European Journal of Radiology 121 (2019) 108647

Contents lists available at ScienceDirect

European Journal of Radiology

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

Development and validation of an MRI-based radiomic signature for the preoperative prediction of treatment response in patients with invasive functional pituitary adenoma

Yanghua Fan[a,1], Zhenyu Liu[b,c,1], Bo Hou[d,1], Longfei Li[e,1], Xiaohai Liu[a], Zehua Liu[e], Renzhi Wang[a], Yusong Lin[e], Feng Feng[d,***], Jie Tian[b,c,f,g,**], Ming Feng[a,*]

[a] Department of Neurosurgery, Peking Union Medical College Hospital, Chinese Academy of Medical Sciences & Peking Union Medical College, Beijing, 100032, China
[b] Key Laboratory of Molecular Imaging, Institute of Automation, Chinese Academy of Sciences, Beijing, 100190, China
[c] University of Chinese Academy of Sciences, Beijing, 100080, China
[d] Department of Radiology, Peking Union Medical College Hospital, Chinese Academy of Medical Sciences & Peking Union Medical College, Beijing, 100032, China
[e] Collaborative Innovation Center for Internet Healthcare, Zhengzhou University, Zhengzhou 450052, Henan, China
[f] Beijing Advanced Innovation Center for Big Data-Based Precision Medicine, School of Medicine, Beihang University, Beijing, 100191, China
[g] Engineering Research Center of Molecular and Neuro Imaging of Ministry of Education, School of Life Science and Technology, Xidian University, Xi'an, Shanxi, 710126, China

ARTICLE INFO

Keywords:
Invasive functional pituitary adenoma
Treatment response
Magnetic resonance imaging
Radiomics

ABSTRACT

Purpose: The preoperative prediction of treatment response is important for determining individual treatment strategies for invasive functional pituitary adenoma (IFPA). This study aimed to develop and validate a magnetic resonance imaging (MRI)-based radiomic signature for preoperative prediction of treatment response in IFPA.
Method: One hundred and sixty-three patients with IFPA were enrolled and divided into primary (n = 108) and validation cohorts (n = 55) according to time point. IFPA patients were divided into remission and non-remission according to postoperative hormone levels. Radiomic features were extracted from their MR images and a radiomic signature was built using a support vector machine. Subsequently, multivariable logistic regression analysis was used to select the most informative clinical features, and a radiomic model incorporating the radiomic signature and selected clinical features was constructed and used as the final predictive model.
Results: Seven radiomic features were selected to construct the radiomic signature, which achieved an area under the curve (AUC) of 0.834 and 0.808 on the primary and validation cohorts respectively. The radiomic model incorporating the radiomic signature and Knosp grade showed good discrimination abilities and calibration, with AUCs of 0.832 and 0.811 for the primary and validation cohorts respectively. The radiomic signature and radiomic model better estimated the treatment responses of patients with IFPA than our clinical features model. Decision curve analysis showed the radiomic model was clinically useful.
Conclusions: This radiomic model may help neurosurgeons predict the treatment responses of patients with IFPA before surgery and determine individual treatment strategies.

Abbreviations: IFPA, invasive functional pituitary adenoma; MRI, magnetic resonance imaging; AUC, area under the curve; GH, growth hormone; ACTH, adrenocorticotropic hormone; TMD, tumor maximal diameter; T2WI, T2-weighted imaging; T1WI, T1-weighted imaging; CE-T1, contrast-enhanced T1 weighted imaging; TR/TE, repetition time/echo time; FOV, field of view; ROIs, regions of interests; SVM, support vector machine; ROC, receiver operating characteristic; ACC, accuracy; PPV, positive predictive value; NPV, negative predictive value; SSA, somatostatin analogs

* Corresponding author at: Department of Neurosurgery, Peking Union Medical College Hospital, Chinese Academy of Medical Sciences & Peking Union Medical College, Beijing, China.
** Corresponding author at: Key Laboratory of Molecular Imaging, Institute of Automation, Chinese Academy of Sciences, Beijing, China.
*** Corresponding author at: Department of Radiology, Peking Union Medical College Hospital, Chinese Academy of Medical Sciences & Peking Union Medical College, Beijing, China.
E-mail addresses: fengfeng@vip.163.com (F. Feng), jie.tian@ia.ac.cn (J. Tian), pumchfengming@163.com (M. Feng).
[1] These authors contributed equally to this work.

https://doi.org/10.1016/j.ejrad.2019.108647
Received 9 January 2019; Received in revised form 9 July 2019; Accepted 6 September 2019

1. Introduction

Invasive functional pituitary adenomas (IFPAs) are the most difficult to treat and most harmful type of pituitary adenoma [1]. Preoperative definition of IFPA requires two conditions: abnormal hormone secretion and a Knosp grade of III–IV [2–4]. Because of their abnormal hormone secretion, IFPAs have multiple complications, and the resulting multiple comorbidities impair patients' quality of life and reduce life expectancy [5]. In clinical practice, a Knosp grade of III–IV means true invasion of the tumor into the cavernous sinus before surgery, and therefore complete removal of the IFPA with standard surgical treatment is difficult [2,4,6]. If there is even a small residual tumor, the patient will still be in a status of elevated hormone levels and persistent symptoms [6]. Postoperative high hormone-status indicates a poor treatment response to surgery, with a likelihood of recurrence and a poor prognosis [5]. If clinicians could accurately predict the treatment responses of patients with IFPA before surgery, those with a poor outcome prediction could undergo appropriate preoperative adjuvant therapy, which might greatly improve the effectiveness of neurosurgery and their ensuing quality of life [7]. Thus, accurate preoperative identification of treatment response in patients with IFPA would be helpful for prognosis and decisions on treatment strategy.

Currently, most pituitary adenoma-related prognostic studies have focused on the retrospective analysis of clinical risk factors and their associations with treatment response after surgery [8,9]. Previous analysis demonstrated that MRI texture was useful for predicting recurrence or progression in patients with non-functioning pituitary adenomas (NFPAs) [10]. However, to the best of our knowledge, no previous study has investigated the presurgical prediction of treatment response and prognosis of IFPA. Moreover, the prognosis should not be determined by only a single risk factor, with it being recognized that a combined analysis with multiple features has more value, and may be powerful enough to change the clinical management [11,12]. Thus, the development of a more comprehensive, effective, and widely applicable preoperative prediction method is conducive to the treatment of IFPAs.

Radiomics is an emerging methodology that allows high-throughput mining of quantitative imaging features from normal medical images followed by their combined analysis for clinical-decision support system development [13,14]. Recently, studies have demonstrated that radiomics-based signatures can be used to develop treatment strategies and improve diagnostic, prognostic, and predictive accuracy in the workup of tumor patients [15]. However, a radiomic signature for predicting treatment responses in patients with pituitary adenoma has not yet been reported, especially for IFPA. Therefore, if a radiomic signature based on quantitative MRI features could preoperatively predict the treatment response and prognosis in IFPA, it would be particularly beneficial, as MRI is non-invasive and would not require extra financial cost, as MRI is generally acquired as part of the current workup.

Therefore, the aim of this study was to develop an MRI-based radiomic signature to estimate treatment response in patients with IFPA. The predictive model so developed incorporated a radiomic signature and clinical features for preoperative prediction of treatment response in patients with IFPA.

