



MRI-based radiomics analysis to predict preoperative lymph node metastasis in papillary thyroid carcinoma

Wenjuan Hu¹, Hao Wang¹, Ran Wei¹, Lanyun Wang¹, Zedong Dai¹, Shaofeng Duan², Yaqiong Ge², Pu-Yeh Wu³, Bin Song¹

¹Department of Radiology, Minhang Hospital, Fudan University, Minhang District, Shanghai, China; ²GE Healthcare, China, Pudong New Town, Shanghai, China; ³GE Healthcare, MR Research China, Beijing, China

Contributions: (I) Conception and design: W Hu; (II) Administrative support: B Song; (III) Provision of study materials or patients: All authors; (IV) Collection and assembly of data: All authors; (V) Data analysis and interpretation: W Hu; (VI) Manuscript writing: All authors; (VII) Final approval of manuscript: All authors.

Correspondence to: Bin Song. Department of Radiology, Minhang Hospital, Fudan University, Minhang District, Shanghai 201199, China. Email: songbin@fudan.edu.cn.

Background: The aim of the present study was to develop a magnetic resonance imaging (MRI) radiomics model and evaluate its clinical value in predicting preoperative lymph node metastasis (LNM) in patients with papillary thyroid carcinoma (PTC).

Methods: Data of 129 patients with histopathologically confirmed PTC were retrospectively reviewed in our study (90 in training group and 39 in testing group). 395 radiomics features were extracted from T2 weighted imaging (T2WI), diffusion weighted imaging (DWI) and T1 weighted multiphase contrast enhancement imaging (T1C+) respectively. Minimum redundancy maximum relevance (mRMR) was used to eliminate irrelevant and redundant features and least absolute shrinkage and selection operator (LASSO), to additionally select an optimized features' subset to construct the radiomics signature. Predictive performance was validated using receiver operating characteristic curve (ROC) analysis, while decision curve analyses (DCA) were conducted to evaluate the clinical worth of the four models according to different sequences. A radiomics nomogram was built using multivariate logistic regression model. The nomogram's performance was assessed and validated in the training and validation cohorts, respectively.

Results: Seven key features were selected from T2WI, five from DWI, ten from T1C+ and seven from the combined images. The scores (Rad-scores) of patients with LNM were significantly higher than patients with non-LNM in both the training cohort and the validation cohort. The combined model performed better than the T2WI, DWI, and T1C+ models alone in both cohorts. In the training cohort, the area under the ROC (AUC) values of T2WI, DWI, T1C+ and combined features were 0.819, 0.826, 0.808, and 0.835, respectively; corresponding values in the validation cohort were 0.798, 0.798, 0.789, and 0.830. The clinical utility of the combined model was confirmed using the radiomics nomogram and DCA.

Conclusions: MRI radiomic model based on anatomical and functional MRI images could be used as a non-invasive biomarker to identify PTC patients at high risk of LNM, which could help to develop individualized treatment strategies in clinical practice.

Keywords: Papillary thyroid carcinoma (PTC); magnetic resonance imaging (MRI); radiomics; lymph node metastasis (LNM)

Submitted Apr 29, 2020. Accepted for publication Sep 10, 2020.

doi: 10.21037/gs-20-479

View this article at: <http://dx.doi.org/10.21037/gs-20-479>

Introduction

Papillary thyroid carcinoma (PTC) is the most common endocrine malignant neoplasm and its incidence is persistently increasing worldwide (1,2). The increasing incidence of PTC in the past few decades is partly because of the improved sensitivity of ultrasound (US)- and fine-needle aspiration biopsy (FNAB)-based PTC detection (3,4). The majority of PTC are differentiated cancers with excellent prognosis and long-term survival (5). However, lymph node metastasis (LNM) is very common in differentiated thyroid cancer and has been found in about 30–80% of PTC patients by pathologic examination (6,7). In some high-risk groups, LNM is considered a risk factor for local recurrence, distant metastases, and decreased survival rates (8-10). While prophylactic lymph node dissection (LND) for PTC patients has been recommended to improve disease-specific survival (11) and local recurrence (12), its clinical significance remains controversial. According to the 2015 American Thyroid Association (ATA) guidelines, prophylactic LND has no significant benefit for improving the long-term survival (13). Additionally, it might increase the risk of complications, such as permanent hypoparathyroidism and permanent recurrent laryngeal nerve injury (5,14). Therefore, it is of importance to preoperatively evaluate cervical LNM in PTC patients to avoid the reoperation for recurrent tumors and reduce the surgical complications.

Preoperative US and US-guided FNAB are the ATA-recommended guidelines for assessment of lymph node involvement in PTC patients (13). It is well known that US is operator-dependent. Previous studies have shown that the estimated sensitivity of US in LNM prediction is low (38–59%) (15-17). Moreover, US is also limited in assessing central lymph nodes and superior mediastinal lymph nodes (18). FNAB is an invasive approach and its sensitivity in evaluating LNM varies and is specific to the operator (19). Computed tomography (CT) is superior to US in detecting center LNM, but inferior to US in predicting lateral LNM (18). Despite having a similar sensitivity and specificity as US (20-22), CT is not recommended for preoperative diagnosis of cervical LNM in PTC patients because of iodinated contrast media use.

Advances in imaging and computational fields lead to the rapid development of radiomics in recent years (23). Radiomics refers to the high-throughput extraction of extensive quantitative features to transform medical images

into utilizable high-dimensional data that could likely be used as prognostic, diagnostic, or predictive biomarkers and support clinical decision-making (23-25). Accordingly, a series of studies have investigated the preoperative predictive value of radiomics for LNM in various types of cancer (26-29). Furthermore, preoperative US- and CT thyroid imaging-based radiomics have also been reported in recent studies to predict cervical LNM in PTC patients (30,31).

As a non-invasive and non-radiative imaging approach, magnetic resonance imaging (MRI) provides not only anatomical information with high soft tissue contrast, but also functional information such as diffusion weighted imaging (DWI). This technique produces qualitative and quantitative information about that changes in cell levels and is widely used in tumor detection, differentiation, treatment response monitoring and prognosis evaluation. DWI has been shown to offer significant diagnostic value in thyroid nodules (32). Some studies have shown that MR based radiomics can improve the prediction of LNM and cancer staging (26,33-35). However, to our best knowledge, the clinical value of MRI-based radiomics for preoperative LNM prediction in PTC patients has not yet been studied in detail.

Therefore, we aimed to develop a radiomics model based on T2-weighted imaging (T2WI), DWI, and T1-weighted multiphase contrast-enhanced images (T1C+) to evaluate its performance in the prediction of cervical LNM in patients with thyroid carcinoma. We present the following article in accordance with the STROBE reporting checklist (available at <http://dx.doi.org/10.21037/gs-20-479>).

