Wang *et al.*
European Journal of Medical Research (2025) 30:164
https://doi.org/10.1186/s40001-025-02438-1

European Journal of Medical Research

RESEARCH

Open Access

Prediction of peripheral lymph node metastasis (LNM) in thyroid cancer using delta radiomics derived from enhanced CT combined with multiple machine learning algorithms

Wenzhi Wang[1†], Feng Jin[1†], Lina Song[1†], Jinfang Yang[1,2], Yingjian Ye[1], Junjie Liu[1], Lei Xu[1,2*] and Peng An[1,2*]

Abstract

Objectives This study aimed to develop a model for predicting peripheral lymph node metastasis (LNM) in thyroid cancer patients by combining enhanced CT radiomic features with machine learning algorithms. It increased the clinical utility and interpretability of the predictions through SHAP (SHapley Additive exPlanation) values and nomograms for model explanation and visualization.

Methods Clinical and enhanced CT image data from 375 patients with thyroid cancer confirmed by postoperative pathology at Xiangyang No. 1 People's Hospital were collected from January 2015 to July 2023. Among them, there were 88 patients in the LNM group and 287 patients in the non-LNM group. The delta radiomic features of the tumours were extracted. Various machine learning algorithms (such as SVM, GBM, RF, XGBoost, KNN, and LightGBM) were trained on the clinical and radiomic feature data sets and used to construct a reliable prediction model. During model training, cross-validation was used to evaluate model performance, and the optimal model was selected. In addition, SHAP values were used to interpret the prediction results of the optimal model, analyse the contribution of each feature to the prediction results, and further develop a nomogram to visually display the prediction results.

Results Univariate analysis confirmed that sex, Hashimoto's disease, tumour adjacency to the thyroid capsule, pathological subtype, Delta Radscore, and Radscore 1 are risk factors for peripheral lymph node metastasis in thyroid cancer patients. The machine learning model based on enhanced CT radiomics performed well in predicting peripheral lymph node metastasis in thyroid cancer patients. In the test set, the optimal model, SVM, achieved high AUC (0.879), sensitivity (0.849), and specificity (0.769) values. Through SHAP value analysis, the importance and contribution of tumour adjacency to the thyroid capsule, pathological subtype, Delta Radscore, and Radscore 1 in the prediction were clarified, providing a more detailed and intuitive basis for clinical decision-making. The nomogram illustrated the model prediction process, facilitating understanding and application by clinicians.

†Wenzhi Wang, Feng Jin and Lina Song contributed equally to this work.

*Correspondence:
Lei Xu
xyyy2025@foxmail.com
Peng An
drpengan@foxmail.com
Full list of author information is available at the end of the article

Conclusions This study successfully constructed a model for predicting peripheral lymph node metastasis in thyroid cancer patients on the basis of enhanced CT radiomics combined with machine learning and improved the interpretability and clinical utility of the model through SHAP values and nomograms. The model not only improves the accuracy of predictions but also provides a more scientific and intuitive basis for clinical decision-making, with potential clinical application value.

Keywords Thyroid cancer, Machine learning, Nomogram, Radiomics

Introduction

Thyroid cancer is one of the most common endocrine malignancies, with an increasing incidence rate (19.42 per 100,000) in China. According to the assessment report released by the International Agency for Research on Cancer (IARC) of the World Health Organization, there were over 820,000 new cases of thyroid cancer worldwide in 2022, ranking seventh in cancer incidence. Although the mortality rate of thyroid cancer is relatively low, its threat to public health cannot be ignored, as approximately 20–30% of thyroid cancers are accompanied by peripheral lymph node metastasis, requiring extensive surgical dissection and radiation therapy.

The epidemiological characteristics of thyroid cancer indicate that its incidence increases with age and that the prevalence is significantly greater in women than in men. This phenomenon may be related to hormonal dysfunction, and studies have shown that biological changes during pregnancy may increase the risk of thyroid cancer. In addition, iodine deficiency or excess may also be associated with the onset of thyroid cancer [1–3]. The pathological types of thyroid cancer are diverse and include mainly papillary carcinoma, follicular carcinoma, medullary carcinoma, and anaplastic carcinoma. Among them, papillary carcinoma is the most common, accounting for 60–70% of thyroid cancers; it has a good prognosis, especially in young patients. Follicular carcinoma accounts for 20–25%, with a relatively poor prognosis and a tendency for haematogenous metastasis. Medullary carcinoma accounts for 5–10% of cases, with low malignancy but the potential for lymph node metastasis. Anaplastic carcinoma accounts for 5–10%, with the highest degree of malignancy and a very poor prognosis. The likelihood of lymph node metastasis varies greatly among different types of thyroid cancer. Therefore, accurate early prediction of peripheral lymph node metastasis in thyroid cancer is crucial for its treatment and prognosis [4–6]. Radiomics, as an emerging image analysis method, converts medical images into high-throughput and quantitative features through computer algorithms and has shown great potential in the diagnosis and prognostic evaluation of various tumours. In thyroid research, radiomics has been used to differentiate benign and malignant thyroid nodules with significant success. However, studies on the prediction of peripheral lymph node metastasis in thyroid cancer patients via delta radiomics combined with machine learning have not been reported. This study aimed to explore the feasibility and accuracy of predicting peripheral lymph node metastasis in thyroid cancer patients via enhanced CT radiomics combined with machine learning. By extracting and analysing delta radiomic features from enhanced CT images and combining them with machine learning algorithms, multiple prediction models have been established to provide new auxiliary means for the clinical diagnosis and treatment of thyroid cancer [7, 8].

Methods

Patient enrolment

Clinical and enhanced CT data were collected from 512 patients with thyroid cancer confirmed by postoperative pathology at Xiangyang No. 1 People's Hospital from January 2015 to July 2023. The inclusion criteria were as follows: ① patients who underwent enhanced CT scanning preoperatively and had a single thyroid nodule; ② patients whose postoperative results confirmed thyroid cancer, including papillary carcinoma, follicular carcinoma, medullary carcinoma, and anaplastic carcinoma; ③ patients who did not undergo preoperative biopsy, radiotherapy, or chemotherapy; and ④ patients whose single thyroid mass had a diameter $\geq$ 6 mm. The exclusion criteria were as follows: ① patients with multiple thyroid nodules; ② patients with a tumour diameter < 6 mm; ③ patients with poor-quality enhanced CT images or significant motion artefacts; ④ patients with a history of other malignant tumours; ⑤ patients who were allergic to CT contrast agents; and ⑥ pregnant or lactating women. Ultimately, 375 patients with thyroid cancer were enrolled, including 88 in the lymph node metastasis group and 287 in the no lymph node metastasis group. They were randomly divided into a training set and a test set at a ratio of 7:3. The collected clinical data included age; sex; microcalcifications; aspect ratio of the tumour; maximum tumour diameter; Hashimoto's disease; tumour location; tumour adjacency to the thyroid capsule; pathological subtype; ultrasound elastography score; history of hypertension, diabetes, smoking and alcohol consumption; BMI; T3, T4, and TSH levels; and nodule CT values (unenhanced, arterial, and venous phases) [9–11] (Fig. 1). This study was approved by the Ethics Review Committee of Xiangyang No. 1 People's Hospital, with ethics approval number XYYYE20240011, and patient-identifiable information was removed from the CT images. Before enrolment, all selected patients were informed of the study purpose, methods, and potential risks and signed informed consent forms. During the study, we strictly adhered to ethical principles to protect patients' privacy and rights.