2. Materials and methods

2.1. Patients

This retrospective study was approved by the ethical review committee of the Peking Union Medical College Hospital, and the requirement for informed patient consent was waved. A total of 163 patients presenting with IFPAs between April 2012 and May 2018 were identified for this analysis, and all patient data and personal information were anonymized prior to analysis. The patients were divided into a primary cohort and validation cohort according to time point. The primary cohort consisted of 108 patients recruited from April 19, 2012 to December 26, 2016 (36 males and 72 females; mean age, 35 years; age range, 28–47 years). A subsequent cohort of 55 consecutive patients recruited from December 27, 2016 to May 22, 2018 were allocated to the validation cohort (20 males and 35 females; mean age, 34 years; age range, 29–46 years). The primary cohort was used for model building, while the validation cohort was set aside for internal validation of the model. The inclusion and exclusion criteria are presented in supplementary material S1.

All patients underwent neurosurgery to remove the IFPA, and the following clinical characteristics were collected: age, gender, type of diagnosis (growth hormone [GH], prolactin, adrenocorticotropic hormone [ACTH], thyroid-stimulating hormone, luteinizing hormone/follicle-stimulating hormone secreting IFPAs), tumor maximal diameter (TMD), bilateral Knosp classification (grade III or grade IV) [2], and treatment response.

The type of diagnosis was determined by the preoperative hormone level [1]. Knosp classification and TMD were defined using preoperative contrast-enhanced T1-weighted imaging (CE-T1), with preoperative grades III and IV defining invasion of the tumor into the cavernous sinus [2,4]. Patients were divided into remission or non-remission groups according to their different treatment responses. The remission criteria for IFPAs were as follows. Postoperative remission of GH-secreting IFPAs was defined as random serum GH < 1 ng/ml or a GH nadir < 0.4 ng/ml during an oral glucose tolerance test (OGTT) at 12 weeks after surgical treatment [16,17]. The OGTT was carried out with 75 g of oral glucose and subsequent measurements of glucose and GH every 30 min over 2 h. Postoperative remission of ACTH-secreting IFPAs was defined as a morning serum cortisol measurement of less than 138 nmol/L (5ug/dl) within one week after surgical treatment [18–20]. The postoperative remission criteria for PRL, TSH, and LH/FSH secreting IFPAs were that the corresponding hormones fell below normal levels. A flowchart of this study is shown in Fig. 1.

2.2. MRI protocol and image acquisition

All patients underwent preoperative brain MRI, with the MRI protocol consisting of T2-weighted imaging (T2WI), T1-weighted imaging (T1WI), and CE-T1. MRI was acquired in the head-first supine position on a 3.0-T scanner (GE, General Electric Company). The acquisition parameters for the T2WI sequence included repetition time/echo time (TR/TE): 4200/103 ms; flip angle: 90°; acquisition matrix: 320 × 224; field of view (FOV): 200 × 200 mm; slice thickness: 4 mm and slice spacing: 1 mm. The acquisition parameters for T1WI sequences included TR/TE: 400/9 ms; flip angle: 90°; acquisition matrix: 288 × 192; FOV: 200 × 200 mm; slice thickness: 3 mm and slice spacing: 0.6 mm. The CE-T1 sequences used parameters consistent with the other T1WI sequence and were preformed immediately after the rapid injection of the gadolinium-DTPA contrast agent (0.1 mmol/kg Gadovist). For this study, coronal T2WI, T1WI, and CE-T1 were used, with these being retrieved from the hospital's Picture Archiving and Communication System.

2.3. Tumor masking and feature extraction

Three-dimensional regions of interests (ROIs) delineating the tumors of all IFPA patients were drawn manually on the above-mentioned MRI images by a neuro-radiologist (with seven years of experience in the study of IFPAs) using ITK-SNAP software (University of Pennsylvania, www.itksnap.org). These ROIs were manually verified by another expert neuroradiologist (with twelve years of experience in the study of IFPAs), without prior knowledge of the operating records. Any disagreements were resolved through negotiation between the two experts. The radiomic features were then extracted from the ROIs.

A total of 13 950 quantitative features were automatically extracted from the delineated tumors on the coronal T2WI, T1WI, and CE-T1;

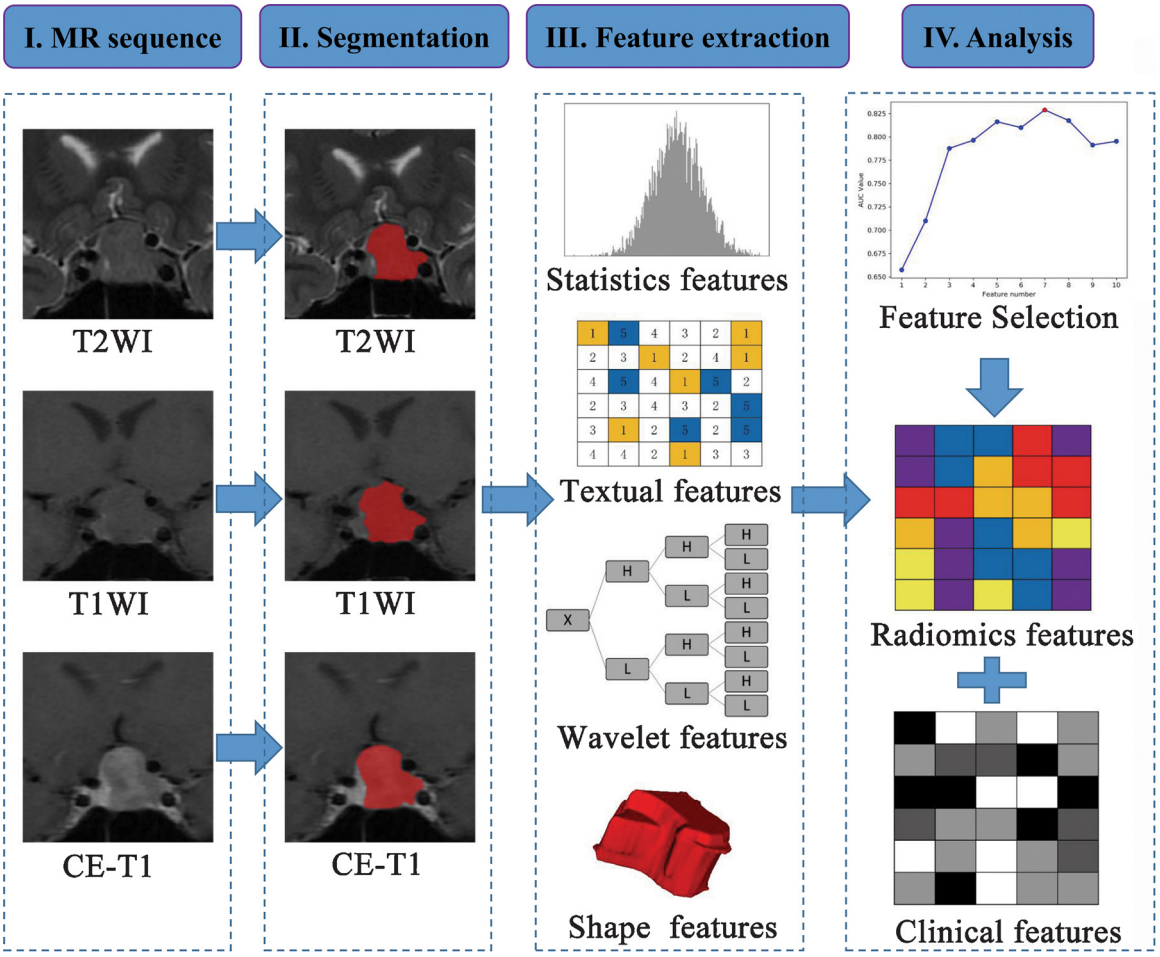

Fig. 1. Flowchart illustrating the procedures in this study. (I) Original MR images: coronal T2WI, T1WI, and CE-T1 sequences. (II) Segmentation was conducted to define the ROIs. (III) Radiomic features were extracted from the ROIs, including statistics features, textual features, wavelet features, and shape features. (IV) Analysis of the radiomic and clinical features after feature selection.