Methods

Patients

The study was conducted in accordance with the Declaration of Helsinki (as revised in 2013). This retrospective study was approved by the institutional ethics committee of Minhang Hospital affiliated to Fudan University School of Medicine (2015_51). Individual consent for this retrospective analysis was waived. From January 2016 to February 2018, 203 patients with suspected PTC were consecutively reviewed. The inclusion criteria were as follows: (I) pathologically confirmed PTC; (II) with central lymph node dissection (CLND) and/or lateral compartmental LND procedure; (III) with preoperative MRI examination; (IV) no prior thyroid surgery, biopsy, head and neck cancer, and history of neck radiation therapy before MRI. The exclusion criteria were as follows: (I) no LND; (II) small tumor volume

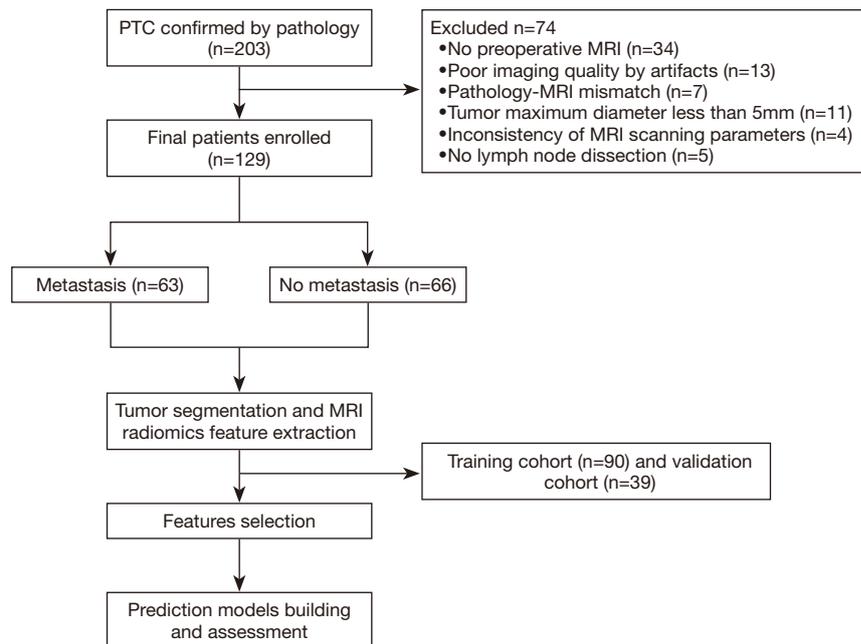


Figure 1 Flow chart of the patient selection process. PTC, papillary thyroid carcinoma; MRI, magnetic resonance imaging.

(maximum diameter <5 mm); (III) poor MR image quality; (IV) pathology-MRI mismatch; and (V) inconsistency of MR scanning parameters. Finally, 129 patients (37 male and 92 females; age: 45.01 ± 13.75 years) were included; of these, 63 patients had pathologically confirmed central LNM (Figure 1). Based on the experience of Logistic regression analysis in constructing the radiomics model, the sample size can meet the experimental requirements.

MRI acquisition

All patients underwent MRI scanning in the 1–2 weeks before surgery. Images were obtained on a 1.5T whole body scanner (EXCITE HD, GE Healthcare, Milwaukee, WI) equipped with an 8-channel special neck surface coil. The scan sequences included axial T1WI, axial T2WI, axial DWI with b values of 0 and 800 s/mm^2 , and axial T1C+. For contrast-enhanced scanning, intravenous gadolinium (Magnevist, Bayer HealthCare Pharmaceuticals, Montville, NJ) at a flow rate of 3 mL/s was administered using an automated injector. Multiphase contrast-enhanced images were acquired at 30, 60, 120, 180, 240, 300 seconds after contrast-medium injection. All patients underwent respiratory training before the MRI and were asked to hold their breath at the same level in all acquisition phases.

Tumor segmentation

All MR images were exported for tumor segmentation. The tumor region was first determined based on preoperative T2WI, DWI, and T1C+ images by two radiologists (B Song, W Hu, with 10 and 6 years of experience, respectively, in thyroid MRI) in consensus. The entire volumes of interest (VOIs) of the primary tumor on both T2WI/DWI and T1C+ images were then manually segmented along the tumor boundaries layer by layer by radiologist W HU with the ITK-SNAP software package version 3.4.0 (www.itksnap.org). For multifocal PTC, we depicted the largest lesion for further analysis.

Interobserver and intra-observer reproducibility

The intra-observer ICCs and inter-observer ICCs were calculated to estimate the reproducibility of radiomics features. Thirty cases were randomly selected and double-blinded for comparing manual segmentation by two radiologists (B Song, W Hu). To evaluate the intra-observer reproducibility, W HU used ITKSNAP (<http://www.itksnap.org>) software manually sketched the ROI, 2 weeks later, W Hu performed the second ROI manual segmentation. Moreover, B Song performed the ROI

segmentation independently to evaluate the intra-observer correlation. An ICC value of >0.75 was considered to indicate good agreement. The features with ICCs greater than 0.75 of first sketch of W Hu were retained.

Feature extraction and selection

The Artificial Intelligence Kit version 3.0.0 (GE Healthcare) was used for features extraction from T2WI, DWI, and T1C+ images. These radiomics features were divided into the following six groups: (I) gray-level histogram features; (II) shape features; (III) texture features; (IV) gray-level cooccurrence matrix (GLCM) features; (V) gray-level run length matrix (GLRLM) features; and (VI) gray-level size zone matrix (GLSZM) features.

All patients were randomly assigned to training and validation groups (ratio, 7:3). To establish radiomics prediction models in the training cohort, we used two feature selection methods to select the most informative radiomics features and avoid overfitting: minimum redundancy maximum relevance (mRMR) and least absolute shrinkage and selection operator (LASSO) algorithm. mRMR was first performed to eliminate all redundant and irrelevant features; finally, 30 features were retained. Then, LASSO was performed to further select the optimized subset of features through regularization to construct the final model in the training group. The LASSO includes choosing the regular parameter λ , coefficients of partial candidate feature were compressed to zero, non-zero coefficients features were finally retained.

Model construction

In order to better evaluate the performance of radiomics features in LNM prediction, four prediction models were constructed based on features extracted from images of T2WI, DWI, T1C+, and the combination of T2WI, DWI, and T1C+ (combined). Radiomics scores (Rad-scores) corresponding to the four models was calculated for each patient by adding the selected features weighted by their coefficients in the LASSO logistic regression model in the training cohort. The Rad-score calculation formula is presented in the *Supplementary files*.

Model validation

Receiver operating characteristic (ROC) curve analysis was used to assess the performance of the prediction model

based on the following indices: the area under the curve (AUC), sensitivity, specificity, positive predictive value (PPV), negative predictive value (NPV), and accuracy. The prediction performance of the model was further tested in the validation cohort. Finally, the decision curve was used to evaluate the model's clinical value.

Construction and assessment of radiomics nomogram model

The Rad-scores corresponding to combined model and the clinical variables were used to construct the radiomics nomogram model. Firstly, filter the clinical variables using the univariate logistic regression method to select the statistically significant feature. Next, the remaining features were entered into the multivariate logistic regression model to construct the clinical model and nomogram. The ROC analysis was calculated to quantify the discriminative performance of the nomogram.

Statistical analysis

Continuous variables were presented as mean \pm standard deviation (SD). To compare the characteristics of patients, independent two-sample *t*-test and chi-square test, respectively, were performed on continuous and categorical variables using SPSS 23.0 (IBM, Armonk, NY, USA). For all tests, $P < 0.05$ was considered statistically significant. Other statistical analyses were performed using R statistical software version 3.5.1. The 'mRMRe' package was applied to the mRMR to first screen the radiomic features. The "Glmnet" package was applied to analyze the LASSO logistic regression model. ROC curves were plotted by using "pROC" package. Decision curve analyses (DCA) was conducted using "dca.R". Nomogram construction was performed using the "rms" package. A two-sided *P* value of <0.05 was considered significant.