CT scanning method

CT scanning equipment and parameter settings: This study employed an advanced Toshiba 320-slice spiral CT scanner to ensure image clarity and accuracy. Patients underwent routine thin-slice plain and enhanced scans of the neck. The scanning parameters were set according to clinical standards and research requirements, with a tube voltage of 120 kV, a tube current of 150–300 mAs, a slice thickness of 5 mm, and a 1 mm thin-slice reconstruction with a slice interval of 1 mm. The window width was 250–350 HU, and the window level was 30–50 HU. For

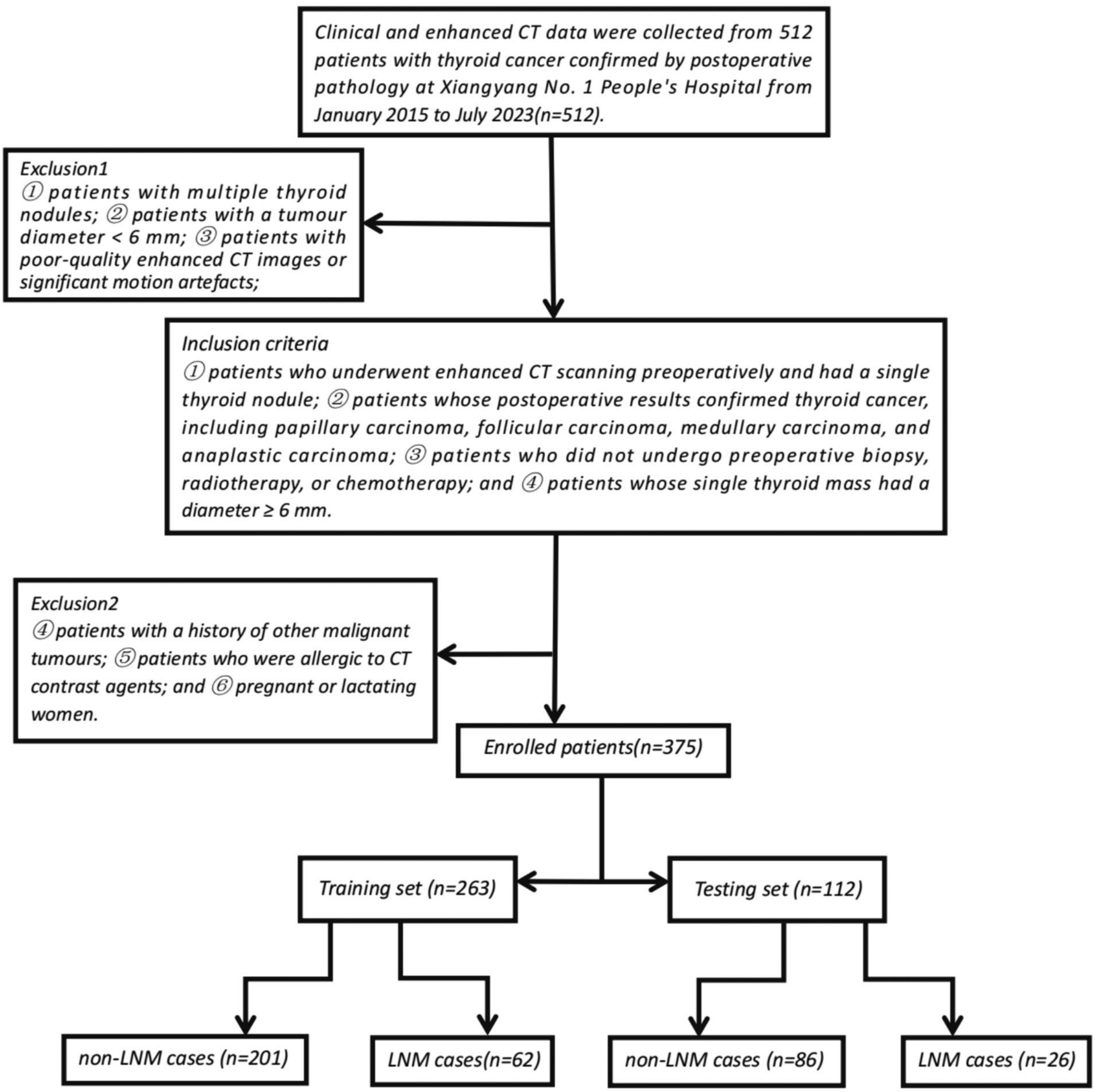

Fig. 1 Schematic diagram of case grouping and division of the training set and test set in this study

enhanced scanning, a nonionic contrast agent, iohexol (containing 300 mg/mL iodine), was used at a dose of 1.0–1.5 mL/kg, with an injection rate of 1 mL/s. The scan was triggered when the aortic threshold reached 180 HU, and a venous phase scan was performed with a delay of 30–60 s. The tube current was automatically adjusted according to the patient's body type to ensure image quality [11, 12].

Image acquisition and processing, region of interest (ROI) delineation, and radiomic feature selection

During the scanning process, the patient was asked to maintain a stable supine position to minimize the impact of motion artefacts on image quality. After scanning, the raw image data were transferred to an image postprocessing workstation for image reconstruction and preprocessing. Reconstruction algorithms such as filtered back projection or more advanced iterative reconstruction algorithms were employed to increase the signal-to-noise ratio and contrast of the images. Preprocessing steps included image denoising, correction, and normalization to ensure the accuracy and reliability of the subsequent radiomic analyses. Following image preprocessing, professional medical image processing software (3D Slicer) was used to manually delineate the region of interest (ROI) on the CT image slice displaying the maximum lesion area (Fig. 2). The ROI delineation was performed by two experienced professionals, a diagnostic physician with 8 years of experience and a radiographer with 10 years of experience, to ensure accuracy and consistency. The intraclass correlation coefficient (ICC) was used to evaluate the consistency of CT features delineated by the two individuals, and features with an ICC less than 0.75 were excluded. The radiomic feature selection process involved the following steps: ① Z score normalization was applied to the extracted radiomic features; ② significant features related to lymph node metastasis were retained through t tests or U tests ($P < 0.05$); ③ Pearson correlation coefficients were used to analyse the relationships between features, and for feature pairs with a high correlation (correlation coefficient > 0.9), only one was retained; and ④ the least absolute selection and shrinkage operator (LASSO) algorithm was used to reduce the dimensionality of the radiomic features. Then, the optimal log(λ) value was determined through tenfold cross-validation to select the best features and generate a radiomic score, the Radscore. The Delta Radscore represents the radiomic score generated by subtracting the texture parameters derived from the plain scan phase from those extracted from the arterial phase CT of thyroid cancer, whereas Radscore 1 represents

