



Magnetic Resonance Imaging-Based Radiomics Features Associated with Depth of Invasion Predicted Lymph Node Metastasis and Prognosis in Tongue Cancer

Fei Wang, MD,^{1,2}  Rukeng Tan, MD,^{1,2} Kun Feng, MD,^{1,2} Jing Hu, MD,³ Zehang Zhuang, MD, PhD,^{1,2} Cheng Wang, MD, PhD,^{1,2} Jinsong Hou, MD, PhD,^{1,2*} and Xiqiang Liu, MD, PhD^{3*} 

Background: Adequate safe margin in tongue cancer radical surgery is one of the most important prognostic factors. However, the role of peritumoral tissues in predicting lymph node metastasis (LNM) and prognosis using radiomics analysis remains unclear. **Purpose:** To investigate whether magnetic resonance imaging (MRI)-based radiomics analysis with peritumoral extensions contributes toward the prediction of LNM and prognosis in tongue cancer.

Study type: Retrospective.

Population: Two hundred and thirty-six patients (38.56% female) with tongue cancer (training set, N = 157; testing set, N = 79; 37.58% and 40.51% female for each).

Field Strength/Sequence: 1.5 T; T2-weighted turbo spin-echo images.

Assessment: Radiomics models (R_{prim} , $R_{\text{prim}+3}$, $R_{\text{prim}+5}$, $R_{\text{prim}+10}$, $R_{\text{prim}+15}$) were developed with features extracted from the primary tumor without or with peritumoral extensions (3, 5, 10, and 15 mm, respectively). Clinicopathological characteristics selected from univariate analysis, including MRI-reported LN status, radiological extrinsic lingual muscle invasion, and pathological depth of invasion (DOI) were further incorporated into radiomics models to develop combined radiomics models (CR_{prim} , $CR_{\text{prim}+3}$, $CR_{\text{prim}+5}$, $CR_{\text{prim}+10}$, $CR_{\text{prim}+15}$). Finally, the model performance was validated in the testing set. DOI was measured from the adjacent normal mucosa to the deepest point of tumor invasion.

Statistical Tests: Chi-square test, regression analysis, receiver operating characteristic curve (ROC) analysis, decision analysis, spearman correlation analysis. The DeLong test was used to compare area under the ROC (AUC). $P < 0.05$ was considered statistically significant.

Results: Of all the models, the $CR_{\text{prim}+10}$ reached the highest AUC of 0.995 in the training set and 0.872 in the testing set. Radiomics features were significantly correlated with pathological DOI (correlation coefficients, -0.157 to -0.336). The $CR_{\text{prim}+10}$ was an independent indicator for poor disease-free survival (hazard ratio, 5.250) and overall survival (hazard ratio, 17.464) in the testing set.

Data Conclusion: Radiomics analysis with a 10-mm peritumoral extension had excellent power to predict LNM and prognosis in tongue cancer.

J. MAGN. RESON. IMAGING 2022;56:196–209.

Oral cancer (OC) is a highly heterogeneous disease with aggressive behavior, and accounting for nearly 377,713 new cases and 177,757 deaths worldwide in 2020.¹ Despite great development in treatment strategies, the 5-year survival rate of OC patients remains at 50% to 60%.^{2,3} Regional lymph node metastasis (LNM) is considered to be the most

View this article online at wileyonlinelibrary.com. DOI: 10.1002/jmri.28019

Received Oct 6, 2021, Accepted for publication Nov 23, 2021.

*Address reprint requests to: J.H., Department of Oral and Maxillofacial Surgery, Hospital of Stomatology, Guanghua School of Stomatology, Sun Yat-Sen University, No. 56 Lingyuan West Road, Guangzhou 510055, China. E-mail: houjs@mail.sysu.edu.cn, or X.L., Department of Oral and Maxillofacial Surgery, Nanfang Hospital, Southern Medical University, No. 1838 North Guangzhou Avenue, Guangzhou 510515, China. E-mail: liuxiqiang@smu.edu.cn

From the ¹Department of Oral and Maxillofacial Surgery, Hospital of Stomatology, Guanghua School of Stomatology, Sun Yat-Sen University, Guangzhou, China; ²Guangdong Provincial Key Laboratory of Stomatology, Guangzhou, China; and ³Department of Oral and Maxillofacial Surgery, Nanfang Hospital, Southern Medical University, Guangzhou, China

Additional supporting information may be found in the online version of this article

This is an open access article under the terms of the [Creative Commons Attribution-NonCommercial-NoDerivs](https://creativecommons.org/licenses/by-nc-nd/4.0/) License, which permits use and distribution in any medium, provided the original work is properly cited, the use is non-commercial and no modifications or adaptations are made.

important factor contributing to poor prognosis.^{3,4} With the presence of cervical LNM, the survival rate dramatically decreases by approximately 50%.⁵ Hence, accurate preoperative identification of cervical LN involvement is crucial for optimal treatment strategy and prognosis prediction.

Depth of invasion (DOI) is a well-known predictor of neck nodal metastasis, local recurrence, and survival in patients with OC.^{6–8} Tongue cancer, the most common OC, has a clear anatomical boundary. Because of the specialized function of the tongue (eg, tongue movement due to speech and swallowing), tumor cells tend to migrate along the muscle fibers, making the measurement of DOI at multiple spatial dimensions very important.⁹ The peritumoral tissues, a crucial component of the microenvironment, carry tumor aggressiveness-associated information and play an important role in metastasis and recurrence.^{10,11} However, it is still a challenging task to effectively find out the “true” margin of the primary tumor and excavate subtle change of peritumoral tissues using conventional medical images, such as magnetic resonance imaging (MRI), computed tomography, and so on. Therefore, a comprehensive evaluation of tumoral and peritumoral features may provide valuable information in the prediction of LNM.

MRI has been routinely used to preoperatively characterize primary tumors and assess clinical staging to establish treatment strategy for patients with OC. Tumor-associated qualitative variables measured on MRI may contribute to the diagnosis of LN status; however, the reported sensitivity and specificity have been variable.^{12,13} Moreover, conventional MRI such as T1-weighted images and T2-weighted images were insufficient to detect micrometastases of less than 3 mm.¹⁴ Hence, there is an urgent need to develop a more effective and noninvasive technique for mining information from MRI and facilitating the prediction of LNM.

“Radiomics” refers to a process that extracts and mines quantitative features from medical images via high-throughput computational techniques, thus uncovering disease characteristics that fail to be detected by the naked eye.¹⁵ Compared with invasive biopsy, radiomics characterizes complete tumor and its microenvironment heterogeneity in a noninvasive and cost-effective way without spatial and temporal limitations. Previous radiomics studies in intratumoral and peritumoral regions have shown to predict staging, treatment response, LNM, and survival.^{16–19} Furthermore, the radiomics nomogram incorporating radiomics signature and clinicopathological characteristics, including LN stage and molecular subtype, performed better than radiomics signature in disease-free survival estimation.²⁰

The aims of this study were: 1) to construct model based on radiomics features extracted from primary tumor to predict LNM of patients with tongue cancer; 2) to explore

the incremental value of peritumoral extensions and clinicopathological characteristics in LNM prediction; 3) to assess the predictive power of radiomics model in MRI-reported LN negative subgroup; 4) to explore the correlation between three-dimensional (3D) radiomics features and pathological DOI; 5) to investigate the association between the optimal radiomics model and survival outcomes.

Materials and Methods

Patients

This study was approved by our Institutional Review Board with written informed consent waived (ERC-[2016]-30). In total, 861 patients who had undergone radical resection glossectomy with or without neck dissection between September 2012 and November 2019 were recruited consecutively. The patient recruitment scheme is presented in Fig. S1 in the Supplemental Material. The inclusion criteria were as follows: 1) patients who had undergone standard head and neck MRI less than 30 days before radical resection; 2) availability of complete clinical and pathological information; and 3) the presence of newly diagnosed and surgically treated tongue cancer. The exclusion criteria were as follows: 1) missing survival information; 2) the presence of a concurrent malignant tumor other than on the tongue; and 3) MRI either degraded or with obvious artifacts. Finally, 236 patients conformed to the inclusion and exclusion criteria, and were randomly divided into two groups, a training set (N = 157) and an independent testing set (N = 79). The two sets were used to train and validate the radiomics models to preoperatively identify patients with high-risk LNM.

