Eur Radiol
https://doi.org/10.1007/s00330-017-5180-6

ONCOLOGY

Non-invasive radiomics approach potentially predicts non-functioning pituitary adenomas subtypes before surgery

Shuaitong Zhang[1,2] · **Guidong Song**[3] · **Yali Zang**[1,2] · **Jian Jia**[3] · **Chao Wang**[1,4] · **Chuzhong Li**[3,5] · **Jie Tian**[1,2,6] · **Di Dong**[1,2] · **Yazhuo Zhang**[3,5]

Received: 10 July 2017 / Revised: 2 November 2017 / Accepted: 6 November 2017

Abstract

Purpose To make individualised preoperative prediction of non-functioning pituitary adenoma (NFPAs) subtypes between null cell adenomas (NCAs) and other subtypes using a radiomics approach.

Methods We enrolled 112 patients (training set: n = 75; test set: n = 37) with complete T1-weighted magnetic resonance imaging (MRI) and contrast-enhanced T1-weighted MRI (CE-T1). A total of 1482 quantitative imaging features were extracted from T1 and CE-T1 images. Support vector machine trained a predictive model that was validated using a receiver operating characteristics (ROC) analysis on an independent test set. Moreover, a nomogram was constructed incorporating clinical characteristics and the radiomics signature for individual prediction.

Results T1 image features yielded area under the curve (AUC) values of 0.8314 and 0.8042 for the training and test sets, respectively, while CE-T1 image features provided no additional contribution to the predictive model. The nomogram incorporating sex and the T1 radiomics signature yielded good calibration in the training and test sets (concordance index (CI) = 0.854 and 0.857, respectively).

Conclusion This study focused on the preoperative prediction of NFPA subtypes between NCAs and others using a radiomics approach. The developed model yielded good performance, indicating that radiomics had good potential for the preoperative diagnosis of NFPAs.

Key points

- *MRI may help in the pre-operative diagnosis of NFPAs subtypes*
- *Retrospective study showed T1-weighted MRI more useful than CE-T1 in NCAs diagnosis*
- *Treatment decision making becomes more individualised*
- *Radiomics approach had potential for classification of NFPAs*

Shuaitong Zhang and Guidong Song contributed equally to this work.

Electronic supplementary material The online version of this article (https://doi.org/10.1007/s00330-017-5180-6) contains supplementary material, which is available to authorized users.

✉ Jie Tian
jie.tian@ia.ac.cn

✉ Di Dong
di.dong@ia.ac.cn

✉ Yazhuo Zhang
zyz2004520@yeah.net

[1] CAS Key Laboratory of Molecular Imaging, Institute of Automation, Chinese Academy of Sciences, Beijing 100190, China

[2] University of Chinese Academy of Sciences, Beijing 100080, China

[3] Beijing Neurosurgical Institute, Capital Medical University, Beijing 100050, China

[4] Electrical Engineering School, Harbin University of Science & Technology, Harbin 150040, China

[5] Department of Neurosurgery, Beijing Tiantan Hospital, Capital Medical University, Beijing 100050, China

[6] CAS Center for Excellence in Brain Science and Intelligence Technology, Institute of Automation, Chinese Academy of Sciences, Beijing 100190, China

Keywords Non-functioning pituitary adenomas · Null cell adenomas · Radiomics · Support vector machine · Nomograms

Abbreviations

NFPAs	Non-functioning pituitary adenomas
NCAs	Null cell adenomas
CE-T1	Contrast-enhanced T1-weighted
SVM	Support vector machine
AUC	Area under the curve
ROC	Receiver operating characteristic
ICCs	Inter-observer correlation coefficients
mRMR	minimum-Redundancy Maximum-Relevancy
BIC	Bayesian information criterion
C-index	Concordance index
NRI	Net reclassification improvement

Introduction

Pituitary adenomas account for 15–20% of all intracranial neoplasms, with an incidence of 80–90 cases per 100,000 population [1–3]. Non-functioning pituitary adenomas (NFPAs) are not associated with any clinical hormonal syndrome and constitute nearly one third of pituitary adenomas [4, 5]. Because of hormonal inactivity, NFPAs are usually macroadenomas at the time of diagnosis, and growing macroadenomas can compress adjacent structures and result in headache, visual field defects, and/or various degrees of hypopituitarism [6, 7]. As a highly heterogeneous group, NFPAs can be divided into null cell adenomas (NCAs), oncocytomas, and gonadotrophic adenomas based on ultrastructural and immunohistochemical characteristics [6]. Yet, there is no preoperative diagnostic method for distinguishing these subtypes.

Previous studies have evaluated the effectiveness of traditional radiotherapy for different NFPAs subtypes. Hakan et al. concluded that radiotherapy was only effective for NCAs and invasive somatotropinomas [8]. Breen et al. [9] showed a similar result, indicating that NCAs were more sensitive to radiotherapy than oncocytomas; however, the use of radiotherapy is restricted by potential complications such as hypopituitarism [10, 11]. New radiotherapy techniques such as stereotactic radiotherapy have demonstrated improved safety [12, 13]. Importantly, the current diagnosis of NFPAs subtypes is performed postoperatively by electron microscopy, limiting the early implementation of radiotherapy. The preoperative diagnosis of NFPAs can substantially benefit patients with NCAs who are more likely to respond to neo-adjuvant radiotherapy. Addressing this problem, here we developed a predictive model to discriminate NCAs from others preoperatively using a radiomics approach.