Results

Patient characteristics

In all, 129 patients with PTC (63 nodes positive and 69 nodes negative) were included in our study. The detailed patient characteristics are presented in *Table 1*. No significant intergroup differences were found with respect to age and sex. The tumor diameter of LNM group was larger than that of non-LNM group. Bilateral PTCs and

Table 1 Clinical characteristics of PTCs

Characteristic	PTCs with LNM		P
	No (n=66)	Yes (n=63)	
Age, years	47.85±13.36	41.82±13.41	0.835
Diameter, cm	1.01±0.41	1.51±0.69	<0.001*
Sex, n (%)			0.717
Female	48 (72.7)	44 (69.8)	
Male	18 (27.3)	19 (30.2)	
Location, n (%)			0.039*
Right lobe	36 (54.5)	26 (41.3)	
Left lobe	29 (43.9)	28 (44.4)	
Isthmus	0 (0.0)	2 (3.2)	
Bilateral	1 (1.5)	7 (11.1)	
Multifocality, n (%)			0.006*
No	62 (93.9)	48 (76.2)	
Yes	4 (6.1)	15 (23.8)	
Short diameter of LN, mm	3.92±1.93	5.02±2.06	0.002*
Long diameter of LN, mm	5.04±2.63	6.91±3.45	0.005*
Long/short diameter ratio of LN	1.29±0.46	1.37±0.30	0.258
Number of LN excised	3.53±2.90	8.73±8.16	<0.001*
Number of LN metastasized	0	3.54±3.19	<0.001*
Histological subtype, n (%)			<0.001*
Classic PTC	57 (86.36)	55 (87.30)	
Follicular PTC	8 (12.12)	2 (3.17)	
Special subtype of PTC	1 (1.52)	6 (9.52)	
ETE, n (%)			<0.001*
No	58 (87.88)	35 (55.56)	
Yes	8 (12.12)	28 (44.44)	
TNM stage, n (%)			0.008*
I	58 (87.88)	35 (55.56)	
II	8 (12.12)	28 (44.44)	

*, P<0.05 was considered statistically significant. PTC, papillary thyroid carcinoma; LN, lymph nodes; TNM, tumor node metastasis; ETE, extrathyroidal extension.

multifocal PTCs were more common in the LNM group. There were significant differences in both the size of lymph nodes and the number of lymph nodes between the LNM group and the non-LNM group. But the ratio of long diameter to short diameter of lymph nodes had no

significant difference. According to postoperative pathology, extrathyroidal extension, special subtypes and TNM stage II were more common in LNM group. There were 6 cases of special subtypes (tall-cell 1, hobnail/micropapillary 3 and solid variants 1) in LNM group and 1 special subtypes in

non-LNM group (hobnail/micropapillary).

Features selection and radiomics signature construction

A total of 395 features were extracted from T2WI, DWI and T1C+ images. After the inter- and intra-observer agreement assessment, we reserved 357 features from T2WI images, 330 features from DWI images, and 373 features from T1C+ images for subsequent analysis (<https://cdn.amegroups.com/static/application/7ff4ec46d3721730c2800ca134c913ad/gs-20-479-01.pdf>).

After the mRMR and LASSO process, we finally selected seven features from T2WI images, five features from DWI images, ten features from T1C+ images and seven features from combined images. Among the seven features of combined images, four, two and one features were respectively from DWI, T2WI and T1C+ images, while the one feature from T1C+ images took the largest weight. The relative feature importance of all extracted features calculated by LASSO is shown in *Figure 2*. The odds ratio and 95% confidence interval of the feature are detailed in the *Table S1*.

Radiomics signature validation

A significant correlation between Rad-scores from four models and LNM was found in the training group. Rad-scores in the LNM group were higher than those of the non-LNM group (*Figure 3*). The validation group showed similar results, indicating that these models can help to stratify node-positive and node-negative PTC.

Model validation

The ROC analysis was used to evaluate performance of all four prediction models. In the training cohort, the respective values of AUC, accuracy, specificity, sensitivity, PPV, and NPV were 0.819, 0.750, 0.872, 0.622, 0.707, and 0.824 in the T2WI model; 0.826, 0.783, 0.702, 0.867, 0.846, and 0.736 in the DWI model; 0.808, 0.761, 0.830, 0.689, 0.736, and 0.795 in the T1C+ model; and 0.835, 0.804, 0.723, 0.889, 0.872, and 0.755 in the combined model. The respective predictive performance of the validation cohort was comparable. The AUC, accuracy, specificity, sensitivity, PPV, and NPV were respectively 0.798, 0.784, 0.789, 0.778, 0.789 and 0.778 in T2WI model; 0.798, 0.784, 0.789, 0.778, 0.789 and 0.778 in DWI model; 0.789, 0.784, 0.737, 0.833, 0.824 and 0.750 in T1C+ model, and 0.830, 0.811, 0.842,

0.778, 0.800 and 0.824 in combined model. The AUCs were slightly higher in combined model than that of models with features from single image type. The ROC curve in distinguishing PTC with and without LNM is shown in *Figure 4*. DCA showed that the combined prediction model was more advantageous (*Figure 5*).

Performance of the radiomics nomogram

Multivariate logistic regression identified Rad-score as an independent predictor of LNM in PTC patients. Our nomogram performed well in both cohort [AUC values: 0.835 (95% CI, 0.751–0.918) in the training cohort and 0.830 (95% CI, 0.691–0.970) in the validation cohort] (*Figure 6*).

Discussion

Cervical LNM conferred independent risk and affects survival outcomes in patients with PTC (10). In this study, we investigated the value of MRI-based radiomics in predicting LNM in PTC patients. Our results revealed that all three models based on features extracted from T2WI, DWI and T1C+ images had good performance in predicting LNM for PTC patients. The model constructed using combined features from three image types achieved better predictive performance for LNM in PTC patients, as shown by higher AUC (0.830), sensitivity (0.842), specificity (0.778), PPV (0.800), and NPV (0.824) in the validation cohort. The application of radiomics based on anatomical and diffusion MRI provides a new method for preoperative differential diagnosis of LNM. This approach could be helpful for preoperative risk stratification of PTC patients and optimal selection of surgical options, while avoiding unnecessary prophylactic lymphadenectomy in low-risk patients.

Cervical LNM is an significant prognosis factor in PTC (9). The identification and resection of cervical LNM is of great importance for reducing the risk of local recurrence (15). US and FNAB are considered as the preferred methods for assessing LNM. Recent studies have reported a few non-invasive and independent predictors of LNM in PTC, such as tumor capsular invasion, distance from capsule, and extrathyroidal extension (ETE) (36–38). In our previous study, we also confirmed that poorly defined tumor margins combined with thyroid contour protrusion sign on MRI could help to identify high-risk patients with LNM (39). However, considering that these image predictors are subjective and qualitative parameters and the low and varied sensitivity of them, the clinical significance