Table 1
Clinical characteristics of IFPAs patients (n = 163).

Characteristic	Whole Set (n = 163)	Primary cohort(n = 108)	Validation cohort(n = 55)	P-value
Gender				
Male	56(34.4%)	36(33.3%)	20(36.4%)	0.7
Female	107(65.6%)	72(64.3%)	35(63.6)	
age (mean ± SD, year)	37.56 ± 11.85	37.89 ± 12.51	36.91 ± 10.52	0.765
Diagnosis				
GH	115(70.6%)	76(70.4%)	39(70.9%)	0.429
PRL	25(15.3%)	18(16.7%)	7(12.7%)	
ACTH	16(9.8%)	11(10.2%)	5(9.1%)	
TSH	6(3.7%)	2(1.9%)	4(7.3%)	
LH/FSH	1(0.6%)	1(0.9%)	0	
Tumor maximal diameter (mean ± SD, mm)	23.39 ± 9.97	24.71 ± 9.90	21.84 ± 10.00	0.106
Knosp classification				
Grade 3	96(58.9%)	60(55.6%)	36(65.5%)	0.225
Grade 4	67(41.1%)	48(44.4%)	19(34.5%)	
Treatment responses				
Non-remission	97(66.3%)	66(61.1%)	31(56.4%)	0.559
Remission	66(33.7%)	42(38.9%)	24(43.6%)	

Abbreviations: SD = standard deviation.
Note: Categorical variables were presented as the number (percentage). Continuous variables consistent with a normal distribution were presented as mean ± standard deviation. Chi-Square or Fisher Exact tests, as appropriate, were used to compare the differences in categorical variables, while the independent sample *t*-test was used to compare the differences in continuous variables.

Table 2
Univariate analysis of clinical characteristics of patients and tumors in the primary cohort and validation cohort.

Characteristic	Primary cohort(n = 108)		P-value	Validation cohort(n = 55)		P-value
	Non-remission	Remission		Non-remission	Remission	
Gender						
Male	22(33.3%)	14(33.3%)	1	12(38.7%)	8(33.3%)	0.681
Female	44(66.7%)	28(66.7%)		19(61.3%)	16(66.7%)	
age (mean ± SD, year)	37.21 ± 12.90	38.95 ± 11.95	0.484	37.90 ± 10.51	35.63 ± 10.62	0.431
Diagnosis						
GH	46(69.7%)	30(71.4%)	0.134	25(80.6%)	14(58.3%)	0.306
PRL	14(21.2%)	4(9.5%)		3(9.7%)	4(16.7%)	
ACTH	6(9.1%)	5(11.9%)		2(6.5%)	3(12.5%)	
TSH	0	2(4.8%)		1(3.2%)	3(12.5%)	
LH/FSH	0	1(2.4%)		0	0	
Tumor maximal diameter (mean ± SD, mm)	26.59 ± 10.26	20.38 ± 8.05	0.001*	25.69 ± 10.76	16.86 ± 6.19	0.001*
Knosp classification						
Grade 3	28(42.4%)	32(76.2%)	0.001*	14(45.2%)	22(91.7%)	0.000*
Grade 4	38(57.6%)	10(23.8%)		17(54.8%)	2(8.3%)	

Abbreviations: SD = standard deviation.
Note: Categorical variables were presented as the number (percentage). Continuous variables consistent with a normal distribution were presented as mean ± standard deviation. Chi-Square or Fisher Exact tests, as appropriate, were used to compare the differences in categorical variables, while the independent sample *t*-test was used to compare the differences in continuous variables.

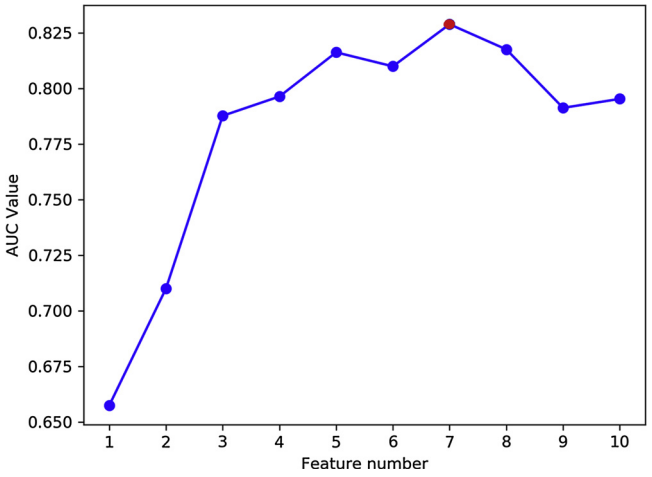

Fig. 2. Ten-fold cross-validation was used to estimate the average AUC value of the different feature sets on the primary cohort, and the best performance was achieved in the set with the first seven features.

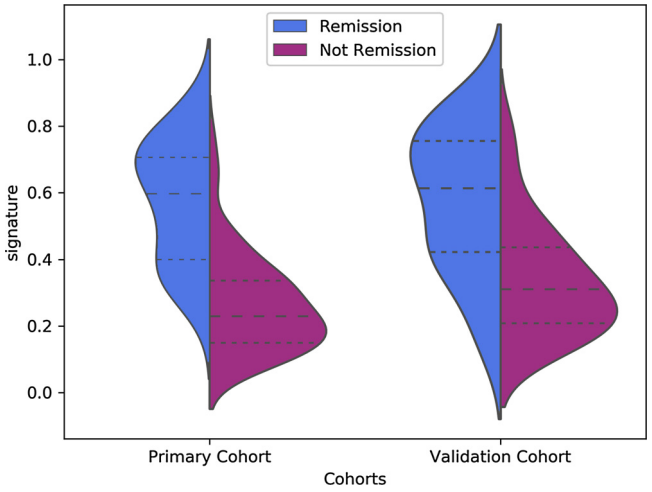

Fig. 3. A violin plot comparing the distribution of the signatures of remission and non-remission patients in the primary and validation cohorts. This plot is a combination of a boxplot and kernel density estimate. The signature distributions of each cohort were compared using independent samples *t*-tests.

