our model can use probabilities to account for the classification, for example, for any sample $X_i^m \in \mathbb{R}^{1\times(d+1)}$, the probability that the sample is classified as label 1 is

$$p(y_i = 1|X_i^m) = \frac{1}{1 + e^{-X_i^m(s+b^m)}} \tag{6}$$

Next, the model can perform the tumor classification task independently by simply picking the appropriate threshold

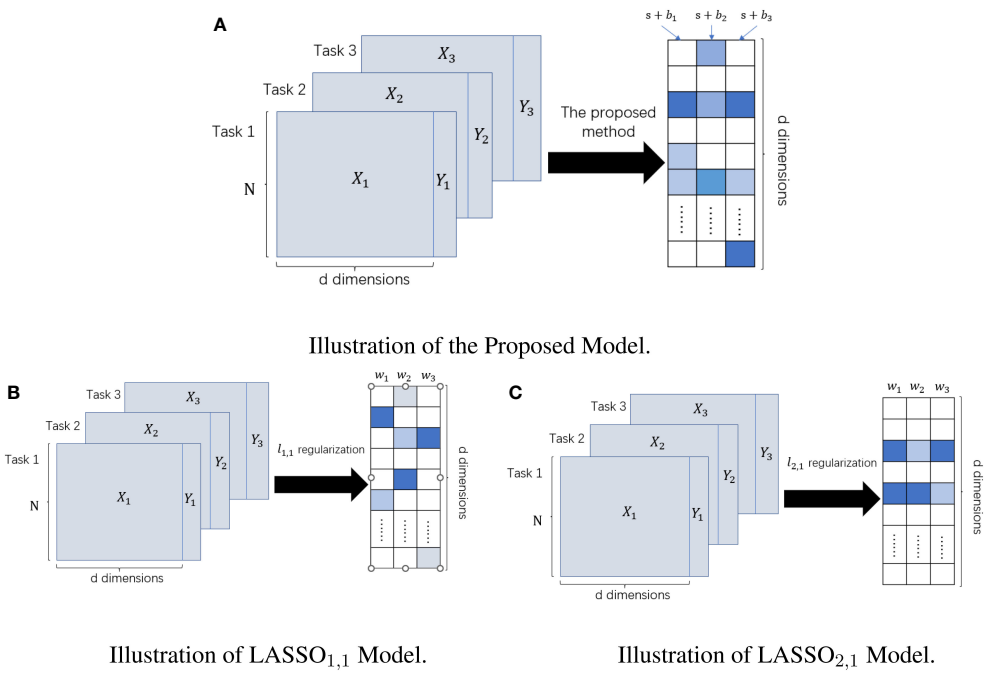


FIGURE 2
The difference between LASSO$_{1,1}$ Model, LASSO$_{2,1}$ Model, and Our Proposed Model. The left panel represents the input datasets; the right panel represents the learned weight matrix. **(A)** Schematic of our method for feature selection. **(B)** Schematic diagram of feature selection based on multi-task Lasso model under $\ell_{1,1}$ regularization. **(C)** Schematic diagram of feature selection based on multi-task Lasso model under $\ell_{2,1}$ regularization.

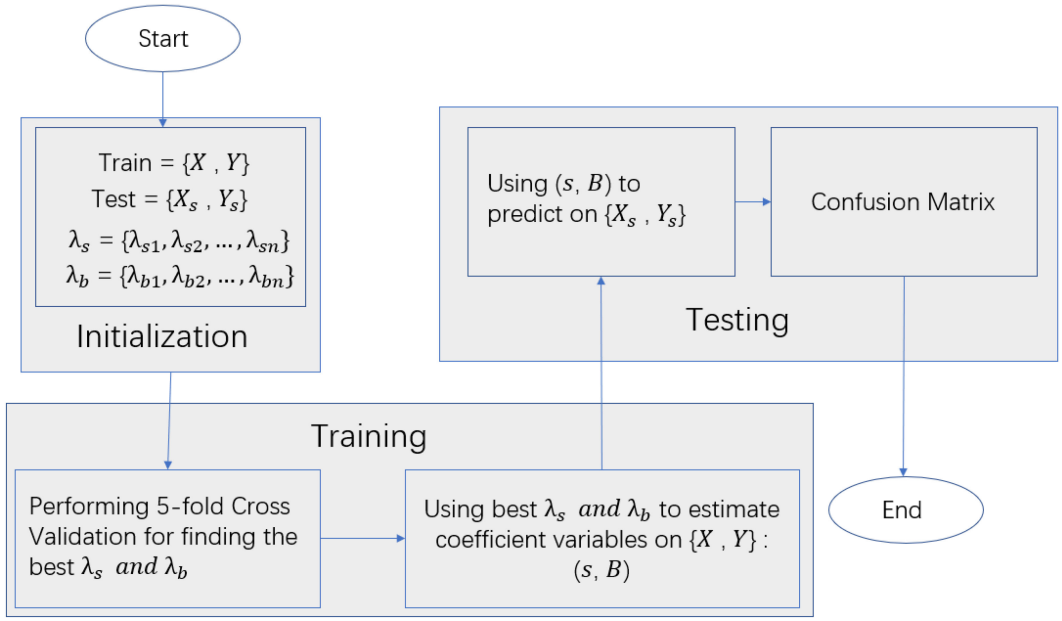


FIGURE 3
Experimental flowchart.

value. Its calculation process will be described in the section of experimental procedure.

2.3.4 Experimental procedure

This experiment was performed using MATLAB 9.11. The features of each sample and the labels are used as input, and the optimal penalty parameters λ_s and λ_b are obtained on the training set using 5-fold cross-validation. The ratio $\frac{\lambda_s}{\lambda_b}$ of the two penalty parameters is always kept as the following six values: 1.25, 1.5, 1.75, 2, 2.25, 2.5. Moreover, the mean AUC values under the optimal penalty parameters are recorded. Then, the data were computed and trained using the alternating miniaturization algorithm and the fast-iterative shrinkage threshold algorithm. Note that during the fast-iterative shrinkage threshold algorithm solution, our parameters $L_0 = 1, \eta = 1.1$. To ensure the confidence of the results, we repeatedly performed 5-fold cross-validation 10 times. Finally, the model"s performance was evaluated by average accuracy, average recall, average precision, and average AUC(The flow chart of our model solution is given in Figure 3).

3 Results

Additional analyses were conducted to compare our model with four diagnostic models based on single-task and multi-task learning. The single-task-based models are the diagnostic model with Lasso model for feature selection and SVM for classification (Lasso+SVM); the diagnostic model with Lasso model for feature selection and logistic regression (LR) algorithm for classification (Lasso+LR); the logistic regression as loss function and ℓ_1 norm constraint of "LR_1" model for feature selection and SVM for classification (LR_1+SVM); and a diagnostic model with LR_1 model for feature selection and classification (LR_1).

The models based on multi-task learning are the diagnostic model with the Lasso multi-task model with $\ell_{1,1}$norm as the regular term for feature selection and the SVM method for classification($Lasso_{1,1}$+SVM); the diagnostic model using the $Lasso_{1,1}$ multi-task model for feature selection and logistic regression(LR) for classification($Lasso_{1,1}$+LR); diagnostic model with feature selection using the Lasso multi-task model with $\ell_{2,1}$ norm as the regular term and classification by SVM method($Lasso_{2,1}$+SVM); diagnostic model with feature selection using $Lasso_{2,1}$ multi-task model and classification by LR method ($Lasso_{2,1}$+LR); our model as a diagnostic model with feature selection method and SVM method for classification (Ours+SVM) and our model as a diagnostic model(Ours).

Table 2 shows the confusion matrices of 4 single-task models based on T1, T1C, and T2 sequences and 6 diagnostic models based on multi-task learning. The values in the table are taken as an average of the results of the 5-fold cross-validation 10 times. In the table, TP (True Positive) represents the number of GBM predicted as GBM, FN (False Negative) means the number of GBM predicted as SBM, FP (False Positive) means the number of SBM predicted as GBM, and TN (True Negative) represents

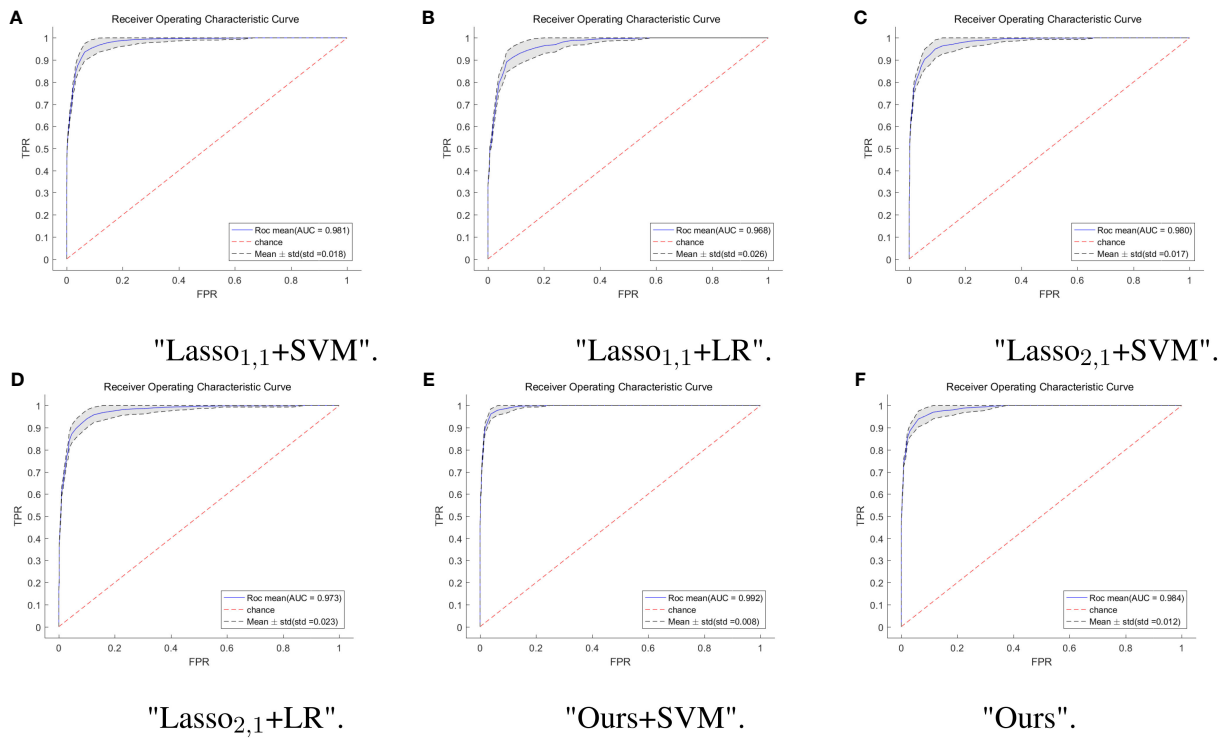


FIGURE 4
ROC curves of multi-task model in six. The horizontal coordinate indicates the false positive rate and the vertical coordinate indicates the true positive rate. **(A)** ROC curves for the $Lasso_{1,1}$+SVM model, the blue solid line indicates the average ROC curve and the black dashed line indicates the mean ± std. **(B)** ROC curves for the $Lasso_{1,1}$+LR model, the blue solid line indicates the average ROC curve and the black dashed line indicates the mean ± std. **(C)** ROC curves for the $Lasso_{2,1}$+SVM model, the blue solid line indicates the average ROC curve and the black dashed line indicates the mean ± std. **(D)** ROC curves for the $Lasso_{2,1}$+LR model, the blue solid line indicates the average ROC curve and the black dashed line indicates the mean ± std. **(E)** ROC curves for the Ours+SVM model, the blue solid line indicates the average ROC curve and the black dashed line indicates the mean ± std. **(F)** ROC curves for the Ours model, the blue solid line indicates the average ROC curve and the black dashed line indicates the mean ± std.

the number of SBM predicted as SBM. From the experimental results, it can be seen that the value of TP or TN is significantly higher than the value of FN or FP. This shows that our model is meaningful and feasible.

To further verify the feasibility of our model, the ROC (receiver operating characteristics) curves of our model are compared with those of four other multi-task models, and the values of the AUC and the standard deviation are also given. Usually, the closer to a value of 1 the AUC is, the better the model performance is; the smaller the standard deviation is, the more stable the model is. For the test dataset, our model combined with SVM classification, multi-task: Ours+SVM, had the most significant AUC of 0.992, however, our multi-task model alone had the second highest AUC of 0.984, with an AUC difference of 0.008 from the optimal model. The standard deviation (std) of the multi-task: Ours+SVM model is also the smallest among the six models, at 0.008, indicating that the model is the most stable. The std of our multi-task model alone was second only to the multi-task: Ours+SVM model. The results are shown in Figure 4.

Feature selection can effectively reduce the dimensionality of the data, thereby reducing the amount of computation and improving the efficiency of problem-solving. Figure 5 shows the distribution of features when selecting 20 features for six multitasking models. Our method not only selected the same features for the three sequences of T1, T1C, and T2, but also selected the characteristics that are unique to each sequence, which may have a large impact on only one of the sequences and not very much on the other sequences.

In brain tumor classification experiments, we compared the classification performance of the single-task-based and multi-task learning-based diagnostic models. In the single-task model experiments, we used four single-task feature selection models to classify the data of T1, T1C, and T2 sequences. Finally, we used 5-fold cross-validation method 10 times to obtain the average AUC on the training set and the average AUC, accuracy, precision, and recall on the test set. In the classification experiment based on multi-task learning, we treated the three sequences of T1, T1C, and T2 as 3 tasks, trained and tested them

10.3389/fonc.2022.1000471

TABLE 2 The mean of the confusion matrix of each model after 5-fold cross-validation 10 times.

Group	Model	Test			
		T P	F N	FP	TN
T1	Lasso+SVM	19.94	4.06	3.98	20.02
	Lasso+LR	19.78	4.22	4.54	19.46
	LR1+SVM	20.60	3.40	3.56	20.44
	LR_1	20.58	3.42	3.86	20.14
T1C	Lasso+SVM	20.36	3.64	2.58	21.42
	Lasso+LR	19.64	4.32	3.40	20.60
	LR1+SVM	19.40	4.60	1.90	22.10
	LR_1	19.18	4.82	1.98	22.02
T2	Lasso+SVM	20.50	3.50	3.38	20.62
	Lasso+LR	19.86	4.14	5.30	18.70
	LR1+SVM	20.60	3.40	2.56	21.44
	LR_1	20.24	3.76	2.96	21.94
Multi-task	$Lasso_{1,1}$+SVM	21.44	2.56	2.22	21.78
	$Lasso_{1,1}$+LR	20.48	3.52	2.70	21.30
	$Lasso_{2,1}$+SVM	20.16	3.84	1.64	22.36
	$Lasso_{2,1}$+LR	20.32	3.68	2.64	21.36
	Ours+SVM	20.90	3.10	0.76	23.24
	Ours	20.12	3.88	1.16	22.84

through 6 multi-task models, and obtained the above 4 evaluation indicators. The results were shown in Table 3.

In the single-task experiment, the LR_1+ SVM model based on the T2 sequence achieved the highest average AUC of 0.973 on the training set, and the average AUC also reached the highest 0.969 on the test set, and accuracy and recall also reached the highest in the single-task model, accuracy = 0.876, recall = 0.886. In all single-task experiments, the maximum value of precision is 0.922.

The multi-task learning model was introduced into the classification of brain tumors, and the classification performance of each model was significantly improved. As can be seen from the data in the table, the values of the indicators of the multi-task model are significantly better than those of the corresponding single-task model. Using our model, the multi-task: Ours+SVM model, its average AUC, accuracy, and precision are all at their highest, and the AUC on the test set reaches 0.992. Our model is not only used as a feature selection method, but also as a classification method. Although its metrics are not optimal, it outperforms the traditional diagnostic models ($Lasso_{1,1}$+SVM, $Lasso_{1,1}$+LR, $Lasso_{2,1}$+SVM, $Lasso_{2,1}$+LR), and we use only one model, thus improving the efficiency of diagnosis.

4 Discussion

Accurate classification of GBM and SBM is a challenging clinical problem. Different sequences of MRI provide unique information, and the rational fusion of multiple sequences can complement information from different dimensions (32). Thus, we proposed a new multi-task learning model to enable an accurate and fast diagnosis method for clinical usage.

This study introduced T1, T1C, and T2 sequences into the multi-task learning model. The feature weights were represented as the sum of shared and private weights. In turn, when filtering radiomic features, we can fully use the correlation between MR images of different sequences while still retaining the differences between the sequences and selecting features that have an essential impact on a specific task. Based on the above multi-task model, we also replaced the data matching term with a logistic regression function, which resulted in efficient model feature selection and classification of brain tumors.

We used an alternating minimization algorithm and a fast-iterative shrinkage threshold algorithm to train the data in model solving. We used the 5-fold cross-validation method to select optimal parameters for the selection of parameters. To ensure the accuracy and credibility of the data results, we conducted a 5-fold cross-validation 10 times in training and testing, and the final metrics were selected as the average value. The optimum model with an average AUC of 0.992 on the test set was found when our model performed feature selection and the SVM method performed classification. As a feature selection and classification method, our method alone reached the second highest average AUC of 0.984 on the test set. Multi-task learning enhances the robustness of our model, thus providing a stable predictive model for clinical diagnosis while ensuring accuracy

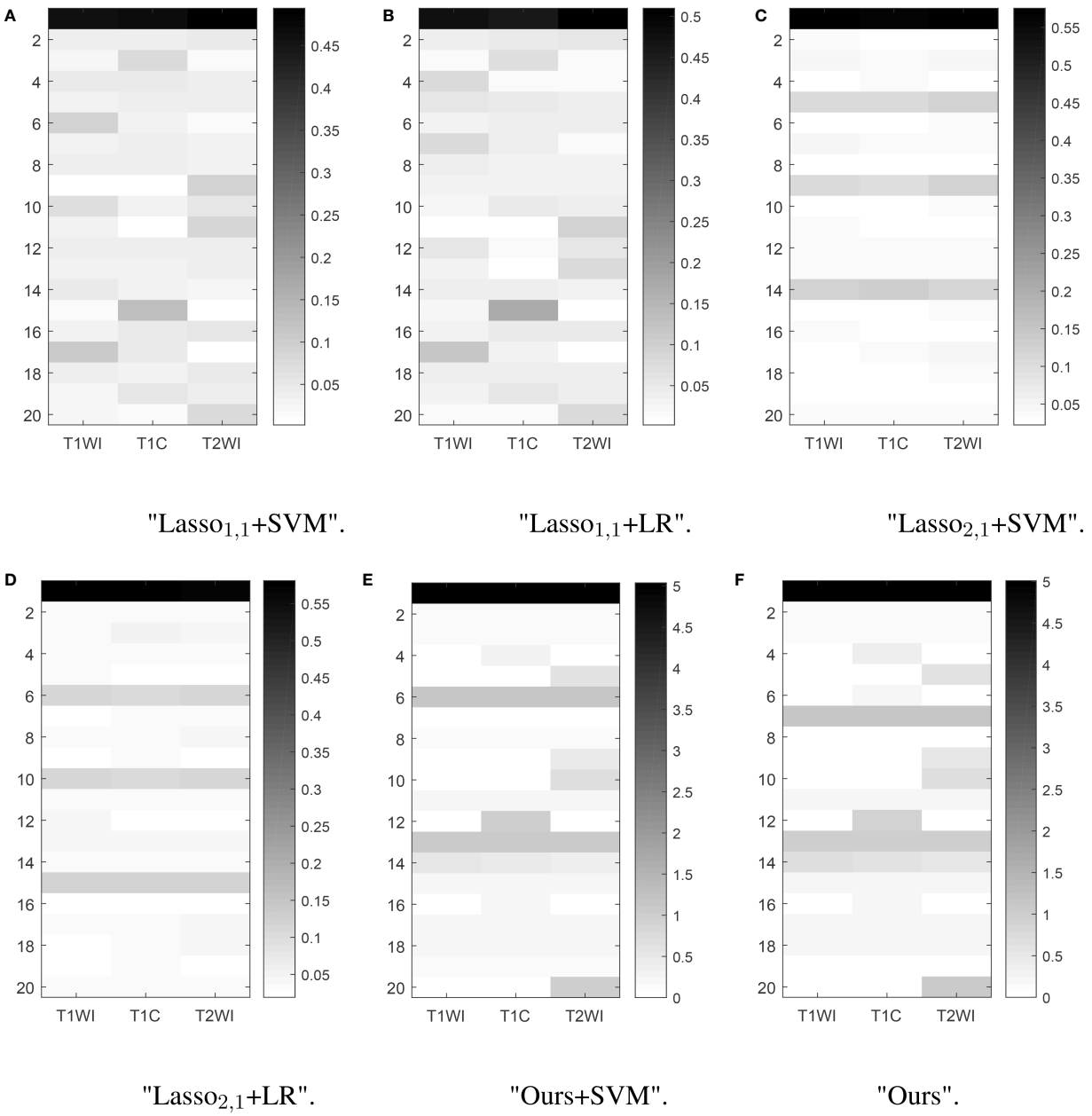


FIGURE 5
Learned weight matrix. The color bar in the right side indicates the values of matrix.The horizontal coordinates indicate the different tasks T1, T1C, and T2. The vertical coordinates indicate the number of features screened.

and making diagnoses possible with just one model, improving the efficiency of diagnostics.

Our model still has a comparative advantage over a single sequence classification task. For example, in a previous study (21), we used the same dataset with the random forest method as the feature selection method. Then six machine learning models were used for classification. The final result is that the random forest did the best classification job, obtaining an AUC of 0.97 on the test set but was limited to the T1C sequences. The single-task model, which does not consider the relationship between different sequences, does not take advantage of complementarity of information, which leads to the final classification effect being relatively not very good. On the other hand, the model proposed in this paper can make full use of the complementary information between different sequences and improve the accuracy and robustness of the prediction model. Our model is comparable with the classical multi-task model based on

TABLE 3 The values of various metrics for each method on the training and test sets.

		Train	Test			
Group	**Model**	**AUC**	**AUC**	**Accuracy**	**Precision**	**Recall**
T1	Lasso+SVM	0.936	0.933	0.832	0.851	0.830
	Lasso+LR	0.940	0.935	0.817	0.840	0.824
	LR_1+SVM	**0.952**	**0.951**	**0.855**	**0.869**	**0.858**
	LR_1	**0.951**	0.947	0.848	0.861	0.857
T1C	Lasso+SVM	**0.955**	0.953	**0.870**	0.901	**0.848**
	Lasso+LR	0.950	0.947	0.838	0.886	0.818
	LR_1+SVM	**0.961**	**0.959**	0.864	**0.922**	0.818
	LR_1	0.958	0.954	0.858	0.919	0.799
T2	Lasso+SVM	0.954	0.953	0.856	0.873	0.854
	Lasso+LR	0.947	0.942	0.803	0.828	0.854
	LR_1+SVM	**0.973**	**0.969**	**0.876**	**0.828**	**0.858**
	LR_1	0.968	0.967	0.860	0.892	0.843
Multi-task	$Lasso_{1,1}$+SVM	0.984	0.981	0.900	0.918	**0.893**
	$Lasso_{1,1}$+LR	0.893	0.981	0.870	0.903	0.853
	$Lasso_{2,1}$+SVM	0.981	0.980	0.886	0.941	0.840
	$Lasso_{2,1}$+LR	0.975	0.973	0.868	0.913	0.847
	Ours+SVM	**0.993**	**0.992**	**0.920**	**0.969**	0.847
	Ours	0.987	0.984	0.895	0.954	0.838

Bold values indicate the largest metrics indifferent models in the same sequence task.

$\ell_{2,1}$ regularization (24), but it can extract not only the same features for each sequence but also features that are important for a specific task. Moreover, compared with the general feature selection model, our model also integrates feature selection and classification to improve efficiency and convenience for diagnosis.

The present study does have some limitations. First, this study used a manual method to segment ROI, which is time consuming. Additionally, although two to three researchers have been involved in the segmentation process, it is very challenging to eliminate all bias in the results. Second, the data sample is small and cannot be extrapolated from this particular population to the general population. Third, the multitask learning model proposed in the present study requires the same features among tasks and does not apply to all multitask problems. Lastly, in the medical imaging part, ROI segmentation requires two neurologists with 5 to 10 years of experience. In our future study, we plan to use deep learning algorithms or image segmentation methods to automatically delineate the ROI to improve the efficiency of our work.

5 Conclusion

In this work, we proposed a feature selection model based on the multi-task learning model for SBM and GBM classification. The feature selection model uses a logistic regression function as a loss function, which makes the classification function of the model possible. Most of the current brain tumor classification studies have been performed using single-task models, which do not take advantage of the correlation between different sequences of MR images and therefore, the performance is not optimized to utilize all available information. Furthermore, the traditional multi-task Lasso model does not fully consider the correlation between different tasks. In contrast, our model makes full use of the correlation between MR images of different sequences while selecting the features that have an essential impact on a specific task. It is possible to select different combinations of features for different tasks, thus improving the classification performance of the model to some extent. In conclusion, our model generally outperforms the traditional multi-task Lasso model.

Our model as a feature selection method and paired with an SVM classification method has a great advantage over other methods of the same type. Our proposed model is also a good choice as a classification method. Although it has inferior performance to that of using our method with other classification algorithms, it improves the convenience of tumor classification. Thus, our model is advantageous in classifying SBM and GBM using MR images with multiple sequences.

Data availability statement

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

Ethics statement

Studies involving human subjects were reviewed and approved by the ethics committees of Xiangya Hospital, Yunnan Cancer Hospital, and Southern Cancer Hospital. The requirement for informed consent was waived.

Author contributions

YH: Writing and Data Analysis. SH: Data Analysis. ZL: Founding acquisition; Writing review and editing. All authors contributed to the article and approved the submitted version.

Funding

The work is supported by Natural Science Foundation of Hunan Province of China with the grant NO.2022JJ30944.

Acknowledgments

We thank the High-Performance Computing Center of Central South University for providing the computation resource.

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

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

1. Pei L, Jones KA, Shboul ZA, Chen JY, Iftekharuddin KM. Deep neural network analysis of pathology images with integrated molecular data for enhanced glioma classification and grading. *Front Oncol* (2021) 11:668694. doi: 10.3389/fonc.2021.668694

2. Alex B, Eva GV, Garth C, Tom R. Primary and metastatic brain tumours in adults: summary of nice guidance. *BMJ* (2018) 362:k2924. doi: 10.1136/bmj.k2924

3. Chen C, Zheng A, Ou X, Wang J, Ma X. Comparison of radiomics-based machine-learning classifiers in diagnosis of glioblastoma from primary central nervous system lymphoma. *Front Oncol* (2020) 10:1151. doi: 10.3389/fonc.2020.01151

4. Nayak L, Lee EQ, Wen PY. Epidemiology of brain metastases. *Curr Oncol Rep* (2012) 14:48–54. doi: 10.1007/s11912-011-0203-y

5. Chen C, Ou X, Wang J, Guo W, Ma X. Radiomics-based machine learning in differentiation between glioblastoma and metastatic brain tumors. *Front Oncol* (2019) 9:806. doi: 10.3389/fonc.2019.00806

6. Cha S, Lupo J, Chen MH, Lamborn K, McDermott M, Berger M, et al. Differentiation of glioblastoma multiforme and single brain metastasis by peak height and percentage of signal intensity recovery derived from dynamic susceptibility-weighted contrast-enhanced perfusion mr imaging. *Am J Neuroradiol* (2007) 28:1078–84. doi: 10.3174/ajnr.a0484

7. Bae S, An C, Ahn SS, Kim H, Han K, Kim SW, et al. Robust performance of deep learning for distinguishing glioblastoma from single brain metastasis using radiomic features: model development and validation. *Sci Rep* (2020) 10:1–10. doi: 10.1038/s41598-020-68980-6

8. Lambin P, Rios-Velazquez E, Leijenaar R, Carvalho S, Van Stiphout RG, Granton P, et al. Radiomics: extracting more information from medical images using advanced feature analysis. *Eur J Cancer* (2012) 48:441–6. doi: 10.1016/j.ejca.2011.11.036

9. Lambin P, Leijenaar RT, Deist TM, Peerlings J, De Jong EE, Van Timmeren J, et al. Radiomics: The bridge between medical imaging and personalized medicine. *Nat Rev Clin Oncol* (2017) 14:749–62. doi: 10.1038/nrclinonc.2017.141

10. Kocher M, Ruge MI, Galldiks N, Lohmann P. Applications of radiomics and machine learning for radiotherapy of malignant brain tumors. *Strahlentherapie und Onkol* (2020) 196:856–67. doi: 10.1007/s00066-020-01626-8

11. Louis DN, Perry A, Reifenberger G, Deimling AV, Figarella-Branger D, Cavenee WK, et al. The 2016 world health organization classification of tumors of the central nervous system: A summary. *Acta Neuropathol* (2016) 131:803–20. doi: 10.1007/s00401-016-1545-1

12. Rathore S, Akbari H, Doshi J, Shukla G, Rozycki M, Bilello M, et al. Radiomic signature of infiltration in peritumoral edema predicts subsequent recurrence in glioblastoma: Implications for personalized radiotherapy planning. *J Med Imaging* (2018) 12:021219. doi: 10.1117/1.JMI.5.2.021219

13. Austin D, Zhang B, Taimur S, Andrew P, Nicholas L, Mary GS, et al. Diagnostic accuracy of mri texture analysis for grading gliomas. *J Neuro-Oncol* (2018) 140:583–9. doi: 10.1007/s11060-018-2984-4

14. Zhao Z, Liu H. Spectral feature selection for supervised and unsupervised learning. *Proc 24th Int Conf Mach Learn* (2007) 227:1151–7. doi: 10.1145/1273496.1273641

15. Peng H, Long F, Ding C. Feature selection based on mutual information criteria of max-dependency, max-relevance, and min-redundancy. *IEEE Trans Pattern Anal Mach Intell* (2005) 27:1226–38. doi: 10.1109/TPAMI.2005.159

16. Constantinopoulos C, Titsias MK, Likas A. Bayesian Feature and model selection for gaussian mixture models. *IEEE Trans Pattern Anal Mach Intell* (2006) 28:1013–8. doi: 10.1109/TPAMI.2006.111

17. Sun Z, Yu Y. Robust multi-class feature selection *via* $\ell_{2,0}$-norm regularization minimization. *Intelligent Data Anal* (2022) 26:57–73. doi: 10.3233/IDA-205724

18. Qian Z, Li Y, Wang Y, Li L, Li R, Wang K, et al. Differentiation of glioblastoma from solitary brain metastases using radiomic machine-learning classifiers. *Cancer Lett* (2019) 451:128–35. doi: 10.1016/j.canlet.2019.02.054

19. Cho HH, Lee SH, Kim J, Park H. Classification of the glioma grading using radiomics analysis. *PeerJ* (2018) 6:e5982. doi: 10.7717/peerj.5982

20. Artzi M, Bressler I, Ben Bashat D. Differentiation between glioblastoma, brain metastasis and subtypes using radiomics analysis. *J Magnetic Resonance Imaging* (2019) 50:519–28. doi: 10.1002/jmri.26643

21. Liu Z, Jiang Z, Meng L, Yang J, Zhou R. Handcrafted and deep learning-based radiomic models can distinguish gbm from brain metastasis. *J Oncol* (2021) 2021:1–10. doi: 10.1155/2021/5518717

22. Bi J, Xiong T, Yu S, Dundar M, Rao RB. An improved multi-task learning approach with applications in medical diagnosis. In: *Joint European conference on machine learning and knowledge discovery in databases*. Berlin, Heidelberg: Springer (2008). p. 117–32. doi: 10.1007/978-3-540-87479-926

23. Fourure D, Emonet R, Fromont E, Muselet D, Neverova N, Tremeau A, et al. Multi-task, multi-domain learning: Application to semantic segmentation and pose regression. *Neurocomputing* (2017) 251:68–80. doi: 10.1016/j.neucom.2017.04.014

24. Obozinski G, Taskar B, Jordan M. *Multi-task feature selection*. Statistics Department, UC Berkeley: Citeseer (2006) 2:2.

25. Obozinski G, Taskar B, Jordan M. *Joint covariate selection for grouped classification*. Dordrecht, Netherlands: Dept. of Statistics, University of California Berkeley (2007).

26. Yin P, Lou Y, He Q, Xin J. Minimization of ℓ_{1-2} for compressed sensing. *SIAM J Sci Computing* (2015) 37:A536–63. doi: 10.1137/140952363

27. Liang C, Li Y, Luo J. A novel method to detect functional microrna regulatory modules by bicliques merging. *IEEE/ACM Trans Comput Biol Bioinf* (2016) 13:1–1. doi: 10.1109/TCBB.2015.2462370

28. Hu Z, Zhuang Q, Xiao Y, Wu G, Shi Z, Chen L, et al. Mil normalization-prerequisites for accurate mri radiomics analysis. *Comput Biol Med* (2021) 104403. doi: 10.1016/j.compbiomed.2021.104403

29. Chawla NV, Bowyer KW, Hall LO, Kegelmeyer WP. Smote: Synthetic minority over-sampling technique. *J Artif Int Res* (2002) 16:321–57. doi: 10.1613/jair.953

30. Beck A, Teboulle M. A fast iterative shrinkage-thresholding algorithm for linear inverse problems. *SIAM J Imaging Sci* (2009) 2:183–202. doi: 10.1137/080716542

31. Obozinski G, Taskar B, Jordan MI. Joint covariate selection and joint subspace selection for multiple classification problems. *Stat Computing* (2010) 20:231–52. doi: 10.1007/s11222-008-9111-x

32. Bathla G, Priya S, Liu Y, Ward C, Le N, Soni N, et al. Radiomics-based differentiation between glioblastoma and primary central nervous system lymphoma: A comparison of diagnostic performance across different mri sequences and machine learning techniques. *Eur Radiol* (2021) 31:1–11. doi: 10.1007/s00330-021-07845-6

frontiers | Frontiers in Oncology

TYPE Original Research
PUBLISHED 03 October 2022
DOI 10.3389/fonc.2022.941786

OPEN ACCESS

EDITED BY
Sweet Ping Ng,
University of Melbourne, Australia

REVIEWED BY
Tyler McKechnie,
McMaster University, Canada
Ji Zhu,
Fudan University, China

*CORRESPONDENCE
Benhua Xu
benhuaxu@163.com
Guoxian Guan
fjxhggx@163.com

†These authors have contributed equally to this work

SPECIALTY SECTION
This article was submitted to Cancer Imaging and Image-directed Interventions, a section of the journal Frontiers in Oncology

RECEIVED 11 May 2022
ACCEPTED 12 September 2022
PUBLISHED 03 October 2022

CITATION
Guan B, Huang X, Xia H, Guan G and Xu B (2022) Prognostic value of mesorectal package area in patients with locally advanced rectal cancer following neoadjuvant chemoradiotherapy: A retrospective cohort study. *Front. Oncol.* 12:941786. doi: 10.3389/fonc.2022.941786

Prognostic value of mesorectal package area in patients with locally advanced rectal cancer following neoadjuvant chemoradiotherapy: A retrospective cohort study

Bingjie Guan[1†], Xinmin Huang[2†], Huang Xia[1], Guoxian Guan[3*] and Benhua Xu[1,4,5*]

[1]Department of Radiation Oncology, Fujian Medical University Union Hospital, Fuzhou, China, [2]Department of Radiology, Fujian Medical University Union Hospital, Fuzhou, China, [3]Department of Colorectal Surgery, The First Affiliated Hospital of Fujian Medical University, Fuzhou, China, [4]Fujian Key Laboratory of Intelligent Imaging and Precision Radiotherapy for Tumors, Fujian Medical University, Fuzhou, China, [5]Clinical Research Center for Radiology and Radiotherapy of Fujian Province, Fuzhou, China

Background: The aim of this study is to explore the most effective inflammation, magnetic resonance imaging (MRI), and nutrition markers for survival and pathology complete response (pCR) in patients with locally advanced rectal cancer (LARC).

Methods: A total of 278 patients with LARC undergoing neoadjuvant chemoradiotherapy (NCRT) and radical surgery from 2016 to 2019 were included. The X-tile method was used to select the optimal cutoff points for the mesorectal package area (MPA), advanced lung cancer inflammation index (ALI), prognostic nutritional index (PNI), systemic immune-inflammation index (SII), neutrophil-to-lymphocyte ratio (NLR), platelet-to-lymphocyte ratio (PLR), and monocyte-to-lymphocyte ratio (MLR) scores. Cox regression analysis was used to identify risk factors of disease-free survival (DFS). To discover pCR risk factors, logistic regression analysis was employed. A predictive nomogram for DFS was constructed.

Results: According to the least absolute shrinkage and selection operator analysis, the MPA was the only significant predictor for the DFS in patients with LARC. Kaplan-Meier (K-M) analysis demonstrated that groups with higher MPA, PNI, SII, NLR, MLR, and ALI score had improved DFS (all $P < 0.05$). Receiver operating characteristic (ROC) analysis revealed that the MPA and PNI could accurately predict the pCR in patients with LARC after NCRT. The MPA score and NLR score were found to be independent predictors of DFS after NCRT using Cox regression analysis. Logistical regression analysis demonstrated that the MPA score, PNI score, and pre-NCRT cN stage were all independent predictors of pCR in patients with LARC after NCRT. Recursive partitioning

analysis and time-independent ROC curve analysis demonstrated that MPA score was the most important predictor of pCR and prognosis in patients with LARC after NCRT.

Conclusions: MPA was identified as the most effective marker for MRI, and the prognostic value was further confirmed by time–ROC analysis. More intense adjuvant treatment could be considered for lower–MPA score patients with LARC after NCRT. Obesity in the pelvis encourages the understanding of the prognosis prediction of patients with LARC after NCRT.

KEYWORDS

locally advanced rectal cancer (LARC), neoadjuvant chemoradiotherapy, inflammation biomarkers, mesorectal package area, prognosis, pathology complete response

Introduction

The neoadjuvant chemoradiotherapy (NCRT) has been the standard treatment for locally advanced rectal cancer (LARC). The NCRT benefited from a higher likelihood of tumor shrinking and downstaging, enhanced tumor resectability, and better local tumor control (1–3). NCRT could contribute to pathological complete response (pCR) in 15%–27% of patients with LARC and 20%–30% near pCR in patients with LARC (4). Patients with pCR or near pCR could adopt the "watch and wait" strategy or local excision to reduce surgery-related morbidity and increase organ preservation when compared to the total mesorectal excision (TME) surgery (5–7). However, more than 30% of patients with LARC were resistant to NCRT and experiencing NCRT adverse effects (8, 9). Currently, it is still challenging to reliably estimate treatment outcomes for patients with LARC after NCRT.

The rates of obesity have risen in the recent years, and obesity contributes to a variety of chronic morbidities (10). Numerous studies have shown that obesity is associated with the occurrence and progression of colorectal cancer (11–13). However, the influence of obesity on NCRT response of LARC remains controversial (14–17). Body mass index (BMI) is the most common tool for assessing obesity, although Asians typically have normal BMI levels and abdominal obesity, which could lead to an inaccuracy evaluation. Instead of the BMI, the NCRT response may be related to the pelvic fat. Investigating pelvic fat may provide an answer to the question of whether obesity affects the NCRT responsiveness. High-resolution pelvic or rectal magnetic resonance imaging (MRI) may accurately quantify the fat in the pelvis and rectal mesorectal thickness to predict the NCRT response. Several studies found that obesity was contributing to the inflammatory response, which influenced tumor development, prognosis, and therapy response (18–20). The inflammatory indexes in the peripheral blood, NLR, MLR, PLR, and SII have been used as markers of predicting efficacy and toxicity of NCRT in patients with LARC in our previous study (21). To explore the relationship among the obesity, inflammatory response and NCRT response were important.

To address the gap in the literature, the present study aimed to explore the most effective marker of MRI measurements, systematic inflammatory, and nutrition in patients with LARC in terms of survival outcome and NCRT response.

Patients and treatment methods

Patients

In this study, we retrospectively analyzed 278 patients with LARC after NCRT who underwent pelvic MRI before NCRT in our hospital and radical resection between 2016 and 2019. The patient inclusion criteria and exclusion criteria were reported in our previous study (8, 22). The evaluation of the tumor staging was according to the American Joint Committee on Cancer (AJCC) (23). The TME was following the NCRT regimen, which has been described in our previous study. According to the National Comprehensive Cancer Network (NCCN) guidelines, the patients received postoperative adjuvant chemotherapy for 6 months about 1 month after surgery (24). All laboratory results and pelvic MRI images were collected within 1 week before NCRT. The last cutoff date for follow-up was 31 December 2021.

Neoadjuvant chemoradiotherapy

Two chemotherapeutic regimens with dosages were given as follows: (1) Capox: oxaliplatin at 130 mg/m^2 intravenously guttae, day 1; capecitabine at 825 mg/m^2 twice daily oral, days

1–14; every 3 weeks, for two cycles during concurrent radiotherapy; another two cycles were performed during the interval from the end of radiation to surgery; (2) capecitabine alone: capecitabine at 825 mg/m^2 twice daily oral, during the whole period of radiotherapy; another one cycle increased dosages to 1,250 mg/m^2 was performed in 2 weeks during the waiting period.

The gross tumor volume (GTV) was calculated on the basis of clinical information, including digital rectal examination, endoscopy ultrasound, and abdominopelvic MRI. The clinical target volume (CTV) included all mesorectum, presacral soft tissue, obturator, and internal iliac lymphatic drainage regions. The planning target volume (PTV) was defined as the GTV or CTV with uniform margins of 10 mm. The neoadjuvant radiotherapy regimens consisted of three-dimensional conformal radiotherapy (3D-CRT) and intensity-modulated radiation therapy (IMRT). A dose of 50.4 Gy was delivered to PTV-GTV with 3D-CRT in 28 fractions, whereas 50 Gy was delivered with IMRT in 25 fractions. In addition, 45 Gy was delivered to PTV-CTV in 25 fractions for both types of regimens (24).

Definitions

The pathological tumor regression grade (TRG) (23) was used as the evaluation criterion of tumor response to NCRT. No residual tumor cells in the resected specimen, including the primary site and lymph nodes, were regarded as pathological complete response (pCR). Venous blood samples were obtained within 1 week before NCRT. The following formulae were employed to determine the systematic inflammatory markers: The systemic immune-inflammation index (SII) = platelet count × neutrophil count/lymphocyte count, neutrophil-to-lymphocyte ratio (NLR) = neutrophil count/lymphocyte count, platelet-to-lymphocyte ratio (PLR) = platelet count/lymphocyte count and monocyte-to-lymphocyte ratio (MLR) = monocyte count/lymphocyte count. Advanced lung cancer inflammation index (ALI) = BMI (kg/m^2) × albumin (g/L)/NLR, prognostic nutritional index (PNI) = serum albumin (g/L) + 5 × lymphocyte count (10^9/L).

Pelvic MRI measurements

MRI was performed using either a 1.5-T General Electric (450 W, software version 25) or a 3-T Phillips (Achieva, software version 3.2.3.5) system. Large field-of-view FT2-weighted axial images with a slice thickness of 5 mm were downloaded from the PACS system and analyzed with publicly available software (3D Slicer©, Version 4.11; Bethesda, MD) (25) that was supported by National Institutes of Health. The measuring procedure is shown in Figure 1, as described by McKechnie et al. (26).

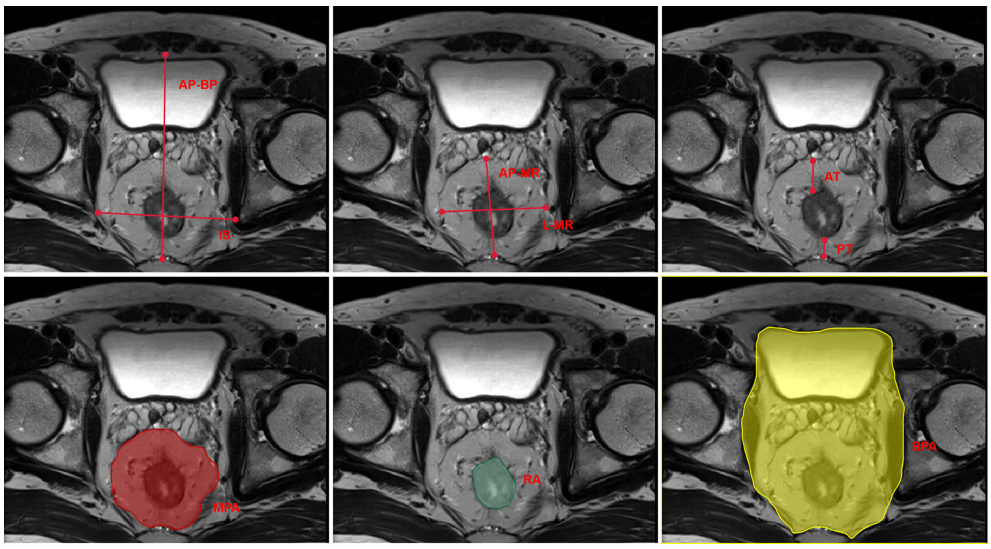


FIGURE 1
Schematic plot of the pelvic MRI measurements. MRI, magnetic resonance imaging; IS, interspinous distance; AP-BP, anterior–posterior bony pelvis span; L-MR, lateral mesorectal span; AP-MR, anterior–posterior mesorectal span; AT, anterior mesorectal thickness; PT, posterior mesorectal thickness; MPA, mesorectalpackage area; RA, rectal area; BPA, bony pelvis area.

Statistical analysis

The Statistic Package for Social Science (SPSS, version 23.0) and R software packages version 4.0.1 were used to perform the statistical analyses. The X-tile program (http://www.tissuearray.org/rimmlab/) was used to calculate and determine the best cutoff points for the SII, NLR, PLR, MLR, ALI, and PNI counts (27). The Kaplan–Meier method and log-rank test were performed to evaluate the survival outcomes. The risk factors for overall survival (OS) and disease-free survival (DFS) were calculated by the Cox proportional hazards model. On the basis of the Cox regression model analysis, a nomogram was developed by using the R project. Time-dependent receiver operating characteristic (ROC) curves were used to evaluate the performance of the biomarkers. Least absolute shrinkage and selection operator (LASSO) Cox regression model was applied to determine the ideal coefficient for each prognostic feature and estimate the likelihood deviance (28, 29). Recursive partitioning analysis (RPA) was used to construct a decision tree that divides patients into different homogeneous risk groups by using the R project (30). Statistical significance was defined as $P < 0.05$.

Result

Patient characteristics

A total of 278 patients with LARC after NCRT were eligible for this analysis. There were 181 (181 of 278, 65.1%) male patients, with a mean age of 53.97 ± 10.11 years. Tables 1, 2 list the baseline clinicopathological characteristics of the patients.

The LASSO analysis

LASSO analysis was used to explore significant predictors in MRI measurement markers for DFS in the patients with LARC after NCRT. The result demonstrated that the mesorectal package area (MPA) was the only factor that mattered (Figures 2A, B). Furthermore, the X-tile plot was employed to select the optimal cutoff point for the MPA, with the outcome revealing the cutoff values of 23 for MPA (Supplementary Figure 1). In addition, the best optimal cut-off point for the

TABLE 1 Baseline characteristics in patients with LARC after NCRT stratified by MAP.