Radiomics is an emerging approach that extracts a large number of image features from medical images and quantitatively evaluates potential associations between different objects [14, 15]. The effectiveness of this approach has been validated for several tumour types; Huang et al. [16] focused on the preoperative prediction of lymph node metastasis in colorectal cancer and yielded a concordance index (C-index) of 0.778, which was a satisfactory outcome compared with the low true positive rate of lymph node metastasis clinically. Radiomics approaches have also been validated in non-small cell lung cancer [17, 18], prostate cancer [19], and rectal cancer [20].

In the present study, we sought to develop and validate a predictive model for the preoperative prediction of NFPAs subtypes (i.e., NCAs and other subtypes) using T1-weighted (T1) and contrast-enhanced T1-weighted (CE-T1) magnetic resonance imaging (MRI). To provide a more convenient and individualised predictive tool for clinicians, a quantitative nomogram was constructed incorporating clinical features and the radiomics signature.

Materials and methods

Estimation of sample sizes for the training and test sets

To develop a stable predictive model, an adequate number of training samples is required. First, we estimated the robustness of the feature selection algorithm to the training set size. Stratified random sampling was performed to generate sample sets with different sizes. A similarity index was used to measure the similarity between two sets of feature selection results produced by the same feature selection algorithm under training data variations. Second, we estimated the robustness of the SVM model to the training set size. Stratified random sampling was performed to generate sample sets with different sizes, based on which, our-fold cross-validation was performed with SVM, and a permutation test was performed to determine the minimal training set size. We found that 60 cases were sufficient to learn a stable model in this study. While for the test sample size, we performed a power calculation and found that the minimum sample size is 23. In this study, 112 patients (75 training data and 37 test data) were enrolled, which is sufficient. Details of the estimation process are shown in Supplementary S6.

Patients

Informed consent was waived and ethical approval was provided for this retrospective study by the institutional review board of Beijing Tiantan Hospital Affiliated to Capital Medical University. We retrospectively enrolled patients diagnosed as NFPAs between January 2011 and April 2016.

Applying the inclusion and exclusion criteria (Supplementary S1), a total of 112 patients were finally included in the study. These patients were sorted according to the MRI acquisition time, and then assigned to the training set and the test set at a ratio of 2:1. The training set included 75 patients who were imaged before June 2014 (inclusive) and was used for imaging signatures building and model development, while the test set included 37 patients with MRI acquired after June 2014 and was used for model validation.

Quantitative analysis of MR images

The MR imaging protocol consists of unenhanced T1-weighted and T2-weighted images and contrast-enhanced T1-weighted images, and the details of T1 and CE-T1 imaging acquisition parameters are shown in Supplementary S2. The flow of our quantitative images analysis included three steps: region of interest delineation, feature extraction, and analysis of inter-observer reproducibility. Tumour segmentation was conducted on T1 and CE-T1 MRI for all patients by an experienced radiologist (Reader 1, with a ten-year experience in the study of pituitary adenomas) using ITK-SNAP software (University of Pennsylvania, www.itksnap.org).

A total of 741 quantitative features were automatically extracted from delineated tumours for each set of T1 and CE-T1 images to describe tumour phenotype characteristics. These features were divided into four groups: (I) tumour image intensity [21], (II) shape and size features, (III) textual features [22–26], and (IV) wavelet features. Features are shown in Supplementary S3**.** Feature generation programming was implemented in Matlab 2012a.

To assess the stability of the identified features for delineation inaccuracies, 50 patients were randomly selected from the whole sample and another experienced radiologist (Reader 2, with an eight-year experience in the study of pituitary adenomas) delineated tumour regions on both T1 and CE-T1 images. The same feature generation procedure was then repeated on these delineated tumours and inter-observer correlation coefficients (ICCs) were computed to evaluate the agreement of all quantitative features extracted from T1 and CE-T1 images.

Radiomics signature building and predictive model development

We first normalised both T1 features and CE-T1 features to the range of [−1, 1]; subsequent feature selection and model training were based on these normalised features. Then, a support vector machine (SVM) was fit with the imaging features in the training set for T1, CE-T1, and T1&CE-T1 images, respectively. To avoid over-fitting caused by high degrees of redundancy and irrelevance, we applied the minimum-redundancy maximum-relevancy (mRMR) feature selection algorithm based on mutual information [27] to generate the rank of all features for T1, CE-T1, and T1&CE-T1 images. This features rank identified the best combination of features for optimally characterising the target rather than ranking individual best features. To mitigate bias and improve the robustness of our results, a stratified random sampling approach was performed; for 250 times we applied the mRMR feature selection algorithm based on the randomly selected samples, thus generating 250 different feature ranks for T1, CE-T1, and T1&CE-T1 images. Next, these different ranks of features were aggregated into an ensemble feature rank using Borda [28]. We assumed that the top ten features contained sufficient information and assigned weights of 10, 9, 8, …, 1 to the top ten features, respectively. After obtaining the aggregated feature rank, we regarded the top 20 features as a candidate feature selection result. The flow chart of feature selection was shown in Fig. 1.

