European Journal of Radiology 117 (2019) 102–111

Contents lists available at ScienceDirect

European Journal of Radiology

journal homepage: www.elsevier.com/locate/ejrad

Research article

Differentiation of atypical non-functional pancreatic neuroendocrine tumor and pancreatic ductal adenocarcinoma using CT based radiomics

Ming He[a,1], Zhenyu Liu[b,d,1], Yusong Lin[c], Jianzhong Wan[c], Juan Li[a], Kai Xu[a], Yun Wang[a], Zhengyu Jin[a], Jie Tian[b,d,e,f,*], Huadan Xue[a,*]

[a] From the Department of Radiology, Peking Union Medical College Hospital Beijing, China
[b] From CAS Key Laboratory of Molecular Imaging, Beijing Key Laboratory of Molecular Imaging, Institute of Automation, Chinese Academy of Sciences, Beijing and China and University of Chinese Academy of Sciences, Beijing, 100049, China
[c] From the Software Technology School of Zhengzhou University, Zhengzhou, China
[d] University of Chinese Academy of Sciences, Beijing, 100080, China
[e] Beijing Advanced Innovation Center for Big Data-Based Precision Medicine, School of Medicine, Beihang University, Beijing, 100191, China
[f] Engineering Research Center of Molecular and Neuro Imaging of Ministry of Education, School of Life Science and Technology, Xidian University, Xi'an, Shaanxi, 710126, China

ARTICLE INFO

Keywords:
Pancreatic ductal carcinoma
Neuroendocrine tumor
Multidetector computed tomography
Radiomics
Differential diagnosis

ABSTRACT

Purpose: To develop and validate an effective model to differentiate NF-pNET from PDAC.

Materials and methods: Between July 2014 and December 2017, 147 patients (80 patients with PDAC and 67 patients with atypical NF-pNET) with pathology results and enhanced CT were consecutively enrolled and chronologically divided into primary and validation cohorts. Three models were built to differentiate atypical NF-pNET from PDAC, including a model based on radiomic signature alone, one based on clinicoradiological features alone and one that integrated the two. The diagnostic performance of the three models was estimated and compared with the area under the receiver operating characteristic curve (AUC) in the validation cohort. A nomogram was used to represent the model with the best performance, and the associated calibration was also assessed.

Results: In the validation cohort, the AUC for differential diagnosis was 0.884 with the integrated model, which was significantly improved over that of the model based on clinicoradiological features (AUC = 0.775, p value = 0.004) and was comparable to that of the model based on the radiomic signature (AUC = 0.873, p value = 0.512). The nomogram representing the integrated model achieved good discrimination performances in both the primary and validation cohorts, with C-indices of 0.960 and 0.884, respectively.

Conclusion: The integrated model outperformed the model based on clinicoradiological features alone and was comparable to the model based on the radiomic signature alone with respect to the differential diagnosis of atypical NF-pNET and PDAC. The nomogram achieved an optimal preoperative, noninvasive differential diagnosis between atypical pNET and PDAC, which can better inform therapeutic choice in clinical practice.

1. Introduction

Pancreatic ductal adenocarcinoma(PDAC) and pancreatic neuroendocrine tumors(pNET) are the two most frequently encountered pancreatic solid lesions [1]. pNET can be classified into functional and nonfunctional types according to their clinical symptoms [2]. Nonfunctional neuroendocrine tumors(NF-pNET) represent 68–80% of all pNET [3–5], and their reported incidence has been increasing steadily due to the development of radiological techniques during the past decades [4–6]. Approximately 60–70% of pNET typically present an imaging appearance as a well-circumscribed hypervascular solid mass on arterial phase [7–9], which can be easily differentiated from PDAC. However, pNET has various atypical characteristics, including hypo-/iso-arterial phase enhancement, unclear margin, dilatation of

Abbreviations: NF-pNET, non-functional pancreatic neuroendocrine tumor; PDAC, pancreatic ductal adenocarcinoma; LASSO, least absolute shrinkage and selection operator; SVM, support vector machine; RF, random forest; AUC, area under the receiver operating characteristic curve; CI, confidential interval

* Corresponding authors.
E-mail addresses: jie.tian@ia.ac.cn (J. Tian), bjdanna95@hotmail.com (H. Xue).
[1] Equal contributions.

https://doi.org/10.1016/j.ejrad.2019.05.024
Received 19 December 2018; Received in revised form 4 April 2019; Accepted 30 May 2019

pancreatic and/or biliary duct, and encasement of peripheral vessels, which makes it difficult to differentiate pNET from PDAC [9–11]. The therapeutic strategies and prognoses differ significantly between these two major pancreatic solid lesion subtypes, which make the correct differentiation of PDAC from pNET a major issue in clinical practice, especially for atypical cases [12,13]. For pNET, enucleation is possible, and patients with liver metastasis and with preoperative vascular abutment or invasion can still benefit from surgical resection [14–16]. For PDAC, more radical surgery is needed, which entails higher postoperative complications and risks; surgery is contraindicated for patients with liver metastasis or vascular invasion [17].

In 2017, Jeon et al. [10] indicated that a well-defined margin and hyper/isoenhancement in the portal venous phase were useful MR imaging features for discriminating nonhypervascular pNET and PDAC with a pooled sensitivity of 82% [18] and specificity of 90%, respectively. In 2016, Kim et al. [19] reported that the diagnostic performance for differentiating uncommon pNET from adenocarcinoma was 0.760-0.806. Both of these studies included functional and nonfunctional pNET. However, NF-pNET usually shows no symptoms of endocrinal malfunction; thus, for these cases, it was more challenging to differentiate between atypical NF-pNET and PDAC with conventional images.

Radiomics is a new image postprocessing technology that creates the potential for more detailed descriptions of tumor characteristics and has shown promising results in terms of differential diagnosis [20]. *In addition, the nomogram is a tool to intergrated several factors together and present the model in user-friendly digital interfaces, which aid the clinical decision making. It has been widely used in clinical practice providing more optimal identification characteristics and facilitating evidence-based and personalized decision-making* [21–24].

In the present study, our hypothesis was that a radiomics-based model represented with a nomogram that integrated clinicoradiological features and the radiomic signature would improve the differential diagnostic performance between atypical NF-pNET and PDAC, which is difficult to achieve in clinical practice. Therefore, we aimed to develop and validate an effective model and represent it with a nomogram to differentiate NF-pNET from PDAC.

2. Materials and methods

This retrospective study was approved by our local institutional review board, and the need for informed consent was waived.

2.1. Patient recruitment

Considering the rarity of the pNET and the consistent of the imaging protocol and scanner used for contrast enhance CT of pancreas in our hospital, the enrollment period is from October 2014 to December 2017 for the atypical NF-pNET cohort and from January 2016 to December 2017 for PDAC cohort. All patients pathologically proven as NF-pNET or PDAC were identified and consecutively recruited through our hospital histopathological database. The inclusion criteria were as follows: (a) patients went preoperative enhanced CT within 2 months before surgery or biopsy. (b) The CT images of the patients did not have severe motion artifacts that affect the assessment of the lesion or no metal artifacts after implantation of the stent. This process was evaluated by one of the radiologist(M.H.). (c) patients who did not undergo local or systemic treatment before surgery or biopsy. For NF-pNET, we excluded patients who did not have at least three of the following findings: ill-defined margin, iso- or hypoenhancement on arterial phase, pancreatic atrophy, pancreatic or bile duct dilatation, peripheral organ involvement, or vessel involvement and metastasis. The results of lab examinations, including CA 19-9, CA 12-5 and CEA, were also collected. Recruited patients were divided chronologically into primary and validation cohorts at a ratio of 2:1. The prevalence and percentage of each clinicoradiological feature were compared to figure out whether the difference between the primary and validation cohorts. Fig. 1 shows the flowchart of the patients' recruitment.