The baseline demographic information was obtained from medical records. To evaluate DOI and tumor budding, hematoxylin and eosin-stained tumor specimens were reviewed independently by two oral surgeons (J.S.H. and X.Q.L., with more than 20 years of experience) who were blinded to the clinicopathological data. Any disagreement was resolved by consensus. The LN status was determined according to pathological report in patients had undergone neck dissection at initial surgery. For patients had undergone radical excision of the primary tumor only, the LN status was evaluated during 2-year follow-up period. Patients who developed suspicious LNM during follow-up and were subsequently subjected to fine-needle aspiration, biopsy, or neck dissection. MRI parameters, including DOI, LN status, neurovascular invasion (NVI), extrinsic lingual muscle invasion (EMI), and lingual-septum invasion (LSI), were assessed blinded by three observers (J.S.H., X.Q.L., and Xiangbo Wan, two oral surgeons and one radiologist with more than 20 years of experience reading MRI). The assessment criteria are described in the Supplementary Information.

MRI Acquisition

Before undergoing radical resection, all patients had undergone standard head and neck MRI with a 1.5T Avanto MR scanner (Siemens Medical Solutions, Erlangen, Germany). Axial T2-weighted turbo spin-echo images were retrieved for feature extraction. The image acquisition parameters were as follows: repetition/echo time, 5800/87 or 6500/78 msec; number of averages, 2; echo train length, 13; flip angle, 15 degree; acquisition matrix, 224 * 320; reconstruction matrix, 320 * 320; number of slices, 36; slice thickness, 4.5 or

5 mm; spacing between slices, 6.5 mm; and field of view, 246 * 246 mm².

Region of Interest Segmentation and Image Preprocessing

As depicted in Fig. 1, the contour of the primary tumor and tongue body were manually delineated on consecutive axial slices with the open-source software ITK-SNAP (version 3.6.0, www.itksnap.org) by two trained oral surgeons (F.W. and K.F., each with 5 years of experience) independently in a blinded fashion. For avoiding possible dilatation beyond the structure of tongue and mouth floor, we limited the surface of tongue body as the upper boundary. As for the mouth floor, the inner side of musculus mylohyoideus was the lateral boundary, with musculus genioglossus and sublingual gland, tongue base, and hyoid bone as the anterior, posterior, and inferior boundaries, respectively. Then, the ROIs were 3D dilated (3, 5, 10, and 15 mm, respectively) outside the primary tumor with AIMED software (version: 1.6.4, Blot Info & Tech Co. Ltd, Beijing, China). Finally, the extended regions intersected with extension boundaries to produce the true ROIs for the subsequent radiomics analysis. Differences in image acquisition parameters, such as slice thickness, may introduce noise interference into the data. Therefore, the images were normalized in the whole slice images and resampled to a voxel resolution of 1 × 1 × 1 mm³.

Feature Extraction and Selection

To extract comprehensive information from the tumor and peritumoral microenvironment, we regarded the ROIs as a whole to extract radiomics features instead of analysis of intratumoral and peritumoral radiomics features separately. Radiomics features were extracted from ROIs with Python (version 3.5.2; <https://www.python.org>). In total, 786 radiomics features were extracted from each ROI, including 540 histograms of oriented gradient (HOG) features, 42 texture features, 156 statistical features, and 48 wavelet features. The normalization procedure was performed to change the values of extracted features to a range of 0 to 1. To reduce dimensionality and redundant features, a three-step procedure was performed to select LNM relevant radiomics features from the training set: 1) invalid features, including infinite values, null values, and values with variance of zero were removed; 2) intraclass correlation coefficient (ICC1) and interclass correlation coefficient (ICC2) of each radiomics feature were calculated to quantify feature extraction reproducibility (Supplementary Information). Features with ICC1 and ICC2 values greater than 0.8 were considered reproducible and robust, and therefore retained for analysis; 3) Subsequently, principal component analysis was conducted to select a small number of unique features, which is one of the most widely used data dimension reduction algorithms. The process of principal component analysis was described in the Supplementary Information.

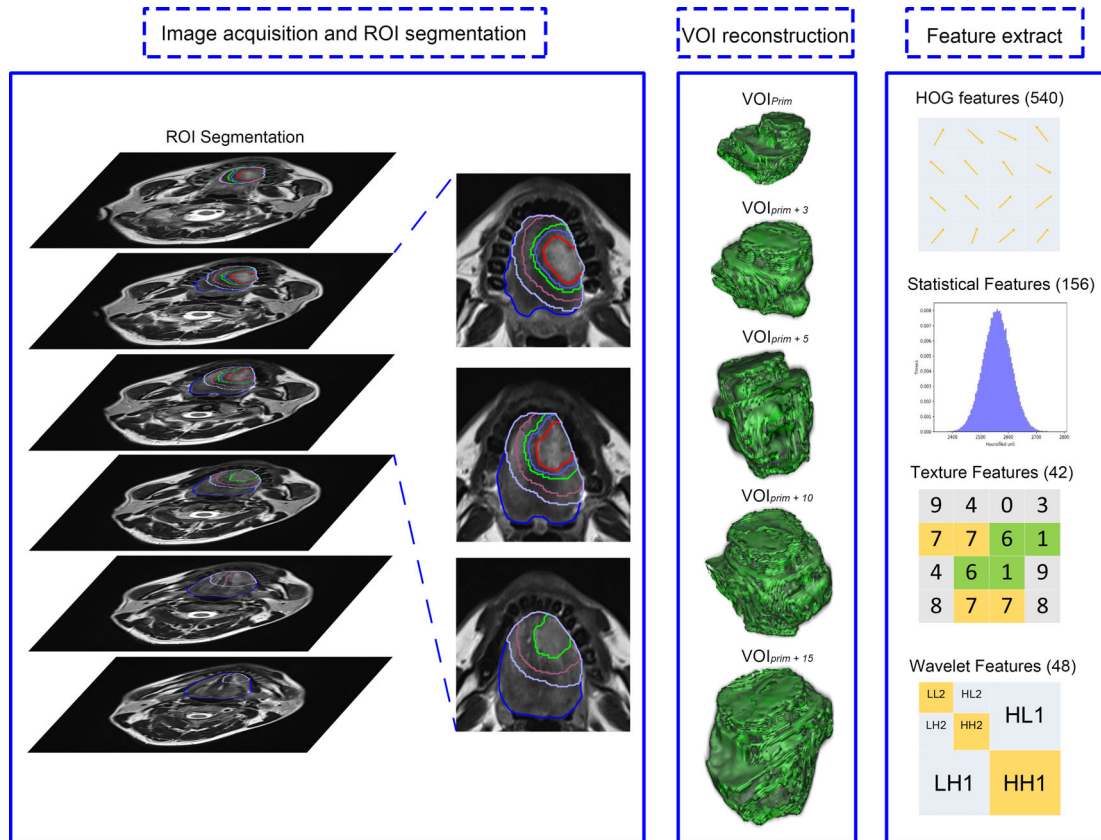


Figure 1: Workflow of region of interest (ROI) segmentation and feature extraction. First, ROIs of the primary tumor and extension boundaries were segmented on T2-weighted images. Then, the ROIs were three-dimensional dilated (3, 5, 10, and 15 mm, respectively) outside the primary tumors. Finally, the extended regions intersected with extension boundaries to produce the true ROIs for following radiomics analysis. Three-dimensional radiomics features, including 540 histogram of oriented gradient (HOG) features, 156 statistical features, 42 texture features, and 48 wavelet features were extracted to quantify shape, intensity, texture, and wavelet.

Model Development and Validation

The support vector machine (SVM) was performed on the training set to evaluate and optimize model hyperparameters employing 5-fold cross-validation. The radiomics models (R) were developed based on selected radiomics features. The radiomics models with 3D features extracted from the primary tumor or dilated ROIs were called R_{prim} , $R_{\text{prim}+3}$, $R_{\text{prim}+5}$, $R_{\text{prim}+10}$, $R_{\text{prim}+15}$, respectively. To evaluate the incremental value of clinicopathological characteristics including age, sex, T-staging, MRI-reported LN status, and so on in LNM risk prediction, we conducted a univariate analysis. Characteristics with P values <0.05 were incorporated into R models to build combined radiomics models (CR), which were called CR_{prim} , $CR_{\text{prim}+3}$, $CR_{\text{prim}+5}$, $CR_{\text{prim}+10}$, $CR_{\text{prim}+15}$, respectively. Besides, clinicopathological model (CP) was constructed using the selected clinicopathological characteristics. The predictive performance was then validated in the independent testing set.

