Precision Medicine and Imaging

Clinical Cancer Research

A Combined Nomogram Model to Preoperatively Predict Histologic Grade in Pancreatic Neuroendocrine Tumors

Wenjie Liang[1,2], Pengfei Yang[3,4,5], Rui Huang[5], Lei Xu[3,4], Jiawei Wang[6], Weihai Liu[7], Lele Zhang[8,9,10], Dalong Wan[10], Qiang Huang[1], Yao Lu[11], Yu Kuang[11], and Tianye Niu[3,4]

Abstract

Purpose: The purpose of this study is to develop and validate a nomogram model combing radiomics features and clinical characteristics to preoperatively differentiate grade 1 and grade 2/3 tumors in patients with pancreatic neuroendocrine tumors (pNET).

Experimental Design: A total of 137 patients who underwent contrast-enhanced CT from two hospitals were included in this study. The patients from the second hospital ($n = 51$) were selected as an independent validation set. The arterial phase in contrast-enhanced CT was selected for radiomics feature extraction. The Mann–Whitney U test and least absolute shrinkage and selection operator regression were applied for feature selection and radiomics signature construction. A combined nomogram model was developed by incorporating the radiomics signature with clinical factors. The association between the nomogram model and the Ki-67 index and rate of nuclear mitosis were also investigated respectively. The utility of the proposed model was evaluated using the ROC, area under ROC curve (AUC), calibration curve, and decision curve analysis (DCA). The Kaplan–Meier (KM) analysis was used for survival analysis.

Results: An eight-feature–combined radiomics signature was constructed as a tumor grade predictor. The nomogram model combining the radiomics signature with clinical stage showed the best performance (training set: AUC = 0.907; validation set: AUC = 0.891). The calibration curve and DCA demonstrated the clinical usefulness of the proposed nomogram. A significant correlation was observed between the developed nomogram and Ki-67 index and rate of nuclear mitosis, respectively. The KM analysis showed a significant difference between the survival of predicted grade 1 and grade 2/3 groups ($P = 0.002$).

Conclusions: The combined nomogram model developed could be useful in differentiating grade 1 and grade 2/3 tumor in patients with pNETs.

Introduction

Pancreatic neuroendocrine tumors (pNET) are heterogeneous neoplasm, which only accounts for about less than 5% of all pancreatic tumors (1). In the last two decades, we have seen a considerable increase of pNETs in the incidence and morbidity, especially due to the significant use of diagnostic imaging with contrast-enhanced detection rate for small nonfunctional pNETs (2, 3). Despite growing experience in the diagnosis and treatment of pNETs, the prognosis of patients with pNETs still varied as the nonfunctional tumors tend to present at a higher histologic grade with notable symptoms causing mass effect and/or tumor metastases (4).

Due to the differences in tumor proliferative pattern, functional status, and biology between well-differentiated and poorly differentiated pNETs, treatment decisions for patients with pNETs are usually guided after staging of the disease has been

[1]Department of Radiology, the First Affiliated Hospital, Zhejiang University School of Medicine, Hangzhou, Zhejiang, China. [2]Key Laboratory of Precision Diagnosis and Treatment for Hepatobiliary and Pancreatic Tumor of Zhejiang Province, the First Affiliated Hospital, Zhejiang University School of Medicine, Hangzhou, Zhejiang, China. [3]Department of Radiation Oncology, Sir Run Run Shaw Hospital, Zhejiang University School of Medicine, Hangzhou, Zhejiang, China. [4]Institute of Translational Medicine, Zhejiang University, Hangzhou, Zhejiang, China. [5]College of Biomedical Engineering and Instrument Science, Zhejiang University, Hangzhou, Zhejiang, China. [6]Department of Radiology, the Second Affiliated Hospital, Zhejiang University School of Medicine, Hangzhou, Zhejiang, China. [7]Department of Radiology, Beilun Branch Hospital of the First Affiliated Hospital, Zhejiang University School of Medicine, the People's Hospital of Beilun District, Ningbo, Zhejiang, China. [8]Collaborative Innovation Center for Diagnosis and Treatment of Infectious Diseases, the First Affiliated Hospital, Zhejiang University School of Medicine, Hangzhou, Zhejiang, China. [9]Key Lab of Combined Multi-Organ Transplantation, Ministry of Public Health, the First Affiliated Hospital, Zhejiang University School of Medicine, Hangzhou, Zhejiang, China. [10]Department of Hepatobiliary and Pancreatic Surgery, the First Affiliated Hospital, Zhejiang University School of Medicine, Hangzhou, Zhejiang, China. [11]Department of Medical Physics, University of Nevada, Las Vegas, Las Vegas, Nevada.

Note: Supplementary data for this article are available at Clinical Cancer Research Online (http://clincancerres.aacrjournals.org/).

W. Liang and P. Yang contributed equally to this article.

Corresponding Authors: Tianye Niu, Sir Run Run Shaw Hospital, Zhejiang University School of Medicine, Institute of Translational Medicine, Huajiachi Campus, Zhejiang University, Hangzhou, Zhejiang 310000, China. Phone: 86-571-88981576; E-mail: tyniu@zju.edu.cn; Yu Kuang, 4505 South Maryland Parkway, Box 453037, Las Vegas, NV 89121. Phone: 1-702-895-3555; Fax: 1-702-895-4819; E-mail: yu.kuang@unlv.edu; and Wenjie Liang, 79 # Qingchun Road, Hangzhou City, Zhejiang Province 310003, China. Phone: 86-571-87236114; Fax: 86-571-8723611; E-mail: baduen@zju.edu.cn

doi: 10.1158/1078-0432.CCR-18-1305

Translational Relevance

Despite growing experience in the diagnosis and treatment of pancreatic neuroendocrine tumors (pNET), the prognosis of patients with pNETs still varies as the preoperative prognostic stratification methods remain suboptimal. The significant limiting factor, tumor grade indexed by the proliferative indicator Ki-67 for prognosis within each tumor–node–metastasis stage, is usually assessed on postoperative specimens, thus hindering the individualized therapeutic decision making in clinical practice. A clinically translatable nomogram model incorporating the radiomics signature with tumor clinical stage developed in this study can effectively predict the pathologic grade of pNETs preoperatively (grade 1 vs. grade 2/3). The model also demonstrated a utility in predicting the postoperative prognosis of patients with pNETs. Therefore, the predictive nomogram model could serve as a preoperative, noninvasive, and precise evaluation tool for patients with pNETs, which may help clinicians tailor the treatment protocol for each individual patient and achieve a better clinical outcome in the future.

established. Surgery for localized disease is still the only curative treatment modality, but there are also other treatment options, e.g., targeted therapy or chemotherapy, for advanced diseases based on determined tumor grades.

Although the staging systems such as tumor diameter, depth of tumor invasion, the presence or absence of distant metastases, and range of tumor involvement have been used to guide the selection of treatment strategies, the preoperative prognostic stratification methods remain suboptimal. The most important factor, tumor grade is indexed by the proliferative indicator Ki-67 for prognosis within each tumor–node–metastasis stage. The latter is usually based on postoperative specimens, and the accuracy of preoperative fine-needle aspiration biopsy to determine tumor grade is still challenging in clinical practice. Thus, the method of preoperatively predicting the pathologic grade of pNETs is still urgently needed to help establish individualized therapeutic decisions.