2.2. CT protocol

All scans were performed using a 128-detector CT scanner (Siemens SOMATOM Definition Flash, Siemens Healthcare, Forchheim, Germany). *The tube voltage was 120 kVp, and automatic tube current modulation was used.* The gantry rotation time was 0.5 s, and the table increment was 46.8 mm per rotation. The matrix was 512 × 512. Both slice thickness and intervals were 5.0 mm. Nonionic contrast media (Ultravist, 370 mgI/ml, Schering, Berlin, Germany) were injected with 1.5 mL/kg at a rate of 3.0 mL/s. An automatic power injector was used, and bolus tracking was applied. There is a 5s-delay of scan after the aortic enhancement of 100 HU. *A pancreatic arterial phase (PAP) scan was performed approximately 15 s after the aortic enhancement of 100 HU,* and a portal venous phase (PVP) scan was started 30 s after the PAP phase imaging acquisition.

The workflow of the construction and validation of the three models was presented in Fig. 2

3. Construction of the three models

3.1. The model based on radiomics signature alone

3.1.1. ROI segmentation

The ROI segmentation was done by one junior radiologist(M.H.) and under anther senior radiologist's (H.D.X.) supervisor (H.D.X. and M.H. had 15 years and 5 years of experience in interpreting abdominal CT respectively). Regions of interest (ROIs) were created manually via ITK-SNAP software (www.itksnap.org) on the axed PAP imaging and all the slices can see the lesion were segmented to included the whole tumor while avoiding blood vessels and surrounding adipose tissue. The number of the slices of ROIs range from 3 to 20.

3.1.2. Radiomics features extraction and selection

In calculation of the tumor's radiomics signatures, each feature is calculated layer by layer, and the characteristics of each layer was added and averaged to present the tumor's radiomics signatures. Four groups of imaging features were extracted from the pancreatic artery phase data in the primary cohort, including (i) 8 shape features, (ii) 17 first-order statistical features, (iii) 54 texture features and (iv) 568 wavelet features (see Supplement). Finally, seven radiomic features (supplement Table 1) were selected to build to the radiomics signature model.

3.1.3. Construction of the radiomics signature model

Based on the selected 7 radiomics features, three methods including LASSO, support vector machine (SVM), and random forest (RF) were used to built the radiomics signature model. *The Net Reclassification Improvement (NRI) test* [25–27] *was applied to select the final radiomic signature model ((NRI test was used to compared the diagnosis performance of two diagnostic tests).* The NRI test showed that the differential diagnostic performance of the three radiomic signature models was comparable ($p > 0.05$) in the validation cohort (see supplement Table 2). Ultimately, the LASSO-based model was chosen as the radiomic model.

3.2. The model based on clinic-radiological features

The evaluation of CT findings was performed by two radiologists (H.D.X. and M.H.) who were blinded to the patients' surgical and pathological information. *Any discrepancy was resolved through consensus by discussion and the results in consensus were used to built the clinic-radiological model.* The evaluation of the CT findings including size, location, and morphology (regular or irregular) on all available imaging. Regular was defined as a round or oval tumor shape, but the tumor margin did not have to be clear; otherwise, the tumor

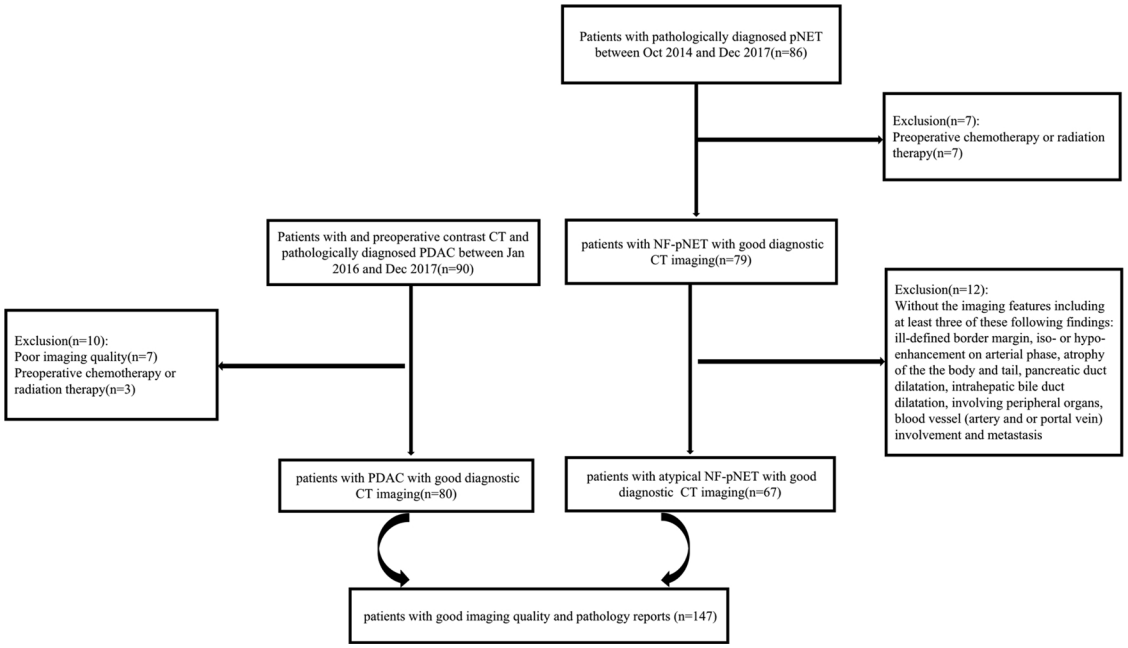

Fig. 1. The flow chart of the study population.

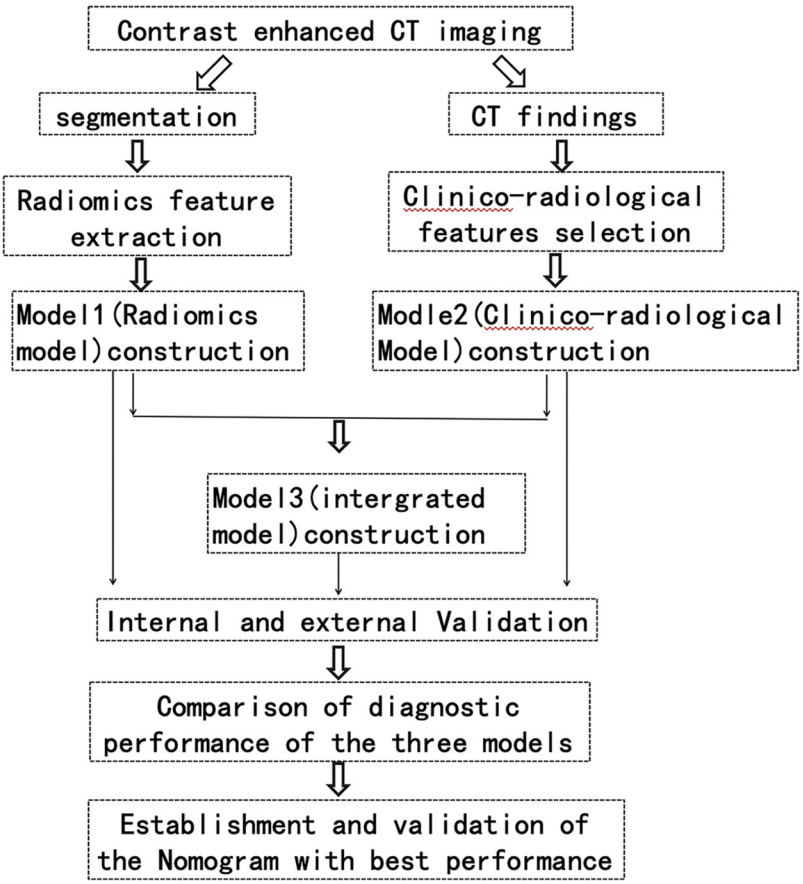

Fig. 2. The workflow.

morphology was considered irregular. Additional features assessed were margin, protruding the pancreas or not, texture (solid or cystic), calcification, hyperenhancement on arterial phase or not (hyperenhancement was defined as the enhancement of a tumor over that of background pancreas) [28], vessel encasement (defined as an attachment of the lesion to vessel of greater than 180°) [29], lymph node enlargement (short axis longer than 1 cm) [28], liver metastasis, pancreatic parenchymal atrophy, dilatation of the main pancreatic duct [30] (defined as an MPD diameter greater than 2 mm), dilation of the bile duct (BD) (defined as a BD diameter greater than 10 mm). Then the independent clinic-radiological features were selected to built the model based on clinic-radiological features alone.