Statistical Analysis

Statistical analyses were implemented with R software (version 3.5.2; R Foundation for Statistical Computing, Vienna, Austria) and SPSS (version 24.0; IBM Corporation, Armonk, NY, USA). The chi-square test was performed for the analysis of categorical variables. Patients with a prediction probability greater than 0.5 were considered to be at high-risk for LNM. The predictive ability of the prognostic models was evaluated by the area under the receiver operating characteristic curve (AUC) values and AUC's 95% confidence interval (CI). The DeLong test was performed to compare AUC values between prognostic models.²¹ Calibration curves were plotted to compare the predicted and actual values. In addition, accuracy rate, sensitivity, and specificity value were calculated. To calculate the clinical usefulness of the optimal prognostic model, decision curve analysis was conducted to quantify the net benefits at different threshold probabilities in the testing set.²² Spearman correlation coefficient was used to calculate correlations between pathological DOI and radiomics features. Kaplan–Meier analysis and the log-rank test were performed to depict and compare disease-free survival (DFS) and overall survival (OS) curves, respectively. Univariate and multivariate Cox regression analysis was performed to investigate the association between model-predicted LNM and survival outcomes. A two-sided P value <0.05 was considered statistically significant.

Results

Patient Characteristics

The clinicopathological characteristics of patients with tongue cancer in the training and independent testing sets are presented in Table 1. Patients of 61.44% (145/236) were men, and the average age of all patients were 50.83 ± 13.58 years (range, 19–86 years). All patients received radical excision of the primary tumors, and 166 underwent neck dissection simultaneously. Of the 99 patients with LNM, 56 patients were diagnosed at surgery. The remaining 43 patients developed LNM during 2-year follow-up period. The LN-positive patients constituted 42.04% (66/157) and 41.78% (33/79) in the training and testing sets, respectively ($P = 0.969$). Clinicopathological characteristics showed no

differences between the two sets, indicating balanced distribution.

Tumor Boundary Visualization

As shown in Fig. 2a, the tumor was infiltrated towards the lingual-septum and mouth floor, which reflects the invasiveness of tongue cancer. The 3D extension toward mouth floor was visualized using anatomical diagram in Fig. 2b. Not only peritumoral tissue on the MR cross section images, but also tissue in mouth floor was included to extract radiomics features. The image of hematoxylin and eosin-stained tongue squamous cell carcinoma was shown in Fig. 2c, which reflects the infiltrative growth pattern of tumor.

Clinicopathological Model

As shown in Table 1, univariate analysis in the training set showed MRI-reported LN status, radiological EMI, and pathological DOI were significantly associated with LNM. Multivariate logistic regression analysis identified MRI-reported LN status (OR 2.432, 95% CI, 1.093–5.411) as an independent predictor of LNM. As shown in Table 2, the MRI-reported LN status had an AUC of 0.643 (95% CI, 0.545–0.741, Fig. 3a) in the training set, with the accuracy, sensitivity, and specificity of 64.97%, 40.91%, and 82.42%, respectively. In the testing set, the AUC, accuracy, sensitivity, and specificity were 0.649 (95% CI, 0.522–0.777, Fig. 3b), 69.62%, 36.36%, and 93.48%, respectively. The SVM-based CP model, which included MRI-reported LN status, radiological EMI, and pathological DOI, yielded an AUC of 0.933 (95% CI, 0.873–0.993, Fig. 3a) in the training set, with the accuracy, sensitivity, and specificity of 86.06%, 84.84%, and 86.87%, respectively. In the testing set, the performance declined, with the AUC, accuracy, sensitivity, and specificity of 0.700 (95% CI, 0.660–0.740, Fig. 3b), 72.15%, 60.61%, and 80.44%, respectively. Nevertheless, the SVM-based CP model performed significantly better than the MRI-reported LN status, with the AUC elevated 0.051 (0.700 vs. 0.649) in the testing set.

Radiomics Model with Peritumoral Extensions

After feature reduction, top 15, 17, 18, 25, and 10 radiomics features from the primary tumor, and the primary tumor with 3D peritumoral extensions (3, 5, 10, and 15 mm, respectively) were selected to build R models, respectively. The importance of selected features to R models are presented in Figs. S2–S6 in the Supplemental Material. In the training set, the R_{prim} showed a favorable predictive performance, with an AUC of 0.978 (95% CI, 0.963–0.993, Fig. 3a). After different peritumoral extensions, the $R_{\text{prim}+3}$, $R_{\text{prim}+5}$, $R_{\text{prim}+10}$, and $R_{\text{prim}+15}$, achieved 0.978 (95% CI, 0.966–0.990), 0.985 (95% CI, 0.975–0.995), 0.992 (95% CI, 0.984–1.000), and 0.987 (95% CI, 0.978–0.996), respectively.

TABLE 1. Baseline Characteristics of Patients in the Training and Testing Sets

Characteristics	Training Set (N = 157)		P Value	Testing Set (N = 79)		P Value*	P Value†
	LN (-) (N = 91)	LN (+) (N = 66)		LN (-) (N = 46)	LN (+) (N = 33)		
Age, years			0.811			0.152	0.978
<60	66	49		31	27		
≥60	25	17		15	6		
Sex			0.947			0.525	0.663
Female	34	25		20	12		
Male	57	41		26	21		
T-staging			0.220			0.058	0.107
T1–T2	49	29		32	16		
T3–T4	42	37		14	17		
MRI-reported LN status			0.001			0.001	0.157
Negative	75	39		43	21		
Positive	16	27		3	12		
Radiological NVI			0.155			0.058	0.446
Negative	70	44		39	22		
Positive	21	22		7	11		
Radiological DOI			0.240			0.058	0.154
≤10 mm	50	30		32	16		
>10 mm	41	36		14	17		
Radiological EMI			0.014			0.012	0.207
Negative	76	44		43	23		
Positive	15	22		3	10		
Radiological LSI			0.995			0.595	0.874
Negative	80	58		42	28		
Positive	11	8		4	5		
Tumor budding			0.050			0.013	0.088
Low	46	23		31	13		
High	45	43		15	20		
Pathological DOI			0.006			0.102	0.394
≤10 mm	72	39		38	22		
>10 mm	19	27		8	11		

LN = lymph node; NVI = neurovascular invasion; DOI = depth of invasion; EMI = extrinsic lingual muscle invasion; LSI = lingual-septum invasion.

*P value represents the difference in clinicopathological characteristics between LN-positive and LN-negative cohorts in the training and testing sets; †P value represents the comparison between training and testing sets.

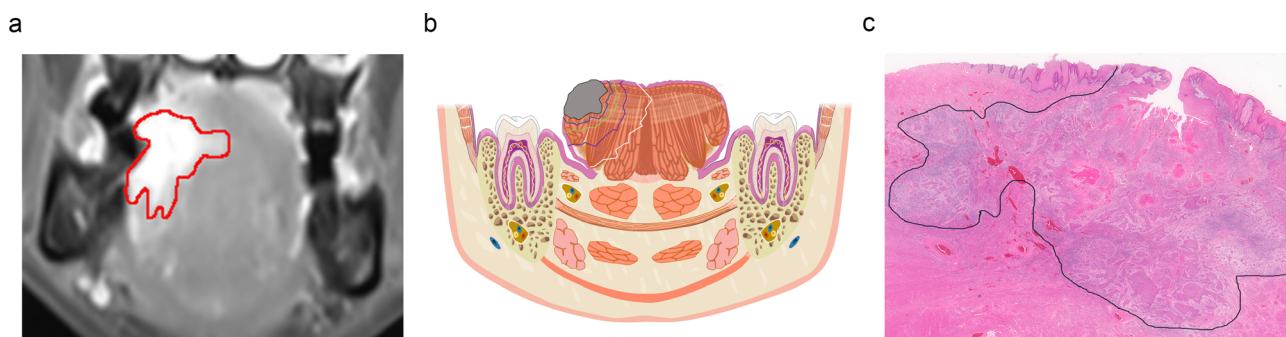


Figure 2: MR image, anatomical diagram, and H&E-stained slide of tongue squamous cell carcinoma. (a) The tumor is infiltrated toward the lingual-septum and mouth floor. (b) The primary tumor and peritumoral extensions are visualized on coronal section using anatomical diagram. (c) The image of H&E-stained tongue squamous cell carcinoma.