Characteristics	MAP < 23 (n = 31)	MAP ≥ 23 (n = 247)	*P*-value
Sex (%)			1.000
Male	20 (64.5)	161 (65.2)	
Female	11 (35.5)	86 (34.8)	
Age (years)	54.35 ± 11.13	53.92 ± 10.68	0.831
ASA score (%)			0.161
1	24 (77.4)	210 (85.0.)	
2	6 (19.4)	36 (14.6)	
3	1 (3.2)	1 (0.4)	
Distance from the anal verge (cm)	6.57 ± 2.17	6.16 ± 2.36	0.357
Interval time between NCRT and surgery (weeks)	9.78 ± 1.98	9.77 ± 3.69	0.987
Pre-NCRT cT stage (%)			0.858
T2	0 (0.0)	4 (1.6)	
T3	12 (38.7)	83 (33.6)	
T4	19 (61.3)	160 (64.8)	
Pre-NCRT cN stage (%)			0.513
N0	1 (3.2)	19 (7.7)	
N+	30 (96.8)	228 (92.3)	
Pre-NCRT CEA (%)			0.443
<5.0 ng/ml	15 (48.4)	141 (57.1)	
≥5.0 ng/ml	16 (51.6)	106 (42.9)	
Pre-NCRT CA19-9 (%)			0.340
<37.0 U/ml	23 (74.2)	201 (81.4)	
≥37.0 U/ml	8 (25.8)	46 (18.6)	
Anemia (%)	6 (19.4)	38 (15.4)	0.601
Hypoproteinemia (%)	1 (3.2)	14 (5.7)	1.000

NLR, neutrophil-to-lymphocyte ratio; LARC, locally advanced rectal cancer; NCRT, neoadjuvant chemoradiotherapy; ASA, American Society of Anesthesiologists; CEA, carcinoembryonic antigen; CA19-9, carbohydrate antigen 19-9.

10.3389/fonc.2022.941786

TABLE 2 Operative and postoperative outcomes in patients with LARC after NCRT stratified by MPA.

Characteristics	MPA < 23 (n = 31)	MPA ≥ 23 (n = 247)	*P*-value
Operative time (min)	215.32 ± 50.51	225.15 ± 66.67	0.429
Estimated blood loss (ml)	60.48 ± 29.84	80.12 ± 108.45	0.317
Surgery approach (%)			0.157
Laparoscopic	28 (90.3)	195 (78.9)	
Open	3 (9.7)	52 (21.1)	
Tumor differentiation (%)			**0.014**
Well to moderately differentiated	21 (67.7)	214(86.6)	
Poorly differentiated and others	10 (32.3)	33 (13.4)	
Postoperative hospital stay (days)	9.03 ± 5.26	8.26 ± 4.55	0.379
Postoperative complications (%)	6(19.4)	42 (17.0)	0.801
BMI			**0.034**
<18	1(3.7)	10 (4.0)	
18~24	26 (83.9)	149 (60.3)	
>24	4 (12.9)	88 (35.6)	
Organ preservation (%)	30 (96.8)	230 (93.1)	0.703
Tumor size (cm)	2.53 ± 1.01	2.61 ± 1.26	0.767
Pathological T stage (%)			**0.001**
T0	1 (3.2)	70 (28.3)	
T1	0 (0.0)	16 (6.5)	
T2	11 (35.5)	60 (24.3)	
T3	16 (51.6)	97 (39.3)	
T4	3 (9.7)	4 (1.6)	
Pathological N stage (%)			0.122
N0	19 (61.3)	185 (74.9)	
N1	8 (25.8)	50 (20.2)	
N2	4 (12.9)	12 (4.9)	
Pathological M stage (%)			**0.011**
M0	27 (87.1)	242 (98.0)	
M1	4 (12.9)	5 (2.0)	
TRG (%)			**0.022**
0	1 (3.2)	67 (27.1)	
1	11 (35.5)	80 (32.4)	
2	17 (54.8)	85 (34.4)	
3	2 (6.5)	15 (6.1)	
pCR rates (%)	1 (3.2)	67 (27.1)	**0.002**
Nerval invasion (%)	5 (16.1)	18 (7.3)	0.155
Vascular invasion (%)	2 (6.5)	10 (4.0)	0.630
NLR score	4.71 ± 7.87	2.65 ± 2.10	**0.001**
SII score	1276.87 ± 2761.05	691.51 ± 693.10	**0.007**
MLR score	0.34 ± 0.21	0.27± 0.15	**0.022**
PLR score	161.34 ± 82.56	149.79 ± 97.47	0.528
PNI score	48.41 ± 4.20	49.75 ± 5.13	0.164
ALI score	42.97 ± 34.38	49.51 ± 26.42	0.212

LARC, locally advanced rectal cancer; NCRT, neoadjuvant chemoradiotherapy; NLR, neutrophil-to-lymphocyte ratio; TRG, tumor regression grade; pCR, pathological complete respons; SII, systemic immune-inflammation index; MLR, monocyte-to-lymphocyte ratio; PLR, platelet-to-lymphocyte ratio. P<0.05 was statistically significant in bold.

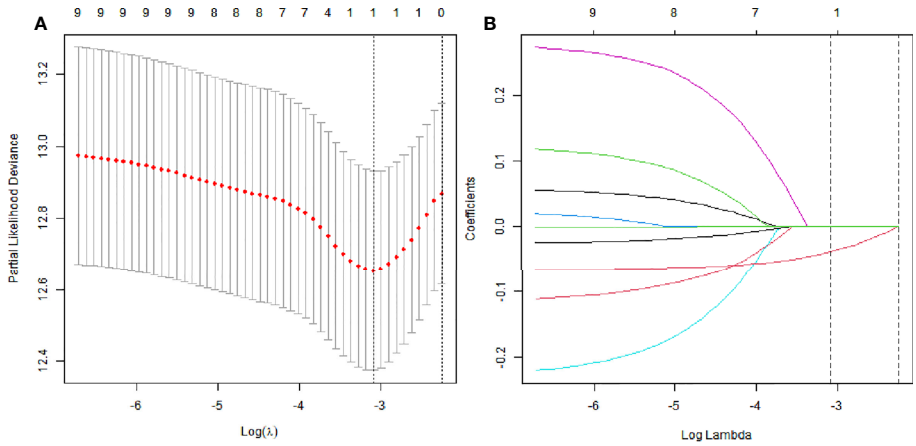


FIGURE 2
The least absolute shrinkage and selection operator (LASSO) analysis and risk score system were constructed. **(A)** The area under the ROC curve (AUC) was estimated with a cross-validation technique, and the largest lambda value was chosen when the cross-validation error was within one standard error of the minimum. **(B)** LASSO coefficient profiles of the eight factors.

MPA was enrolled in the next analysis. The result demonstrated that a high value of the MPA had better prognosis in the patients with LARC (DFS, P < 0.01, Figure 3G; OS, P = 0.05, Figure 4G).

Association of inflammation and nutrition biomarkers with survival

On the basis of the DFS, the X-tile plots were constructed and identified 540, 4.9, 0.268, 165, 46.8, and 31.8 as the cutoff values for SII, NLR, MLR, PLR, PNI, and ALI, respectively. Then, we divided the entire cohort into low and high subgroups. As shown in Figure 3, higher PNI and ALI scores were associated with better DFS in patients with LARC. DFS rates were significantly greater in the high PNI and ALI score group at 3 years, at 85.9% and 85.9%, respectively, than in the low PNI and ALI score group, at 67.7% and 54.5%, respectively (all P = 0.01; Figures 3E, F). Moreover, a high score of the SII, NLR, PLR, and MLR was correlated with worse DFS in patients with LARC compared with the low SII, NLR, PLR, and MLR score group, as shown in Figures 3A–D. The DFS rates at 3 years for the high SII, NLR, PLR, and MLR group were 73.2%, 53.8%, 76.3%, and 74.5%, respectively, significantly lower than 88.8%, 77.6%, 85.1%, and 87.1% in the low SII, NLR, PLR, and MLR groups,

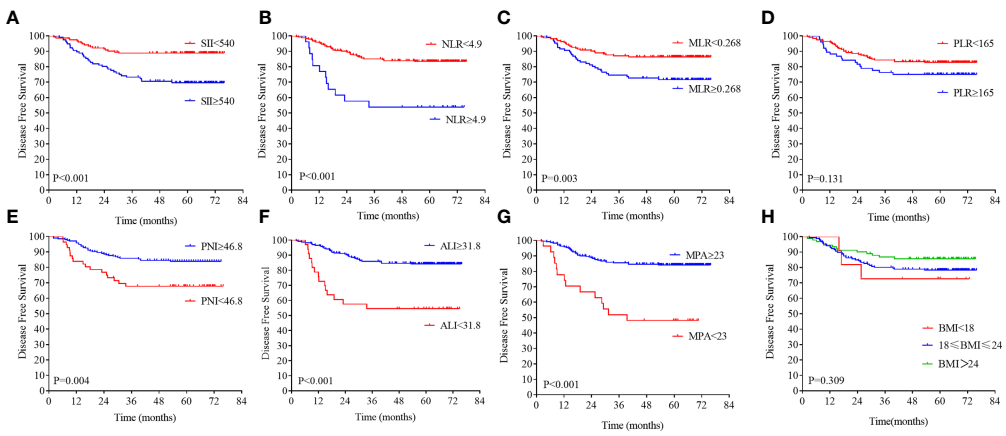


FIGURE 3
Kaplan–Meier analysis of the NLR, SII, PLR, MLR, PNI, ALI, MPA, and BMI level. The disease-free survival for the optimal cutoff point of the SII **(A)**, NLR **(B)**, MLR **(C)**, PLR **(D)**, PNI **(E)**, ALI **(F)**, MPA **(G)**, and BMI level **(H)**.

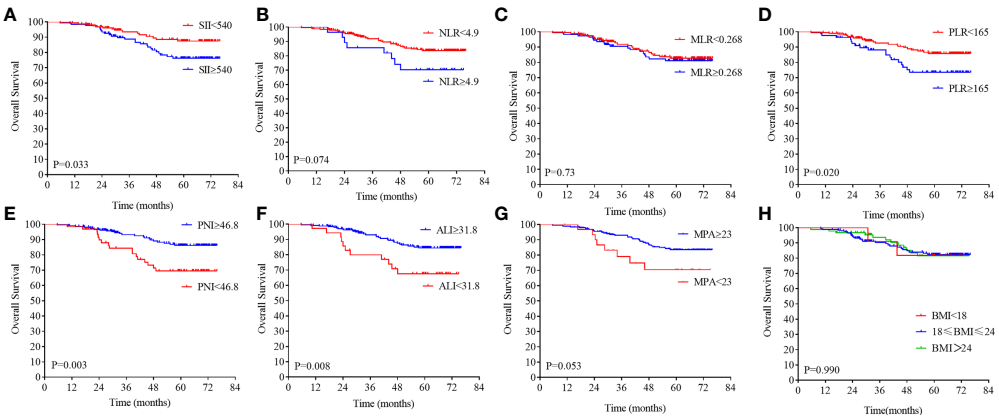


FIGURE 4
Kaplan–Meier analysis of the NLR, SII, PLR, MLR, PNI, ALI, MPA, and BMI level. The overall survival for the optimal cutoff point of the SII **(A)**, NLR **(B)**. MLR **(C)**, PLR **(D)**, PNI **(E)**, ALI **(F)**, MPA **(G)**, BMI **(H)**.

respectively (P < 0.01, P < 0.01, P = 0.13, and P < 0.01, respectively). Noticeably, the high PNI and ALI score groups had better OS compared with the low score group, as shown in Figures 4E, F (all P < 0.01). In addition, low SII and PLR score group had significantly better OS than the high score group (SII: P = 0.03, Figure 4A; PLR: P = 0.02, Figure 4D). There was no statistical difference between the low NLR and MLR score groups and the high NLR and MLR score groups (NLR: P = 0.07, Figure 4B; MLR: P = 0.73, Figure 4C). Moreover, the BMI level was not associated with the DFS and OS in the patients with LARC (DFS: P = 0.31, Figure 3H; OS: P = 0.99, Figure 4H).

Association of biomarkers with pCR

The correlations between pathological complete response (pCR) and MRI, inflammatory and nutritional biomarkers were further explored. The ROC analysis was performed to verify the predicting ability of the biomarkers. The MPA and PNI scores had powerful ability to predict the pCR in the patients with LARC [PNI: area under the ROC curve (AUC) = 0.62, P < 0.01, Figure 5E; MPA: AUC = 0.70, P < 0.01, Figure 5G]. However, the other biomarkers could not predict the pCR in the patients with LARC after NCRT (NLR: AUC = 0.53, P = 0.51, Figure 5A; MLR: AUC = 0.57, P =

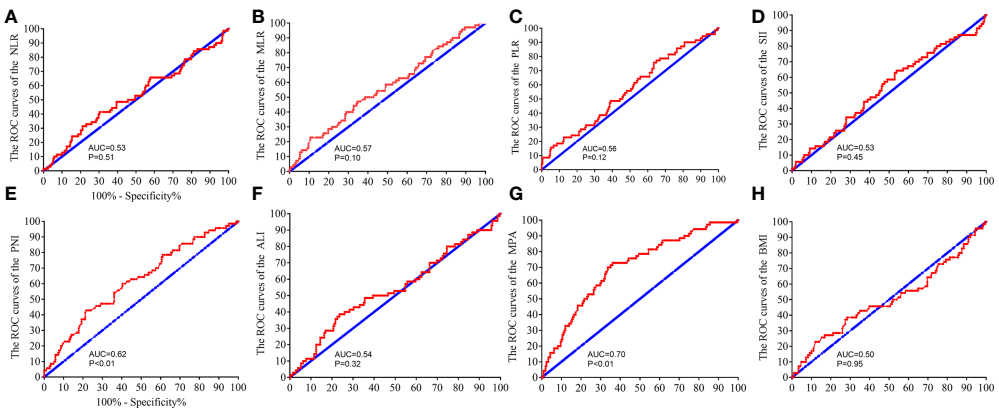


FIGURE 5
Receiver operating characteristic (ROC) analysis to evaluate the predictive efficiency of the NLR, SII, PLR, MLR, PNI, ALI, MPA, and BMI level in patients with LARC for NCRT response. The ROC analysis for NLR **(A)**, MLR **(B)**, PLR **(C)**, SII **(D)**, PNI **(E)**, ALI **(F)**, MPA **(G)**, and BMI level **(H)**.

0.10, Figure 5B; PLR: AUC = 0.56, P = 0.12, Figure 5C; SII: AUC = 0.53, P = 0.45, Figure 5D; ALI: AUC = 0.54, P = 0.32, Figure 5F; BMI: AUC = 0.50, P = 0.94, Figure 5H).

Association of MPA with clinicopathological characteristics

On the basis of the optimal cutoff value, these patients were dichotomized into the low-MPA group (n = 31, 11.1%) and the high-ALI group (n = 247, 88.9%). No significant differences were found between the groups regarding baseline characteristics, such as sex, age, American Society of Anaesthesiologists (ASA) score, preoperative carcinoembryonic antigen (CEA) level, preoperative CA19-9 level, distance from the anal verge, interval time between NCRT and surgery, pre-NCRT cT stage, pre-NCRT cN stage, hypoproteinemia, estimated blood loss, operative time, postoperative hospital stay, organ preservation, tumor size, BMI level, and anemia (all P > 0.05, Tables 1, 2). As expected, a lower MAP level was associated with poorly tumor differentiation, higher pathology TNM stage, higher AJCC TRG stage, higher NLR score, higher MLR score, and higher SII score (all P < 0.05).

Prognostic value of the biomarkers

To explore the prognostic impact of the biomarkers on DFS in patients with LARC, we performed a Cox regression model analysis. On univariate analysis, pathological T stage (P < 0.001), pathological N stage (P < 0.001), AJCC TRG grade (P = 0.001), pre-NCRT CEA level (P = 0.017), MPA score (P < 0.001), NLR score (P < 0.001), MLR score (P < 0.001), PNI score (P = 0.005), ALI score (P = 0.001), nerval invasion (P = 0.006), and tumor differentiation (P = 0.002) were independently associated with DFS in patients with LARC after NCRT and TME (Table 3). Results from the multivariate Cox regression model demonstrated that MPA score [hazard ratio (HR) = 0.954; 95% confidence interval (CI), 0.921–0.988; P = 0.009] and NLR level (HR = 1.058; 95% CI, 1.004–1.115; P = 0.034) were independent predictors of DFS after NCRT (Table 3).

Univariate and multivariate analysis of pCR

The score of the biomarkers in the pCR group and non-pCR group was compared by the T-test, as shown in Figure 6A. The result demonstrated that the MPA and PNI scores were significantly higher in the pCR group compared with that in the non-pCR group (MPA: pCR, 37.98 ± 9.02, vs. non-pCR, 31.86 ± 8.62, P < 0.01; PNI: pCR, 51.31 ± 5.33, vs. non-pCR, 49.05 ± 4.84, P < 0.01). However, another biomarkers score was no significant association with the pCR or non-pCR group (NLR: pCR, 2.89 ± 2.94, vs. non-pCR, 2.88 ± 3.45, P = 0.94; SII: pCR, 783.12 ± 1,090.10, vs. non-pCR, 748.26 ± 1,150.48, P = 0.81; PLR: pCR, 141.4 ± 98.22, vs. non-pCR, 154.22 ± 95.11, P = 0.35; MLR: pCR, 0.25 ± 0.12, vs. non-pCR, 0.29 ± 0.17, P = 0.07; ALI: pCR, 52.46 ± 30.4, vs. non-pCR, 47.59 ± 26.36, P = 0.30). To explore the impact of the biomarkers on pCR in patients with LARC, we performed a logistical regression model analysis. On univariate analysis, pre-NCRT cT stage (P = 0.023), MPA score (P < 0.001), pre-NCRT cN stage (P = 0.006), and PNI score (P = 0.002) were independently associated with pCR in patients with LARC (Table 4). The multivariate logistical regression model demonstrated that MPA score (OR = 0.926; 95% CI, 0.895–0.958; P < 0.001), PNI score (OR = 0.925; 95% CI, 0.871–0.983; P = 0.011), and pre-NCRT cN stage (OR = 1.634; 95% CI, 1.177–2.269; P = 0.034) were independent predictors of pCR after NCRT (Table 4).

Predictive models for DFS

The time-dependent ROC curves of the biomarkers showed that all the AUCs were relatively stable after surgery during the observation period. However, the AUC of the MPA tended to be higher than the other biomarkers at all times tested (Figure 6B). Based on the above important factors of logistics regression, a nomogram was constructed to predict DFS in LARC patients (Figure 6C). The 3-year DFS predictive probabilities were obtained by drawing a straight line after summing up the score of each variable (Figures 6D). Patients with a higher total score tended to have lower DFS rate.

RPA to identify high-risk and low-risk groups of pCR

On the basis of the results of the multivariate logistical regression analysis, RPA was performed, and patients with LARC after NCRT were divided into different pCR rate groups (Figure 7). The independent risk factors included in the RPA were MPA score, PNI score, and pre-NCRT CEA level. On the basis of the above three factors, the patients were divided into four groups. The model showed that the MPA score was the most important factor affecting pCR. When the MPA score is under 33, the pCR rate remains at 14.8%. In contrast, the MPA score of more than 33, the pCR rate was 37.7%. Moreover, we found the similar result that the PNI score is more than 46, resulting in the pCR rates of 42.7%. In addition, on the basis of the pre-NCRT CEA level, the patients with LARC were divided into two groups. Finally, the pCR rates in the low-risk group patients were 51.7%, whereas the pCR rates in the high-risk group patients were 17.7%, and the difference was significant (p < 0.001).

TABLE 3 Cox regression analysis of predictive factors for disease-free survival in patients with LARC after NCRT (n = 278).

Variables	Univariate analysis			Multivariate analysis		
	HR	95% CI	*P*-value	HR	95% CI	*P*-value
Sex, male/female	1.049	0.622–1.769	0.858			
Age	0.986	0.964–1.009	0.232			
ASA	0.978	0.508–1.883	0.948			
Postoperative hospital stay	1.002	0.949–1.057	0.947			
Distance from the anal verge	0.999	0.898–1.112	0.989			
Tumor size	0.843	0.679–1.048	0.125			
Pathological T stage	1.734	1.332–2.258	**<0.001**	1.380	0.945–2.015	0.095
Pathological N stage	2.023	1.426–2.870	**<0.001**	1.153	0.751–1.771	0.515
AJCC TRG grade	1.675	1.246–2.253	**0.001**	1.020	0.633–1.645	0.935
Interval time between NCRT and surgery	0.908	0.778–1.061	0.224			
Pre-NCRT cT stage	1.180	0.722–1.931	0.509			
Pre-NCRT cN stage	0.869	0.668–1.131	0.295			
Operative time	1.003	0.999–1.007	0.113			
Estimated blood loss	0.999	0.996–1.002	0.553			
Pre-NCRT CEA level	1.860	1.119–3.089	**0.017**	1.506	0.883–2.569	0.133
Pre-NCRT CA19-9 level	1.458	0.814–2.612	0.205			
Anemia	1.369	0.728–2.574	0.329			
Hypoproteinemia	1.660	0.665–4.144	0.278			
MPA score	0.926	0.896–0.957	**<0.001**	0.954	0.921–0.988	**0.009**
BMI	0.928	0.852–1.010	0.082			
NLR score	1.090	1.053–1.128	**<0.001**	1.058	1.004–1.115	**0.034**
SII score	1.000	1.000–1.001	0.078			
MLR score	9.954	3.015–32.856	**<0.001**	1.157	0.222–6.029	0.862
PLR score	1.001	1.000–1.003	0.099			
PNI score	0.930	0.884–0.979	**0.005**	0.990	0.975–1.004	0.156
ALI score	0.980	0.969–0.992	**0.001**	1.013	0.949–1.080	0.707
Organ preservation	1.353	0.424–4.319	0.609			
Postoperative complications	1.661	0.972–2.975	0.088			
Nerval invasion	2.594	1.316–5.113	**0.006**	1.616	0.778–3.358	0.198
Vascular invasion	1.692	0.614–4.663	0.310			
Tumor differentiation	2.374	1.356–4.157	**0.002**	1.550	0.860–2.791	0.145

LARC, locally advanced rectal cancer; NCRT, neoadjuvant chemoradiotherapy; HR, hazard ratio; CI, confidential interval; ASA, American Society of Anesthesiologists; AJCC, American Joint Committee on Cancer; CEA, carcinoembryonic antigen; CA19-9, carbohydrate antigen 19-9; NLR, neutrophil-to-lymphocyte ratio; SII, systemic immune-inflammation index; MLR, monocyte-to-lymphocyte ratio; PLR, platelet-to-lymphocyte ratio. P<0.05 was statistically significant in bold.

Discussion

The occurrence rate of obesity is increasing in the worldwide especially in China (31, 32). Recently, many studies reported that obesity contributed to developing multiple cancers and a worse prognosis (18, 33). Controversially, several studies revealed that patients with obesity had greater NCRT response when they had rectal cancer (34, 35). To assess obesity, a number of measurements are available, including BMI, waistline, and visceral adipose tissue. The above measurements are aimed to determining body fat. However, the NCRT range of irradiation for patients with LARC was limited to the pelvic and rectal. Thus, whether the body fat can instead of the pelvic and rectal fat is yet uncertain. The current study aims to evaluate the pelvic and rectal fat to predict NCRT response and prognosis in patients with LARC.

There was not standard for correctly assessing pelvic and rectal fat until now. McKechnie et al. (26) reported a better way to evaluate pelvic and rectal fat using MRI to measure the area of the fat in the pelvic. Moreover, radiomics shows multiple advantages in evaluating NCRT response in LARC (36–38). According to the NCCN and ESMO guidelines, high-resolution pelvic or rectal MRI could be an efficient routine imaging tool for evaluating clinical tumor stage and NCRT

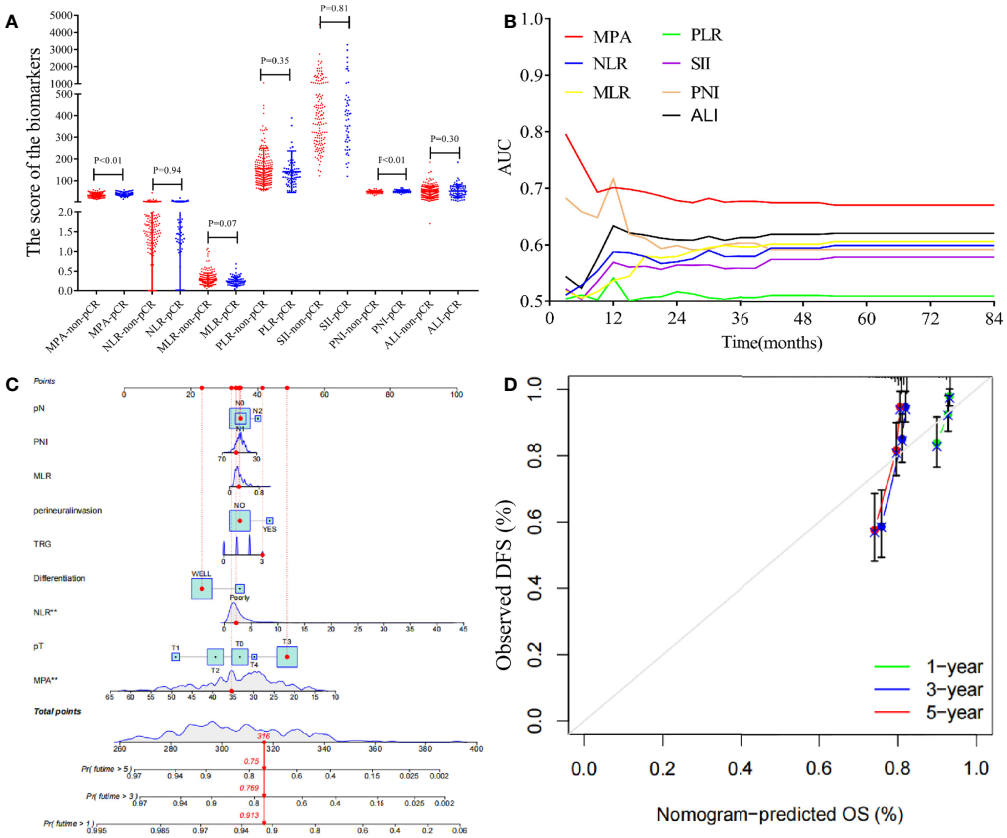


FIGURE 6
Analysis the biomarkers in the patients with LARC and construction a nomogram model for the disease-free survival. **(A)** The biomarkers value in the pCR group and non-pCR group (MPA: pCR, 37.98 ± 9.02, vs. non-pCR, 31.86 ± 8.62, P < 0.01; NLR: pCR, 2.89 ± 2.94, vs. non-pCR, 2.88 ± 3.45, P = 0.94; SII: pCR, 783.12 ± 1,090.10, vs. non-pCR, 748.26 ± 1,150.48, P = 0.81; PLR: pCR, 141.4 ± 98.22, vs. non-pCR, 154.22 ± 95.11, P = 0.35; MLR: pCR, 0.25 ± 0.12, vs. non-pCR, 0.29 ± 0.17, P = 0.07; PNI: pCR, 51.31 ± 5.33, vs. non-pCR, 49.05 ± 4.84, P < 0.01; ALI: pCR, 52.46 ± 30.4, vs. non-pCR, 47.59 ± 26.36, P = 0.30). **(B)** Time-dependent AUC curves of the NLR, SII, PLR, MLR, PNI, ALI, and MPA, for the prediction of disease-free survival. **(C)** Nomogram developed for prediction of disease-free survival. **(D)** Calibration curves for 1-, 3-, and 5-year DFS for the nomogram in patients with LARC after NCRT.

response. In the present study, we analyzed the pelvic parameters combined with radiomics, and on the basis of the LASSO analysis, the MPA was selected as the effective biomarker to predict the prognosis.

MPA includes the mesorectal and rectal thicknesses, and the rectal thickness is usually steady, so the mesorectal thickness determined the MPA score. Posterior mesorectal thickness is an important factor influencing operative complexity in rectal surgery (26, 39). Patients with the hypertrophy mesorectal might obscure anatomic dissection planes or limit access to the pelvis, potentially increasing the technical challenge of rectal surgery particularly for the patients with LARC after NCRT (40). The MPA score is connected with the BMI level and may represent pelvic obesity in the current study. Investigating the impact of the MPA score in patients with LARC may shed light on the involvement of obesity in the pelvis. On the basis of the MPA high- and low score groups, we found that the high MPA score group was associated with a low pathology TNM stage and high rates of the pCR. However, the BMI level could not distinguish the above results well. The result revealed that, rather than BMI, the pelvic obesity may contribute to the NCRT response. Moreover, the Cox regression and logistical regression also identified that the MPA score was crucial in predicting NCRT response and prognosis in patients with LARC.

Several studies found that nutrition and inflammation are related to tumor development and progression (41, 42). Obesity and albumin have been recognized as essential parameters for evaluating the nutritional status of patients with cancer (43). At present, obesity is associated with inflammatory response and

10.3389/fonc.2022.941786

TABLE 4 Logistic regression analysis of predictive factors for pCR in patients with LARC after NCRT (n = 278).

Variables	Univariate analysis			Multivariate analysis		
	OR	95% CI	*P*-value	OR	95% CI	*P*-value
Sex, male/female	0687	0.393–0.198	0.186			
Age	0.991	0.966–1.017	0.508			
ASA	0.950	0.478–1.886	0.883			
Distance from the anal verge	0.976	0.870–1.096	0.682			
Tumor size	0.964	0.775–1.199	0.741			
MPA score	0.926	0.897–0.956	**<0.001**	0.926	0.895–0.958	**<0.001**
BMI	1.001	0.917–1.093	0.978			
Interval time between NCRT and surgery	0.996	0.925–1.073	0.924			
Pre-NCRT cT stage	1.777	1.082–2.919	**0.023**	1.587	0.920–2.735	0.097
Pre-NCRT cN stage	1.524	1.129–2.057	**0.006**	1.634	1.177–2.269	**0.003**
Postoperative hospital stay	0.995	0.939–1.054	0.856			
Pre-NCRT CEA level	1.856	1.052–3.276	**0.033**	1.550	0.835–2.878	0.165
Pre-NCRT CA19-9 level	1.398	0.676–2.889	0.366			
Anemia	1.622	0.715–3.680	0.247			
Hypoproteinemia	1.367	0.374–4.993	0.636			
NLR score	0.997	0.920–1.080	0.943			
SII score	1.000	1.000–1.001	0.809			
MLR score	6.905	0.844–56.527	0.072			
PLR score	1.002	0.998–1.005	0.354			
PNI score	0.913	0.863–0.967	**0.002**	0.925	0.871–0.983	**0.011**
ALI score	0.995	0.985–1.005	0.301			
Tumor differentiation	2.308	0.930–5.729	0.071			

CI, confidential interval; ASA, American Society of Anesthesiologists; AJCC, American Joint Committee on Cancer; CEA, carcinoembryonic antigen; CA19-9, carbohydrate antigen 19-9; NLR, neutrophil-to-lymphocyte ratio; SII, systemic immune-inflammation index; MLR, monocyte-to-lymphocyte ratio; PLR, platelet-to-lymphocyte ratio. P<0.05 was statistically significant in bold.

affects the efficiency and toxicity of chemotherapy and radiotherapy in patients with cancer (44–46). The inflammatory indexes in the peripheral blood, NLR, MLR, PLR, and SII have been used as markers of predicting efficacy and toxicity of NCRT in patients with LARC in our previous study (21). Furthermore, mounting evidence suggested that obesity was contributing to the inflammatory response, which influenced tumor development, prognosis, and therapy response (18–20). As a result, we hypothesize that the pelvic fat increases the inflammatory response to affect the NCRT response. To further explore the relationship between the pelvic obesity and inflammatory response, we analyze the relationship between the MPA score and NLR, PLR, MLR, SII, PNI, and ALI. The result showed that the MPA score was associated with the NLR, PLR, and MLR score. There were more pieces of evidence that pelvic fat was related with the inflammatory response. PNI is a novel index to reflect the nutritional and inflammatory status of patients, and its clinical efficacy as a predictive factor in different malignancies has been established (47, 48). In the present study, we found that both the PNI and MPA scores were effective at predicting NCRT response in patients with LARC. However, only MPA was associated with NCRT response in logistic regression analysis. This could imply that pelvic fat modulates inflammatory response to elicit NCRT response.

To predict the pCR rates of the patients with LARC after NCRT, the RPA was performed to classify patients with LARC into different risk groups. RPA was a useful statistical method for predicting patient risk in a number of cancers, including colorectal cancer, nasopharyngeal cancer, cervical cancer, and breast cancer, which could assist clinicians to determine the best medication regimen (30, 47–49). However, few studies used the RPA to forecast the NCRT response in patients with LARC. In the present study, the MPA score, PNI score, and pre-NCRT CEA level play an important role in dividing the patients into the different risk groups. Among the affecting criteria, the MPA score has the most significant influence. On the basis of the MPA score, we distinguished over half of the patients with LARC in the first step and then selected 20% of patients as the low-risk patients, who may accept a greater pCR rates than the high-risk group, based on the PNI score and pre-NCRT CEA level. The results mentioned provide fresh treatment options for patients with LARC after NCRT.

Several limitations warrant discussion. First, the present study was subjected to potential selection bias due to the

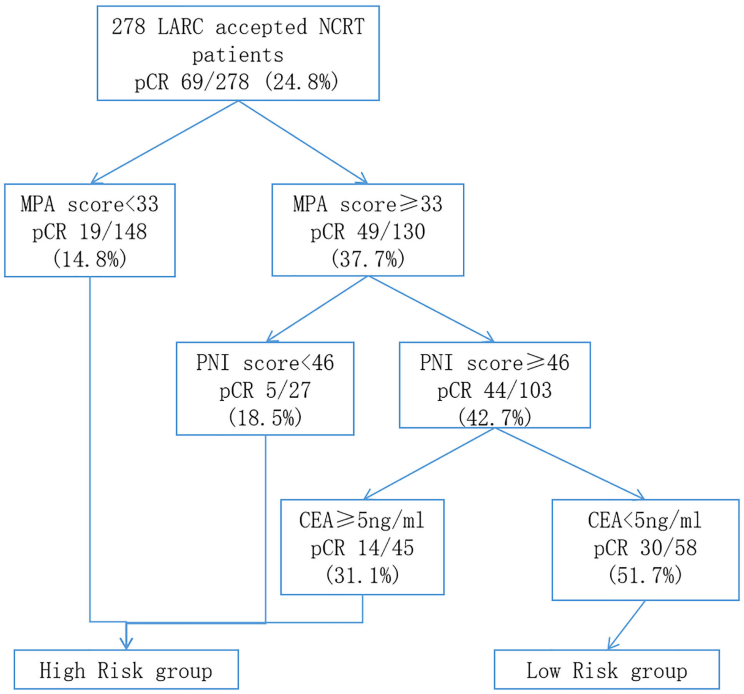


FIGURE 7
Classification tree identifying the groups at high- and low-risk for pCR of patients with LARC after NCRT. pCR, pathology complete response; LARC, locally advanced rectal cancer; NCRT, neo-chemoradiotherapy.

retrospective design. In addition, limitations in statistical methods resulted in imbalanced grouping of the groups. Second, peripheral blood cell analysis results might be affected by factors, such as blood circulation capacity, infection, and nutritional status. Third, the impact of gene profiling and tumor microenvironment inflammation was not assessed, owing to the lack of complete medical records. Despite these limitations, we believe that this study adds to the understanding of the impact of pelvic obesity on the oncological outcomes in patients with LARC after NCRT.

In conclusion, a higher MPA score was associated with poorer DFS and OS in patients with LARC after NCRT. In addition, MPA score was identified to be the most reliable marker, and the prognostic value was further confirmed by time–ROC analysis. Finally, an RPA was constructed to predict the DFS outcomes. Patients in the high-risk group who have LARC after NCRT may benefit from more intensive adjuvant therapy. Larger-scale prospective clinical trials are warranted to support the above findings.

Data availability statement

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

Ethics statement

Written informed consent was obtained from the individual(s) for the publication of any potentially identifiable images or data included in this article.

Author contributions

All authors made a significant contribution to the present study, whether that is in the conception, study design, execution, acquisition of data, analysis and interpretation, or in all these areas; took part in drafting, revising, or critically reviewing the article; gave final approval of the version to be published; have agreed on the journal to which the article has been submitted; and agree to be accountable for all aspects of the work.

Funding

This study was sponsored by Fujian provincial health technology project (2021CXA011 and 2021CXB009); National Foundation of China (No. 82172800); Science Foundation of the Fujian Province (No. 2019J01161); Special Financial Foundation of Fujian Provincial (No. 2020B019); Joint Funds for the

innovation of science and Technology, Fujian Province (2020Y9125); and the Talent programs granted from The First Affiliated Hospital of Fujian Medical University (YJRC3600).

Acknowledgments

The authors thank all the staff in the Fujian Medical University Union Hospital, Fuzhou, Fujian Province, People's Republic of China.

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

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fonc.2022.941786/full#supplementary-material

References

1. Roh MS, Colangelo LH, O'Connell MJ, Yothers G, Deutsch M, Allegra CJ, et al. Preoperative multimodality therapy improves disease-free survival in patients with carcinoma of the rectum: NSABP r-03. *J Clin Oncol* (2009) 27:5124–30. doi: 10.1200/JCO.2009.22.0467

2. Sauer R, Liersch T, Merkel S, Fietkau R, Hohenberger W, Hess C, et al. Preoperative versus postoperative chemoradiotherapy for locally advanced rectal cancer: Results of the German CAO/ARO/AIO-94 randomized phase III trial after a median follow-up of 11 years. *J Clin Oncol* (2012) 30:1926–33. doi: 10.1200/JCO.2011.40.1836

3. van Gijn W, Marijnen CA, Nagtegaal ID, Kranenbarg EM, Putter H, Wiggers T, et al. Preoperative radiotherapy combined with total mesorectal excision for resectable rectal cancer: 12-year follow-up of the multicentre, randomised controlled TME trial. *Lancet Oncol* (2011) 12:575–82. doi: 10.1016/S1470-2045(11)70097-3

4. Maas M, Nelemans PJ, Valentini V, Das P, Rödel C, Kuo LJ, et al. Long-term outcome in patients with a pathological complete response after chemoradiation for rectal cancer: a pooled analysis of individual patient data. *Lancet Oncol* (2010) 11:835–44. doi: 10.1016/S1470-2045(10)70172-8

5. Habr-Gama A, Perez RO, Proscurshim I, Campos FG, Nadalin W, Kiss D, et al. Patterns of failure and survival for nonoperative treatment of stage c0 distal rectal cancer following neoadjuvant chemoradiation therapy. *J Gastrointest Surg* (2006) 10:1319–28; discussion 1328-9. doi: 10.1016/j.gassur.2006.09.005

6. Renehan AG, Malcomson L, Emsley R, Gollins S, Maw A, Myint AS, et al. Watch-and-wait approach versus surgical resection after chemoradiotherapy for patients with rectal cancer (the OnCoRe project): A propensity-score matched cohort analysis. *Lancet Oncol* (2016) 17:174–83. doi: 10.1016/S1470-2045(15)00467-2

7. Maas M, Beets-Tan RG, Lambregts DM, Lammering G, Nelemans PJ, Engelen SM, et al. Wait-and-see policy for clinical complete responders after chemoradiation for rectal cancer. *J Clin Oncol* (2011) 29:4633–40. doi: 10.1200/JCO.2011.37.7176

8. Zhang Y, Sun Y, Xu Z, Chi P, Lu X. Is neoadjuvant chemoradiotherapy always necessary for mid/high local advanced rectal cancer: A comparative analysis after propensity score matching. *Eur J Surg Oncol* (2017) 43:1440–6. doi: 10.1016/j.ejso.2017.04.007

9. Zhang Y, Yan L, Wu Y, Xu M, Liu X, Guan G. Worse treatment response to neoadjuvant chemoradiotherapy in young patients with locally advanced rectal cancer. *BMC Cancer* (2020) 20:854. doi: 10.1186/s12885-020-07359-2

10. Calle EE, Rodriguez C, Walker-Thurmond K, Thun MJ. Overweight, obesity, and mortality from cancer in a prospectively studied cohort of U.S. adults. *N Engl J Med* (2003) 348:1625–38. doi: 10.1056/NEJMoa021423

11. Renehan AG, Tyson M, Egger M, Heller RF, Zwahlen M. Body-mass index and incidence of cancer: A systematic review and meta-analysis of prospective observational studies. *Lancet* (2008) 371:569–78. doi: 10.1016/S0140-6736(08)60269-X

12. Moghaddam AA, Woodward M, Huxley R. Obesity and risk of colorectal cancer: a meta-analysis of 31 studies with 70,000 events. *Cancer Epidemiol Biomarkers Prev* (2007) 16:2533–47. doi: 10.1158/1055-9965.EPI-07-0708

13. Ma Y, Yang Y, Wang F, Zhang P, Shi C, Zou Y, et al. Obesity and risk of colorectal cancer: A systematic review of prospective studies. *PloS One* (2013) 8:e53916. doi: 10.1371/journal.pone.0053916

14. Kang J, Baek SE, Kim T, Hur H, Min BS, Lim JS, et al. Impact of fat obesity on laparoscopic total mesorectal excision: More reliable indicator than body mass index. *Int J Colorectal Dis* (2012) 27:497–505. doi: 10.1007/s00384-011-1333-2

15. Ballian N, Yamane B, Leverson G, Harms B, Heise CP, Foley EF, et al. Body mass index does not affect postoperative morbidity and oncologic outcomes of total mesorectal excision for rectal adenocarcinoma. *Ann Surg Oncol* (2010) 17:1606–13. doi: 10.1245/s10434-010-0908-4

16. Seishima R, Okabayashi K, Hasegawa H, Sugiyama D, Ishii Y, Tsuruta M, et al. Obesity was associated with a decreased postoperative recurrence of rectal cancer in a Japanese population. *Surg Today* (2014) 44:2324–31. doi: 10.1007/s00595-014-0899-z

17. Meyerhardt JA, Tepper JE, Niedzwiecki D, Hollis DR, McCollum AD, Brady D, et al. Impact of body mass index on outcomes and treatment-related toxicity in patients with stage II and III rectal cancer: Findings from intergroup trial 0114. *J Clin Oncol* (2004) 22:648–57. doi: 10.1200/JCO.2004.07.121

18. Zhang X, Liu Y, Shao H, Zheng X. Obesity paradox in lung cancer prognosis: Evolving biological insights and clinical implications. *J Thorac Oncol* (2017) 12:1478–88. doi: 10.1016/j.jtho.2017.07.022

19. Gallo M, Adinolfi V, Barucca V, Prinzi N, Renzelli V, Barrea L, et al. Expected and paradoxical effects of obesity on cancer treatment response. *Rev Endocr Metab Disord* (2021) 22:681–702. doi: 10.1007/s11154-020-09597-y

20. Yang Z, Wei X, Pan Y, Min Z, Xu J, Yu B. Colon cancer combined with obesity indicates improved survival- research on relevant mechanism. *Aging (Albany NY)* (2020) 12:23778–94. doi: 10.18632/aging.103972

21. Zhang Y, Liu X, Xu M, Chen K, Li S, Guan G. Prognostic value of pretreatment systemic inflammatory markers in patients with locally advanced rectal cancer following neoadjuvant chemoradiotherapy. *Sci Rep* (2020) 10:8017. doi: 10.1038/s41598-020-64684-z

22. Sun Y, Zhang Y, Wu X, Lin H, Lu X, Huang Y, et al. Prognostic significance of neoadjuvant rectal score in locally advanced rectal cancer after neoadjuvant chemoradiotherapy and construction of a prediction model. *J Surg Oncol* (2018) 117:737–44. doi: 10.1002/jso.24907

23. Ryan R, Gibbons D, Hyland JM, Treanor D, White A, Mulcahy HE, et al. Pathological response following long-course neoadjuvant chemoradiotherapy for locally advanced rectal cancer. *Histopathology* (2005) 47:141–6. doi: 10.1111/j.1365-2559.2005.02176.x

24. Benson AB, Venook AP, Al-Hawary MM, Cederquist L, Chen YJ, Ciombor KK, et al. Rectal cancer, version 2.2018, NCCN clinical practice guidelines in

oncology. *J Natl Compr Canc Netw* (2018) 16:874–901. doi: 10.6004/jnccn.2018.0061

25. Parmar C, Rios Velazquez E, Leijenaar R, Jermoumi M, Carvalho S, Mak RH, et al. Robust radiomics feature quantification using semiautomatic volumetric segmentation. *PloS One* (2014) 9:e102107. doi: 10.1371/journal.pone.0102107

26. McKechnie T, Ramji K, Kruse C, Jaffer H, Rebello R, Amin N, et al. Posterior mesorectal thickness as a predictor of increased operative time in rectal cancer surgery: A retrospective cohort study. *Surg Endosc* (2022) 36:3520–32. doi: 10.1007/s00464-021-08674-w

27. Camp RL, Dolled-Filhart M, Rimm DL. X-Tile: a new bio-informatics tool for biomarker assessment and outcome-based cut-point optimization. *Clin Cancer Res* (2004) 10:7252–9. doi: 10.1158/1078-0432.CCR-04-0713

28. Zhang Y, Wang Y, Liu X, Chen B, Zhuang J, Li S, et al. Worse prognosis in young patients with locally advanced rectal cancer following neoadjuvant chemoradiotherapy: A comparative study. *Med (Baltimore)* (2020) 99:e21304. doi: 10.1097/MD.0000000000021304

29. Zhang Y, Xu M, Sun Y, Chen Y, Chi P, Xu Z, et al. Identification of LncRNAs associated with FOLFOX chemoresistance in mCRC and construction of a predictive model. *Front Cell Dev Biol* (2020) 8:609832. doi: 10.3389/fcell.2020.609832

30. Wang Y, Zhang Y, Lin H, Xu M, Zhou X, Zhuang J, et al. Risk factors for lymph node metastasis in rectal neuroendocrine tumors: A recursive partitioning analysis based on multicenter data. *J Surg Oncol* (2021) 124:1098–105. doi: 10.1002/jso.26615

31. Kadam I, Neupane S, Wei J, Fullington LA, Li T, An R, et al. A systematic review of diet quality index and obesity among Chinese adults. *Nutrients* (2021) 13 (10):3555. doi: 10.3390/nu13103555

32. Zheng R, Li M, Xu M, Lu J, Wang T, Dai M, et al. Chinese Adults are more susceptible to effects of overall obesity and fat distribution on cardiometabolic risk factors. *J Clin Endocrinol Metab* (2021) 106:e2775–2775e2788. doi: 10.1210/clinem/dgab049

33. Cacho-Díaz B, Spínola-Maroño H, Reynoso N, González-Aguilar A, Mohar-Betancourt A. Role of overweight, obesity, and comorbidities in the prognosis of patients with breast cancer with brain metastases. *Clin Breast Cancer* (2019) 19: e394–394e398. doi: 10.1016/j.clbc.2018.12.018

34. Sun Y, Xu Z, Lin H, Lu X, Huang Y, Huang S, et al. Impact of body mass index on treatment outcome of neoadjuvant chemoradiotherapy in locally advanced rectal cancer. *Eur J Surg Oncol* (2017) 43:1828–34. doi: 10.1016/j.ejso.2017.07.022

35. Sun Y, Chi P. Impact of body mass index on surgical and oncological outcomes in laparoscopic total mesorectal excision for locally advanced rectal cancer after neoadjuvant 5-Fluorouracil-Based chemoradiotherapy. *Gastroenterol Res Pract* (2017) 2017:1509140. doi: 10.1155/2017/1509140

36. Liu Z, Zhang XY, Shi YJ, Wang L, Zhu HT, Tang Z, et al. Radiomics analysis for evaluation of pathological complete response to neoadjuvant chemoradiotherapy in locally advanced rectal cancer. *Clin Cancer Res* (2017) 23:7253–62. doi: 10.1158/1078-0432.CCR-17-1038

37. Nie K, Shi L, Chen Q, Hu X, Jabbour SK, Yue N, et al. Rectal cancer: Assessment of neoadjuvant chemoradiation outcome based on radiomics of multiparametric MRI. *Clin Cancer Res* (2016) 22:5256–64. doi: 10.1158/1078-0432.CCR-15-2997

38. Cui Y, Yang X, Shi Z, Yang Z, Du X, Zhao Z, et al. Radiomics analysis of multiparametric MRI for prediction of pathological complete response to neoadjuvant chemoradiotherapy in locally advanced rectal cancer. *Eur Radiol* (2019) 29:1211–20. doi: 10.1007/s00330-018-5683-9

39. Kim YW, Cha SW, Pyo J, Kim NK, Min BS, Kim MJ, et al. Factors related to preoperative assessment of the circumferential resection margin and the extent of mesorectal invasion by magnetic resonance imaging in rectal cancer: A prospective comparison study. *World J Surg* (2009) 33:1952–60. doi: 10.1007/s00268-009-0126-z

40. Fang Y, Sheng C, Ding F, Zhao W, Guan G, Liu X. Adding consolidation capecitabine to neoadjuvant chemoradiotherapy for locally advanced rectal cancer: A propensity-matched comparative study. *Front Surg* (2021) 8:770767. doi: 10.3389/fsurg.2021.770767

41. Diakos CI, Charles KA, McMillan DC, Clarke SJ. Cancer-related inflammation and treatment effectiveness. *Lancet Oncol* (2014) 15:e493–503. doi: 10.1016/S1470-2045(14)70263-3

42. Ravasco P. Nutrition in cancer patients. *J Clin Med* (2019) 8(8):1211. doi: 10.3390/jcm8081211

43. Mantzorou M, Koutelidakis A, Theocharis S, Giaginis C. Clinical value of nutritional status in cancer: What is its impact and how it affects disease progression and prognosis. *Nutr Cancer* (2017) 69:1151–76. doi: 10.1080/01635581.2017.1367947

44. Del Cornò M, Varì R, Scazzocchio B, Varano B, Masella R, Conti L. Dietary fatty acids at the crossroad between obesity and colorectal cancer: Fine regulators of adipose tissue homeostasis and immune response. *Cells* (2021) 10(7):1738. doi: 10.3390/cells10071738

45. Boi SK, Orlandella RM, Gibson JT, Turbitt WJ, Wald G, Thomas L, et al. Obesity diminishes response to PD-1-based immunotherapies in renal cancer. *J Immunother Cancer* (2020) 8(2):e000725. doi: 10.1136/jitc-2020-000725

46. Assumpção J, Pasquarelli-do-Nascimento G, Duarte M, Bonamino MH, Magalhães KG. The ambiguous role of obesity in oncology by promoting cancer but boosting antitumor immunotherapy. *J BioMed Sci* (2022) 29:12. doi: 10.1186/s12929-022-00796-0

47. Zhang Q, Chen J, Yu X, Ma J, Cai G, Yang Z, et al. Systemic treatment after whole-brain radiotherapy may improve survival in RPA class II/III breast cancer patients with brain metastasis. *J Neurooncol* (2013) 114:181–9. doi: 10.1007/s11060-013-1169-4

48. Xu M, Xie X, Cai L, Xie Y, Gao Q, Sun P. Risk factor assessment of lymph node metastasis in patients with FIGO stage IB1 cervical cancer. *Front Oncol* (2022) 12:809159. doi: 10.3389/fonc.2022.809159

49. Yang ZC, Hu YY, Liu LT, Guo SS, Du CC, Liang YJ, et al. Determining the suitability of definitive radiation therapy in patients with metastatic nasopharyngeal carcinoma based on PET/CT: A large cohort study. *Eur Radiol* (2022). doi: 10.1007/s00330-022-08814-3

frontiers | Frontiers in Oncology

TYPE Original Research
PUBLISHED 04 October 2022
DOI 10.3389/fonc.2022.937277

OPEN ACCESS

EDITED BY
Pei Yang,
Hunan Cancer Hospital, Central South University, China

REVIEWED BY
Liyue Shen,
Stanford University, United States
Lina Zhao,
Fourth Military Medical University, China
Yao Ding,
University of Texas MD Anderson Cancer Center, United States

*CORRESPONDENCE
Hai Jun Wu
wuhaijun@csu.edu.cn

†These authors have contributed equally to this work and share first authorship

SPECIALTY SECTION
This article was submitted to Cancer Imaging and Image-directed Interventions, a section of the journal Frontiers in Oncology

RECEIVED 06 May 2022
ACCEPTED 30 August 2022
PUBLISHED 04 October 2022

CITATION
Lin Q, Wu HJ, Song QS and Tang YK (2022) CT-based radiomics in predicting pathological response in non-small cell lung cancer patients receiving neoadjuvant immunotherapy.
Front. Oncol. 12:937277.
doi: 10.3389/fonc.2022.937277

CT-based radiomics in predicting pathological response in non-small cell lung cancer patients receiving neoadjuvant immunotherapy

Qian Lin†, Hai Jun Wu*, Qi Shi Song† and Yu Kai Tang†

Department of Oncology, Xiangya Hospital, Central South University, Changsha, China

Objectives: In radiomics, high-throughput algorithms extract objective quantitative features from medical images. In this study, we evaluated CT-based radiomics features, clinical features, in-depth learning features, and a combination of features for predicting a good pathological response (GPR) in non-small cell lung cancer (NSCLC) patients receiving immunotherapy-based neoadjuvant therapy (NAT).