4650 features for each sequence. These features were divided into four groups: (I) tumor image intensity features, (II) shape and size features, (III) textural features, and (IV) wavelet features. The radiomic features of the primary cohort were calculated from the ROIs using an inhouse program written in Matlab 2015b (Mathworks Inc. Natick, MA, USA). The mathematical definitions and details of the features and different wavelets are given in supplementary information S2.

2.4. Feature selection

High-dimensional data may involve a high degree of irrelevant and redundant information that can lead to overfitting and can greatly reduce the performance of the learning algorithm; therefore, a feature selection process is necessary. In this study, feature selection was performed in three stages, based on results obtained on the primary cohort. The radiomic features were first normalized to the range of 0–1. Second, features were preliminarily selected using the SelectKBest algorithm [21] to remove the majority of the redundancy. The SelectKBest algorithm selects the most appropriate radiomic feature based on the mutual information between the features and labels. Finally, the most important features for the optimal feature subset were identified using the recursive feature elimination (RFE) algorithm [22] with ten-fold cross-validation. The RFE algorithm is a commonly used feature selection method, which sorts features according to their importance and finds the feature set with the best prediction performance by deleting the least important features [23].

2.5. Development and validation of the radiomic signature

After obtaining the selected features, a support vector machine (SVM) model was used to build the radiomic signature for the primary cohort. The SVM was trained using a radial basis function (RBF) kernel, with the parameter C being determined by a grid search with ten-fold cross-validation. Finally, the radiomic signature was calculated by transformation of the representative features for subsequent use. A violin plot was used to compare differences between the distributions of the signatures of remission and non-remission patients in the primary and validation cohorts, while a receiver operating characteristic (ROC) [24] curve was plotted to show the predictive performance of the radiomic features.

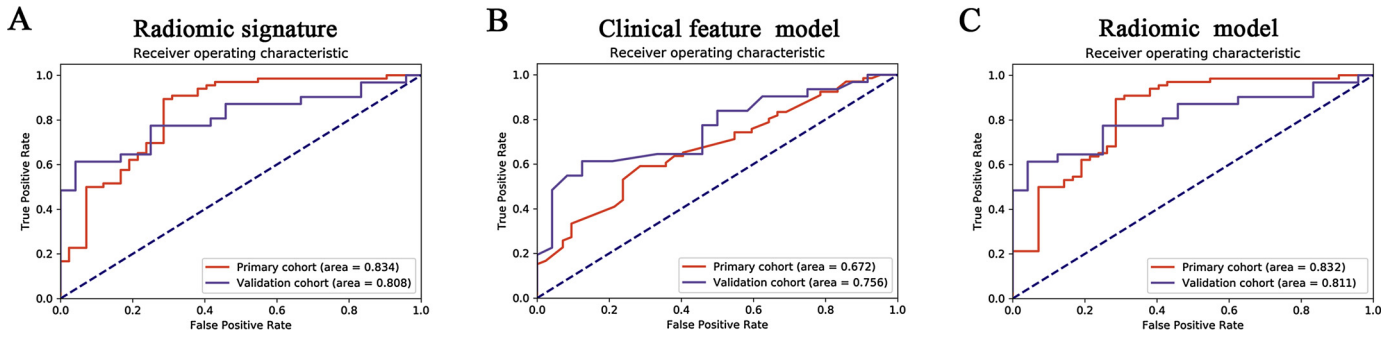

Fig. 4. ROC curves for the performance of the three models on the primary and validation cohorts. The performance of the models on the primary and validation cohorts was assessed using the area under the curve (AUC). A. radiomic signature; B. clinical feature model; C. radiomic model.

Table 3
Performance of radiomic signature, clinical features and radiomic model.

Model	Performance	AUC(95% CI)	ACC	SN	SP	PPV	NPV
Radiomic signature	Primary cohort	0.834 (0.807-0.861)	81.4%	87.9%	71.4%	82.9%	78.9%
	Validation cohort	0.808 (0.782-0.834)	72.7%	77.4%	66.7%	75.0%	69.6%
Clinical feature model	Primary cohort	0.672 (0.570-0.774)	62.0%	78.79%	35.71%	65.82%	51.72%
	Validation cohort	0.756 (0.606-0.902)	65.5%	64.52%	66.67%	71.43%	59.26%
Radiomic model	Primary cohort	0.832 (0.749-0.915)	81.5%	89.4%	69.0%	81.9%	80.6%
	Validation cohort	0.811 (0.696-0.925)	74.5%	61.3%	91.7%	70.5%	64.7%

Abbreviations: AUC, area under the curve; ACC accuracy, SN sensitivity; SP, specificity; PPV, positive predictive value; NPV, negative predictive value.

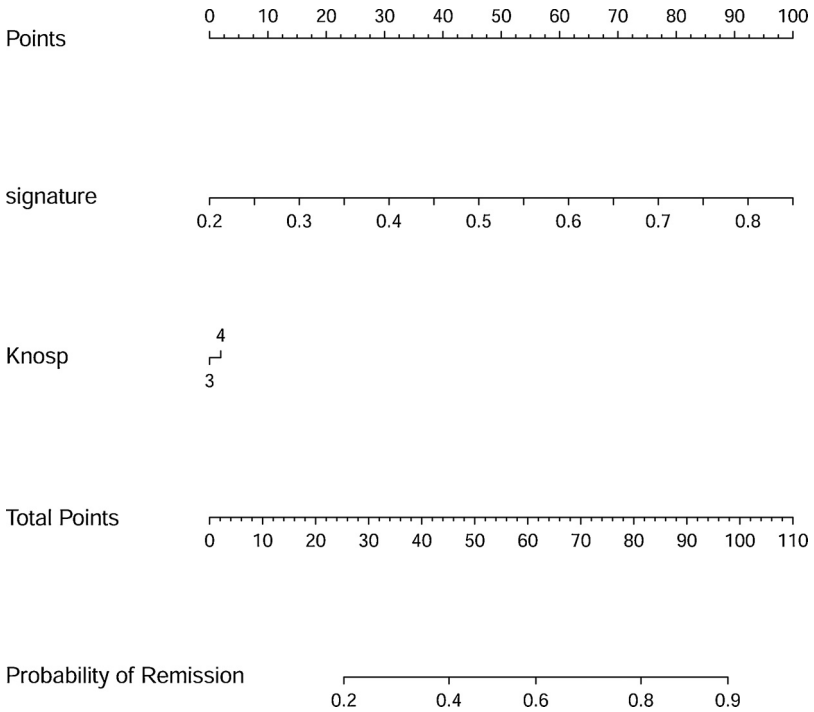

Fig. 5. A nomogram derived from the radiomic model applied to the primary cohort. The radiomic model was developed using the primary cohort, and incorporated the Knosp grade and radiomic signature.

2.6. Development and validation of the clinical features model

Multivariate logistic regression analysis was used to select the clinical features showing the strongest associations between the clinical information and treatment responses. The five preoperative clinical features of age, sex, diagnosis type, TMD, and Knosp classification were evaluated. Backward step-wise selection was conducted using the likelihood ratio test with Akaike's information criterion (AIC) [25] as the stopping rule. The selected clinical features were then used to develop a clinical features model for the preoperative prediction of treatment response in the primary cohort. Subsequently, the validation cohort was used to validate the performance of this clinical features model. The area under the curve (AUC) from the ROC analysis was used to assess the discriminatory efficacy of the clinical features model.