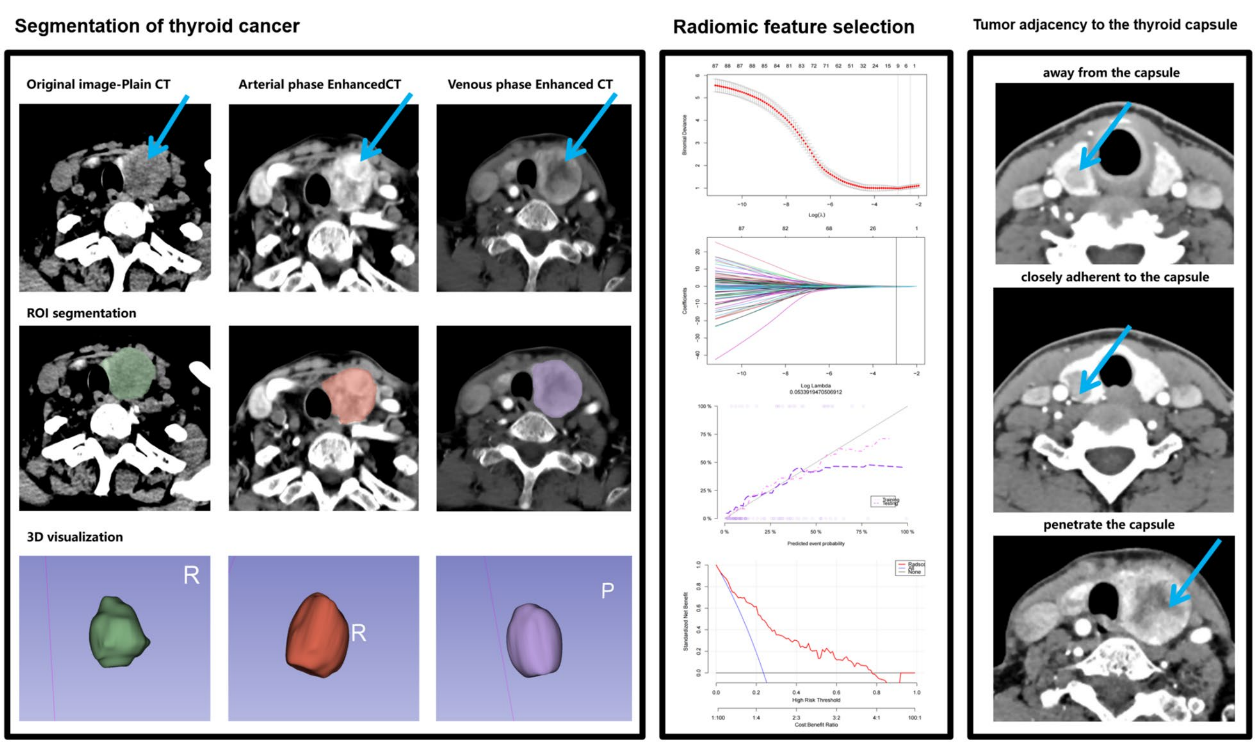

Fig. 2 Flowchart of radiomics feature extraction for thyroid cancer in this study and an illustrative diagram of the relationship between a thyroid tumour and its capsule

the radiomic score generated by the texture parameters extracted from the arterial phase CT of thyroid cancer [13–15].

Overview of statistics

In this study, statistical analyses were performed using R software. The measurement data for the two groups, if normally distributed, are presented as $X \pm S$ values and were analysed using independent sample t tests. The measurement data for the two groups, if not normally distributed, are presented as ranges (median, 25–75%) and were analysed using rank-sum tests. Categorical variables are presented as specific case counts and were analysed using chi-square tests or Fisher's exact tests. The cases were randomly divided into a training set and a test set at a ratio of 7:3. In both the training and test sets, meaningful clinical parameters and radiomic features were selected and used to construct models using ten machine learning algorithms: logistic regression (LR), support vector machine (SVM), gradient boosting machine (GBM), neural network, random forest (RF), extreme gradient boosting (XGBoost), K-nearest neighbour (KNN), adaptive boosting (AdaBoost), light gradient boosting machine (LightGBM), and CatBoost. The optimal model was selected in the test set, and the performances of the multimodal and unimodal models were compared and analysed via receiver operating characteristic (ROC) curves, decision curves, and calibration curves. Finally, the SHAP plot and corresponding sensitive and efficient SHAP values were derived from the optimal model in the test set. Ultimately, reliable factors selected on the basis of SHAP values were used to construct a novel nomogram for corresponding clinical application. A P value < 0.05 indicates a statistically significant difference [16, 17].

Results

This study aimed to predict the status of perithyroidal lymph node metastasis in thyroid cancer patients by combining enhanced CT radiomic features with machine learning algorithms and to conduct an in-depth analysis of the prediction results. The main findings of this study are as follows.

Analysis of clinical risk factors

By analysing the clinical data of the thyroid cancer patients included in this study, we identified several risk factors associated with perithyroidal lymph node metastasis in thyroid cancer patients. These factors include patient sex, Hashimoto's thyroiditis, tumour proximity to the thyroid capsule, and pathological subtype. Among these patients, those with tumours breaking through the adjacent thyroid capsule and those with well-differentiated pathological subtypes had a significantly increased risk of perithyroidal lymph node metastasis. However, common factors such as microcalcifications and age were not significantly correlated with lymph node metastasis in this study (Tables 1 and 2).

Radiomic feature analysis

Radiomic features extracted from enhanced CT images included shape features, texture features, wavelet features, and grey level co-occurrence matrix features. Through the LASSO algorithm combined with tenfold cross-validation for feature selection and dimensionality reduction, we identified the most predictive subset of radiomic features, including Inverse Variance...49, Grey Level Non-Uniformity...88, Size Zone Non-Uniformity Normalized...97, Run Length Non-Uniformity Normalized...733, Small Area Emphasis...749, Zone Variance...754, Skewness...774, Dependence Non-Uniformity...152, Dependence Non-Uniformity...431, Grey Level Non-Uniformity...460, Size Zone Non-Uniformity Normalized...748, Large Dependence High Grey Level Emphasis...810, and Inverse Variance...514. These features reflect the intratumoral heterogeneity and structural changes in surrounding tissues and are closely related to perithyroidal lymph node metastasis in thyroid cancer patients.

Comparison of model prediction performance

We employed various machine learning algorithms (support vector machines (SVMs), random forests, gradient boosting trees, etc.) to construct prediction models and evaluated their prediction performance through cross-validation. The results showed that in the training set, the random forest and LightGBM models exhibited overfitting, whereas the other models demonstrated good prediction efficacy. In the test set, the model based on the SVM algorithm exhibited the best performance in predicting perithyroidal lymph node metastasis in thyroid cancer patients, with a sensitivity of 0.849, a specificity of 0.769, and an AUC value of 0.879, outperforming the other algorithms. These findings indicate that the SVM algorithm has significant advantages in handling complex radiomic features and clinical factors (Figs. 3, 4, and 5).