Materials and methods: We reviewed 62 patients with NSCLC who received surgery after immunotherapy-based NAT and collected clinicopathological data and CT images before and after immunotherapy-based NAT. A series of image preprocessing was carried out on CT scanning images: tumor segmentation, conventional radiomics feature extraction, deep learning feature extraction, and normalization. Spearman correlation coefficient, principal component analysis (PCA), and least absolute shrinkage and selection operator (LASSO) were used to screen features. The pretreatment traditional radiomics combined with clinical characteristics (before_rad_cil) model and pretreatment deep learning characteristics (before_dl) model were constructed according to the data collected before treatment. The data collected after NAT created the after_rad_cil model and after_dl model. The entire model was jointly constructed by all clinical features, conventional radiomics features, and deep learning features before and after neoadjuvant treatment. Finally, according to the data obtained before and after treatment, the before_nomogram and after_nomogram were constructed.

Results: In the before_rad_cil model, four traditional radiomics features ("original_shape_flatness," "wavelet hhl_firer_skewness," "wavelet hlh_firer_skewness," and "wavelet lll_glcm_correlation") and two clinical features ("gender" and "N stage") were screened out to predict a GPR. The average prediction accuracy (ACC) after modeling with k-nearest neighbor (KNN) was 0.707. In the after_rad_cil model, nine features predictive of GPR were obtained after feature screening, among which seven were traditional radiomics features: "exponential_firer_skewness," "exponential_glrlm_runentropy," "log- sigma-5-0-mm-3d_firer_kurtosis," "logarithm_skewness," "original_shape_elongation,"

"original_shape_brilliance," and "wavelet llh_glcm_clustershade"; two were clinical features: "after_CRP" and "after lymphocyte percentage." The ACC after modeling with support vector machine (SVM) was 0.682. The before_dl model and after_dl model were modeled by SVM, and the ACC was 0.629 and 0.603, respectively. After feature screening, the entire model was constructed by multilayer perceptron (MLP), and the ACC of the GPR was the highest, 0.805. The calibration curve showed that the predictions of the GPR by the before_nomogram and after_nomogram were in consensus with the actual GPR.

Conclusion: CT-based radiomics has a good predictive ability for a GPR in NSCLC patients receiving immunotherapy-based NAT. Among the radiomics features combined with the clinicopathological information model, deep learning feature model, and the entire model, the entire model had the highest prediction accuracy.

KEYWORDS

radiomics, pathological response, NSCLC, biomarkers, lung cancer, immunotherapy, neoadjuvant therapy

Introduction

Lung cancer is a heterogeneous malignant disease arising from the bronchial epithelium or alveolar tissue, usually caused by smoking, varying environmental exposures, and underlying genetic susceptibility (1). According to the 2020 global burden of cancer statistics provided by the International Agency for Research on Cancer (IARC), lung cancer ranks second in global incidence and first in mortality. In China, lung cancer is the most common cancer with the highest incidence and mortality rate. The most common subtype is non-small cell lung cancer (NSCLC) with an incidence rate of about 85%, while the 5-year survival rate is only 10%–20% (2, 3).

Currently, surgical treatment remains the mainstay of treatment for early-stage and locally advanced (stages I and II and some with stages IIIA and IIIB) NSCLC(NCCN) (4, 5). However, patients experience high rates of local and distant recurrence postoperatively, suggesting that systemic therapy is necessary to improve cure rates. Neoadjuvant therapy (NAT) is a form of cancer treatment and refers to systemic therapy given before surgery, including neoadjuvant chemotherapy, chemoradiotherapy, targeted therapy, and immunotherapy. However, the technical definition of NAT usually refers exclusively to neoadjuvant chemotherapy and is distinguished from adjuvant chemotherapy after surgery. NAT (6) reduces the rate of distant disease recurrence by taking advantage of the damaged lymphatics and vasculature resulting from surgery, thereby increasing local drug concentration. Effective antitumor therapy can shrink the primary lesion and downstage the tumor stage, reducing the need for extensive surgery leading to organ preservation and improved quality of life. Moreover, patients are generally in a better situation and less likely to experience acute toxicity before surgery, and currently, receiving systemic therapy is more well-tolerated. By observing the radiological and pathological responses following NAT, tumor sensitivity to chemotherapeutic drugs can be understood, which provides a reference for the choice of the postoperative treatment regimen. Effective NAT can minimize the proliferative capacity of tumors at the time of surgery and reduce the risk of intraoperative dissemination of cancer cells, and even a small proportion of patients can experience major pathological response (MPR) and complete pathological response (CPR) (7).

Lately, with the wide use of immune checkpoint inhibitors (ICIs) such as anti-programmed death-ligand-1 (PD-L1) and anti-programmed death-1 (PD-1) antibodies in advanced NSCLC, patients have had significantly improved quality of life and good prognosis, making immunotherapy (8–10) a new option for the treatment of resectable and potentially resectable NSCLC. In previous clinical studies, either single-agent ICI or immune doublet combination NAT significantly led to higher MPR and CPR rates and lower complication rates than neoadjuvant chemotherapy (11, 12).

In an exciting phase II study of ICI with chemotherapy, the MPR and CPR rates were 83% and 63%, respectively, with 90% of patients who underwent resection achieving clinical stage (13). Although the large-scale phase III study is still ongoing from the results of the abovementioned phase II studies, patients treated with neoadjuvant ICI and then surgery had a similar adverse event as compared to chemotherapy combination but

with better pathological remission (residual tumor cells in tumor bed ≤10%) and improved quality of life (14).

In general, the higher the MPR/CPR rate, the better efficacy of NAT (15, 16). If the pathological response information can be evaluated before surgical resection, it will guide the type of surgery. Therefore, developing noninvasive assessment response models can help identify patients with good responses who may benefit from local excision. Those who achieve CPR may benefit from the watch-and-wait or nonsurgical strategies.

Imaging modalities such as computed tomography (CT), magnetic resonance imaging (MRI), and positron emission tomography (PET) have become routine in the clinical management of patients with tumors such as lung cancer. Lesions detected through these imaging modalities are described and analyzed only based on simple qualitative (e.g., shape, location, spiculated lesion, and lobulation) and quantitative (e.g., size, volume, density, signal, and standardized uptake values) features. The radiological diagnostic accuracy is closely related to the radiologists' experience, with marked subjective differences. In 2012, Dutch investigators (17) first proposed the concept of radiomics, hypothesizing that the cellular and molecular heterogeneity of tumor cells can be reflected by quantitative imaging microheterogeneity. They also concluded that when extracted with radiomics, these features can transform image data of interest regions from medical imaging into quantitative data through high-throughput algorithms. The general procedure of radiomics is as follows: 1) Acquisition of medical imaging data; 2) Region of interest (ROI) segmentation and feature extraction; 3) Feature selection, model building, and validation; 4) Statistical data analysis. Since then, the concept of radiomics has been widely studied in the differentiation of benign and malignant lesions (18), in the preoperative prediction of lymph node metastasis in lung cancer (19), and in the assessment of the mutational status of genes such as Epidermal Growth Factor Receptor (EGFR) (20) and anaplastic lymphoma kinase (ALK) (21). More radiomics studies include the prediction of treatment effects and prognosis in cancer (22, 23) and even in non-neoplastic diseases such as the early diagnosis of Alzheimer's disease (24) and the rapid radiological diagnosis of coronavirus disease 2019 (COVID-19) pneumonia (25, 26).

With the development of computer software, computational power has significantly improved, and in recent years, artificial intelligence (AI) technology based on deep learning (DL) algorithms has been vigorously developed and has gradually begun to be applied in medical research. Currently, it is mainly based on medical images using computer vision technology to solve clinical tasks such as lesion segmentation and disease classification (27–29). In the processing of medical images, the most widely used DL network is the convolutional neural network (CNN). A CNN is a computational method for learning relevant features from image signal intensities proposed based on the working principle of the human nervous system. The ability to directly utilize high-dimensional numerical information in images from a large enough number of training data and identify image features with a high degree of representativeness creates and selects a large amount of abstract information at the hidden layer, which is defined as DL. DL features can be used more comprehensively to accomplish segmentation, classification, and other targeted tasks (30).

Building entirely new DL models requires large amounts of annotated data to be used as training data; however, the sample size of most radiological data is often limited. The method of using a trained DL model on other data sets for a target data set and for extracting the features of the target data set is called transfer learning. Transfer learning offers the possibility of DL for small samples of medical data. Many studies have confirmed that this is an effective and superior way to conventional machine learning (27, 31).

This study aims to extract radiomics features and DL features from CT images of patients with NSCLC before NAT with ICIs and after NAT with chemotherapy, then combine the features with clinicopathological information of patients.

Combining clinicopathological feature signature and DL feature signature that could predict pathological remission after NAT with ICIs in NSCLC patients was done through feature screening. Also, binary logistic regression analysis constructed a prediction model integrating traditional radiomics feature labels, DL feature labels, and clinicopathological information.

Finally, a nomogram was constructed to visualize the model, achieving a precise assessment of pathological remission after NAT with ICIs in NSCLC, thereby providing an adjuvant tool for developing individualized treatment regimens for patients.

Materials and methods

Study participants

Clinicopathological information was retrospectively collected from 83 patients with pathologically confirmed NSCLC and treated with immunotherapy-based NAT between 1 March 2020 and 1 January 2022.

Inclusion criteria include patients with pathologically diagnosed NSCLC through either image-guided biopsy or bronchoscopy-directed biopsy; potentially operable stage Ib–III NSCLC as per the 2017 Union for International Cancer Control (UICC)/American Joint Committee on Cancer (AJCC) Eighth Edition; no history of other tumors or other antitumor therapy; patients who received immunotherapy-based NAT with planned surgery; patients with at least 2 chest CT results available performed within 2 weeks (up to a maximum of 1 month) before immunotherapy-based NAT and 1 week (up to a maximum of half a month) before surgery.

Exclusion criteria include CT scans not done in our hospital or outside the study timeline (n = 11); Surgical cases done outside our hospital (n = 10).

Finally, the data of 62 patients were included in the study. They then were randomly divided into the training group (for the establishment of the radiomics label and model) and a validation group (for the verification of the radiomics label and model) at a ratio of 7:3 or 8:2.

Pathology

Preoperative pathological information

The preoperative pathology was mainly determined by lung biopsy or bronchoscopy biopsy, and cases biopsied outside our hospital were all reconfirmed by our pathologists.

Preoperative pathologic information included common immunohistochemistry, genetic testing, and PD-L1 testing, Ki67 (percentage), chromogranin A (CgA) (negative/positive), Syn (negative/positive), p63 (strong/moderate/weak), cancer embroyonic antigen (CEA) (negative/positive), thyroid transcription factor 1 (TTF-1) (negative/positive), P40 (negative/positive), p53 (negative/positive), napsin-a (negative/positive), cytokeratin 5/6 (CK5/6) (negative/positive), cytokeratin 7 (CK7) (negative/positive), pan-cytokeratin (CK-Pan) (strong/moderate/weak), ALK control X3/echinoderm microtubule-associated protein-like 4 (EML4-ALK) (Ventana) (negative/positive), gene mutation (negative/positive), PD-L1 (22c3)-Ventana (total positive score) (negative/positive).

Postoperative pathological information

The postoperative pathology was confirmed from surgical resection samples after NAT immunotherapy, and the efficacy of NAT in NSCLC was evaluated by the pathologists of our hospital depending on tumor bed, lymph node tumor remission, and residual disease according to an expert consensus issued by the expert committee on lung cancer quality control at the National Cancer Quality Control Center (NCQCC) with the following criteria: MPR as viable tumor cell residual ≤10% from the tumor bed, CPR as no viable tumor cells remaining in the tumor bed and lymph nodes, and partial pathological response (PPR) as >10% viable tumor cells remaining in the tumor bed (14).

Collected postoperative pathology information, including PPR, MPR, and CPR, was defined as a good and poor pathological response (CPR, MPR) and poor pathological response (PPR).

Patient clinical data

Clinical information was derived from the electronic medical record and was cross-checked by two independent investigators.

General clinical information collected includes age (years), gender (men/women), smoking status (yes/no), tumor differentiation grade (low/intermediate/well-differentiated), tumor type (squamous/adenocarcinoma), T stage (stage I/II/III/IV), N stage (stage I/II), clinical Tumor, Node, and Metastasis (TNM) stage (stage Ib–III), smoking history (yes/no), family history of tumor (yes/no), history of chronic comorbities (yes/no), height (m), number of immunotherapy cycles (two cycles/three cycles/four cycles), and immunodrugs (domestic/imported).

Clinical information before and after treatment was also collected, including body weight (kg), body mass index (kg/m^2), lactate dehydrogenase (high/low), albumin (g/L), C-reactive protein (mg/L), white blood cells (10^9/L), percentage of lymphocytes (%), tumor markers (normal/abnormal), thyroid function (normal/abnormal), T helper/induced T cells (%), and inhibitory T cells/cytotoxic T cells (%).

Processing of missing data

The filling method for missing data is called data interpolation, which can be divided into the single filling and multiple filling methods. The single imputation method only yields one set of imputation results for an incomplete data set and greatly impacts the data distribution; for example, the mean filling method is to fill in missing values using the mean of all available data. The multiple imputation method uses the existing data of the incomplete data set to fill the missing value at any time to generate multiple complete data sets.

Random forest-based chained equations with multiple imputations (MICEforest, multiple imputations based on forest by chained equations) enable the generation of multiple groups of data with mean and variance that are all like the original data set according to the method of random forest, with imputation completed by comparing a selected group of data with the smallest difference from the original data set. Technically, any predictive model can be used for MICEforest.

This study deleted categories with missing values greater than 25%, and MICEforest multiple interpolation methods were used to fill the missing values. Category “thyroid function” was also removed because of the large difference between the data after filling and the original value.

Processing of clinicopathological information

After deleting items with a lot of missing data and multiple imputing items with a small amount of missing data, the final remaining clinicopathological data used in this study were used to compare the distribution of variables in the two groups of good pathological response (GPR) vs. bad pathological response

(BPR) using the R language program. The chi-square test was used for categorical variables, the Wilcoxon rank and t-test were used for continuous variables, and a p-value <0.05 was considered statistically significant.

CT scan protocol

Patients were instructed to hold their breath after deep inspiration and complete the scan with one breath-hold using one of three CT scanners: Toshiba Aquilion ONE (Toshiba Medical System Corporation, Japan), Siemens SOMATOM Drive (Siemens Medical System Co., Ltd., Germany), and GE Revolution (General Electric Medical System Co., Ltd., America) with the scan ranging from at least the thoracic inlet to the level of the costophrenic angle, including the whole lung. Scanning parameters of different CT scanners are shown in Table 1.

Tumor segmentation

From the picture archiving and communication system (PACS), the CT images of each patient containing at least a lung window and mediastinal window were exported, and the enhanced CT was also exported synchronously if it existed. All cases were performed with the open-source software ITK-SNAP (version 3.8.0, http://www.itk-snap.org) in high-resolution lung windows (window width 1,500–2,000 Hu, window position −450 to −600 Hu). In all CT images containing tumor lesions, the ROI was manually outlined along the contour of the lesion layer by layer to try to keep the ROI containing only the entire tumor and does not contain other distinguishable tissues, such as air and obvious blood vessels. If there is a simultaneously enhanced CT or PET, it will be compared layer by layer to make it as accurate as possible. The final ROI is shown in Figure 1.

Radiomics conventional feature extraction and model construction

The manually drawn ROI was used to extract the traditional quantitative features of each patient using the Pyradiomics package (version 3.0.1, https://pyradiomics.readthedocs.io) in the Python program (version 3.6.13, https://www.python.org). The extracted radiomics traditional features included first-order features, two-dimensional (2D) shape features, 3D shape features, gray-level size zone matrix (GLSZM) features, gray-level co-occurrence matrix (GLCM) features, gray-level dependence matrix (GLDM) features, gray-level run-length matrix (GLRLM) features, and wavelet transform features. The extracted traditional features of radiomics and the clinicopathological information after deletion and imputation according to whether they were acquired before neoadjuvant immunotherapy or acquired after NAT were used to construct models separately. CT images acquired before NAT were extracted for radiomics; traditional quantitative features named before_rad_data, together with clinicopathological information that was available immediately before NAT, were used to construct the model, which is referred to as before_rad_cil. After 2–4 cycles of NAT, CT images obtained before surgery were used to extract radiomics quantitative features named after_rad_ data. Combined with the clinical information after NAT, a model named after_ rad_cil was constructed.

Before modeling, the data are transformed into structured data with 0 mean 1 variance by standardization for subsequent processing to eliminate the differences of different eigenvalues on

TABLE 1 Scanning parameters of the different CT scanners.

	Toshiba Aquilion ONE	Siemens SOMATOM Drive	GE Revolution
tube current	automatic tube current	80-350mA automatic tube current	80-350mA automatic tube current
tube voltage	120 kV	120 kV	120 kV
FOV	320.3	307	\
construction algorithm	standard algorithm, lung algorithm and soft tissue algorithm	standard algorithm, lung algorithm and soft tissue algorithm	standard algorithm, lung algorithm and soft tissue algorithm
slice thickness	1mm	1mm	1mm
slice separation	0.8mm	1mm	1mm
matrix	512×512	512×512	512×512
construction slice separation	1mm	1mm	1
Construction slice thickness	1mm	1mm	1
revolution speed	\	\	158.75mm/s
detector width	\	\	80mm

FOV, field of view.

FIGURE 1
The final ROI. **(A)** Segmentation results of the cross section. **(B)** Segmentation results of the sagittal plane. **(C)** Three-dimensional (3D) visualization effect of the tumor area. **(D)** Segmentation results of the coronal plane. ROI, region of interest.

the scale. After data normalization, features were screened by the Spearman correlation coefficient, and only one feature was retained in features with a high correlation (correlation coefficient >0.9).

For the filtered features, the least absolute shrinkage and selection operator (LASSO) was used to select features to construct the LASSO equation and calculate the feature weights, respectively. These features with a feature coefficient >0 were randomly divided into training and test sets in a ratio of 7:3 and then modeled with one of eight machine learning algorithms: support vector machine (SVM), k-nearest neighbor (KNN), decision tree (DecisionTree), random forest (RandomForest), extreme gradient boosting (XGBoost), multilayer perceptron (MLP), extremely randomized trees (ExtraTrees), and light gradient boosting machine (LightGBM), And to compare the predictive accuracy of each model. The accuracy, area under the curve (AUC), sensitivity, and specificity of 5-fold cross-validation after random grouping were used as evaluation indexes. Finally, the model results with the best prediction efficiency after 100 random groupings were selected to construct the pretreatment radiomics features combined with the clinicopathological feature label (before_rad_cil_signature) and the posttreatment radiomics features combined with the clinicopathological feature label (after_rad_cil_signature). The flow of the ROI delineation, conventional feature extraction, data analysis, model building, and comparison is shown in Figure 2.

Tools used in combined clinicopathological feature and radiomics feature screening and model building were as follows: Python program (version 3.6.13, https://www.python.org/) and the packages of scikit-learn (version 1.0, https://scikit-learn.org) and pandas (version 1.1.5, https://pandas.pydata.org).

Radiomics deep learning feature extraction and model construction

Usually, training a DL model requires large amounts of annotated data and an excellent performance hardware platform. Due to this study's limited sample size and hardware platform, transfer learning and fine-tuning are used to overcome this limitation. Transfer learning can transfer knowledge learned from previous tasks to new tasks and avoid retraining new tasks to improve the learning efficiency of new tasks. Fine-tuning is usually used with transfer learning.

The pretrained model constructed by previous tasks is used to learn the task data to avoid the large amounts of human, computational, material, and financial resources required to retrain the model while ensuring its effectiveness. The pretrained model used in our study was resnet50, and the model network was pretrained on Imagenet to determine all network parameters through parameter adjustment and optimization.

The CT images for fine-tuning and DL feature extraction in this study were obtained by intercepting the tumor area at the layer of the maximum ROI on the cross section. The steps of DL feature extraction are shown in Figure 3: 1) Manually outline the

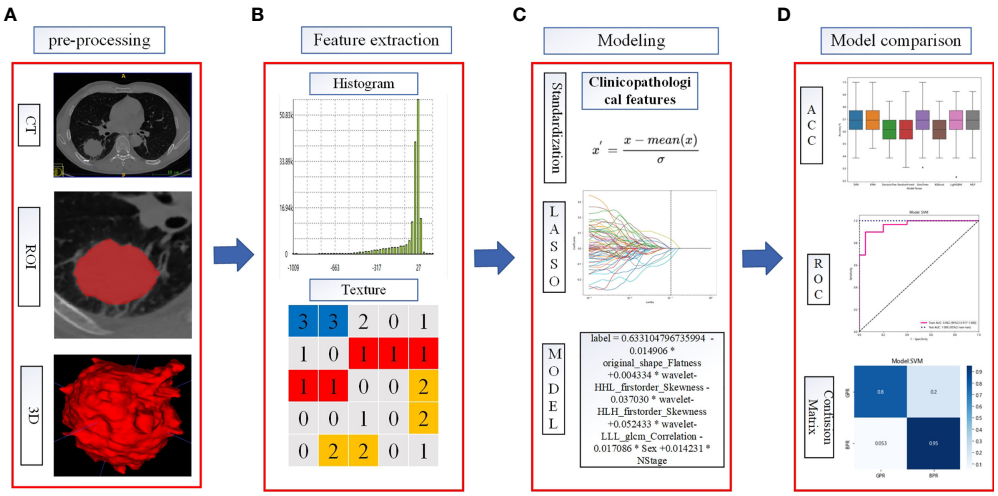


FIGURE 2
Flowchart of radiomics analysis. **(A)** The ROI was manually segmented on CT images before and after neoadjuvant therapy. **(B)** Quantitative features in the ROI were calculated, including features such as shape, texture, and wavelet filtering. **(C)** Combined clinicopathological features, standardized processing of data, feature filtering, and model building. **(D)** To evaluate the predictive model efficacy, the evaluation indexes of each model were compared to select the better performing model to construct the before_rad_cil_signature and after_rad_cil_signature. *multiply.

ROI layer by layer using ITK-SNAP software. 2) The CT 3D image containing the ROI was automatically read using software written by ourselves, identifying the level at which the largest ROI was located in the cross section and clipping out the CT image at which the tumor was located with a rectangular box. 3) The cropped CT pictures were randomly split into training and validation groups in an 8:2 ratio, fine-tuned using the pretrained model resnet50. 4) Select the most accurate grouping model for

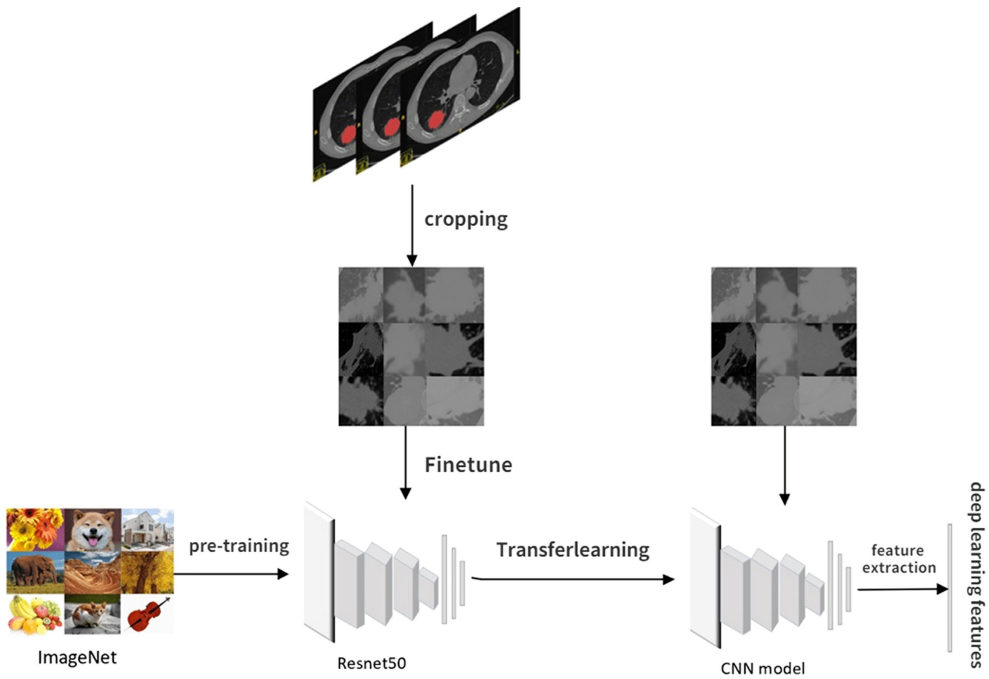


FIGURE 3
Flowchart of deep learning feature extraction.

DL feature extraction. 5) The output of DL features from the last layer before the full connection layer is usually the most concise. It is also the general DL feature selection layer in the industry. Therefore, our study selected the data of the last layer before the full connection layer, avgpool, as the DL features for subsequent studies, with a total of 2,048 features. 6) Like the above steps, the CT images obtained before and after NAT were separately subjected to DL feature extraction.

According to the CT images obtained before NAT, the extracted DL feature is named before_dl_data. According to the CT images obtained before surgery after 2–4 cycles of NAT, the DL feature is named after_dl_data.

Before modeling, the above data are converted into structured data with 0 mean 1 variance by standardization for subsequent processing. The data dimension is reduced to 62 dimensions by principal component analysis (PCA) after data standardization. The data set after dimension reduction was randomly divided into a training set and a test set in an 8:2 ratio. Then, this was modeled with one of eight machine learning algorithms: SVM, KNN, DecisionTree, RandomForest, XGBoost, MLP, ExtraTrees, and LightGBM, And to compare the predictive accuracy of each model. The accuracy, AUC, sensitivity, and specificity of 5-fold cross-validation after random grouping were used as evaluation indexes. Finally, the model results with the best prediction efficiency after 100 random groupings were selected to construct the pretreatment DL label (before_dl_signature) and the posttreatment DL label (after_dl_signature).

Tools used in DL feature extraction and model building were as follows: ITK-SNAP (version 3.8.0, http://www.itk-snap.org), Python program (version 3.6.13, https://www.python.org), and packages: Pytorch (version 1.9.0, https://pytorch.org), scikit-learn (version 1.0, https://scikit-learn.org), and pandas (version 1.1.5, https://pandas.pydata.org).

Combined model construction

After deletion and missing value processing, all previously extracted radiomics traditional quantitative features, radiomics DL features, and clinicopathological features were combined into a 7,421-dimensional joint data set. Before modeling, the joint data were transformed into structured data with 0 mean 1 variance by standardization to eliminate the scale difference of different eigenvalues. After data normalization, features were screened by the Spearman correlation coefficient, and only one was retained in features with a high correlation (correlation coefficient >0.9). LASSO further screens the selected features, and the LASSO equation is constructed for the features with feature coefficient >0, and the feature weights are calculated. These features with feature coefficient >0 were randomly divided into training and test sets in a ratio of 7:3. Then, these were modeled with one of eight machine learning algorithms: SVM, KNN, DecisionTree, RandomForest, XGBoost, MLP, ExtraTrees, LightGBM, And to compare the predictive accuracy of each model. The accuracy, AUC, sensitivity, and specificity of 5-fold cross-validation after random grouping were used as evaluation indexes. Finally, the model results with the best prediction efficiency after 100 random groupings were selected to construct the entire feature label (entire_signature).

Tools used in the entire feature screening and model building were as follows: Python program (version 3.6.13, https://www.python.org) and packages: scikit-learn (version 1.0, https://scikit-learn.org) and pandas (version 1.1.5, https://pandas.pydata.org).

Nomogram construction

The nomogram provides a simple graphical presentation of a clinical prediction model, allowing calculation of the probability of a certain target event based on individualized information for the patient. Its simple graphical interface promotes the wide application of nomograms. Our study used previously constructed before_rad_clinic_signature, before_dl_signature, and clinical features screened by LASSO to construct a pre-NAT nomogram (before_nomogram); before_rad_clinic_signature, before_dl_signature, after_rad_clinic_signature, after_dl_signature, and entire_signature were used to jointly construct a posttreatment nomogram. The procedure of nomogram construction is shown in Figure 4. Both nomograms were based on logistic regression models by plotting calibration curves to compare predicted vs. actual outcome events.

Tools used in nomogram construction were as follows: the R program (version 3.6.1, https://www.r-project.org) based on the RMS package (version 5.1-3.1).

Results

Clinicopathological information of patients

The categories, totals, and missing numbers of clinicopathological information collected are shown in Tables 2 and 3. The categories with missing values greater than 30% were CK, Caudal-related homeobox transcription factor-2 (CDX-2), CEA, p53, CD56, CgA, Syn, p63, CK-Pan, Ki67, CK7, PD-L1 expression, Driving gene mutation, p40, Napsin-A, CK5/6, TTF-1, after_assisted/induced T, after_inhibited/cytotoxic T, before_inhibited/cytotoxic T, and before_assisted/induced T. These categories were deleted.

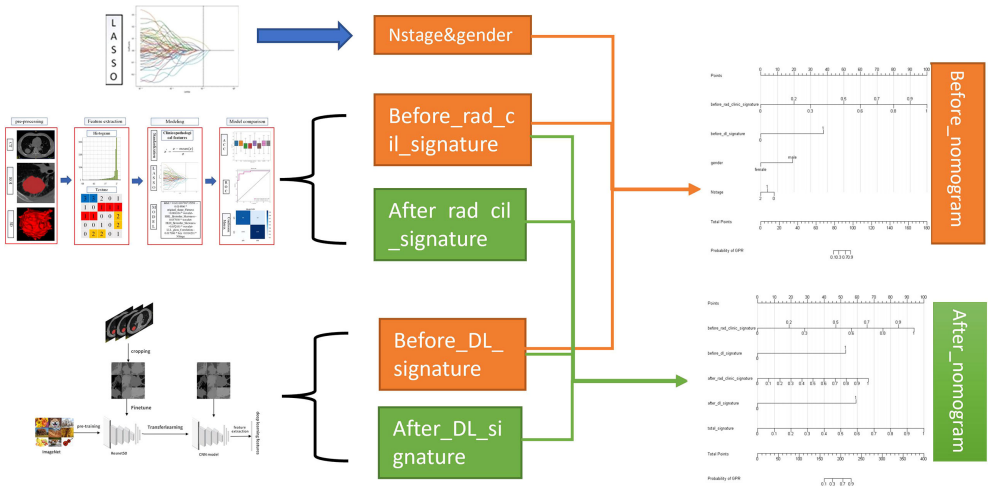


FIGURE 4
Nomogram construction flowchart.

Missing data processing

After deleting the categories with missing values greater than 25%, the items and missing ratios that needed MICEforest for multiple interpolations are before_CRP, 24.19%; before_thyroid function, 19.35%; after_thyroid function, 11.29%; after_CRP, 6.45%; differentiation, 6.45%; after tumor markers, 4.84%; and before_LDH, 4.84%. Because the difference between the thyroid function data using multiple interpolations and the original data is more than 10%, they were deleted. The items and proportions of missing data that were ultimately included in the study and subjected to multiple imputations using the MICEforest method are shown in Figure 5. After multiple imputations, the mean difference between imputed and raw data was between 0.01% and 3%. Figure 5 shows the fit of the imputed data to the original data.

TABLE 2 Categories with missing values, numbers missing, and proportions of missingness in the clinicopathological data collected.

Category	Numbers missing	Proportions of missingness	Category	Numbers missing	Proportions of missingness
CK	60	0.968	Napsin-A	36	0.581
CDX-2	60	0.968	CK5/6	36	0.581
CEA	59	0.952	TTF-1	31	0.5
P53	57	0.919	after _ assisted / induced T	30	0.484
CD56	50	0.806	after _ inhibited / cytotoxic T	30	0.484
CgA	50	0.806	before _ inhibited / cytotoxic T	28	0.452
Syn	49	0.79	before _ assisted / induced T	28	0.452
P63	48	0.774	before_CRP	15	0.242
CK-Pan	45	0.726	before_ thyroid function	12	0.194
Ki67	43	0.694	after_ thyroid function	7	0.113
CK7	43	0.694	after_CRP	4	0.065
PD-L1 expression	40	0.645	differentiation	4	0.065
Driving gene mutation	39	0.629	after tumor markers	3	0.048
P40	36	0.581	before_LDH	3	0.048

TABLE 3 Categories with no missing values in the collected clinicopathological data.

Category	Numbers missing	Proportions of missingness	Category	Numbers missing	Proportions of missingness
before tumor markers	0	0	before_ albumin	0	0
after_L%	0	0	before_BMI	0	0
after_WBC	0	0	before_ weight	0	0
after_ albumin	0	0	height	0	0
after_LDH	0	0	chronic history	0	0
after_ weight	0	0	gender	0	0
after_BMI	0	0	smoking status	0	0
number of immunotherapy cycles	0	0	Evaluation of postoperative clinical TNM stage	0	0
family history of tumor	0	0	N stage	0	0
before_L%	0	0	T stage	0	0
before_WBC	0	0	tumor type	0	0
Classification of immunodrugs	0	0	age	0	0

General information on clinical pathology

The distribution of clinicopathological information used for data analysis in this study was not statistically significant between GPR and BPR except for gender, as shown in Table 4.

Radiomics conventional features and model building

Before_rad_data extraction and before_rad_cil model construction

Python Pyradiomics package extracted radiomics Conventional quantitative features from CT images, extracting 1,648 features per patient. Features extracted from all CT images before NAT were constructed as the before_rad_data. There were 197 first-order features (First order), 13 2D shape features, 231 3D shape features, 242 GLCM features, 84 GLDM features, 96 GLRLM features, 96 GLSZM features, and 688 wavelet transform features in the before_rad_data.

Clinicopathological characteristics before neoadjuvant treatment are shown in Table 5, used with the before_rad_data, and jointly constructed into a 2D array with a feature number of 1,667. After screening by the Spearman correlation coefficient, 238 features related to the efficacy of neoadjuvant immunotherapy were obtained. The screened features were randomly divided into training and test groups in a 7:3 ratio and further screened by the LASSO regression model. At Lamba = 0.1048 (Figure 6), six features highly correlated with the efficacy of neoadjuvant immunotherapy

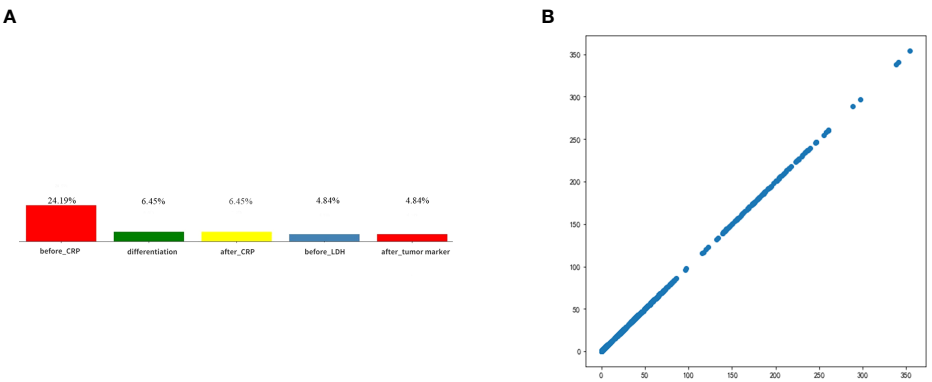


FIGURE 5
(A) Categories of missing values and ratios that need interpolation. **(B)** The fit of the imputed data to the original data.

10.3389/fonc.2022.937277

TABLE 4 General information of clinicopathological information.

Category	Sub-classification	Total number	BPR	GPR	p-value	Test methods
		62	23	39		
gender	Women	11 (17.7%)	8 (34.8%)	3 (7.7%)	0.019	chi-square test
	Men	51 (82.3%)	15 (65.2%)	36 (92.3%)		
age(year)		58.5 (54.2-64.0)	58.0 (53.0-62.0)	59.0 (56.0-64.5)	0.321	Wilcoxon signed rank test
Type of tumor	adenocarcinoma	18 (29.0%)	10 (43.5%)	8 (20.5%)	0.102	chi-square test
	squamous	44 (71.0%)	13 (56.5%)	31 (79.5%)		
differentiation	intermediate	24 (38.7%)	12 (52.2%)	12 (30.8%)	0.226	chi-square test
	low	16 (25.8%)	4 (17.4%)	12 (30.8%)		
	well	22 (35.5%)	7 (30.4%)	15 (38.5%)		
T stage	III stage	16 (25.8%)	5 (21.7%)	11 (28.2%)	0.847	chi-square test
	II stage	26 (41.9%)	10 (43.5%)	16 (41.0%)		
	IV stage	16 (25.8%)	7 (30.4%)	9 (23.1%)		
	I stage	4 (6.5%)	1 (4.3%)	3 (7.7%)		
N stage	0 stage	16 (25.8%)	7 (30.4%)	9 (23.1%)	0.615	chi-square test
	II stage	32 (51.6%)	10 (43.5%)	22 (56.4%)		
	I stage	14 (22.6%)	6 (26.1%)	8 (20.5%)		
clinical stages	I/II stage	14 (22.6%)	6 (26.1%)	8 (20.5%)	0.721	chi-square test
	IIIA stage	34 (54.8%)	13 (56.5%)	21 (53.8%)		
	IIIB stage	14 (22.6%)	4 (17.4%)	10 (25.6%)		
smoking history	No	17 (27.4%)	9 (39.1%)	8 (20.5%)	0.196	chi-square test
	Yes	45 (72.6%)	14 (60.9%)	31 (79.5%)		
chronic history	No	48 (77.4%)	16 (69.6%)	32 (82.1%)	0.411	chi-square test
	Yes	14 (22.6%)	7 (30.4%)	7 (17.9%)		
height (m)		1.7 (1.6-1.7)	1.6 (1.6-1.7)	1.7 (1.6-1.7)	0.088	Wilcoxon signed rank test
before weight(kg)		61.2 (56.0-69.6)	60.0 (53.5-65.8)	64.0 (57.0-70.5)	0.117	Wilcoxon signed rank test
before_BMI(kg/m^2)		22.6 (21.1-24.2)	22.6 (20.8-23.7)	22.7 (21.1-24.5)	0.6	Wilcoxon signed rank test
before_LDH(U/L)		175.1 (150.8-191.4)	178.5 (165.9-189.0)	175.0 (150.0-192.8)	0.925	Wilcoxon signed rank test
Before_albumin(g/L)		39.3 (37.2-41.6)	38.9 (36.9-41.5)	39.5 (37.3-41.5)	0.62	Wilcoxon signed rank test
before_CRP(mg/L)		9.5 (2.3-21.9)	7.0 (2.0-16.4)	10.4 (3.1-24.0)	0.166	Wilcoxon signed rank test
before_WBC(10^9/L)		6.7 (5.1-7.7)	5.7 (4.7-6.8)	7.0 (5.4-7.8)	0.022	Wilcoxon signed rank test
Before_lymphocyte percentage(%)		0.2 (0.2-0.3)	0.3 (0.2-0.3)	0.2 (0.2-0.3)	0.16	Wilcoxon signed rank test
Before_tumor markers	abnormal	34 (54.8%)	14 (60.9%)	20 (51.3%)	0.639	chi-square test
	normal	28 (45.2%)	9 (39.1%)	19 (48.7%)		
after_BMI(kg/m^2)		23.2 (21.7-24.9)	22.8 (21.1-24.7)	23.6 (21.8-25.2)	0.336	Wilcoxon signed rank test
After_weight(kg)		64.0 (57.2-70.0)	60.0 (55.8-66.5)	65.5 (58.5-71.2)	0.085	Wilcoxon signed rank test
after_LDH(U/L)		199.2 (170.0-224.3)	204.0 (176.2-221.2)	193.0 (166.7-223.8)	0.359	Wilcoxon signed rank test
After_albumin(g/L)		40.0 (38.9-42.5)	41.4 (39.0-42.7)	40.0 (38.8-42.1)	0.517	Wilcoxon signed rank test
after_CRP(mg/L)		3.1 (2.2-5.7)	4.2 (2.1-6.6)	2.9 (2.3-5.2)	0.503	Wilcoxon signed rank test
after_WBC(10^9/L)		4.7 (3.9-5.6)	4.7 (4.0-5.3)	4.6 (4.0-5.8)	0.793	Wilcoxon signed rank test
after_lymphocyte percentage(%)		0.3 (0.2-0.4)	0.3 (0.2-0.4)	0.3 (0.3-0.4)	0.431	Wilcoxon signed rank test
after_tumor markers	abnormal	13 (21.0%)	7 (30.4%)	6 (15.4%)	0.279	chi-square test
	normal	49 (79.0%)	16 (69.6%)	33 (84.6%)		
number of immunotherapy cycles	two cycles	5 (8.1%)	2 (8.7%)	3 (7.7%)	0.982	chi-square test
	three cycles	54 (87.1%)	20 (87.0%)	34 (87.2%)		
	four cycles	3 (4.8%)	1 (4.3%)	2 (5.1%)		
Classification of immunodrugs	domestic	28 (45.2%)	11 (47.8%)	17 (43.6%)	0.952	chi-square test
	imported	34 (54.8%)	12 (52.2%)	22 (56.4%)		
history of tumor	no	60 (96.8%)	23 (100.0%)	37 (94.9%)	0.719	chi-square test
	yes	2 (3.2%)	0 (0.0%)	2 (5.1%)		

TABLE 5 Clinicopathological information converged with the before_rad_data.

Gender (men/women)	Age (years)	Tumor type (squamous / adenocarcinoma)	Differentiation grade (low/intermediate/well differentiated)
smoking history (yes / no)	history of tumor (yes / no)	chronic history (yes / no)	beforeweight(kg)
before_LDH(U/L)	before albumin(g/L)	before_CRP(mg/L)	before_WBC(10^9/L)
T stage(I/II/III/IVstage)	N stage(I/II stage)	Clinical TNM stage(IB-III stage)	before tumor markers (normal / abnormal)
before_BMI(kg/m^2)	before_ percentage of lymphocytes (%)	height(m)	

(four radiomics traditional features and two clinical features) were obtained, and the selected six features were combined into one label using a generalized linear model, and the label score was calculated for each patient. The label score is calculated as follows: label score = 0.633104796735994 - 0.014906 * original_shape_Flatness + 0.004334 * wavelet-HHL_firstorder_Skewness - 0.037030 * wavelet-HLH_firstorder_Skewness + 0.052433 * wavelet-LLL_glcm_Correlation - 0.017086 * gender + 0.014231 * N_stage.

The six features screened by LASSO regression are randomly divided into a training group and a test group according to the 7:3 ratio. Then, eight common machine learning algorithms are used to model. The accuracy of the model obtained by 100 random grouping modeling is shown in Figure 7 and Table 6. Three of the eight models (SVM, KNN, ExtraTrees) had a maximum accuracy of 1. The mean accuracy of all models was 0.708–0.599.