Table 1
Demographic Characteristics and lab results comparison between primary and validation cohort.

characteristics	primary cohort(n = 100)	validation cohort (n = 47)	p value	kappa value
age(y)*	58(52, 65)	57(52, 64)	0.806	
sex†			0.598	
Female	50(50)	21(45)		
Male	50(50)	26(55)		
pathology type†			0.861	
PDAC	55(55)	25(53)		
NF-pNET	45(45)	22(47)		
location†			0.114	
Head	50(50)	31(66)		
Body	22(22)	4(9)		
Tail	23(23)	8(17)		
Multiple	5(5)	4(9)		
Size(mm)*	3.7(2.9, 5.2)	3.7(2.7, 4.5)	0.865	
CA 19-9#	45/98(46)	28/46(61)	0.109	
CA 12-5#	16/75(21)	3/33(9)	0.172	
CEA#	23/98(24)	14/47(30)	0.423	
Morphology†			0.854	0.758
Regular	37(37)	16(34)		
Irregular	63(63)	31(66)		
Intensity†			0.216	0.848
Homogeneous	38(38)	23(49)		
Heterogeneous	62(62)	24(51)		
Texture†			0.855	0.785
solid	64(64)	31(66)		
cystic	36(36)	16(34)		
Margin†			0.681	0.842
well defined	26(26)	10(21)		
ill defined	74(74)	37(79)		
confined to the pancreas†			0.715	0.814
no	62(62)	31(66)		
yes	38(38)	16(34)		
calcification†	14(14)	3(6)	0.269	0.937
arteries invasion†	47(47)	13(28)	0.026	0.859
portal vein invasion†	57(57)	25(53)	0.665	0.863
lymph node†	41(41)	27(57)	0.077	0.877
liver metastasis†	19(19)	6(13)	0.481	0.86
parenchymal atrophy†	37(37)	11(23)	0.132	0.86
MPD dilatation†	53(53)	22(47)	0.64	0.838
BD dilation†	27(27)	15(32)	0.113	0.901
non hyper enhanced on PAP†	45(45)	24(51)	0.595	0.892

*Data are median with interquartile range in the parentheses. *P* value was calculated with Mann-Whitney U test.
† Data are number of patients, with the percentage in parentheses. *P* value was calculated with the x2 or Fisher exact-test.
#Data are presented as number of the patient with elevated results/all the patients who had done the examination, with percentage in the parentheses.
Laboratory analysis of CEA CA 19-9 and CA 12-5 was done via routine blood tests within 1 week before surgery. The threshold value for elevated for CEA, CA 19-9 and CA 12-5 was > 5 ng/mL, > 39 U/mL and > 35 U/mL respectively.
PDAC: pancreatic ductal adenocarcinoma, NF-pNET: non-functional pancreatic neuroendocrine tumor, PAP: pancreatic artery phase, CEA: carcinoembryonic antigen, CA 19-9: carbohydrate antigen 19-9, CA 125: carbohydrate antigen 125.

3.3. The integrated model

Among the useful radiomics signature and clinic-radiological features, the independent predictors for atypical NF-pNETs were selected and then used to built the intergrated model.

3.4. Validation of the three model

All three models were internally validated by leave-one-out cross-validation in the primary cohort and then externally validated in the independent validation cohort.

3.5. Diagnostic Comparison of the three models

The diagnostic performance was evaluated and compared the three models, including the radiomic signature, clinico-radiological features and the integrated model.

3.6. Establishment and validation of the nomogram

To provide clinicians with a quantitative tool to predict the individual probability of atypical NF-pNET, we present the model with best diagnostic performance in the nomogram. Calibration curves were plotted to assess the calibration of the nomogram, accompanied by the Hosmer-Lemeshow test. The C-index and calibration curve were derived based on regression analysis in both the primary and validation cohorts.

3.7. Statistic analysis

The radiomics signature extraction and building were implemented on MatlabR2014b (Math Works Inc.). While the built of clinic-radiological model and the intergrated model and all the statistic analysis were conducted with R software (version 3.0.1; http://www.Rproject.org). Two-sided P less than 0.05 was considered as significant difference. The prevalence and percentage of each clinic-radiological feature were compared by a $\chi 2$ test or the Fisher exact-test for categorical variables and a Mann-Whitney U test for continuous variables. A two-sample *t*-test and the least absolute shrinkage and selection operator (LASSO) [30] were applied to select the radiomic features. Univariate and multivariate logistic regression models were used to select the independent clinic-radiological features. Multivariable logistic regression analysis (stepwise) was applied to selected independent predictors for the integrated model and using the likelihood ratio test with Akaike's information criterion [31] as a stopping rule. Receiver operating characteristics curve (ROC) analysis was performed to ascertain the diagnostic performance. The NRI test was performed to compare performance among models in the validation cohort. Inter-observer agreement for the findings was evaluated by weighted kappa statistics.

4. Results

4.1. Clinical characteristics and CT findings

Ultimately, 147 patients including 67 patients with atypical NF-pNET (mean age, 53.4 years; range, 25–79 years) and 80 patients with PDAC (mean age, 59.9 years; range, 35–79 years) were enrolled in this study (Flow chart in Fig. 1). In total, 100 patients including 45 atypical NF-pNET and 55 PDAC comprised the primary cohort, and 47 patients including 22 atypical NF-pNET and 25 PDAC comprised the validation cohort. The proportion of PDAC in the primary cohort and the validation cohort was 55% and 53%, respectively, and there was no significant difference between them ($P > 0.05$). There were no significant differences in clinical characteristics or CT findings between the primary and validation cohorts except with respect to artery invasion (Table 1).

4.2. The clinico-radiological model

Kappa values corresponding to the interobserver agreement of CT findings for atypical NF-pNET and PDAC between the two radiologists were all greater than 0.75. The kappa values ranged from 0.758 to 0.937 (Table 1).