Medical imaging plays a vital role in the preoperative evaluation of pNETs. Previous studies suggested that radiologic features derived from MRI and CT favor the identification of pNETs' pathologic grades (5–10). The low-to-intermediate signal intensity on preoperative T2-weighted MRI images with ill-defined borders and lower apparent diffusion coefficient (ADC) values suggesting tumor invasiveness and increased cellularity patterns were significantly correlated with high-grade pNETs. However, thresholds of ADC values for the differentiation of pNETs grades remain poorly defined, and the prediction accuracy using ADC also varies.

Use of CT to image the calcifications of pNETs and/or contrast enhancement patterns can predict the biological aggressiveness of pNETs associated with different tumor grades. Although using Hounsfield units to assess the degree of contrast enhancement patterns indeed sheds lights on the prognostic significance of CT for pNETs, the evaluation is not quantitative with a limitation of missing the spatial information within a single tumor due to the cystic or heterogeneous nature in most of pNETs.

Recently, radiomics studies using CT texture analysis of cancer images allowed to extract a series of quantitative imaging characteristics to pinpoint the valuable radiomics signatures through machine learning methods for disease diagnosis, tumor staging, and evaluation of curative effects. This strategy could be more useful for differentiating pathologic grading in patients with pNETs than routine CT image features alone (11, 12). However, to the best of our knowledge, a noninvasive optimal method to incorporate imaging biomarkers with clinical characteristics, e.g., tumor diameter, distant metastases status, and hormone secretion status, to improve preoperative prognostic stratification and predict the pathologic grade of the more aggressive pNETs preoperatively for better treatment strategies guidance has yet to be developed.

Therefore, we aimed to develop and validate a combined nomogram model that integrates radiomics signature derived from contrast-enhanced CT arterial phase images with clinical characteristics for personalized preoperative prediction of pathologic grades [grade 1 (G1) or grade 2/3 (G2/3)] in patients with pNETs.

Materials and Methods

Workflow

The workflow of the analysis is summarized in Fig. 1 and can be divided into four parts: image acquisition, region of interest (ROI) segmentation, feature extraction, and tumor pathologic grade classifier construction. Contrast-enhanced CT images were acquired, and tumor regions were manually contoured by radiologists on all image slices. Quantitative radiomics features were then extracted from the contoured ROIs to build a machine learning–based model to classify the tumor pathologic grade.

Two separated datasets were used to develop and validate the machine learning–based tumor pathologic grade classifier. The data from Institution I ($n = 86$) were used as a training dataset to derive the tumor pathologic grade classifier. The data from Institution II ($n = 51$) were used as an independent validation dataset to verify the classifier developed.

After the quantitative radiomics features had been extracted from the contoured tumor regions, the Mann–Whitney *U* test and least absolute shrinkage and selection operator (LASSO) regression were applied to select the optimal radiomics features to build a radiomics signature. The radiomics signature generated was then integrated with clinical characteristics to generate a tumor pathologic grade classifier through the multivariable logistic regression method.

Patients

This retrospective study was approved by the Institutional Review Board of the First Affiliated Hospital and the Second Affiliated Hospital, Zhejiang University School of Medicine (Zhejiang, China). The signed informed consent forms were waived. This study was conducted according to the Declaration of Helsinki. The inclusion criteria for patients were as follows: (i) patients diagnosed with pNETs had surgical tumor specimens; (ii) patients underwent preoperative pancreatic contrast-enhanced CT scan within 1 month before surgery; and (iii) patients had complete clinical imaging data and pathologic specimens available for reevaluation. The exclusion criteria for patients included: (i) patients had a pNET that was too small to display clearly on CT; and (ii) patients had a pNET that displayed an

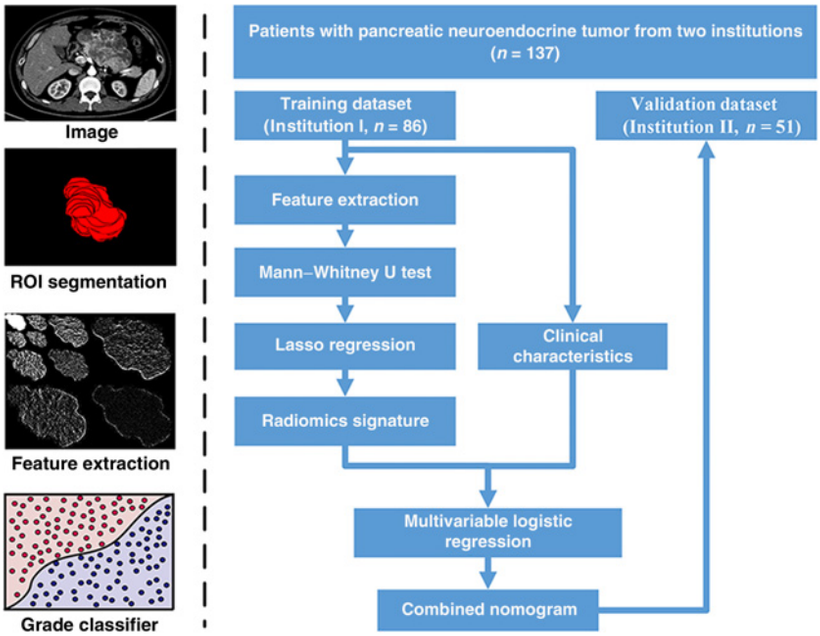

Figure 1.
Workflow in this study.

isodense pattern in the arterial phase of the contrast-enhanced CT scan. The specific patient selection pathway is shown in Supplementary Fig. S1.

Clinical characteristics [gender, age, endocrine symptoms (yes or no), multiple tumors (yes or no), maximum diameter, and clinical stage (I/IIA or IIB/III) of the tumor] were obtained through review of clinical data by one surgeon with more than 10 years of clinical experience. Two pathologists with more than 10 years of experience in the diagnosis of abdominal tumor evaluated the pathologic grade (1 to 3) according to the 2010 WHO classification system (13). The two pathologists agreed on the final pathologic grading of tumors. The clinical stage of the tumor was determined preoperatively according to the *American Joint Committee on Cancer TNM Staging System Manual, 7th edition* (14). Follow-up data for the major patients with pNETs were also obtained through clinic visit or telephone communications.

Image acquisition

All patients underwent an abdominal contrast-enhanced CT scan preoperatively. Contrast-enhanced CT scan in Institution I was performed on three CT scanners including a 16-slice CT (Toshiba Medical Systems), a 64-, and a 256-slice CT (Philips Healthcare). The contrast-enhanced CT scan in Institution II was undertaken on two CT scanners, including a 40-slice CT (Siemens AG) and a 320-slice CT (Toshiba Medical Systems). CT scans in the two institutions used the same CT scanning parameters: tube voltage of 120 kVp, tube current of 125 to 300 mAs, pitch of 0.6 to 1.25mm, slice thickness of 3 to 5 mm, and reconstruction interval of 3 to 5 mm. The nonionic contrast agent Ultravist (Bayer Schering Pharma) was bolus-injected (1.5 mL/kg) with a high-pressure syringe at 3.0 mL/s. CT scans of the arterial phase and portal vein phase were carried out at 25 to 35 seconds and 55 to 75 seconds after injection, respectively.

ROI segmentation and radiomics feature extraction

On all slices, the entire tumor was manually contoured using ITK-SNAP (http://www.itksnap.org/pmwiki/pmwiki.php; ref. 15). According to the studies reported previously, when multiple pNETs are present, the tumor with the largest diameter is chosen for analysis (12, 16). The tumor boundary was contoured by a radiologist and validated by another radiologist. The two radiologists were blinded to the final pathologic result before ROI segmentation.