In the 100 random groupings, the performances of the best grouping in each model are shown in Figure 7 and Table 7. The best model was Extratrees. The accuracy, AUC, sensitivity, and specificity in both training and test sets were 1. So, Extratrees model data were selected to build the before_rad_cil_signature for subsequent research.

After_rad_data extraction and after_rad_cil model construction

The extraction method for the after_rad_data was consistent with the before_rad_data, and the data for extracting quantitative features were derived from CT images of all patients after NAT. Finally, the resulting quantitative feature categories and numbers are consistent with the before_rad_data.

Clinicopathological characteristics after neoadjuvant treatment are shown in Table 8, fused with the after_rad_data and jointly constructed into a 2D array with a feature number of 1,658. After screening by the Spearman correlation coefficient, 237 features related to the efficacy of neoadjuvant immunotherapy were obtained. The screened features were randomly divided into training and testing groups in a 7:3 ratio, further screened by the LASSO regression model. At Lamba = 0.06866 (Figure 8), nine features highly correlated with the efficacy of neoadjuvant immunotherapy (seven radiomics traditional features and two clinical features) were obtained. The selected nine features were combined into one label using a generalized linear model, and the label score was calculated for each patient. The label score is calculated as follows: Label score = 0.6110867308836403 + 0.012900 * exponential_firstorder_Skewness - 0.150077 *

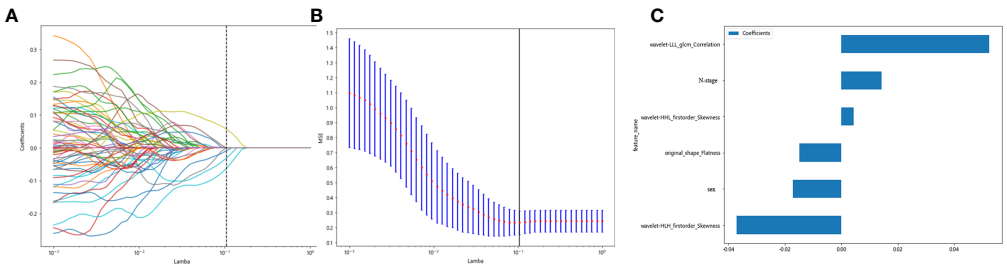


FIGURE 6
Regression feature screening. **(A)** Feature selection plot for the LASSO regression, which was adjusted by a super parameter (Lamba), to achieve the purpose of screening the optimal features. The vertical dashed line indicates that the corresponding optimal Lamba value when obtaining the minimum deviation value is Lamba = 0.1048. **(B)** The convergence graph of characteristic coefficients for feature selection by cross-validation. Features with non-zero coefficients were screened out corresponding to the vertical lines in the plot, with a total of six best features selected. **(C)** Features and weights of LASSO regression screening. The six features screened by LASSO were "original_shape_Flatness," "wavelet-HHL_firstorder_Skewness," "wavelet-HLH_firstorder_Skewness," "wavelet-LLL_glcm_Correlation," "gender," and "N_stage".

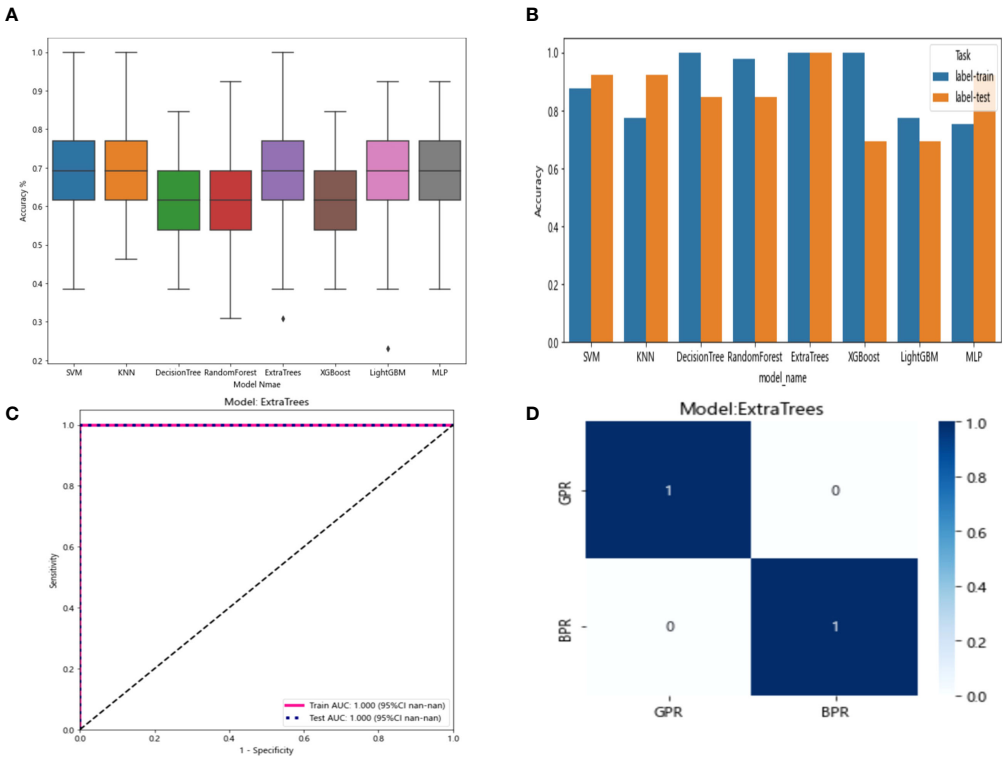


FIGURE 7
(A) Distribution of the accuracy of each model randomly tested 100 times by eight common machine learning algorithms. **(B)** Comparison of the accuracy of the best performing one out of 100 randomizations in eight machine learning algorithms. **(C)** AUCs of the training set and test set in the Extratrees model. **(D)** Confusion matrix for the Extratrees model.

exponential_glrlm_RunEntropy + 0.015404 * log-sigma-5-0-mm-3D_firstorder_Kurtosis + 0.033240 * logarithm_firstorder_Skewness - 0.003723 * original_shape_Elongation - 0.027763 * original_shape_Flatness - 0.006209 * wavelet-LLH_glcm_ClusterShade - 0.025090 * after_CRP + 0.022022 * after_percentage of lymphocytes.

The nine features screened by LASSO regression were randomly divided into a training group and a testing group according to the ratio of 7:3. Then, eight common machine learning algorithms are used to model. The accuracy of the model obtained by 100 random grouping modeling is shown in Figure 9 and Table 9. One of the eight models (MLP) had a maximum accuracy of 1. The maximum accuracy of the other models ranged from 0.846 to 0.923. The mean accuracy of all models was 0.602–0.682.

In the 100 random groupings, the performances of the best grouping in each model are shown in Figure 9 and Table 10. The best model was XGBoost. The accuracy, AUC, sensitivity, and specificity in the training set were 1 and in the testing set were 0.923, 0.9, 0.7, and 1, respectively. So, XGBoost model data were selected to build the after_rad_cil_signature for subsequent research.

TABLE 6 Distribution of the accuracy of each model randomly tested 100 times by eight common machine learning algorithms.

model Names	SVM	KNN	DecisionTree	RandomForest	ExtraTrees	XGBoost	LightGBM	MLP
Number of tests	100	100	100	100	100	100	100	100
Mean accuracy	0.668	0.708	0.599	0.631	0.667	0.621	0.664	0.686
25 % quantile accuracy	0.615	0.615	0.538	0.538	0.615	0.538	0.615	0.615
50% quantile accuracy	0.692	0.692	0.615	0.615	0.692	0.615	0.692	0.692
75% quantile accuracy	0.769	0.769	0.692	0.692	0.769	0.692	0.769	0.769
Maximum accuracy	1	1	0.846	0.923	1	0.846	0.923	0.923

TABLE 7 Each evaluation index of the best performance of 100 random groupings in eight machine learning algorithms.

Model names	Accuracy	AUC	Sensitivity	Specificity	Sets
SVM	0.878	0.962	0.897	0.95	train
SVM	0.923	1	1	1	test
KNN	0.776	0.863	0.828	0.7	train
KNN	0.923	0.95	0.9	1	test
DecisionTree	1	1	1	1	train
DecisionTree	0.846	0.9	0.8	1	test
RandomForest	0.980	0.997	0.966	1	train
RandomForest	0.846	0.883	0.8	1	test
ExtraTrees	1	1	1	1	train
ExtraTrees	1	1	1	1	test
XGBoost	1	1	1	1	train
XGBoost	0.692	0.9	0.7	1	test
LightGBM	0.776	0.803	0.862	0.7	train
LightGBM	0.692	0.833	1	0.667	test
MLP	0.755	0.916	0.931	0.75	train
MLP	0.923	0.833	1	0.667	test

TABLE 8 Clinicopathological information converged with the before_rad_data.

After_BMI(Kg/m2)	After_weight (kg)	After_LDH(U/L)	After_albumin(g/L)
after_CRP(mg/L)	after_WBC(10^9/L)	After_ percentage of lymphocytes (%)	After_ tumor markers (normal / abnormal)
number of immunotherapy cycles (two cycles / three cycles / four cycles)	Classification of immunodrugs (domestic / imported)		

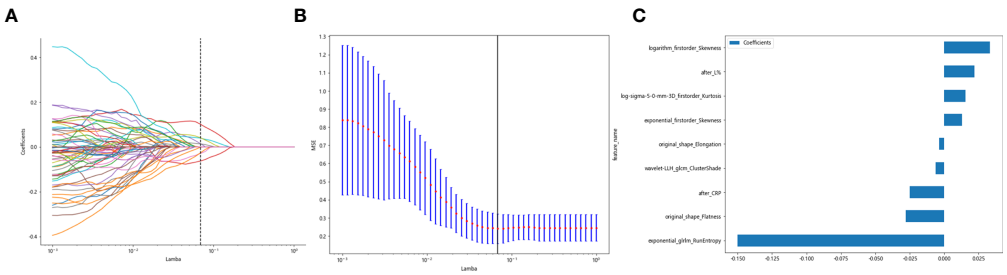


FIGURE 8
Regression feature screening. **(A)** Feature selection plot for the LASSO regression, which was adjusted by a super parameter (Lamba), to achieve the purpose of screening the optimal features. The vertical dashed line indicates that the corresponding optimal Lamba value when obtaining the minimum deviation value is Lamba = 0.06866. **(B)** The convergence graph of characteristic coefficients for feature selection by cross-validation. Features with non-zero coefficients were screened out corresponding to the vertical lines in the plot, with a total of nine best features selected. **(C)** Twelve features and weights of LASSO regression screening. The nine features screened were "exponential_firstorder_Skewness," "exponential_glrlm_RunEntropy," "log-sigma-5-0-mm-3D_firstorder_Kurtosis," "logarithm_firstorder_Skewness," "original_shape_Elongation," "original_shape_Flatness," "wavelet-LLH_glcm_ClusterShade," "after_CRP," and "after_percentage of lymphocytes".

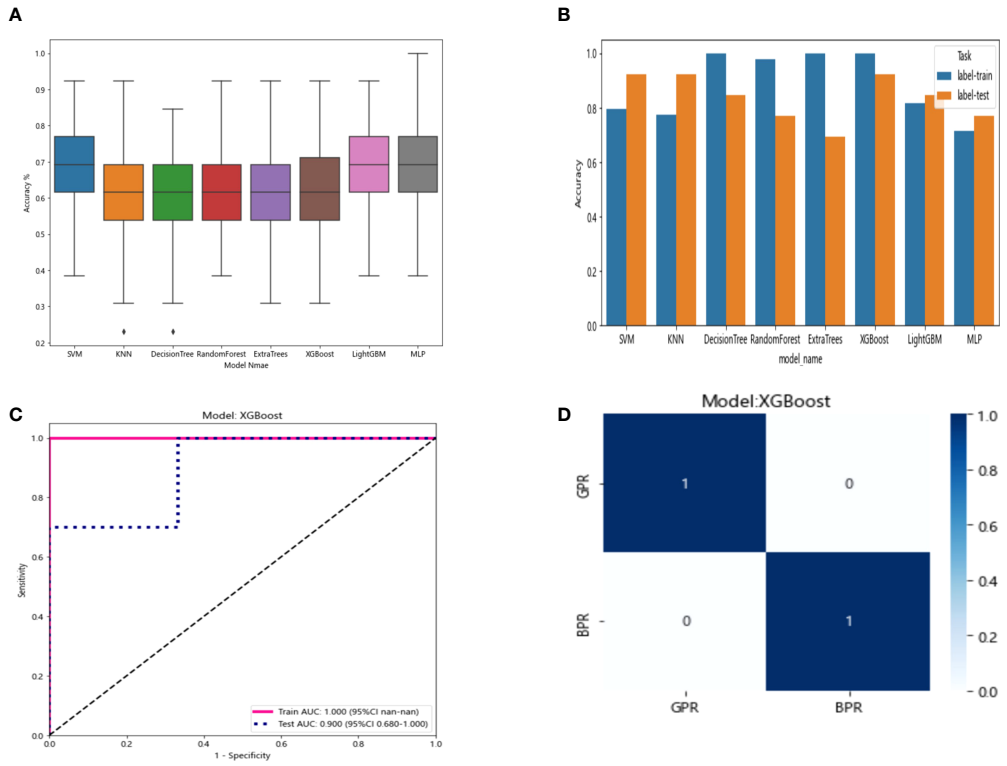


FIGURE 9
(A) Distribution of the accuracy of each model randomly tested 100 times by eight common machine learning algorithms. **(B)** Comparison of the accuracy of the best performing one out of 100 randomizations in eight machine learning algorithms. **(C)** AUCs of the training set and test set in the XGBoost model. **(D)** Confusion matrix for the XGBoost model.

Radiomics deep learning model building

Before_dl model building

The resnet50 DL model after fine-tuning was used to extract features, obtaining 2,048-dimensional DL features from all CT images obtained before NAT by intercepting the lung window images at the maximum level of the tumor in the cross section. The 62-dimensional data set was obtained by PCA dimensionality reduction and randomly divided into training and testing groups according to the ratio of 8:2. Then, eight common machine learning algorithms were used to model. The accuracy of the model obtained by 100 random grouping modeling is shown in Figure 10 and Table 11. One of the eight models (DecisionTree) had a maximum accuracy of 0.923. The mean accuracy of all models was 0.469–0.629.

In the 100 random groupings, the performances of the best grouping in each model were shown in Figure 10 and Table 12. The best model was DecisionTree. The accuracy, AUC, sensitivity, and specificity in the training set were 1 and in the testing set were 0.923, 0.9, 1, and 1. So, DecisionTree model data were selected to build the before_dl_signature for subsequent research.

TABLE 9 Distribution of the accuracy of each model randomly tested 100 times by eight common machine learning algorithms.

Model Names	SVM	KNN	DecisionTree	RandomForest	ExtraTrees	XGBoost	LightGBM	MLP
Number of tests	100	100	100	100	100	100	100	100
Mean accuracy	0.682	0.602	0.611	0.627	0.637	0.645	0.675	0.68
25 % quantile accuracy	0.616	0.538	0.538	0.538	0.538	0.538	0.615	0.615
50% quantile accuracy	0.692	0.615	0.615	0.615	0.615	0.615	0.692	0.692
75% quantile accuracy	0.769	0.692	0.692	0.692	0.692	0.712	0.769	0.769
Maximum accuracy	0.923	0.923	0.846	0.923	0.923	0.923	0.923	1

TABLE 10 Each evaluation index of the best performance of 100 random groupings in eight machine learning algorithms.

Model names	Accuracy	AUC	Sensitivity	Specificity	Sets
SVM	0.796	0.928	0.828	0.9	train
SVM	0.923	0.967	0.9	1	test
KNN	0.776	0.8	0.793	0.75	train
KNN	0.923	0.967	0.8	1	test
DecisionTree	1	1	1	1	train
DecisionTree	0.846	0.783	0.9	1	test
RandomForest	0.980	0.999	0.966	1	train
RandomForest	0.769	0.833	0.5	1	test
ExtraTrees	1	1	1	1	train
ExtraTrees	0.692	0.8	0.5	1	test
XGBoost	1	1	1	1	train
XGBoost	0.923	0.9	0.7	1	test
LightGBM	0.816	0.872	0.862	0.85	train
LightGBM	0.846	0.833	1	0.667	test
MLP	0.714	0.910	0.862	0.9	train
MLP	0.769	1	1	1	test

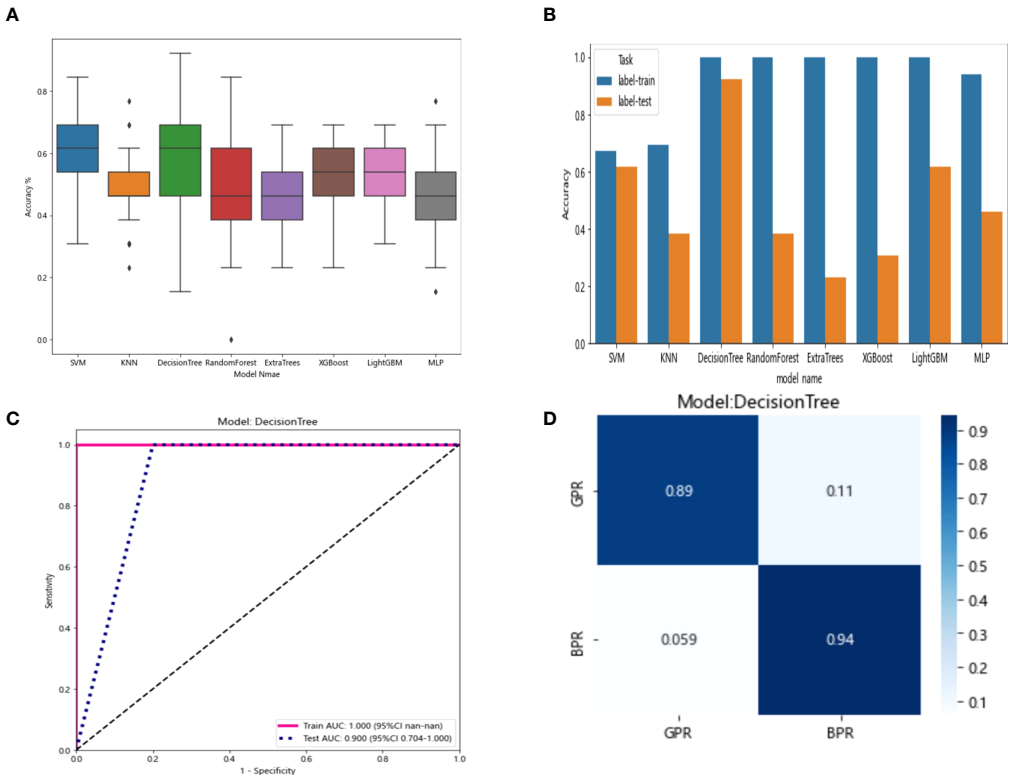


FIGURE 10
(A) Each model's accuracy distribution was randomly tested 100 times by eight common machine learning algorithms. **(B)** Comparison of the accuracy of the best performing one out of 100 randomizations in eight machine learning algorithms. **(C)** AUCs of the training set and test set in the DecisionTree model. **(D)** Confusion matrix for the DecisionTree model.

TABLE 11 Distribution of the accuracy of each model randomly tested 100 times by eight common machine learning algorithms.

model Names	SVM	KNN	DecisionTree	RandomForest	ExtraTrees	XGBoost	LightGBM	MLP
Number of tests	100	100	100	100	100	100	100	100
Mean accuracy	0.629	0.492	0.599	0.489	0.469	0.511	0.515	0.472
25 % quantile accuracy	0.538	0.462	0.462	0.385	0.385	0.462	0.462	0.385
50% quantile accuracy	0.615	0.462	0.615	0.462	0.462	0.538	0.538	0.462
75% quantile accuracy	0.692	0.538	0.692	0.615	0.538	0.615	0.615	0.538
Maximum accuracy	0.846	0.769	0.923	0.846	0.692	0.692	0.692	0.769

After_dl model building

The resnet50 DL model after fine-tuning was used to extract features, obtaining 2,048-dimensional DL features from all CT images obtained after NAT by intercepting the lung window images at the maximum level of the tumor in the cross section. The 62-dimensional data set was obtained by PCA dimensionality reduction and randomly divided into training and testing groups according to the ratio of 8:2. Then, eight common machine learning algorithms were used to model. The accuracy of the model obtained by 100 random grouping modeling is shown in Figure 11 and Table 13. Three of the eight models (DecisionTree, XGBoost, and LightGBM) had a maximum accuracy of 0.923. The mean accuracy of all models was 0.509–0.603.

In the 100 random groupings, the performances of the best grouping in each model are shown in Figure 11 and Table 14. The best model was DecisionTree. The accuracy, AUC, sensitivity, and specificity in the training set were 1 and in the testing set were 0.923, 0.833, 1, and 1. So, DecisionTree model data were selected to build the after_dl_signature for subsequent research.

Entire model

All previously extracted radiomics traditional quantitative features, radiomics DL features, and clinicopathological features after deletion and missing value processing were combined into a 7,421-dimensional joint data set. The combined data set obtained 4,266 characteristics related to the efficacy of immune NAT after standardization and Spearman correlation coefficient screening. The screened features were randomly divided into training and test groups in a 7:3 ratio, which was further screened by the LASSO regression model. At Lamba = 0.0596 (Figure 12), 20 features highly correlated with the efficacy of immune NAT (six radiomics traditional features, 11 DL features, and three clinical features) were obtained, and the selected 20 features were combined into one label using generalized linear model, and the label score was calculated for each patient. The label score is calculated as follows: Label score = 0.6143861851029497 + 0.003364 * exponential_gldm _DependenceEntropy_before - 0.045913 * wavelet-

TABLE 12 Each evaluation index of the best performance of 100 random groupings in eight machine learning algorithms.

Model names	Accuracy	AUC	Sensitivity	Specificity	Sets
SVM	0.673	0.113	0.968	0.059	train
SVM	0.615	0.375	0.5	0.75	test
KNN	0.694	0.659	0.871	0.5	train
KNN	0.385	0.263	1	0	test
DecisionTree	1	1	1	1	train
DecisionTree	0.923	0.9	1	1	test
RandomForest	1	1	1	1	train
RandomForest	0.385	0.425	0.5	0.75	test
ExtraTrees	1	1	1	1	train
ExtraTrees	0.231	0.063	1	0	test
XGBoost	1	1	1	1	train
XGBoost	0.308	0.35	0.5	0.75	test
LightGBM	1	1	1	1	train
LightGBM	0.615	0.2	1	0	test
MLP	0.939	0.996	0.968	1	train
MLP	0.462	0.325	0.125	1	test

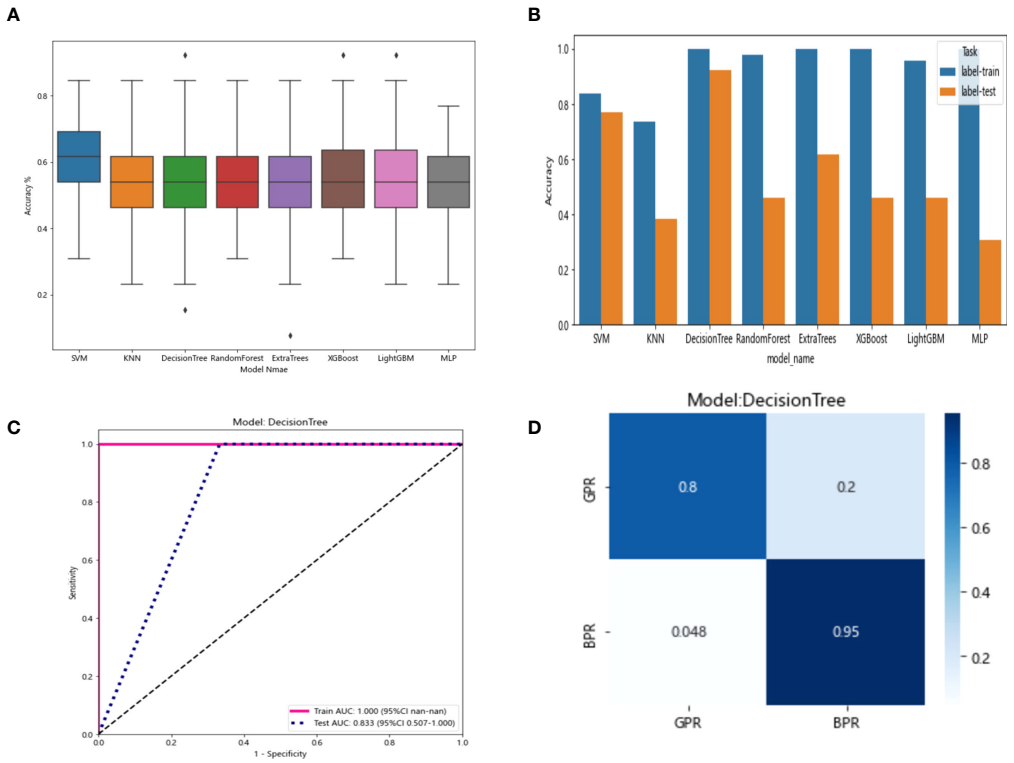


FIGURE 11
(A) Each model's accuracy distribution was randomly tested 100 times by eight common machine learning algorithms. **(B)** Comparison of the accuracy of the best performing one out of 100 randomizations in eight machine learning algorithms. **(C)** AUCs of the training set and test set in the DecisionTree model. **(D)** Confusion matrix for the DecisionTree model.

HLH_firstorder_Skewness_before - 0.008996 * wavelet-HLH_glcm_Correlation_before + 0.058934 * wavelet-LLL_glcm_Correlation_before + 0.028862 * square_glszm_SmallAreaLowGrayLevelEmphasis_after - 0.001464 * wavelet-LHL_glcm_ClusterShade_after - 0.051536 * gender + 0.002784 * age + 0.011369 * N_stage + 0.021636 * 642_before + 0.009017 * 61_after + 0.122782 * 80_after - 0.138473 * 199_after +0.021755 * 284_after + 0.036575 * 802_after - 0.008443 * 965_after + 0.038793 * 1508_after +0.034689 * 1538_after + 0.045343 * 1553_after + 0.000306 * 2030_after.

Note: "_before" represents the radiomics features before treatment, and "_after" represents the radiomics features after treatment. Naming is like "exponential_gldm_dependenceenterprise_before" for radiomics traditional features, and the naming is like "80_after" for DL features.

The 20 features screened by LASSO regression are randomly divided into a training group and a testing group according to the ratio of 7:3. Then, eight common machine learning algorithms are used to model. The accuracy of the model obtained by 100 random grouping modeling is shown in

TABLE 13 Distribution of the accuracy of each model randomly tested 100 times by eight common machine learning algorithms.

model Names	SVM	KNN	DecisionTree	RandomForest	ExtraTrees	XGBoost	LightGBM	MLP
Number of tests	100	100	100	100	100	100	100	100
Mean accuracy	0.603	0.545	0.548	0.547	0.545	0.567	0.568	0.509
25 % quantile accuracy	0.538	0.462	0.462	0.462	0.462	0.462	0.462	0.462
50% quantile accuracy	0.615	0.538	0.538	0.538	0.538	0.538	0.538	0.538
75% quantile accuracy	0.692	0.615	0.615	0.615	0.615	0.635	0.635	0.615
Maximum accuracy	0.846	0.846	0.923	0.846	0.846	0.923	0.923	0.769

TABLE 14 Each evaluation index of the best performance of 100 random groupings in eight machine learning algorithms.

Model names	Accuracy	AUC	Sensitivity	Specificity	Sets
SVM	0.837	0.017	1	0	train
SVM	0.769	0.733	0.6	1	test
KNN	0.735	0.769	0.793	0.65	train
KNN	0.385	0.25	1	NaN	test
DecisionTree	1	1	1	1	train
DecisionTree	0.923	0.833	1	1	test
RandomForest	0.980	0.999	0.966	1	train
RandomForest	0.462	0.45	0.5	0.667	test
ExtraTrees	1	1	1	1	train
ExtraTrees	0.615	0.367	1	0	test
XGBoost	1	1	1	1	train
XGBoost	0.462	0.433	0.4	1	test
LightGBM	0.959	0.984	0.966	1	train
LightGBM	0.462	0.433	0.4	1	test
MLP	1	1	1	1	train
MLP	0.308	0.033	1	0	test

Figure 13 and Table 15. Four of the eight models (SVM, XGBoost, LightGBM, and MLP) had a maximum accuracy of 1. The mean accuracy of all models was 0.628–0.805.

In the 100 random groupings, the performances of the best grouping in each model are shown in Figure 13 and Table 16. The best model was SVM and XGBoost. The accuracy, AUC, sensitivity, and specificity in both training and test sets were 1. However, the average accuracy of model SVM in 100 random grouping tests is higher. So, SVM model data were selected to build the entire_signature for further research.

Nomogram

Drawing and calibration of the before_nomogram

To provide a simple graphical presentation of a clinical prediction model, our study used previously constructed before_rad_clinic_signature, before_dl_signature, and clinical features screened by LASSO to construct a pre-NAT nomogram (before_nomogram). According to the characteristics in Figure 14A, a patient's corresponding

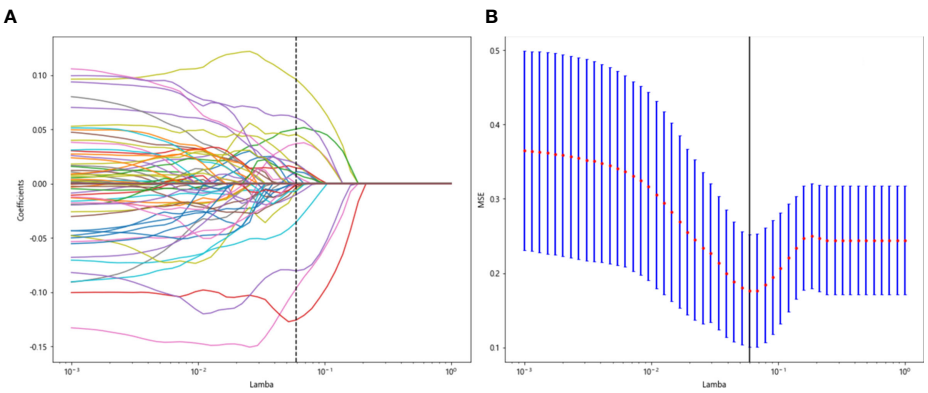


FIGURE 12
Regression feature screening. **(A)** Feature selection plot for the LASSO regression, which was adjusted by a super parameter (Lamba) to achieve the purpose of screening the optimal features. The vertical dashed line indicates that the corresponding optimal Lamba value when obtaining the minimum deviation value is Lamba = 0.0596. **(B)** The convergence graph of characteristic coefficients for feature selection by cross-validation. Features with non-zero coefficients were screened out corresponding to the vertical lines in the plot, with a total of 20 best features selected.

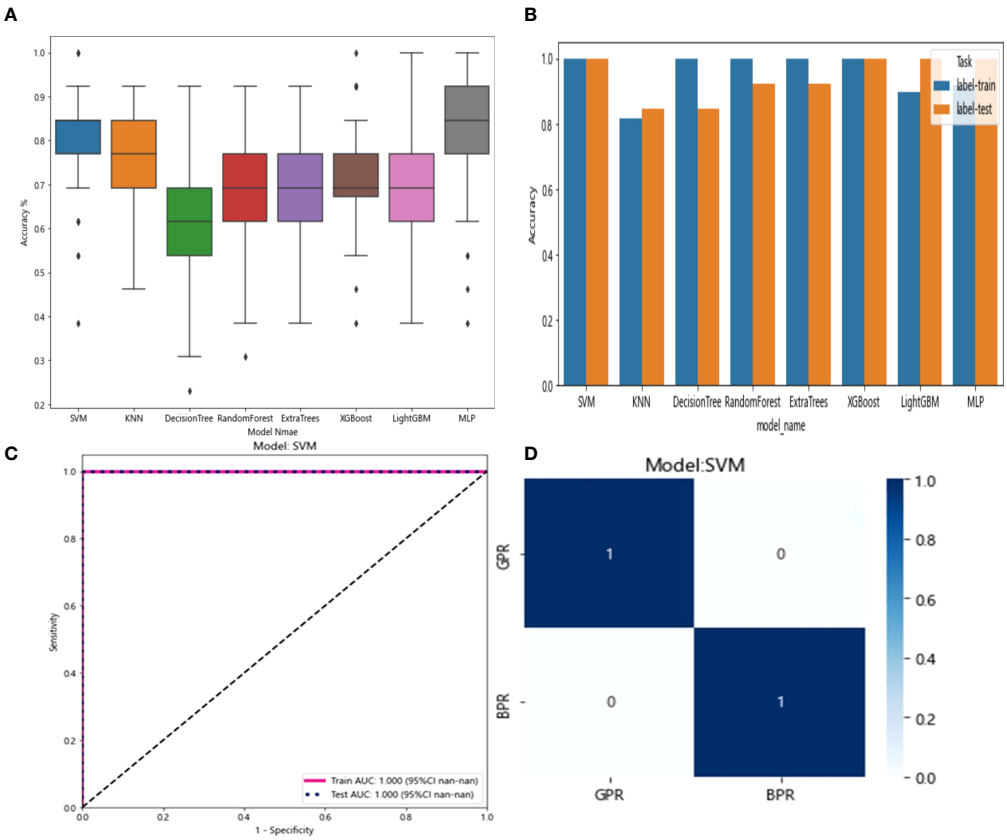


FIGURE 13
(A) Distribution of the accuracy of each model randomly tested 100 times by eight common machine learning algorithms. **(B)** Comparison of the accuracy of the best performing one out of 100 randomizations in eight machine learning algorithms. **(C)** AUCs of the training set and test set in the SVM model. **(D)** Confusion matrix for the SVM model.

probability of GPR after treatment can be calculated. Table 17 shows the features in the before_nomogram and the corresponding scores. Table 18 shows the GPR probability corresponding to different total scores in the before_nomogram. Figure 14B shows that the probability of GPR predicted by the nomogram after 1,000 repeated samplings and the actual GPR is consistent.

Drawing and calibration of the after_nomogram

In our study, we used previously constructed before_rad_clinic_signature, before_dl_signature, after_rad_clinic_signature, after_dl_signature, and entire_signature to jointly construct a posttreatment nomogram (after_nomogram). According to the characteristics

TABLE 15 Distribution of the accuracy of each model randomly tested 100 times by eight common machine learning algorithms.

model Names	SVM	KNN	DecisionTree	RandomForest	ExtraTrees	XGBoost	LightGBM	MLP
Number of tests	100	100	100	100	100	100	100	100
Mean accuracy	0.801	0.732	0.628	0.714	0.714	0.723	0.702	0.805
25 % quantile accuracy	0.769	0.692	0.538	0.615	0.615	0.673	0.615	0.769
50% quantile accuracy	0.846	0.769	0.615	0.692	0.692	0.692	0.692	0.846
75% quantile accuracy	0.846	0.846	0.692	0.769	0.769	0.769	0.769	0.923
Maximum accuracy	1	0.923	0.923	0.923	0.923	1	1	1

TABLE 16 Each evaluation index of the best performance of 100 random groupings in eight machine learning algorithms.

Model names	Accuracy	AUC	Sensitivity	Specificity	Sets
SVM	1	1	1	1	train
SVM	1	1	1	1	test
KNN	0.816	0.872	0.821	0.810	train
KNN	0.846	0.977	0.909	1	test
DecisionTree	1	1	1	1	train
DecisionTree	0.846	0.705	0.909	1	test
RandomForest	1	1	1	1	train
RandomForest	0.923	1	1	1	test
ExtraTrees	1	1	1	1	train
ExtraTrees	0.923	0.932	0.909	1	test
XGBoost	1	1	1	1	train
XGBoost	1	1	1	1	test
LightGBM	0.898	0.927	0.929	0.857	train
LightGBM	1	1	1	1	test
MLP	0.918	0.986	0.929	0.952	train
MLP	1	1	1	1	test

in Figure 15A, a patient's corresponding probability of GPR after treatment can be calculated. Table 19 shows the features in the after_nomogram corresponding scores. Table 20 shows the GPR probability to different total scores in the after_nomogram. Figure 15B shows that the probability of GPR predicted by the nomogram after 1,000 repeated samplings and the actual GPR is consistent.

Discussion

Lung cancer is still the leading cancer in the world and in China, where NSCLC is the most common, with an incidence rate of about 85% (2, 3). Surgery is still the main treatment for early and locally advanced NSCLC (I, II and IIIA, IIIB)(NCCN) (4, 5). However, great progress has been made in recent years with the application of

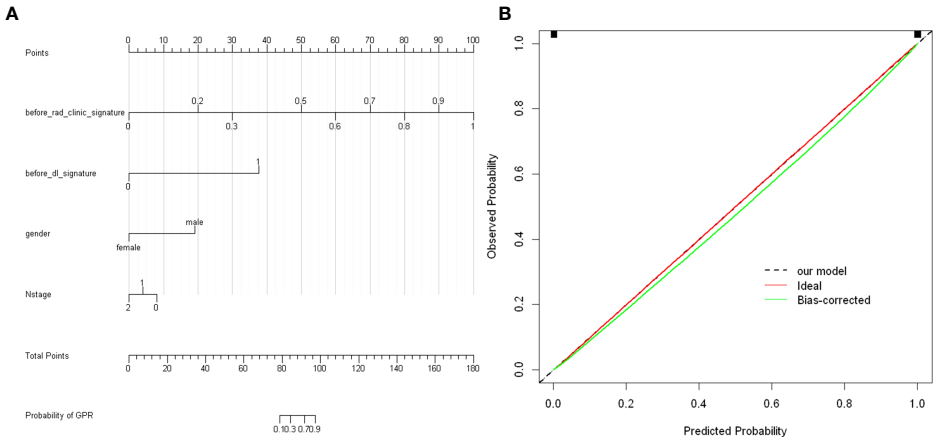


FIGURE 14
(A) before_nomogram.Locate on the before_rad_clinic_signature, before_dl_signature, gender, and N stage coordinate axis. Draw a straight line perpendicular to the first points, calculate and sum the scores corresponding to each straight line, locate on the total points coordinate axis, and draw a straight line perpendicular to the horizontal axis of the probability of GPR. The corresponding value is the probability of pathological response to GPR in patients with non-small cell lung cancer after neoadjuvant immunotherapy. **(B)** Calibration curve corresponding to the before_nomogram. The consistency between the probability of GPR predicted by the nomogram after 1,000 repeated samplings and the actual GPR. The 45° red line represents the ideal prediction performance. The black dotted line and the solid green line represent the prediction performance of the nomogram and the correction of the deviation of the nomogram, respectively. The closer the black dotted line is to the 45° ideal red line, the higher the model's prediction accuracy.

TABLE 17 The features in the nomogram and the corresponding scores.

Before_rad_clinic_signature	Points	Before_dl_signature	Points
0	0	0	0
0.2	20	1	38
0.3	30	gender	Points
0.5	50	women	0
0.6	60	men	19
0.7	70	N stage	Points
0.8	80	0	8
0.9	90	1	4
1	100	2	0

TABLE 18 GPR probability corresponding to different total scores in the before_nomogram.

Total Points	Probability of GPR
79	0.1
84	0.3
92	0.7
97	0.9

ICIs as NAT approach. Immunotherapy-based NAT (nivolumab 360 mg and platinum-containing chemotherapy, once every 3 weeks, for three cycles) has been included in the latest NCCN guidelines(NCCN) (4, 5). Although PD-L1 and tumor mutation burden (TMB) can indicate the effect of immunotherapy to some extent (32), different types of immune cells in the tumor microenvironment, such as CD8+ T-cell infiltration, usually indicate that immunotherapy confers a good response and prognosis (33).

The combination of different immune cells, such as CD3/CD8/CD45RO combined with immune score (34), has a certain suggestive effect on the efficacy of immunotherapy. However, there is currently no reliable indicator to predict the exact efficacy of immunotherapy-based NAT. The common endpoints of clinical trials are progression-free survival (PFS) and overall survival (OS). Although the pathological response can evaluate the treatment benefit earlier than the traditional clinical trial endpoints, the histopathological evaluation can only be determined after the pathological results of surgical resection. CT chest is a common diagnostic imaging modality of lung cancer.

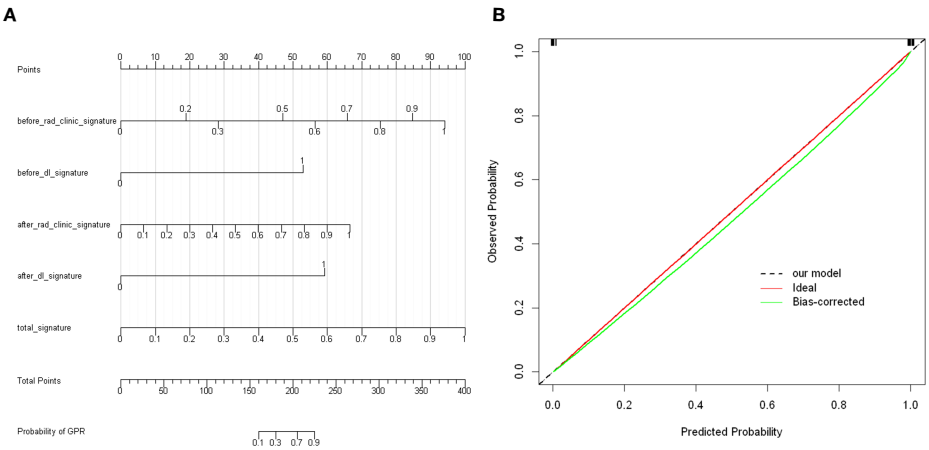


FIGURE 15
(A) After_nomogram.Locate on the before_rad_clinic_signature, before_dl_signature, after_rad_clinic_signature, after_dl_signature, and entire_signature coordinate axis. Draw a straight line perpendicular to the first points, calculate and sum the scores corresponding to each straight line, locate on the total points coordinate axis, and draw a straight line perpendicular to the horizontal axis of the probability of GPR. The corresponding value is the probability of pathological response to GPR in patients with non-small cell lung cancer after neoadjuvant immunotherapy. **(B)** Calibration curve corresponding to after_nomogram. The consistency between the probability of GPR predicted by the nomogram after 1,000 repeated samplings and the actual GPR. The 45° red line represents the ideal prediction performance. The black dotted line and the solid green line represent the prediction performance of the nomogram and the correction of the deviation of the nomogram, respectively. The closer the black dotted line is to the 45° ideal red line, the higher the model's prediction accuracy.

TABLE 19 The features in the nomogram and the corresponding scores.

Before_rad_clinic_signature	Points	After_rad_clinic_signature	Points
0	0	0	0
0.2	19	0.1	7
0.3	28	0.2	13
0.5	47	0.3	20
0.6	56	0.4	27
0.7	66	0.5	33
0.8	75	0.6	40
0.9	85	0.7	47
1	94	0.8	53
		0.9	60
		1	67
after_dl_signature	Points	total_signature	Points
0	0	0	0
1	59	0.1	10
before_dl_signature	Points	0.2	20
0	0	0.3	30
1	53	0.4	40
		0.5	50
		0.6	60
		0.7	70
		0.8	80
		0.9	90
		1	100

TABLE 20 GPR probability corresponding to different total scores in the after_nomogram.

Total Points	Probability of GPR
160	0.1
180	0.3
205	0.7
225	0.9

In contrast, another common efficacy evaluation standard is RECIST 1.1, which is mainly based on the 2D evaluation of the number and size of tumors shown on CT (35). From a quantitative point of view, this method is basic and ignores a large amount of information in medical images. Based on saving diagnostic financial costs, radiomics uses high-throughput technology to extract the conventional features and/or DL features of medical images together with molecular biological information such as genes, proteins, and tumor metabolism, which is then transformed into quantitative features. Combined with machine learning or/and DL, it can be used for disease diagnosis, therapeutic efficacy prediction, and prognosis analysis (22, 23, 25, 26).

Accurate prediction of treatment response is of great significance for the stratification and selection of patients benefiting from immunotherapy. Yang et al. (36) selected 88 radiomics features from the CT images of 92 patients with lung cancer before immunotherapy and constructed a random forest model. Combined with clinicopathological information, they successfully predicted the patients who would benefit from ICI treatment (the AUCs of the training and validation groups were 0.848 and 0.795, respectively). Similarly, Barabino et al. (37) extracted the radiomics features of lung lesions from CT scans at baseline and the first evaluation and calculated their changes by absolute difference and relative reduction (Delta, Δ). After feature screening and model construction, 27 delta features were identified, which were able to distinguish the response to NSCLC immunotherapy with statistically significant accuracy. Moreover, it was found that the changes in the other nine features were significantly correlated with false progression. Another report by Shen et al. (38) predicted the effect of immunotherapy in NSCLC patients through texture feature extraction and texture analysis of lung enhanced CT before treatment. The highest prediction efficiency was 88.2% (sensitivity), 76.3% (specificity), and 81.9% accuracy. These studies suggest that radiomics can help predict and select the right NSCLC patient population for immunotherapy. Before immunotherapy was widely used in the clinic, there have been studies using radiomics to predict the pathological response after concurrent neoadjuvant chemoradiotherapy. A study by

Coroller et al. (39) showed that both radiomics features of primary tumors and lymph node phenotypic information could predict pathological responses. Also, in another study (40), they found that seven features could predict pathological gross residual lesions (AUC >0.6, $p < 0.05$), in which one indicator could predict pathological complete response (AUC = 0.63, $p = 0.01$), and tumors with poor response to neoadjuvant chemoradiotherapy were more likely to show well-circumscribed (spherical nonproportional, AUC = 0.63, $p = 0.009$) and spiculated lesions (LoG 5 mm 3D-GLCM entropy, AUC = 0.61, $p = 0.03$). At present, there is no relevant research on radiomics that predicts the efficacy of neoadjuvant immunotherapy. To explore the role of radiomics in predicting the pathological remission of NSCLC after neoadjuvant immunotherapy, this study extracted the conventional radiological features and DL features from the CT images of NSCLC patients before and after neoadjuvant immunotherapy combined with clinicopathological information to construct models that can predict the pathological remission of NSCLC patients after immunotherapy-based NAT.

The before_rad_cil model was constructed after feature screening. Four radiomics traditional features ("original_shape_Flatness, " "wavelet-HHL_firstorder_Skewness, " "wavelet-HLH_firstorder_Skewness, " and "wavelet-LLL_glcm_Correlation") and two clinical features ("sex" and "N stage") were obtained. Eight common machine learning algorithms model the selected features: SVM, KNN, DecisionTree, RandomForest, ExtraTrees, XGBoost, LightGBM, and MLP. After 100 random groupings of 5-fold cross-validation, the average prediction accuracy of the KNN model was the highest, 0.708. Conventional radiomics features were extracted from the CT images of NSCLC patients after immunotherapy-based NAT and combined with the clinicopathological information obtained after immunotherapy-based NAT. The after_rad_cil model was constructed. After feature screening, seven radiomics traditional features ("exponential_firstorder_Skewness," "exponential_glrlm_RunEntropy," "log-sigma-5-0-mm-3D_firstorder_Kurtosis," "logarithm_firstorder_Skewness," "original_shape_Elongation," "original_shape_Flatness," and "wavelet-LLH_glcm_ClusterShade") and two clinical features ("after_CRP" and "after_percentage of lymphocytes") were obtained.

After the selected features were modeled and cross-verified by eight common machine learning algorithms, the average prediction accuracy of the SVM model was the highest (0.682). After fine-tuning resnet50, the before_dl model extracted DL features from the CT images of NSCLC patients before immunotherapy-based NAT. The DL features only reflected the relationship between features and outcomes without exact physical meaning. Eight common machine learning algorithms then modeled the features after dimensionality reduction by PCA: SVM, KNN, DecisionTree, RandomForest, ExtraTrees, XGBoost, LightGBM, and MLP. After 100 random groupings of 5-fold cross-validation, the average prediction accuracy of the SVM model was the highest, 0.629.