TABLE 2. Performances of Prognostic Models in the Training and Testing Sets

Model	Set	AUC	Accuracy (%)	Sensitivity (%)	Specificity (%)
MR	Training	0.643 (0.545–0.741)	64.97 (55.36–73.26)	40.91 (32.68–48.94)	82.42 (73.56–90.02)
	Testing	0.649 (0.522–0.777)	69.62 (60.16–76.38)	36.36 (30.74–44.37)	93.48 (96.33–89.79)
CP	Training	0.933 (0.873–0.993)	86.06 (80.86–91.26)	84.84 (80.84–88.84)	86.87 (88.85–92.89)
	Testing	0.700 (0.660–0.740)	72.15 (66.15–78.15)	60.61 (52.61–68.61)	80.44 (75.43–85.45)
R_{prim}	Training	0.978 (0.963–0.993)	92.99 (89.15–86.09)	89.39 (85.19–93.59)	95.6 (93.12–98.18)
	Testing	0.710 (0.680–0.740)	73.42 (68.3–78.54)	60.6 (56.00–65.20)	82.6 (78.59–86.61)
$R_{\text{prim}+3}$	Training	0.978 (0.966–0.990)	94.9 (92.1–97.71)	92.42 (88.42–96.42)	96.70 (94.97–98.46)
	Testing	0.793 (0.765–0.821)	77.21 (73.06–81.36)	66.67 (62.47–70.87)	84.78 (81.27–88.29)
$R_{\text{prim}+5}$	Training	0.985 (0.975–0.995)	96.17 (93.57–98.77)	93.93 (90.23–97.63)	97.8 (96.32–99.31)
	Testing	0.784 (0.759–0.809)	81.01 (77.01–85.01)	72.72 (68.92–76.52)	86.96 (83.73–90.19)
$R_{\text{prim}+10}$	Training	0.992 (0.984–1.00)	97.45 (94.95–99.95)	95.45 (92.25–98.65)	98.9 (98.00–99.83)
	Testing	0.855 (0.833–0.877)	86.08 (82.50–89.66)	78.79 (75.59–81.99)	91.30 (88.73–93.87)
$R_{\text{prim}+15}$	Training	0.987 (0.978–0.996)	95.51 (93.31–97.71)	95.38 (92.38–98.38)	95.60 (93.44–97.83)
	Testing	0.823 (0.795–0.851)	83.54 (79.81–87.27)	75.75 (72.15–79.35)	89.13 (85.88–92.38)
CR_{prim}	Training	0.987 (0.974–0.992)	93.63 (90.73–96.53)	90.9 (86.9–94.9)	95.6 (93.63–97.57)

TABLE 2. Continued

Model	Set	AUC	Accuracy (%)	Sensitivity (%)	Specificity (%)
	Testing	0.720 (0.680–0.760)	78.48 (73.47–83.49)	66.67 (62.57–70.77)	86.96 (83.45–90.47)
CR _{prim+3}	Training	0.988 (0.979–0.997)	96.15 (93.35–98.95)	93.93 (90.43–97.28)	97.77 (96.57–98.96)
	Testing	0.787 (0.757–0.817)	81.01 (76.96–85.06)	69.69 (65.69–73.26)	89.13 (85.93–92.33)
CR _{prim+5}	Training	0.992 (0.985–0.999)	96.81 (94.31–99.31)	95.45 (92.25–98.65)	97.8 (96.75–98.78)
	Testing	0.793 (0.764–0.823)	84.81 (80.96–88.66)	75.76 (72.56–78.88)	91.30 (88.77–93.83)
CR _{prim+10}	Training	0.995 (0.991–0.999)	97.45 (95.35–99.95)	95.45 (93.39–98.76)	98.9 (98.12–99.82)
	Testing	0.872 (0.847–0.897)	87.34 (84.32–90.36)	78.78 (75.78–81.63)	93.47 (91.21–95.73)
CR _{prim+15}	Training	0.991 (0.984–0.998)	95.54 (93.14–97.94)	92.42 (89.92–94.82)	97.80 (96.61–99.01)
	Testing	0.808 (0.788–0.828)	80.24 (76.72–83.76)	68.57 (62.38–74.57)	89.13 (86.32–91.94)

MR = MRI-reported lymph node status; SVM-CP = support vector machine-based clinicopathological model; R = radiomics models; CR = combined radiomics models; AUC = area under the receiver operating characteristic curve.

In the testing set, the AUC of the R_{prim} declined to 0.710 (95% CI, 0.680–0.740, Fig. 3b); however, AUC was still significantly higher than that of MRI-reported LN status and SVM-based CP. With the increase of the scale of peritumoral extension, the AUC increased and reached the maximum value at 10-mm peritumoral extension. Therefore, the R_{prim+10} yielded the highest AUC of 0.855 (95% CI, 0.833–0.877), which was significantly higher than the AUC of other radiomics models, including R_{prim}, R_{prim+3} (AUC = 0.793, 95% CI, 0.765–0.821), R_{prim+5} (AUC = 0.784, 95% CI, 0.759–0.809), R_{prim+15} (AUC = 0.823, 95% CI, 0.795–0.851). Meanwhile, the accuracy, sensitivity, and specificity of the R_{prim+10} were higher than the above models with values of 86.08%, 78.79%, and 91.30%, respectively.

Radiomics Model with Clinicopathological Characteristics Integrated

MRI-reported LN status, radiological EMI, and pathological DOI were integrated into R models to construct CR models. The importance of selected features to CR models is presented in Figs. S7–S11 in the Supplemental Material. In the training set, CR_{prim}, CR_{prim+3}, CR_{prim+5}, and CR_{prim+15} achieved AUCs of 0.987 (95% CI, 0.974–0.992), 0.988 (95% CI, 0.979–0.997), 0.992 (95% CI, 0.985–0.999), and 0.991 (95%

CI, 0.984–0.998), respectively, while the CR_{prim+10} achieved the AUC of 0.995 (95% CI, 0.991–0.999, Fig. 3a), significantly higher than other CR models. In the testing set, the CR_{prim+10} still achieved the highest predictive power with the AUC of 0.872 (95% CI, 0.847–0.897, Fig. 3b), which was significantly higher than the AUC of CR_{prim} (AUC = 0.720, 95% CI, 0.680–0.760), CR_{prim+3} (AUC = 0.787, 95% CI, 0.757–0.817), CR_{prim+5} (AUC = 0.793, 95% CI, 0.764–0.823), and CR_{prim+15} (AUC = 0.808, 95% CI, 0.788–0.828). The accuracy, sensitivity, and specificity of the CR_{prim+10} were 87.34%, 78.78%, and 93.47%, respectively, which were consistently higher than other models. The detailed results are summarized in Table 2. The calibration curves of prediction models, especially the CR_{prim+10}, showed good agreement between model-predicted probability and the actual LNM rate in both the training and independent testing sets (Fig. 3c,d). Compared with the R_{prim+10}, R_{prim}, CP model, treat-all, and treat-none scheme, the CR_{prim+10} provided the greatest net benefits for predicting LNM for threshold probabilities of 0.415–0.806 (Fig. 4a).

Discriminative Ability of the CR_{prim+10} in the MRI-Reported LN-Negative Subgroup

Of all cases, 60.61% (60/99) with pathologically positive nodes were understaged and 13.87% (19/137) with