To normalize different image specifications due to the utilization of different CT scanners, image resampling and gray-level normalization were performed before radiomics feature extraction from image textures (17). All image data were resampled to a $1 \times 1 \times 1$ mm voxel space size, and the gray level was normalized to 64 levels for the calculation of radiomics features.

A total of 467 radiomics features were extracted from three-dimensional ROIs using an in-house–developed software with MATLAB 2016a (MathWorks Inc.). The radiomics features extracted included 6 histogram features, 22 gray-level co-occurrence matrix (GLCM) features, 13 gray-level run-length matrix (GLRLM) features, 13 gray-level size zone matrix (GLSZM) features, 5 neighborhood gray-tone difference matrix features, and 408 wavelet-based features. Details of the procedures for extraction of radiomics feature are described in Supplementary II.

Radiomics features selection and radiomics signature building

The training dataset was used to build the pathologic grade classifier. To normalize the different scales used in variables processed, all radiomics features in the training dataset were individually subtracted by the mean value of each feature and divided by their respective SD values. The same normalization method was then applied to the validation dataset using the mean values and SD values derived from the training dataset.

To build a realistic radiomics signature with the most suitable radiomics features combined, the LASSO regression method was

used to select the most robust and nonredundant radiomics features from the features extracted (18). The complexity of LASSO regression is controlled by a tuning parameter lambda (λ) with the rule that as the value of λ increases, the penalty for each variable coefficient also increases. Only variables with nonzero coefficients were selected in this method. Details of the LASSO method are described in Supplementary III.

The binomial deviance in the logistic regression model fitting method was used as the criterion to select the best value of λ (18, 19). The iterative selection process was undertaken by conducting 10-fold cross-validation method 100 times. The λ value with the least binomial deviance was used for the final LASSO regression. A newly-assembled radiomics signature was created by summing the radiomics features selected by LASSO, multiplied with their respective coefficients. The ROC curve and area under the ROC curve (AUC) were employed to evaluate the predictive accuracy of the radiomics signature developed.

Development of a radiomics model as a pathologic grade classifier

To consider the potential influence of clinical characteristics for each patient, a multivariable logistic regression analysis was applied to integrate the developed radiomics signature with the clinical characteristics, which were significantly different between G1 group and G2/3 group [$P < 0.01$, 99% confidence interval (CI)]. The combinations of developed radiomics signature with different clinical characteristics were tested using a multivariable logistic regression method.

The backward search method with Akaike Information Criterion (AIC) score was used to select the optimal combination, which assessed the quality of developed model with comprehensive consideration of the influences of the binomial deviance and the number of variables in the selection process (20). The model with the lowest AIC score was selected as a combined radiomics model as the final tumor pathologic grade classifier. Based on the tumor pathologic grade classifier determined, a combined nomogram was also generated.

Validation of the radiomics signature and the pathologic grade classifier developed

The radiomics signature and combined nomogram model developed from the training dataset were validated on the independent validation dataset. The ROC curves and AUC values derived from the independent validation dataset were respectively generated to further evaluate the predictive accuracy of radiomics signature and nomogram model developed. The efficiency of the combined nomogram model, radiomics signature, and clinical stage alone in predicting pathologic grades in pNETs was also evaluated using both datasets. A quantitative value to represent the radiomics signature of each patient in both datasets was calculated by the radiomics signature formula developed in the training dataset.

To demonstrate the overall improvement of the radiomics model combining radiomics signature and clinical factor as compared with the clinical factors alone, a best clinical model was also constructed based on the validation dataset to assess the added value of radiomics signature to the best clinical model. Multivariable logistic regression with a backward search method was used to construct the best clinical model using the factors gender, age, endocrine symptoms, multiple tumors, maximum diameter, and clinical stage of tumor. Then, a radiomics model was constructed by combining the best clinical model and the radiomics signature in the validation dataset. A comparison of the performances in histologic grade prediction between the best clinical model and the radiomics model was assessed in terms of AUC.

Calibration curves were applied to evaluate the predictive accuracy of the combined nomogram model generated. The calibration curve is the curve with the nomogram-predicted probability of G2/3 tumors as abscissa and the actual rate acquired by bootstrapping method as ordinate. The degree of overlap between the calibration curve and the diagonal in the graph reflects the predictive accuracy of the combined nomogram model.

Decision curve analysis (DCA) was employed to evaluate the clinical utility of the combined nomogram model developed in the training dataset. The x axis of the decision curve is the threshold of the predicted probability using the combined nomogram to classify G1 patients and G2/3 patients. The y axis shows the clinical decision net benefit for patients based on the classification result in this threshold. The decision curves of the treat-all scheme and the treat-none scheme are used as references in the DCA. The definitions of net benefit, treat-all, and treat-none scheme were described in Supplementary IV. The area under the decision curve showed the clinical utility of the combined nomogram tested.

Clinical and biological significance analysis

To assess the clinical significance of radiomics features, a correlation analysis between the radiomics features included in the radiomics signature and the clinical characteristics (tumor pathologic grade, endocrine symptoms, clinical stage) was conducted using the Spearman rank correlation method. To assess the biological significance of the radiomics features, an association between radiomics features and Ki-67 index and the rate of nuclear mitosis which recognize a core antigen present in proliferating cells but absent in quiescent cells was also performed using the Spearman rank correlation method. The details of the Ki-67 index and rate of nuclear mitosis were described in Supplementary V. To assess the significance of the histologic grades predicted by the radiomics signature and the nomogram model, a correlation analysis between the probability of having G2/3 pNETs predicted by the radiomics signature and the nomogram model and Ki-67 index and rate of nuclear mitosis was also performed.

Survival analysis

Survival analysis was performed to explore the potential of the tumor pathologic grade classifier in survival prediction. Patients from the two institutions were divided into the G1 group and the G2/3 group according to the prediction results using the threshold computed from the training dataset through the Youden Index. The Kaplan–Meier (KM) method was used for the survival analysis of predicted G1 group and G2/3 group.

Statistical analysis

Differences of clinical characteristics between the training dataset and the validation dataset as well as between G1 group and G2/3 group in their respective datasets were assessed using independent sample t test, Mann–Whitney U test, or χ^2 test with a statistical significance level set at 0.01 where appropriate. The Mann–Whitney U test was applied to select radiomics features that were significantly different between G1 and G2/3 groups ($P <$

Table 1. Characteristics of patients with pNETs in the grade 1 group and grade 2/3 group

Characteristics	Training set			Validation set		
	Grade 1 ($n = 42$)	Grade 2/3 ($n = 44$)	*p*	Grade 1 ($n = 28$)	Grade 2/3 ($n = 23$)	*p*
Gender			0.286			0.080
Male	20	26		9	13	
Female	22	18		19	10	
Age (years, range)	29–81	25–78	0.646	29–82	35–79	0.784
Endocrine symptom			0.026			0.391
With	13	5		6	2	
Without	29	39		22	21	
Multiple tumor			0.529			0.466
Yes	2	1		1	3	
No	40	43		27	20	
Maximum diameter (cm, range)	2.0 (0.8–6.5)	4.1 (1–14)	<0.001[a]	2.3 (0.8–7.5)	5.7 (1.4–16.0)	<0.001[a]
Clinical stage						
I/IIA	42	29	<0.001[a]	28	14	0.001[a]
IIB/III	0	15		0	9	

[a]Significant difference threshold ($P < 0.01$).