Because mRMR feature selection ranks all input features, there is no indication about the size of feature set required for optimal classification performance [29]. To select the best feature set size from the candidate obtained above, a four fold cross-validation was conducted in the training set where a support vector machine with the radial basis function kernel (RBF-SVM) [30] was used as a classifier. To avoid over-fitting in the training set, we utilised the Bayesian information criterion (BIC) calculated by the average accuracy in the four fold cross-validation to select the optimal feature set size.

We next developed the predictive model in the entire training set using RBF-SVM for T1, CE-T1, and T1&CE-T1 images. Finally, we obtained radiomics signatures corresponding to optimal feature set sizes and predictive models for classifying NCAs and other NFPA subtypes using T1, CE-T1, and T1&CE-T1 images.

Validation of the radiomics model

The performance of the three final predictive radiomics models were validated in a completely independent test set. Prediction accuracy, area under the curve (AUC), sensitivity, and specificity were calculated for both the training set and validation set. ROC curves were also plotted to illustrate diagnostic performance. A Net Reclassification Improvement (NRI) was calculated to measure the prediction increment of CE-T1 MR imaging features. Finally, the best predictive model was determined by performance comparison and the radiomics signature corresponding to the best predictive model was selected as the final signature for subsequent use.

Development and validation of individualised predictive nomogram

To provide a more individualised predictive model, a nomogram based on the multivariate logistic regression

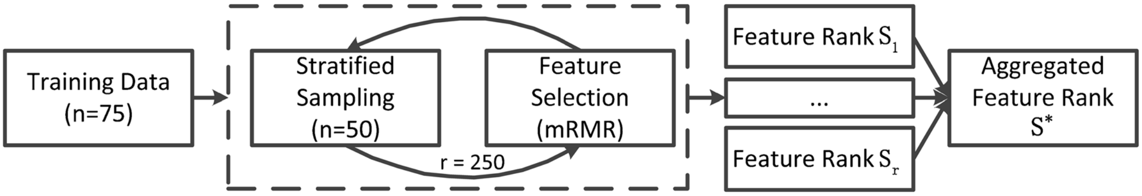


Fig. 1 The flow chart of ensemble feature selection

combining both the final radiomics signature and clinical characteristics (continuous variables: age and tumour volume; and categorical variables: sex and Knosp grade) were built in the training set. The calibration curve was plotted for the constructed individual nomogram and the Hosmer-Lemeshow test was conducted to assess similarity between the predicted and expected probabilities. The performance of the nomogram was first assessed in the training set and then validated in the completely independent test set. C-indices and 95% confidence intervals (CIs) were computed in both the training and test sets to quantify performance of the nomogram.

Statistical analysis

All statistical analyses were conducted using PASW Statistics 18.0.0 (SPSS Company). Continuous variables such as age and tumour volume were presented as the mean ± standard deviation. Categorical variables such as sex and Knosp grade were presented as the number (percentage). Group differences were analysed for using Student's t test or Pearson χ^2 test, and a two-sided p-value < 0.05 was considered to be statistically significant.

Results

Patient characteristics

Of 112 patients included in the study (age, 50.11 ± 11.85 years), 46 (45.1%) were diagnosed as NCAs and 66 (54.9%) as other NFPAs subtypes based on electron microscopy findings. Patient characteristics in the training and test sets are listed in Table 1. There were no significant differences between training set and validation set in age (p = 0.328) or sex (p = 0.185). This indicated the justifiability using these two sets as training set and test set. Univariate analysis of clinical characteristics in the training and test sets is shown in Table 2.

Satisfactory inter-observer reproducibility was achieved for both T1 and CE-T1 imaging features, with a calculated ICC range of 0.752–0.951 for T1 features and 0.780–0.973 for CE-T1 features. Subsequent processes were based on the segmentation results of the Reader 1, who had a longer work experience than Reader 2.

Radiomics signature building and predictive model development

In the phase of feature selection, the optimal feature set size was identified by averaging accuracy values across a four fold cross-validation and corresponding BIC in the training set with SVMs. Figure 2A shows the relationship between average accuracy and the number of features for T1, CE-T1, and T1&CE-T1 images. Figure 2B shows the relationship between BIC value and the number of features for T1, CE-T1, and T1&CE-T1 images. The top three features of T1 candidates were selected in accordance with BIC value as the radiomics signature for T1 images; the top two features were selected for CE-T1 images; and top three features were selected for T1&CE-T1 images. For T1 images, the three features selected were wavelet HH 'Informational Measure of Correlation 1_45°' (HH_IMC1_45°), 'Sphericity' and wavelet HL 'Informational Measure of Correlation 2_45°' (HL_IMC2_45°). For CE-T1 images, the two features selected were 'Low Grey Level Run Emphasis_135°' (LGLRE_135°) and wavelet LH 'Informational Measure of Correlation 1_0°' (LH_IMC1_0°). For T1&CE-T1 images, the three features selected were the same as those selected from T1 images. This suggested better performance of the T1 radiomics signature than the CE-T1 radiomics signature and equal performance of the T1 and T1&CE-T1 radiomics signatures. The details of these selected features are shown in Supplementary S3.