2.7. Development and validation of the radiomic model

To provide a comprehensive model for clinicians and patients to use to predict the treatment responses of IFPAs, a radiomic model [26] incorporating the radiomic signature and clinical features was

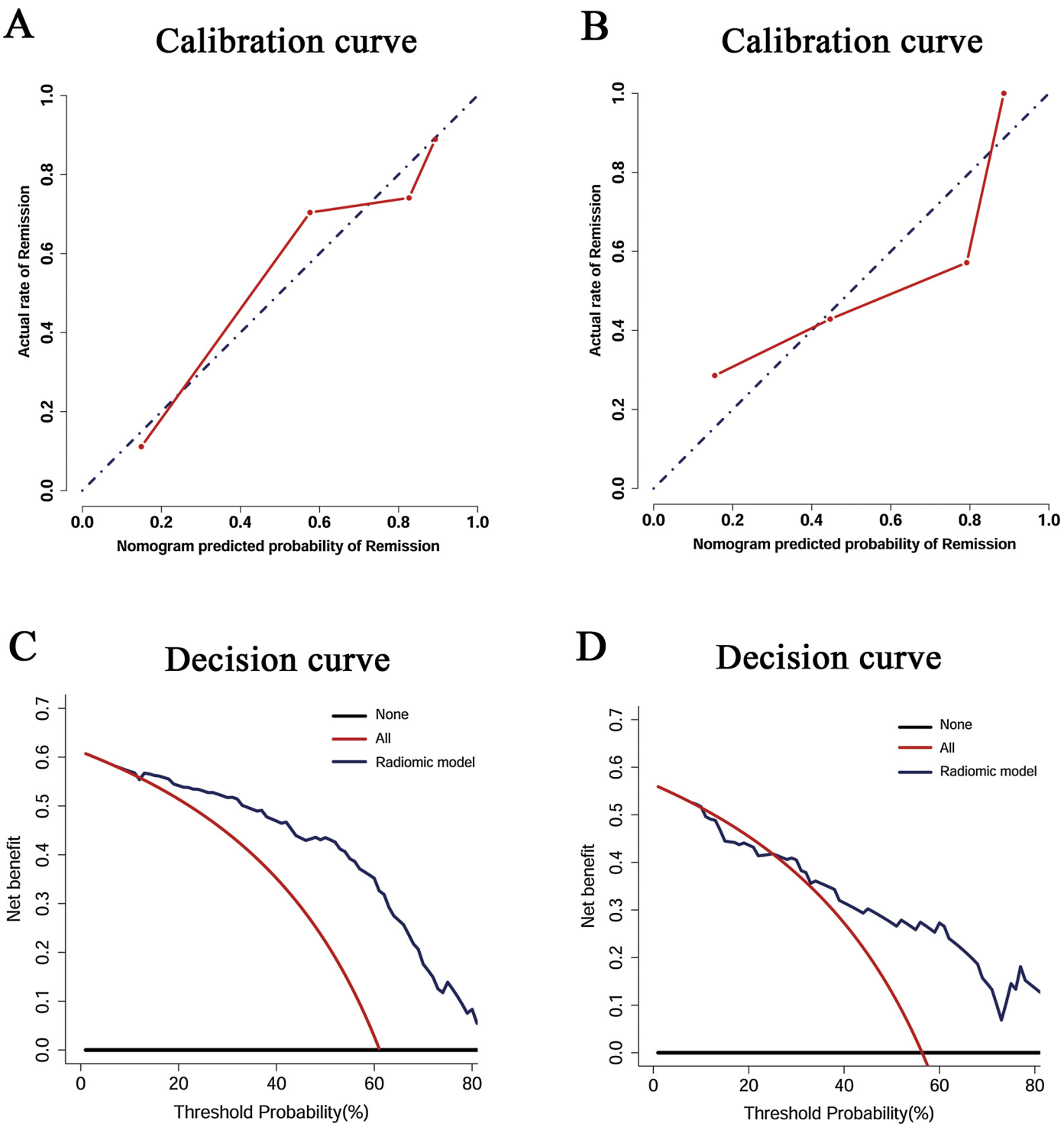

Fig. 6. Calibration curve and decision curve analysis for the radiomic model. A–B. Calibration curves of the radiomic model in the primary (A) and validation (B) cohorts. Calibration curves depict the calibration of each model in terms of the agreement between the predicted and actual probability of the remission rate. The Y axis represents the actual rate. The X axis represents the predicted probability. The diagonal blue line represents a perfect prediction by an ideal model. The red lines represent the performance of the radiomic model on the primary and validation cohorts respectively, with a closer fit to the diagonal blue line representing a better prediction. C–D. Decision curve analysis for the radiomic model. C: primary cohort, D: validation cohort. The Y axis measures the net benefit. The blue line represents the radiomic model. The red line represents the assumption that all patients showed remission. The black line represents the assumption that no patients showed remission.

constructed. The clinical features used in the radiomic model were selected from four preoperative clinical features (age, sex, diagnosis type, and Knosp classification) using a multivariate logistic regression algorithm with backward step-wise selection and AIC. A nomogram was derived from the radiomic model.

ROC analyses and associated classification measures (AUC, accuracy [ACC], sensitivity, specificity, positive predictive value [PPV], and negative predictive value [NPV]) were used to compare the discriminatory efficacies of the radiomic model, radiomic signature, and clinical features model, on both the primary and validation cohorts.

Calibration curves were plotted, and along with the Hosmer-Lemeshow test, were used to evaluate the similarity between the predicted and observed remission probabilities [27].

2.8. Clinical use

Decision curve analysis (DCA) was performed to assess the clinical usefulness of the radiomic model by quantifying the net benefits at different threshold probabilities [28].

2.9. Statistical analysis

A two-sided p-value < 0.05 was considered to be statistically significant. All statistical analyses were performed by a dedicated statistician using R statistical software (version 3.4.1; R Foundation for Statistical Computing, Vienna, Austria). The “rms” package was used to create the combined model, the calibration plot was analyzed using the “hdnom” packages, and decision curve analysis was analyzed with the function “dca.R”.

3. Results

3.1. Clinical characteristics

There were 42 (38.9%) and 24 (43.60%) patients in remission in the primary and validation cohorts respectively, with there being no significant difference in the prevalence of remission patients between the primary and validation cohorts (p = 0.559). In addition, no significant differences (p = 0.106–0.765) were found in the clinical characteristics between the primary and validation cohorts. The characteristics of all patients are shown in Table 1. These results justify the use of these two sets as a primary cohort and validation cohort. A univariate analysis was conducted on the primary and validation cohorts (Table 2), and TMD and Knosp classification were found to be significantly different between the non-remission and remission patients.

3.2. Feature selection and construction of the radiomic signature

Two steps were used to select the representative radiomic features of the primary cohort. First, 300 radiomic features were selected using the SelectKBest algorithm. Second, the importance of each feature was assessed using the RFE algorithm, with the 300 radiomic features being sorted according to their level of importance and the best performance being determined by ten-fold cross-validation. The seven features thus identified as giving the best performance were selected as the final features for subsequent use (Fig. 2).