In the after_dl model constructed by the DL features extracted from the CT images of patients after immunotherapy-based NAT like the before_dl model, the average prediction accuracy of SVM was the highest (0.603). The entire model was a prediction model constructed by combining the conventional features of radiomics, DL features, and clinicopathological features before and after NAT. After being modeled by eight common machine learning algorithms and 100 random groupings of 5-fold cross-validation, the average prediction accuracy of the MLP model was the highest, which was 0.805.

Nomograms can graphically describe biological information, characteristics, and clinical variables as a statistical prediction model and estimate the individualized risk according to the characteristics of patients and diseases. It is a simple, easy-to-understand, and user-friendly clinical decision-making tool (41) and was widely used in individualized prognostic evaluation of breast cancer (42), rectal cancer (43), prostate cancer (44), glioma (45), and lung adenocarcinoma (46).

We combined prognostic variables obtained in the before_rad_clinic model, gender, N stage, before_ rad_ clinic_ signature, and before_dl_signature, to construct the before_nomogram, while the before_rad_clinic_signature, before_dl_signature, after_rad_clinic_signature, after_dl_signature, and entire_signature were used to construct the after_nomogram. The calibration curve showed that the nomogram before and after treatment had a good predictive effect on the GPR.

Our results suggest that radiomics can predict the pathological remission of NSCLC after immunotherapy-based NAT. Similar to previous studies (19, 47–49), the prediction efficiency of the entire model is higher than that of the single DL model and the radiomics traditional features combined clinical features model. CT images and clinicopathological information obtained before NAT were constructed as the before_rad_cil model. After classification using the KNN algorithm, the average prediction accuracy was 0.708. Combining GPR-related clinical variables was done to construct the before_nomogram. It shows that clinicians can judge the probability of achieving GPR before treatment in each patient who intends to receive immunotherapy-based NAT.

The entire model had the highest predictive efficacy after classification using the MLP algorithm with an average predictive accuracy of 0.805, combined with the before_rad_clinic_signature, before_dl_signature, after_rad_clinic_signature, and after_dl_signature to construct

the after_nomogram, which can predict the probability that patients are obtaining GPR after immunotherapy-based NAT ahead of surgery. If the follow-up data show that patients who achieved GPR after immunotherapy-based NAT have significant survival benefits after surgery or not, reducing the scope of surgery or avoiding surgical treatment altogether may have the same survival benefits as patients with total resection or extended resection.

Although the results are satisfactory, our study also has some limitations. First, the sample size of this retrospective study is limited, and the consistency test between observers was not carried out. Although we used various methods for feature selection and compared the prediction results of various machine learning methods to select the best model, the limited data set may lead to insufficient generalization of the model. Therefore, in future research, we will strive to conduct a multicenter study and aim to construct large samples with diversified data sets to evaluate the proposed model and verify the robustness and effectiveness of our model through prospective studies.

Secondly, previous studies (50) have shown that adenocarcinoma (ADC) and squamous cell carcinoma (SCC) have different imaging phenotypes on CT scans: peripheral hair glass shadows are more common in ADC, and SCC is more likely to show necrosis. Different imaging phenotypes may lead to different prediction performances. Due to the limited number of cases in this study, the data sets of different histopathological types (ADC and SCC) cannot be hierarchically modeled and verified.

The model constructed by merging the two tumors may lead to a decline in prediction efficiency. Therefore, larger data sets should be used in future research, and the two histological subtypes should be hierarchically modeled and verified. This study only analyzed the predictive effect of some clinicopathological features, conventional radiomics features, and DL features extracted from CT images before and after immunotherapy-based NAT for NSCLC. Studies have shown that the pathological characteristics (51) of patients, genes (52), and protein expression (53) can also affect the prognosis. In future studies, if we combine pathology, genomics, proteomics, and comprehensive clinical information, it is expected to further improve the prediction efficiency of the model. Finally, another limitation of this study is manually sketching the ROI, in which the operator may have different sketching regions, which is time-consuming and laborious. Some studies (54, 55) have begun to attempt to automatically sketch ROIs and automatically extract conventional radiomics features and DL features to construct end-to-end models (56) to complete the research objectives. In case the constructed model has stable performance and accurate efficacy, it may give patients a relatively accurate prediction within a few minutes after obtaining patient-related information, which is convenient for clinical application and makes precise individualized treatment possible.

Conclusion

CT-based radiomics has a good predictive ability for GPR in NSCLC patients receiving immunotherapy-based NAT. Among the radiomics features combined with the clinicopathological information model, DL feature model, and the entire model, the entire model had the highest prediction accuracy.

Data availability statement

The original contributions presented in the study are included in the article/Supplementary Material. Further inquiries can be directed to the corresponding author.

Ethics statement

Ethical review and approval was not required for the study on human participants in accordance with the local legislation and institutional requirements. Written informed consent for participation was not required for this study in accordance with the national legislation and the institutional requirements.

Author contributions

LQ, YK and HJ contributed to conception and design of the study. QS and YK organized the database. LQ performed the statistical analysis. LQ wrote the draft of the manuscript. All authors contributed to the article and approved the submitted version.

Funding

This work was supported by the Natural Science Foundation of Hunan Province, China (Grand NO.2020JJ4915) and the Postgraduate Independent Exploration and Innovation Project of Central South University (No. 2022ZZTS0905).

Acknowledgments

For editing the language of this manuscript, we thank Omar, Ph.D., and for advice regarding the code used in this manuscript, we thank the OnekeyAl platform and its developers.

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.

The handling editor PY declared a shared parent affiliation with the authors at the time of review.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fonc.2022.937277/full#supplementary-material

References

1. Yang D, Liu Y, Bai C, Wang X, Powell CA. Epidemiology of lung cancer and lung cancer screening programs in China and the united states. *Cancer Lett* (2020) 468:82–7. doi: 10.1016/j.canlet.2019.10.009

2. Chen W, Zheng R, Baade PD, Zhang S, Zeng H, Bray F, et al. Cancer statistics in China 2015. *CA Cancer J Clin* (2016) 66(2):115–32. doi: 10.3322/caac.21338

3. Sung H, Ferlay J, Siegel RL, Laversanne M, Soerjomataram I, Jemal A, et al. Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin* (2021) 71(3):209–49. doi: 10.3322/caac.21660

4. NCCN. *NCCN clinical practice guidelines in oncology non-small cell lung Cancer(Version 2.2022–march 7,2022)*. USA: NCCN (2022). Available at: http://www.nccn.org/.

5. NCCN. *"NCCN clinical practice guidelines in oncology non-small cell lung Cancer(Version 3.2022–march 16,2022)"*. USA: NCCN (2022). Available at: http://www.nccn.org/.

6. Bunn P, Schenk E, Pacheco J, Dimou AJO. New developments in neoadjuvant therapy for lung cancer. *Oncology* (2019) 33(3):101–106, 109.

7. Qu Y, Emoto K, Eguchi T, Aly RG, Zheng H, Chaft JE, et al. Pathologic assessment after neoadjuvant chemotherapy for NSCLC: Importance and implications of distinguishing adenocarcinoma from squamous cell carcinoma. *J Thorac Oncol* (2019) 14(3):482–93. doi: 10.1016/j.jtho.2018.11.017

8. de Silva M, Itchins M, Pavlakis N. Breakthrough 5-year survival with pembrolizumab in keynote-001 study: Horizon shifting in advanced non-small cell lung cancer with immune check point inhibition. *Ann Transl Med* (2020) 8 (8):555. doi: 10.21037/atm.2020.01.87

9. Socinski MA, Jotte RM, Cappuzzo F, Orlandi F, Stroyakovskiy D, Nogami N, et al. Atezolizumab for first-line treatment of metastatic nonsquamous NSCLC. *N Engl J Med* (2018) 378(24):2288–301. doi: 10.1056/NEJMoa1716948

10. Ikeda S, Kato T, Kenmotsu H, Ogura T, Iwasawa S, Iwasawa T, et al. A phase II study of atezolizumab for pretreated advanced/recurrent non-small cell lung cancer with idiopathic interstitial pneumonias: rationale and design for the TORG1936/AMBITIOUS study. *Ther Adv Med Oncol* (2020) 12:1758835920922022. doi: 10.1177/1758835920922022

11. Forde PM, Chaft JE, Smith KN, Anagnostou V, Cottrell TR, Hellmann MD, et al. Neoadjuvant PD-1 blockade in resectable lung cancer. *N Engl J Med* (2018) 378(21):1976–86. doi: 10.1056/NEJMoa1716078

12. Cascone T, William WN, Weissferdt A, Lin HY, Sepesi B. Neoadjuvant nivolumab (N) or nivolumab plus ipilimumab (NI) for resectable non-small cell lung cancer (NSCLC): Clinical and correlative results from the NEOSTAR study. *J Clin Oncol* (2019) 37(Suppl 15):8504–4. doi: 10.1200/JCO.2019.37.15_suppl.8504

13. Provencio M, Nadal E, Insa A, García-Campelo M. Neoadjuvant chemotherapy and nivolumab in resectable non-small-cell lung cancer (NADIM): an open-label, multicentre, single-arm, phase 2 trial. *Oncology* (2020) 21(11):1413–22. doi: 10.1016/S1470-2045(20)30453-8

14. Zhi Z.B.L.X.Z. Expert consensus on the pathological evaluation of neoadjuvant therapy efficacy for non-small cell lung cancer. *Chin J Pathol* (2021) 50(9):1002–7. doi: 10.3760/cma.j.cn112151-20210429-00335

15. Hellmann MD, Chaft JE, William WNJr., Rusch V, Pisters KM, Kalhor N, et al. Pathological response after neoadjuvant chemotherapy in resectable non-small-cell lung cancers: proposal for the use of major pathological response as a surrogate endpoint. *Lancet Oncol* (2014) 15(1):e42–50. doi: 10.1016/s1470-2045(13)70334-6

16. Travis WD, Dacic S, Wistuba I, Sholl L, Adusumilli P, Bubendorf L, et al. IASLC multidisciplinary recommendations for pathologic assessment of lung cancer resection specimens after neoadjuvant therapy. *J Thorac Oncol* (2020) 15 (5):709–40. doi: 10.1016/j.jtho.2020.01.005

17. Lambin P, Rios-Velazquez E, Leijenaar R, Carvalho S, van Stiphout RG, Granton P, et al. Radiomics: extracting more information from medical images using advanced feature analysis. *Eur J Cancer* (2012) 48(4):441–6. doi: 10.1016/j.ejca.2011.11.036

18. Guo J, Liu Z, Shen C, Li Z, Yan F, Tian J, et al. MR-based radiomics signature in differentiating ocular adnexal lymphoma from idiopathic orbital inflammation. *Eur Radiol* (2018) 28(9):3872–81. doi: 10.1007/s00330-018-5381-7

19. Lv J, Chen X, Liu X, Du D, Lv W, Lu L, et al. Imbalanced data correction based PET/CT radiomics model for predicting lymph node metastasis in clinical stage T1 lung adenocarcinoma. *Front Oncol* (2022) 12:788968. doi: 10.3389/fonc.2022.788968

20. Rossi G, Barabino E, Fedeli A, Ficarra G, Coco S, Russo A, et al. Radiomic detection of EGFR mutations in NSCLC. *Cancer Res* (2021) 81(3):724–31. doi: 10.1158/0008-5472.Can-20-0999

21. Song L, Zhu Z, Mao L, Li X, Han W, Du H, et al. Clinical, conventional CT and radiomic feature-based machine learning models for predicting ALK rearrangement status in lung adenocarcinoma patients. *Front Oncol* (2020) 10:369. doi: 10.3389/fonc.2020.00369

22. Hirose TA, Arimura H, Ninomiya K, Yoshitake T, Fukunaga JI, Shioyama Y. Radiomic prediction of radiation pneumonitis on pretreatment planning computed tomography images prior to lung cancer stereotactic body radiation therapy. *Sci Rep* (2020) 10(1):20424. doi: 10.1038/s41598-020-77552-7

23. Bortolotto C, Lancia A, Stelitano C, Montesano M, Merizzoli E, Agustoni F, et al. Radiomics features as predictive and prognostic biomarkers in NSCLC. *Expert Rev Anticancer Ther* (2021) 21(3):257–66. doi: 10.1080/14737140.2021.1852935

24. Wu Y, Li T, Han Y, Jiang J. Use of radiomic features and support vector machine to discriminate subjective cognitive decline and healthy controls(). *Annu Int Conf IEEE Eng Med Biol Soc* (2020) 2020:1762–5. doi: 10.1109/embc44109.2020.9175840

25. Wang D, Huang C, Bao S, Fan T, Sun Z, Wang Y, et al. Study on the prognosis predictive model of COVID-19 patients based on CT radiomics. *Sci Rep* (2021) 11(1):11591. doi: 10.1038/s41598-021-90991-0

26. Wang H, Wang L, Lee EH, Zheng J, Zhang W, Halabi S, et al. Decoding COVID-19 pneumonia: comparison of deep learning and radiomics CT image signatures. *Eur J Nucl Med Mol Imaging* (2021) 48(5):1478–86. doi: 10.1007/s00259-020-05075-4

27. Wang S, Shi J, Ye Z, Dong D, Yu D, Zhou M, et al. Predicting EGFR mutation status in lung adenocarcinoma on computed tomography image using deep learning. *Eur Respir J* (2019) 53(3):1800986. doi: 10.1183/13993003.00986-2018

28. Choi YS, Bae S, Chang JH, Kang SG, Kim SH, Kim J, et al. Fully automated hybrid approach to predict the IDH mutation status of gliomas *via* deep learning and radiomics. *Neuro Oncol* (2021) 23(2):304–13. doi: 10.1093/neuonc/noaa177

29. Uhm KH, Jung SW, Choi MH, Shin HK, Yoo JI, Oh SW, et al. Deep learning for end-to-end kidney cancer diagnosis on multi-phase abdominal computed tomography. *NPJ Precis Oncol* (2021) 5(1):54. doi: 10.1038/s41698-021-00195-y

30. Saba L, Biswas M, Kuppili V, Cuadrado Godia E, Suri HS, Edla DR, et al. The present and future of deep learning in radiology. *Eur J Radiol* (2019) 114:14–24. doi: 10.1016/j.ejrad.2019.02.038

31. Bo L, Zhang Z, Jiang Z, Yang C, Huang P, Chen T, et al. Differentiation of brain abscess from cystic glioma using conventional MRI based on deep transfer learning features and hand-crafted radiomics features. *Front Med (Lausanne)* (2021) 8:748144. doi: 10.3389/fmed.2021.748144

32. Kaderbhaï C, Tharin Z, Ghiringhelli F. The role of molecular profiling to predict the response to immune checkpoint inhibitors in lung cancer. *Cancers (Basel)* (2019) 11(2):201. doi: 10.3390/cancers11020201

33. Jiang X, Xu J, Liu M, Xing H, Wang Z, Huang L, et al. Adoptive CD8(+) T cell therapy against cancer:Challenges and opportunities. *Cancer Lett* (2019) 462:23–32. doi: 10.1016/j.canlet.2019.07.017

34. Nosho K, Baba Y, Tanaka N, Shima K, Hayashi M, Meyerhardt JA, et al. Tumour-infiltrating T-cell subsets, molecular changes in colorectal cancer, and prognosis: cohort study and literature review. *J Pathol* (2010) 222(4):350–66. doi: 10.1002/path.2774

35. Eisenhauer EA, Therasse P, Bogaerts J, Schwartz LH, Sargent D, Ford R, et al. New response evaluation criteria in solid tumours: revised RECIST guideline (version 1.1). *Eur J Cancer* (2009) 45(2):228–47. doi: 10.1016/j.ejca.2008.10.026

36. Yang B, Zhou L, Zhong J, Lv T, Li A, Ma L, et al. Combination of computed tomography imaging-based radiomics and clinicopathological characteristics for predicting the clinical benefits of immune checkpoint inhibitors in lung cancer. *Respir Res* (2021) 22(1):189. doi: 10.1186/s12931-021-01780-2

37. Barabino E, Rossi G, Pamparino S, Fiannacca M, Caprioli S, Fedeli A, et al. Exploring response to immunotherapy in non-small cell lung cancer using delta-radiomics. *Cancers (Basel)* (2022) 14(2):350. doi: 10.3390/cancers14020350

38. Shen L, Fu H, Tao G, Liu X, Yuan Z, Ye X. Pre-immunotherapy contrast-enhanced CT texture-based classification: A useful approach to non-small cell lung cancer immunotherapy efficacy prediction. *Front Oncol* (2021) 11:591106. doi: 10.3389/fonc.2021.591106

39. Coroller TP, Agrawal V, Huynh E, Narayan V, Lee SW, Mak RH, et al. Radiomic-based pathological response prediction from primary tumors and lymph nodes in NSCLC. *J Thorac Oncol* (2017) 12(3):467–76. doi: 10.1016/j.jtho.2016.11.2226

40. Coroller TP, Agrawal V, Narayan V, Hou Y, Grossmann P, Lee SW, et al. Radiomic phenotype features predict pathological response in non-small cell lung cancer. *Radiother Oncol* (2016) 119(3):480–6. doi: 10.1016/j.radonc.2016.04.004

41. Balachandran VP, Gonen M, Smith JJ, DeMatteo RP. Nomograms in oncology: more than meets the eye. *Lancet Oncol* (2015) 16(4):e173–180. doi: 10.1016/s1470-2045(14)71116-7

42. Li F, Zhao Y, Wei Y, Xi Y, Bu H. Tumor-infiltrating lymphocytes improve magee equation-based prediction of pathologic complete response in HR-Positive/HER2-Negative breast cancer. *Am J Clin Pathol* (2022) 158(2):291–9. doi: 10.1093/ajcp/aqac041

43. Jiang W, Wang S, Wan J, Zheng J, Dong X, Liu Z, et al. Association of the collagen signature with pathological complete response in rectal cancer patients. *Cancer Sci* (2022) 113(7):2409–24. doi: 10.1111/cas.15385

44. Rocco B, Sighinolfi MC, Sandri M, Puliatti S, Bianchi G. A novel nomogram for predicting ECE of prostate cancer. *BJU Int* (2018) 122(6):916–8. doi: 10.1111/bju.14503

45. Li J, Wang J, Ding Y, Zhao J, Wang W. Prognostic biomarker SGSM1 and its correlation with immune infiltration in gliomas. *BMC Cancer* (2022) 22(1):466. doi: 10.1186/s12885-022-09548-7

46. Zhou M, Zhu X. Construction and validation of a robust ferroptosis-associated gene signature predictive of prognosis in lung adenocarcinoma. *Med (Baltimore)* (2022) 101(16):e29068. doi: 10.1097/md.0000000000029068

47. Cong M, Feng H, Ren JL, Xu Q, Cong L, Hou Z, et al. Developing a predictive radiomics model for lymph node metastases in pre-surgical CT-based stage IA non-small cell lung cancer. *Lung Cancer* (2020) 139:73–9. doi: 10.1016/j.lungcan.2019.11.003

48. Hu X, Gong J, Zhou W, Li H, Wang S, Wei M, et al. Computer-aided diagnosis of ground glass pulmonary nodule by fusing deep learning and radiomics features. *Phys Med Biol* (2021) 66(6):065015. doi: 10.1088/1361-6560/abe735

49. Ren C, Zhang J, Qi M, Zhang J, Zhang Y, Song S, et al. Machine learning based on clinico-biological features integrated (18)F-FDG PET/CT radiomics for distinguishing squamous cell carcinoma from adenocarcinoma of lung. *Eur J Nucl Med Mol Imaging* (2021) 48(5):1538–49. doi: 10.1007/s00259-020-05065-6

50. Liu J, Cui J, Liu F, Yuan Y, Guo F, Zhang G. Multi-subtype classification model for non-small cell lung cancer based on radiomics: SLS model. *Med Phys* (2019) 46(7):3091–100. doi: 10.1002/mp.13551

51. Bera K, Schalper KA, Rimm DL, Velcheti V, Madabhushi A. Artificial intelligence in digital pathology - new tools for diagnosis and precision oncology. *Nat Rev Clin Oncol* (2019) 16(11):703–15. doi: 10.1038/s41571-019-0252-y

52. McDonald BR, Contente-Cuomo T, Sammut SJ, Odenheimer-Bergman A, Ernst B, Perdigones N, et al. Personalized circulating tumor DNA analysis to detect residual disease after neoadjuvant therapy in breast cancer. *Sci Transl Med* (2019) 11(504):eaax7392. doi: 10.1126/scitranslmed.aax7392

53. Huang CM, Huang CW, Ma CJ, Yeh YS, Su WC, Chang TK, et al. Predictive value of FOLFOX-based regimen, long interval, hemoglobin levels and clinical negative nodal status, and postchemoradiotherapy CEA levels for pathological complete response in patients with locally advanced rectal cancer after neoadjuvant chemoradiotherapy. *J Oncol* (2020) 2020:9437684. doi: 10.1155/2020/9437684

54. Heutink F, Koch V, Verbist B, van der Woude WJ, Mylanus E, Huinck W, et al. Multi-scale deep learning framework for cochlea localization, segmentation and analysis on clinical ultra-high-resolution CT images. *Comput Methods Programs BioMed* (2020) 191:105387. doi: 10.1016/j.cmpb.2020.105387

55. Zhou X. Automatic segmentation of multiple organs on 3D CT images by using deep learning approaches. *Adv Exp Med Biol* (2020) 1213:135–47. doi: 10.1007/978-3-030-33128-3_9

56. Jin C, Yu H, Ke J, Ding P, Yi Y, Jiang X, et al. Predicting treatment response from longitudinal images using multi-task deep learning. *Nat Commun* (2021) 12(1):1851. doi: 10.1038/s41467-021-22188-y

frontiers | Frontiers in Oncology

TYPE Review
PUBLISHED 04 January 2023
DOI 10.3389/fonc.2022.998222

OPEN ACCESS

EDITED BY
Sweet Ping Ng,
University of Melbourne, Australia

REVIEWED BY
Liyue Shen,
Stanford University, United States
Andrea Barucci,
National Research Council (CNR), Italy

*CORRESPONDENCE
Pengfei Rong
rongpengfei66@163.com
Wei Wang
cjr.wangwei@vip.163.com
Wei Li
weililx@csu.edu.cn
Li Zhou
zhouli723@csu.edu.cn

SPECIALTY SECTION
This article was submitted to Cancer Imaging and Image-directed Interventions, a section of the journal Frontiers in Oncology

RECEIVED 19 July 2022
ACCEPTED 22 November 2022
PUBLISHED 04 January 2023

CITATION
Liao J, Li X, Gan Y, Han S, Rong P, Wang W, Li W and Zhou L (2023) Artificial intelligence assists precision medicine in cancer treatment. *Front. Oncol.* 12:998222. doi: 10.3389/fonc.2022.998222

Artificial intelligence assists precision medicine in cancer treatment

Jinzhuang Liao[1], Xiaoying Li[1], Yu Gan[1], Shuangze Han[1], Pengfei Rong[1,2]*, Wei Wang[1,2]*, Wei Li[1,2]* and Li Zhou[1,2,3]*

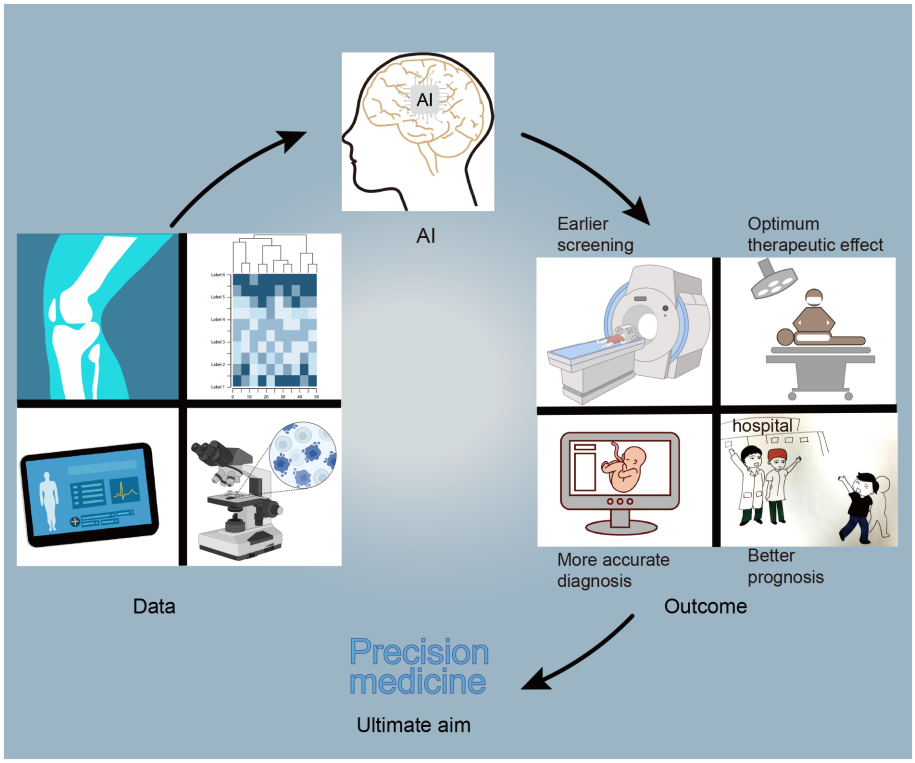
[1]Department of Radiology, The Third Xiangya Hospital of Central South University, Changsha, Hunan, China, [2]Cell Transplantation and Gene Therapy Institute, The Third Xiangya Hospital, Central South University, Changsha, Hunan, China, [3]Department of Pathology, The Xiangya Hospital of Central South University, Changsha, Hunan, China

Cancer is a major medical problem worldwide. Due to its high heterogeneity, the use of the same drugs or surgical methods in patients with the same tumor may have different curative effects, leading to the need for more accurate treatment methods for tumors and personalized treatments for patients. The precise treatment of tumors is essential, which renders obtaining an in-depth understanding of the changes that tumors undergo urgent, including changes in their genes, proteins and cancer cell phenotypes, in order to develop targeted treatment strategies for patients. Artificial intelligence (AI) based on big data can extract the hidden patterns, important information, and corresponding knowledge behind the enormous amount of data. For example, the ML and deep learning of subsets of AI can be used to mine the deep-level information in genomics, transcriptomics, proteomics, radiomics, digital pathological images, and other data, which can make clinicians synthetically and comprehensively understand tumors. In addition, AI can find new biomarkers from data to assist tumor screening, detection, diagnosis, treatment and prognosis prediction, so as to providing the best treatment for individual patients and improving their clinical outcomes.

KEYWORDS

artificial intelligence, precision medicine, omics, cancer, medical imaging

Abbreviations: AI, artificial intelligence; ML, machine learning; DL, deep learning; PM, precision medicine; DSS, decision support system; NGS, next-generation sequencing; WSI, whole slide imaging; CADs, computer-aided diagnosis system; CT, computed tomography; MRI, magnetic resonance imaging; CNN, convolutional neural networks; US, ultrasound; PET/CT, positron emission tomography/computed tomography; DLR, deep learning radiomics; CGHub, Cancer Genomics Hub; TCGA, The Cancer Genome Atlas; CCLE, Cancer Cell Line Encyclopedia; ICGC, International Cancer Genome Consortium; EGA, European Genome-phenome Archive; COSMIC, Catalogue Of Somatic Mutations In Cancer; SomamiR, Somatic mutations altering microRNA-ceRNA interactions; CTRP, Cancer Therapeutics Response Portal; gCSI, The Genentech Cell Line Screening Initiative; GDSC, Genomics of Drug Sensitivity in Cancer; NCI, National Cancer Institute; DepMap, Dependency Map; TCIA, The Cancer Immunome Database.



GRAPHICAL ABSTRACT

Introduction

Cancer is a severe threat to human health with a high mortality and a rising incidence rate (1). Several types of cancer can be cured if they are diagnosed and treated early. However, the treatment of cancer is not ideal at present. Cancer mortality rates remain high and continue to rise, including for prostate, colorectal, and cervical cancer (2). These tumors lack effective screening and treatment methods, resulting in patients not getting timely and effective treatment. Secondly, the heterogeneity of tumors is high, which can create great challenges in their treatment (3). Therefore, new diagnostic and treatment methods that are tailored to individual patients are needed. Precision medicine (PM) is a promising approach that takes individual genetics, environment and lifestyle into account and concentrates on clarifying, diagnosing and treating diseases to create a customized treatment plan for patients through obtaining multi-omics or multi-mode information from individuals (4). Furthermore, artificial intelligence (AI) uses computers or machines to carry out tasks by mimicking or emulating human intelligence, which mainly includes machine learning (ML) and deep learning (DL) (5). AI can process an enormous amount of information to promote the brand-new discovery of PM. AI has shown extraordinary potential in processing, mining and analyzing data and can use the data to develop different models to help achieve PM.

Tumors are generally caught sight of in the following two situations: one is the screening of high-risk groups (6). The other one is the discovery of tumors with clinical manifestations. After the cancer is detected, patients will receive further examinations, such as physical examination, imaging, pathology, and serum tumor markers (6). Based on these results, tumors will be accurately diagnosed, staged, and classified to help the patients benefit from precision treatment. AI can play a part in tumor prevention, screening, diagnosis, treatment, and prognosis prediction (7–10). After AI is injected into the clinical process, it will improve the detection rate of lesions and make the screening method more effective. Secondly, AI can promote the level of diagnosis by helping doctors distinguish between true and false disease progression (7). Finally, AI can calculate the advantages and disadvantages of each treatment scheme and provide the best treatment for patients. In addition, a framework diagram (Figure 1) is added to this article, which shows a series of processes from the discovery of tumor patients to the end of their diagnosis, treatment, and the changes that AI can bring.

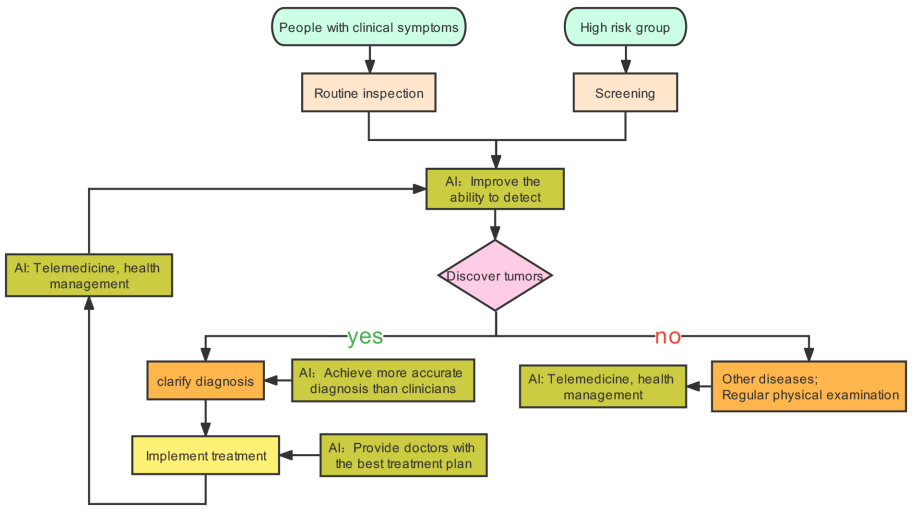


FIGURE 1
Possible changes caused by AI injection into clinical practice.

With the development of next-generation sequencing (NGS) technology, omics data, such as genomics, proteomics and transcriptomics, have been accumulated (11). Meanwhile, the massive growth and wide availability of patients' clinical data such as electronic medical records, clinical trial data, and medical images have led to the era of "big data" (12). The best analysis method is data analysis based on AI, since ML and DL can extract the hidden patterns, important information, and corresponding knowledge behind the data. Based on extracted data, information about the disease is obtained to help clinical analysis. For example, ML and DL can be used to analyze omics data to establish models, generate biomarkers related to diagnosis, classification, and prognosis, provide molecular changes such as DNA, RNA and protein, predict drug efficacy and therapeutic response, and develop targeted drugs (13). Furthermore, as compared with single-omics, multi-omics provides an opportunity to understand the information flow behind a disease (14). Multi-omics integration is crucial to the comprehensive understanding of complex biological processes. Combined with the new longitudinal experimental design, multi-omics can clarify the dynamic relationship between all layers of omics, distinguish the key roles or interactions in system exploitation or complicated phenotypes, clarify the causal relationship and functional mechanisms of complicated diseases, and promote the discovery of PM (15, 16). Quantitative image analysis is a suitable candidate for PM and can assist PM for cancer. ML and DL have been used for quantitatively extracting image features to establish models for diagnosis, monitoring, and predicting recurrence and metastasis, biomarkers and prognosis (17–21). AI can integrate the above data for comprehensive analysis of tumors for the development of a clinical decision support system (DSS) (22). With the continuous improvement of AI algorithms and the improvement of computer software and hardware, AI will mature and will be used more extensively in the medical field (Figure 2) in the future. Therefore, PM for tumors will great evolve.

In the present review, we first introduced the application of AI in omics, and then in pathology and medical imaging, and expanded on how these applications assist PM. Finally, we described the challenges and future directions of AI assisted PM for tumors.

AI-based big data assists PM for cancer

Big data technology mainly includes data analysis, mining, and sharing. It may play a revolutionary role in cancer diagnosis, treatment, prevention, and prognosis, but transforming data into available information to benefit patients is almost at a standstill (23, 24). A major reason for this is that data analysis significantly lags behind data generation (24). The reforms caused by "big data" have affected nearly all aspects of tumor research. For example, the technology can analyze data generated by NGS to discover commonly mutated genes, abnormal gene expression, and biomarkers in tumors for accurate diagnosis and prognosis prediction or to determine the cause of disease and develop targeted drugs for treatment (23, 24). The technology can analyze features that humans can and cannot see in medical images, and mine and filter these features to determine information related to diagnosis, treatment, and prognosis (25,

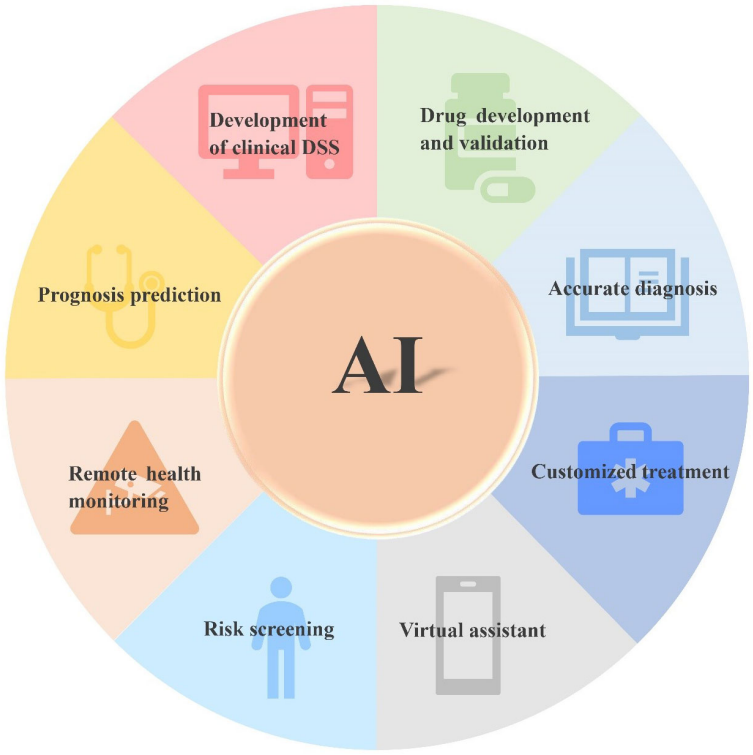


FIGURE 2
Application prospect of AI in tumor.

26). In addition, the technology can analyze patients' demographic and clinical data, as well as outcome information to predict the factors affecting the prognosis of cancer patients (27). In addition, AI is used to analyze, mine and process tumor-related data, build a health care provider platform based on a significant quantity of tumor-related data, efficiently solve the problem of difficult medical treatment for patients and reduce the waste of unnecessary medical resources (28). Big data reanalysis has been not been sufficiently taken advantage of so far, but we cannot ignore its potential. It can analyze the data in an existing database and provide new insights. For example, Borziak et al. discovered the dedifferentiation markers of liver cancer by using data from existing databases (29). The current big data technology is mainly used in certain fields, such as omics, pathological imaging and medical imaging. However, it does not combine data from multiple fields for data analysis, mining, and sharing, which leads to data not being comprehensively utilized and not meeting clinicians' and patients' needs. The challenges in the diagnosis, treatment and monitoring of cancer can be overcome by integrating omics and non-omics data. AI can play an important role in analyzing high-dimensional data-sets with complexity and heterogeneity, especially in multi-omics, intergroup methods and data integration, thus setting forth the cancer molecular mechanism, and identifying new dynamic diagnostic and prognostic biomarkers to provide accurate cancer care (30).

There are certain problems with the current data, such as poor data quality, unstructured databases, inadequate analytics, and lack of delivery (23, 31). Therefore, there is a need for a more authoritative and reliable prospective database. In addition, a longitudinal database is also needed to understand the cancer dynamics of patients in the whole study care continuum (23). Establishing a patient-centered collection of various data-sets will be crucial in the future (32). On this basis, AI-based big data analysis may automatically generate patient diagnosis, personalized treatment plans, and key information for prognostic prediction, thereby helping clinicians provide the best treatment for their patients.

AI assists tumor PM in omics

A large amount of data resources (Table 1) generated by NGS can provide key information about tumors. Combining the information with AI will help clarify the etiology and pathogenesis of tumors, and assist the accurate diagnosis, risk stratification, and disease subtype analysis (30, 33). Moreover, AI can identify new therapeutic targets, evaluate the sensitivity and

TABLE 1 Comprehensive omics database resources for building AI models.

Name	Main features	Web link
CGHub	Overall data repository; enormous data	https://cghub.ucsc.edu/
TCGA	Comprehensive database; enormous data	https://www.cancer.gov/about-nci/organization/ccg/research/structural-genomics/tcga
CCLE	Comprehensive database; enormous data	https://sites.broadinstitute.org/ccle
EGA	Overall data repository; enormous data	https://ega-archive.org/
ICGC	Comprehensive genomics data	https://dcc.icgc.org/
DepMap	High data quality; visualization	https://depmap.org/portal/
SomamiR	Correlation between cancer somatic mutation and miRNA	https://compbio.uthsc.edu/SomamiR/
COSMIC	largest and most comprehensive somatic mutation database; regularly-updated	https://cancer.sanger.ac.uk/cosmic
MethyCancer	integrated data of DNA methylation, cancer-related gene, mutation and cancer information	http://methycancer.psych.ac.cn/
CTRP	connecting sensitivity to cancer feature	https://portals.broadinstitute.org/ctrp/
gCSI	Large amount of transcriptomics data	https://pharmacodb.pmgenomics.ca/datasets/4
GDSC	Drug response data; genomics markers of drug sensitivity; update irregularly	https://www.cancerrxgene.org/
NCI60	Large amount of drug data and genomics data	https://discover.nci.nih.gov/cellminer/loadDownload.do https://dtp.cancer.gov/databases_tools/bulk_data.htm
canSAR	Comprehensive database; discovery drug	https://cansarblack.icr.ac.uk/
cBioPortal	Large amount of available data	https://www.cbioportal.org/datasets
UCSC	Synthetical genomics information	https://genome.ucsc.edu/
dbNSFP	Predictive data	http://bib.oxfordjournals.org/
NONCODE	database dedicated to non-coding RNAs	http://www.noncode.org/
CSD	The positive and negative training sets	http://bib.oxfordjournals.org/
TCIA	A great quantity of medical related image data sets	https://www.tcia.at/home
MSKCC	Cancer mutation databases	http://www.cbioportal.org/
ARCHS4	comprehensive processed mRNA expression data	https://maayanlab.cloud/archs4/

resistance of anticancer drugs, develop new targeted drugs, improve cancer immunotherapy, monitor the recurrence and evolution of the tumor, discover new biomarkers, and predict the prognosis and survival analysis of tumor patients (Figure 3) (34–40). In a few words, AI enables PM for cancer patients, bridging the distance between omics and the clinic. Since NGS produces high-dimensional and complex data, NGS methods for cancer diagnosis usually need higher-dimensional and deeper-seated data coverage to enhance the possibility of detecting a small number of tumor cell mutations and improve the sensitivity and accuracy of AI algorithms (41).

AI assists tumor PM in genomics

In recent years, genomics, which relies on nucleotide sequences for data analysis, has become more closely combined with clinical practice (30). The significant

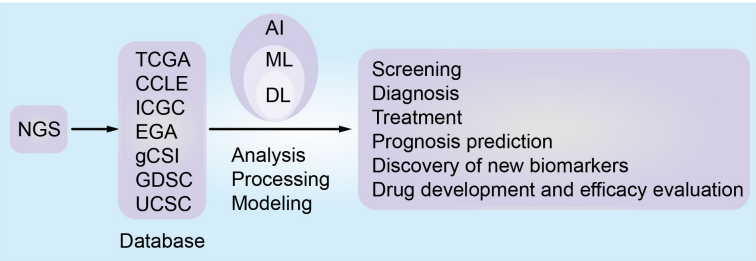


FIGURE 3
The role of artificial intelligence based on omics database in tumor precision medicine.

accumulation of data has improved the understanding of cancer vulnerability and has enabled us to increasingly anticipate noticeable treatment effects for tumor patients (42). The use of spatial and single-cell genomics may reconstruct the process of tumorigenesis to facilitate a more comprehensive understanding of tumor, decipher the unclear pathogenesis of human beings and develop targeted drugs based on this mechanism (43–47). The combination of ML and genomics data can assist the diagnosis of cancer subtypes, discovering new markers and drug targets, and understanding cancer-driving genes better, which contributes to providing customized treatment for patients (48). For example, Wang et al. developed a compounded deep network model that can diagnose lung cancer subtypes by mixing image-genomics data and can help biomedical professionals determine the potential therapeutic targets by the attention weights of the model (49). In addition, Vanderbilt et al. developed and validated a brand-new approach to identify DNA viruses from corresponding normal or tumor NGS specimens and inquire about virus-tumor type relevance without carrying out extra sequencing. Data on these viruses can provide information for the diagnosis and care of tumor patients. Their study illustrated the function of DNA viruses in the tumor (50). Sudhakar et al. used cancer genomics data and built a pan-cancer model to forecast and identify new driver genes (51). The identification of driving genes can help understand the carcinogenic mechanism and the design of treatment strategies, which has an important biological and clinical significance (52).

AI assists PM for tumors in transcriptomics

Transcriptomics is a powerful means to evaluate all transcripts produced during metabolism (30). Transcriptomics have expanded our knowledge of cancer occurrence and development, tumor microenvironment, and immune-oncology, and can directly determine gene expression levels and analyze the activation of related molecular pathways (53, 54). Transcriptomics is a bridge between genomics and proteomics, mainly involving quantitative reverse-transcription-polymerase chain reaction, microarrays and NGS (RNA-sequencing) (55). Since RNA-sequencing has a higher accuracy in measuring gene expression, it is considered the gold standard for high-throughput gene expression screening (55, 56). Through data mining or more complex mathematical approaches using ML or DL, the features are extracted to facilitate tumor screening and early diagnosis, discover new or previously unknown cancer biomarkers and potential therapeutic targets, as well as drug prioritization, and predict cancer drug sensitivity and prognosis (53–55, 57–62). For example, Warnat-Herresthal et al. found that ML-based transcriptomics can assist in the diagnosis of acute myeloid leukemia (63). Moreover, Ben Azzouz et al. used an ML approach based on transcriptomics data to calculate triple-negative breast cancer subtypes, in order to overcome the barrier of heterogeneity in the treatment of the disease (64). Finally, some ML-based transcriptomics have also been used in the development of prognostic biomarkers for prostate cancer (65), the diagnosis of colorectal cancer (66), and the prediction of immune response (67).

AI assists tumor PM in proteomics

Proteomics can provide comprehensive and quantitative information about proteins in tissues, blood and cell samples (68). Protein expression profiles generated by proteomics and ML-based profile analysis can identify more specific and sensitive protein biomarkers than other single-omics. These biomarkers can diagnose cancer, predict prognosis (69), reveal critical signaling pathways behind disease mechanisms (70, 71), determine new therapeutic targets, evaluate drug therapy efficacy and toxicity (72), and predict therapeutic responses, recurrence, and metastasis (73, 74). Recently, Henry et al. proposed a method of drug ranking using ML to predict drug response using proteomics data, and prioritize drugs in order to identify the most suitable drug for each patient (75). In addition, Federica et al. built a clearer and more transparent DSS to assist in diagnosing high-grade serous ovarian cancer (76). Therefore, AI-based proteomics may play an important role in the accurate diagnosis and treatment of tumors in the future.

Besides the widely used omics data mentioned above, other omics data (metabolomics, immunomics and microbiome data) are also used (77). For instance, disposing of metabolomics data by AI can assist the diagnosis (78, 79), the of treatment response evaluation (80–82), discovery of new biomarkers (83, 84), and determination of patient tolerance (85) and cancer status (invasive or non-invasive) (86). Moreover, the AI model based on immunomics data can forecast the emergency immune characteristics of tumor patients (87).

AI assists tumor PM in multi-omics

Although the current single-omics data can be used for diagnosis, treatment and prediction, they cannot thoroughly and systematically reflect the molecular changes of a tumor (88). Therefore, it is necessary to integrate multi-omics data to comprehensively understand the tumor information and its dynamic development process to screen and accurately diagnose patients, develop tailored treatment strategies, predict prognosis, and monitor recurrence and metastasis (89–92). Some approaches and algorithms of using AI to analyze multi-omics data comprehensively include clustering, factorization, feature transformation, networks-based means and feature extraction (89). These approaches can be used for stratifying medicine,

discovering biomarkers (93), pathway analysis, and drug reuse or discovery (89, 94) (Table 2). For example, Ma et al. introduced a new approach that can analyze multi-omics information and related knowledge to reveal the complex relationship between molecular features and clinical characteristics (114). In addition, Wang et al. developed a molecular algorithm for early cancer detection, which is used to confirm malignant cellular tumors according to the spectrum of changes in single-cell copy numbers based on doubtful cells in humoral, resulting in a well-defined cancer diagnosis (115). Furthermore, Olivier B et al. have developed an integrated framework of DL and ML, which can use multi-omics data to accurately predict survival and prognosis (95). Furthermore, except for the above commonly used omics, studies have also focused on linking radiomics with genomics and transcriptomics for accurate diagnosis (116). A multi-task DL framework called OmiEmbed, which can analyze and process several kinds of omics data and simultaneously handle multiple tasks has recently emerged. This disruptive technological breakthrough will significantly promote the development of PM (117). The application of AI to integrate multi-omics data is shown in Table 2.

AI in pathology assists the accurate diagnosis of tumors

Pathological analysis is considered the gold standard of the clinical diagnosis of tumors (118). However, the current shortage of clinical pathologists and their reliance on subjective consciousness for diagnosis leads to low repeatability and unequal diagnostic levels of clinical pathologists, which is not helpful for clinicians' decision-making with regards to treatment (119). Computational pathology has seen significant developments from the use of improved AI algorithms and computing power. With the use of image analysis of digital pathology, ML and DL, AI has been used to evaluate whole slide imaging (WSI) and produce computer-aided diagnosis systems (CADs), as well as aid cancer prognosis prediction (120–124). At present, the diagnostic ability of the AI-based diagnostic model can be comparable to or even surpass that of experts (125). In combination with human experts, the precision of diagnosis can be even better. It also has the advantages of being less time-consuming, and having a high efficiency and repeatability. Therefore, an increasing number of AI models are being developed to assist clinical pathologists and reduce their workload (120). For example, Ho et al. proposed incorporating AI models into the pathological workflow as the first reader, the second reader, triage, and pre-screening (120, 126, 127).