0.01, 99% CI). The Youden Index was used to determine the best threshold in the ROC analysis. The Hosmer–Lemeshow test was carried out to examine the goodness of fit for the developed logistic regression models. The Spearman rank correlation was used in the correlation analysis, and a corresponding *P* value of 0.05 was used as a cutoff value for a significant correlation. The log-rank test was used to examine the difference between the survival curves of G1 group and G2/3 group.

All statistical analysis was performed with R 3.4.1 (www.R-project.org, 2016) and MedCalc 15.2.2 (MedCalc Inc.). The LASSO logistic regression was performed using the "glmnet" package in R. The nomogram and calibration curve were plotted using the "rms" package. DCA was performed using "dca.r" package.

Results

Patients' characteristics

Based on the criteria for patient selection, 137 patients diagnosed with pNETs between July 2010 and June 2017 from the First Affiliated Hospital, Zhejiang University School of Medicine (Institution I), and the Second Affiliated Hospital, Zhejiang University School of Medicine (Institution II), were included in this study. Eighty-six patients from Institution I were taken as the training dataset, and the other 51 patients from Institution II were used for the independent validation dataset.

The training dataset and validation dataset had an even distribution in patient characteristics (Supplementary Table S1). No significant difference was found in pNETs pathologic grade and clinical characteristics (gender, age, endocrine symptoms, multiple tumors, maximum diameter, and clinical stage of the tumor) between the training dataset and validation dataset. The detailed distribution of clinical characteristics in the G1 group and G2/3 group was summarized in Table 1. The maximum diameter and clinical stage had a significant difference between the G1 group and G2/3 group both in the training dataset and validation dataset. Fifteen patients (10.9%) were confirmed deceased in this study, and their survival time ranged from 2 months to 50 months.

Radiomics features selection and radiomics signature building

Two hundred and thirty-three features with statistical significance ($P < 0.01$) between the G1 and G2/3 groups were preliminarily selected from the 467 radiomics features in the training dataset. A radiomics signature was further constructed based on eight features with respective nonzero coefficients selected from these 233 features through LASSO regression method (Equation 1). The coefficient for each feature selected was derived from the LASSO regression method. A quantitative value to represent the radiomics signature (Equation 1) includes one run-length variance feature of GLRLM and seven wavelet-based features. Details of the procedure for construction of the radiomics signature are described in Supplementary Fig. S3.

$$\text{Radiomics signature} = 1.95970843 + 0.08039944 \times \text{RLV} + 0.32532329 \times \text{LLH\_GLRLM\_RLV} + 0.15006665 \times \text{LLH\_GLSZM\_ZP} - 0.03743996 \times \text{LHH\_GLCM\_indnc} - 0.01169231 \times \text{HLH\_GLCM\_corrm} + 0.29796698 \times \text{HLH\_GLRLM\_GLV} - 0.03900033 \times \text{HHL\_GLSZM\_SZE} - 0.19033698 \times \text{HHL\_GLSZM\_GLV} \quad (1)$$

The utility of histologic grade prediction using developed radiomics signature

The developed radiomics signature model showed a favorable result in predicting the histologic grade (G1 vs. G2/3) that produced an AUC of 0.870 in the training set (95% CI, 0.780–0.933) and 0.862 in the validation set (95% CI, 0.736–0.942), respectively. The ROC curves of radiomics signature derived from the two datasets were shown in Fig. 2A and B. To demonstrate the effectiveness of radiomics signature model at the individual scale, the quantitative values of radiomics signature for each patient regarding the classification of G1 and G2/3 groups were shown in Fig. 2C and D. With the calculated threshold using Youden Index for classification, the sensitivity in the training and validation set was 88.10% and 89.29%, respectively. This result demonstrated the high accuracy of the developed radiomics signature for the classification of G1 and G2/3 pNETs.

Combined nomogram construction

A radiomics model incorporating the developed radiomics signature with clinical stage with the lowest AIC score was chosen as the best tumor pathologic grade classifier. The results showed that radiomics signature ($P < 0.001$) and clinical stage ($P < 0.001$) were significant, independent factors in the training cohort (Supplementary VI and Supplementary Table S4). The specific process

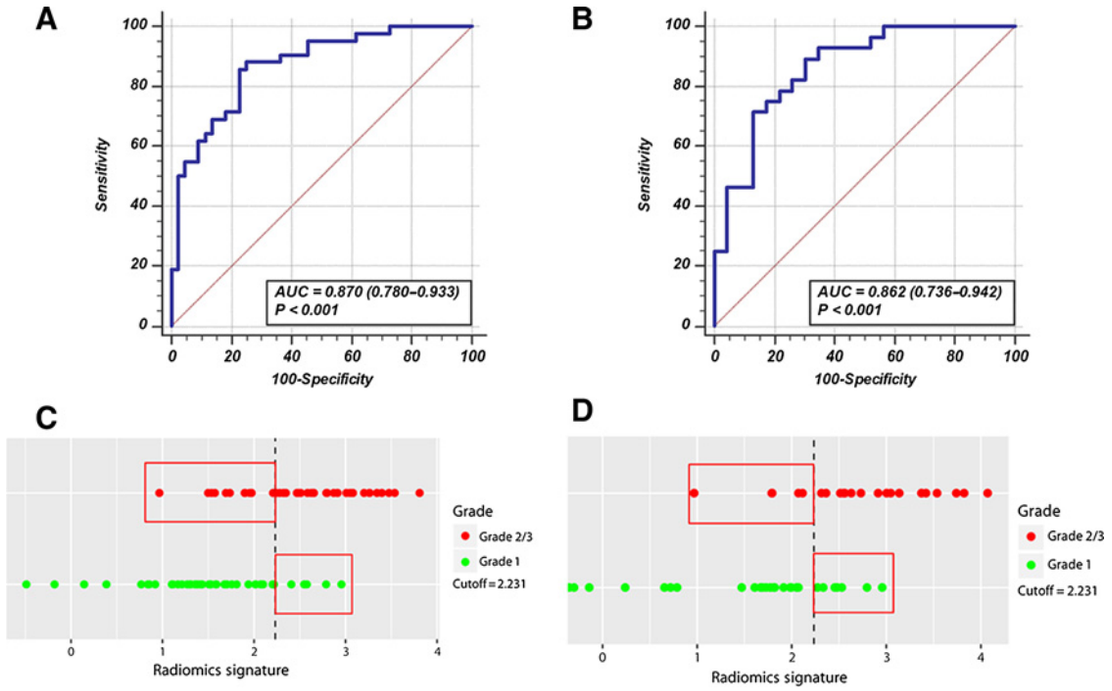

Figure 2.

The ROC curves for the radiomics signature in the training set (**A**) and the validation set (**B**). Radiomics signature distribution in the training (**C**) and validation (**D**) datasets. The red markers indicate G2/3 patients, whereas the green ones indicate G1 patients. The dotted line presents the calculated threshold (cutoff value = 2.231) by Youden Index to classify G1 and G2/3 groups. The markers in the red rectangle indicate the patients with incorrect tumor grade discrimination.

of the backward search method for selecting the best combination was described in Supplementary VI.