A few clinicoradiological features were significantly different between the atypical NF-pNET and PDAC groups in both the primary and

Table 2
Demographic Characteristics and lab results of patients with atypical NF-pNET and PDAC.

characteristics	primary cohort(n = 100) atypical NF-pNET (n = 45)	PDAC (n = 55)	p value	validation cohort(n = 47) atypical NF-pNET (n = 22)	PDAC (n = 25)	p value
Age(y)*	52(48-63)	61(55-67)	0.001	57(49-61)	61(53-65)	0.501
Sex†			0.228			0.202
Female	26(58)	31(56)		12(55)	9(36)	
Male	19(42)	24(44)		10(45)	16(64)	
location†			0.01			0.162
Head	22(49)	28(51)		13(59)	18(72)	
Body	6(13)	16(29)		2(9)	2(8)	
Tail	12(27)	11(20)		3(14)	5(20)	
Multiple	5(11)	0		4(18)	0	
Size(mm)*	46(29-62)	36(29-41)	0.012	43(35-57)	32(26-43)	0.029
CA 19-9#	10/43(23)	35/55(64)	0	6/21(29)	22/25(88)	0
CA 12-5#	5/37(14)	11/38(29)	0.158	2/19(11)	1/14(7)	1
CEA#	6/37(14)	17/55(31)	0.058	8/22(36)	6/25(24)	0.355
Morphology			0			0.03
Regular	28(62)	9(16)		11(50)	5(20)	
Irregular	17(38)	46(84)		11(50)	20(80)	
Intensity			0.001			0.028
Homogeneous	9(20)	29(53)		7(32)	16(64)	
Heterogeneous	36(80)	26(47)		15(68)	9(36)	
Texture			0.112			0.121
solid	25(56)	39(71)		12(55)	19(76)	
cystic	20(44)	16(29)		10(45)	6(24)	
Margin			0.13			0
well defined	15(33)	11(20)		10(45)	0	
ill defined	30(67)	44(80)		12(55)	25(100)	
confined to the pancreas						0.351
no	40(89)	22(40)	0	13(59)	18(72)	
yes	5(11)	33(60)		9(41)	7(28)	
calcification	11(24)	3(6)	0.006	2(9)	1(4)	0.593
arteries invasion	18(40)	29(53)	0.205	4(18)	9(36)	0.173
portal vein invasion	24(53)	23(60)	0.503	12(55)	13(52)	0.861
lymph node	17(38)	24(44)	0.553	7(32)	20(80)	0.001
liver metastasis	14(31)	5(9)	0.005	5(23)	1(4)	0.085
parenchymal atrophy	13(29)	24(44)	0.129	3(14)	8(32)	0.138
MPD dilatation	20(44)	33(60)	0.111	5(23)	17(68)	0.002
BD dilation	8(18)	19(35)	0.073	3(14)	10(40)	0.021
non hyper enhanced on PAP	16(24)	29(53)	0.107	8(36)	16(67)	0.059

*Data are mean (standard deviation) or median(interquartile range). *P* value was calculated with student *t*-test or nonparametric test.
† Data are number of patients, with the percentage in parentheses. *P* value was calculated with the x2 or Fisher exact-test.
#Data are number of the patient with elevated results/all the patients who had done the examination, with percentage in the parentheses.
Laboratory analysis of CEA CA 19-9 and CA 12-5 was done via routine blood tests within 1 week before surgery. The threshold value for elevated for CEA, CA 19-9 and CA 12-5 was > 5 ng/mL, > 39 U/mL and > 35 U/mL respectively.
PDAC: pancreatic ductal adenocarcinoma, NF-pNET: non-functional pancreatic neuroendocrine tumor, PAP: pancreatic artery phase, CEA: carcinoembryonic antigen, CA 19-9: carbohydrate antigen 19-9, CA 12-5: carbohydrate antigen 12-5.

validation cohorts, including size, CA 19-9 level, morphology and intensity (Table 2). Five features including low CA 19-9 level, regular morphology, protruding pancreas, calcification and liver metastasis were found to be independent predictors ($p < 0.05$) for atypical NF-pNET by multivariate logistic regression analysis in the primary cohort. When we applied these predictors in the multivariable logistic regression model, an AUC of 0.937 (95% CI: 0.913, 0.959) and a classification accuracy of 86.7% (95% CI: 83.4%, 90.2%) were achieved in the primary cohort, and an AUC of 0.775 (95% CI: 0.702, 0.848) and accuracy of 60.9% (95% CI: 53.4%, 67.8%) were achieved in the validation group with respect to the differential diagnosis between atypical NF-pNET and PDAC. Supplemental Figures A.2 and Figure A.3 show two cases of atypical NF-pNET and PDAC.

4.3. The radiomic signature-based model

According to the diagnostic performance, the LASSO-based radiomic signature yielded an AUC of 0.890 (95% CI: 0.849, 0.919) and an accuracy of 82.0% (95% CI: 78.0%, 84.9%) in the primary cohort and an AUC of 0.873 (95% CI, 0.796, 0.884) and an accuracy of 76.6% (95% CI: 71.7%, 79.2%) in the validation cohort (Fig. 3 and Table 3).

4.4. The integrated model

All five independent clinicoradiological features and LASSO-based radiomic signatures were screened and sorted by β coefficient values (absolute value) in the primary cohort. Finally, the four independent variables including CA 19-9 level, regular morphology, protruding the pancreas and LASSO-based radiomic signatures with 7 features were selected to build the integrated model (Table 4). The integrated model yielded an AUC of 0.960 (95% CI: 0.942, 0.979) and accuracy of 89.8% (95% CI: 86.9%, 92.5%) in the primary cohort and an AUC of 0.884 (95% CI, 0.0.831, 0.927) and accuracy of 80.4% (95% CI: 74.1%, 86.5%) in the validation cohort (Table 3). The integrated model achieved an NPV of 92.7% (95% CI, 89.3%, and 96.1%) in the primary cohort and of 84.0% (95% CI, 76.3%, and 91.1%) in the validation cohort.

4.5. Comparison of diagnostic performance

In the primary cohort, the integrated model achieved considerably better discrimination ability than the radiomic signature model (NRI test, $p < 0.001$). The differential diagnostic performance of the

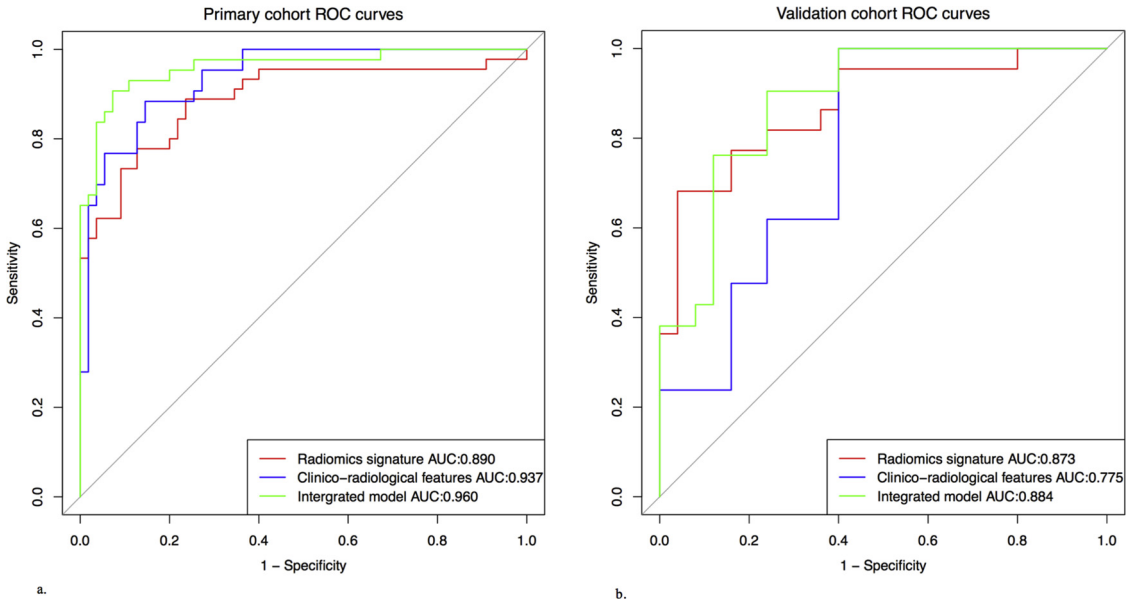

Fig. 3. Graphs show receiver operation characteristic curve analysis of diagnostic performance of the LASSO-based radiomics signature (red line), the clinico-radiological features (blue line) and the integrated model (green line) in the primary cohort(a) and validation cohort (b)for atypical NF-pNET.
AUC: area under the curve, NF-pNET: non-functional pancreatic neuroendocrine tumor.

clinicoradiological features model was comparable to that of the radiomic signature model and the integrated model (NRI test, p value was 0.405 and 0.262, respectively).