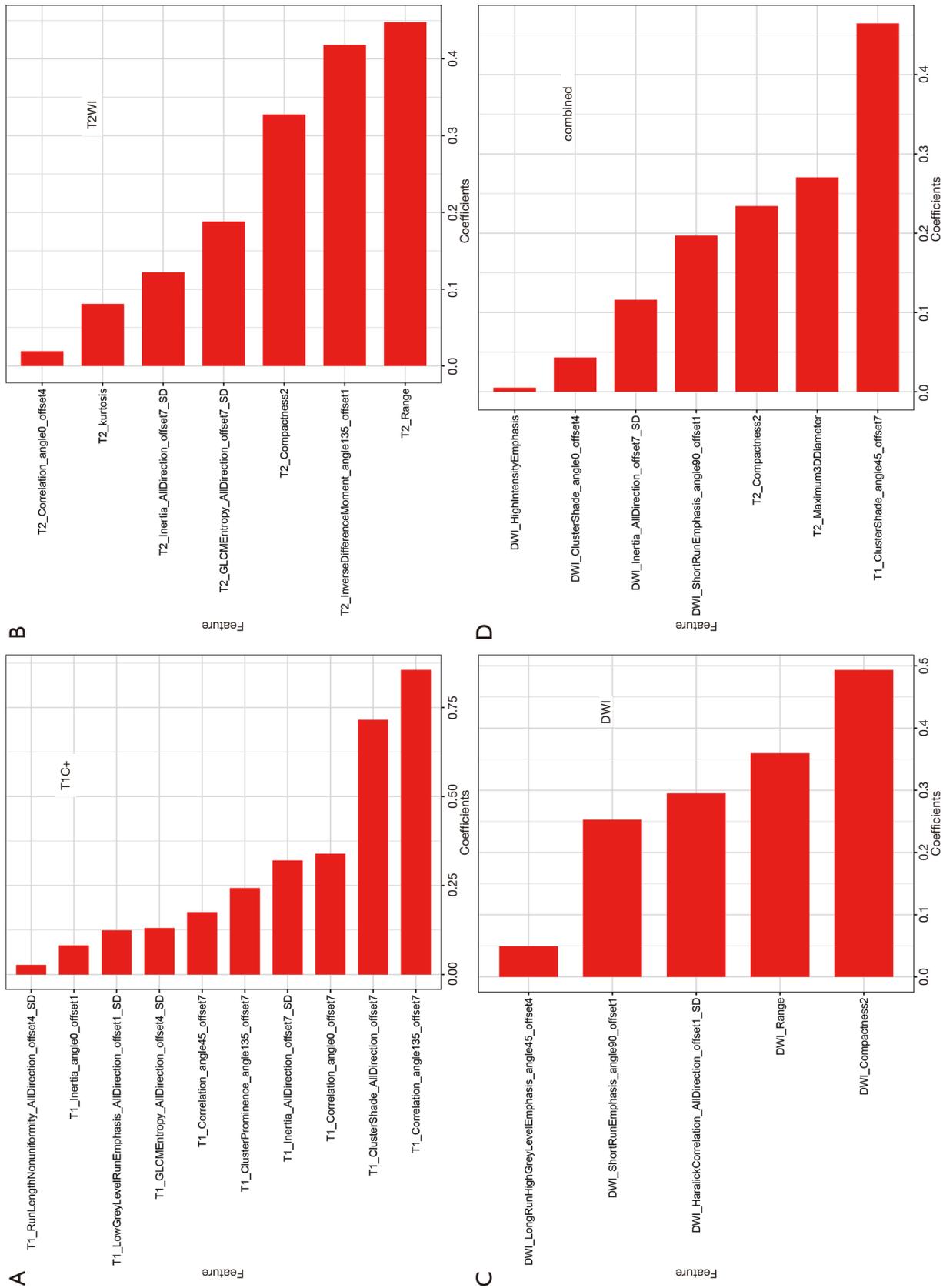


Figure 2 The final selected features extracted from T2WI, DWI, T1C+ and combined images to distinguish LNM from non LNM patients were shown in (A,B,C,D), respectively. T2WI, T2 weighted imaging; DWI, diffusion weighted imaging; T1C+, T1 weighted multiphase contrast enhancement imaging.

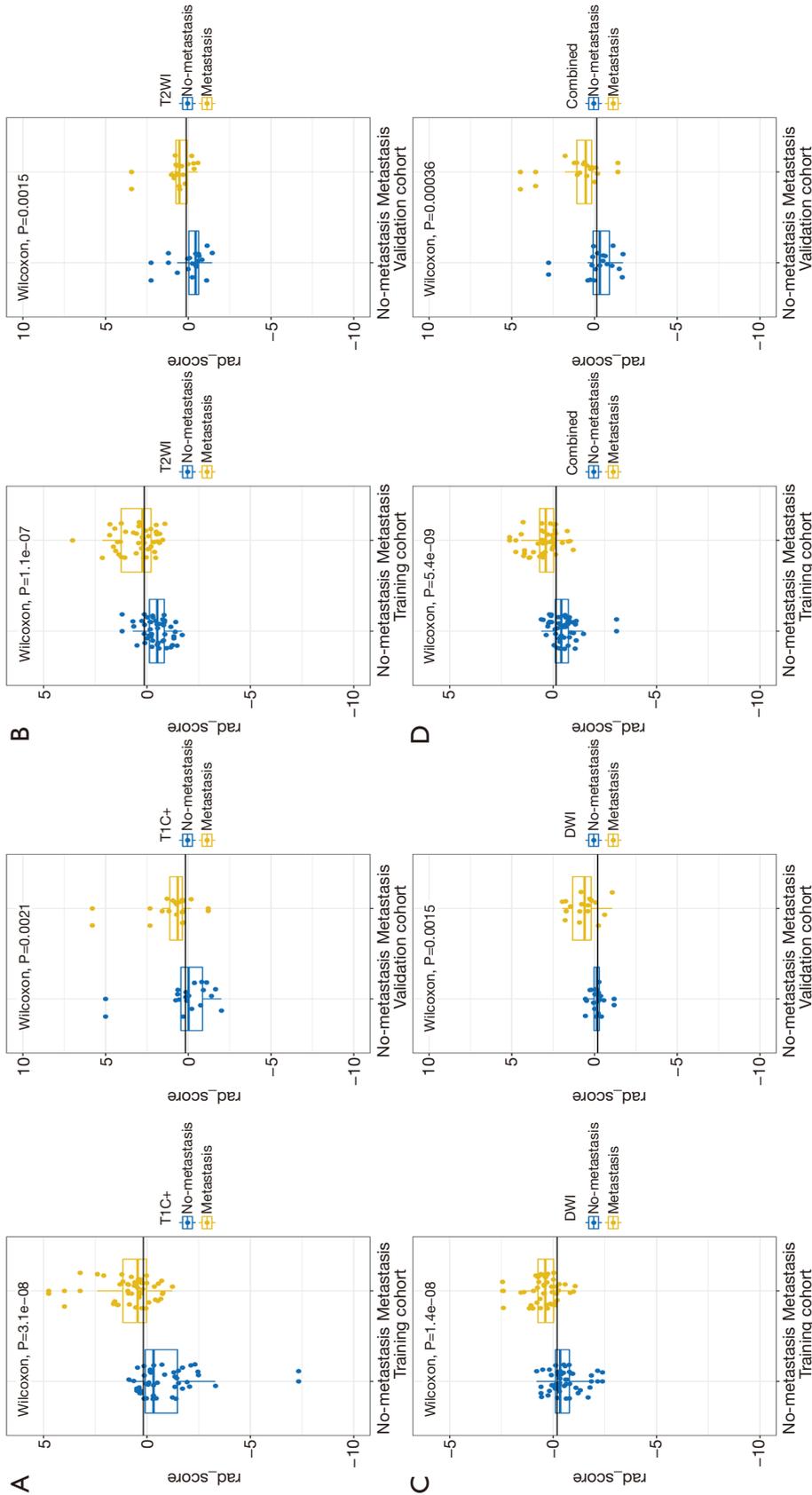


Figure 3 Comparison of the Rad-scores from T2WI, DWI, T1C+ and combined model between LNM and non-LNM group were shown in A-D. In all the four models, the Rad-scores of LNM patients were significantly higher than those of non-LNM patients in both training and validation cohort. T2WI, T2 weighted imaging; DWI, diffusion weighted imaging; T1C+, T1 weighted multiphase contrast enhancement imaging; LNM, lymph node metastasis.