Aiding diagnosis through ML and DL mainly includes three steps: The first step includes data preprocessing, such as image sharpening, masking and smoothing, image graying and color normalization, data standardization, and data annotation. The second step includes the division of nucleus/tissue. During the third step, models are established for training and verification, diagnosis and prediction (128, 129). For example, the computer-aided diagnosis and prognosis prediction model of WSI based on hematoxylin and eosin staining, can screen, classify and grade tumors (130), and identify micro-metastasis in lymph nodes (25), and microsatellite instability (128). It can also predict the changes at the molecular level (131), the risk of metastasis and recurrence after surgical resection (132) and disease-specific survival (133). Moreover, Armin et al. used a DL model based on digital images of immunohistochemistry to calculate the risk of mortality (134). The role of AI in the digital pathological image is summarized in Figure 4.

Traditional ML methods analyze pathological tissue by manually extracting mainly morphological, textural (135) and spatial features (118). It is easier to understand and explain than DL, and its training sample size is small, especially suitable for the analysis of rare tumor subtypes with a limited sample size. However, manually extracted features have the following limitations (118): They are extracted in an unsupervised way and have nothing to do with the subsequent WSI analysis tasks (136). Only the surface features of the input image can be learned, which is not enough to show the complex features of WSI. It is exceptionally arduous to process multiple WSI images at the same time and the processing speed is slow (137). Compared with ML, DL can automatically extract the features in the image for analysis and can also efficiently process a considerable amount of data (129). The DL model has good scalability (138), but it is easy to overfit, resulting in the low generalization ability of the model. Furthermore, it is characterized by low interpretability and cannot be trusted by clinical pathologists (139). Recently, diagnostic models combining various methods of traditional ML, DL have been developed to integrate their advantages for accurate diagnosis and prediction (129). For example, Sengupta et al. have proposed a novel deep hybrid learning model based on nuclear morphology for accurately diagnosing ovarian cancer (129). It can be safely assumed that in the near future more high-performance prediction models will be developed and enter the clinic to assist clinical pathologists in accurate diagnosis and prognosis prediction, consequently providing accurate and personalized medical care for patients.

AI assists PM for tumors in medical imaging

Imaging is one of the indispensable tools for screening, diagnosis, treatment and follow-ups for several types of tumors. At present, the performance of imaging examination equipment such as thin-layer computed tomography (CT) and multi-parameter magnetic resonance imaging (MRI), is continuously improved, detecting more subtle lesions and

TABLE 2 Application of AI in the Integration of Multi-omics.

Clinical application	Data	Model/Algorithm	Performance	References
cancer prognosis and survival prediction	RNA-Seq, Methylation, and miRNA	semi-supervised flexible hybrid machine-learning framework	Not applicable	Poirion, O.B., et al. (95)
breast cancer subtype identification	mRNA expression, miRNA expression and DNA methylation	deep learning fusion clustering framework	0.664	Shuangshuang, L., et al. (91)
cancer susceptibility prediction	copy number variations, miRNA expression, and gene expression	multimodal convolutional autoencoder model	0.9625	Karim, M.R., et al. (96)
identifying Neuroblastoma subtypes	gene expression, copy number alterations, Sequencing Quality Control project	deep learning	0.74	Zhang, L., et al. (97)
predict the survival of patients with lung cancer	TCGA	unsupervised learning	0.99	Takahashi, S., et al. (90)
survival stratification of gastric cancer	transcriptomics and epigenomics	bidirectional deep neural networks	0.76	Xu, J.M., et al. (98)
pan-cancer metastasis prediction	RNA-Seq, microRNA sequencing, and DNA methylation	deep learning	0.8885	Albaradei, S., et al. (92)
ovarian cancer subtypes identification	mRNA-seq, miRNA-seq, copy number variation, and the clinical information	deep learning	0.583	Guo, L. Y., et al. (99)
drug repurposing	copy number alteration, DNA methylation, gene expression, pharmacological characteristics for cancer cell lines	deep learning	0.84	Wang, Y., et al. (94)
predicting lung adenocarcinoma prognostication	mRNA, miRNA, DNA methylation and copy number variations	deep learning	0.65	Lee, T.-Y., et al. (100)
Diagnostic Classification of Lung Cancer	mRNA expression, miRNA-seq data, and DNA methylation data	deep transfer Learning	0.824	Zhu, R., et al. (101)
predicting effective therapeutic agents for breast cancer	copy number variations, miRNA, mutation, RNA, protein expression and methylation	deep learning	0.94	Khan, D. and S. Shedole (102)
predicting survival prognosis for glioma patients	transcription profile, miRNA expression, somatic mutations, copy number variation, DNA methylation, and protein expression	deep learning	0.990	Pan, X., et al. (103)
Diagnostic classification of cancers	mRNA expression, miRNA-seq, DNA methylation data and clinical information	XGBoost	0.595-0.872	Ma, B., et al. (104)
identify tumor molecular subtypes	copy number, mRNA, miRNA, DNA methylation and other omics data	consensus clustering and the Gaussian Mixture model	Not applicable	Yang, H., et al. (105)
predicting outcome for patients with hepatocellular carcinoma	DNA methylation and mRNA expression data	unsupervised machine-learning	Not applicable	Huang, G. J., et al. (106)
predicting the Gleason score levels of prostate cancer and the tumor stage in breast cancer	gene expression, DNA methylation, and copy number alteration	gene similarity network based on uniform manifold approximation and projection and convolutional neural networks	0.99	ElKarami, B., et al. (107)
patient classification, tumor grade classification, cancer subtype classification	mRNA expression, DNA methylation, and microRNA expression data	Multi-Omics Graph cOnvolutional NETworks	Not applicable	Wang, T. X., et al. (108)
cancer prognosis prediction	mRNA, miRNA, DNA methylation, and copy number variation	denoising Autoencoder	Not applicable	Chai, H., et al. (109)
cancer subtype classification	gene expression, miRNA expression and DNA methylation data	hierarchical integration deep flexible neural forest framework	0.885	Xu, J., et al. (110)
Prediction of prognosis of cancer	single nucleotide polymorphism, copy number variant, gene expression, and DNA methylation data	deep learning	0.67-0.88	Park, C., et al. (111)
tumor Stratification	deoxyribonucleic acid methylation, messenger ribonucleic acid expression data, and protein–protein interactions	Network Embedding; supervised learning; unsupervised clustering algorithm	0.91	Li, F., et al. (112)
discovery of cancer subtypes	mRNA expression, miRNA expression, DNA methylation, and copy number alterations	end-to-end variational deep learning-based clustering method; Variational Bayes	Not applicable	Rong, Z., et al. (113)

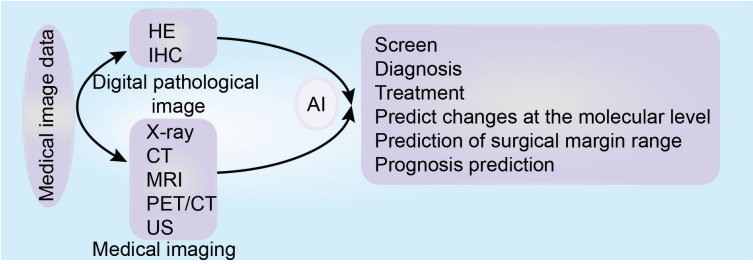


FIGURE 4
The role of artificial intelligence based on medical images (digital pathological image and medical imaging) in tumor precision medicine.

producing increasingly complex data. However, this requires more time and effort from radiologists to make a diagnosis, increasing their workload. AI-based data analysis can effectively process huge amounts of data, with the CADs model based on medical images exhibiting high precision and standard (140–142). Introducing AI into clinical practice will help radiologists make diagnostic decisions quickly, accurately, and efficiently, will help focus their energy on advanced decision-making, and promoting accurate medical treatment and personalized treatment for tumors (141, 143). The role of AI in medical imaging is summarized in Figure 4.

AI assists radiologists in accurately diagnosing tumors

AI has three main tasks in tumor imaging: Detecting, characterizing and monitoring tumors (144). Detection refers to the location of the region of interest in the image. Characterization includes tumor diagnosis, and staging. Monitoring refers to the monitoring of the changes in tumors with time (144). The process of ML-assisted tumor detection and diagnosis is as follows: Image data acquisition, image preprocessing, segmentation of regions of interest, feature selection, establishing the model and carrying out training, verification and testing (145). Among them, feature selection is the most important step, since it is most related to the model's performance (145). Moreover, DL can automatically extract the feature from the image. Therefore, recent research has increasingly focused on the DL to build "fusion" models for the diagnosis of tumor lesions, including classification, grading and staging, which have been proven to be effective (146–148). For example, Chougrad et al. have built CADs based on deep convolutional neural networks (CNN) to aid radiologists in categorizing breast X-ray masses (149). Moreover, Misra et al. have developed a highly robust DL model for categorizing benign and malignant neoplastic lesions of the breast. The model can improve the accuracy of breast cancer classification by correcting patients whose traditional methods are misclassified (150). Overall, these models can make radiologists more effective in detecting and diagnosing tumors faster, and will likely be popularized and applied to clinical medical treatment soon.

AI assists PM for tumors diagnosed *via* medical imaging

The choice of treatment depends on the outcome of the diagnosis. For example, if the detected lesion is benign, it can reduce unnecessary surgical resection and other treatments, and provide more targeted medical management for patients. In addition, the use of radiation imaging, a non-invasive diagnostic method, can protect patients from the discomfort caused by biopsy and avoid the risk of implant metastasis in pathological biopsy (151). Moreover, preoperative evaluation of tumor grading prediction using radiology can help select more appropriate treatment options for patients and avoid unnecessary surgery, thereby reducing the patient's medical burden and avoiding excessive medical treatment (152, 153).

In addition to assisting in accurate diagnosis, AI can also play a significant role in prognosis prediction and treatment of patients. It can predict patient viability based on imaging features and determine the level of treatment needed to achieve optimal survival. The prediction of recurrence, metastasis, surgical margins and therapeutic responses can be used to formulate an optimal therapeutic strategy for individual patients (21, 26).

Accurately identifying and evaluating lesion before an operation can help create appropriate treatment plans for patients and avoid unnecessary treatment measures such as surgery, postoperative radiotherapy, and chemotherapy, which is beneficial to both patients and doctors. For instance, Zhao et al. built a neoplasm grade forecast model of pre-operative G1/2 assessment of nonfunctional pancreatic neuroendocrine tumors by using radiomics to analyze the multi-slice helical

CT images (152). In addition, Xie et al. used a CT-based radiomics ML method to distinguish pancreatic mucinous cystic neoplasm from atypical serous cystadenomas prior to surgery (154). The classification and types of tumors are different, and their treatment methods are inconsistent. For example, according to the model established by Zhao and Xie, if the preoperative prediction result is a high-risk or high-grade tumor, it is necessary to strengthen the follow-up treatment, such as postoperative neoadjuvant chemotherapy or radiotherapy (152, 154). Depending on whether the lymph node is metastatic or not, clinicians will choose different treatment options for patients. Therefore, the detection of lymph node metastasis is extremely important. For example, Song et al. established and verified a radiomics nomogram based on dynamic contrast-enhanced MRI, which can predict metastasis of axillary lymph nodes in mastocarcinoma (155). Similarly, Eresen et al. used the radiomics-derived model established by ML to detect metastatic lymph nodes in colorectal cancer patients (156). Predicting preoperative tumor markers and imaging biomarkers can lead to better clinical decision-making and help provide the best treatment for patients. For example, Guo et al. developed LR and LR-SVMSMOTE models based on CT radiomics to predict thyroid cartilage invasion in certain cancers types, such as hypopharyngeal squamous cell carcinoma and laryngeal carcinoma (157). Similarly, Akbari et al. combined the advanced mode analysis and ML method of multi-parameter MRI to provide the prediction space map of tumor invasion and early recurrence possibility to provide more targeted surgical and radiotherapy strategies for tumor patients, aiming to maximize the treatment effect while maintaining neurological function (158).

The DL model based on one of the subsets of AI can assist radiotherapy or oncology doctors in accurately outlining tumor targets, reducing the time doctors take to manually segment images as well as reducing the variation between observers (159–163). The model can predict and verify the therapeutic dose, and allows for the dose prescription to be changed in time to reduce the impact on the surrounding normal tissue, prevent unnecessary radiation, and reduce the occurrence of adverse reactions (164, 165). The model can evaluate the efficacy of radiotherapy and chemotherapy, as well as the therapeutic response, so as to achieve better-personalized prescriptions for patients (166–168). For example, Ermiş et al. have used DL to depict fully automatic brain resection cavity delineation in patients with glioblastoma (169). Likewise, Zhou et al. have developed and tested a three-dimensional DL model capable of predicting the dosage distribution of three-dimensional volume units to carry out intensity-modulated radiation therapy (170). Establishing a dose distribution model prior to treatment helps adjust the dose distribution in advance and reduce the probability of complications from radiotherapy. ML and DL can also assist in post-radiotherapy management, such as distinguishing between the true and false progression of the tumor, radiation necrosis and tumor recurrence, and promoting clinical medical decision-making, thus improving PM (171–174). In addition, the image-based AI model can also assist radiologists in treatment evaluation, including predicting the response of individual cancer patients to chemotherapy or immunotherapy, and monitoring recurrence and metastasis (175–177). Several radiomics-based ML and DL models can predict patient prognosis, such as recrudesce-free and progression-free survival, survival rate, mortality, surgical results, postoperative metastasis and recurrence. According to the prediction results of the model, the corresponding processing is carried out to create a customized treatment scheme for patients and improve the treatment effect and later quality of life. For example, patients with lower overall survival prediction need more intensive treatment. Patients with poor surgical results may want to consider changing the surgical method or choosing non-surgical treatment. Patients with a higher risk of tumor recurrence and metastasis should continue to receive neoadjuvant radiotherapy and chemotherapy (21, 26, 178–180).

Current challenges and future prospects

Although AI is expected to help improve a series of clinical applications against cancer, it does have some challenges and limitations.

One limitation is the lack of standards and imbalance in the data used to build the model (181, 182). These disordered data will lead to the low robustness of the model and be unfit for constructing a DL model with high generalization and precision (129). For example, medical imaging data is generated under different parameters for different devices (5). Digital pathological images are produced by staining with different dyes. The non-standard operation of pathological specimen collection will also affect the quality of pathological images (122). Irregularities in data collection lead to bias. Omics data are also noisy and heterogeneous (183). These data sets, which are generated by different technologies and standards limit the promotion and generalization of AI models, thus limiting their application in clinical practice. In addition, the sample size of training samples and verification samples used to establish the AI model is small, which can easily cause the overfitting of the model (120). Finally, integrating various types of data, such as genomics, transcriptomics, proteomics, metabolomics, immunomics, electronic health records, clinical medical records, pathology and medical images, will help evaluate the tumor comprehensively and develop the best treatment plan for patients (77, 184, 185). Therefore, it is necessary to establish an extensive comprehensive standardized database. However,

many types of data have a multi-scale nature, which makes the mechanical connection between data elusive. The biological knowledge of connecting all these variables in a single model is limited, so many data variables will be omitted from the model development process (186). Recently, some studies have combined dynamic modeling and ML to promote the integration of mathematics and clinical oncology. This method can integrate multiple types of data for personalized prediction to assist PM (187). More in-depth research to promote the combination of mechanical modeling and ML approaches is required in the future, so that mathematical oncology can be introduced into clinics. Building deep fusion models such as multi-modality DL is the primary method to develop AI models that can effectively integrate multimodal data information. However, the current method mainly focuses on representation fusion (feature- and decision-level fusion). The main challenge of this method is that the data is highly dimensional, noisy, heterogeneous, and has a small sample size, and there will be data loss during processing (91, 96, 188, 189). Here are some methods to address these barriers: T-distributed stochastic neighborhood embedding, autoencoder, random forest deep feature selection, a stacked autoencoder, gradient descent method, multi-view factorization autoencoder, co-expression network analysis, and regulation techniques (88, 91, 114, 190–193).

In addition, data from patients are governed by privacy laws (194). The lack of supervision of these data may lead to breaching patient privacy rules; therefore, appropriate intervention and the improvement of laws and regulations are required. Certainly, studies have focused on solving the privacy problem with regards to patient data. Under the same performance, the privacy vulnerability is reduced by vocabulary selection means (194). At the same time, fusing these data should comply with the principles of medical ethics.

Another limitation is that AI algorithms have been regarded as "black boxes" (139), since the process of their output results is unknown and unexplained, which makes clinicians have low trust in AI and a low willingness to introduce it into the clinical workflow (181). Developing a knowledge-embedded DL model for multi-dimensional data fusion is a hopeful means for this problem (195). An increasing number of studies on AI interpretability and transparency are being conducted. The research aims to make AI transparent and interpretable so that its results can be convincing and easy to introduce into the clinical (128). Traditional ML and DL have their advantages and disadvantages, prompting more research on hybrid learning methods. The current research results show that the hybrid learning model exhibits a better performance, better interpretability, higher transparency and more accurate prediction (129, 148). Although AI has shown the ability to surpass people, since it cannot produce 100% correct results, doctors' participation is still required for the final diagnosis and treatment decisions (183). Future research will focus on improving the interpretability and performance of AI, because it is an important step for AI to realize clinical application (31).

AI model may be necessary to carry out clinical experiments similar to clinical drug trials because when AI models are initially applied to clinical practice, unexpected clinical conditions will inevitably occur. Only through continuous practice can we better find problems and solve them to improve AI models. However, as AI differs from drugs, its clinical trial plan should also be distinguished from drug clinical trials.

AI can be deployed before, during, and after diagnosis, which respectively stands for cancer prevention, screening, diagnosis, and treatment. For example, before diagnosis, AI can be combined with gene detection, endoscopic examination, and other technology to predict the risk of disease occurrence earlier and carry out risk management for patients to reduce the possibility of disease occurrence (196). Augmented or virtual reality can simulate experiences to improve patient compliance (165). During diagnosis, AI can roughly ask patients for relevant information and process it. Secondly, AI can analyze medical image, blood biochemistry, and other clinical overall data to automatically generate a diagnosis report and a variety of feasible as well as optimal treatment methods. Furthermore, after diagnosis, AI can assist clinicians cut down the damage and maximizing the benefits for patients in surgery, radiotherapy, and chemotherapy. The deployment of AI in clinical practice will improve the efficiency of clinicians, reduce the possibility of clinical errors, improve the medical status in areas with low medical levels, and reduce unnecessary procedures, interventions, and medical costs. In a word, patients and doctors will benefit from AI to achieve a win-win situation.

Most people believe that AI cannot replace doctors (7). AI is an assistant in clinical practice, so the final decision must be made by doctors; the responsibility should also be borne by doctors. However, the clinician cannot control AI because it can make self-development and its development process illegible. Therefore, doctors should not be fully responsible for AI errors. Despite that, when using AI, clinicians should not lose their ability to doubt AI to make accurate diagnoses and treatments and develops the doctor-patient relationship in a sound direction.

Conclusion

AI has shown promising results in certain fields of oncology, including tumor screening, detection, diagnosis, treatment, and prognosis prediction. With the progress of AI, the improvement of computer performance, and the explosive growth of various data, new learning methods, such as the hybrid learning method, will continue to emerge, further improving the overall performance of the model, such as efficient data analysis and accurate prediction. The recent model generated by the ML and DL that can analyze various data sets will also improve the

prospects of PM. In conclusion, AI-assisted PM can help detect, diagnose and treat cancer early, as well as assist in the selection of the best treatment scheme, consequently improving the prognosis of patients and improving their treatment results.

Author contributions

JL collected the related papers and was a significant contributor to writing the manuscript. XL, YG, and SH made figures and tables. PR, WW, WL and LZ revised the article. PR, WW, WL and LZ initiated the study and revised the manuscript. All authors contributed to the article and approved the submitted version.

Funding

This work was supported by China's National Natural Science Foundation [grant numbers 81972837, 82003203, 82071986, 81771827, 81971721], and the Natural Science Foundation of Hunan Province [grant numbers 2021JJ31011, 2021JJ41058, 2021RC4017].

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. Bray F, Ren JS, Masuyer E, Ferlay J. Global estimates of cancer prevalence for 27 sites in the adult population in 2008. *Int J Cancer* (2013) 132(5):1133–45. doi: 10.1002/ijc.27711

2. Torre LA, Siegel RL, Ward EM, Jemal A. Global cancer incidence and mortality rates and trends–an update. *Cancer Epidemiol Biomarkers Prev* (2016) 25(1):16–27. doi: 10.1158/1055-9965.EPI-15-0578

3. Consortium, I.T.P.-C.A.o.W.G. Pan-cancer analysis of whole genomes. *Nature* (2020) 578(7793):82–93.

4. MacEachern SJ, Forkert ND. Machine learning for precision medicine. *Genome* (2021) 64(4):416–25. doi: 10.1139/gen-2020-0131

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

6. Zhou J, Sun H, Wang Z, Cong W, Wang J, Zeng M, et al. Guidelines for the diagnosis and treatment of hepatocellular carcinoma (2019 edition). *Liver Cancer* (2020) 9(6):682–720. doi: 10.1159/000509424

7. Le Berre C, Sandborn WJ, Aridhi S, Devignes M-D, Fournier L, Smail-Tabbone M, et al. Application of artificial intelligence to gastroenterology and hepatology. *Gastroenterology* (2020) 158(1):76–+. doi: 10.1053/j.gastro.2019.08.058

8. Senan EM, Jadhav ME, Rassem TH, Aljaloud AS, Mohammed BA, Al-Mekhlafi ZG. Early diagnosis of brain tumour MRI images using hybrid techniques between deep and machine learning. *Comput Math Methods Med* (2022) 2022. doi: 10.1155/2022/8330833

9. Rajput G, Agrawal S, Biyani K, Vishvakarma SK. Early breast cancer diagnosis using cogent activation function-based deep learning implementation on screened mammograms. *Int J Imaging Syst Technol* (2022) 32(4):1101–18. doi: 10.1002/ima.22701

10. Qiu H, Ding S, Liu J, Wang L, Wang X. Applications of artificial intelligence in screening, diagnosis, treatment, and prognosis of colorectal cancer. *Curr Oncol* (2022) 29(3):1773–95. doi: 10.3390/curroncol29030146

11. Kaczmarek E, Pyman B, Nanayakkara J, Tuschl T, Tyryshkin K, Renwick N, et al. Discriminating neoplastic from nonneoplastic tissues using an miRNA-based deep cancer classifier. *Am J Pathol* (2022) 192(2):344–52. doi: 10.1016/j.ajpath.2021.10.012

12. Hwang K-T. Clinical databases for breast cancer research. In: Noh D-Y, Han W, Toi M, editors. *Translational research in breast cancer*. Singapore: Springer Singapore (2021). p. 493–509.

13. Picard M, Scott-Boyer MP, Bodein A, Perin O, Droit A. Integration strategies of multi-omics data for machine learning analysis. *Comput Struct Biotechnol J* (2021) 19:3735–46. doi: 10.1016/j.csbj.2021.06.030

14. Hasin Y, Seldin M, Lusis A. Multi-omics approaches to disease. *Genome Biol* (2017) 18(1):83. doi: 10.1186/s13059-017-1215-1

15. Bodein A, Scott-Boyer M-P, Perin O, Lê Cao K-A, Droit A. Interpretation of network-based integration from multi-omics longitudinal data. *Nucleic Acids Res* (2022) 50(5):e27–7. doi: 10.1093/nar/gkab1200

16. Song M, Greenbaum J, Luttrell J, Zhou WH, Wu C, Shen H, et al. A review of integrative imputation for multi-omics datasets. *Front Genet* (2020) 11. doi: 10.3389/fgene.2020.570255

17. Lambin P, Leijenaar RTH, Deist TM, Peerlings J, de Jong EEC, van Timmeren J, et al. Radiomics: the bridge between medical imaging and personalized medicine. *Nat Rev Clin Oncol* (2017) 14(12):749–62. doi: 10.1038/nrclinonc.2017.141

18. Ibrahim A, Primakov S, Beuque M, Woodruff HC, Halilaj I, Wu G, et al. Radiomics for precision medicine: Current challenges, future prospects, and the proposal of a new framework. *Methods* (2021) 188:20–9. doi: 10.1016/j.ymeth.2020.05.022

19. Refaee T, Wu G, Ibrahim A, Halilaj I, Leijenaar RTH, Rogers W, et al. The emerging role of radiomics in COPD and lung cancer. *Respiration* (2020) 99(2):99–107. doi: 10.1159/000505429

20. Gillies RJ, Kinahan PE, Hricak H. Radiomics: Images are more than pictures, they are data. *Radiology* (2015) 278(2):563–77.

21. Preuss K, Thach N, Liang X, Baine M, Chen J, Zhang C, et al. Using quantitative imaging for personalized medicine in pancreatic cancer: A review of radiomics and deep learning applications. *Cancers* (2022) 14(7). doi: 10.3390/cancers14071654

22. Xu Q, Zhai JC, Huo CQ, Li Y, Dong XJ, Li DF, et al. OncoPDSS: an evidence-based clinical decision support system for oncology pharmacotherapy at the individual level. *BMC Cancer* (2020) 20(1):740. doi: 10.1186/s12885-020-07221-5

23. Barker AD, Lee JS. Translating "Big data" in oncology for clinical benefit: Progress or paralysis. *Cancer Res* (2022) 82(11):2072–5. doi: 10.1158/0008-5472.CAN-22-0100

24. Malone ER, et al. Molecular profiling for precision cancer therapies. *Genome Med* (2020) 12(1):8. doi: 10.1186/s13073-019-0703-1

25. Chuang W-Y, Chen C-C, Yu W-H, Yeh C-J, Chang S-H, Ueng S-H, et al. Identification of nodal micrometastasis in colorectal cancer using deep learning on annotation-free whole-slide images. *Modern Pathol* (2021) 34(10):1901–11. doi: 10.1038/s41379-021-00838-2

26. Luo H-S, Chen Y-Y, Huang W-Z, Wu S-X, Huang S-F, Xu H-Y, et al. Development and validation of a radiomics-based model to predict local progression-free survival after chemo-radiotherapy in patients with esophageal squamous cell cancer. *Radiat Oncol* (2021) 16(1):201. doi: 10.1186/s13014-021-01925-z

27. Wu W, Yang J, Li D, Huang Q, Zhao F, Feng X, et al. Competitive risk analysis of prognosis in patients with cecum cancer: A population-based study. *Cancer Control* (2021) 28:1073274821989316. doi: 10.1177/1073274821989316

28. Zheng J, Gao ZJ, Pu LX, He MJ, Fan JP, Wang S, et al. Analysis of tumor disease patterns based on medical big data. *J Med Imaging Health Inf* (2021) 11 (2):478–86. doi: 10.1166/jmihi.2021.3306

29. Borziak K, Finkelstein J. Utilizing shared big data to identify liver cancer dedifferentiation markers. *Stud Health Technol Inform* (2022) 289:73–6. doi: 10.3233/SHTI210862

30. Patel SK, George B, Rai V. Artificial intelligence to decode cancer mechanism: Beyond patient stratification for precision oncology. *Front Pharmacol* (2020) 11. doi: 10.3389/fphar.2020.01177

31. Kondylakis H, Ciarrocchi E, Cerda-Alberich L, Chouvarda I, Fromont LA, Garcia-Aznar JM, et al. Position of the AI for health imaging (AI4HI) network on metadata models for imaging biobanks. *Eur Radiol Exp* (2022) 6(1). doi: 10.1186/s41747-022-00281-1

32. Marti-Bonmati L, Alberich-Bayarri A, Ladenstein R, Blanquer I, Segrelles JD, Cerda-Alberich L, et al. PRIMAGE project: predictive in silico multiscale analytics to support childhood cancer personalised evaluation empowered by imaging biomarkers. *Eur Radiol Exp* (2020) 4(1):22. doi: 10.1186/s41747-020-00150-9

33. Song H, Ruan C, Xu Y, Xu T, Fan R, Jiang T, et al. Survival stratification for colorectal cancer via multi-omics integration using an autoencoder-based model. *Exp Biol Med* (2022) 247(11):898–909. doi: 10.1177/15353702211065010

34. Trubicka J, Grajkowska W, Dembowska-Baginska B. Molecular markers of pediatric solid tumors-diagnosis, optimizing treatments, and determining susceptibility: Current state and future directions. *Cells* (2022) 11(7):1238. doi: 10.3390/cells11071238

35. Kawazu M, Ueno T, Saeki K, Sax N, Togashi Y, Kanaseki T, et al. HLA class I analysis provides insight into the genetic and epigenetic background of immune evasion in colorectal cancer with high microsatellite instability. *Gastroenterology* (2022) 162(3):799–812. doi: 10.1053/j.gastro.2021.10.010

36. Xu F, Chen JH, Huang DH. Pan-cancer analysis identifies FAM49B as an immune-related prognostic maker for hepatocellular carcinoma. *J Cancer* (2022) 13 (1):278–89. doi: 10.7150/jca.65421

37. Johnson BE, Creason AL, Stommel JM, Keck JM, Parmar S, Betts CB, et al. An omic and multidimensional spatial atlas from serial biopsies of an evolving metastatic breast cancer. *Cell Rep Med* (2022) 3(2):100525. doi: 10.1016/j.xcrm.2022.100525

38. Soundararajan M, Eswaran J. Atypical GTPases as drug targets. *Anticancer Agents Med Chem* (2012) 12(1):19–28. doi: 10.2174/187152012798764705

39. Kucab JE, Zou XQ, Morganella S, Joel M, Nanda AS, Nagy E, et al. A compendium of mutational signatures of environmental agents. *Cell* (2019) 177 (4):821–+. doi: 10.1016/j.cell.2019.03.001

40. van de Geer WS, Hoogstrate Y, Draaisma K, Robe PA, Bins S, Mathijssen RHJ, et al. Landscape of driver gene events, biomarkers, and druggable targets identified by whole-genome sequencing of glioblastomas. *Neuro-oncology Adv* (2022) 4(1):vdab177–vdab177. doi: 10.1093/noajnl/vdab177

41. Griffith M, Miller CA, Griffith OL, Krysiak K, Skidmore ZL, Ramu A, et al. Optimizing cancer genome sequencing and analysis. *Cell Syst* (2015) 1(3):210–23. doi: 10.1016/j.cels.2015.08.015

42. Berger MF, Mardis ER. The emerging clinical relevance of genomics in cancer medicine. *Nat Rev Clin Oncol* (2018) 15(6):353–65. doi: 10.1038/s41571-018-0002-6

43. Gaublomme JT, Li B, McCabe C, Knecht A, Yang Y, Drokhlyansky E, et al. Nuclei multiplexing with barcoded antibodies for single-nucleus genomics. *Nat Commun* (2019) 10(1):2907. doi: 10.1038/s41467-019-10756-2

44. Giesen C, Wang HA, Schapiro D, Zivanovic N, Jacobs A, Hattendorf B, et al. Highly multiplexed imaging of tumor tissues with subcellular resolution by mass cytometry. *Nat Methods* (2014) 11(4):417–22. doi: 10.1038/nmeth.2869

45. Macosko EZ, Basu A, Satija R, Nemesh J, Shekhar K, Goldman M, et al. Highly parallel genome-wide expression profiling of individual cells using nanoliter droplets. *Cell* (2015) 161(5):1202–14. doi: 10.1016/j.cell.2015.05.002

46. Moffitt JR, Hao J, Wang G, Chen KH, Babcock HP, Zhuang X, et al. High-throughput single-cell gene-expression profiling with multiplexed error-robust fluorescence in situ hybridization. *Proc Natl Acad Sci U.S.A.* (2016) 113 (39):11046–51. doi: 10.1073/pnas.1612826113

47. Stoeckius M, Zheng S, Houck-Loomis B, Hao S, Yeung BZ, Mauck WM. Cell hashing with barcoded antibodies enables multiplexing and doublet detection for single cell genomics. *Genome Biol* (2018) 19(1):224. doi: 10.1186/s13059-018-1603-1

48. Huang S, Cai N, Pacheco PP, Narrandes S, Wang Y, Xu W, et al. Applications of support vector machine (SVM) learning in cancer genomics. *Cancer Genomics Proteomics* (2018) 15(1):41–51.

49. Wang X, Yu G, Yan Z, Wan L, Wang W, Lizhen LCC. Lung cancer subtype diagnosis by fusing image-genomics data and hybrid deep networks. *IEEE/ACM Trans Comput Biol Bioinf* (2021) p:1–1. doi: 10.1109/TCBB.2021.3132292

50. Vanderbilt CM, Bowman AS, Middha S, Petrova-Drus K, Tang Y-W, Chen X, et al. Defining novel DNA virus-tumor associations and genomic correlates using prospective clinical Tumor/Normal matched sequencing data. *J Mol Diagnostics* (2022). doi: 10.1016/j.jmoldx.2022.01.011

51. Sudhakar M, Rengaswamy R, Raman K. Novel ratio-metric features enable the identification of new driver genes across cancer types. *Sci Rep* (2022) 12(1):5. doi: 10.1038/s41598-021-04015-y

52. Wu HH, Hua X, Shi JX, Chatterjee N, Zhu B. Utilizing patient information to identify subtype heterogeneity of cancer driver genes. *Stat Methods Med Res* (2022) 31(3):510–9. doi: 10.1177/09622802211055854

53. Buzdin A, Sorokin M, Garazha A, Sekacheva M, Kim E, Zhukov N, et al. Molecular pathway activation - new type of biomarkers for tumor morphology and personalized selection of target drugs. *Semin Cancer Biol* (2018) 53:110–24. doi: 10.1016/j.semcancer.2018.06.003

54. Supplitt S, Karpinski P, Sasiadek M, Laczmanska I. Current achievements and applications of transcriptomics in personalized cancer medicine. *Int J Mol Sci* (2021) 22(3). doi: 10.3390/ijms22031422

55. Buzdin A, Sorokin M, Garazha A, Glusker A, Aleshin A, Poddubskaya E, et al. RNA Sequencing for research and diagnostics in clinical oncology. *Semin Cancer Biol* (2020) 60:311–23. doi: 10.1016/j.semcancer.2019.07.010

56. Consortium SM-I. A comprehensive assessment of RNA-seq accuracy, reproducibility and information content by the sequencing quality control consortium. *Nat Biotechnol* (2014) 32(9):903–14. doi: 10.1038/nbt.2957

57. Ma L, Liang Z, Zhou H, Qu L. Applications of RNA indexes for precision oncology in breast cancer. *Genomics Proteomics Bioinf* (2018) 16(2):108–19. doi: 10.1016/j.gpb.2018.03.002

58. Clarke PA, te Poele R, Workman P. Gene expression microarray technologies in the development of new therapeutic agents. *Eur J Cancer* (2004) 40(17):2560–91. doi: 10.1016/j.ejca.2004.07.024

59. Abdul Aziz NA, Mokhtar NM, Harun R, Mollah MM, Mohamed Rose I, Sagap I, et al. A 19-gene expression signature as a predictor of survival in colorectal cancer. *BMC Med Genomics* (2016) 9(1):58.

60. Montani F, Marzi MJ, Dezi F, Dama E, Carletti RM, Bonizzi G, et al. miR-test: a blood test for lung cancer early detection. *J Natl Cancer Inst* (2015) 107(6): djv063. doi: 10.1093/jnci/djv063

61. Nadal E, Truini A, Nakata A, Lin J, Reddy RM, Chang AC, et al. A novel serum 4-microRNA signature for lung cancer detection. *Sci Rep* (2015) 5:12464. doi: 10.1038/srep12464

62. Hira ZM, Gillies DF. A review of feature selection and feature extraction methods applied on microarray data. *Adv Bioinf* (2015) 2015:198363.

63. Warnat-Herresthal S, Perrakis K, Taschler B, Becker M, Bassler K, Beyer M, et al. Scalable prediction of acute myeloid leukemia using high-dimensional machine learning and blood transcriptomics. *iScience* (2020) 23(1).

64. Ben Azzouz F, Michel B, Lasla H, Gouraud W, Francois AF, Girka F, et al. Development of an absolute assignment predictor for triple-negative breast cancer subtyping using machine learning approaches. *Comput Biol Med* (2021) 129.

65. Alkhateeb A, Rezaeian I, Singireddy S, Cavallo-Medved D, Porter LA, Rueda L. Transcriptomics signature from next-generation sequencing data reveals new transcriptomic biomarkers related to prostate cancer. *Cancer Inf* (2019) 18.

66. Long NP, Park S, Anh NH, Nghi TD, Yoon SJ, Park JH, et al. High-throughput omics and statistical learning integration for the discovery and validation of novel diagnostic signatures in colorectal cancer. *Int J Mol Sci* (2019) 20(2).

67. Torang A, Gupta P, Klinke DJ. An elastic-net logistic regression approach to generate classifiers and gene signatures for types of immune cells and T helper cell subsets. *BMC Bioinf* (2019) 20(1):433. doi: 10.1186/s12859-019-2994-z

68. Masuda T, Mori A, Ito S, Ohtsuki S. Quantitative and targeted proteomics-based identification and validation of drug efficacy biomarkers. *Drug Metab Pharmacokinet* (2021) 36:100361. doi: 10.1016/j.dmpk.2020.09.006

69. Khan MZI, Tam SY, Law HKW. Advances in high throughput proteomics profiling in establishing potential biomarkers for gastrointestinal cancer. *CELLS* (2022) 11(6):973.

70. Li N, Long Y, Fan X, Liu H, Li C, Chen L, et al. Proteomic analysis of differentially expressed proteins in hepatitis b virus-related hepatocellular carcinoma tissues. *J Exp Clin Cancer Res* (2009) 28(1):122. doi: 10.1186/1756-9966-28-122

71. Murata M, Matsuzaki K, Yoshida K, Sekimoto G, Tahashi Y, Mori S, et al. Hepatitis b virus X protein shifts human hepatic transforming growth factor (TGF)-beta signaling from tumor suppression to oncogenesis in early chronic hepatitis b. *Hepatology* (2009) 49(4):1203–17. doi: 10.1002/hep.22765

72. Wu Z, Doondeea JB, Gholami AM, Janning MC, Lemeer S, Kramer K, et al. Quantitative chemical proteomics reveals new potential drug targets in head and neck cancer. *Mol Cell Proteomics* (2011) 10(12):M111.011635. doi: 10.1074/mcp.M111.011635

73. Nambu M, Masuda T, Ito S, Kato K, Kojima T, Daiko H, et al. Leucine-rich alpha-2-Glycoprotein 1 in serum is a possible biomarker to predict response to preoperative chemoradiotherapy for esophageal cancer. *Biol Pharm Bull* (2019) 42 (10):1766–71. doi: 10.1248/bpb.b19-00395

74. Huang X, Zeng Y, Xing X, Zeng J, Gao Y, Cai Z, et al. Quantitative proteomics analysis of early recurrence/metastasis of huge hepatocellular carcinoma following radical resection. *Proteome Sci* (2014) 12:22. doi: 10.1186/1477-5956-12-22

75. Gerdes H, Casado P, Dokal A, Hijazi M, Akhtar N, Osuntola R, et al. Drug ranking using machine learning systematically predicts the efficacy of anti-cancer drugs. *Nat Commun* (2021) 12(1). doi: 10.1038/s41467-021-22170-8

76. Farinella F, Merone M, Bacco L, Capirchio A, Ciccozzi M, Caligiore D, et al. Machine learning analysis of high-grade serous ovarian cancer proteomic dataset reveals novel candidate biomarkers. *Sci Rep* (2022) 12(1):3041. doi: 10.1038/s41598-022-06788-2

77. Ntzioni E, Chouvarda I. Combining machine learning and network analysis pipelines: The case of microbiome and metabolomics data in colorectal cancer. *Stud Health Technol Inf* (2022) 289:489–90. doi: 10.3233/SHTI210965

78. Gupta A, Sagar G, Siddiqui Z, Rao KVS, Nayak S, Saquib N, et al. A non-invasive method for concurrent detection of early-stage women-specific cancers. *Sci Rep* (2022) 12(1):2301. doi: 10.1038/s41598-022-06274-9

79. Prade VM, Sun N, Shen J, Feuchtinger A, Kunzke T, Buck A, et al. The synergism of spatial metabolomics and morphometry improves machine learning-based renal tumour subtype classification. *Clin Trans Med* (2022) 12(2). doi: 10.1002/ctm2.666

80. Shen J, Sun N, Zens P, Kunzke T, Buck A, Prade VM, et al. Spatial metabolomics for evaluating response to neoadjuvant therapy in non-small cell lung cancer patients. *Cancer Commun (London England)* (2022) 42(6):517–35. doi: 10.1002/cac2.12310

81. Liang T-L, Li R-Z, Mai C-T, Guan X-X, Li J-X, Wang X-R, et al. A method establishment and comparison of in vivo lung cancer model development platforms for evaluation of tumour metabolism and pharmaceutical efficacy. *Phytomedicine* (2022) 96:153831. doi: 10.1016/j.phymed.2021.153831

82. Miller HA, Yin X, Smith SA, Hu X, Zhang X, Yan J, et al. Evaluation of disease staging and chemotherapeutic response in non-small cell lung cancer from patient tumor-derived metabolomic data. *Lung Cancer* (2021) 156:20–30. doi: 10.1016/j.lungcan.2021.04.012

83. Evangelista EB, Kwee SA, Sato MM, Wang L, Rettenmeier C, Xie G, et al. Phospholipids are a potentially important source of tissue biomarkers for hepatocellular carcinoma: Results of a pilot study involving targeted metabolomics. *Diagnostics* (2019) 9(4):167. doi: 10.3390/diagnostics9040167

84. Xie Y, Meng W-Y, Li R-Z, Wang Y-W, Qian X, Chan C, et al. Early lung cancer diagnostic biomarker discovery by machine learning methods. *Trans Oncol* (2021) 14(1):100907. doi: 10.1016/j.tranon.2020.100907

85. Dykstra MA, Switzer N, Eisner R, Tso V, Foshaug R, Ismond K, et al. Urine metabolomics as a predictor of patient tolerance and response to adjuvant chemotherapy in colorectal cancer. *Mol Clin Oncol* (2017) 7(5):767–70. doi: 10.3892/mco.2017.1407

86. Bifarin OO, Gaul DA, Sah S, Arnold RS, Ogan K, Master VA, et al. Machine learning-enabled renal cell carcinoma status prediction using multiplatform urine-based metabolomics. *J Proteome Res* (2021) 20(7):3629–41. doi: 10.1021/acs.jproteome.1c00213

87. Biswas N, Chakrabarti S. Artificial intelligence (AI)-based systems biology approaches in multi-omics data analysis of cancer. *Front Oncol* (2020) 10. doi: 10.3389/fonc.2020.588221

88. Tong L, Wu H, Wang MD. Integrating multi-omics data by learning modality invariant representations for improved prediction of overall survival of cancer. *Methods* (2021) 189:74–85. doi: 10.1016/j.ymeth.2020.07.008

89. Nicora G, Vitali F, Dagliati A, Geifman N, Bellazzi R. Integrated multi-omics analyses in oncology: A review of machine learning methods and tools. *Front Oncol* (2020) 10. doi: 10.3389/fonc.2020.01030

90. Takahashi S, Asada K, Takasawa K, Shimoyama R, Sakai A, Bolatkan A, et al. Predicting deep learning based multi-omics parallel integration survival subtypes in lung cancer using reverse phase protein array data. *Biomolecules* (2020) 10(10):1460. doi: 10.3390/biom10101460

91. Shuangshuang L, et al. (2020). A deep learning fusion clustering framework for breast cancer subtypes identification by integrating multi-omics data, in: *2020 5TH International Conference on Mechanical, Control and Computer Engineering (ICMCCE 2020)*, . pp. 1710–4.

92. Albaradei S, Napolitano F, Thafar MA, Gojobori T, Essack M, Gao X. MetaCancer: A deep learning-based pan-cancer metastasis prediction model developed using multi-omics data. *Comput Struct Biotechnol J* (2021) 19:4404–11. doi: 10.1016/j.csbj.2021.08.006

93. Lewis JE, Kemp ML. Integration of machine learning and genome-scale metabolic modeling identifies multi-omics biomarkers for radiation resistance. *Nat Commun* (2021) 12(1):2700. doi: 10.1038/s41467-021-22989-1

94. Wang Y, Yang YX, Chen SL, Wang JG. DeepDRK: a deep learning framework for drug repurposing through kernel-based multi-omics integration. *Briefings Bioinf* (2021) 22(5):bbab048. doi: 10.1093/bib/bbab048

95. Poirion OB, Jing Z, Chaudhary K, Huang SJ, Garmire LX. DeepProg: an ensemble of deep-learning and machine-learning models for prognosis prediction using multi-omics data. *Genome Med* (2021) 13(1). doi: 10.1186/s13073-021-00930-x

96. Karim MR, Islam T, Lange C, Rebholz-Schuhmann D, Decker S. Adversary-aware multimodal neural networks for cancer susceptibility prediction from multiomics data. *IEEE Access* (2022) 10:54386–409. doi: 10.1109/ACCESS.2022.3175816

97. Zhang L, Lv CK, Jin YQ, Cheng GQ, Fu YB, Yuan DS, et al. Deep learning-based multi-omics data integration reveals two prognostic subtypes in high-risk neuroblastoma. *Front Genet* (2018) 9. doi: 10.3389/fgene.2018.00477

98. Xu JM, Xu BH, Li YP, Su ZJ, Yao YP. Unsupervised learning of cross-modal mappings in multi-omics data for survival stratification of gastric cancer. *Future Oncol* (2021) 18(2):215–30.

99. Guo LY, Wu AH, Wang YX, Zhang LP, Chai H, Liang XF. Deep learning-based ovarian cancer subtypes identification using multi-omics data. *Biodata Min* (2020) 13(1):10. doi: 10.1186/s13040-020-00222-x

100. Lee T-Y, Huang K-Y, Chuang C-H, Lee C-Y, Chang T-H. Incorporating deep learning and multi-omics autoencoding for analysis of lung adenocarcinoma prognostication. *Comput Biol Chem* (2020) 87:107277. doi: 10.1016/j.compbiolchem.2020.107277

101. Zhu R, Dai LY, Liu JX, Guo Y. Diagnostic classification of lung cancer using deep transfer learning technology and multi-omics data. *Chin J Electron* (2021) 30(5):843–52.

102. Khan D, Shedole S. Leveraging deep learning techniques and integrated omics data for tailored treatment of breast cancer. *J Personalized Med* (2022) 12(5). doi: 10.3390/jpm12050674

103. Pan X, Burgman B, Wu E, Huang JH, Sahni N, Stephen Yi S, et al. I-modern: Integrated multi-omics network model identifies potential therapeutic targets in glioma by deep learning with interpretability. *Comput Struct Biotechnol J* (2022) 20:3511–21. doi: 10.1016/j.csbj.2022.06.058

104. Ma B, Meng F, Yan G, Yan H, Chai B, Song F. Diagnostic classification of cancers using extreme gradient boosting algorithm and multi-omics data. *Comput Biol Med* (2020) 121:103761. doi: 10.1016/j.compbiomed.2020.103761

105. Yang H, Chen R, Li D, Wang Z. Subtype-GAN: a deep learning approach for integrative cancer subtyping of multi-omics data. *Bioinformatics* (2021) 37 (16):2231–7. doi: 10.1093/bioinformatics/btab109

106. Huang GJ, Wang C, Fu X. Bidirectional deep neural networks to integrate RNA and DNA data for predicting outcome for patients with hepatocellular carcinoma. *Future Oncol* (2021) 17(33):4481–95. doi: 10.2217/fon-2021-0659

107. ElKarami B, Alkhateeb A, Qattous H, Alshomali L, Shahrrava B. Multi-omics data integration model based on UMAP embedding and convolutional neural network. *Cancer Inf* (2022) 21:11769351221124205. doi: 10.1177/11769351221124205

108. Wang TX, Shao W, Huang Z, Tang HX, Zhang J, Ding ZM, et al. MOGONET integrates multi-omics data using graph convolutional networks allowing patient classification and biomarker identification. *Nat Commun* (2021) 12(1):3445. doi: 10.1038/s41467-021-23774-w

109. Chai H, Zhou X, Zhang Z, Rao J, Zhao H, Yang Y, et al. Integrating multi-omics data through deep learning for accurate cancer prognosis prediction. *Comput Biol Med* (2021) 134:104481. doi: 10.1016/j.compbiomed.2021.104481

110. Xu J, Wu P, Chen YH, Meng QF, Dawood H, Dawood H. A hierarchical integration deep flexible neural forest framework for cancer subtype classification by integrating multi-omics data. *BMC Bioinf* (2019) 20(1):527. doi: 10.1186/s12859-019-3116-7

111. Park C, Oh I, Choi J, Ko S, Ahn J. Improved prediction of cancer outcome using graph-embedded generative adversarial networks. *IEEE Access* (2021) 9:20076–88. doi: 10.1109/ACCESS.2021.3054894

112. Li F, Sun ZS, Liu JX, Shang JL, Dai LY, Liu XK. NESM: a network embedding method for tumor stratification by integrating multi-omics data. *G3-Genes Genomes Genet*.