Validation of the predictive model

The performance of each predictive model was first assessed in the entire training set and then validated in a completely independent test set. Both the T1 and T1&CE-T1 predictive models yielded AUC values of 0.8314 in the training set and 0.8042 in the test set; the CE-T1 predictive model yielded an AUC value of 0.634 in the training set and 0.510 in the test set. To quantitatively evaluate the prediction increment of CE-T1 images, NRI was calculated to be 0, indicating that the 741 CE-T1 imaging features had no additional contribution to the prediction of NCAs based on T1 imaging features. Because the T1&CE-T1 model was the same as T1 model, we only showed results for the T1 and CE-T1 models. AUC, sensitivity, specificity, and accuracy of T1 and CE-T1 models are shown in Table 3**.** ROC curve for each predictive model are

Table 1 Clinical characteristics of patients (*n*=112)

Characteristic	Training Set (n=75)	Test Set (n=37)	Whole Set (n=112)	P-value
Subtype(No.)				0.087†
NCAs	35 (57.33%)	11 (29.73%)	46 (41.07%)	
Others	40 (42.67%)	26 (70.27%)	66 (58.93%)	
Age (yr, mean ± std)	49.31 ± 12.54	49.31 ± 10.35	50.11 ± 11.85	0.328*
Gender(No.)				0.185†
Male	43 (57.33%)	26 (70.27%)	69 (61.61%)	
Female	32 (42.67%)	11 (29.73%)	43 (38.39%)	
Tumor Volume (cm^3, mean ± std)	9.21 ± 9.88	6.69 ± 6.45	8.35 ± 8.90	0.176*
Knosp Grade (No [%])				0.085†
Grade 0-2	43 (57.33%)	29 (78.38%)	72 (64.29%)	
Grade 3-4	32 (42.67%)	8 (21.62%)	40 (35.71%)	

Note: yr, year; std, standard deviation. P-value < 0.05 was considered as a significant difference

*Student's t test

†Pearson's test

shown in Fig. 3. Finally, the T1 predictive model was chosen as the final predictive model for distinguishing NCAs from other NFPA subtypes. The formula of the final model is shown in Supplementary S4**.** To show the model more clearly, we also plotted it in three-dimensional space in Fig. 4.

Development and validation of individualised predictive nomogram

Incorporating sex and the radiomics signature, a nomogram was developed in the training set (Fig. 5A). The calibration curve of the training set showed good agreement between the predicted and expected probabilities for other NFPA subtypes (Fig. 5B), and the Hosmer-Lemeshow test showed good similarity (p = 0.286). The calibration curve of the test set also confirmed good model performance (Fig. 5C). The C-index of the nomogram was 0.8536 (95% CI = 0.833–0.874) in the training set and 0.857 (95% CI = 0.816–0.897) in the test set.

Discussion

To date, few studies have tried to discriminate NFPAs subtypes using MR images. In the present study, we employed a quantitative radiomics approach to provide preoperative

Table 2 Univariate analysis of clinical characteristics of patients in the training set and test set

Characteristic	Training Set (N = 75)		P-value	Test Set (N = 37)		P-value
	NCAs	Others		NCAs	Others	
Age (yr, mean±std)	47.85 ± 12.22	50.65 ± 12.83	0.352*	47.85 ± 10.89	53.80 ± 9.57	0.065*
Gender (No [%])			0.005†			0.032†
Male	14 (40.0%)	29 (72.5%)		5 (45.5%)	21 (80.8%)	
Female	21 (60.0%)	11 (27.5%)		6 (54.5%)	5 (19.2%)	
Tumor Volume (cm^3, mean ± std)	10.89 ± 12.08	7.38 ± 6.43	0.149*	5.00 ± 3.12	7.47 ± 7.43	0.299*
Knosp Grade (No [%])			0.465†			0.124†
Grade 0-2	19 (54.29%)	24 (60.00%)		7 (63.64%)	22 (84.62%)	
Grade 3-4	16 (45.71%)	16 (40.00%)		4 (36.36%)	4 (15.38%)	

Note: yr, year; std, standard deviation. P-value < 0.05 was considered as a significant difference