The seven selected features of ‘glszm_ZSV’ of the gabor 9, ‘glszm_GLNU’ of the gabor 27, ‘glszm_ZSV’ of the gabor 31, ‘glrlm_SRLGLE’ of the gabor 33, ‘fos_minimum’ of the gabor 34, ‘glszm_GLNU’ of the gabor 34, and ‘glszm_ZSV’ of the gabor 34 were all extracted from CE-T1 images. These selected features were then entered into an SVM to build a radiomic signature with the best differentiation of non-remission from remission patients in the primary cohort.

3.3. Performance of the radiomic signature

The violin plot showed significant differences in the distribution of the radiomic signature between remission and non-remission patients in both the primary and validation cohorts (p < 0.01, Fig. 3). The model performance was first assessed in the primary cohort and then validated using the internal validation cohort. As shown in Fig. 4A, the radiomic signature showed favorable discrimination, with an AUC of 0.834 (95% confidence interval [CI], 0.807–0.861) in the primary cohort and 0.808 (95% CI, 0.782–0.834) in the validation cohort. The results demonstrated that the radiomic signature could successfully categorize the remission and non-remission IFPA patients in both the primary and validation cohorts.

3.4. Radiomic model construction and performance assessment

TMD was selected as the most important discriminatory factor from the 5 clinical features and yielded AUC values of 0.672 (95% CI, 0.570–0.774) and 0.756 (95% CI, 0.606–0.902) for the primary and validation cohorts respectively (Fig. 4B, Table 3).

Knosp grade, together with the radiomic signature, were selected according to the AIC to build the radiomic model, which yielded an AUC of 0.832 (95% CI, 0.749–0.915) for the primary cohort and 0.811 (95% CI, 0.696–0.925) for the validation cohort (Fig. 4C, Table 3). A nomogram was derived from the radiomic model and is presented in Fig. 5. The ACC, sensitivity, specificity, PPV, and NPV of the three models for the primary and validation cohorts are shown in Table 3. The Delong test showed that the radiomic model and radiomic signature performed significantly better than the TMD, but there was no significant difference between the radiomic model and radiomic signature.

Furthermore, a calibration curve for treatment response according to the radiomic model demonstrated good agreement between observations and predictions in both the primary (p = 0.10, Fig. 6A) and validation cohorts (p = 0.29, Fig. 6B). The Hosmer-Lemeshow test showed no statistical significance, indicating no significant departure from a perfect fit.

Good discrimination and good calibration were observed with the radiomic model, with the results showing that it predicted the treatment responses of IFPA patients more accurately than the clinical features model.

3.5. Clinical usefulness of the radiomic model

The decision curve analysis for the radiomic model is shown in Fig. 6C-D. The radiomic model offered a net benefit over presumption of remission or non-remission at a threshold probability > 13% in the primary cohort and > 25% in the validation cohort, thereby indicating that the radiomic model is clinically useful. The decision curve showed a relatively good performance for this radiomic model in terms of clinical application, and the use of this model to predict remission showed advantages over assuming either all patients would achieve remission, or none would achieve remission.

4. Discussion

In the present study, we developed and validated a radiomic model using a radiomic signature and clinical features for the preoperative individualized prediction of treatment response in patients with IFPA. This radiomic model showed favorable discrimination and calibration, and provides an effective non-invasive tool for deciding on individual treatment strategies in patients with IFPA.

IFPAs have lower postoperative remission rates than other pituitary adenomas, with the main factor hindering remission being the presence of cavernous sinus invasion. Unlike non-functioning pituitary adenomas, even a small residual IFPA tumor portion left in place can lead to the continuous hypersecretion of the corresponding hormone, and the patient’s symptoms will persist [6]. An increase in GH levels can cause cardiovascular and cerebrovascular diseases, and GH-secreting IFPAs carry a mortality rate at least twice as high as in the general population [29]. The chronic hypercortisolism caused by ACTH-secreting IFPAs can lead to serious complications and death, with a reported four-fold increased risk of mortality [30]. Therefore, in addition to the mass effect caused by the tumor size, IFPA patients may need individualized treatment for symptomatic elevated serum hormone levels.

At present, the need for preoperative adjuvant treatments for IFPA is subject to some controversy [31]. Some studies have shown that the

long-term preoperative application of somatostatin analogs (SSA) can improve the surgical curative rate of invasive GH-secreting IFPAs [32]. However, other scholars do not recommend the routine use of preoperative SSA for the improvement of postoperative remission rate [17,33]. It is currently considered that for IFPA patients with a preoperative prediction of a poor treatment response, preoperative adjuvant therapy can be performed to achieve optimal surgical conditions and improve the surgical outcome and patient's quality of life [7].

Considering the above findings, accurate preoperative prediction can change individualized treatment strategies and provide a basis for judging whether preoperative adjuvant therapy is needed. However, there has been no relevant study to predict the treatment response and prognosis of IFPA before surgery. Therefore, there is an urgent need for a more effective and widely applicable preoperative prediction method. Recently, the field of radiomics has undergone rapid development, with radiomic approaches having been used in the diagnosis or prognosis of multiple cancer types [34–37]. A radiomic model established by Zhang et al. showed good performance in predicting NFPA subtypes before surgery, indicating that radiomic methods have good potential for the preoperative diagnosis of pituitary adenomas [38]. Furthermore, Niu et al. demonstrated that the radiomic method may be helpful in predicting cavernous sinus invasion before pituitary adenoma surgery [39]. These studies suggest that the use of radiomic analysis is very effective in pituitary adenomas. Therefore, through the radiomic approach outlined in this study, we aimed to predict the treatment responses of IFPAs before surgery.

Thus, we first constructed a radiomic signature consisting of seven selected features and showed that this radiomic signature could successfully categorize remission and non-remission IFPA patients. Consistent with a previous radiomic study of non-functioning pituitary adenomas [39], all the included features were selected from CE-T1 images. This may be explained by the fact that each slice image of the CE-T1 acquisition has high resolution, and the contrast difference between normal and tumor tissues will increase after injection of contrast agent [40,41]. Secondly, we incorporated Knosp grade into a radiomic model along with the radiomic signature. Encouragingly, the radiomic model yielded an AUC of 0.832 in the primary cohort and 0.811 in the validation cohort, and showed favorable calibration and discrimination. Therefore, the model could help clinicians predict treatment response before operation.

We believe that this study on the development of a radiomic model for the preoperative prediction of treatment response in patients with IFPA offers several advantages. First, because of the generally lower proportion and higher hazard of IFPAs in comparison with PAs in general, the 163 eligible IFPA patients with complete clinical information and follow-up data present a valuable resource. Second, compared with a prior postoperative retrospective study of treatment response-related clinical risk factors for IFPAs, this radiomics-based preoperative prediction model containing clinical and imaging features was more effective and widely applicable. Third, compared with a prior MRI texture analysis [10], we extracted higher-dimensional features from the MRI images for the development of the radiomic signature, and the resulting radiomic signature more effectively revealed the heterogeneity of the IFPAs. Moreover, the radiomic signature was more reliable because the MRI images used to develop it were stored digitally and did not suffer from any degradation. Thus, our radiomic model may serve as a non-invasive tool for the preoperative prediction of treatment response in IFPA patients.