113. Rong Z, Liu Z, Song J, Cao L, Yu Y, Qiu M, et al. MCluster-VAEs: An end-to-end variational deep learning-based clustering method for subtype discovery using multi-omics data. *Comput Biol Med* (2022) 150:106085. doi: 10.1016/j.compbiomed.2022.106085

114. Ma TL, Zhang AD. Integrate multi-omics data with biological interaction networks using multi-view factorization AutoEncoder (MAE). *BMC Genomics* (2019) 20:944. doi: 10.1186/s12864-019-6285-x

115. Wang Z, Zhao Y, Shen X, Zhao Y, Zhang Z, Yin H, et al. Single-cell genomics-based molecular algorithm for early cancer detection. *Analytical Chem* (2022) 94(5):2607–14. doi: 10.1021/acs.analchem.1c04968

116. Trivizakis E, Souglakos J, Karantanas A, Marias K. Deep radiotranscriptomics of non-small cell lung carcinoma for assessing molecular and histology subtypes with a data-driven analysis. *Diagnostics* (2021) 11(12):2383. doi: 10.3390/diagnostics11122383

117. Zhang X, Xing Y, Sun K, Guo Y. OmiEmbed: A unified multi-task deep learning framework for multi-omics data. *Cancers* (2021) 13(12):3047. doi: 10.3390/cancers13123047

118. Wu Y, Cheng M, Huang S, Pei Z, Zuo Y, Liu J, et al. Recent advances of deep learning for computational histopathology: Principles and applications. *Cancers* (2022) 14(5):1199. doi: 10.3390/cancers14051199

119. Robertson S, Azizpour H, Smith K, Hartman J. Digital image analysis in breast pathology–from image processing techniques to artificial intelligence. *Trans Res* (2018) 194:19–35. doi: 10.1016/j.trsl.2017.10.010

120. Ho C, Zhao Z, Chen XF, Sauer J, Saraf SA, Jialdasani R, et al. A promising deep learning-assistive algorithm for histopathological screening of colorectal cancer. *Sci Rep* (2022) 12(1):2222. doi: 10.1038/s41598-022-06264-x

121. Srinidhi CL, Ciga O, Martel AL. Deep neural network models for computational histopathology: A survey. *Med Image Anal* (2021) 67:101813. doi: 10.1016/j.media.2020.101813

122. Iizuka O, Kanavati F, Kato K, Rambeau M, Arihiro K, Tsuneki M. Deep learning models for histopathological classification of gastric and colonic epithelial tumours. *Sci Rep* (2020) 10(1):1504. doi: 10.1038/s41598-020-58467-9

123. Kainz P, Pfeiffer M, Urschler M. Segmentation and classification of colon glands with deep convolutional neural networks and total variation regularization. *PeerJ* (2017) 5:e3874. doi: 10.7717/peerj.3874

124. Lui TKL, Guo CG, Leung WK. Accuracy of artificial intelligence on histology prediction and detection of colorectal polyps: A systematic review and meta-analysis. *Gastrointest Endosc* (2020) 92(1):11–22.e6. doi: 10.1016/j.gie.2020.02.033

125. Ehteshami Bejnordi B, Veta M, Johannes van Diest P, van Ginneken B, Karssemeijer N, Litjens G, et al. Diagnostic assessment of deep learning algorithms for detection of lymph node metastases in women with breast cancer. *Jama* (2017) 318(22):2199–210. doi: 10.1001/jama.2017.14585

126. Dembrower K, Wåhlin E, Liu Y, Salim M, Smith K, Lindholm P, et al. Effect of artificial intelligence-based triaging of breast cancer screening mammograms on cancer detection and radiologist workload: a retrospective simulation study. *Lancet Digit Health* (2020) 2(9):e468–74. doi: 10.1016/S2589-7500(20)30185-0

127. Stenzinger A, Alber M, Allgäuer M, Jurmeister P, Bockmayr M, Budczies J, et al. Artificial intelligence and pathology: From principles to practice and future applications in histomorphology and molecular profiling. *Semin Cancer Biol* (2022) 84:129–43. doi: 10.1016/j.semcancer.2021.02.011

128. Su F, Li J, Zhao X, Wang B, Hu Y, Sun Y, et al. Interpretable tumor differentiation grade and microsatellite instability recognition in gastric cancer using deep learning. *Lab Invest* (2022) 102(6):641–9. doi: 10.1038/s41374-022-00742-6

129. Sengupta D, Ali SN, Bhattacharya A, Mustafi J, Mukhopadhyay A, Sengupta K. A deep hybrid learning pipeline for accurate diagnosis of ovarian cancer based on nuclear morphology. *PloS One* (2022) 17(1):e0261181. doi: 10.1371/journal.pone.0261181

130. Bulten W, Balkenhol M, Belinga JA, Brilhante A, Çakır A, Egevad L, et al. Artificial intelligence assistance significantly improves Gleason grading of prostate biopsies by pathologists. *Mod Pathol* (2021) 34(3):660–71. doi: 10.1038/s41379-020-0640-y

131. Coudray N, Ocampo PS, Sakellaropoulos T, Narula N, Snuderl M, Fenyö D, et al. Classification and mutation prediction from non-small cell lung cancer histopathology images using deep learning. *Nat Med* (2018) 24(10):1559–67. doi: 10.1038/s41591-018-0177-5

132. Ye ZX, Zhang YX, Liang YB, Lang JD, Zhang XL, Zang GL, et al. Cervical cancer metastasis and recurrence risk prediction based on deep convolutional neural network. *Curr Bioinf* (2022) 17(2):164–73. doi: 10.2174/1574893616666210708143556

133. Kulkarni PM, Robinson EJ, Sarin Pradhan J, Gartrell-Corrado RD, Rohr BR, Trager MH, et al. Deep learning based on standard H&E images of primary melanoma tumors identifies patients at risk for visceral recurrence and death. *Clin Cancer Res* (2020) 26(5):1126–34. doi: 10.1158/1078-0432.CCR-19-1495

134. Meier A, Nekolla K, Hewitt LC, Earle S, Yoshikawa T, Oshima T, et al. Hypothesis-free deep survival learning applied to the tumour microenvironment in gastric cancer. *J Pathol Clin Res* (2020) 6(4):273–82. doi: 10.1002/cjp2.170

135. Peikari M, Gangeh MJ, Zubovits J, Clarke G, Martel AL. Triaging diagnostically relevant regions from pathology whole slides of breast cancer: A texture based approach. *IEEE Trans Med Imaging* (2016) 35(1):307–15. doi: 10.1109/TMI.2015.2470529

136. Mao B, Zhang L, Ning P, Ding F, Wu F, Lu G, et al. Preoperative prediction for pathological grade of hepatocellular carcinoma via machine learning-based radiomics. *Eur Radiol* (2020) 30(12):6924–32. doi: 10.1007/s00330-020-07056-5

137. Komura D, Ishikawa S. Machine learning approaches for pathologic diagnosis. *Virchows Arch* (2019) 475(2):131–8. doi: 10.1007/s00428-019-02594-w

138. Fakoor R, Ladhak F, Nazi A, Huber M. Using deep learning to enhance cancer diagnosis and classification. In: *Proc Int Conf Mach* (2013) 28:3937–49.

139. Mobadersany P, Yousefi S, Amgad M, Gutman DA, Barnholtz-Sloan JS, Velázquez Vega JE, et al. Predicting cancer outcomes from histology and genomics using convolutional networks. *Proc Natl Acad Sci United States America* (2018) 115 (13):E2970–9. doi: 10.1073/pnas.1717139115

140. Al-Antari MA, Han SM, Kim TS. Evaluation of deep learning detection and classification towards computer-aided diagnosis of breast lesions in digital X-ray mammograms. *Comput Methods Programs BioMed* (2020) 196:105584. doi: 10.1016/j.cmpb.2020.105584

141. Hussein S, Kandel P, Bolan CW, Wallace MB, Bagci U. Lung and pancreatic tumor characterization in the deep learning era: Novel supervised and unsupervised learning approaches. *IEEE Trans Med Imaging* (2019) 38(8):1777–87. doi: 10.1109/TMI.2019.2894349

142. Bertelli E, Mercatelli L, Marzi C, Pachetti E, Baccini M, Barucci A, et al. Machine and deep learning prediction of prostate cancer aggressiveness using multiparametric MRI. *Front Oncol* (2021) 11:802964. doi: 10.3389/fonc.2021.802964

143. Ayalew YA, Fante KA, Mohammed MA. Modified U-net for liver cancer segmentation from computed tomography images with a new class balancing method. *BMC Biomed Eng* (2021) 3(1):4–4. doi: 10.1186/s42490-021-00050-y

144. Bi WL, Hosny A, Schabath MB, Giger ML, Birkbak NJ, Mehrtash A, et al. Artificial intelligence in cancer imaging: Clinical challenges and applications. *CA Cancer J Clin* (2019) 69(2):127–57. doi: 10.3322/caac.21552

145. Dhahri H, Al Maghayreh E, Mahmood A, Elkilani W, Faisal Nagi M. Automated breast cancer diagnosis based on machine learning algorithms. *J Healthcare Eng* (2019) 2019:4253641. doi: 10.1155/2019/4253641

146. Liu RY, Pan DR, Xu Y, Zeng H, He ZL, Lin JB, et al. A deep learning-machine learning fusion approach for the classification of benign, malignant, and intermediate bone tumors. *Eur Radiol* (2022) 32(2):1371–83. doi: 10.1007/s00330-021-08195-z

147. Alanazi MF, Ali MU, Hussain SJ, Zafar A, Mohatram M, Irfan M, et al. Brain Tumor/Mass classification framework using magnetic-Resonance-Imaging-Based isolated and developed transfer deep-learning model. *Sensors* (2022) 22 (1):372. doi: 10.3390/s22010372

148. Zahoor MM, Qureshi SA, Bibi S, Khan SH, Khan A, Ghafoor U, et al. A new deep hybrid boosted and ensemble learning-based brain tumor analysis using MRI. *Sensors* (2022) 22(7):2726. doi: 10.3390/s22072726

149. Chougrad H, Zouaki H, Alheyane O. Deep convolutional neural networks for breast cancer screening. *Comput Methods Programs BioMed* (2018) 157:19–30. doi: 10.1016/j.cmpb.2018.01.011

150. Misra S, Jeon S, Managuli R, Lee S, Kim G, Yoon C, et al. Bi-modal transfer learning for classifying breast cancers via combined b-mode and ultrasound strain imaging. *IEEE Trans Ultrasonics Ferroelectrics Frequency Control* (2022) 69 (1):222–32. doi: 10.1109/TUFFC.2021.3119251

151. Brunese L, Mercaldo F, Reginelli A, Santone A. Radiomics for Gleason score detection through deep learning. *Sensors (Basel Switzerland)* (2020) 20 (18):5411. doi: 10.3390/s20185411

152. Zhao Z, Bian Y, Jiang H, Fang X, Li J, Cao K, et al. CT-radiomic approach to predict G1/2 nonfunctional pancreatic neuroendocrine tumor. *Acad Radiol* (2020) 27(12):e272–81. doi: 10.1016/j.acra.2020.01.002

153. Gurgitano M, Angileri SA, Roda GM, Liguori A, Pandolfi M, Lerardi AM, et al. Interventional radiology ex-machina: impact of artificial intelligence on practice. *Radiologia Med* (2021) 126(7):998–1006. doi: 10.1007/s11547-021-01351-x

154. Xie T, Wang X, Zhang Z, Zhou Z. CT-based radiomics analysis for preoperative diagnosis of pancreatic mucinous cystic neoplasm and atypical

serous cystadenomas. *Front Oncol* (2021) 11:621520. doi: 10.3389/fonc.2021.621520

155. Song D, Yang F, Zhang Y, Guo Y, Qu Y, Zhang X, et al. Dynamic contrast-enhanced MRI radiomics nomogram for predicting axillary lymph node metastasis in breast cancer. *Cancer Imaging* (2022) 22(1):17. doi: 10.1186/s40644-022-00450-w

156. Eresen A, Li Y, Yang J, Shangguan J, Velichko Y, Yaghmai V, et al. Preoperative assessment of lymph node metastasis in colon cancer patients using machine learning: a pilot study. *Cancer Imaging* (2020) 20(1):30. doi: 10.1186/s40644-020-00308-z

157. Guo R, Guo J, Zhang L, Qu X, Dai S, Peng R, et al. CT-based radiomics features in the prediction of thyroid cartilage invasion from laryngeal and hypopharyngeal squamous cell carcinoma. *Cancer Imaging* (2020) 20(1):81. doi: 10.1186/s40644-020-00359-2

158. Akbari H, Macyszyn L, Da X, Bilello M, Wolf RL, Martinez-Lage M, et al. Imaging surrogates of infiltration obtained via multiparametric imaging pattern analysis predict subsequent location of recurrence of glioblastoma. *Neurosurgery* (2016) 78(4):572–80. doi: 10.1227/NEU.0000000000001202

159. Fung NTC, Hung WM, Sze CK, Lee MCH, Ng WT. Automatic segmentation for adaptive planning in nasopharyngeal carcinoma IMRT: Time, geometrical, and dosimetric analysis. *Med Dosim* (2020) 45(1):60–5. doi: 10.1016/j.meddos.2019.06.002

160. Lustberg T, van Soest J, Gooding M, Peressutti D, Aljabar P, van der Stoep J, et al. Clinical evaluation of atlas and deep learning based automatic contouring for lung cancer. *Radiother Oncol* (2018) 126(2):312–7. doi: 10.1016/j.radonc.2017.11.012

161. Elguindi S, Zelefsky MJ, Jiang J, Veeraraghavan H, Deasy JO, Hunt MA, et al. Deep learning-based auto-segmentation of targets and organs-at-risk for magnetic resonance imaging only planning of prostate radiotherapy. *Phys Imaging Radiat Oncol* (2019) 12:80–6. doi: 10.1016/j.phro.2019.11.006

162. Schouten JPE, Noteboom S, Martens RM, Mes SW, Leemans CR, de Graaf P, et al. Automatic segmentation of head and neck primary tumors on MRI using a multi-view CNN. *Cancer Imaging* (2022) 22(1):8. doi: 10.1186/s40644-022-00445-7

163. Lin L, Dou Q, Jin YM, Zhou GQ, Tang YQ, Chen WL, et al. Deep learning for automated contouring of primary tumor volumes by MRI for nasopharyngeal carcinoma. *Radiology* (2019) 291(3):677–86. doi: 10.1148/radiol.2019182012

164. Zhang J, Cheng Z, Fan Z, Zhang Q, Zhang X, Yang R, et al. A feasibility study for in vivo treatment verification of IMRT using Monte Carlo dose calculation and deep learning-based modelling of EPID detector response. *Radiat Oncol* (2022) 17(1):31. doi: 10.1186/s13014-022-01999-3

165. Auloge P, Cazzato RL, Ramamurthy N, de Marini P, Rousseau C, Garnon J, et al. Augmented reality and artificial intelligence-based navigation during percutaneous vertebroplasty: a pilot randomised clinical trial. *Eur Spine J* (2020) 29(7):1580–9. doi: 10.1007/s00586-019-06054-6

166. Tomaszewski MR, Latifi K, Boyer E, Palm RF, El Naqa I, Moros EG, et al. Delta radiomics analysis of magnetic resonance guided radiotherapy imaging data can enable treatment response prediction in pancreatic cancer. *Radiat Oncol* (2021) 16(1):237. doi: 10.1186/s13014-021-01957-5

167. Bousabarah K, Blanck O, Temming S, Wilhelm M-L, Hoevels M, Baus WW, et al. Radiomics for prediction of radiation-induced lung injury and oncologic outcome after robotic stereotactic body radiotherapy of lung cancer: results from two independent institutions. *Radiat Oncol* (2021) 16(1):74. doi: 10.1186/s13014-021-01805-6

168. Zhang DH, Duan YC, Guo J, Wang YW, Yang Y, Li ZH, et al. Using multi-scale convolutional neural network based on multi-instance learning to predict the efficacy of neoadjuvant chemoradiotherapy for rectal cancer. *IEEE J Trans Eng Health Med* (2022) 10. doi: 10.1109/JTEHM.2022.3156851

169. Ermiş E, Jungo A, Poel R, Blatti-Moreno M, Meier R, Knecht U, et al. Fully automated brain resection cavity delineation for radiation target volume definition in glioblastoma patients using deep learning. *Radiat Oncol* (2020) 15(1):100. doi: 10.1186/s13014-020-01553-z

170. Zhou J, Peng Z, Song Y, Chang Y, Pei X, Sheng L, et al. A method of using deep learning to predict three-dimensional dose distributions for intensity-modulated radiotherapy of rectal cancer. *J Appl Clin Med Phys* (2020) 21(5):26–37. doi: 10.1002/acm2.12849

171. Peng L, Parekh V, Huang P, Lin DD, Sheikh K, Baker B, et al. Distinguishing true progression from radionecrosis after stereotactic radiation therapy for brain metastases with machine learning and radiomics. *Int J Radiat Oncol Biol Phys* (2018) 102(4):1236–43. doi: 10.1016/j.ijrobp.2018.05.041

172. Zhang Z, Yang J, Ho A, Jiang W, Logan J, Wang X, et al. A predictive model for distinguishing radiation necrosis from tumour progression after gamma knife radiosurgery based on radiomic features from MR images. *Eur Radiol* (2018) 28(6):2255–63. doi: 10.1007/s00330-017-5154-8

173. Shaver MM, et al. Optimizing neuro-oncology imaging: A review of deep learning approaches for glioma imaging. *Cancers* (2019) 11(6):829. doi: 10.3390/cancers11060829

174. Wang K, Qiao Z, Zhao X, Li X, Wang X, Wu T, et al. Individualized discrimination of tumor recurrence from radiation necrosis in glioma patients using an integrated radiomics-based model. *Eur J Nucl Med Mol Imaging* (2020) 47(6):1400–11. doi: 10.1007/s00259-019-04604-0

175. Cha KH, Hadjiiski L, Chan HP, Weizer AZ, Alva A, Cohan RH, et al. Bladder cancer treatment response assessment in CT using radiomics with deep-learning. *Sci Rep* (2017) 7(1):8738. doi: 10.1038/s41598-017-09315-w

176. Gong J, Bao X, Wang T, Liu JY, Peng WJ, Shi JY, et al. A short-term follow-up CT based radiomics approach to predict response to immunotherapy in advanced non-small-cell lung cancer. *Oncoimmunology* (2022) 11(1):2028962. doi: 10.1080/2162402X.2022.2028962

177. Bhardwaj D, Dasgupta A, DiCenzo D, Brade S, Fatima K, Quiaoit K, et al. Early changes in quantitative ultrasound imaging parameters during neoadjuvant chemotherapy to predict recurrence in patients with locally advanced breast cancer. *Cancers* (2022) 14(5):1247. doi: 10.3390/cancers14051247

178. Liu G, Poon M, Zapala MA, Temple WC, Vo KT, Matthay KK, et al. Incorporating radiomics into machine learning models to predict outcomes of neuroblastoma. *J Digital Imaging* (2022) 35(3):605–12. doi: 10.1007/s10278-022-00607-w

179. Cheng SH, Cheng YJ, Jin ZY, Xue HD. Unresectable pancreatic ductal adenocarcinoma: Role of CT quantitative imaging biomarkers for predicting outcomes of patients treated with chemotherapy. *Eur J Radiol* (2019) 113:188–97. doi: 10.1016/j.ejrad.2019.02.009

180. Parr E, Du Q, Zhang C, Lin C, Kamal A, McAlister J, et al. Radiomics-based outcome prediction for pancreatic cancer following stereotactic body radiotherapy. *Cancers (Basel)* (2020) 12(4):1051. doi: 10.3390/cancers12041051

181. Schilling KG, Landman BA. AI In MRI: A case for grassroots deep learning. *Magn Reson Imaging* (2019) 64:1–3. doi: 10.1016/j.mri.2019.07.004

182. Li BJ, Wang TY, Nabavi S, Assoc Comp M. (2021). Cancer molecular subtype classification by graph convolutional networks on multi-omics data, in: *12TH ACM Conference on Bioinformatics, Computational Biology, and Health Informatics (ACM-BCB 2021)*, .

183. Wu WT, Li YJ, Feng AZ, Li L, Huang T, Xu AD, et al. Data mining in clinical big data: the frequently used databases, steps, and methodological models. *Mil Med Res* (2021) 8(1):44. doi: 10.1186/s40779-021-00338-z

184. Yu KH, Berry GJ, Rubin DL, Ré C, Altman RB, Snyder M. Association of omics features with histopathology patterns in lung adenocarcinoma. *Cell Syst* (2017) 5(6):620–627.e3. doi: 10.1016/j.cels.2017.10.014

185. James TA, Fan B. ASO author reflections: Using tumor genomics to predict axillary response to chemotherapy in breast cancer. *Ann Surg Oncol* (2021) 28(3):1326–7. doi: 10.1245/s10434-020-09455-w

186. Hatzikirou H. Combining dynamic modeling with machine learning can be the key for the integration of mathematical and clinical oncology: Comment on "Improving cancer treatments via dynamical biophysical models" by m. kuznetsov, j. clairambault, v. volpert. *Phys Life Rev* (2022) 40:1–2. doi: 10.1016/j.plrev.2022.01.002

187. Kuznetsov M, Clairambault J, Volpert V. Improving cancer treatments via dynamical biophysical models. *Phys Life Rev* (2021) 39:1–48. doi: 10.1016/j.plrev.2021.10.001

188. Chen RJ, Lu MY, Wang J, Williamson DFK, Rodig SJ, Lindeman NI, et al. Pathomic fusion: An integrated framework for fusing histopathology and genomic features for cancer diagnosis and prognosis. *IEEE Trans Med Imaging* (2022) 41(4):757–70. doi: 10.1109/TMI.2020.3021387

189. Sun D, Wang M, Li A. A multimodal deep neural network for human breast cancer prognosis prediction by integrating multi-dimensional data. *IEEE/ACM Trans Comput Biol Bioinform* (2019) 16(3):841–50. doi: 10.1109/TCBB.2018.2806438

190. Huang Z, Zhan X, Xiang S, Johnson TS, Helm B, Yu CY, et al. SALMON: Survival analysis learning with multi-omics neural networks on breast cancer. *Front Genet* (2019) 10:166. doi: 10.3389/fgene.2019.00166

191. Khan AN, Ihalage AA, Ma Y, Liu B, Liu Y, Hao Y. Deep learning framework for subject-independent emotion detection using wireless signals. *PloS One* (2021) 16(2):e0242946. doi: 10.1371/journal.pone.0242946

192. Davatzikos C. Machine learning in neuroimaging: Progress and challenges. *Neuroimage* (2019) 197:652–6. doi: 10.1016/j.neuroimage.2018.10.003

193. Hu Y, Zhao L, Li Z, Dong X, Xu T, Zhao Y. Classifying the multi-omics data of gastric cancer using a deep feature selection method. *Expert Syst Appl* (2022) 200:116813. doi: 10.1016/j.eswa.2022.116813

194. Yoon HJ, Stanley C, Christian JB, Klasky HB, Blanchard AE, Durbin EB, et al. Optimal vocabulary selection approaches for privacy-preserving deep NLP model training for information extraction and cancer epidemiology. *Cancer Biomarkers* (2022) 33(2):185–98. doi: 10.3233/CBM-210306

195. Leng DJ, Zheng LY, Wen YQ, Zhang YH, Wu LL, Wang J, et al. A benchmark study of deep learning-based multi-omics data fusion methods for cancer. *Genome Biol* (2022) 23(1):171. doi: 10.1186/s13059-022-02739-2

196. Huang LM, Yang WJ, Huang ZY, Tang CW, Li J. Artificial intelligence technique in detection of early esophageal cancer. *World J Gastroenterol* (2020) 26(39):5959–69. doi: 10.3748/wjg.v26.i39.5959

frontiers | Frontiers in Oncology

TYPE Original Research
PUBLISHED 01 March 2023
DOI 10.3389/fonc.2023.979437

OPEN ACCESS

EDITED BY
Sweet Ping Ng,
University of Melbourne, Australia

REVIEWED BY
Alessandro Stefano,
National Research Council (CNR), Italy
Bilgin Kadri Aribas,
Bülent Ecevit University, Türkiye

*CORRESPONDENCE
Daowei Li
daoweili024@163.com

SPECIALTY SECTION
This article was submitted to Cancer Imaging and Image-directed Interventions, a section of the journal Frontiers in Oncology

RECEIVED 27 June 2022
ACCEPTED 20 February 2023
PUBLISHED 01 March 2023

CITATION
Lu J, Jiang N, Zhang Y and Li D (2023) A CT based radiomics nomogram for differentiation between focal-type autoimmune pancreatitis and pancreatic ductal adenocarcinoma.
Front. Oncol. 13:979437.
doi: 10.3389/fonc.2023.979437

A CT based radiomics nomogram for differentiation between focal-type autoimmune pancreatitis and pancreatic ductal adenocarcinoma

Jia Lu[1], Nannan Jiang[2], Yuqing Zhang[1] and Daowei Li[1]*

[1]Department of Radiology, The People's Hospital of China Medical University and The People's Hospital of Liaoning Province, Shenyang, China, [2]Department of Radiology, The People's Hospital of Liaoning Province, Shenyang, China

Objectives: The purpose of this study was to develop and validate an CT-based radiomics nomogram for the preoperative differentiation of focal-type autoimmune pancreatitis from pancreatic ductal adenocarcinoma.

Methods: 96 patients with focal-type autoimmune pancreatitis and pancreatic ductal adenocarcinoma have been enrolled in the study (32 and 64 cases respectively). All cases have been confirmed by imaging, clinical follow-up and/or pathology. The imaging data were considered as: 70% training cohort and 30% test cohort. Pancreatic lesions have been manually delineated by two radiologists and image segmentation was performed to extract radiomic features from the CT images. Independent-sample T tests and LASSO regression were used for feature selection. The training cohort was classified using a variety of machine learning-based classifiers, and 5-fold cross-validation has been performed. The classification performance was evaluated using the test cohort. Multivariate logistic regression analysis was then used to develop a radiomics nomogram model, containing the CT findings and Rad-Score. Calibration curves have been plotted showing the agreement between the predicted and actual probabilities of the radiomics nomogram model. Different patients have been selected to test and evaluate the model prediction process. Finally, receiver operating characteristic curves and decision curves were plotted, and the radiomics nomogram model was compared with a single model to visually assess its diagnostic ability.

Results: A total of 158 radiomics features were extracted from each image. 7 features were selected to construct the radiomics model, then a variety of classifiers were used for classification and multinomial logistic regression (MLR) was selected to be the optimal classifier. Combining CT findings with radiomics model, a prediction model based on CT findings and radiomics was finally obtained. The nomogram model showed a good sensitivity and specificity with AUCs of 0.87 and 0.83 in training and test cohorts, respectively. The areas under the curve and decision curve analysis showed that the radiomics nomogram model may provide better diagnostic performance than the single model and

achieve greater clinical net benefits than the CT finding model and radiomics signature model individually.

Conclusions: The CT image-based radiomics nomogram model can accurately distinguish between focal-type autoimmune pancreatitis and pancreatic ductal adenocarcinoma patients and provide additional clinical benefits.

KEYWORDS

radiomics, focal-type autoimmune pancreatitis, pancreatic ductal adenocarcinoma, differential, machine learning

1 Introduction

The concept of autoimmune pancreatitis (AIP) was first proposed by Yoshida et al. in 1995 (1). As a rare chronic disease, AIP usually presents as recurrent acute pancreatitis with abundant pathological lymphoplasmacytic infiltration (2, 3). The current study classifies AIP into two types: diffuse pancreatitis and focal pancreatitis (4). Focal-type autoimmune pancreatitis (fAIP) presents with segmental involvement of the pancreatic parenchyma, accounting for approximately 28-41% of AIP cases (5, 6). The imaging and clinical features of fAIP and pancreatic ductal adenocarcinoma (PDAC) are very similar, including focal or mass-like enlargement of the pancreas and obstructive jaundice, making their differential diagnosis very difficult. In addition, the treatment and prognosis of the two diseases vary widely. AIP is a benign fibro-inflammatory disease that responds to steroid therapy within one month in 90% of cases (7), whereas PDAC requires surgical resection to cure. Studies have shown that nearly 16% of cases of AIP are misdiagnosed as PDAC and undergo unnecessary pancreatectomy, with approximately 5-21% of cases undergoing pancreatectomy being ultimately confirmed as AIP. Currently, the only reference standard for the differential diagnosis of fAIP from PDAC is post-operative histology. The imaging examination is lacking clear reference standards for definitive diagnosis (7). Therefore, it is crucial to develop a non-invasive and effective methods to distinguish fAIP from PDAC preoperatively, enabling clinicians in the selection of appropriate treatment strategies.

As an emerging technology in the field of medical imaging, radiomics has provided a large amount of quantitative high-throughput information on radiographic images, helping to describe the tumor heterogeneity and the corresponding microenvironment (8). In this way, more predictive information can be obtained from medical imaging data than just the traditional visual interpretation (9), and provides a new way of approaching clinical diagnosis. In the field of abdominal radiology, radiomics techniques have been extensively studied, aiming to predict the tumor grade, survival and response to treatment, and to distinguish benign from malignant lesions. Therefore, it has the potential to be a non-invasive diagnostic method with performance close to biopsy. Some studies have applied this technique to pancreatic diseases (10–17), with a few studies reporting that the radiomic features extracted from enhanced CT images have certain value in the identification of AIP and PDAC. However, a more accurate integrative analysis of radiomics nomogram models to discriminate between fAIP and PDAC has not been fully developed.

Therefore, this study aims to develop and validate a non-invasive, reproducible and personalized radiomics-based nomogram method for preoperative identification of fAIP and PDAC based on contrast-enhanced CT images.

2 Materials and methods

2.1 Patients

The patients with fAIP between January 2011 and January 2021 in our hospital have been considered for this study. These patients were included according to the 2011 International Consensus Diagnostic Criteria (ICDC). The exclusion criteria were as follows (1): Contrast CT was not performed prior to steroid therapy or surgery; (2) The mass involving the pancreas is greater than 1/2 the length of the pancreas; (3) Significant autoimmune processes outside the pancreas, including sclerosing cholangitis, renal involvement, and retroperitoneal fibrosis, which may suggest fAIP; (4) CT images have severe artifacts. Finally, 32 patients with fAIP were included in our study (23 males, 9 females; mean ± SD: 60 ± 12.1 years; range: 43–82 years). Other patients from our hospital with PDAC pathologically confirmed between January 2017 and January 2022 were also considered. The exclusion criteria were as follows: (1) Received any type of treatment (radiation, chemotherapy, or chemoradiation) prior to the imaging study; (2) Enhanced CT scan was not performed within 1 month before surgery; (3) History of other malignancies; (4) CT images have severe artifacts. Finally, 64 patients with PDAC were included in our study (47 males, 17 females; mean ± SD: 60.1 ± 9.8 years; range: 40–88 years). Then, all patients were randomly divided into training cohort and test cohort at ratio of 7:3 (Figure 1). The clinical data were derived from medical records.

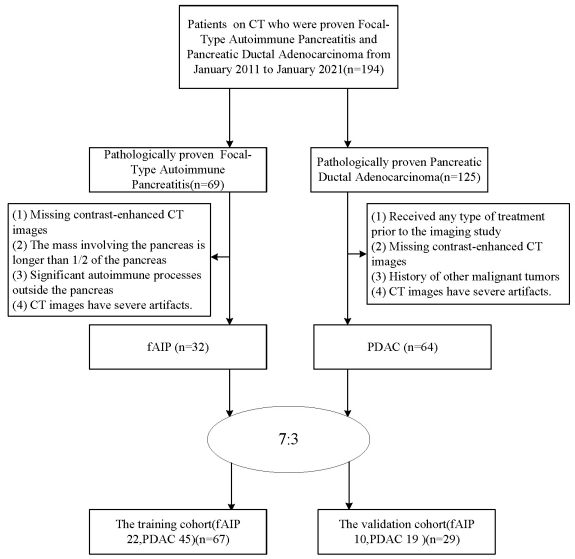


FIGURE 1
Patient Screening Flowchart.

2.2 CT image acquisition

All patients were scanned with a 64-slice multidetector CT (SOMATOM, Definition AS+, Siemens, Forchheim, Germany). The parameters involved were as follows: 120 kVp; effective 180 mA; rotation time, 0.5 s; detector collimation, 32 × 1.2 mm; field of view, 350 × 350 mm; matrix, 512 × 512; section thickness, 5 mm; and reconstruction section thickness, 1.5 mm. All patients were required to fast for at least 6 hours and drink 500 to 800 mL water before the examination. Contrast-enhanced CT images were obtained after intravenous administration of nonionic contrast medium (Ultravist 300 mg I/mL; Bayer Schering Pharma AG, Berlin, Germany) at an injection rate of 2.5-3.0 mL/s using a power injector (1.5 mL/kg). The arterial phase images were scanned at 7 seconds after the attenuation value of abdominal aorta reached 100 Hounsfield units. The portal venous phase images were scanned at 40 seconds after the completion of the arterial phase scanning.

2.3 CT findings evaluation

The pancreas lesions CT images for each patient were independently evaluated and recorded in a blinded manner by two experienced abdominal radiologists (10 and 20 years of experience in the interpretation of abdominal radiology). If there was a discrepancy between the two radiologists for some cases, they would reach a consensus after reviewing the images again and consulting. CT quantitative parameters were based on the mean values recorded by two radiologists. Due to the different size of the lesions, the slices of patients are also different, each lesion is segmented into approximately 30 slices, each slice 1mm thick.

The CT images were analyzed, considering: (1) Location of lesions (head-neck and body-tail of the pancreas); (2) The size of the lesion (the largest diameter of the tumor in cross section) (3) Capsule-like rim; (4) Pancreatic atrophy (5) Biliary wall thickening(thickness≥3 mm); (6) Peripancreatic vascular involvement (invasion of the common hepatic artery, splenic artery and vein, gastroduodenal artery, superior mesenteric artery and vein, portal vein; the standard is vascular occlusion, stenosis, or more than half of the circumference is in contact with the tumor); (7) Regional lymph node swelling (Lymph node short diameter≥1 cm); (8) Abrupt bile duct cut-off; (9) Pancreatic ductal cut-off; (10) MPD dilatation upstream (Upstream PD expansion ≥ 5 mm).

2.4 Segmentation and feature extraction

The construction process of the radiomics nomogram model is shown in Figure 2. The whole process includes: (A) Data acquisition and segmentation; (B) Feature extraction; (C) Feature screening and (D) Radiomics nomogram construction and evaluation.

2.4.1 Image segmentation, feature extraction, and data preprocessing

We used the open-source software LIFEx (https://www.lifexsoft.org/index.php) to manually draw the three-dimensional volume of interest (VOI) of CT venous phase lesions.

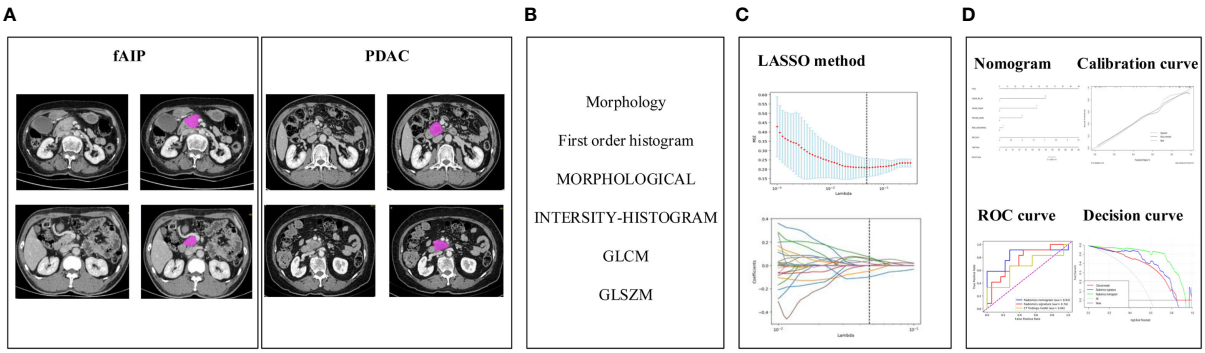


FIGURE 2
The research process including: **(A)** Data acquisition and segmentation; **(B)** Feature extraction; **(C)** Feature screening and **(D)** Radiomics nomogram construction and evaluation.

Particular care was taken to avoid the common bile duct and blood vessels while drawing the VOI. The segmentation process was performed by two experienced radiologists (10 and 20 years of experience in abdominal imaging), both of whom were blinded to the clinicopathological information of the patients, except the tumor location. The segmentation was finally completed with the consensus of the two radiologists.

LIFEx software is an open infrastructure software platform that flexibly supports common radiomics workflow tasks and is widely used in radiomics analysis. In our study, some parameters of LIFEx are as follows: In intensity discretization, nb of grep levels=400.0 and size of bins=10.0. In intensity Rescaling, min bound=-1000.0, max bound=3000.0 (18). We used LIFEx to extract 158 quantitative radiomics features. For each image, these features included six categories: morphological features, intensity features, grey-level cooccurrence matrix features (GLCM), grey-level distance zone-based features (GLDM), grey-level run-length matrix features (GLRLM), gray-level size zone matrix features (GLSZM), neighborhood grey tone difference-based features (NGTDM). The list of specific features we extracted is shown in Table 1. During data collection and image screening, we performed a normalization to ensure the reproducibility of our results.

TABLE 1 The extracted features using LIFEx toolbox.

Feature type	Feature name
MORPHOLOGICAL	MORPHOLOGICAL_Volume(IBSI : RNU0)
	MORPHOLOGICAL_ApproximateVolume(IBSI : YEKZ)
	MORPHOLOGICAL_voxelsCounting(IBSI : No)
	MORPHOLOGICAL_Compactness1(IBSI : SKGS)
	MORPHOLOGICAL_Compactness2(IBSI : BQWJ)
	MORPHOLOGICAL_SphericalDisproportion(IBSI : KRCK)
	MORPHOLOGICAL_Sphericity(IBSI : QCFX)
	MORPHOLOGICAL_Asphericity(IBSI:25C7)
	MORPHOLOGICAL_MaxValueCoordinates(IBSI : No)
	MORPHOLOGICAL_CenterOfMass(IBSI : No)
	MORPHOLOGICAL_WeightedCenterOfMass(IBSI : No)
	MORPHOLOGICAL_Hoc(IBSI : No)
	MORPHOLOGICAL_NormalizedHocRadiusRoi(IBSI : No)
	MORPHOLOGICAL_NormalizedHocRadiusSphere(IBSI : No)
	MORPHOLOGICAL_CentreOfMassShift(IBSI : KLMA)
	MORPHOLOGICAL_NormalizedHocRadiusRoi(IBSI : No)
	MORPHOLOGICAL_NormalizedHocRadiusSphere(IBSI : No)
	MORPHOLOGICAL_CentreOfMassShift(IBSI : KLMA)
	MORPHOLOGICAL_NormalizedHocRadiusRoi(IBSI : No)
	MORPHOLOGICAL_NormalizedHocRadiusSphere(IBSI : No)
INTENSITY	INTENSITY-BASED_Mean(HU)IBSI:Q4LE
	INTENSITY-BASED_Variance(HU)IBSI : ECT3
	INTENSITY-BASED_Skewness(HU)IBSI : KE2A
	INTENSITY-BASED_Kurtosis(HU)IBSI : IPH6
	INTENSITY-BASED_Median(HU)IBSI:Y12H
	INTENSITY-BASED_MinimumGreyLevel(HU)IBSI:1GSF
	INTENSITY-BASED_10thPercentile(HU)IBSI : QG58
	INTENSITY-BASED_25thPercentile(HU)IBSI : No
	INTENSITY-BASED_50thPercentile(HU)IBSI:Y12H
	INTENSITY-BASED_75thPercentile(HU)IBSI : No

(Continued)

TABLE 1 Continued

Feature type	Feature name
	INTENSITY-BASED_90thPercentile(HU)IBSI:8DWT
	INTENSITY-BASED_StandardDeviation(HU)IBSI : No
	INTENSITY-BASED_MaximumGreyLevel(HU)IBSI:84IY
	INTENSITY-BASED_InterquartileRange(HU)IBSI : SALO
	INTENSITY-BASED_Range(HU)IBSI:2OJQ
	INTENSITY-BASED_MeanAbsoluteDeviation(HU)IBSI:4FUA
	INTENSITY-BASED_RobustMeanAbsoluteDeviation(HU)IBSI:1128
	INTENSITY-BASED_MedianAbsoluteDeviation(HU)IBSI:N72L
	INTENSITY-BASED_CoefficientOfVariation(HU)IBSI:7TET
	INTENSITY-BASED_QuartileCoefficientOfDispersion(HU)IBSI:9S40
	INTENSITY-BASED_AreaUnderCurveCsh(HU)IBSI : No
	INTENSITY-BASED_Energy(HU)IBSI:N8CA
	INTENSITY-BASED_RootMeanSquare(HU)IBSI:5ZWQ
	INTENSITY-BASED_TotalLesionGlycolysis(HU)IBSI : No
	INTENSITY-BASED_TotalCalciumScoreIBSI : No
	LOCAL_INTENSITY_BASED_IntensityPeakDiscretizedVolumeSought(0.5mL)(mL)IBSI : No
	LOCAL_INTENSITY_BASED_GlobalIntensityPeak(0.5mL)(HU)IBSI : No
	LOCAL_INTENSITY_BASED_IntensityPeakDiscretizedVolumeSought(1mL)(mL)IBSI : No
	LOCAL_INTENSITY_BASED_GlobalIntensityPeak(1mL)(HU)IBSI:0F91
	LOCAL_INTENSITY_BASED_LocalIntensityPeak(HU)IBSI : VJGA
	INTENSITY-BASED-RIM_Min(HU)IBSI : No
	INTENSITY-BASED-RIM_Mean(HU)IBSI : No
	INTENSITY-BASED-RIM_Stdev(HU)IBSI : No
	INTENSITY-BASED-RIM_Max(HU)IBSI : No
	INTENSITY-BASED-RIM_CountingVoxels(#vx)IBSI : No
	INTENSITY-BASED-RIM_ApproximateVolume(mL)IBSI : No
	INTENSITY-BASED-RIM_Sum(HU)IBSI : No
	INTENSITY-HISTOGRAM_IntensityHistogramMean(HU)IBSI:X6K6
	INTENSITY-HISTOGRAM_IntensityHistogramVariance(HU)IBSI : CH89
	INTENSITY-HISTOGRAM_IntensityHistogramSkewness(HU)IBSI:88K1
	INTENSITY-HISTOGRAM_IntensityHistogramKurtosis(HU)IBSI:C3I7
	INTENSITY-HISTOGRAM_IntensityHistogramMedian(HU)IBSI : WIFQ
	INTENSITY-HISTOGRAM_IntensityHistogramMinimumGreyLevel(HU)IBSI:1PR8
	INTENSITY-HISTOGRAM_IntensityHistogram10thPercentile(HU)IBSI : GPMT
	INTENSITY-HISTOGRAM_IntensityHistogram25thPercentile(HU)IBSI : No
	INTENSITY-HISTOGRAM_IntensityHistogram50thPercentile(HU)IBSI : No
	INTENSITY-HISTOGRAM_IntensityHistogram75thPercentile(HU)IBSI : No
	INTENSITY-HISTOGRAM_IntensityHistogram90thPercentile(HU)IBSI : OZ0C
	INTENSITY-HISTOGRAM_IntensityHistogramStandardDeviation(HU)IBSI : No

(Continued)

TABLE 1 Continued

Feature type	Feature name
	INTENSITY-HISTOGRAM_IntensityHistogramMaximumGreyLevel(HU)IBSI:3NCY
	INTENSITY-HISTOGRAM_IntensityHistogramMode(HU)IBSI : AMMC
	INTENSITY-HISTOGRAM_IntensityHistogramInterquartileRange(HU)IBSI : WR0O
	INTENSITY-HISTOGRAM_IntensityHistogramRange(HU)IBSI:5Z3W
	INTENSITY-HISTOGRAM_IntensityHistogramMeanAbsoluteDeviation(HU)IBSI:D2ZX
	INTENSITY-HISTOGRAM_IntensityHistogramRobustMeanAbsoluteDeviation(HU)IBSI : WRZB
	INTENSITY-HISTOGRAM_IntensityHistogramMedianAbsoluteDeviation(HU)IBSI:4RNL
	INTENSITY-HISTOGRAM_IntensityHistogramCoefficientOfVariation(HU)IBSI : CWYJ
	INTENSITY-HISTOGRAM_IntensityHistogramQuartileCoefficientOfDispersion(HU)IBSI : SLWD
	INTENSITY-HISTOGRAM_IntensityHistogramEntropyLog10(HU)IBSI : No
	INTENSITY-HISTOGRAM_IntensityHistogramEntropyLog2(HU)IBSI : TLU2
	INTENSITY-HISTOGRAM_AreaUnderCurveCsh(HU)IBSI : No
	INTENSITY-HISTOGRAM_MaximumHistogramGradient(HU)IBSI:12CE
	INTENSITY-HISTOGRAM_MaximumHistogramGradientGreyLevel(HU)IBSI:8E6O
	INTENSITY-HISTOGRAM_MinimumHistogramGradient(HU)IBSI : VQB3
	INTENSITY-HISTOGRAM_MinimumHistogramGradientGreyLevel(HU)IBSI : RHQZ
	LOCAL_INTENSITY_HISTOGRAM_IntensityPeakDiscretizedVolumeSought(0.5mL)(mL)IBSI : No
	LOCAL_INTENSITY_HISTOGRAM_GlobalIntensityPeak(0.5mL)(HU)IBSI : No
	LOCAL_INTENSITY_HISTOGRAM_IntensityPeakDiscretizedVolumeSought(1mL)(mL)IBSI : No
	LOCAL_INTENSITY_HISTOGRAM_GlobalIntensityPeak(1mL)(HU)IBSI : No
	LOCAL_INTENSITY_HISTOGRAM_LocalIntensityPeak(HU)IBSI : No
	INTENSITY-HISTOGRAM-RIM_Min(HU)IBSI : No
	INTENSITY-HISTOGRAM-RIM_Mean(HU)IBSI : No
	INTENSITY-HISTOGRAM-RIM_Stdev(HU)IBSI : No
	INTENSITY-HISTOGRAM-RIM_Max(HU)IBSI : No
	INTENSITY-HISTOGRAM-RIM_CountingVoxels(#vx)IBSI : No
	INTENSITY-HISTOGRAM-RIM_ApproximateVolume(mL)IBSI : No
	INTENSITY-HISTOGRAM-RIM_Sum(HU)IBSI : No
GLCM	GLCM_JointMaximum(IBSI : GYBY)
	GLCM_JointAverage(IBSI:60VM)
	GLCM_JointVariance(IBSI : UR99)
	GLCM_JointEntropyLog2(IBSI : TU9B)
	GLCM_JointEntropyLog10(IBSI : No)
	GLCM_DifferenceAverage(IBSI : TF7R)
	GLCM_DifferenceVariance(IBSI:D3YU)
	GLCM_DifferenceEntropy(IBSI : NTRS)
	GLCM_SumAverage(IBSI : ZGXS)
	GLCM_SumVariance(IBSI : OEEB)
	GLCM_SumEntropy(IBSI:P6QZ)

(Continued)

TABLE 1 Continued

Feature type	Feature name
	GLCM_AngularSecondMoment(IBSI:8ZQL)
	GLCM_Contrast(IBSI : ACUI)
	GLCM_Dissimilarity(IBSI:8S9J)
	GLCM_InverseDifference(IBSI : IB1Z)
	GLCM_NormalisedInverseDifference(IBSI : NDRX)
	GLCM_InverseDifferenceMoment(IBSI : WF0Z)
	GLCM_NormalisedInverseDifferenceMoment(IBSI:1QCO)
	GLCM_InverseVariance(IBSI:E8JP)
	GLCM_Correlation(IBSI : NI2N)
	GLCM_Autocorrelation(IBSI : QWB0)
	GLCM_ClusterTendency(IBSI : DG8W)
	GLCM_ClusterShade(IBSI:7NFM)
	GLCM_ClusterProminence(IBSI : AE86)
GLRLM	GLRLM_ShortRunsEmphasis(IBSI:22OV)
	GLRLM_LongRunsEmphasis(IBSI:W4KF)
	GLRLM_LowGreyLevelRunEmphasis(IBSI:V3SW)
	GLRLM_HighGreyLevelRunEmphasis(IBSI:G3QZ)
	GLRLM_ShortRunLowGreyLevelEmphasis(IBSI : HTZT)
	GLRLM_ShortRunHighGreyLevelEmphasis(IBSI : GD3A)
	GLRLM_LongRunLowGreyLevelEmphasis(IBSI : IVPO)
	GLRLM_LongRunHighGreyLevelEmphasis(IBSI:3KUM)
	GLRLM_GreyLevelNonUniformity(IBSI:R5YN)
	GLRLM_RunLengthNonUniformity(IBSI:W92Y)
	GLRLM_RunPercentage(IBSI:9ZK5)
NGTDM	NGTDM_Coarseness(IBSI : QCDE)
	NGTDM_Contrast(IBSI:65HE)
	NGTDM_Busyness(IBSI : NQ30)
	NGTDM_Complexity(IBSI : HDEZ)
	NGTDM_Strength(IBSI:1X9X)
GLSZM	GLSZM_SmallZoneEmphasis(IBSI:5QRC)
	GLSZM_LargeZoneEmphasis(IBSI:48P8)
	GLSZM_LowGrayLevelZoneEmphasis(IBSI : XMSY)
	GLSZM_HighGrayLevelZoneEmphasis(IBSI:5GN9)
	GLSZM_SmallZoneLowGreyLevelEmphasis(IBSI:5RAI)
	GLSZM_SmallZoneHighGreyLevelEmphasis(IBSI : HW1V)
	GLSZM_LargeZoneLowGreyLevelEmphasis(IBSI : YH51)
	GLSZM_LargeZoneHighGreyLevelEmphasis(IBSI:J17V)
	GLSZM_GreyLevelNonUniformity(IBSI : JNSA)
	GLSZM_NormalisedGreyLevelNonUniformity(IBSI:Y1RO)

(Continued)

TABLE 1 Continued

Feature type	Feature name
	GLSZM_ZoneSizeNonUniformity(IBSI:4JP3)
	GLSZM_NormalisedZoneSizeNonUniformity(IBSI : VB3A)
	GLSZM_ZonePercentage(IBSI:P30P)
	GLSZM_GreyLevelVariance(IBSI : BYLV)
	GLSZM_ZoneSizeVariance(IBSI:3NSA)
	GLSZM_ZoneSizeEntropy(IBSI : GU8N)

2.4.2 Intra- and inter-observer reliability

To assess inter-observer reliability, blinded two radiologists performed VOI segmentation. For intra-observer reliability, features were extracted twice by the first observer at a one-month interval. Reliability was calculated using the intraclass correlation coefficient (ICC). Radiomics signatures with both intra- and inter-observer ICC values greater than 0.75 (indicating excellent stability) were selected for follow-up investigations.