*Student's t test

†Pearson's test

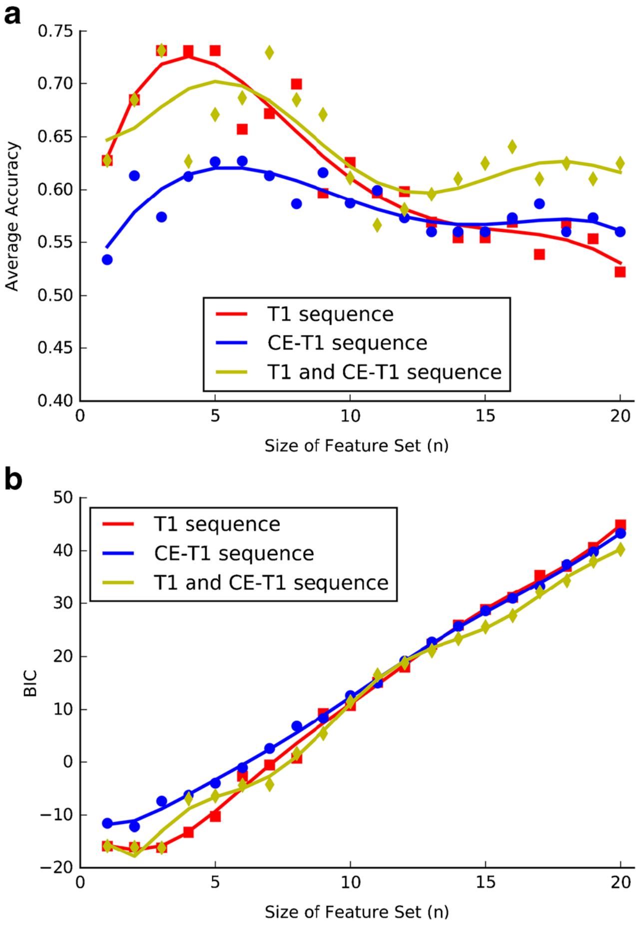


Fig. 2 Choose the best size of feature set. (A) The relationship between the size of feature set and the average of performance in the training set for T1, CE-T1 and combination of T1 and CE-T1 MR images. (B) The relationship between the size of feature set and BIC in the training set for T1, CE-T1 and combination of T1 and CE-T1 MR images. Note: The smooth curves were fitted based on those points

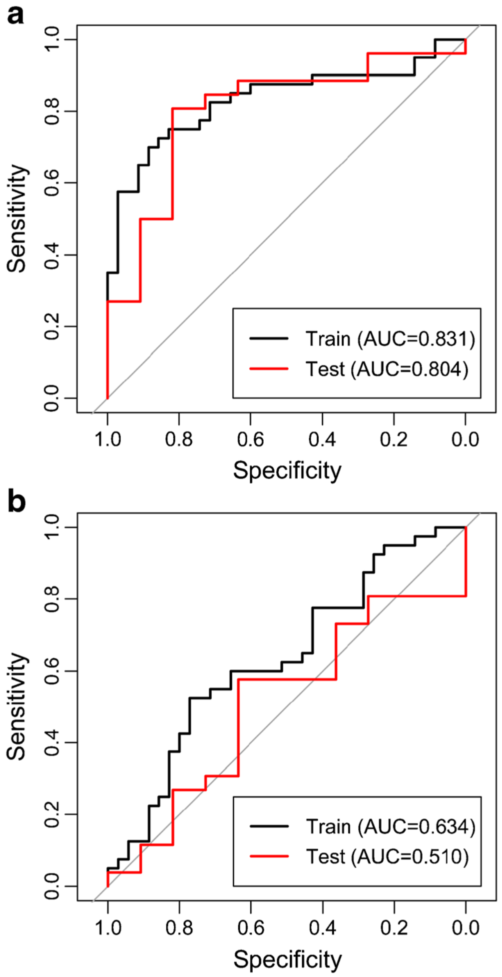


Fig. 3 The ROC curves of T1 signature (A) and CE-T1 signature (B) in the training set and test set

prediction of NFPAs subtypes, specifically distinguishing between NCAs and other subtypes, using T1 and CE-T1 images.

In our cohort, there were more male patients with NFPAs (*n* = 69, 61.61%) than female patients (*n* = 43, 38.39%), consistent with the findings of Fernandez et al. [3] and Greenman et al. [6]. Additionally, NCAs were slightly more common in female patients (*n* = 27, 58.7%) than in male patients (*n*=19, 41.3%), while other subtypes were more common in male patients (*n* = 50, 75.76%) than in female patients (*n* = 16, 24.24%). These findings were in agreement with those of Feng et al. [32] and Blogun et al. [33], but diverged slightly from that of Schaller [31], who found that NCAs were slightly more prevalent in males; this discrepancy may be related to

Table 3 The performance of models based on T1 weighted MR images and contrast enhanced T1 weighted MR images

Variables & models	Training set				Test set			
	AUC (95%CI)	SEN	SPE	ACC	AUC (95%CI)	SEN	SPE	ACC
T1 Radiomics Signature	0.8314 (0.8290, 0.8354)	0.750	0.829	0.787	0.8042 (0.7905, 0.8009)	0.808	0.818	0.811
CE-T1 Radiomics Signature	0.634 (0.6338, 0.6418)	0.550	0.714	0.627	0.510 (0.5016, 0.5148)	0.577	0.455	0.541