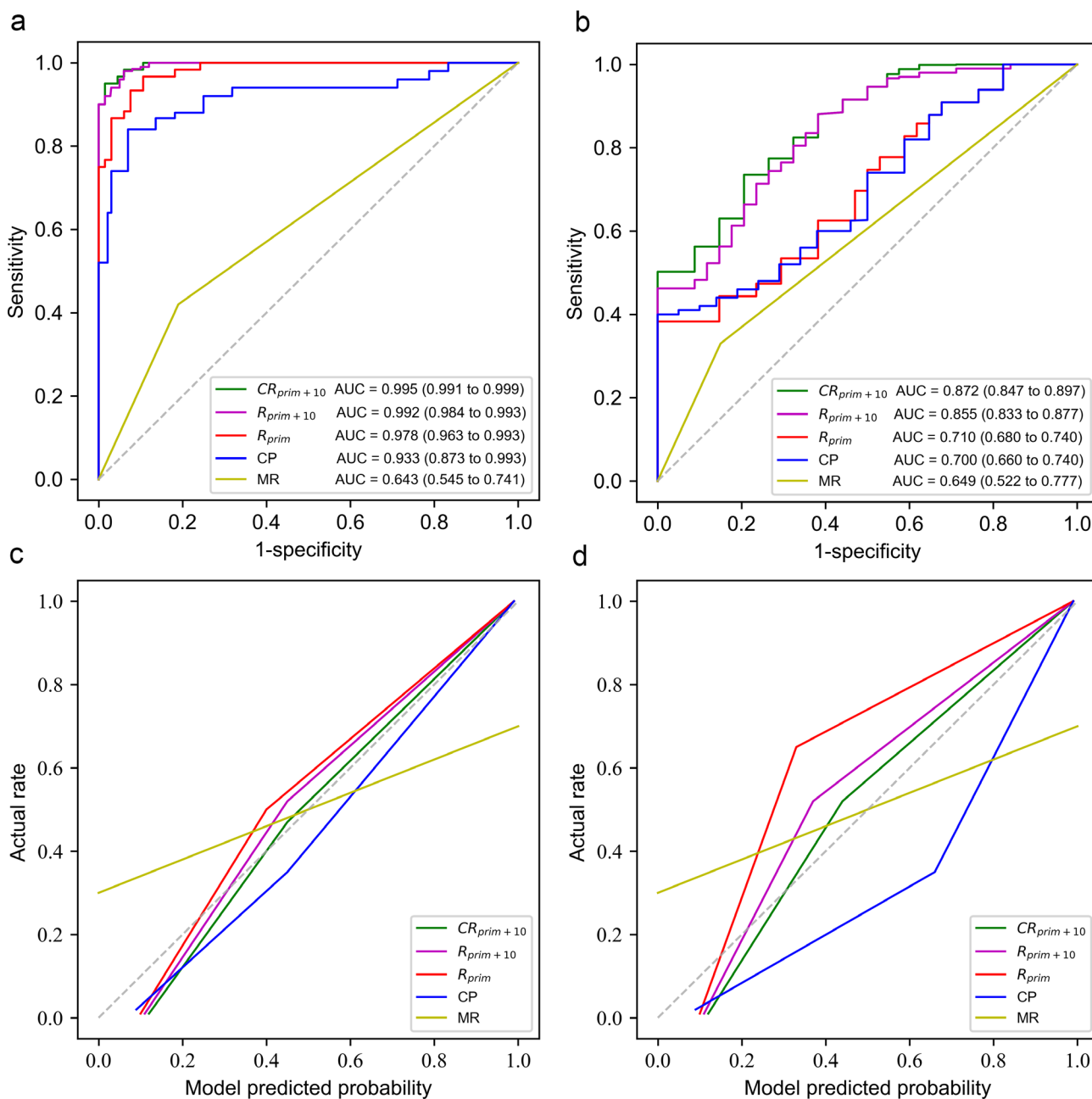


Figure 3: Receiver operating characteristic curves of different diagnostic models in the training set (a) and testing set (b). The combined radiomics model of the primary tumor with a 10-mm peritumoral extension ($CR_{prim+10}$) reached the highest AUC of 0.995 (95% CI, 0.991–0.999) in the training set and 0.872 (95% CI, 0.847–0.897) in the testing set. Calibration curves of prognostic models in the training set (c) and testing set (d). The $CR_{prim+10}$ showed good agreement between model-predicted probability and the actual LNM rate. The lines of diagnostic model have a closer fit to the 45° dotted grey line, indicating their better predictive accuracy.

pathologically negative nodes were overstaged according to MRI reports, which suggested that MRI reports had high specificity but low sensitivity. In the MRI-reported LN-negative subgroup, $CR_{prim+10}$ presented a favorable predictive performance with an AUC of 0.883 (95% CI, 0.862–0.902, Fig. 4b). The LN status of 87.5% (56/64) were correctly identified, with only three positive and five negative patients misclassified, showing discriminative capability of the $CR_{prim+10}$ (Fig. 4c).

Correlation Analysis

In NVI-negative subgroup, there was a significant negative correlation between pathological DOI and statistic_1_0_1-Homogeneity ($r = -0.157$, Fig. 5a). In NVI-positive subgroup, pathological DOI was negatively correlated with statistic_0_1_0-MaxProbability ($r = -0.278$, Fig. 5b). In radiological EMI-positive subgroup, a significant negative correlation between pathological DOI and statistic_0_1_0-MaxProbability ($r = -0.336$, Fig. 5c) was found. In tumor budding number

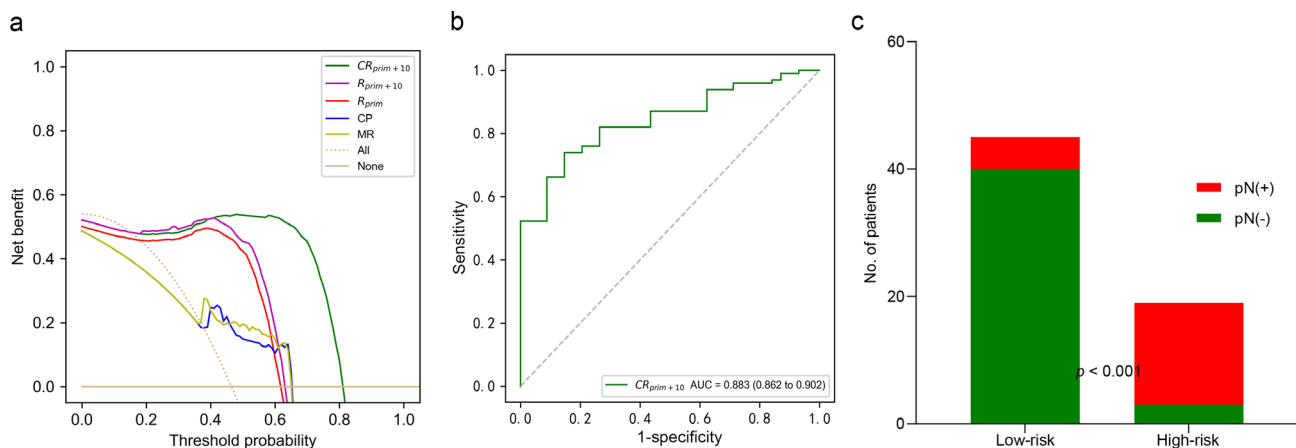


Figure 4: (a) Decision curve analysis of prognostic models in the testing set. The x-axis represents the threshold probability. The y-axis represents the net benefit. The decision curve demonstrated that in threshold probabilities of 0.415–0.806, applying the $CR_{prim+10}$ to predict lymph node metastasis (LNM) adds more benefit than other models. (b) The $CR_{prim+10}$ presented a favorable predictive performance in the MRI-reported LN-negative subgroup, with an AUC of 0.883. (c) The distribution of the $CR_{prim+10}$ -predicted LN status in the MRI-reported LN-negative subgroup, indicated that the $CR_{prim+10}$ had a good discriminate ability.

≥ 5 subgroup, there was a significant positive correlation between pathological DOI and statistic_1_1_-1-ClusterShade ($r = -0.185$, Fig. 5d). Also, there were significant correlations between pathological DOI and some HOG features (Fig. S12 in the Supplemental Material).

Representative LN-negative and LN-positive cases were showed in Fig. 6a–f. The signal intensity within ROI was relatively higher in the LN-positive patients as compared with the LN-negative patients, which indicated distinct textural pattern on MR images between LN-negative and LN-positive patients with tongue cancer.

The Predictive Value of the $CR_{prim+10}$ in Survival Outcomes

The median follow-up period was 41 months (range: 4–95 months). The DFS rates and OS rates at 2, 3, and 5 years were 66.67%, 55.70%, 44.79%, and 77.78%, 67.13%, 48.96%, respectively. A significant association between the $CR_{prim+10}$ and survival outcome was identified, and the LNM high-risk group had worse DFS and OS compared with the low-risk group identified by the $CR_{prim+10}$ (Fig. 6g,h). Tumor budding and $CR_{prim+10}$ -predicted LN status were incorporated into the multivariate Cox analysis of DFS. MRI-reported LN status, radiological lingual-septum invasion, and $CR_{prim+10}$ -predicted LN status were incorporated into the multivariate Cox analysis of OS. Multivariate Cox regression analysis showed that the $CR_{prim+10}$ was an independent pre-operative indicator for poor DFS (hazard ratio, 5.250 [95% CI, 2.152–12.808]) and OS (hazard ratio, 17.464 [95% CI, 3.660–83.330]) (Table 3).

Discussion

In this study, we developed radiomics models with 3D peritumoral extensions to predict LNM and prognosis preoperatively. Our results showed that the $CR_{prim+10}$ had the

greatest predictive performance, especially in the MRI-reported LN-negative subgroup. The 3D radiomics features were significantly correlated with pathological DOI. Importantly, the $CR_{prim+10}$ was an independent indicator for poor DFS and OS.