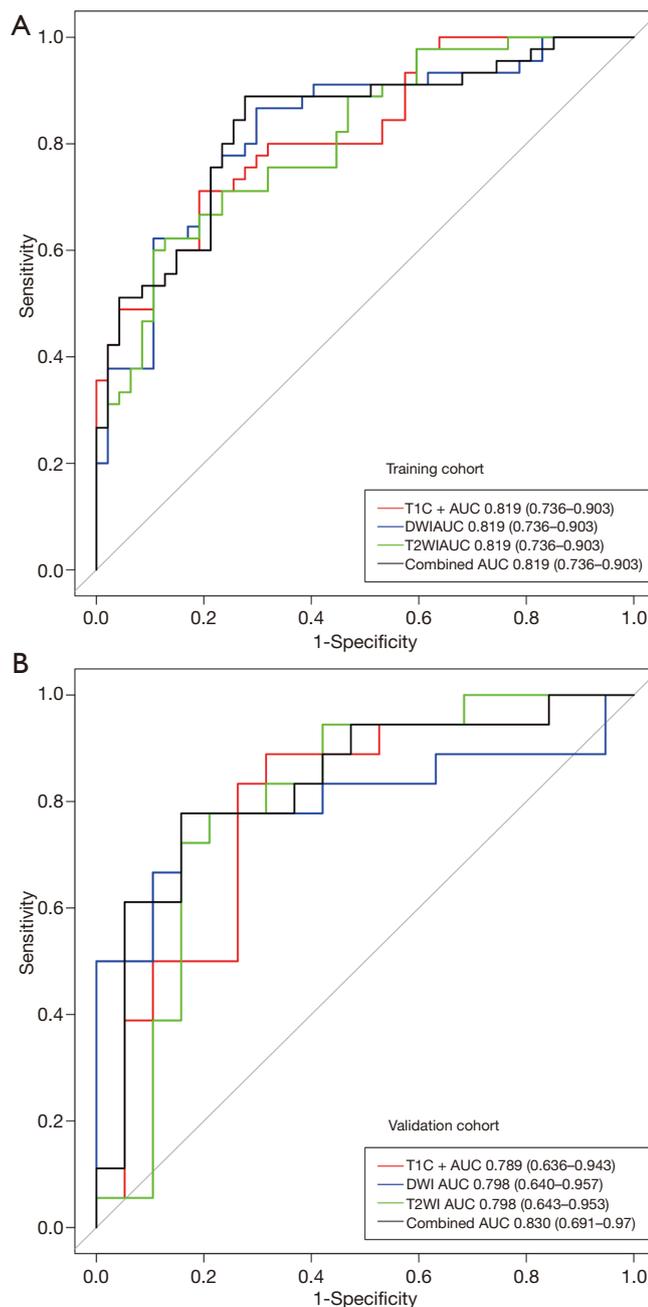


Figure 4 ROC curves for T2WI, DWI, T1C+ and combined model in distinguishing PTC with and without LNM in the (A) training and (B) validation cohorts. T2WI, T2 weighted imaging; DWI, diffusion weighted imaging; T1C+, T1 weighted multiphase contrast enhancement imaging; LNM, lymph node metastasis; ROC, receiver operating characteristic; PTC, papillary thyroid carcinoma.

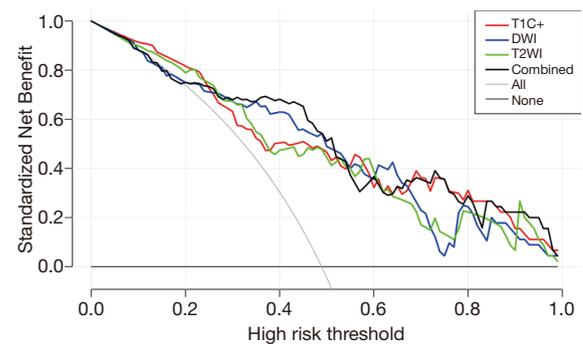


Figure 5 DCA for the four radiomics models. The x-axis represents the threshold probability and y-axis measures the net benefit. The black line represents the combined radiomics model. The DCA analysis showed that the combined model is more beneficial than the other three radiomic model. T2WI, T2 weighted imaging; DWI, diffusion weighted imaging; T1C+, T1 weighted multiphase contrast enhancement imaging; DCA, decision curve analysis; AUC, the area under the ROC.

was somewhat limited.

We investigated the value of radiomics based on MRI in predicting LNM in PTC patients in this study. Completely randomized grouping design and the consistency analysis between Interobserver and intra-observer were done to reduce the bias. We chose mRMR and LASSO to select features, due to their high usage rate and effectiveness according to previous studies (36,40-43). The defining principle of mRMR is to maximize the correlation between features and classification variables, while minimizing the correlation among features, which help to select features with less redundancies (44,45). Therefore, we carried out mRMR in the first stage to eliminate all redundant and irrelevant features. In the subsequent stage, we conducted LASSO to choose the optimized subset of features to construct the final model. The most outstanding advantage of LASSO regression is that the relatively unimportant variable coefficients were excluded by penalized regression. All independent variables can be processed at the same time, which greatly improves the stability of modeling. Radiomics is a novel and non-invasive method, which extract high-throughput features from medical images to establish appropriate model according to the heterogeneity of tumor improve diagnostic, prognostic, and predictive accuracy (46). To our best knowledge, few radiomics studies have focused