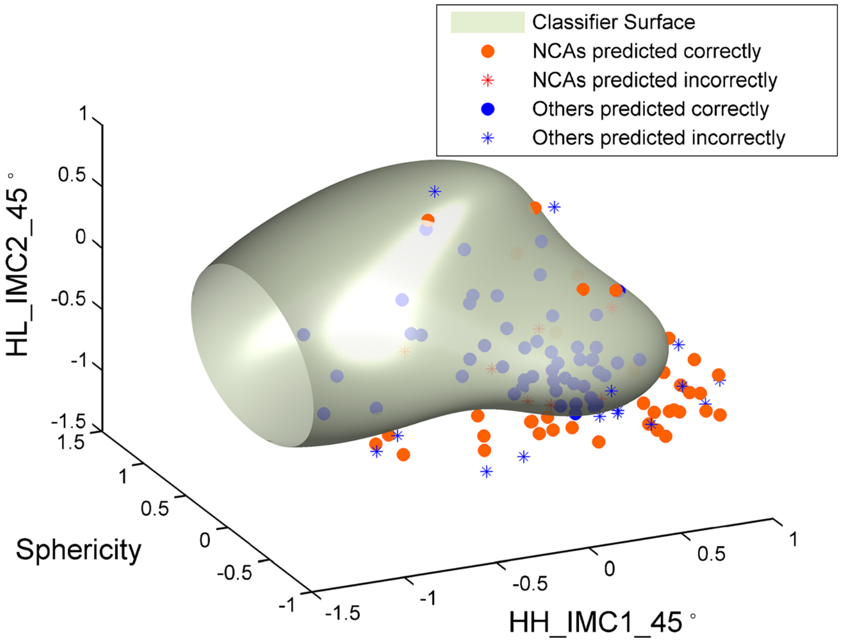


Fig. 4 The radiomics predictive model. This model was plotted to facilitate comprehension. The model was built based on T1 radiomics signature using SVMs. The x, y and z axis represents the feature named HH_IMC1_45°, Sphericity and HL_IMC2_45° respectively. These three features were normalized to the range of -1 to 1. The grey cream surface represents the classifier surface. The orange points outside were predicted as NCAs, while the blue points inside were predicted as others. The orange solid points represent the patients with NCAs that were predicted correctly; the blue solid points represent the patients with others that were predicted correctly. The patients with NCAs which was predicted incorrectly were shown as orange asterisk; the patients with others which were predicted incorrectly were shown as blue asterisk

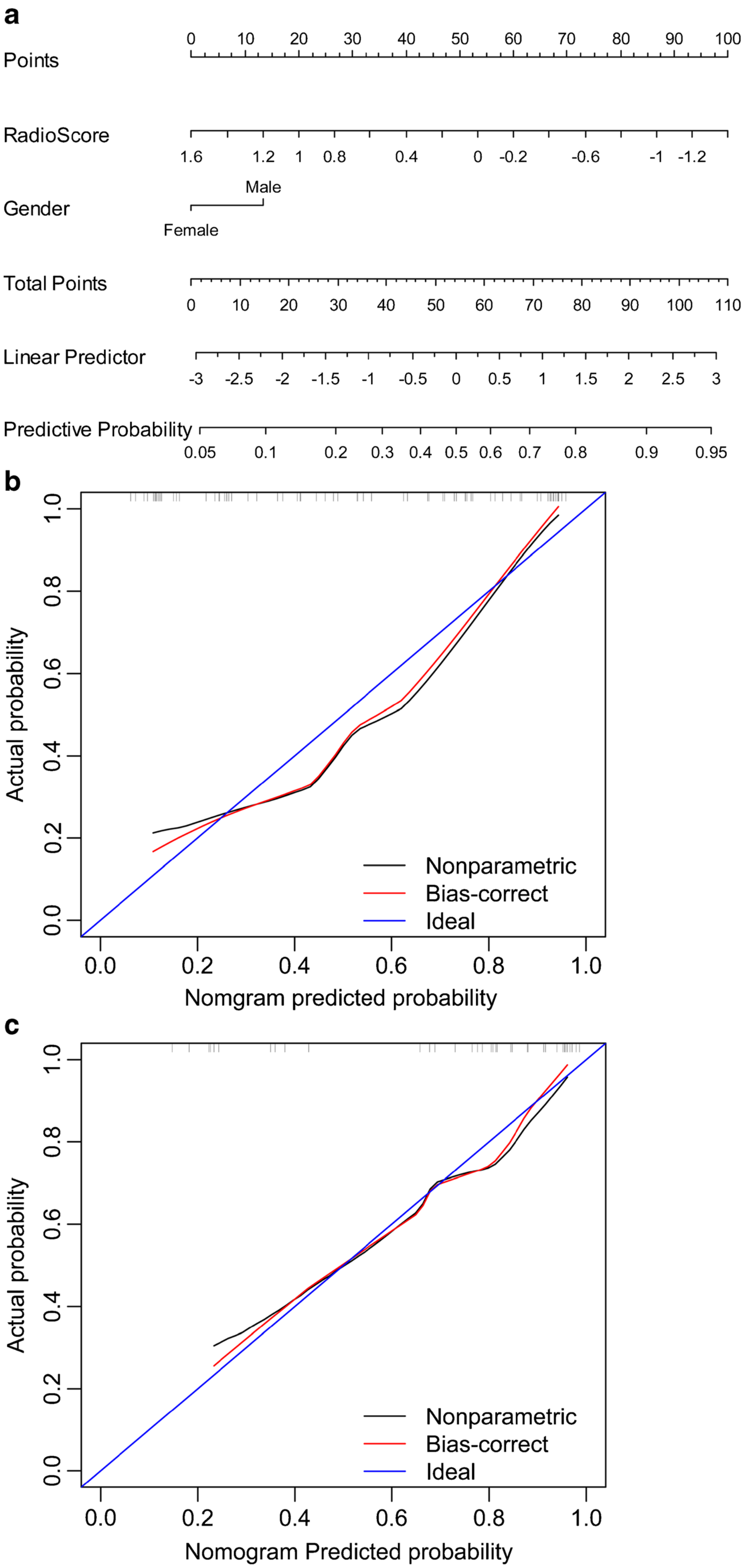
differences in cohort, races, or study time period, but need further validation.