To develop the radiomics models, we investigated tongue cancer, the most common cancer of the oral cavity. The tongue is a unique muscular organ with legible 3D anatomy architecture, making it relatively leisurely to determine the tongue boundary and tumor border with MR imaging. Therefore, the tongue is an appropriate model to study the infiltrative growth pattern of tumor cells and 3D peritumoral extension of primary tumors using radiomics analysis.

Previous radiomics studies in OC have focused on the primary tumor or LN without taking peritumoral regions into account.^{23,24} In our study, we focused not only on the primary tumor, but also the peritumoral extensions. Furthermore, instead of simple 2D peritumoral extensions on the MR cross section images, here, we conducted 3D extensions to extract comprehensive information from the cancer nest and peritumoral microenvironment. Our results verified that a 10-mm peritumoral extension improved the predictive performance of the radiomics model. Peritumoral tissue has morphological and phenotypic differences from non-tumor-bearing healthy tissue.^{25–27} These substantial phenotypic and genetic changes are apparent up to 10 mm from the tumor margins. Lymphatic vessel and immune cells in peritumoral regions are an accurate predictor of occult LNM and prognosis.^{28–30} Also, a large number of studies have shown the power of peritumoral radiomics features for predicting tumor phenotype, recurrence, LNM, and prognosis.^{16–19} Similarity, a recent study demonstrated that features from the intratumoral and 10-mm peritumoral region provided the best performance for pretreatment prediction of response to chemotherapy in hepatocellular carcinoma.³¹ Our findings agree with prior studies in that the

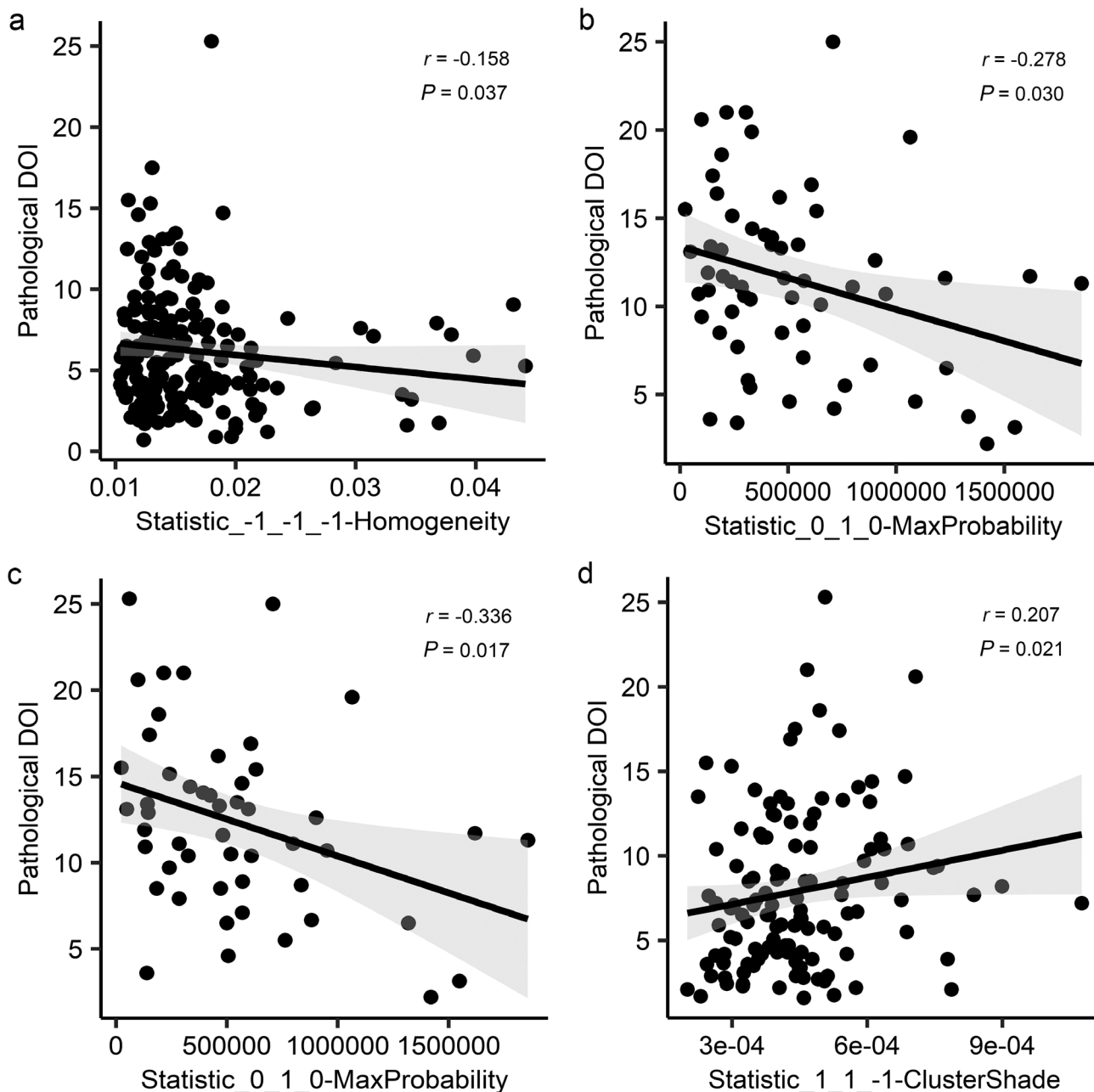


Figure 5: Spearman correlation analysis showed a significant correlation between pathological depth of invasion and three-dimensional radiomics features in radiological neurovascular invasion (NVI) negative subgroup (a), radiological NVI positive subgroup (b), radiological extrinsic lingual muscle invasion (EMI)-positive subgroup (c) and tumor budding number ≥ 5 subgroup (d), respectively.

peritumoral extension could potentially improve the ability of the radiomics model to predict LNM. According to the Eighth American Joint Committee on Cancer Guidelines, a “safe” excision of tongue cancer is routinely recommended at least 15–20 mm away from the primary tumor border.³² Studies have demonstrated that insufficient excision of tumor tissues can result in a poorer outcome.^{33,34} Also, our results verified that the en bloc resection with a margin more than 10 mm was relatively safe for patients with tongue cancer.

Numerous studies have suggested that DOI is an independent predictor of neck nodal metastasis, local recurrence,

and survival in patients with OC.^{6–8} For this reason, now DOI is included in routinely postoperative pathology reports.³² In our study, DOI was integrated into radiomics analysis. Results demonstrated that clinicopathological characteristics, including pathological DOI, enhanced the predictive power of radiomics model in LNM prediction. Similarly, previous radiomics studies have shown that clinicopathological characteristics have incremental value in LNM and prognosis prediction.^{20,35} We further conducted the correlation analysis between pathological DOI and radiomics features. Although no significant correlation was observed in all samples,