Machine learning SHAP value analysis, nomogram construction, and validation

To further understand the prediction mechanism of the model, we used SHAP values (SHapley Additive exPlanations) to interpret the model's prediction results, ultimately confirming that the tumour adjacency to the thyroid capsule, pathological subtype, Delta Radscore, and Radscore 1 are correlated with perithyroidal lymph node metastasis in thyroid cancer. The SHAP value

Table 1 Comparison of the clinical and imaging data between the two groups

Factors	LNM group (N = 88)	non-LNM group (N = 287)	X^2, Z or t value	P
Sex			3.98	0.04*
Male	49	125		
Female	39	162		
Hashimoto's thyroiditis			4.87	0.03*
No	67	247		
Yes	21	40		
Microcalcifications			0.29	0.58
No	64	217		
Yes	24	70		
Tumour aspect ratio			1.56	0.21
1 (< 1)	63	224		
2 (> 1)	25	63		
Tumour location			1.39	0.16
1 (Upper lobe)	38	144		
2 (Middle lobe)	17	58		
3 (Inferior lobe)	33	85		
Tumour adjacency to the thyroid capsule			3.33	< 0.05*
1 (away from the capsule)	44	203		
2 (closely adherent to the capsule)	32	61		
3 (penetrated the capsule)	12	23		
Pathological subtype			4.72	< 0.05*
3 (papillary carcinoma)	41	199		
2 (follicular carcinoma)	32	75		
1 (medullary and anaplastic carcinoma)	15	13		
Grading of tumour ultrasound elastography			1.75	0.08
1	42	165		
2	17	61		
3	21	36		
4	4	16		
5	4	9		
Hypertension history	8.73 ± 7.26	13.31(0,11.13)	0.47	0.63
Diabetes history	3.31(0,6.23)	3.00(0,7.01)	0.52	0.61
Smoking history	8.78 ± 7.21	8.13 ± 6.89	0.76	0.45
Drinking history	9.98 ± 8.92	8.66 ± 8.77	1.23	0.22
Tumour diameter	11.37 ± 1.56	11.21 ± 1.64	0.84	0.39
BMI	22.70 ± 2.61	22.45 ± 2.62	0.79	0.42
Triiodothyronine (T3) level	1.26 ± 0.31	1.32 ± 0.35	1.49	0.14
Thyroxine(T4) level	9.38 ± 2.78	9.64 ± 2.91	0.75	0.45
Thyroid-stimulating hormone (TSH) level	3.75 ± 1.57	3.90 ± 1.68	0.73	0.46
Plain CT value	50.03 ± 10.16	54.90 ± 12.55	1.27	0.21
Arterial phase CT value	108.44 ± 20.67	103.60(86.70,114.31)	1.77	0.08
Venous phase CT value	102.63 ± 19.92	103.60(83.70,112.40)	1.81	0.07
Age	46.51 ± 7.62	46.04 ± 6.58	0.56	0.57
Delta Radscore	0.32 ± 0.08	0.21 ± 0.09	9.62	< 0.05*
Radscore 1	0.32 ± 0.09	0.21 ± 0.11	8.53	< 0.05*

*P < 0.05 indicates a statistically significant difference

Table 2 Results of the multivariate logistic regression model based on the abovementioned characteristics for predicting LNM in thyroid cancer patients; **P* < 0.05 indicates a statistically significant difference

Logistic regression model	Univariate analysis		Multivariate analysis	
Factors	***P***	**Hazard ratio**	***P***	**Hazard ratio**
Sex	0.04*	1.63(1.01–2.63)		
Hashimoto's thyroiditis	0.03*	1.94(1.07–3.50)		
Tumour adjacency to the thyroid capsule	< 0.05*	1.75(1.24–2.46)	0.04*	1.61(1.01–2.57)
Pathological subtype	< 0.05*	2.26(1.57–3.25)	< 0.05*	2.63(1.55–4.49)
Delta Radscore	< 0.05*	3.32(2.44–4.52)	< 0.05*	4.04(2.76–5.94)
Radscore 1	< 0.05*	2.62(2.01–3.43)	< 0.05*	3.02(2.15–4.25)

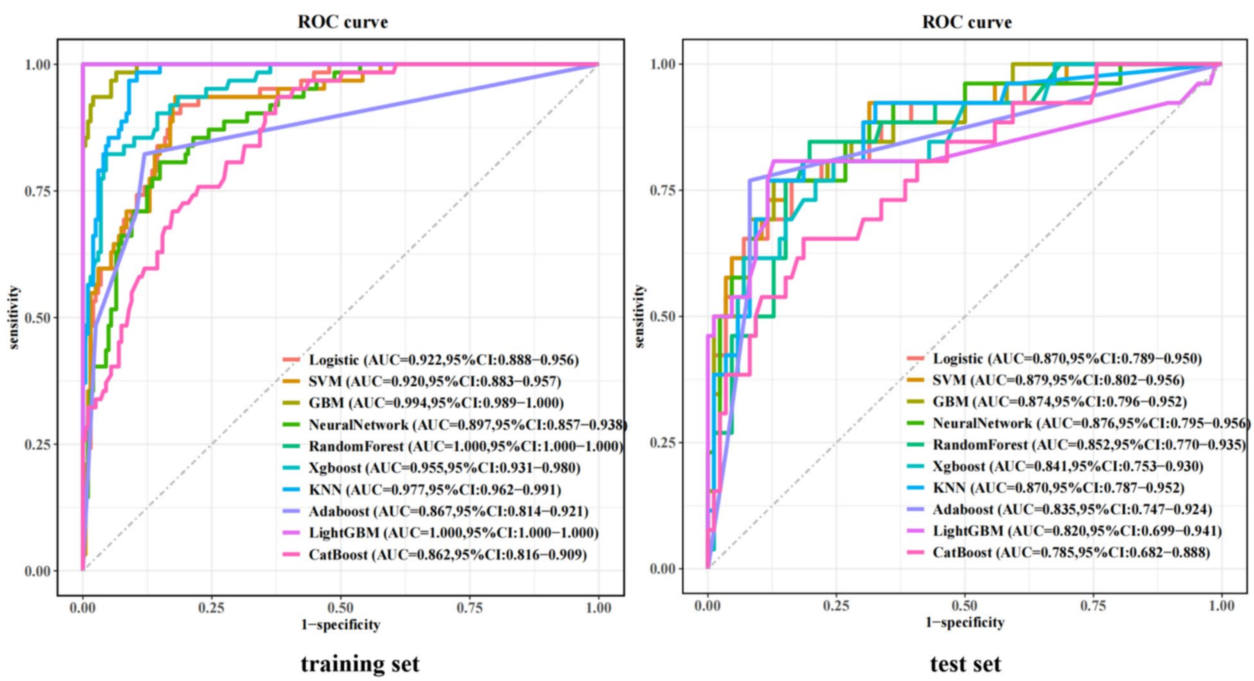

Fig. 3 In both the training and test sets, the ROC evaluation results of the machine learning model combining clinical features and the radiomics score revealed that, in the test set, the AUC value of the SVM as the optimal model was significantly greater than those of the other models

analysis revealed the contribution of clinical and radiomic parameters to the prediction results, allowing us to more clearly understand which features play a key role in predicting perithyroidal lymph node metastasis in thyroid cancer patients. On the basis of the prediction results of the SVM model and the SHAP value analysis, we constructed a novel nomogram to visually present the probability of perithyroidal lymph node metastasis in thyroid cancer patients. The nomogram combines clinical risk factors and radiomic features, providing clinicians with a simple and easy-to-use prediction tool. Through validation with the test set data, we found that the nomogram has good prediction performance and can accurately predict the risk of perithyroidal lymph node metastasis in thyroid cancer patients (Figs. 6 and 7).