To visualize the multivariable logistic regression model, a combined nomogram was constructed from the radiomics model as shown in Fig. 3A. The usefulness of combined nomogram was also confirmed in the ROC analysis with an AUC of 0.906 (95% CI, 0.824–0.959) for the training set and an AUC of 0.891 (95% CI, 0.772–0.961) for the validation set (Fig. 3B and C). The AUC value revealed the high performance of tumor grade discrimination using the combined nomogram.

The results of using the combined nomogram to predict the tumor pathologic grades with the recommended threshold were shown in Fig. 3D and E. As calculated by the Youden Index, the threshold to differentiate tumor pathologic grades was 0.505 for the training dataset.

The calibration curve and the Hosmer–Lemeshow test showed a high accuracy of the nomogram for predicting tumor pathologic grades both in the training dataset ($P = 0.9513$, 95% CI) and validation dataset ($P = 0.8592$, 95% CI; Fig. 4A and B). The DCA was used to demonstrate clinical decision utility of the combined nomogram. The area under the decision curve in Fig. 4C and D showed the clinical utility of corresponding strategies. The combined nomogram (red) showed more area than that using the radiomics signature alone (blue), which were better than the "treat all" (blue) or "treat none" (black) strategies, in both the training set and validation set.

The combined nomogram showed a better performance in predicting the tumor pathologic grades (AUC = 0.894: 95% CI, 0.830–0.940) than the radiomics signature (AUC = 0.857: 95% CI, 0.787–0.911) and clinical stage (AUC = 0.679: 95% CI, 0.594–0.756) alone (Fig. 5A). Specifically, the combined nomogram showed a significant improvement than the radiomics signature alone in the Delong Test ($P = 0.0065$). The high predictive performance of the proposed nomogram model evaluated in the lumped group of both datasets was consistent with that in the separated training and validation sets.

For comparison purpose, the clinical stage and maximum diameter of the tumor were selected through backward search method to build the best clinical model. The radiomics nomogram model incorporating the best clinical model and radiomics signature showed an AUC of 0.885 (95% CI, 0.765–0.957), higher than the best clinical model alone (AUC = 0.856: 95% CI, 0.730–0.939). The radiomic signature alone performed comparably to clinical features. The improvement of the combined radiomics nomogram model over clinical features alone, when used optimally, is modest.

The correlation analysis demonstrated the association between the selected radiomics features and the tumor pathologic grades, endocrine symptoms, and clinical stages. Both the radiomics signature and nomogram were associated with Ki-67 expression level and the rate of nuclear mitosis ($P < 0.001$), suggesting a correlation of radiomics features with cell proliferation of tumors. The radiomics nomogram had a higher correlation coefficient with Ki-67 index and the rate of nuclear mitosis than the radiomics signature in the correlation analysis (Supplementary V).

The KM survival analysis (Fig. 5B) showed a significant difference between the nomogram-predicted G1 group and

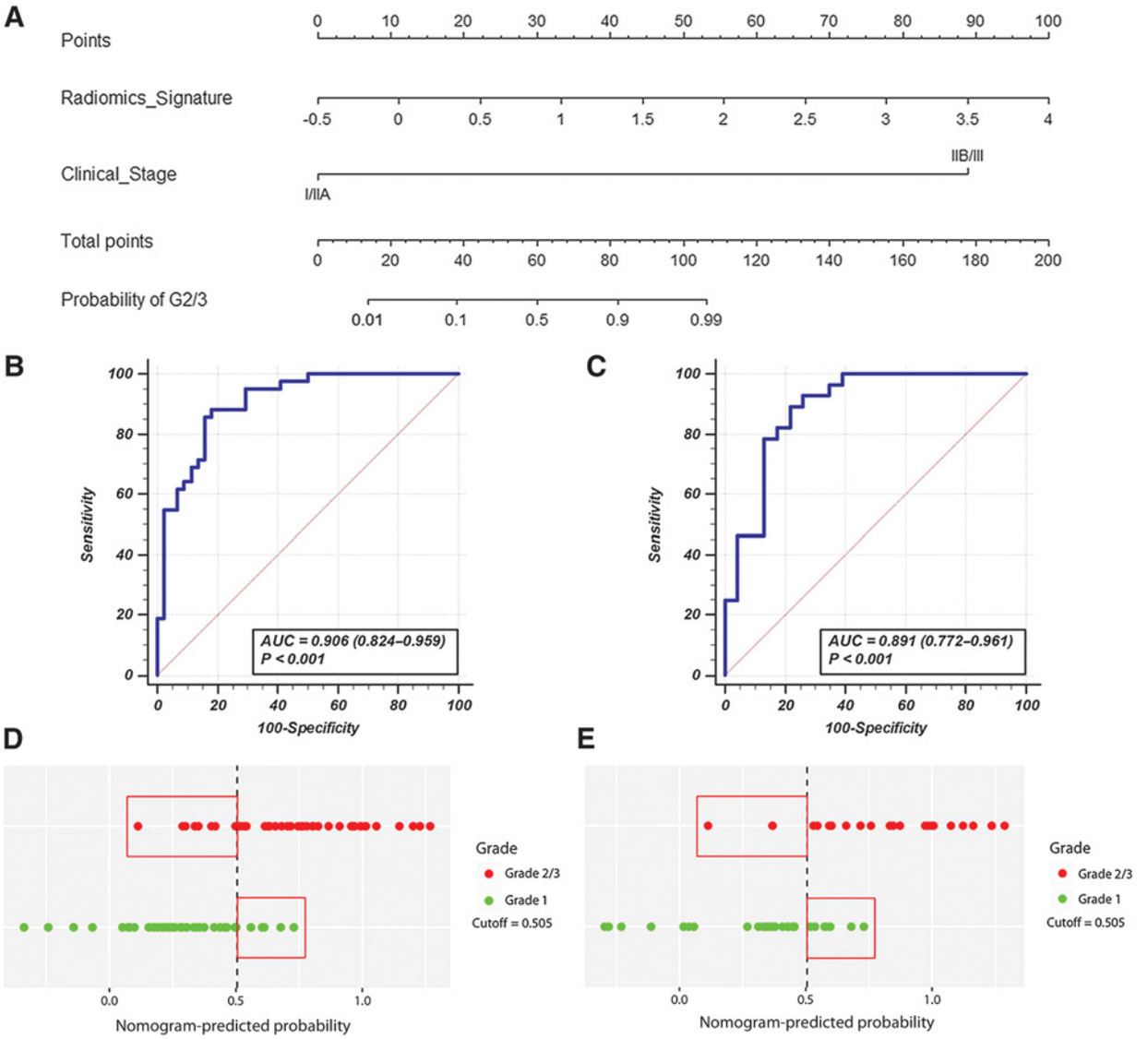

Figure 3.

A, Combined nomogram incorporating radiomics signature and clinical stage. The ROC curves for the combined nomogram in the training (**B**) and validation (**C**) set. The tumor pathologic grade prediction results in the training (**D**) and validation (**E**) sets.

nomogram-predicted G2/3 group, which suggested the prognostic value of the combined nomogram ($P = 0.0002$).

Discussion

We investigated the utility of a combined nomogram model to preoperatively predict tumor pathologic grades in patients with pNETs. An eight-feature–based radiomics signature was found to be effective for tumor grade classification. This signature could stratify patients into G1 and G2/3 groups with an AUC of 0.857. The predictive performance was further significantly improved by combining the radiomics signature with clinical stage as a combined nomogram model, achieving an AUC of 0.894. The combined nomogram developed was also validated with the independent dataset from the other institution, suggesting the reproducibility and reliability of the developed prediction model.