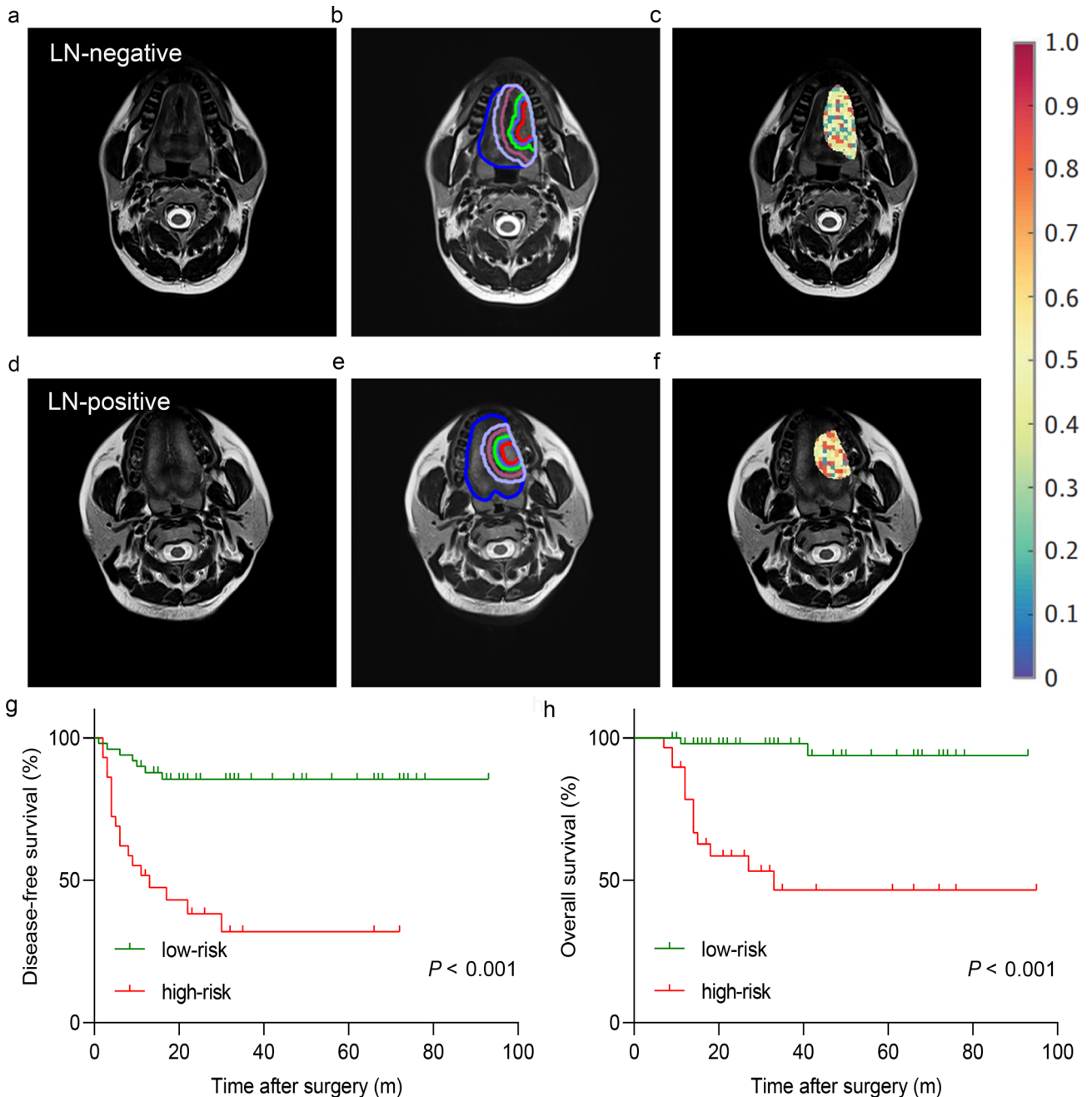


Figure 6: Axial MR images of tongue cancer in a 33-year-old patient without LNM and a 43-year-old patient with LNM, respectively (a,d). Segmentation and three-dimensional peritumoral extensions of the primary tumor (b,e). Heat map of radiomics feature extracted from the primary tumor with a 10-mm peritumoral extension showed differences between LN-negative and LN-positive patients (c,f). Kaplan–Meier survival analysis indicated that the $CR_{prim+10}$ -predicted high-risk of nodal metastasis had a poor disease-free survival and overall survival in the testing set (g,h).

pathological DOI were significantly correlated with homogeneity, maximum probability, and cluster shade in different subgroups, including radiological NVI negative, radiological NVI positive, radiological EMI-positive and tumor budding number ≥ 5 subgroups. Homogeneity feature and maximum probability feature have been associated with survival and metastasis in glioblastoma patients.³⁶ Cluster shade feature has been significantly correlated with recurrence-free survival in breast cancer.³⁷ These results indicate that radiomics features could evaluate tumor heterogeneity to some extent.

The traditional MRI reports have the disadvantage of low sensitivity and may lead to delayed treatment and a poor prognosis for OC patients with occult LNM. A previous study has demonstrated that radiomics analysis is able to classify LN status in clinically LN-negative subgroup.³⁸ Similarly, in our study, the $CR_{prim+10}$ showed good predictive properties in the MRI-reported LN-negative subgroup, indicating its ability to detect micrometastasis. The $CR_{prim+10}$ greatly improved the sensitivity in LNM prediction, suggesting that patients with tongue cancer would

TABLE 3. Univariate and Multivariate Cox Regression Analysis of DFS and OS in the Testing Set

Characteristics	DFS		OS					
	Univariate analysis		Multivariate analysis		Univariate analysis		Multivariate analysis	
	Hazard ratio	<i>P</i> value	Hazard ratio	<i>P</i> value	Hazard ratio	<i>P</i> value	Hazard ratio	<i>P</i> value
Age (<60/≥60)	0.465 (0.159–1.354)	0.160			0.743 (0.210–2.64)	0.646		
Sex (Female/male)	1.557 (0.672–3.610)	0.302			0.905 (0.322–2.55)	0.850		
T-staging (T1–T2/T3–T4)	1.238 (0.562–2.727)	0.596			1.841 (0.667–5.079)	0.238		
MRI-reported LN status (negative/positive)	1.629 (0.680–3.901)	0.273			3.256 (1.151–9.208)	0.026	0.979 (0.245–3.906)	0.976
Radiological NVI (negative/positive)	0.772 (0.290–2.059)	0.606			1.339 (0.426–4.212)	0.618		
Radiological DOI (≤10 mm/>10 mm)	1.238 (0.562–2.727)	0.596			1.841 (0.667–5.079)	0.238		
Radiological EMI (negative/positive)	1.170 (0.439–3.120)	0.754			1.897 (0.602–5.980)	0.275		
Radiological LSI (negative/positive)	1.559 (0.535–4.545)	0.416			3.201 (1.018–10.068)	0.047	4.501 (0.983–20.599)	0.053
Tumor budding (low/high)	3.255 (1.402–7.559)	0.006	2.389 (1.015, 5.624)	0.046	2.760 (0.943–8.079)	0.064		
Pathological DOI (≤10 mm/>10 mm)	0.975 (0.389–2.442)	0.957			0.883 (0.249–3.133)	0.847		
CR _{prim+10} -predicted LN status (negative/positive)	6.230 (2.587–15.004)	<0.001	5.250 (2.152–12.808)	<0.001	15.130 (3.393, 67.472)	<0.001	17.464 (3.660–83.330)	<0.001

DFS = disease-free survival; OS = overall survival; LN = lymph node; NVI = neurovascular invasion; DOI = depth of invasion; EMI = extrinsic lingual muscle invasion; LSI = lingual-septum invasion; CR_{prim+10} = combined radiomics model with a 10-mm peritumoral extension.

greatly benefit from the radiomics models in terms of treatment strategy.

Limitations

First, this was a retrospective study with relatively small sample size. A prospective multicenter validation with larger

sample size is needed. Considering that the tumor border could be easily identified based on the most commonly used T2-weighted images, we only chose the representative T2-weighted sequence to build radiomics models. However, this approach may lose complementary information on tumor heterogeneity, which is available from other sequences, such

as T1-weighted sequences, contrast-enhanced sequences. Third, studies have shown the potential of radiomics integrated with genomics.^{39,40} However, the unavailability of genomic information from this retrospective study hindered us to further enhance our model.

Conclusions

We have proposed a radiomics model with a 10-mm peritumoral extension and clinicopathological characteristics incorporated to accurately predict LNM and prognosis in tongue cancer. This model exhibited powerful discrimination ability and sensitivity in LNM prediction, especially in MRI-reported LN-negative patients.

Acknowledgments

We thank Dr. Zhi Wang and Jin Liu from the Blot Info & Tech (Beijing) Co. Ltd for radiomics analysis. We thank Professor Xiangbo Wan from the Department of Radiology, the Sixth Affiliated Hospital of Sun Yat-Sen University, for data analysis and guidance. This work was supported by the National Natural Science Foundation of China (82172752, 81772889, 81874128, 82072994), Sun Yat-Sen University Clinical Research 5010 Program (2015018), and High-Level Introduction of Talent Research Start-Up Fund of Nanfang Hospital, Southern Medical University.

Conflict of interest

The authors declare no potential conflict of interest.