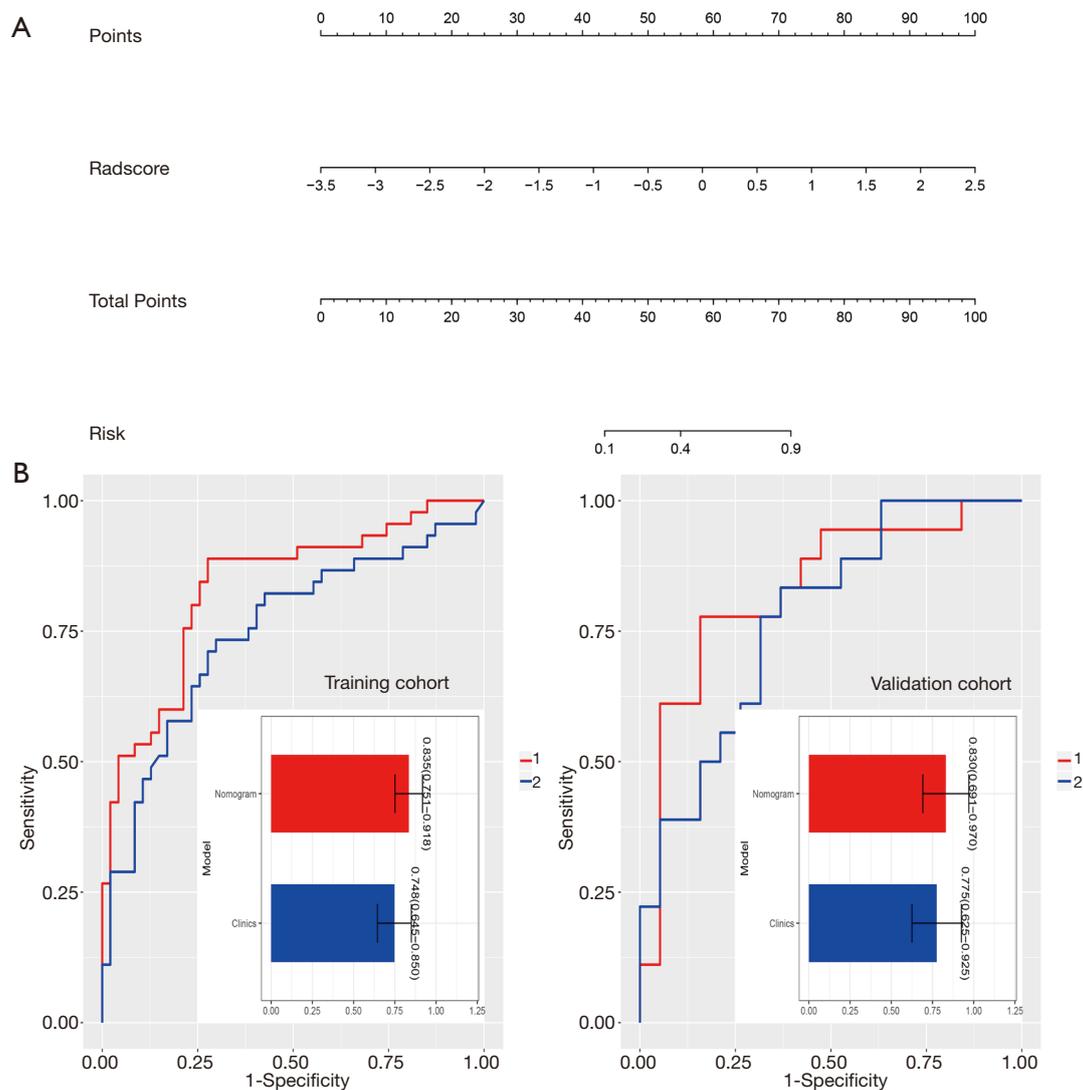


Figure 6 The radiomics nomogram by multivariable logistics regression analysis (A). The ROC analysis of the clinical model and the radiomic nomogram in the training cohort and validation cohort (B). The prediction performance of the radiomic nomogram for preoperative LNM of PTC was better than that of the clinical model in both the training and validation cohort. LNM, lymph node metastasis; ROC, receiver operating characteristic; PTC, papillary thyroid carcinoma.

on LNM prediction in PTC. Lu *et al.* (30). indicated that radiomics signatures extracted from CT images showed significant association with LNM with an AUC of 0.706 in the validation cohort. Liu *et al.* (31). showed that radiomics based on preoperative US images has the potential to predict LNM with an AUC and accuracy of 0.727 and 0.710, respectively, in the validation cohort. Here, we studied the value of radiomics on the basis of different MRI image types in the prediction of cervical LNM in PTC. Our results show that all radiomics models based on features

from T2WI, DWI and T1C+ images had good prediction performance, which were better than that reported by the two previous studies. Furthermore, it was encouraging to find that the model using combined features from three image types had a better prediction performance (AUC, 0.835 *vs.* 0.830), and a good balance between sensitivity and specificity in both study cohorts (sensitivity, 0.723 *vs.* 0.842; specificity, 0.889 *vs.* 0.778). In this study, we incorporated diffusion MRI in addition to conventional morphological images to predict LNM in PTC patients. A previous study

confirmed that whole ADC tumor volume is a valuable index for tumor biology in thyroid carcinoma (47). Our previous study showed that ADC is a reliable predictor for the separation of high invasive PTC lesions from low invasive lesions (48).

Hu *et al.* reported that the stability of quantitative features may be affected by segmentation results, machine models, and machine settings. Nevertheless, some features such as morphological features, intensity features, and GLCM features remained stable (49). In the combined images of our study, we extracted seven features including one texture feature from T1C+, two morphological features from T2WI and four features (two GLCM features, one intensity features and one texture features) from DWI. This result may indicate that our model is stable and cannot be easily influenced.

According to univariate analysis, although tumor size and short lymph nodes' diameter were significantly associated with LNM, they were not independent risk factors in the multivariate analysis. Rad-score was the most valuable predictor and retained in the radiomic nomogram.

The limitations of our study should be acknowledged. First, this study is a single-center study with a small sample size. Future multi-center studies with a large cohort population are needed to further validate the results. Second, for multifocal PTC, we only selected the largest tumor, while we cannot ensure that the pathogenic lesion is the largest one. Thirdly, lesions less than 5 mm were excluded in this study due to the indistinct boundary. Future advances in MRI acquisition may improve the detection of these small lesions. Finally, features were extracted from primary tumors without considering information of lymph nodes. A synthetic radiomics analysis of both primary tumor and lymph node may improve the prediction performance.

Conclusions

In conclusion, we established a noninvasive radiomics prediction model based on multimodality MRI. Overall, good accuracy was obtained in the preoperative prediction of the LNM status in PTC patients. This approach may be a convenient tool for clinicians to estimate individual risk of LNM and guide personalized treatment in PTC.

Acknowledgments

Funding: This work was supported by the Nature Science Foundation of Shanghai (grant number: 19ZR1446200) and the Municipal Commission of Health and Family Planning

(grant number: 201740242).

Footnote

Reporting Checklist: The authors have completed the STROBE reporting checklist. Available at <http://dx.doi.org/10.21037/gs-20-479>

Data Sharing Statement: Available at <http://dx.doi.org/10.21037/gs-20-479>

Conflicts of Interest: All authors have completed the ICMJE uniform disclosure form (available at <http://dx.doi.org/10.21037/gs-20-479>). The authors have no conflicts of interest to declare.

Ethical Statement: The authors are accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved. The study was conducted in accordance with the Declaration of Helsinki (as revised in 2013). This retrospective study was approved by the institutional ethics committee of Minhang Hospital affiliated to Fu Dan University School of Medicine (2015_51). Individual consent for this retrospective analysis was waived.

Open Access Statement: This is an Open Access article distributed in accordance with the Creative Commons Attribution-NonCommercial-NoDerivs 4.0 International License (CC BY-NC-ND 4.0), which permits the non-commercial replication and distribution of the article with the strict proviso that no changes or edits are made and the original work is properly cited (including links to both the formal publication through the relevant DOI and the license). See: <https://creativecommons.org/licenses/by-nc-nd/4.0/>.

References

1. Wiltshire JJ, Drake TM, Uttley L, et al. Systematic Review of Trends in the Incidence Rates of Thyroid Cancer. *Thyroid* 2016;26:1541-52.
2. Lim H, Devesa SS, Sosa JA, et al. Trends in Thyroid Cancer Incidence and Mortality in the United States, 1974-2013. *JAMA* 2017;317:1338-48.
3. Li N, Du XL, Reitzel LR, et al. Impact of enhanced detection on the increase in thyroid cancer incidence in the United States: review of incidence trends by socioeconomic