Previous studies suggested the proteogenomics and tumor morphology could be reflected on the medical images (21). In clinics, the tumor grade is routinely determined by Ki-67 expression which is a crucial component with intratumoral heterogeneity in the complex proteogenomics of tumors (22). Patrick and colleagues investigated the biological basis of radiomics phenotypes in lung cancer. They showed that radiomics approaches permit noninvasive assessment of both molecular and clinical characteristics of tumors in lung cancer (23). We demonstrated the association of the developed radiomics signature and nomogram with Ki-67 expression and rate of nuclear mitosis in pNETs. The results demonstrated the capability of radiomics nomogram

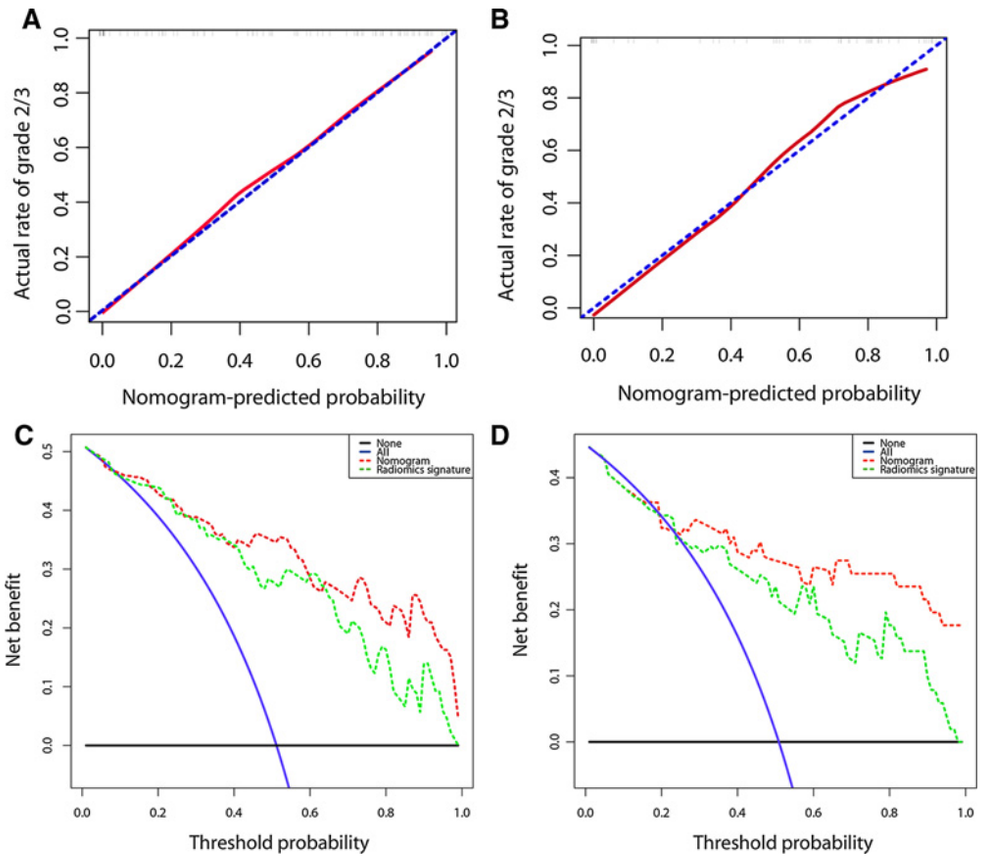

Figure 4.
The calibration curve (training set: **A**; validation set: **B**) and decision curve (training set: **C**; validation set: **D**) of the combined nomogram.

model in reflecting the underlying biological mechanisms within tumors.

In the optimization process of LASSO method for radiomics feature selection, the wavelet features had the highest weights in the radiomics signature, suggesting the vital role of wavelet-based features in the prediction model. This observation is consistent with previous studies which included wavelet-based features in the radiomics model construction (24–26). The wavelet transformation splits imaging data into different frequency components on three axis of the tumor region which may further explore the spatial heterogeneity at multiple scales within tumor regions (24).

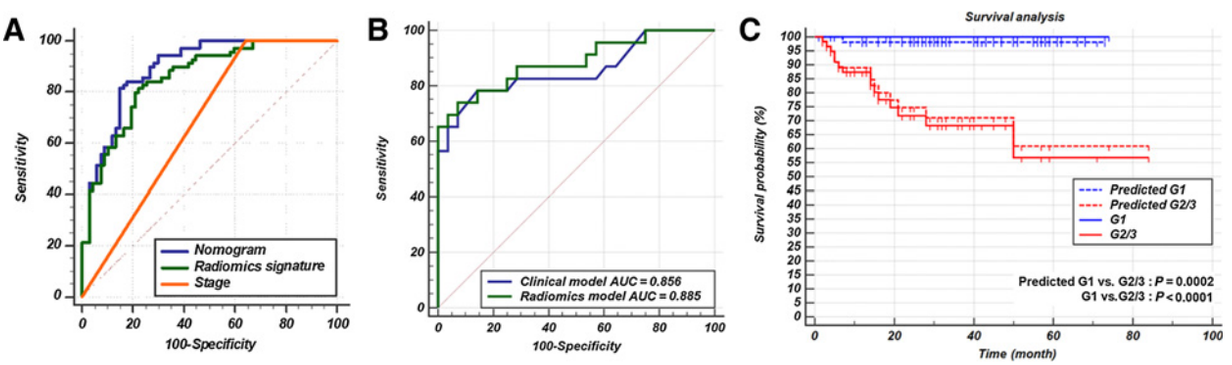

Figure 5.
A, ROC curves for the nomogram, radiomics signature, and clinical stage in both datasets. **B,** ROC curves for the best clinical model and radiomics model in the validation dataset. **C,** Survival analysis using the known grades and nomogram-predicted tumor grades. The KM analysis shows a significant difference between the predicted G1 group and G2/3 group ($P = 0.0002$).

The capability of the combined nomogram for preoperative prediction of the pathologic grade in pNETs may facilitate personalized treatment decisions (27). For functional pNETs, parenchyma-sparing pancreatic resection is a routine treatment strategy. However, the postoperative relapse rate varied within the patients with functional pNETs suggesting that parenchyma-sparing pancreatic resection might be insufficient for this subset group of patients. The previous studies indicated that a high risk of postoperative recurrence exists in patients with G2/3 tumors (28). The dilemma is that the tumor grading information is invisible preoperatively in clinical practice, thus compromising the effectiveness of surgery for pNETs patients.

The combined nomogram developed could effectively identify the more aggressive functional pNETs before operations. As such, it could stratify the patients with pNETs into G1 and G2/3 groups, in which G1 group could receive parenchyma-sparing pancreatic resection while G2/3 group might undergo comprehensive treatment strategies including radical surgical resection and systematic therapy to improve the long-term prognosis (29).

For instance, in the operative management of insulinomas, malignant insulinomas should be precisely differentiated from benign insulinomas before operation as malignant insulinomas would require relative extensive surgical resections instead of a minimally invasive surgery (28, 30). At the same time, systematic treatment is also needed for these patients with malignant insulinoma. The combined nomogram developed could provide clinicians such combined radiomics feature and clinical stage-derived grading information for clinical decision making of insulinomas so that a tailored treatment strategy could be determined preoperatively.