Discussion

In recent years, due to factors, such as environmental pollution, drug abuse, and increased life stress in mainland China, the incidence of thyroid cancer has been increasing annually, reaching an annual growth rate of 22.86%. The incidence of thyroid cancer is particularly prominent in East Asia and North America. In 2022 alone, China reported 466,000 new cases, accounting for more than half of the global disease burden [1, 3, 18]. Although the majority of newly diagnosed thyroid cancers are papillary carcinomas, which typically do not cause symptoms or

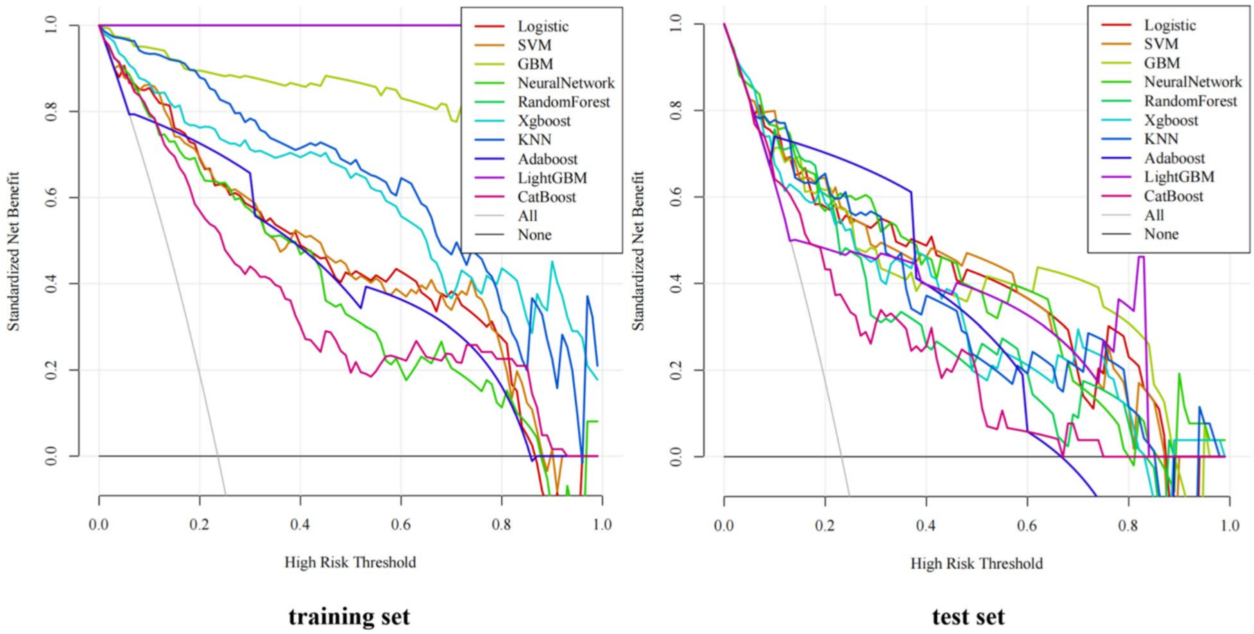

Fig. 4 For both the training and test sets, all the machine learning models demonstrated better net clinical benefits, and the optimal model, SVM, exceeded the average threshold for net clinical benefit in the test set

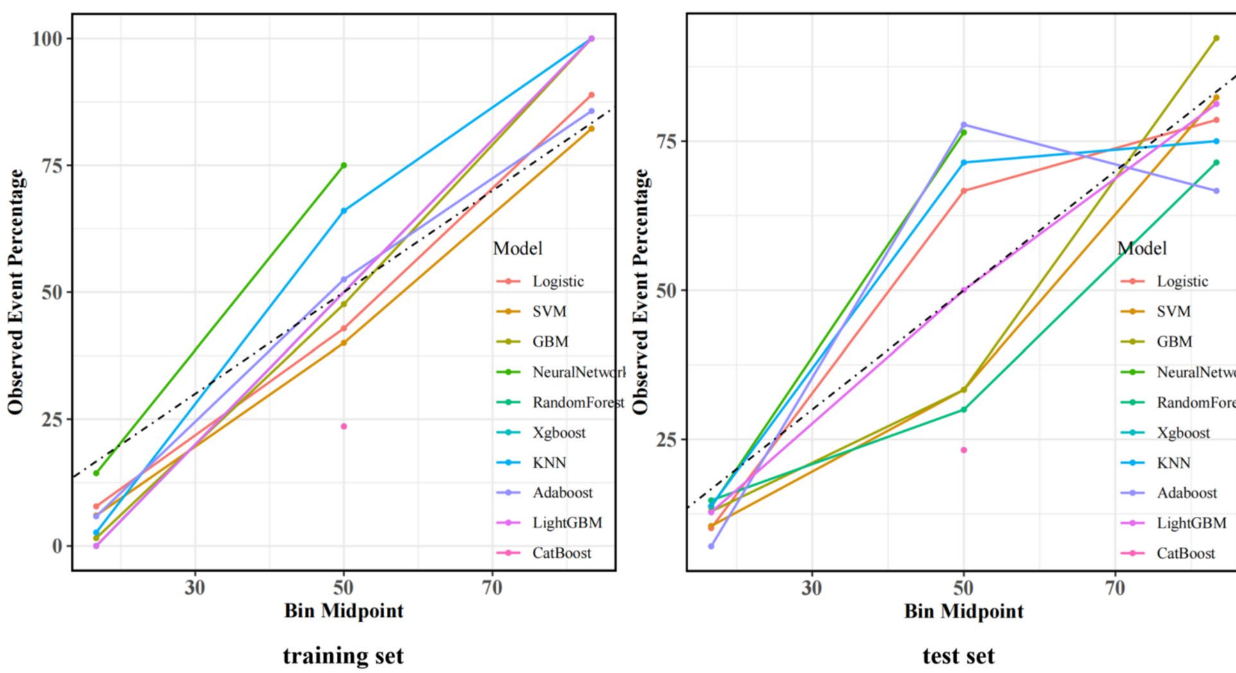

Fig. 5 For both the training and test sets, all the machine learning models achieved good calibration

lead to short-term mortality, the high incidence of lymph node metastasis, the difficulties in treating metastatic disease, and the anxiety and depression experienced by patients cannot be ignored.Cervical lymph node metastasis is a common pathway for thyroid cancer metastasis and is closely related to local recurrence and a poor prognosis. Accurate prediction of lymph node metastasis in thyroid cancer patients is crucial for formulating treatment plans and assessing patient prognosis. However, lymph node metastasis in thyroid cancer often lacks