Abbreviations

3D	three dimensional
AUC	area under the receiver operating characteristic curve
CI	confidence interval
CP	clinicopathological model
CR	combined radiomics models
DFS	disease-free survival
DOI	depth of invasion
EMI	extrinsic lingual muscle invasion
HOG	histogram of oriented gradient
ICC1	intraclass correlation coefficient
ICC2	interclass correlation coefficient
LN	lymph node
LNM	lymph node metastasis
LSI	lingual-septum invasion
MRI	magnetic resonance imaging
NVI	neurovascular invasion
OC	oral cancer
OS	overall survival
R	radiomics models
ROI	region-of-interest
SVM	support vector machine

References

- Sung H, Ferlay J, Siegel RL, et al. Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin* 2021;71:209-249.
- Leemans CR, Braakhuis BJ, Brakenhoff RH. The molecular biology of head and neck cancer. *Nat Rev Cancer* 2011;11:9-22.
- Chinn SB, Myers JN. Oral cavity carcinoma: Current management, controversies, and future directions. *J Clin Oncol* 2015;33:3269-3276.
- Pantel K, Brakenhoff RH. Dissecting the metastatic cascade. *Nat Rev Cancer* 2004;4:448-456.
- Lim SC, Zhang S, Ishii G, et al. Predictive markers for late cervical metastasis in stage I and II invasive squamous cell carcinoma of the oral tongue. *Clin Cancer Res* 2004;10:166-172.
- Tam S, Amit M, Zafereo M, Bell D, Weber RS. Depth of invasion as a predictor of nodal disease and survival in patients with oral tongue squamous cell carcinoma. *Head Neck* 2019;41:177-184.
- Yu P, Wang W, Zhuang Z, et al. A novel prognostic model for tongue squamous cell carcinoma based on the characteristics of tumour and its microenvironment: iBD score. *Histopathology* 2019;74:766-779.
- Xie N, Wang C, Liu X, et al. Tumor budding correlates with occult cervical lymph node metastasis and poor prognosis in clinical early-stage tongue squamous cell carcinoma. *J Oral Pathol Med* 2015;44:266-272.
- Bello IO, Soini Y, Salo T. Prognostic evaluation of oral tongue cancer: Means, markers and perspectives (I). *Oral Oncol* 2010;46:630-635.
- Raudenska M, Sztalmachova M, Gumulec J, et al. Prognostic significance of the tumour-adjacent tissue in head and neck cancers. *Tumour Biol* 2015;36:9929-9939.
- Ganci F, Sacconi A, Manciocco V, et al. Altered peritumoral microRNA expression predicts head and neck cancer patients with a high risk of recurrence. *Mod Pathol* 2017;30:1387-1401.
- de Bondt RB, Nelemans PJ, Hofman PA, et al. Detection of lymph node metastases in head and neck cancer: A meta-analysis comparing US, USgFNAC, CT and MR imaging. *Eur J Radiol* 2007;64:266-272.
- Si J, Huang S, Shi H, et al. Usefulness of 3T diffusion-weighted MRI for discrimination of reactive and metastatic cervical lymph nodes in patients with oral squamous cell carcinoma: A pilot study. *Dentomaxillofac Radiol* 2014;43:20130202.
- Buckley JG, MacLennan K. Cervical node metastases in laryngeal and hypopharyngeal cancer: A prospective analysis of prevalence and distribution. *Head Neck* 2000;22:380-385.
- Lambin P, Rios-Velazquez E, Leijenaar R, et al. Radiomics: Extracting more information from medical images using advanced feature analysis. *Eur J Cancer* 2012;48:441-446.
- Kim S, Shin J, Kim DY, Choi GH, Kim MJ, Choi JY. Radiomics on gadoteric acid-enhanced magnetic resonance imaging for prediction of postoperative early and late recurrence of single hepatocellular carcinoma. *Clin Cancer Res* 2019;25:3847-3855.
- Braman N, Prasanna P, Whitney J, et al. Association of peritumoral radiomics with tumor biology and pathologic response to preoperative targeted therapy for HER2 (ERBB2)-positive breast cancer. *JAMA Netw Open* 2019;2:e192561.
- Hu Y, Xie C, Yang H, et al. Assessment of intratumoral and peritumoral computed tomography radiomics for predicting pathological complete response to neoadjuvant chemoradiation in patients with esophageal squamous cell carcinoma. *JAMA Netw Open* 2020;3:e2015927.
- Wu Q, Wang S, Chen X, et al. Radiomics analysis of magnetic resonance imaging improves diagnostic performance of lymph node metastasis in patients with cervical cancer. *Radiother Oncol* 2019;138:141-148.
- Park H, Lim Y, Ko ES, et al. Radiomics signature on magnetic resonance imaging: Association with disease-free survival in patients with invasive breast cancer. *Clin Cancer Res* 2018;24:4705-4714.

21. DeLong ER, DeLong DM, Clarke-Pearson DL. Comparing the areas under two or more correlated receiver operating characteristic curves: A nonparametric approach. *Biometrics* 1988;44:837-845.
22. Vickers AJ, Cronin AM, Elkin EB, Gonen M. Extensions to decision curve analysis, a novel method for evaluating diagnostic tests, prediction models and molecular markers. *BMC Med Inform Decis Mak* 2008; 8:53.
23. Ho TY, Chao CH, Chin SC, Ng SH, Kang CJ, Tsang NM. Classifying neck lymph nodes of head and neck squamous cell carcinoma in MRI images with radiomic features. *J Digit Imaging* 2020;33:613-618.
24. Yuan Y, Ren J, Tao X. Machine learning-based MRI texture analysis to predict occult lymph node metastasis in early-stage oral tongue squamous cell carcinoma. *Eur Radiol* 2021;31:6429-6437.
25. Gerweck LE, Seetharaman K. Cellular pH gradient in tumor versus normal tissue: Potential exploitation for the treatment of cancer. *Cancer Res* 1996;56:1194-1198.
26. Aran D, Camarda R, Odegaard J, et al. Comprehensive analysis of normal tissue adjacent to tumor transcriptomes. *Nat Commun* 2017;8:1077.
27. Heaphy CM, Bisoffi M, Fordyce CA, et al. Telomere DNA content and allelic imbalance demonstrate field cancerization in histologically normal tissue adjacent to breast tumors. *Int J Cancer* 2006;119:108-116.
28. Mermod M, Bongiovanni M, Petrova TV, et al. Prediction of occult lymph node metastasis in squamous cell carcinoma of the oral cavity and the oropharynx using peritumoral Prospero homeobox protein 1 lymphatic nuclear quantification. *Head Neck* 2016;38:1407-1415.
29. Cho YA, Yoon HJ, Lee JI, Hong SP, Hong SD. Relationship between the expressions of PD-L1 and tumor-infiltrating lymphocytes in oral squamous cell carcinoma. *Oral Oncol* 2011;47:1148-1153.
30. Jardim JF, Gondak R, Galvis MM, Pinto CAL, Kowalski LP. A decreased peritumoral CD1a+ cell number predicts a worse prognosis in oral squamous cell carcinoma. *Histopathology* 2018;72:905-913.
31. Chen M, Cao J, Hu J, et al. Clinical-radiomic analysis for pre-treatment prediction of objective response to first transarterial chemoembolization in hepatocellular carcinoma. *Liver Cancer* 2021;10:38-51.
32. Pfister DG, Spencer S, Adelstein D, et al. Head and neck cancers, version 2.2020, NCCN clinical practice guidelines in oncology. *J Natl Compr Canc Netw* 2020;18:873-898.
33. Kurita H, Nakanishi Y, Nishizawa R, et al. Impact of different surgical margin conditions on local recurrence of oral squamous cell carcinoma. *Oral Oncol* 2010;46:814-817.
34. Slootweg PJ, Hordijk GJ, Schade Y, van Es RJJ, Koole R. Treatment failure and margin status in head and neck cancer. A critical view on the potential value of molecular pathology. *Oral Oncol* 2002;38: 500-503.
35. Yang J, Wu Q, Xu L, et al. Integrating tumor and nodal radiomics to predict lymph node metastasis in gastric cancer. *Radiother Oncol* 2020;150:89-96.
36. Beckers RCJ, Trebeschi S, Maas M, et al. CT texture analysis in colorectal liver metastases and the surrounding liver parenchyma and its potential as an imaging biomarker of disease aggressiveness, response and survival. *Eur J Radiol* 2018;102:15-21.
37. Mazurowski MA, Saha A, Harowicz MR, Cain EH, Marks JR, Marcom PK. Association of distant recurrence-free survival with algorithmically extracted MRI characteristics in breast cancer. *J Magn Reson Imaging* 2019;49:e231-e240.
38. 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;23:6904-6911.
39. Chaddad A, Daniel P, Sabri S, Desrosiers C, Abdulkarim B. Integration of radiomic and multi-omic analyses predicts survival of newly diagnosed IDH1 wild-type glioblastoma. *Cancers (Basel)* 2019;11:1148.
40. Sun R, Limkin EJ, Vakalopoulou M, et al. A radiomics approach to assess tumour-infiltrating CD8 cells and response to anti-PD-1 or anti-PD-L1 immunotherapy: An imaging biomarker, retrospective multi-cohort study. *Lancet Oncol* 2018;19:1180-1191.