For nonfunctional pNETs, to the best of our knowledge, there is no complete consensus as of yet in the clinical treatment scheme of choice. Tumor diameter is a key factor in determining whether or not operative management would be used, as it is closely related to the malignant activities of pNETs (31). In clinical practice, a tumor diameter of 2 cm is conventionally used as a cutoff line to decide the patients with nonfunctional pNETs either receive a standard resection or perform a conservative management (29, 30). In addition, multiple endocrine neoplasia type 1 (MEN 1) patients with pNETs were believed to achieve little survival benefit through surgical excision (32). However, a tumor diameter of 2 cm is not a reliable and reproducible factor to make the treatment decision as a recent comparative study showed that patients with nonfunctional pNETs of diameter > 1.5 cm could also benefit from operative management (33).

The treatment strategies based on tumor diameter are not practically perfect in nonfunctional pNETs. It could be better fully considering tumor grade of pNETs in the treatment of choice for this group of patients. The combined nomogram developed would afford a reliable tool to identify pNETs' grades. Combining tumor diameter with tumor grade might be also be helpful for clinicians to determine personalized treatment strategies preoperatively. Although endoscopic ultrasound-guided fine-needle aspiration can be used for the pathologic grading of pNETs (34, 35), this method is associated with the risk of interventional procedures. The limitation of tumor location also prevents it from being widely used in clinics for preoperatively grading. The proposed combined nomogram for pNETs' pathologic grading is atraumatic, easy to use as compared with the fine-needle aspiration procedure and has the potential for preoperative evaluations with high accuracy in clinics.

Furthermore, the predictive model constructed in this study also shows an effectiveness in predicting the postoperative prognosis of patients with pNETs. The survival curves predicted by the combined nomogram model agreed well with the real survival curves derived from patients' postoperative follow-up data. There was a significant difference in the combined nomogram model–predicted survival curves between G1 patients and G2/3 patients which implied the capability of the combined nomogram as a promising prognostic biomarker.

A study on the prognosis of 3851 cases of resected pNETs confirmed that the independent risk factors for the prognosis are age, pathologic grade, metastatic state, tumor function, and resection mode, wherein the pathologic grade is a significant limiting factor for prognosis (25). A series of studies also suggested that different staging systems, including American Joint Committee on Cancer and European Neuroendocrine Tumor Society (ENETS) staging systems, have strong predictive capabilities for pNETs' prognosis which could be used to guide clinical treatment of pNETs (36–39). More recently, a modified ENETS system was proposed to deliver a better prognostic stratification for pNETs patients (1). The combined nomogram incorporating radiomics signature with clinical stage could be more useful than using the radiomics signature alone to provide prognostic information for different individuals with pNETs.

A multiple factor–based omics method is usually preferable to depict the complex heterogeneity within the pNETs' regions. Using multiple factors to predict the tumor grades could potentially pinpoint the interactions of different features related to tumor growth. However, using multiple factors to predict the tumor grades might also increase the clinical burdens to the patients while collecting different omics datasets. The developed combined nomogram incorporates radiomics features with routinely available clinical characteristics which might afford a clinically translatable paradigm easy to implement in the clinical setting.

As a retrospective study, the limitations of our study include that genomics and proteomics data cannot be incorporated into the nomogram model to classify G1 and G2/3 pNETs due to the fact that the corresponding tumor specimens were not well preserved. In addition, the validation cohort used in this study is relatively small in sample size. Moreover, as G3 group is relatively small in sample number which is not compatible for machine learning algorithms, G2/3 pNETs were not further separated into G2 and G3 patients. A future prospective study to separate G2 and G3 groups is needed to further validate the utility of nomogram model developed, thus facilitating a better personalized treatment strategies selection. Furthermore, due to the relatively short follow-up time, median overall survival for resectable pNETs was not available. We will continue to follow up with these patients to secure a more complete prognosis status.

The developed combined nomogram model using radiomics signature and tumor clinical stage can effectively predict the pathologic grade of pNETs preoperatively. The model also demonstrated a utility in predicting the postoperative prognosis of patients with pNETs. The predictive

nomogram model could serve as a preoperative, noninvasive, precise evaluation tool for patients with pNETs which may assist clinicians to tailor the treatment protocol for each individual patient and achieve a better clinical outcome in the future.

Disclosure of Potential Conflicts of Interest

No potential conflicts of interest were disclosed.

Authors' Contributions

Conception and design: W. Liang, P. Yang, Y. Kuang, T. Niu
Development of methodology: W. Liang, P. Yang, Y. Kuang, T. Niu
Acquisition of data (provided animals, acquired and managed patients, provided facilities, etc.): W. Liang, Y. Lu, Y. Kuang
Analysis and interpretation of data (e.g., statistical analysis, biostatistics, computational analysis): W. Liang, P. Yang, Q. Huang, Y. Lu, Y. Kuang, T. Niu
Writing, review, and/or revision of the manuscript: W. Liang, P. Yang, R. Huang, Y. Kuang, T. Niu
Administrative, technical, or material support (i.e., reporting or organizing data, constructing databases): W. Liang, P. Yang, R. Huang, L. Xu, J. Wang, W. Liu, L. Zhang, D. Wan, Q. Huang, Y. Kuang, T. Niu
Study supervision: W. Liang, Y. Kuang, T. Niu

Acknowledgments

This work was supported by the Zhejiang Provincial Natural Science Foundation of China (LY17H160010: W. Liang, LR16F010001: T. Niu), National High-tech R&D Program for Young Scientists by the Ministry of Science and Technology of China (2015AA020917: T. Niu), National Key Research Plan by the Ministry of Science and Technology of China (2016YFC0104507: T. Niu), Natural Science Foundation of China (81201091: T. Niu), and the NCI of the NIH under Award Number P30CA042014 (Y. Kuang).

The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked advertisement in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

Received April 27, 2018; revised September 28, 2018; accepted October 31, 2018; published first November 5, 2018.