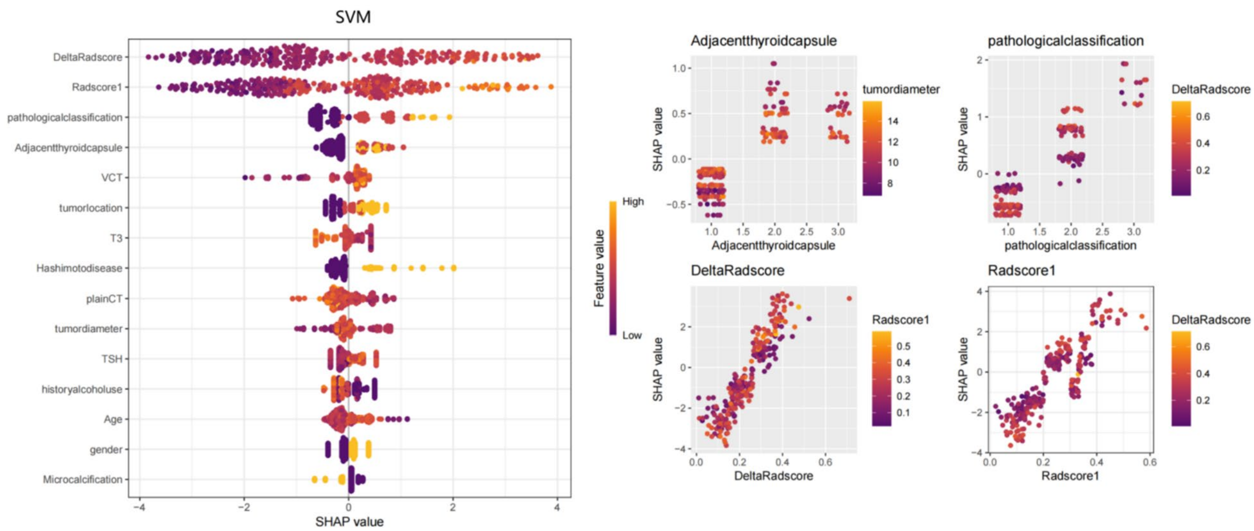

Fig. 6 Interpretable SHAP plot derived from the optimal SVM model in the test set confirmed that tumour adjacency to the thyroid capsule, pathological subtype, Delta Radscore, and Radscore 1 are the best predictors for LNM in thyroid cancer, whereas the SHAP values of subsequent factors mostly show bias on both sides or cluster in the centre, indicating poor predictive ability. Therefore, the correlation plot generated on the basis of the tumour adjacency to the thyroid capsule, pathological subtype, Delta Radscore, and Radscore 1 provides a good explanation for the prediction trend of the SVM model

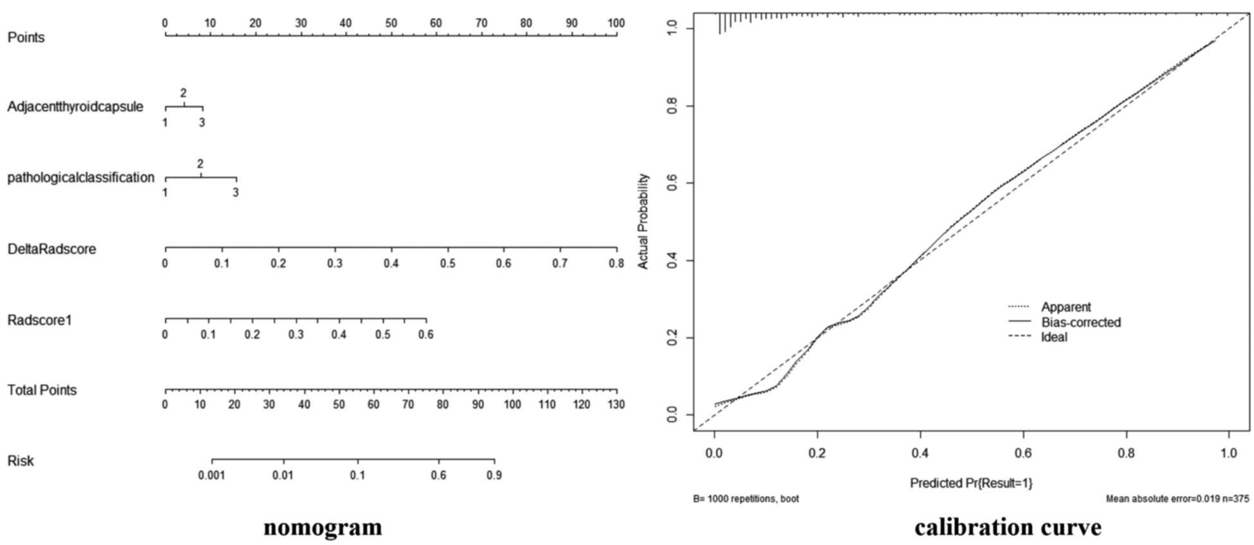

Fig. 7 Nomogram and calibration curve established on the basis of the effective factors with interpretable SHAP values have certain clinical application value. Clinical use: patient 498 from another hospital, with the lesion breaking through the capsule of the thyroid gland (10 points), was histologically classified as papillary carcinoma (18 points), with a Delta Radscore of 0.6 (75 points) and a Radscore 1 of 0.1 (10 points), for a total of 113 points, corresponding to a risk value > 0.9. Postoperative pathology confirmed metastasis to the left level IV cervical lymph nodes

obvious symptoms in the early stages, or the symptoms are easily overlooked. Symptoms such as small lumps in the neck and mild neck discomfort are often mistaken for muscle soreness caused by fatigue or poor posture. This stealthiness increases the difficulty of early diagnosis. Radiomics is increasingly being applied in the diagnosis and treatment decision-making of thyroid diseases. Through machine learning algorithms, radiomics can extract high-throughput and quantitative features from medical images, providing support for disease diagnosis, prognosis assessment, and treatment decision-making [18–20]. This study confirmed that radiomics has demonstrated its potential in predicting peripheral lymph node metastasis in thyroid cancer and is promising for improving the accuracy of diagnosis and the effectiveness of personalized treatment.