- status within the surveillance, epidemiology, and end results registry, 1980–2008. *Thyroid* 2013;23:103–10.
4. Davies L, Welch HG. Current thyroid cancer trends in the United States. *JAMA Otolaryngol Head Neck Surg* 2014;140:317–22.
 5. Kato E, Niebuhr DW. Screening for Thyroid Cancer. *Am Fam Physician* 2018;97:406–7.
 6. Xu JM, Xu XH, Xu HX, et al. Prediction of cervical lymph node metastasis in patients with papillary thyroid cancer using combined conventional ultrasound, strain elastography, and acoustic radiation force impulse (ARFI) elastography. *Eur Radiol* 2016;26:2611–22.
 7. Eskander A, Merdad M, Freeman JL, et al. Pattern of spread to the lateral neck in metastatic well-differentiated thyroid cancer: a systematic review and meta-analysis. *Thyroid* 2013;23:583–92.
 8. Randolph GW, Duh QY, Heller KS, et al. The prognostic significance of nodal metastases from papillary thyroid carcinoma can be stratified based on the size and number of metastatic lymph nodes, as well as the presence of extranodal extension. *Thyroid* 2012;22:1144–52.
 9. Zaydfudim V, Feurer ID, Griffin MR, et al. The impact of lymph node involvement on survival in patients with papillary and follicular thyroid carcinoma. *Surgery* 2008;144:1070–7; discussion 1077–8.
 10. Smith VA, Sessions RB, Lentsch EJ. Cervical lymph node metastasis and papillary thyroid carcinoma: does the compartment involved affect survival? Experience from the SEER database. *J Surg Oncol* 2012;106:357–62.
 11. Barczyński M, Konturek A, Stopa M, et al. Prophylactic central neck dissection for papillary thyroid cancer. *Br J Surg* 2013;100:410–8.
 12. Hartl DM, Mamelle E, Borget I, et al. Influence of prophylactic neck dissection on rate of retreatment for papillary thyroid carcinoma. *World J Surg* 2013;37:1951–8.
 13. Haugen BR, Alexander EK, Bible KC, et al. 2015 American Thyroid Association Management Guidelines for Adult Patients with Thyroid Nodules and Differentiated Thyroid Cancer: The American Thyroid Association Guidelines Task Force on Thyroid Nodules and Differentiated Thyroid Cancer. *Thyroid* 2016;26:1–133.
 14. Rosenbaum MA, McHenry CR. Central neck dissection for papillary thyroid cancer. *Arch Otolaryngol Head Neck Surg* 2009;135:1092–7.
 15. Khokhar MT, Day KM, Sangal RB, et al. Preoperative High-Resolution Ultrasound for the Assessment of Malignant Central Compartment Lymph Nodes in Papillary Thyroid Cancer. *Thyroid* 2015;25:1351–4.
 16. Kim E, Park JS, Son KR, et al. Preoperative diagnosis of cervical metastatic lymph nodes in papillary thyroid carcinoma: comparison of ultrasound, computed tomography, and combined ultrasound with computed tomography. *Thyroid* 2008;18:411–8.
 17. Monteiro R, Han A, Etiwy M, et al. Importance of surgeon-performed ultrasound in the preoperative nodal assessment of patients with potential thyroid malignancy. *Surgery* 2018;163:112–7.
 18. Yang SY, Shin JH, Hahn SY, et al. Comparison of ultrasonography and CT for preoperative nodal assessment of patients with papillary thyroid cancer: diagnostic performance according to primary tumor size. *Acta Radiol* 2020;61:21–7.
 19. Jun HH, Kim SM, Kim BW, et al. Overcoming the limitations of fine needle aspiration biopsy: detection of lateral neck node metastasis in papillary thyroid carcinoma. *Yonsei Med J* 2015;56:182–8.
 20. Cho SJ, Suh CH, Baek JH, et al. Diagnostic performance of CT in detection of metastatic cervical lymph nodes in patients with thyroid cancer: a systematic review and meta-analysis. *Eur Radiol* 2019;29:4635–47.
 21. Li L, Wang Y, Luo DH, et al. Diagnostic value of single-source dual-energy spectral computed tomography for papillary thyroid microcarcinomas. *J Xray Sci Technol* 2017;25:793–802.
 22. Liu X, Ouyang D, Li H, et al. Papillary thyroid cancer: dual-energy spectral CT quantitative parameters for preoperative diagnosis of metastasis to the cervical lymph nodes. *Radiology* 2015;275:167–76.
 23. Gillies RJ, Kinahan PE, Hricak H. Radiomics: Images Are More than Pictures, They Are Data. *Radiology* 2016;278:563–77.
 24. Limkin EJ, Sun R, Dercle L, et al. Promises and challenges for the implementation of computational medical imaging (radiomics) in oncology. *Ann Oncol* 2017;28:1191–206.
 25. Lambin P, Rios-Velazquez E, Leijenaar R, et al. Radiomics: extracting more information from medical images using advanced feature analysis. *Eur J Cancer* 2012;48:441–6.
 26. Han L, Zhu Y, Liu Z, et al. Radiomic nomogram for prediction of axillary lymph node metastasis in breast cancer. *Eur Radiol* 2019;29:3820–9.
 27. Wu S, Zheng J, Li Y, et al. A Radiomics Nomogram for the Preoperative Prediction of Lymph Node Metastasis in Bladder Cancer. *Clin Cancer Res* 2017;23:6904–11.
 28. Shen C, Liu Z, Wang Z, et al. Building CT Radiomics Based Nomogram for Preoperative Esophageal Cancer Patients Lymph Node Metastasis Prediction. *Transl Oncol*

- 2018;11:815-24.
29. Ji GW, Zhang YD, Zhang H, et al. Biliary Tract Cancer at CT: A Radiomics-based Model to Predict Lymph Node Metastasis and Survival Outcomes. *Radiology* 2019;290:90-8.
 30. Lu W, Zhong L, Dong D, et al. Radiomic analysis for preoperative prediction of cervical lymph node metastasis in patients with papillary thyroid carcinoma. *Eur J Radiol* 2019;118:231-8.
 31. Liu T, Zhou S, Yu J, et al. Prediction of Lymph Node Metastasis in Patients With Papillary Thyroid Carcinoma: A Radiomics Method Based on Preoperative Ultrasound Images. *Technol Cancer Res Treat* 2019;18:1533033819831713.
 32. Chen L, Xu J, Bao J, et al. Diffusion-weighted MRI in differentiating malignant from benign thyroid nodules: a meta-analysis. *BMJ Open* 2016;6:e008413.
 33. Dong Y, Feng Q, Yang W, et al. Preoperative prediction of sentinel lymph node metastasis in breast cancer based on radiomics of T2-weighted fat-suppression and diffusion-weighted MRI. *Eur Radiol* 2018;28:582-91.
 34. Liu H, Zhang C, Wang L, et al. MRI radiomics analysis for predicting preoperative synchronous distant metastasis in patients with rectal cancer. *Eur Radiol* 2019;29:4418-26.
 35. Liu C, Ding J, Spuhler K, et al. Preoperative prediction of sentinel lymph node metastasis in breast cancer by radiomic signatures from dynamic contrast-enhanced MRI. *J Magn Reson Imaging* 2019;49:131-40.
 36. Ryu YJ, Kang SJ, Cho JS, et al. Identifying risk factors of lateral lymph node recurrence in clinically node-negative papillary thyroid cancer. *Medicine (Baltimore)* 2018;97:e13435.
 37. Li F, Wu Y, Chen L, et al. Evaluation of clinical risk factors for predicting insidious right central and posterior right recurrent laryngeal nerve lymph node metastasis in papillary thyroid microcarcinoma patients (cN0): experience of a single center. *Ann Transl Med* 2019;7:8.
 38. Seong CY, Chai YJ, Lee SM, et al. Significance of distance between tumor and thyroid capsule as an indicator for central lymph node metastasis in clinically node negative papillary thyroid carcinoma patients. *PLoS One* 2018;13:e0200166.
 39. Wang H, Liu K, Ren J, et al. Magnetic Resonance Imaging Characteristics of Papillary Thyroid Carcinoma for the Prediction of Cervical Central Compartment Lymph Node Metastasis. *J Comput Assist Tomogr* 2019;43:963-9.
 40. Peng H, Long F, Ding C. Feature selection based on mutual information: criteria of max-dependency, max-relevance, and min-redundancy. *IEEE Trans Pattern Anal Mach Intell* 2005;27:1226-38.
 41. Tan X, Ma Z, Yan L, et al. Radiomics nomogram outperforms size criteria in discriminating lymph node metastasis in resectable esophageal squamous cell carcinoma. *Eur Radiol* 2019;29:392-400.
 42. Huang YQ, Liang CH, He L, et al. Development and Validation of a Radiomics Nomogram for Preoperative Prediction of Lymph Node Metastasis in Colorectal Cancer. *J Clin Oncol* 2016;34:2157-64.
 43. Yin P, Mao N, Zhao C, et al. Comparison of radiomics machine-learning classifiers and feature selection for differentiation of sacral chordoma and sacral giant cell tumour based on 3D computed tomography features. *Eur Radiol* 2019;29:1841-7.
 44. Ding C, Peng H. Minimum redundancy feature selection from microarray gene expression data. *J Bioinform Comput Biol* 2005;3:185-205.
 45. Garbarine E, DePasquale J, Gadia V, et al. Information-theoretic approaches to SVM feature selection for metagenome read classification. *Comput Biol Chem* 2011;35:199-209.
 46. Lambin P, Leijenaar RTH, Deist TM, et al. Radiomics: the bridge between medical imaging and personalized medicine. *Nat Rev Clin Oncol* 2017;14:749-62.
 47. Schob S, Meyer HJ, Dieckow J, et al. Histogram Analysis of Diffusion Weighted Imaging at 3T is Useful for Prediction of Lymphatic Metastatic Spread, Proliferative Activity, and Cellularity in Thyroid Cancer. *Int J Mol Sci* 2017;18:821.
 48. Song B, Wang H, Chen Y, et al. Efficacy of apparent diffusion coefficient in predicting aggressive histological features of papillary thyroid carcinoma. *Diagn Interv Radiol* 2018;24:348-56.
 49. Hu Y, Qiao M, Guo Y, et al. Reproducibility of quantitative high-throughput BI-RADS features extracted from ultrasound images of breast cancer. *Med Phys* 2017;44:3676-85.