Unlike previous work [16–20], we used ensemble feature selection to select informative features. This approach may avoid sampling bias and improve the robustness of feature selection results. When selecting optimal feature set size from candidates, we used BIC values instead of average accuracy values from the four fold cross-validation in order to enhance the generalizability of the selected features. Finally, we found that the T1 image signature provided more diagnostic value than did the CE-T1 image signature. In both the T1 and CE-T1 radiomics signatures, almost all features were textural features. We supposed that the different update rates of contrast agent Gd-DTPA between patients might influence the intensity of MR images and the CE-T1 textural features, which might result in that the selected CE-T1 features were of discriminative ability for the training set, but not for the test set. Thus, the CE-T1 radiomics model based on the selected CE-T1 features performed badly for the test set, that is, the CE-T1 model risked overfitting. Additionally, we found that NCAs tended to have smaller sphericity than other subtypes; this observation merits further explanation. We found that sphericity was positively correlated with compactness. We also plotted this relationship in our data and found exactly the same trend (Supplementary S5 Fig. 2). Terada et al. [34] found that fibrosis was more frequently identified in oncocytomas than in NCAs, which may account for increased compactness in oncocytomas and other subtypes in our study.

Our study had some limitations. First, to build the radiomics signature and predictive model, we chose sagittal pre-contrast T1 and CE-T1 images, which were usually referred to clinically; however, the combinations with other sequences such as diffusion-weighted imaging may have provided additional information and improved performance of the predictive model. Second, the dataset in this study was obtained from a single hospital. Many hospitals do not conduct electron microscopy for the post-surgical diagnosis of NFPAs subtypes due to the use of identical treatment protocols for different subtypes, such that a multicentre study was not feasible. A future multicentre trial is required to validate the performance of our model. Third, our model has limitations for clinical use. At present, surgery is the first-line therapy for the treatment of patients with NFPAs; however, the preoperative diagnosis of patients with NCAs may provide the possibility of neo-adjuvant radiotherapy for patients with NCAs that are more likely to benefit from radiotherapy.

In conclusion, we built a quantitative nomogram for preoperative prediction of NFPAs subtypes. The model was easy to

Fig. 5 Developed radiomics nomogram. (A) The nomogram was developed in the training set, incorporating radiomics signature and patients' gender. The predictive probability axis represents the probability of not being NCAs. (B) Calibration curve of the radiomics nomogram in the training set. (C) Calibration curve of the radiomics nomogram in the test set. Calibration curve presented the agreement between the predicted probability of the NCAs and ground truth. The diagonal blue line represents an ideal evaluation, while the black and red lines represent the performance of the nomogram. Closer fit to the diagonal blue line indicates a better evaluation



use for both clinicians and patients and may assist clinicians preoperatively to predict NFPAs subtypes more precisely, which indicated that analysis of medical data using a radiomics method could serve clinicians better.

Funding This study has received funding by National Key Research and Development Program of China (2017YFA0205200, 2017YFC1308700, 2106YFC0103702, 2016YFA0201401, 2017YFC1308701, 2017YFC1309100, 2016CZYD0001), National Natural Science Foundation of China (81227901, 81527805, 61231004, 81501616, 81671851), Beijing Municipal Science & Technology Commission (Z161100002616022, Z171100000117023), the Science and Technology Service Network Initiative of the Chinese Academy of Sciences (KFJ-SW-STS-160), the International Innovation Team of CAS (20140491524), the Instrument Developing Project of the Chinese Academy of Sciences (YZ201502) and National High Technology Research and Development Program of China (2015AA020504).

Compliance with ethical standards

Guarantor The scientific guarantor of this publication is Jie Tian.

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.

Statistics and biometry Dr. Di Dong from the University of Chinese Academy of Sciences, who is one of the authors, has significant statistical expertise.

Informed consent Written informed consent was waived by the Institutional Review Board of Beijing Tiantan Hospital Affiliated to Capital Medical University.

Ethical approval Institutional Review Board approval was obtained.

Methodology
- retrospective
- diagnostic or prognostic study
- performed at one institution