In the validation cohort, both the radiomic signature and the integrated model achieved better discrimination ability than the clinicoradiological features model (NRI test, p value was 0.024 and 0.004, respectively), and the diagnostic performance of the radiomic signature was comparable to that of the integrated model (NRI test, p value was 0.512).

4.6. Establishment and validation of the nomogram

The nomogram was used to represent the integrated model for discriminating atypical NF-pNET and PDAC (Fig. 4). The Hosmer-Lemeshow test showed no statistical significance (P = 0 0.985 and 0.988 in the primary and validation cohorts, respectively). The model displays a C-index of 0.960 (95% CI: 0.942, 0.979) and of 0.884 (95% CI: 0.831, 0.927) in the primary and validation cohorts, respectively, presenting sufficient accuracy to differentiate atypical NF-pNET from PDAC. The optimal calibration curves also revealed agreement between the nomogram outcome and pathologic diagnosis in the primary and validation cohorts (Fig. 3). Fig. 5 presented two cases pathologically diagnosed as atypical NF-pNET and PDAC respectively and the probability of atypical NF-pNET predicted by the three models (Fig. 5).

Table 4
Useful features for differential diagnosis atypical NF-pNET from PDAC in the integrated model.

variable	β	Odds Ratio (95% CI)	P
radiomic features	7.097	19.692(6.372,60.858)	< 0.001
CA 19-9 level	1.754	5.775(2.358,14.144)	< 0.001
regular morphology	2.155	8.625(3.353,22.187)	< 0.001
protruding the pancreas	2.683	14.625(4.576,46.739)	< 0.001

5. Discussion

In this study, to find an effective model to differentiate NF-pNET from PDAC, we developed and validated three models including one based on clinico-radiological features alone, one based on the radiomic signature alone and one which integrated both of them. In an external validation cohort, the diagnostic performance of the integrated model (AUC = 0.884) was significantly better than that of the model based on clinico-radiological features alone (AUC = 0.775) (p value less than 0.05) and was better than that of the model based on the radiomic signature alone (AUC = 0.873) but not reach statistically difference (p value = 0.512). *Considering the facts that in clinical practice, the clinical information like the CA 19-9 level and the traditional imaging features still play an important role in differential diagnosis between PDAC and NF-pNET and the intergrated model was more robust than the former two, we finally*

Table 3
Diagnosis performance of the three models for atypical NF-pNET.

variable	Clinico-radiological features		radiomics signature		integrated model	
	primary cohort	validation cohort	primary cohort	validation cohort	primary cohort	validation cohort
AUC	0.937(0.913,0.959)	0.775(0.702,0.848)	0.890(0.849,0.919)	0.873(0.796,0.884)	0.960(0.942,0.979)	0.884(0.831,0.927)
Accuracy (%)	86.7(85/98)[83.4, 90.2]	60.9(28/46)[53.4,67.8]	82.0(82/100)[78.0,84.9]	76.6(36/47)[71.7,79.2]	89.8(88/98)[86.9,92.5]	80.4(37/46)[74.1,86.5]
Sensitivity (%)	91.7(33/36)[87.4,96.3]	63.6(7/11)[48.8,78.2]	86.5(32/37)[81.0,91.3]	92.3(12/13)[82.6,95.3]	90.2(37/41)[85.5.94.7]	80.0(16/20)[71.5,89.2]
Specificity (%)	83.9(52/62)[79.1,88.6]	60.0(21/35)[52.6,68.3]	79.4(50/63)[74.5,82.9]	70.6(24/34)[66.2,75.2]	89.5(51/57)[85.6,94.7]	80.8(21/26)[73.7,88.6]
PPV (%)	76.7(33/43)[70.5,83.4]	33.3(7/21)[25.0,56.0]	71.1(32/45)[64.0,75.8]	54.5(12/22)[44.5,58.7]	86.1(37/43)[80.8,91.1]	76.2(16/21)[66.8,86.0]
NPV (%)	94.6(52/55)[91.7,97.8]	84.0(21/25)[76.4,91.0]	90.9(50/55)[87.1,94.5]	96.0(24/25)[91.9,97.7]	92.7(51/55)[89.3,96.1]	84.0(21/25)[76.3,91.1]

NOTE. Unless otherwise indicated, data are percentages, with 95% CIs in parentheses.
AUC: area under the curve, PDAC: pancreatic duct adenocarcinoma, PPV: positive predict value, NPV: negative predict value.