Cite this article as: Hu W, Wang H, Wei R, Wang L, Dai Z, Duan S, Ge Y, Wu PY, Song B. MRI-based radiomics analysis to predict preoperative lymph node metastasis in papillary thyroid carcinoma. *Gland Surg* 2020;9(5):1214-1226. doi: 10.21037/gs-20-479

Rad-Score***T2WI***

Rad-Score = $0.328 \times T2_Compactness2 + 0.448 \times T2_Range + -0.188 \times T2_GLCMEntropy_AllDirection_offset7_SD + 0.419 \times T2_InverseDifferenceMoment_angle135_offset1 + -0.122 \times T2_Inertia_AllDirection_offset7_SD + 0.08 \times T2_kurtosis + 0.019 \times T2_Correlation_angle0_offset4 + -0.013$

DWI

Rad-Score = $0.253 \times DWI_ShortRunEmphasis_angle90_offset1 + 0.493 \times DWI_Compactness2 + -0.295 \times DWI_HaralickCorrelation_AllDirection_offset1_SD + -0.049 \times DWI_LongRunHighGreyLevelEmphasis_angle45_offset4 + 0.359 \times DWI_Range + -0.051$

T1WC +

Rad-Score = $-0.488 \times T1_LongRunLowGreyLevelEmphasis_angle0_offset7 + 0.962 \times T1_ClusterShade_angle90_offset7 + -0.502 \times T1_Inertia_angle0_offset1 + 0.14 \times T1_Correlation_AllDirection_offset7 + -0.152 \times T1_Inertia_AllDirection_offset7_SD + 0.213 \times T1_Maximum3DDiameter + 0.035 \times T1_Correlation_angle135_offset4 + -0.084$

Combined

Rad-Score = $0.197 \times DWI_ShortRunEmphasis_angle90_offset1 + -0.116 \times DWI_Inertia_AllDirection_offset7_SD + 0.465 \times T1_ClusterShade_angle45_offset7 + 0.234 \times T2_Compactness2 + 0.043 \times DWI_ClusterShade_angle0_offset4 + 0.27 \times T2_Maximum3DDiameter + 0.005 \times DWI_HighIntensityEmphasis + -0.045$

Table S1 The odds ratio of the feathers in each model

Variable and units	Odds ratio	Lower	Upper	P value
Combined model				
DWI_ShortRunEmphasis_angle90_offset1	2.488431	1.382362	4.479499	0.00237
DWI_Inertia_AllDirection_offset7_SD	0.582117	0.30598	1.107458	0.099164
T1_ClusterShade_angle45_offset7	9.864021	2.088047	46.59804	0.003861
T2_Compactness2	2.084356	1.241057	3.500679	0.005498
DWI_ClusterShade_angle0_offset4	1.746857	0.916475	3.329613	0.090086
T2_Maximum3DDiameter	2.472385	1.436455	4.255398	0.001086
DWI_HighIntensityEmphasis	1.569581	0.974815	2.527235	0.063596
DWI model				
DWI_ShortRunEmphasis_angle90_offset1	2.633214	1.392147	4.980663	0.002907
DWI_Compactness2	2.589255	1.509883	4.440241	5.46E-04
DWI_HaralickCorrelation_AllDirection_offset1_SD	0.718531	0.454622	1.13564	0.156971
DWI_LongRunHighGreyLevelEmphasis_angle45_offset4	0.488622	0.30233	0.789705	0.003457
DWI_Range	1.899267	1.158974	3.112423	0.010915
T1WC+ model				
T1_RunLengthNonuniformity_AllDirection_offset1_SD	31887196	1.20E-05	8.47E+19	0.236527
T1_Inertia_angle0_offset1	0.46882	0.283087	0.776412	0.003248
T1_Clustershade_AllDirection_offset7	3.327627	1.434688	7.718125	0.005097
T1_LowGreyLevelRunEmphasis_AllDirection_offset4_SD	0.017801	1.05E-04	3.020419	0.124059
T1_Inertia_AllDirection_offset7_SD	0.19205	0.04175	0.883428	0.034078
T1_Correlation_angle45_offset7	1.981494	1.101885	3.563276	0.022371
T1_Correlation_angle135_offset7	1.453282	0.924527	2.284442	0.105251
T1_GLCMEntropy_AllDirection_offset7_SD	5.49E-14	3.79E-50	7.94E+22	0.472298
T1_Correlation_angle0_offset7	2.293938	1.213839	4.335128	0.010567
T1_ClusterProminence_angle135_offset7	1.140744	0.747008	1.742012	0.542112
T2WI model				
T2_Compactness2	2.228061	1.285117	3.862884	0.004325
T2_Range	2.22889	1.325111	3.749083	0.00252
T2_GLCMEntropy_AllDirection_offset7_SD	0.397476	0.123443	1.279835	0.122004
T2_InverseDifferenceMoment_angle135_offset1	2.808027	1.44644	5.451327	0.002285
T2_Inertia_AllDirection_offset7_SD	0.076447	0.004193	1.39381	0.082599
T2_kurtosis	1.760236	1.064204	2.911499	0.027642
T2_Correlation_angle0_offset4	1.553793	0.994733	2.427056	0.052775