Our study has some limitations. This is a single center study, and although we divided patients into primary and validation cohorts based on the patients' surgical dates, the model may behave differently if a multicenter dataset with different parameters is used. Additionally, more data and patients from multiple centers could be used to validate the robustness and repeatability of our radiomic model. Moreover, the results of the Delong test showed that there was no significant difference between the radiomic model and radiomic signature, with the specificity of the radiomic model on the validation cohort being 91.7%; thus the radiomic model can be reliably used to determine the patients who can achieve postoperative remission. Because of the low sensitivity of the radiomic model and the high sensitivity of the radiomic signature, the radiomic signature is more suitable for identifying patients who are unlikely to achieve remission. Finally, radiomics can be complementary to transcriptomics, genomics, and proteomics, which could all be combined to build a more accurate predictive model.

In conclusion, this study focused on the preoperative prediction of treatment response in patients with IFPAs, and developed and validated a radiomic model incorporating an MRI-based radiomic signature and clinical features. The radiomic model can serve as an effective non-invasive approach to predict treatment response and can help to determine individual treatment strategies for patients with IFPAs.

Funding

This work was supported by the Graduate Innovation Fund of Peking Union Medical College (2018-1002-01-10); the Chinese Academy of Medical Sciences (grant No. 2017-I2M-3-014); National Natural Science Foundation of China under Grant (81922040, 81772012, 81772009); Beijing Natural Science Foundation under Grant (7182137, 7182109), Scientific and Technological Research Project of Henan Province (Grant 182102310162).

Declaration of Competing Interest

We declare that we have no conflicts of interest.

Acknowledgements

We thank Karl Embleton, PhD, from Liwen Bianji, Edanz Group China (www.liwenbianji.cn/ac), for editing the English text of a draft of this manuscript.

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.2019.108647.

References

[1] M.E. Molitch, Diagnosis and treatment of pituitary adenomas: a review, JAMA 317 (5) (2017) 516–524.
[2] E. Knosp, E. Steiner, K. Kitz, C. Matula, Pituitary adenomas with invasion of the cavernous sinus space: a magnetic resonance imaging classification compared with surgical findings, Neurosurgery 33 (4) (1993) 610–617 discussion 617-618.
[3] M.Y. Kim, J.H. Kim, Y.K. Oh, E. Kim, Long-term outcomes of surgery and radiotherapy for secreting and non-secreting pituitary adenoma, Radiat. Oncol. J. 34 (2) (2016) 121–127.
[4] A. Di Ieva, F. Rotondo, L.V. Syro, M.D. Cusimano, K. Kovacs, Aggressive pituitary adenomas–diagnosis and emerging treatments, Nat. Rev. Endocrinol. 10 (7) (2014) 423–435.
[5] N. Ironside, G. Chatain, D. Asuzu, S. Benzo, M. Lodish, S. Sharma, L. Nieman, C.A. Stratakis, R.R. Lonser, P. Chittiboina, Earlier post-operative hypocortisolemia may predict durable remission from Cushing's disease, Eur. J. Endocrinol. 178 (3) (2018) 255–263.
[6] E.H. Kim, M.C. Oh, J.H. Chang, J.H. Moon, C.R. Ku, W.S. Chang, E.J. Lee, S.H. Kim, Postoperative gamma knife radiosurgery for cavernous sinus-invading growth hormone-secreting pituitary adenomas, World Neurosurg. 110 (2018) e534–e545.
[7] S. Bacigaluppi, F. Gatto, P. Anania, N.L. Bragazzi, D.C. Rossi, G. Benvegnu, E. Nazzari, R. Spaziante, M. Giusti, D. Ferone, G. Zona, Impact of pre-treatment with somatostatin analogs on surgical management of acromegalic patients referred to a single center, Endocrine 51 (3) (2016) 524–533.
[8] S. Yano, N. Shinojima, J. Kawashima, T. Kondo, T. Hide, Intraoperative scoring system to predict postoperative remission in endoscopic endonasal transsphenoidal surgery for growth hormone-secreting pituitary adenomas, World Neurosurg. 105 (2017) 375–385.
[9] W.J. Zhou, S.C. Ma, M. Zhao, C. Liu, X.D. Guan, Z.S. Bao, G.J. Jia, W. Jia, Risk factors and the prognosis of sexual dysfunction in male patients with pituitary adenomas: a multivariate analysis, Asian J. Androl. 20 (1) (2018) 43–49.
[10] B.P. Galm, E.L. Martinez-Salazar, B. Swearingen, M. Torriani, A. Klibanski, M.A. Bredella, N.A. Tritos, MRI texture analysis as a predictor of tumor recurrence