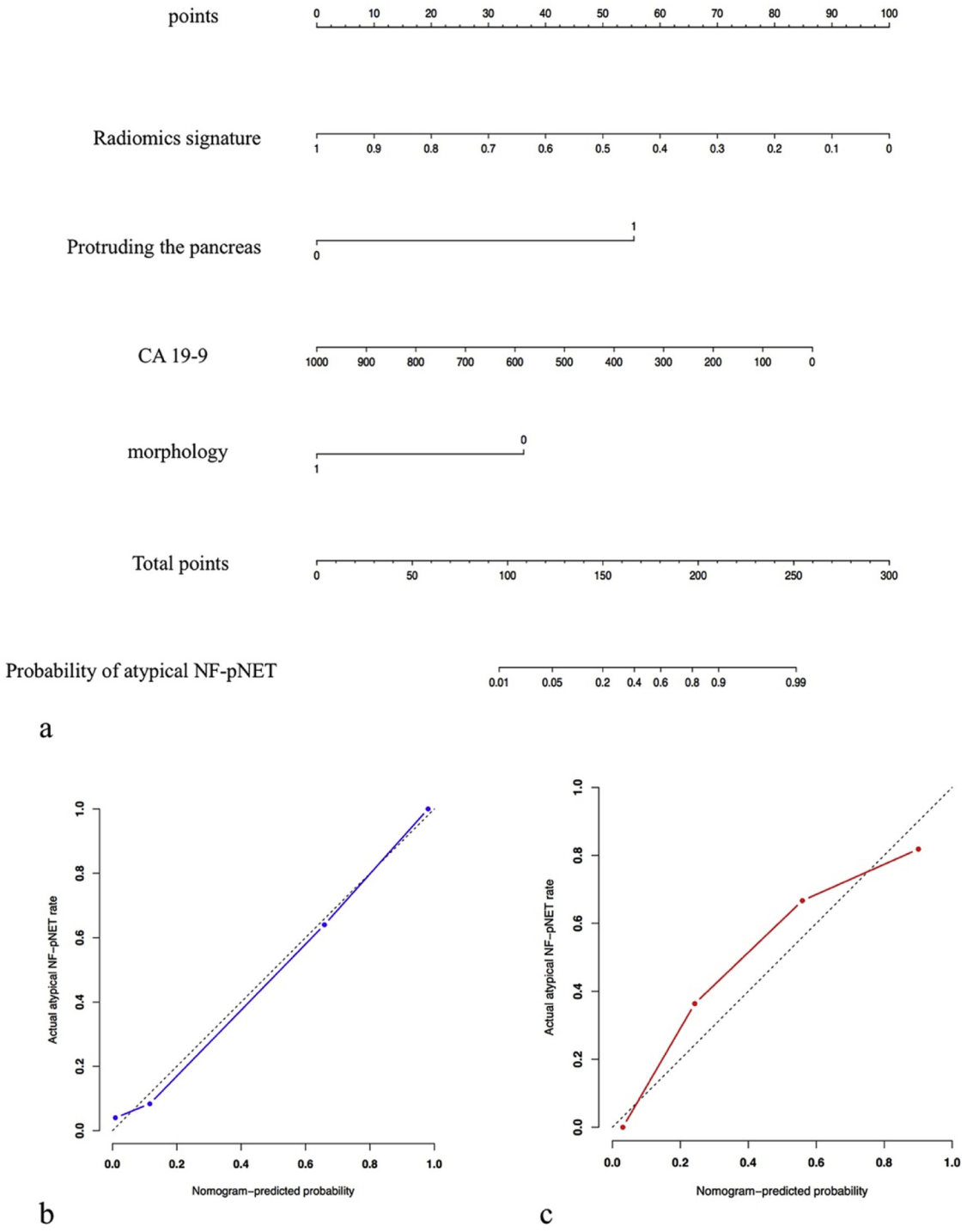

Fig. 4. The nomogram (a) and its calibration curves in primary cohort(b) and validation cohort (c). To use the nomogram, match each patient's results for each parameter to a position on their corresponding axis, then draw a line to the axis of each point at the top of the nomogram to calculate the respective points for each parameter; finally, add the total points from all parameters, and draw a line from the Total Points axis to the Risk Probability axis at the bottom of the nomogram to determine the probability of atypical pNET. Calibration curves depict the calibration of each model in terms of the agreement between the predicted pathology of the pancreatic solid lesion and actual pathology of the pancreatic solid lesion. The y axis represents the actual pathology of the pancreatic solid lesion. The x axis represents the predicted probability of pathology type. The diagonal black dotted line represents a perfect prediction by an ideal model. The blue solid line in (b) represents the performance of the nomogram in the primary cohort and the red solid line in (c) represents the performance of the validation cohort.

selected the intergrated model and represented it with a nomogram. The nomogram achieved optimal preoperative, noninvasive differential diagnosis between atypical pNET and PDAC. This information can help facilitate informed therapeutic choices in clinical practice.

Before the three models were built, all patients were divided into primary and validation cohorts, among which there were no significant differences in clinical characteristics or CT findings, except for artery invasion. As we have proved that there was no difference in the composition of the primary and validation cohorts, then we can draw the conclusion that the difference in primary and validation cohorts was not due to the heterogeneity of the patients.

In selecting the clinicoradiological features, the following were

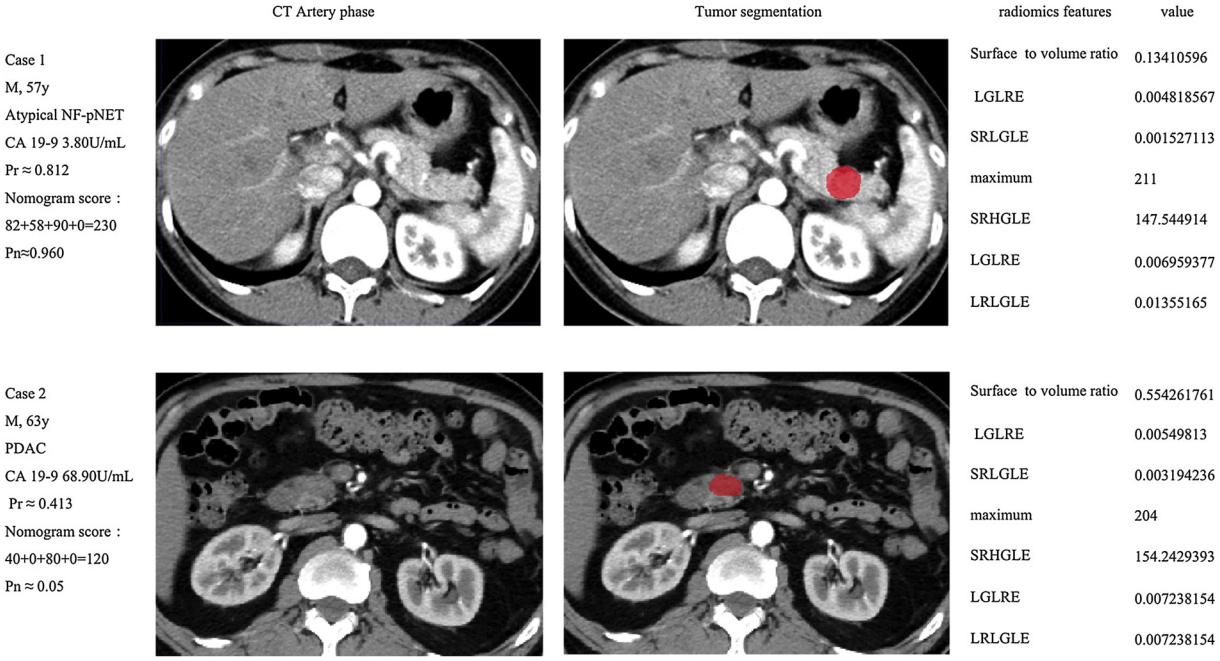

Fig. 5. two cases pathologically diagnosed as atypical NF-pNET and PDAC respectively and their CT imaging and seven selected radiomics features. All the radiomics features are different between atypical NF-pNET and PDAC. Both the NF-pNET and the PDAC were iso-intensity which means the CT value difference between the lesion and the normal pancreatic parenchyma was less than 15Hu.
Note: Pr means the probability of atypical NF-pNET predicted by the model based on radiomics signature. Pn means the probability pf atypical NF-pNET predicted by the radiomics-based nomogram. Nomogram score is valued by the following predictors included in the nomogram: radiomics signature, protruding the pancreas, CA 199 level, and morphology.

useful clinical and imaging features for differential diagnosis: low CA 19-9 level, regular morphology, lack of confinement to the pancreas and the presence of calcification and metastasis; additionally, the pooled sensitivity and specificity were 91.7% and 83.9% in the primary cohort, which are comparable to previous findings [10,12,19]. In the validation cohort, the pooled sensitivity and specificity were 63.6% and 60.0%, respectively, which are lower than those previously found (pooled sensitivity and specificity were 90.0% and 89.0%, respectively). This finding may be explained by numerous reasons. First, the patients enrolled in our study were nonfunctional pNET. Unlike F-pNET patients, their imaging features, like a kaleidoscope, overlap with those of PDAC patients; thus, it was more challenging to differentiate NF-pNET from PDAC [32]. Second, we excluded 12 typical cases and included only NF-pNET case with malignant CT findings, thereby increasing the difficulty of differential diagnosis.

For the construction of the radiomic signature, 7 potential predictors were finally selected from 647 candidate radiomic features with the LASSO method. This method has two advantages; it surpasses the method of choosing predictors based on the strength of their univariable association with outcome. In addition, this method enables the panel of selected features to be combined into a radiomic signature [33,34]. For the modeling of radiomic signatures, we used three different methods, including LASSO, SVM and RF. All of them showed good to excellent differential diagnostic performance, which was comparable in both the primary and validation cohorts. This finding indicates that radiomics is a reliable tool for differentiating atypical NF-pNET and PDAC.