Lymph node metastasis in thyroid cancer is an important factor affecting the prognosis of thyroid cancer patients. This study confirmed that its occurrence is closely related to factors, such as sex, Hashimoto's disease, the proximity of the tumour to the thyroid capsule, pathological subtype, Delta Radscore, and Radscore 1. First, sex plays a certain role in the lymph node metastasis of thyroid cancer. This study indicates that although the incidence of thyroid cancer is lower in men than in women, male patients have a relatively higher risk of lymph node metastasis. This may be related to differences in hormone levels, immune status, and genetic background among male patients. In China, it may also be associated with increased work and life stress among men, a lack of time for medical attention, financial constraints preventing early diagnosis, or ignorance of enlarged neck lymph nodes. Hashimoto's disease, an autoimmune thyroid disease, has also attracted attention because of its relationship with lymph node metastasis in thyroid cancer. It has been reported that thyroid tissue in patients with Hashimoto's disease often exhibits chronic inflammation and immune cell infiltration, and this microenvironment may affect the biological behaviour of thyroid cancer cells, including their invasiveness and metastatic ability [21–23]. The proximity of the tumour to the thyroid capsule is one of the newly identified important factors affecting lymph node metastasis in thyroid cancer patients in this study. When the tumour is adjacent to, has invaded, or has broken through the thyroid capsule, the risk of lymph node metastasis significantly increases. This may be because when the tumour is near the capsule, it is more easily transferred to the neck lymph nodes through the lymphatic system. Furthermore, cancer cells breaking through the thyroid capsule often indicate that the disease has progressed to a certain stage, known as an advanced lesion. Thyroid cancer that invades the capsule is more prone to lymph node metastasis, because after the tumour cells break through the capsule, they more easily invade the surrounding tissues and lymphatic vessels, thereby flowing with the lymph nodes to the neck. Moreover, the pathological subtype of thyroid cancer is closely related to lymph node metastasis. Different pathological types of thyroid cancer present different characteristics in terms of lymph node metastasis. Papillary and follicular thyroid cancers are the most common types of thyroid cancer and are relatively well-differentiated and slow growing, making them difficult to detect and thus prone to long-term infiltrative growth and lymph node metastasis. Medullary and anaplastic thyroid cancers, on the other hand, are highly malignant and grow rapidly, so the opportunity for lymph node metastasis is curtailed during early diagnosis and timely surgical treatment. Therefore, early screening for thyroid cancer remains an important strategy for preventing neck lymph node metastasis [24–26].

Radiomics, as an emerging image analysis method, offers new insights into predicting lymph node metastasis (LNM) in thyroid cancer by extracting many quantitative features from medical images. Radiomic features can reflect biological information, such as tumour heterogeneity, angiogenesis, and cell proliferation, which are closely related to the invasiveness and metastatic capability of thyroid cancer. Therefore, radiomics has significant value in predicting LNM in thyroid cancer patients. In recent years, with the rapid development of artificial intelligence (AI) and machine learning (ML) technologies, their applications in the field of medical image analysis have become increasingly widespread. The combination of ML and radiomics has demonstrated tremendous potential in predicting LNM in thyroid cancer [27–29]. By constructing predictive models based on various ML algorithms, this study can extract the most relevant features associated with LNM in thyroid cancer from a vast array of clinical and radiomic features, thereby enabling precise prediction of LNM. The support vector machine (SVM) model was validated to have good predictive performance in the test set. SVM is a supervised binary classification algorithm whose basic idea is to find an optimal hyperplane that separates samples of different classes and maximizes the distance from the nearest sample points to the hyperplane. This hyperplane is determined by support vectors, which are the sample points closest to the hyperplane and play a decisive role in the final classification result. The SVM algorithm possesses powerful data processing and pattern recognition capabilities, enabling it to automatically learn complex nonlinear relationships from large data sets and provide reliable support for predicting binary classification outcomes. We may use grid search or random search algorithms to find the optimal combination of hyperparameters based on the cross-validation results, thus improving the performance of the SVM model. In predicting LNM in thyroid cancer, the SVM algorithm can process massive amounts of clinical and radiomic feature data, uncover hidden patterns and laws, and thereby enable early prediction of LNM. In addition, ML algorithms have strong generalization capabilities and can handle data of different types and sources, allowing for the prediction of LNM in thyroid cancer across different patients and pathological types. In practical applications, the combination of ML and radiomics can provide more personalized treatment plans for thyroid cancer patients. By analysing and predicting patients' radiomic features, doctors can assess their risk of LNM more accurately and formulate more reasonable treatment plans accordingly. For patients with a greater risk of LNM, a more aggressive surgical treatment plan, including expanded surgical resection, postoperative radiotherapy, and chemotherapy, can be adopted. Conversely, for patients with a lower risk of LNM, a more conservative treatment plan can be employed, reducing unnecessary surgical and medicinal interventions [30–32].

Limitations

Despite the tremendous potential of radiomics, which is based on enhanced CT combined with machine learning, for the prediction of peripheral lymph node metastasis in thyroid cancer patients, it still has certain limitations. First, the predictive performance of machine learning models highly depends on the quantity and quality of the training data, and currently available high-quality, large-scale data sets are relatively limited, which may result in insufficient generalization ability of the models and difficulty in adapting to the prediction needs of different patient populations. The interaction between radiomic features and clinical features (such as pathological subtype) has not been explored, potentially missing important predictive information. Second, the extraction and analysis process of radiomic features is complex and time-consuming, requiring highly specialized technology and equipment support, which limits its widespread application in clinical practice. Furthermore, the interpretability of machine learning models is relatively vague, and their internal working principles and decision-making processes are difficult to understand intuitively, which may lead to scepticism among clinicians and patients regarding their prediction results. Therefore, in future research, it will be necessary to further optimize machine learning algorithms, improve model generalizability, and enhance the clinical interpretability of radiomic features to promote the clinical application of this technology in predicting peripheral lymph node metastasis in thyroid cancer[33,34].

Conclusion

In summary, this study successfully constructed a model for predicting peripheral lymph node metastasis in thyroid cancer patients by combining enhanced CT radiomic features with machine learning algorithms and provided in-depth interpretation and visualization of the model's prediction results through SHAP value analysis and nomogram construction. These results provide strong support for the clinical diagnosis and treatment of thyroid cancer.

Acknowledgements
We thank all the volunteers for generously providing clinical and enhanced CT data.

Author contributions
An. Wang. and Jin. Song. wrote the main manuscript text and Ye. Liu. prepared Figs. 1, 2, 3, 4, 5, 6 and 7. All authors reviewed the manuscript. Writing—original draft, WZW and PA; conceptualization, FJ, PA; writing—review and editing, JFY, YJY, and PA; methodology, NLS, JJL, and LX. All authors have read and agreed to the published version of the manuscript. An and Xu approved the final version for publication and agreed to act as guarantors of the work. Wang, Jin, and Song contributed equally to this study.

Funding
This study was funded by the "323" Major Chronic Disease Project of the Hubei Provincial Health Commission and Xiangyang NO.1 People's Hospital (XYY2022-323).General Program of the Natural Science Foundation of Hubei Province (exploring the prediction of immunophenotyping and prognosis modeling for oropharyngeal cancer through multimodality imaging-pathomics combined with generalized machine learning—JCZRYB202501213-belonged to Peng An). Innovative Research Program of Xiangyang No.1 People's Hospital (XYY2025SD17).

Data availability
The data sets generated during and analyzed during the current study are available from the corresponding author on reasonable request.

Declarations

Ethics approval and consent to participate
The framework of this study was developed in accordance with the guidelines of the Declaration of Helsinki and has been reviewed and approved by the Ethics Committee of Xiangyang NO.1 People's Hospital, Hubei University of Medicine (Approval No. XYYYE20240011).

Consent for publication
Informed consent was obtained from all subjects involved in the research.

Competing interests
The authors declare no competing interests.

Author details
[1]Department of Cardiology, Radiology, and Surgery, Xiangyang No.1 People's Hospital, Hubei University of Medicine, Xiangyang 441000, China. [2]Department of Emergency, Oncology and Epidemiology, Xiangyang Key Laboratory of Maternal-Fetal Medicine on Fetal Congenital Heart Disease, Xiangyang No. 1 People'S Hospital, Hubei University of Medicine, Xiangyang 441000, Hubei, People's Republic of China.