References

1. Melmed S (2011) Pathogenesis of pituitary tumors. Nat Rev Endocrinol 7:257–266
2. Daly AF, Rixhon M, Adam C, Dempegioti A, Tichomirowa MA, Beckers A (2006) High Prevalence of Pituitary Adenomas: A Cross-Sectional Study in the Province of Liège, Belgium. J Clin Endocrinol Metab 91:4769–4775
3. Fernandez A, Karavitaki N, Wass JAH (2010) Prevalence of pituitary adenomas: a community-based, cross-sectional study in Banbury (Oxfordshire, UK). Clin Endocrinol 72: 377–382
4. Asa SL, Ezzat S (2009) The pathogenesis of pituitary tumors. Annu Rev Pathol-Mech 4:97–126
5. Wass JAH, Reddy R, Karavitaki N (2011) The postoperative monitoring of nonfunctioning pituitary adenomas. Nat Rev Endocrinol 7:431–434
6. Greenman Y, Stern N (2009) Non-functioning pituitary adenomas. Best Pract Res Clin Endocrinol Metab 23:625–638
7. Jaffe CA (2006) Clinically non-functioning pituitary adenoma. Pituitary 9:317–321
8. Oruçkaptan HH, Senmevsim Ö, Özcan OE, Özgen T (2000) Pituitary adenomas: results of 684 surgically treated patients and review of the literature. Surg Neurol 53:211–219
9. Breen P, Flickinger JC, Kondziolka D, Martinez AJ (1998) Radiotherapy for nonfunctional pituitary adenoma: analysis of long-term tumor control. J Neurosurg 89:933–938
10. Hashimoto N, Handa H, Yamashita J, Yamagami T (1986) Long-term follow-up of large or invasive pituitary adenomas. Surg Neurol 25:49–54
11. AL-Mefty O, Kersh JE, Routh A, Smith RR (1990) The long-term side effects of radiation therapy for benign brain tumors in adults. J Neurosurg 73:502–512
12. Kanner AA, Corn BW, Greenman Y (2009) Radiotherapy of non-functioning and gonadotroph adenomas. Pituitary 12:15–22
13. Castinetti F, Régis J, Dufour H, Brue T (2010) Role of stereotactic radiotherapy in the management of pituitary adenomas. Nat Rev Endocrinol 6:214–223
14. Lambin P, Rios-Velazquez E, Leijenaar R et al (2012) Radiomics: Extracting more information from medical images using advanced feature analysis. Eur J Cancer 48:441–446
15. Aerts HJWL, Velazquez ER, Leijenaar RT et al (2014) Decoding tumor phenotype by noninvasive imaging using a quantitative radiomics approach. Nat Commun 5:4006–4013
16. Huang YQ, Liang CH, He L et al (2016) Development and validation of a radiomics nomogram for preoperative predicting lymph node metastasis in colorectal cancer. J Clin Oncol 34:2157–2164
17. Wu J, Aguilera T, Shultz D et al (2016) Early-stage non-small cell lung cancer: quantitative imaging characteristics of (18)F Fluorodeoxyglucose PET/CT allow prediction of distant metastasis. Radiology 281:270–278
18. Huang Y, Liu Z, He L et al (2016) Radiomics Signature: A Potential Biomarker for the Prediction of Disease-Free Survival in Early-Stage (I or II) Non-Small Cell Lung Cancer. Radiology 281:947–957
19. Shiradkar R, Podder TK, Algohary A, Viswanath S, Eills RJ, Madabhushi A (2016) Radiomics based targeted radiotherapy planning (Rad-TRaP): a computational framework for prostate cancer treatment planning with MRI. Radiat Oncol 11:148
20. Nie K, Shi LM, Chen Q et al (2016) Rectal cancer: assessment of neoadjuvant chemoradiation outcome based on radiomics of multiparametric MRI. Clin Cancer Res 22:5256–5264
21. White GH (1980) Skewness, kurtosis, and extreme values of northern hemisphere gerpotential heights. Mon Weather Rev 108:1446–1455
22. Haralick RM, Shanmugam K, Dinstein I (1973) Textural features for image classification. IEEE Trans Syst Man Cybern 6:610–621
23. Soh LK, Tsatsoulis C (1999) Texture analysis of SAR sea ice imagery using gray level co-occurrence matrices. IEEE T Geosci Remote 37:780–795
24. Clausi DA (2002) An analysis of co-occurrence texture statistics as a function of grey level quantization. Can J Remote Sens 28:45–62
25. Galloway MM (1975) Texture analysis using gray level run lengths. Comput Graph Image Process 4:172–179
26. Chu A, Sehgal CM, Greenleaf JF (1990) Use of gray value distribution of run lengths for texture analysis. Pattern Recogn Lett 11: 415–419
27. Peng HC, Long FH, Ding C (2005) Feature selection based on mutual information: criteria of max-dependency, max-relevance, and min-redundancy. IEEE T Pattern Anal 27:1226–1238
28. Baharad E, Nitzan S (2003) The borda rule, condorcet consistency and condorcet stability. Economic Theory 22:685–688
29. Mehmood T, Liland KH, Snipen L, Sæbø S (2012) A review of variable selection methods in partial least squares regression. Chemometr Intell Lab 118:62–69

30. Chang CC, Lin CJ (2011) LIBSVM: a library for Support Vector Machines. ACM Intel Syst Tec 2:1–27
31. Schaller B (2003) Gender-related differences in non-functioning pituitary adenomas. Neuroendocrinol Lett 24:425–430
32. Feng J, Hong LC, Wu YG et al (2014) Identification of a subtype-specific ENC1 gene related to invasiveness in human pituitary null cell adenoma and oncocytomas. J Neuro-Oncol 119:307–315
33. Balogun JA, Monsalves E, Juraschka K et al (2015) Null cell adenomas of the pituitary gland: an institutional review of their clinical imaging and behavioral characteristics. Endocr Pathol 26:63–70
34. Terada T, Kovacs K, Stefaneanu L, Horvath E (1995) Incidence, pathology, and recurrence of pituitary adenomas: study of 647 unselected surgical cases. Endocr Pathol 6:301–310