2.4.3 Dimensionality reduction and feature selection

Feature selection consists of two steps: independent samples t-test and least absolute shrinkage and selection operator (LASSO) logistic regression algorithm. Regarding the selection of hyperparameters of the LASSO algorithm, after repeated training, we selected *alphas*=[0.001, 0.05, 50], and the final optimal *alpha* value was 0.04832; we selected *cv*=5, which was determined according to the amount of data, in order to ensure that the number of each sample set divided is more than 15 samples, thereby ensuring the stability of the model; *max_iter*=100000 is selected, to ensure that there are enough iterations for the model to complete the training. The other parameters and their values have been added to the additional file. Finally, each patient's radiomics score (Rad-Score) was calculated using a linear combination of selected features weighted by the respective coefficients.

2.4.4 Machine learning classifier selection

We analyzed the classification performance of the following four most used classifiers: Multivariate Logistic Regression (MLR), Random Forest (RF), Support Vector Machine (SVM), and Decision Tree (DT). These four classifiers were used to train the feature data in the training cohort. The 5-fold cross-validation method was used to ensure the stability and reliability of the training results, the classification performance was evaluated using the test cohort, and the hyperparameters of the four classifiers can be found in the additional file. To ensure that the number of samples in each data set divided is more than 15, and to ensure the training effect of the classifier, k = 5 has been empirically determined through the trial-and-error method (k range: 5–15, step size of 5) (19). To obtain the same percentage of patient status in both training and test datasets, in each training process, although the sample size of training is determined by the total amount of data, the sample size of the two types of data is equal.

2.5 Radiomics nomogram construction and evaluation

We developed combinatorial models combining CT findings and radiomic features. A radiomics nomogram was then generated from the above features by MLR, providing clinicians with the appropriate tool to differentiate between fAIP and PDAC in each patient. We then plotted a calibration curve for the nomogram, graphically showing the agreement between the predicted and actual probabilities of the radiomics nomogram, and presented the prediction process and results with two randomly selected patients and assessment. To further measure the predictive performance of the combined model, we used the receiver operating characteristic (ROC) area under the curve (AUC) to quantify the radiomics nomogram with 95% confidence interval (95% CI) and compared it to the single model. In order to ensure the consistency of the classifiers and then correctly evaluate the predictive ability of each model, we used the MLR classifier on both the CT findings model and the radiomics model for classification, depending on previous studies (20). Finally, the decision curves for the three models were plotted to assess the overall net benefit performance of the radiomics nomogram model.

2.6 Statistical analysis

All the statistical analyses were performed using R software (version 3.6.0, https://www.r-project.org) and Python (version 3.7.0, https://www.python.org). Continuous variables were expressed as mean ± standard deviation and compared by independent t-test with normal distribution or Mann-Whitney U test with abnormal distribution. Differences in categorical variables were analyzed by chi-square test or Fisher's exact test. Multivariate logistic regression analysis was used to select independent predictors in the subjective CT findings model. Values with two-sided $P < 0.05$ were considered statistically significant.

3 Results

3.1 Clinical characteristics and CT findings model

The clinical characteristics of the patients with fAIP and PDAC are listed in Table 2, and the CT findings of patients are shown in Table 3. All clinical characteristics showed no significant difference

between the fAIPs group and the PADCs group (P=0.325~0.873). CT images, including capsule-like rim, pancreatic atrophy, biliary wall thickening and vascular invasion, differed significantly between the two groups (P<0.05), indicating these features have a certain role in the diagnosis of fAIP and PDAC. There is no significant difference between groups in other characteristics (P>0.05).

3.2 Radiomic signature construction and evaluation

The radiomics feature selection process was performed separately at various stages. Based on venous phase CT images, 38 features were initially extracted by independent samples t-test. After removing redundant features, 7 potential features were selected by the LASSO algorithm. Then, a multiparametric radiomics signature based on venous phase images was established (Figure 3), the final filter gets the feature name and its weight performance (Figure 4).

According to the feature data screened by the LASSO algorithm, different classifiers based on machine learning to classify the feature data were used, and 5-fold cross-validation was performed to ensure the stability of the classification results. The test cohort was then used to verify the performance of different classifiers, and obtain the classification results as shown in Table 4.

Specifically, the favorable radiomics signature can be expressed by Rad-score:

$$\begin{aligned} Rad-Score = {} & 16.1747 - (0.009034 * MORPHOLOGICAL\_SurfaceToVolumeRatio) \\ & + (0.093138 * MORPHOLOGICAL\_Spehericity) - (0.019595 * INTENSITY \\ & - HISTOGRAM\_IntensityHistogram90thPercentile) - (0.161556 * INTENSITY \\ & - HISTOGRAM\_MaximumHistogramGradientGreyLevel) - (0.036626 * GLCM \\ & \_DifferenceVariance) + (0.103065 * GLCM\_Correlation) - (0.148918 * GLSZM \\ & \_ZonePercentage) \end{aligned}$$

After the Rad-score calculation for the fAIPs group (median: -0.81; range: -2.50~-0.09) was significantly lower than that of the PDACs group (median: -0.34; range: -0.63~0.62). We tested both sets of data using an independent samples t-test and found p-values < 0.001 for both sets of data.

3.3 Radiomics nomogram construction and validation

Five characteristics including the capsule-like rim, pancreatic atrophy, biliary wall thickening, vascular invasion and Rad-Score

TABLE 2 Clinical characteristics of patients.

Characteristics	fAIPs (n=32)	PDACs (n=64)	*P* value
Age (year), mean ± SD	60 ± 12.1	60.1 ± 9.8	0.635
Gender			0.873
Male	23	47	
Female	9	17	
Location			0.645
Head and neck	21	45	
Body and tail	11	19	
Maximum section diameter, mean ± SD	44.8 ± 16.7	41.6 ± 13.8	0.325

fAIPs, focal-type autoimmune pancreatitis; PDACs, pancreatic ductal adenocarcinoma; SD, standard deviation.

TABLE 3 CT findings of the patients.

Characteristics	fAIPs (n=32)	PDACs (n=64)	*P* value
Capsule-like rim	20	3	<0.001*
Regional lymph node swelling	7	25	0.094
Abrupt bile duct cut-off	2	12	0.104
Pancreatic atrophy	7	35	0.002*
Pancreatic ductal cut-off	4	19	0.064
Biliary wall thickening	17	4	<0.001*
Vascular invasion	2	51	<0.001*
MPD dilatation upstream(>5mm)	13	33	0.317

fAIPs, focal-type autoimmune pancreatitis; PDACs, pancreatic ductal adenocarcinoma; *P<0.05.

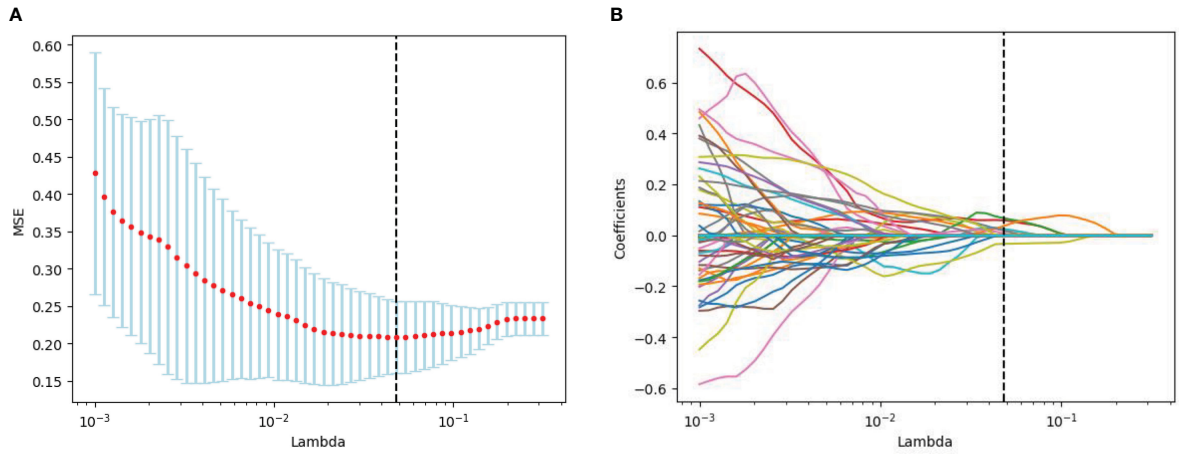


FIGURE 3
Radiomics feature screening by LASSO regression algorithm. **(A)** Plot of polynomial deviation versus λ. The red dots represented the mean deviation value for each model with a given λ, the vertical line was plotted at the best value by using the minimum criterion, where 7 features had non-zero coefficients. **(B)** Distribution of LASSO coefficients for radiological features. Each colored line represents the coefficient of each feature.

were included in the multivariate logistic regression analysis, and a combined model of radiomics nomogram was constructed (Figure 5). Figure 6 shows that the nomogram calibration curve with good agreement between predictions and observations in both cohorts. In addition, we randomly selected two patients and used the radiomics nomogram model for prediction. The prediction process and results are shown in Figures 7A, B.

The Figure 9 presents the DCA curves. We observed that the patients would benefit more from the radiomics nomograms than either the treat-no-patient schemes or the treat-all-patients regimens. Furthermore, the DCA curve showed that the radiomics nomogram had a higher net benefit than the curvilinear CT discovery model and the radiomics model in identifying patients with PDAC.

3.4 Comparison between different models

The ROC curves (Figure 8) analyzed the diagnostic ability of three different models in the training and test cohort. Radiomics nomogram showed the best diagnostic performance in both training (AUC = 0.87) and test cohort (AUC = 0.83), followed by radiomics signature (training cohort, AUC = 0.73; test cohort, AUC = 0.76). Both models outperformed the model based on CT findings in both the training (AUC = 0.67) ($P < 0.05$) and test cohorts (AUC = 0.66) ($P < 0.05$).

4 Discussion

In the present study, we developed and validated a diagnostic radiomics nomogram model combining subjective CT findings and radiomic features as a novel and effective complementary method for preoperative identification of fAIP and PDAC. The calibration curve, ROC curve and decision curve were used to verify the discriminating efficacy of our model. All evaluation metrics show that the nomogram model outperforms the single model in distinguishing fAIP and PDAC, and the nomogram model

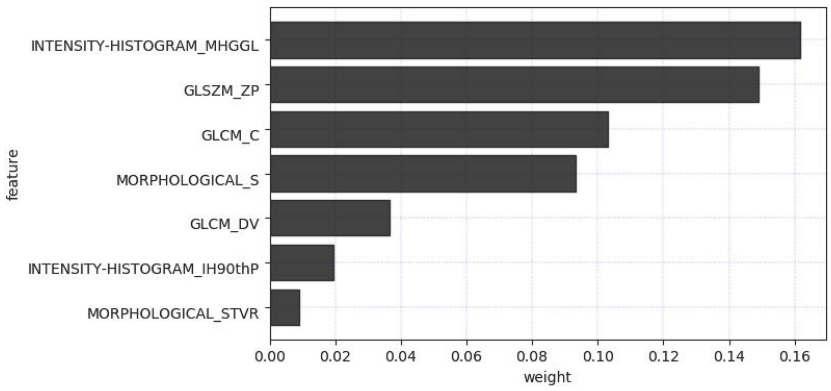


FIGURE 4
Filtered feature names and its weight performance.

TABLE 4 Classification performance of different classifiers.

Classifier	Training cohort (n=67)				Test cohort (n=29)			
	ACC	AUC	Sensitivity	Specificity	ACC	AUC	Sensitivity	Specificity
MLR	0.72	0.73	0.95	0.97	0.71	0.76	0.89	0.93
RF	0.92	0.95	1.0	1.0	0.56	0.60	1.0	0.17
SVM	0.89	0.93	1.0	0.92	0.61	0.69	0.93	0.48
DT	0.69	0.71	0.95	0.90	0.65	0.70	0.87	0.72

MLR, Multivariate Logistic Regression; RF, Random Forest; SVM, Support Vector Machine; DT, Decision Tree; ACC, Accuracy; AUC, Area Under the Curve.

enables model visualization. The nomogram model has potential as a decision tool for the need for surgical resection.

Some previous studies (21) found that some imaging features were more correlated with fAIP in contrast-enhanced CT than in PDAC. These included capsule-like rim with low attenuation (5, 22–26), without atrophic changes in uninvolved pancreatic tissue (23, 27), without MPD dilatation upstream (> 5 mm) (28), and our study showed similar results. Furthermore, we also found that the biliary wall thickening is helpful in differentiating the two diseases. This may be due to the fact that AIP is a systemic fibro-inflammatory disease, most commonly involving the bile ducts, resulting in sclerosing cholangitis (SC), biliary wall thickening and bile duct stricture (29); In PDAC however, there is only external compression, with rare cases of bile wall thickening. However, the diagnostic accuracy of imaging studies depends on the presence or absence of characteristic symptoms and the overall experience of the radiologist.

In recent years, radiomics techniques have rapidly developed, with the radiomics analysis aiming to provide a quantitative measure of intralesional heterogeneity. This is helpful in assessing tumor aggressiveness, treatment response and prognosis, and distinguishing benign from malignant lesions (29). The radiomics value in distinguishing between AIP and PDAC has been previously reported (21, 30–34). By extracting the radiomics features of the venous phase, Park et al (31) could distinguish AIP from PDAC with 89.7% sensitivity, 100% specificity, and 95.2% overall accuracy. The classification effect is better than that of the arterial phase, so in our study, the imaging data of the venous phase was used for the diagnosis of the two diseases. However, the previous study did not focus on fAIP patients, but included both diffuse AIP and fAIP patients. Furthermore, Zhang Y et al (30) and Liu Z et al (33) noninvasively classified PDAC and AIP lesions using PET/CT images using a radiomics-based predictive model. (Mean AUC: 0.9668, Accuracy: 89.91%, Sensitivity: 85.31%, Specificity: 96.04%). The above results show that establishing a radiomics signature model significantly improves the diagnostic efficiency.

To obtain an appropriate model able to distinguish between fAIP and PDAC, we developed and validated three models, and found that the combined nomogram performed better than the radiomics model and the CT findings model (training cohort AUC were 0.87, 0.73 and 0.67, and the test cohort AUCs were 0.83, 0.76, and 0.66). The calibration curves showed good agreement between the predicted values and the actual results. The decision curves showed that the radiomics nomogram model had a higher net benefit than the individual CT findings model and radiomics model respectively. By acquiring high-throughput quantitative features from CT images, radiomics signatures allow the assessment of tumor heterogeneity and the spatial distribution of biologically relevant voxels (9).

In our study, a two-step feature selection process screened 7 best features from 158 radiomic features, suggesting that these 7 features play a relatively important role in identifying fAIP and PDAC. For example, “LoG” (Laplace Gaussian) and “GLCM” (Gray Level Co-occurrence Matrix) are features that have proven useful in

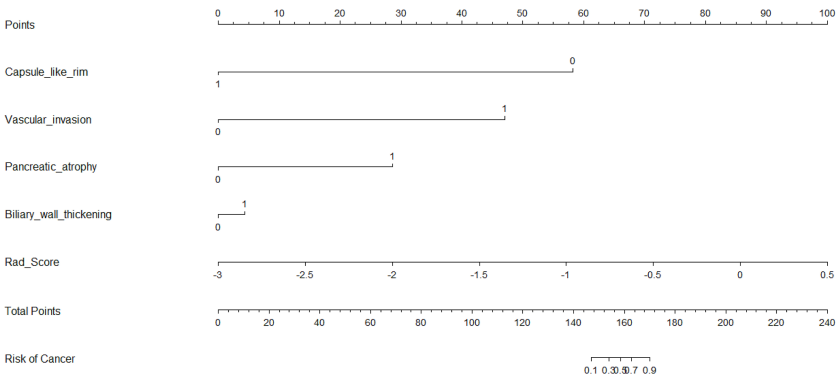


FIGURE 5
Nomogram for differentiating focal-type autoimmune pancreatitis (fAIPs) and pancreatic ductal adenocarcinoma (PDACs). The Capsule-like rim, Pancreatic atrophy, Biliary wall thickening, Vascular invasion and Rad-score were used for building the radiomics nomogram. Plotted the first scale “points” to identify points for each predictor. When the total points were calculated by adding the scores of these five predictors, the corresponding prediction probability was obtained at the last scale.

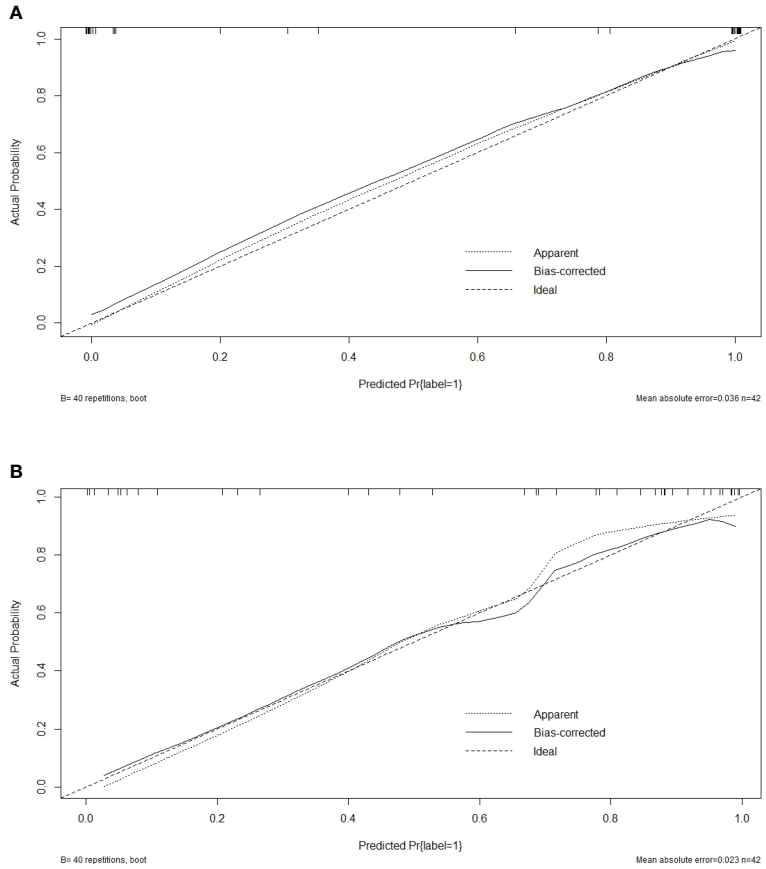


FIGURE 6
Calibration curves of the radiomics nomogram in training cohort **(A)** and test cohort **(B)**.

predicting the pathological features of certain tumor types (12, 35–37). We classified the filtered features using a variety of machine learning-based classifiers, and we chose these methods mainly because they were popular and performed well in previous studies (38). The performance of MLR classification is not the best on the training cohort, but it performed best on the test cohort. The performance of some classifiers in the training cohort and the test cohort is quite different. The preliminary judgment is that due to the problem of data volume and classifiers, RF and SVM were seriously overfitted. The radiomics features composed of the above 7 selected features are then represented by Rad-Score. When a patient has a high Rad-Score through CT image-based radiomic analysis, PDAC can be initially determined after comprehensive consideration. In addition, serum markers such as CA19-9 or IgG4 levels can be further detected, thereby establishing a personalized and convenient diagnostic system.

Histopathology obtained by endoscopic ultrasonography (EUS) guided fine-needle aspiration biopsy (EUS-FNA/B) is the gold standard for the AIP diagnosis. However, a recent multicenter study reported that the diagnosis rate for type I AIP using EUS-FNA/B was only 58.2% (39). EUS-FNB/B may not achieve definitive diagnosis even in the presence of large tissue volumes (40). The nomogram established in our study, combined with CT findings and radiomics features, is a non-invasive predictive tool that can analyze the overall characteristics of the lesion regardless of the location and size of the lesion. This may improve the accuracy of diagnosis, and reduce patient trauma with optimal compliance at the same time.

However, our study still has some limitations. First, CT images of fAIP patients were acquired over 11 years (2011 to 2021), whereas CT images of PDAC patients were acquired in the last 6 years (2017 to 2022). This may affect CT findings and features extracted. Second, due to the low incidence of fAIP, cases over nearly a decade have been included in our study, but there are still not enough cases to validate the proposed radiomics model, and selection bias is inevitable due to matched sampling. In order to verify the performance of our study on multi-center data, we initially selected CT image data of two patients from other hospital, and used the nomogram model to make predictions. The prediction results have been added to Figures 7C, D. It can be preliminarily seen from the prediction results that the nomogram model has good generalization ability and can be applied to new patient and multi-center data. However, the above-mentioned external verification data is seriously insufficient and has certain contingency. In follow-up study, we will continue to collect data and add more external validation data to enrich our study. And to overcome small and unbalanced sample size problems, the method

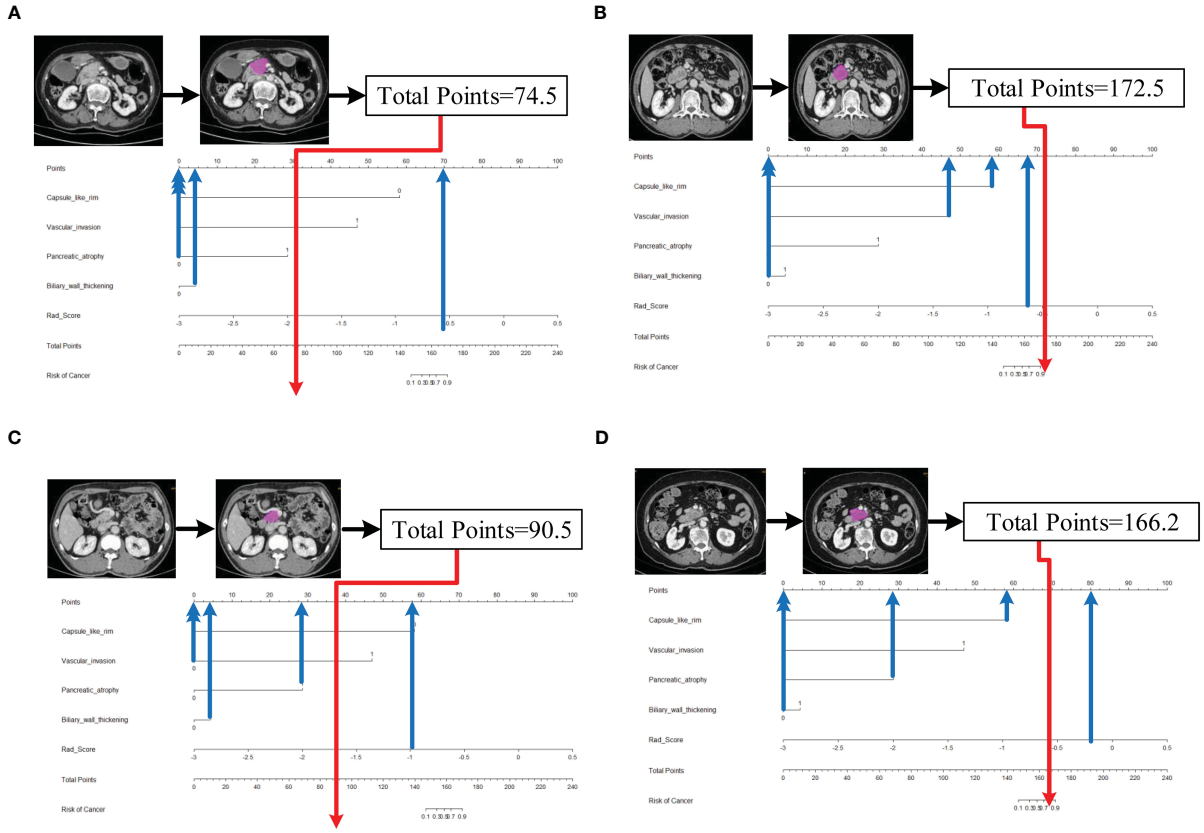


FIGURE 7
Schematic diagram of prediction flow of radiomics nomogram model. **(A)** after VOI delineating, image preprocessing, the value of total points was 74.5, which was calculated by the CT findings and Rad-Score. The result corresponded to <10% probability of a firm consistency. Thus, the patient's disease was predicted to be fAIP, which was confirmed by ICDC. **(B)**The total points was 172.5, which corresponding to >90% probability of a firm consistency. Thus, the patient's disease was predicted to be PDAC, which was confirmed in surgery. **(C)** The external validation data, the total points was 90.5. Thus, the patient's disease was predicted to be fAIP, which was confirmed by ICDC. **(D)** The external validation data, the total points was 166.2. Thus, the patient's disease was predicted to be PDAC, which was confirmed in surgery.

maybe the future directions of our study which was used by Stefano Barone et al (41). Third, the contours of VOIs of pancreatic lesions may have some influence on the performance of our prediction model. In the current study, two radiologists manually delineated the contours of the lesions, and it is a time-consuming process. Therefore, methods requiring less manual intervention should be considered, and the establishment of automated pancreas segmentation software may help improve this situation. Automated

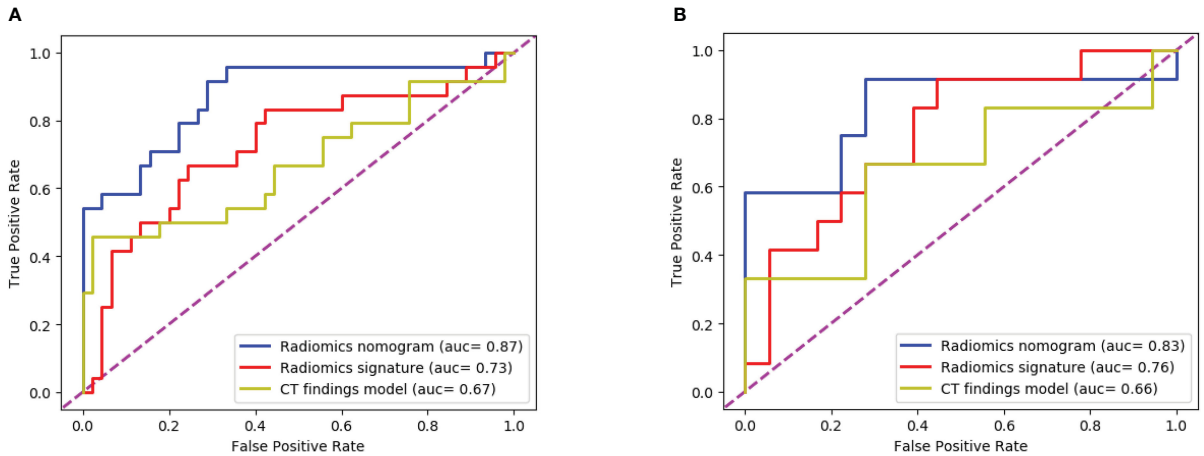


FIGURE 8
The ROC curves of the three models: **(A)** The training cohort; **(B)** The test cohort. The AUC values of the radiomics nomogram model were higher than that of the CT appearance model and the radiomics model in both training cohort and test cohort.

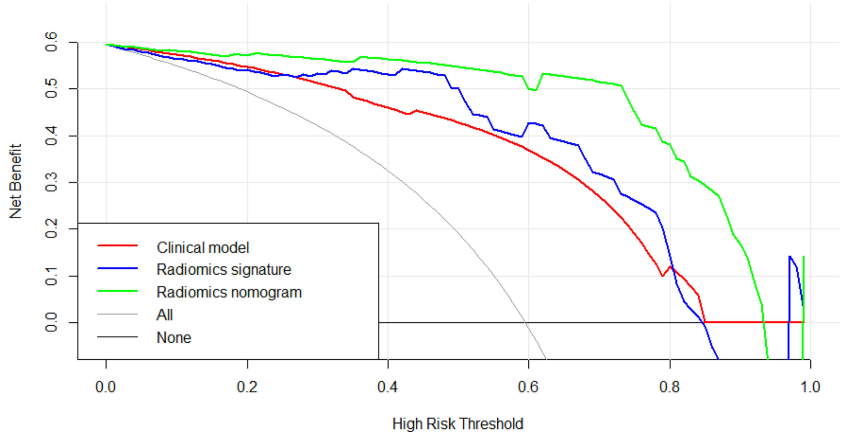


FIGURE 9
The decision curve analysis of CT findings model, radiomics model and radiomics nomogram in test cohort. The green line represents radiomics nomogram. The blue line represents radiomics model. The red line represents CT findings model. The grey line represents the assumption that all patients had pancreatic ductal adenocarcinoma (PDAC) (All). The black line represents the assumption that all patients had focal autoimmune pancreatitis (fAIP) (none). Threshold probabilities in clinical decisions take arbitrary values and patients will benefit more from the radiomics nomogram than either the treat-no-patient or treat-all-patient options.

segmentation is the development trend of lesion segmentation in radiomics (18, 42). But automated segmentation also has its disadvantages. Automated segmentation often requires a large amount of data for training, usually the methods used are based on deep learning, and in the current research status, automated segmentation can only identify disease with evident lesion areas. But for the identification of two types of diseases with complex lesion areas, the results of automated segmentation are often not appropriate. In our study, fAIP is a relatively rare disease, the amount of data is not enough to support the training of automated segmentation software, and the similarity between fAIP and PDAC is high, so some key features may be lost using automated segmentation. Therefore, we chose to use manual segmentation as the method of lesion segmentation. At the same time, to ensure the reproducibility of VOI, we selected two experienced abdominal radiologists to jointly segment VOI. But automated segmentation is still an important direction for our future studies, and we are continuing to collect relevant data to prepare for the construction of automated segmentation software. In theory, the only texture features that resulted to be reliable (ICC>0.75), could lead to the elimination of fundamental features for building the predictive model. This is indeed a limitation of our study, but it is already one of the best methods, and in our actual study, the reproducibility and quality of feature extraction are guaranteed due to the extensive experience of the physicians responsible for VOI segmentation.

In summary, we have developed a preoperative CT imaging-based radiomics nomogram for distinguishing between fAIP and PDAC with high accuracy and clear diagnostic value. Quantitative and noninvasive radiomics analysis may be a useful application to help clinicians develop personalized treatment plans.

Data availability statement

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

Ethics statement

The studies involving human participants were reviewed and approved by the ethics committee of the People's Hospital of China Medical University. Written informed consent for participation was not required for this study in accordance with the national legislation and the institutional requirements.

Author contributions

DL takes responsibility for the integrity of the work as a whole, from inception to published article, JL performed the research, collected and analyzed the data, completed the manuscript. NJ and YZ provided the technology support. All authors contributed to the article and approved the submitted version.

Funding

The People's Hospital of China Medical University and The People's Hospital of Liaoning Province. Shenyang Youth Science and Technology Project(RC200222).

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

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fonc.2023.979437/full#supplementary-material

References

1. Yoshida K, Toki F, Takeuchi T, Watanabe S-I, Shiratori K, Hayashi N. Chronic pancreatitis caused by an autoimmune abnormality: Proposal of the concept of autoimmune pancreatitis. *Digest Dis Sci* (1995) 40:1561–8. doi: 10.1007/BF02285209

2. Kamisawa T, Funata N, Hayashi Y, Eishi Y, Koike M, Tsuruta K, et al. A new clinicopathological entity of IgG4-related autoimmune disease. *J Gastroenterol* (2003) 38:982–4. doi: 10.1007/s00535-003-1175-y

3. Kamisawa T, Funata N, Hayashi Y, Tsuruta K, Okamoto A, Amemiya K, et al. Close relationship between autoimmune pancreatitis and multifocal fibrosclerosis. *Gut* (2003) 52:683–7. doi: 10.1136/gut.52.5.683

4. Sahani DV, Kalva SP, Farrell J, Maher MM, Saini S, Mueller PR, et al. Autoimmune pancreatitis: imaging features. *Radiology* (2004) 233:345–52. doi: 10.1148/radiol.2332031436

5. Muhi A, Ichikawa T, Motosugi U, Sou H, Sano K, Tsukamoto T, et al. Mass-forming autoimmune pancreatitis and pancreatic carcinoma: Differential diagnosis on the basis of computed tomography and magnetic resonance cholangiopancreatography, and diffusion-weighted imaging findings. *J Magn Reson Imaging* (2012) 35:827–36. doi: 10.1002/jmri.22881

6. Al-Hawary MM, Kaza RK, Azar SF, Ruma JA, Francis IR. Mimics of pancreatic ductal adenocarcinoma. *Cancer Imaging* (2013) 13:342–9. doi: 10.1102/1470-7330.2013.9012

7. Shimosegawa T, Chari ST, Frulloni L, Kamisawa T, Kawa S, Mino-Kenudson M, et al. International consensus diagnostic criteria for autoimmune pancreatitis: Guidelines of the international association of pancreatology. *Pancreas* (2011) 40:352–8. doi: 10.1097/MPA.0b013e3182142fd2

8. Mu W, Schabath MB, Gillies RJ. Images are data: Challenges and opportunities in the clinical translation of radiomics. *Cancer Res* (2022) 82:2066–8. doi: 10.1158/0008-5472.CAN-22-1183

9. Gillies RJ, Kinahan PE, Hricak H. Radiomics: Images are more than pictures, they are data. *Radiology* (2016) 278:563–77. doi: 10.1148/radiol.2015151169

10. Chakraborty J, Midya A, Gazit L, Attiyeh M, Langdon-Embry L, Allen PJ, et al. CT radiomics to predict high-risk intraductal papillary mucinous neoplasms of the pancreas. *Med Phys* (2018) 45:5019–29. doi: 10.1002/mp.13159

11. Hanania AN, Bantis LE, Feng Z, Wang H, Tamm EP, Katz MH, et al. Quantitative imaging to evaluate malignant potential of IPMNs. *Oncotarget* (2016) 7:85776–84. doi: 10.18632/oncotarget.11769

12. Canellas R, Burk KS, Parakh A, Sahani DV. Prediction of pancreatic neuroendocrine tumor grade based on CT features and texture analysis. *AJR Am J Roentgenol* (2018) 210:341–6. doi: 10.2214/AJR.17.18417

13. Choi TW, Kim JH, Yu MH, Park SJ, Han JK. Pancreatic neuroendocrine tumor: prediction of the tumor grade using CT findings and computerized texture analysis. *Acta Radiol* (2018) 59:383–92. doi: 10.1177/0284185117725367

14. Ciaravino V, Cardobi N DE, Robertis R, Capelli P, Melisi D, Simionato F, et al. CT texture analysis of ductal adenocarcinoma downstaged after chemotherapy. *Anticancer Res* (2018) 38:4889–95. doi: 10.21873/anticanres.12803

15. Lin X, Xu L, Wu A, Guo C, Chen X, Wang Z. Differentiation of intrapancreatic accessory spleen from small hypervascular neuroendocrine tumor of the pancreas: textural analysis on contrast-enhanced computed tomography. *Acta Radiol* (2019) 60:553–60. doi: 10.1177/0284185118788895

16. Chu LC, Park S, Kawamoto S, Fouladi DF, Shayesteh S, Zinreich ES, et al. Utility of CT radiomics features in differentiation of pancreatic ductal adenocarcinoma from normal pancreatic tissue. *AJR Am J Roentgenol* (2019) 213:349–57. doi: 10.2214/AJR.18.20901

17. Sandrasegaran K, Lin Y, Asare-Sawiri M, Taiyini T, Tann M. CT texture analysis of pancreatic cancer. *Eur Radiol* (2019) 29:1067–73. doi: 10.1007/s00330-018-5662-1

18. Stefano A, Leal A, Richiusa S, Trang P, Comelli A, Benfante V, et al. Robustness of PET radiomics features: Impact of Co-registration with MRI. *Appl Sci* (2021) 11:10170. doi: 10.3390/app112110170

19. Alongi P, Stefano A, Comelli A, Laudicella R, Scalisi S, Arnone G, et al. Radiomics analysis of 18F-choline PET/CT in the prediction of disease outcome in high-risk prostate cancer: an explorative study on machine learning feature classification in 94 patients. *Eur Radiol* (2021) 31:4595–605. doi: 10.1007/s00330-020-07617-8

20. Cheng S, Shi H, Lu M, Wang C, Duan S, Xu Q, et al. Radiomics analysis for predicting malignant potential of intraductal papillary mucinous neoplasms of the pancreas: Comparison of CT and MRI. *Acad Radiol* (2022) 29:367–75. doi: 10.1016/j.acra.2021.04.013

21. E L, Xu Y, Wu Z, Li L, Zhang N, Yang H, et al. Differentiation of focal-type autoimmune pancreatitis from pancreatic ductal adenocarcinoma using radiomics based on multiphasic computed tomography. *J Comput Assist Tomogr* (2020) 44:511–8. doi: 10.1097/RCT.0000000000001049

22. Kamisawa T, Imai M, Yui Chen P, Tu Y, Egawa N, Tsuruta K, et al. Strategy for differentiating autoimmune pancreatitis from pancreatic cancer. *Pancreas* (2008) 37:e62–67. doi: 10.1097/MPA.0b013e318175e3a0

23. Naitoh I, Nakazawa T, Hayashi K, Okumura F, Miyabe K, Shimizu S, et al. Clinical differences between mass-forming autoimmune pancreatitis and pancreatic cancer. *Scandinavian J Gastroenterol* (2012) 47:607–13. doi: 10.3109/00365521.2012.667147

24. Furuhashi N, Suzuki K, Sakurai Y, Ikeda M, Kawai Y, Naganawa S. Differentiation of focal-type autoimmune pancreatitis from pancreatic carcinoma: assessment by multiphase contrast-enhanced CT. *Eur Radiol* (2015) 25:1366–74. doi: 10.1007/s00330-014-3512-3

25. Sun G-F. Focal autoimmune pancreatitis: Radiological characteristics help to distinguish from pancreatic cancer. *WJG* (2013) 19:3634. doi: 10.3748/wjg.v19.i23.3634

26. Lee-Felker SA, Felker ER, Kadell B, Farrell J, Raman SS, Sayre J, et al. Use of MDCT to differentiate autoimmune pancreatitis from ductal adenocarcinoma and interstitial pancreatitis. *AJR Am J Roentgenol* (2015) 205:2–9. doi: 10.2214/AJR.14.14059

27. Chari ST, Longnecker DS, Klöppel G. The diagnosis of autoimmune pancreatitis: a Western perspective. *Pancreas* (2009) 38:846–8. doi: 10.1097/MPA.0b013e3181bba281

28. Chang WI, Kim BJ, Lee JK, Kang P, Lee KH, Lee KT, et al. The clinical and radiological characteristics of focal mass-forming autoimmune pancreatitis: Comparison with chronic pancreatitis and pancreatic cancer. *Pancreas* (2009) 38:401–8. doi: 10.1097/MPA.0b013e31818d92c0

29. Hirano K, Tada M, Isayama H, Yamamoto K, Mizuno S, Yagioka H, et al. Endoscopic evaluation of factors contributing to intrapancreatic biliary stricture in autoimmune pancreatitis. *Gastrointest Endosc* (2010) 71:85–90. doi: 10.1016/j.gie.2009.08.008

30. Zhang Y, Cheng C, Liu Z, Wang L, Pan G, Sun G, et al. Radiomics analysis for the differentiation of autoimmune pancreatitis and pancreatic ductal adenocarcinoma in 18 f-FDG PET/CT. *Med Phys* (2019) 46:4520–30. doi: 10.1002/mp.13733

31. Park S, Chu LC, Hruban RH, Vogelstein B, Kinzler KW, Yuille AL, et al. Differentiating autoimmune pancreatitis from pancreatic ductal adenocarcinoma with CT radiomics features. *Diagn Interv Imaging* (2020) 101:555–64. doi: 10.1016/j.diii.2020.03.002

32. Ziegelmayer S, Kaissis G, Harder F, Jungmann F, Müller T, Makowski M, et al. Deep convolutional neural network-assisted feature extraction for diagnostic discrimination and feature visualization in pancreatic ductal adenocarcinoma

(PDAC) versus autoimmune pancreatitis (AIP). *J Clin Med* (2020) 9:E4013. doi: 10.3390/jcm9124013

33. Liu Z, Li M, Zuo C, Yang Z, Yang X, Ren S, et al. Radiomics model of dual-time 2-[18F]FDG PET/CT imaging to distinguish between pancreatic ductal adenocarcinoma and autoimmune pancreatitis. *Eur Radiol* (2021) 31:6983–91. doi: 10.1007/s00330-021-07778-0

34. Li J, Liu F, Fang X, Cao K, Meng Y, Zhang H, et al. CT radiomics features in differentiation of focal-type autoimmune pancreatitis from pancreatic ductal adenocarcinoma: A propensity score analysis. *Acad Radiol* (2022) 29:358–66. doi: 10.1016/j.acra.2021.04.014

35. Ma Z, Fang M, Huang Y, He L, Chen X, Liang C, et al. CT-based radiomics signature for differentiating borrmann type IV gastric cancer from primary gastric lymphoma. *Eur J Radiol* (2017) 91:142–7. doi: 10.1016/j.ejrad.2017.04.007

36. Yu H, Scalera J, Khalid M, Touret A-S, Bloch N, Li B, et al. Texture analysis as a radiomic marker for differentiating renal tumors. *Abdom Radiol (NY)* (2017) 42:2470–8. doi: 10.1007/s00261-017-1144-1

37. Zhu X, Dong D, Chen Z, Fang M, Zhang L, Song J, et al. Radiomic signature as a diagnostic factor for histologic subtype classification of non-small cell lung cancer. *Eur Radiol* (2018) 28:2772–8. doi: 10.1007/s00330-017-5221-1

38. Zhai Y, Song D, Yang F, Wang Y, Jia X, Wei S, et al. Preoperative prediction of meningioma consistency *via* machine learning-based radiomics. *Front Oncol* (2021) 11:657288. doi: 10.3389/fonc.2021.657288

39. Ishikawa T, Kawashima H, Ohno E, Suhara H, Hayashi D, Hiramatsu T, et al. Usefulness of endoscopic ultrasound-guided fine-needle biopsy for the diagnosis of autoimmune pancreatitis using a 22-gauge franseen needle: a prospective multicenter study. *Endoscopy* (2020) 52:978–85. doi: 10.1055/a-1183-3583

40. Notohara K, Kamisawa T, Kanno A, Naitoh I, Iwasaki E, Shimizu K, et al. Efficacy and limitations of the histological diagnosis of type 1 autoimmune pancreatitis with endoscopic ultrasound-guided fine needle biopsy with large tissue amounts. *Pancreatology* (2020) 20:834–43. doi: 10.1016/j.pan.2020.05.026

41. Barone S, Cannella R, Comelli A, Pellegrino A, Salvaggio G, Stefano A, et al. Hybrid descriptive-inferential method for key feature selection in prostate cancer radiomics. *Appl Stoch Models Bus Ind* (2021) 37:961–72. doi: 10.1002/asmb.2642

42. Stefano A, Comelli A. Customized efficient neural network for COVID-19 infected region identification in CT images. *J Imaging* (2021) 7:131. doi: 10.3390/jimaging7080131

Frontiers in Oncology

Advances knowledge of carcinogenesis and tumor progression for better treatment and management

The third most-cited oncology journal, which highlights research in carcinogenesis and tumor progression, bridging the gap between basic research and applications to imrpove diagnosis, therapeutics and management strategies.

Discover the latest Research Topics

See more →

Frontiers

Avenue du Tribunal-Fédéral 34
1005 Lausanne, Switzerland
frontiersin.org

Contact us

+41 (0)21 510 17 00
frontiersin.org/about/contact