Notably, both the differential diagnostic accuracies of the clinicoradiological features model alone and the radiomic signature model alone were decreased dramatically in the validation cohort. However, there was no difference in the composition of the primary and validation cohorts as we stated before, suggesting that the difference in primary and validation cohorts was not due to the heterogeneity of the patients. *One of the reasons for the difference in the predictive performance of the model in the primary(training) cohort and the validation cohort was that the (primary) training and validation cohorts played different roles in determining a model* [18,35]. *The good predict performance of the primary (training) cohort indicated that the model had a good predictive performance for the fitted data (primary or training cohort). But whether the predict performance of the model is robust should be verified on a new data set, that is, the validation cohort. The good predict performance of the model on the validation cohort indicated that it has predictive value on predicting unfitting new data, thus further proving that the model had good predictive ability and robustness* [36,37]. *In this study, the prediction performance of the clinic-radiological model decrease significantly in validation cohort, which means this model was not stable and reliable in differential diagnosis. As to the model based on radiomics signature alone, it performed well both in the primary and validation cohorts, which means it is reliable in differential diagnose. The intergrated model which had intergrated the clinical-radiological as well as* the *radiomics signatures performed well both in the primary and validation cohort, showing its robustness and reliability in the differential diagnose.* Moreover, both the clinicoradiological features alone and the radiomic signature alone achieved relatively high sensitivity, but their specificity was low, almost comparable to chance. These findings may suggest that the models based on clinicoradiological features or on the radiomic signature alone were poor diagnostic tools for differentiating atypical NF-pNET and PDAC. When clinicoradiological features and radiomic signature were integrated in a model, the specificity of diagnosis increased significantly, while a high sensitivity was maintained. As shown above, the sensitivity and specificity of the integrated model were higher than those of the clinicoradiological features or of the radiomic signature models alone. Therefore, we may conclude that *the intergrated model can improve the specificity of the differential diagnosis* and improve the diagnosis confidence to differentiate atypical NF-pNET from PDAC in clinical practice. This nomogram represented the intergrated model will provide both doctors and patients with a noninvasive scoring system to personalize the risk of atypical NF-pNET.

To the best of our knowledge, previous research mainly focused on the clinicoradiological features that can differentiate pNET from PDAC.

No one has previously used radiomics, which is a promising tool in individualized medicine. In this study, we used three different models to differentiate these two pancreatic solid lesions and compared the diagnostic performance of the models in an independent validation cohort. The integrated nomogram outperformed the models based on clinicoradiological features alone and radiomic features alone with respect to differential diagnosis of atypical NF-pNET from PDAC. We represented the integrated model with a nomogram, which showed high accuracy in both primary (AUC = 0.960) and validation cohorts (AUC = 0.884).

However, our study had several limitations. First, due to the retrospective study design, only the image in 5 mm slice thickness was available for all the patients. So all the radiomics analysis was performed on 5 mm slice. However, in this study, in calculation of tumor's radiomics signatures, each radiomics signature is calculated layer by layer, and the signatures of each layer were added and averaged to present the tumor's radiomics signatures. Therefore, the effects of non-isotropic and layer thickness interference are relatively small. Second, it is a retrospective study. This study only included data from NF-pNET and PDAC, which were confirmed by pathology. We did not include other pancreatic solid lesions, such as solid pseudopapillary tumors; our population therefore did not accurately reflect clinical reality. However, PDAC and NF-pNET are the two most common pancreatic solid lesions. In clinical practice, differential diagnosis between these lesions is a common dilemma for clinicians.Third, recruitment time for patients with these two diseases was different, which may have generated a selection bias. However, the scan protocols of these two cohorts were the same. Forth, we did not collect clinical information on certain factors, such as jaundice and weight loss.

In conclusion, the integrated model outperformed the model based on the clinicoradiological features alone and performed comparably to the model based on the radiomic signature alone in the differential diagnosis of atypical NF-pNET versus PDAC. The nomogram achieved an optimal preoperative, noninvasive differential diagnosis between atypical pNET and PDAC, which facilitates informed therapeutic choices in clinical practice.

Conflict of interest

The authors of this manuscript declare no relationships with any companies, whose products or services may be related to the subject matter of the article.

Appendix A. Supplementary data

Supplementary material related to this article can be found, in the online version, at doi:https://doi.org/10.1016/j.ejrad.2019.05.024.

References

[1] J.L. Tokar, R. Walia, Diagnostic evaluation of solid pancreatic masses, Curr. Gastroenterol. Rep. 15 (10) (2013) 347.

[2] S. Singh, C. Dey, H. Kennecke, W. Kocha, J. Maroun, P. Metrakos, T. Mukhtar, J. Pasieka, D. Rayson, C. Rowsell, L. Sideris, R. Wong, C. Law, Consensus recommendations for the diagnosis and management of pancreatic neuroendocrine tumors: guidelines from a Canadian National Expert Group, Ann. Surg. Oncol. 22 (8) (2015) 2685–2699.

[3] J.Y. Kim, S.M. Hong, J.Y. Ro, Recent updates on grading and classification of neuroendocrine tumors, Ann. Diagn. Pathol. 29 (2017) 11–16.

[4] K. Young, R. Iyer, D. Morganstein, I. Chau, D. Cunningham, N. Starling, Pancreatic neuroendocrine tumors: a review, Fut. Oncol. (Lond., Engl.) 11 (5) (2015) 853–864.

[5] A. Paniccia, B.H. Edil, R.D. Schulick, Pancreatic neuroendocrine tumors: an update, Indian J. Surg. 77 (5) (2015) 395–402.

[6] J.M. Cloyd, G.A. Poultsides, Non-functional neuroendocrine tumors of the pancreas: advances in diagnosis and management, World J. Gastroenterol. 21 (32) (2015) 9512–9525.

[7] M. Falconi, B. Eriksson, G. Kaltsas, D.K. Bartsch, J. Capdevila, M. Caplin, B. Kos-Kudla, D. Kwekkeboom, G. Rindi, G. Kloppel, N. Reed, R. Kianmanesh, R.T. Jensen, C. Vienna Consensus, ENETS consensus guidelines update for the management of patients with functional pancreatic neuroendocrine tumors and non-functional pancreatic neuroendocrine tumors, Neuroendocrinology 103 (2) (2016) 153–171.

[8] P.E. Humphrey, F. Alessandrino, A.M. Bellizzi, K.J. Mortele, Non-hyperfunctioning pancreatic endocrine tumors: multimodality imaging features with histopathological correlation, Abdom. Imaging 40 (7) (2015) 2398–2410.

[9] G. Foti, L. Boninsegna, M. Falconi, R.P. Mucelli, Preoperative assessment of non-functioning pancreatic endocrine tumours: role of MDCT and MRI, Radiol. Med. 118 (7) (2013) 1082–1101.

[10] S.K. Jeon, J.M. Lee, I. Joo, E.S. Lee, H.J. Park, J.Y. Jang, J.K. Ryu, K.B. Lee, J.K. Han, Nonhypervascular pancreatic neuroendocrine tumors: differential diagnosis from pancreatic ductal adenocarcinomas at MR imaging-retrospective cross-sectional study, Radiology 284 (1) (2017) 77–87.

[11] S.P. Raman, R.H. Hruban, J.L. Cameron, C.L. Wolfgang, E.K. Fishman, Pancreatic imaging mimics: part 2, pancreatic neuroendocrine tumors and their mimics, AJR Am. J. Roentgenol. 199 (2) (2012) 309–318.

[12] W. Yang, M.H. Chen, K. Yan, W. Wu, Y. Dai, H. Zhang, Differential diagnosis of non-functional islet cell tumor and pancreatic carcinoma with sonography, Eur. J. Radiol. 62 (3) (2007) 342–351.

[13] F.V. Coakley, K. Hanley-Knutson, J. Mongan, R. Barajas, M. Bucknor, A. Qayyum, Pancreatic imaging mimics: part 1, imaging mimics of pancreatic adenocarcinoma, AJR, Am. J. Roentgenol. 199 (2) (2012) 301–308.

[14] Y. Huang, Z. Liu, L. He, X. Chen, D. Pan, Z. Ma, C. Liang, J. Tian, C. Liang, Radiomics signature: a potential biomarker for the prediction of disease-free survival in early-stage (I or II) non-small cell lung Cancer, Radiology 281 (3) (2016) 947–957.