References

1. Halfdanarson TR, Rabe KG, Rubin J, Petersen GM. Pancreatic neuroendocrine tumors (PNETs): incidence, prognosis and recent trend toward improved survival. Ann Oncol 2008;19:1727–33.
2. Yao JC, Hassan M, Phan A, Dagohoy C, Leary C, Mares JE, et al. One hundred years after "carcinoid": epidemiology of and prognostic factors for neuroendocrine tumors in 35,825 cases in the United States. J Clin Oncol 2008;26:3063–72.
3. Lawrence B, Gustafsson BI, Chan A, Svejda B, Kidd M, Modlin IM. The epidemiology of gastroenteropancreatic neuroendocrine tumors. Endocrinol Metab Clin North America 2011;40:1–18, vii.
4. Metz DC, Jensen RT. Gastrointestinal neuroendocrine tumors: pancreatic endocrine tumors. Gastroenterology 2008;135:1469–92.
5. Kim DW, Kim HJ, Kim KW, Byun JH, Song KB, Kim JH, et al. Neuroendocrine neoplasms of the pancreas at dynamic enhanced CT: comparison between grade 3 neuroendocrine carcinoma and grade 1/2 neuroendocrine tumour. Eur Radiol 2015;25:1375–83.
6. Pereira JA, Rosado E, Bali M, Metens T, Chao SL. Pancreatic neuroendocrine tumors: correlation between histogram analysis of apparent diffusion coefficient maps and tumor grade. Abdominal Imaging 2015;40:3122–8.
7. Kim JH, Eun HW, Kim YJ, Lee JM, Han JK, Choi BI. Pancreatic neuroendocrine tumour (PNET): staging accuracy of MDCT and its diagnostic performance for the differentiation of PNET with uncommon CT findings from pancreatic adenocarcinoma. Eur Radiol 2016;26:1338–47.
8. Toshima F, Inoue D, Komori T, Yoshida K, Yoneda N, Minami T, et al. Is the combination of MR and CT findings useful in determining the tumor grade of pancreatic neuroendocrine tumors? Japanese J Radiol 2017;35:242–53.
9. Guo C, Chen X, Xiao W, Wang Q, Sun K, Wang Z. Pancreatic neuroendocrine neoplasms at magnetic resonance imaging: comparison between grade 3 and grade 1/2 tumors. OncoTargets and therapy 2017;10:1465–74.
10. Lotfalizadeh E, Ronot M, Wagner M, Cros J, Couvelard A, Vullierme MP, et al. Prediction of pancreatic neuroendocrine tumour grade with MR imaging features: added value of diffusion-weighted imaging. Eur Radiol 2017;27:1748–59.
11. 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.
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.
13. Flejou JF.[WHO Classification of digestive tumors: the fourth edition]. Annales de Pathologie 2011;31:S27–31.
14. Luo G, Javed A, Strosberg JR, Jin K, Zhang Y, Liu C, et al. Modified staging classification for pancreatic neuroendocrine tumors on the basis of the American Joint Committee on Cancer and European Neuroendocrine Tumor Society Systems. J Clin Oncol 2017;35:274–80.
15. Yushkevich PA, Piven J, Hazlett HC, Smith RG, Ho S, Gee JC, et al. User-guided 3D active contour segmentation of anatomical structures: significantly improved efficiency and reliability. NeuroImage 2006;31:1116–28.
16. Xia W, Chen Y, Zhang R, Yan Z, Zhou X, Zhang B, et al. Radiogenomics of hepatocellular carcinoma: multiregion analysis-based identification of prognostic imaging biomarkers by integrating gene data-a preliminary study. Phys Med Biol 2018;63:035044.
17. Huynh E, Coroller TP, Narayan V, Agrawal V, Hou Y, Romano J, et al. CT-based radiomic analysis of stereotactic body radiation therapy patients with lung cancer. Radiother Oncol 2016;120:258–66.
18. Vasquez MM, Hu C, Roe DJ, Chen Z, Halonen M, Guerra S. Least absolute shrinkage and selection operator type methods for the identification of serum biomarkers of overweight and obesity: simulation and application. BMC Med Res Methodol 2016;16:154.
19. Kumamaru KK, Saboo SS, Aghayev A, Cai P, Quesada CG, George E, et al. CT pulmonary angiography-based scoring system to predict the prognosis of acute pulmonary embolism. J Cardiovas Comput Tomogr 2016;10:473–9.
20. Akaike H. A new look at the statistical model identification. IEEE Transact Automatic Control 1974;19:716–23.
21. Rutman AM, Kuo MD. Radiogenomics: creating a link between molecular diagnostics and diagnostic imaging. Eur J Radiol 2009;70:232–41.
22. Klimstra DS, Modlin IR, Coppola D, Lloyd RV, Suster S. The pathologic classification of neuroendocrine tumors: a review of nomenclature, grading, and staging systems. Pancreas 2010;39:707–12.
23. Grossmann P, Stringfield O, El-Hachem N, Bui MM, Rios Velazquez E, Parmar C, et al. Defining the biological basis of radiomic phenotypes in lung cancer. eLife 2017;6.
24. Wilson R, Devaraj A. Radiomics of pulmonary nodules and lung cancer. Translat Lung Cancer Res 2017;6:86–91.
25. 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.
26. Wu S, Zheng J, Li Y, Yu H, Shi S, Xie W, et al. A radiomics nomogram for the preoperative prediction of lymph node metastasis in bladder cancer. Clin Cancer Res 2017;23:6904–11.
27. Partelli S, Bartsch DK, Capdevila J, Chen J, Knigge U, Niederle B, et al. ENETS consensus guidelines for standard of care in neuroendocrine tumours: surgery for small intestinal and pancreatic neuroendocrine tumours. Neuroendocrinology 2017;105:255–65.
28. Crippa S, Zerbi A, Boninsegna L, Capitanio V, Partelli S, Balzano G, et al. Surgical management of insulinomas: short- and long-term outcomes after enucleations and pancreatic resections. Arch Surg 2012;147:261–6.
29. Valle JW, Eatock M, Clueit B, Gabriel Z, Ferdinand R, Mitchell S. A systematic review of non-surgical treatments for pancreatic neuroendocrine tumours. Cancer Treat Rev 2014;40:376–89.

30. Tamburrino D, Spoletini G, Partelli S, Muffatti F, Adamenko O, Crippa S, et al. Surgical management of neuroendocrine tumors. Best Pract Res Clin Endocrinol Metab 2016;30:93–102.
31. Bettini R, Partelli S, Boninsegna L, Capelli P, Crippa S, Pederzoli P, et al. Tumor size correlates with malignancy in nonfunctioning pancreatic endocrine tumor. Surgery 2011;150:75–82.
32. Kishi Y, Shimada K, Nara S, Esaki M, Hiraoka N, Kosuge T. Basing treatment strategy for non-functional pancreatic neuroendocrine tumors on tumor size. Ann Surg Oncol 2014;21:2882–8.
33. Zhang IY, Zhao J, Fernandez-Del Castillo C, Braun Y, Razmdjou S, Warshaw AL, et al. Operative versus nonoperative management of nonfunctioning pancreatic neuroendocrine tumors. J Gastrointest Surg 2016;20:277–83.
34. Hijioka S, Hara K, Mizuno N, Imaoka H, Bhatia V, Mekky MA, et al. Diagnostic performance and factors influencing the accuracy of EUS-FNA of pancreatic neuroendocrine neoplasms. J Gastroenterol 2016; 51:923–30.
35. Larghi A, Capurso G, Carnuccio A, Ricci R, Alfieri S, Galasso D, et al. Ki-67 grading of nonfunctioning pancreatic neuroendocrine tumors on histologic samples obtained by EUS-guided fine-needle tissue acquisition: a prospective study. Gastrointest Endos 2012;76:570–7.
36. Strosberg JR, Cheema A, Weber J, Han G, Coppola D, Kvols LK. Prognostic validity of a novel American Joint Committee on Cancer Staging Classification for pancreatic neuroendocrine tumors. J Clin Oncol 2011;29: 3044–9.
37. Ellison TA, Wolfgang CL, Shi C, Cameron JL, Murakami P, Mun LJ, et al. A single institution's 26-year experience with nonfunctional pancreatic neuroendocrine tumors: a validation of current staging systems and a new prognostic nomogram. Ann Surg 2014;259: 204–12.
38. Ekeblad S, Skogseid B, Dunder K, Oberg K, Eriksson B. Prognostic factors and survival in 324 patients with pancreatic endocrine tumor treated at a single institution. Clin Cancer Res 2008;14:7798–803.
39. Scarpa A, Mantovani W, Capelli P, Beghelli S, Boninsegna L, Bettini R, et al. Pancreatic endocrine tumors: improved TNM staging and histopathological grading permit a clinically efficient prognostic stratification of patients. Modern Pathol 2010;23:824–33.