or progression in patients with clinically non-functioning pituitary adenomas, Eur. J. Endocrinol. 179 (3) (2018) 191–198.
[11] M. Birkhahn, A.P. Mitra, R.J. Cote, Molecular markers for bladder cancer: the road to a multimarker approach, Expert Rev. Anticancer Ther. 7 (12) (2007) 1717–1727.
[12] 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.
[13] P. Lambin, R.T.H. Leijenaar, T.M. Deist, J. Peerlings, E.E.C. de Jong, J. van Timmeren, S. Sanduleanu, R. Larue, A.J.G. Even, A. Jochems, Y. van Wijk, H. Woodruff, J. van Soest, T. Lustberg, E. Roelofs, W. van Elmpt, A. Dekker, F.M. Mottaghy, J.E. Wildberger, S. Walsh, Radiomics: the bridge between medical imaging and personalized medicine, Nat. Rev. Clin. Oncol. 14 (12) (2017) 749–762.
[14] 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.
[15] Z. Liu, S. Wang, D. Dong, J. Wei, C. Fang, X. Zhou, K. Sun, L. Li, B. Li, M. Wang, J. Tian, The applications of radiomics in precision diagnosis and treatment of oncology: opportunities and challenges, Theranostics 9 (5) (2019) 1303–1322.
[16] A. Giustina, P. Chanson, M.D. Bronstein, A. Klibanski, S. Lamberts, F.F. Casanueva, P. Trainer, E. Ghigo, K. Ho, S. Melmed, G. Acromegaly Consensus, A consensus on criteria for cure of acromegaly, J. Clin. Endocrinol. Metab. 95 (7) (2010) 3141–3148.
[17] L. Katznelson, E.R. Laws Jr., S. Melmed, M.E. Molitch, M.H. Murad, A. Utz, J.A. Wass, S. Endocrine, Acromegaly: an endocrine society clinical practice guideline, J. Clin. Endocrinol. Metab. 99 (11) (2014) 3933–3951.
[18] B.M. Biller, A.B. Grossman, P.M. Stewart, S. Melmed, X. Bertagna, J. Bertherat, M. Buchfelder, A. Colao, A.R. Hermus, L.J. Hofland, A. Klibanski, A. Lacroix, J.R. Lindsay, J. Newell-Price, L.K. Nieman, S. Petersenn, N. Sonino, G.K. Stalla, B. Swearingen, M.L. Vance, J.A. Wass, M. Boscaro, Treatment of adrenocorticotropin-dependent Cushing's syndrome: a consensus statement, J. Clin. Endocrinol. Metab. 93 (7) (2008) 2454–2462.
[19] K.I. Alexandraki, G.A. Kaltsas, A.M. Isidori, H.L. Storr, F. Afshar, I. Sabin, S.A. Akker, S.L. Chew, W.M. Drake, J.P. Monson, G.M. Besser, A.B. Grossman, Long-term remission and recurrence rates in Cushing's disease: predictive factors in a single-centre study, Eur. J. Endocrinol. 168 (4) (2013) 639–648.
[20] L.K. Nieman, B.M. Biller, J.W. Findling, M.H. Murad, J. Newell-Price, M.O. Savage, A. Tabarin, S. Endocrine, Treatment of Cushing's syndrome: an endocrine society clinical practice guideline, J. Clin. Endocrinol. Metab. 100 (8) (2015) 2807–2831.
[21] S. Kharangarh, H. Sandhu, S. Tangadpalliwar, P. Garg, Predicting inhibitors for multidrug resistance associated Protein-2 transporter by machine learning approach, Comb. Chem. High Throughput Screen. 21 (8) (2018) 557–566.
[22] Y. Ren, X. Zhang, W. Rui, H. Pang, T. Qiu, J. Wang, Q. Xie, T. Jin, H. Zhang, H. Chen, Y. Zhang, H. Lu, Z. Yao, J. Zhang, X. Feng, Noninvasive prediction of IDH1 mutation and ATRX expression loss in low-grade gliomas using multiparametric MR radiomic features, J. Magn. Reson. Imaging 49 (3) (2019) 808–817.
[23] Q. Tian, L.F. Yan, X. Zhang, X. Zhang, Y.C. Hu, Y. Han, Z.C. Liu, H.Y. Nan, Q. Sun, Y.Z. Sun, Y. Yang, Y. Yu, J. Zhang, B. Hu, G. Xiao, P. Chen, S. Tian, J. Xu, W. Wang, G.B. Cui, Radiomics strategy for glioma grading using texture features from multiparametric MRI, J. Magn. Reson. Imaging 48 (6) (2018) 1518–1528.
[24] S.M. Erturk, Receiver operating characteristic analysis, AJR Am. J. Roentgenol. 197 (4) (2011) W784; author reply W785.
[25] W. Pan, Akaike's information criterion in generalized estimating equations, Biometrics 57 (1) (2001) 120–125.
[26] O. Hyder, H. Marques, C. Pulitano, J.W. Marsh, S. Alexandrescu, T.W. Bauer, T.C. Gamblin, G.C. Sotiropoulos, A. Paul, E. Barroso, B.M. Clary, L. Aldrighetti, C.R. Ferrone, A.X. Zhu, I. Popescu, J.F. Gigot, G. Mentha, S. Feng, T.M. Pawlik, A nomogram to predict long-term survival after resection for intrahepatic cholangiocarcinoma: an Eastern and Western experience, JAMA Surg. 149 (5) (2014) 432–438.
[27] A.A. Kramer, J.E. Zimmerman, Assessing the calibration of mortality benchmarks in critical care: the Hosmer-Lemeshow test revisited, Crit. Care Med. 35 (9) (2007) 2052–2056.
[28] A.J. Vickers, A.M. Cronin, E.B. Elkin, M. Gonen, Extensions to decision curve analysis, a novel method for evaluating diagnostic tests, prediction models and molecular markers, BMC Med. Inform. Decis. Mak. 8 (2008) 53.
[29] A. Buliman, L.G. Tataranu, V. Ciubotaru, T.L. Cazac, C. Dumitrache, The multimodal management of GH-secreting pituitary adenomas: predictive factors, strategies and outcomes, J. Med. Life 9 (2) (2016) 187–192.
[30] C. Brichard, E. Costa, E. Fomekong, D. Maiter, C. Raftopoulos, Outcome of transsphenoidal surgery for cushing disease: a single-center experience over 20 years, World Neurosurg. 119 (2018) e106–e117.
[31] J.J. Jacob, J.S. Bevan, Should all patients with acromegaly receive somatostatin analogue therapy before surgery and, if so, for how long? Clin. Endocrinol. (Oxf.) 81 (6) (2014) 812–817.
[32] L. Duan, H. Zhu, B. Xing, F. Gu, Prolonged preoperative treatment of acromegaly with Somatostatin analogs may improve surgical outcome in patients with invasive pituitary macroadenoma (Knosp grades 1-3): a retrospective cohort study conducted at a single center, BMC Endocr. Disord. 17 (1) (2017) 55.
[33] C.B. Newman, S. Melmed, A. George, D. Torigian, M. Duhaney, P. Snyder, W. Young, A. Klibanski, M.E. Molitch, R. Gagel, L. Sheeler, D. Cook, W. Malarkey, I. Jackson, M.L. Vance, A. Barkan, L. Frohman, D.L. Kleinberg, Octreotide as primary therapy for acromegaly, J. Clin. Endocrinol. Metab. 83 (9) (1998) 3034–3040.
[34] Z. Liu, Z. Li, J. Qu, R. Zhang, X. Zhou, L. Li, K. Sun, Z. Tang, H. Jiang, H. Li, Q. Xiong, Y. Ding, X. Zhao, K. Wang, Z. Liu, J. Tian, Radiomics of multi-parametric MRI for pretreatment prediction of pathological complete response to neoadjuvant chemotherapy in breast cancer: a multicenter study, Clin. Cancer Res. (2019).
[35] J. Ding, Z. Xing, Z. Jiang, J. Chen, L. Pan, J. Qiu, W. Xing, CT-based radiomic model predicts high grade of clear cell renal cell carcinoma, Eur. J. Radiol. 103 (2018) 51–56.
[36] C. Shen, Z. Liu, Z. Wang, J. Guo, H. Zhang, Y. Wang, J. Qin, H. Li, M. Fang, Z. Tang, Y. Li, J. Qu, J. Tian, Building CT radiomics based nomogram for preoperative esophageal cancer patients lymph node metastasis prediction, Transl. Oncol. 11 (3) (2018) 815–824.
[37] C. Shen, Z. Liu, M. Guan, J. Song, Y. Lian, S. Wang, Z. Tang, D. Dong, L. Kong, M. Wang, D. Shi, J. Tian, 2D and 3D CT radiomics features prognostic performance comparison in non-small cell lung cancer, Transl. Oncol. 10 (6) (2017) 886–894.
[38] S. Zhang, G. Song, Y. Zang, J. Jia, C. Wang, C. Li, J. Tian, D. Dong, Y. Zhang, Non-invasive radiomics approach potentially predicts non-functioning pituitary adenomas subtypes before surgery, Eur. Radiol. 28 (9) (2018) 3692–3701.
[39] J. Niu, S. Zhang, S. Ma, J. Diao, W. Zhou, J. Tian, Y. Zang, W. Jia, Preoperative prediction of cavernous sinus invasion by pituitary adenomas using a radiomics method based on magnetic resonance images, Eur. Radiol. 29 (3) (2019) 1625–1634.
[40] J.F. Bonneville, F. Bonneville, F. Cattin, Magnetic resonance imaging of pituitary adenomas, Eur. Radiol. 15 (3) (2005) 543–548.
[41] J.F. Bonneville, Magnetic resonance imaging of pituitary tumors, Front. Horm. Res. 45 (2016) 97–120.