[15] S. Partelli, M. Inama, A. Rinke, N. Begum, R. Valente, V. Fendrich, D. Tamburrino, T. Keck, M.E. Caplin, D. Bartsch, C. Thirlwell, G. Fusai, M. Falconi, Long-term outcomes of surgical management of pancreatic neuroendocrine tumors with synchronous liver metastases, Neuroendocrinology 102 (1-2) (2015) 68–76.

[16] A. Perren, A. Couvelard, J.Y. Scoazec, F. Costa, I. Borbath, G. Delle Fave, V. Gorbounova, D. Gross, A. Grossma, R.T. Jense, M. Kulke, K. Oeberg, G. Rindi, H. Sorbye, S. Welin, ENETS consensus guidelines for the standards of care in neuroendocrine tumors: pathology: diagnosis and prognostic stratification, Neuroendocrinology 105 (3) (2017) 196–200.

[17] M.A. Tempero, M.P. Malafa, M. Al-Hawary, H. Asbun, A. Bain, S.W. Behrman, A.B. Benson 3rd, E. Binder, D.B. Cardin, C. Cha, E.G. Chiorean, V. Chung, B. Czito, M. Dillhoff, E. Dotan, C.R. Ferrone, J. Hardacre, W.G. Hawkins, J. Herman, A.H. Ko, S. Komanduri, A. Koong, N. LoConte, A.M. Lowy, C. Moravek, E.K. Nakakura, E.M. O'Reilly, J. Obando, S. Reddy, C. Scaife, S. Thayer, C.D. Weekes, R.A. Wolff, B.M. Wolpin, J. Burns, S. Darlow, Pancreatic adenocarcinoma, version 2.2017, NCCN clinical practice guidelines in oncology, J. Compr. Canc. Netw. 15 (8) (2017) 1028–1061.

[18] G. James, D. Witten, T. Hastie, R. Tibshirani, An Introduction to Statistical Learning, (2013).

[19] J.H. Kim, H.W. Eun, Y.J. Kim, J.M. Lee, J.K. Han, B.I. Choi, 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. 26 (5) (2016) 1338–1347.

[20] J. Guo, Z. Liu, C. Shen, Z. Li, F. Yan, J. Tian, J. Xian, MR-based radiomics signature in differentiating ocular adnexal lymphoma from idiopathic orbital inflammation, Eur. Radiol. (2018).

[21] V.P. Balachandran, M. Gonen, J.J. Smith, R.P. DeMatteo, Nomograms in oncology: more than meets the eye, Lancet Oncol. 16 (4) (2015) e173–80.

[22] Z. Liu, X.Y. Zhang, Y.J. Shi, L. Wang, H.T. Zhu, Z. Tang, S. Wang, X.T. Li, J. Tian, Y.S. Sun, Radiomics analysis for evaluation of pathological complete response to neoadjuvant chemoradiotherapy in locally advanced rectal Cancer, Clin. Cancer Res. 23 (23) (2017) 7253–7262.

[23] M.A. Attiyeh, J. Chakraborty, A. Doussot, L. Langdon-Embry, S. Mainarich, M. Gonen, V.P. Balachandran, M.I. D'Angelica, R.P. DeMatteo, W.R. Jarnagin, T.P. Kingham, P.J. Allen, A.L. Simpson, R.K. Do, Survival prediction in pancreatic ductal adenocarcinoma by quantitative computed tomography image analysis, Ann. Surg. Oncol. 25 (4) (2018) 1034–1042.

[24] Z. Lei, J. Li, D. Wu, Y. Xia, Q. Wang, A. Si, K. Wang, X. Wan, W.Y. Lau, M. Wu, F. Shen, Nomogram for preoperative estimation of microvascular invasion risk in hepatitis B virus-related hepatocellular carcinoma within the milan criteria, JAMA Surg. 151 (4) (2016) 356–363.

[25] M.J. Leening, M.M. Vedder, J.C. Witteman, M.J. Pencina, E.W. Steyerberg, Net reclassification improvement: computation, interpretation, and controversies: a literature review and clinician's guide, Ann. Intern. Med. 160 (2) (2014) 122–131.

[26] A.C. Alba, T. Agoritsas, M. Walsh, S. Hanna, A. Iorio, P.J. Devereaux, T. McGinn, G. Guyatt, Discrimination and calibration of clinical prediction models: users' guides to the medical literature, Jama 318 (14) (2017) 1377–1384.

[27] M.J. Pencina, R.B. D'Agostino Sr, R.B. D'Agostino Jr, R.S. Vasan, Evaluating the added predictive ability of a new marker: from area under the ROC curve to reclassification and beyond, Stat. Med. 27 (2) (2008) 157–172 discussion 207-12.

[28] E.G. Dumlu, D. Karakoc, A. Ozdemir, Nonfunctional pancreatic neuroendocrine tumors: advances in diagnosis, management, and controversies, Int. Surg. 100 (6) (2015) 1089–1097.

[29] M.F. Akisik, K. Sandrasegaran, G.X. Bu, C. Lin, G.D. Hutchins, E.G. Chiorean, Pancreatic Cancer: utility of dynamic contrast-enhanced MR imaging in assessment of antiangiogenic therapy, Radiology 256 (2) (2010) 441–449.

[30] A.J. Aguirre, E.A. Collisson, Advances in the genetics and biology of pancreatic Cancer, Cancer J. 23 (6) (2017) 315–320.

[31] P. Addeo, E. Rosso, P. Fuchshuber, E. Oussoultzoglou, V. De Blasi, G. Simone, C. Belletier, P. Dufour, P. Bachellier, Resection of Borderline Resectable and Locally Advanced Pancreatic Adenocarcinomas after Neoadjuvant Chemotherapy, Oncology 89 (1) (2015) 37–46.

[32] G. Karmazanovsky, E. Belousova, W. Schima, A. Glotov, D. Kalinin, A. Kriger, Nonhypervascular pancreatic neuroendocrine tumors: spectrum of MDCT imaging findings and differentiation from pancreatic ductal adenocarcinoma, Eur. J. Radiol. 110 (2019) 66–73.
[33] J.A. Sparano, R.J. Gray, D.F. Makower, K.I. Pritchard, K.S. Albain, D.F. Hayes, C.E. Geyer Jr, E.C. Dees, E.A. Perez, J.A. Olson Jr, J. Zujewski, T. Lively, S.S. Badve, T.J. Saphner, L.I. Wagner, T.J. Whelan, M.J. Ellis, S. Paik, W.C. Wood, P. Ravdin, M.M. Keane, H.L. Gomez Moreno, P.S. Reddy, T.F. Goggins, I.A. Mayer, A.M. Brufsky, D.L. Toppmeyer, V.G. Kaklamani, J.N. Atkins, J.L. Berenberg, G.W. Sledge, Prospective validation of a 21-Gene expression assay in breast Cancer, N. Engl. J. Med. 373 (21) (2015) 2005–2014.
[34] Y.Q. Huang, C.H. Liang, L. He, J. Tian, C.S. Liang, X. Chen, Z.L. Ma, Z.Y. Liu, Development and validation of a radiomics nomogram for preoperative prediction of lymph node metastasis in colorectal Cancer, J. Clin. Oncol. 34 (18) (2016) 2157–2164.
[35] M. Kuhn, K. Johnson, Applied Predictive Modeling, (2013).
[36] S.J. Russell, P. Norvig, Artificial Intelligence: a Modern Approach, Pearson Education Limited, Malaysia, 2016.
[37] B.D. Ripley, Pattern Recognition and Neural Networks, Cambridge university press, 2007.