Received: 25 January 2025 Accepted: 6 March 2025
Published online: 13 March 2025

References

1. Fei X, Lou Z, Christakos G, Liu Q, Ren Y, Wu J. Contribution of industrial density and socioeconomic status to the spatial distribution of thyroid cancer risk in Hangzhou. China Sci Total Environ. 2018;613–614:679–86. https://doi.org/10.1016/j.scitotenv.2017.08.270.
2. HajiEsmailPoor Z, Kargar Z, Tabnak P. Radiomics diagnostic performance in predicting lymph node metastasis of papillary thyroid carcinoma: a systematic review and meta-analysis. Eur J Radiol. 2023;168: 111129. https://doi.org/10.1016/j.ejrad.2023.111129.
3. Du L, Wang Y, Sun X, et al. Thyroid cancer: trends in incidence, mortality and clinical-pathological patterns in Zhejiang Province, Southeast China. BMC Cancer. 2018;18(1):291. https://doi.org/10.1186/s12885-018-4081-7.
4. Mu J, Cao Y, Zhong X, Diao W, Jia Z. Prediction of cervical lymph node metastasis in differentiated thyroid cancer based on radiomics models. Br J Radiol. 2024;97(1155):526–34. https://doi.org/10.1093/bjr/tqae010.
5. Ren Y, Lu S, Zhang D, Wang X, Agyekum EA, Zhang J, Zhang Q, Xu F, Zhang G, Chen Y, Shen X, Zhang X, Wu T, Hu H, Shan X, Wang J, Qian X. Dual-modal radiomics for predicting cervical lymph node metastasis in papillary thyroid

carcinoma. J Xray Sci Technol. 2023;31(6):1263–80. https://doi.org/10.3233/XST-230091.
6. Peng Y, Zhang ZT, Wang TT, Wang Y, Li CH, Zuo MJ, Lin HS, Gong LG. Prediction of Central Lymph Node Metastasis in cN0 Papillary Thyroid Carcinoma by CT Radiomics. Acad Radiol. 2023;30(7):1400–7. https://doi.org/10.1016/j.acra.2022.09.002.
7. Ren S, Zhao R, Cui W, Qiu W, Guo K, Cao Y, Duan S, Wang Z, Chen R. Computed tomography-based radiomics signature for the preoperative differentiation of pancreatic adenosquamous carcinoma from pancreatic ductal adenocarcinoma. Front Oncol. 2020;25(10):1618. https://doi.org/10.3389/fonc.2020.01618.
8. Dong L, Han X, Yu P, Zhang W, Wang C, Sun Q, Song F, Zhang H, Zheng G, Mao N, Song X. CT Radiomics-based nomogram for predicting the lateral neck lymph node metastasis in papillary thyroid carcinoma: a prospective multicenter study. Acad Radiol. 2023;30(12):3032–46. https://doi.org/10.1016/j.acra.2023.03.039.
9. Wang C, Yu P, Zhang H, Han X, Song Z, Zheng G, Wang G, Zheng H, Mao N, Song X. Artificial intelligence-based prediction of cervical lymph node metastasis in papillary thyroid cancer with CT. Eur Radiol. 2023;33(10):6828–40. https://doi.org/10.1007/s00330-023-09700-2.
10. Lu S, Ren Y, Lu C, Qian X, Liu Y, Zhang J, Shan X, Sun E. Radiomics features from whole thyroid gland tissue for prediction of cervical lymph node metastasis in the patients with papillary thyroid carcinoma. J Cancer Res Clin Oncol. 2023;149(14):13005–16. https://doi.org/10.1007/s00432-023-05184-1.
11. Mou Y, Han X, Li J, Yu P, Wang C, Song Z, Wang X, Zhang M, Zhang H, Mao N, Song X. Development and validation of a computed tomography-based radiomics nomogram for the preoperative prediction of central lymph node metastasis in papillary thyroid microcarcinoma. Acad Radiol. 2024;31(5):1805–17. https://doi.org/10.1016/j.acra.2023.11.030.
12. Feng JW, Liu SQ, Qi GF, Ye J, Hong LZ, Wu WX, Jiang Y. Development and validation of clinical-radiomics nomogram for preoperative prediction of central lymph node metastasis in papillary thyroid carcinoma. Acad Radiol. 2024;31(6):2292–305. https://doi.org/10.1016/j.acra.2023.12.008.
13. Gao Y, Wang W, Yang Y, Xu Z, Lin Y, Lang T, Lei S, Xiao Y, Yang W, Huang W, Li Y. An integrated model incorporating deep learning, hand-crafted radiomics and clinical and US features to diagnose central lymph node metastasis in patients with papillary thyroid cancer. BMC Cancer. 2024;24(1):69. https://doi.org/10.1186/s12885-024-11838-1.
14. Ren S, Qian LC, Cao YY, Daniels MJ, Song LN, Tian Y, Wang ZQ. Computed tomography-based radiomics diagnostic approach for differential diagnosis between early- and late-stage pancreatic ductal adenocarcinoma. World J Gastrointest Oncol. 2024;16(4):1256–67. https://doi.org/10.4251/wjgo.v16.i4.1256.
15. 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):1214–26. https://doi.org/10.21037/gs-20-479.
16. Yang G, Yang F, Zhang F, Wang X, Tan Y, Qiao Y, Zhang H. Radiomics profiling identifies the value of CT features for the preoperative evaluation of lymph node metastasis in papillary thyroid carcinoma. Diagnostics (Basel). 2022;12(5):1119. https://doi.org/10.3390/diagnostics12051119.
17. Li MH, Liu L, Feng L, Zheng LJ, Xu QM, Zhang YJ, Zhang FR, Feng LN. Prediction of cervical lymph node metastasis in solitary papillary thyroid carcinoma based on ultrasound radiomics analysis. Front Oncol. 2024;25(14):1291767. https://doi.org/10.3389/fonc.2024.1291767.
18. Wei T, Wei W, Ma Q, Shen Z, Lu K, Zhu X. Development of a clinical-radiomics nomogram that used contrast-enhanced ultrasound images to anticipate the occurrence of preoperative cervical lymph node metastasis in papillary thyroid carcinoma patients. Int J Gen Med. 2023;29(16):3921–32. https://doi.org/10.2147/IJGM.S424880.
19. Tong Y, Zhang J, Wei Y, Yu J, Zhan W, Xia H, Zhou S, Wang Y, Chang C. Ultrasound-based radiomics analysis for preoperative prediction of central and lateral cervical lymph node metastasis in papillary thyroid carcinoma: a multi-institutional study. BMC Med Imaging. 2022;22(1):82. https://doi.org/10.1186/s12880-022-00809-2.
20. Zhang H, Hu S, Wang X, He J, Liu W, Yu C, Sun Z, Ge Y, Duan S. Prediction of cervical lymph node metastasis using mri radiomics approach in papillary thyroid carcinoma: a feasibility study. Technol Cancer Res Treat. 2020. https://doi.org/10.1177/1533033820969451.
21. Jia W, Cai Y, Wang S, Wang J. Predictive value of an ultrasound-based radiomics model for central lymph node metastasis of papillary thyroid carcinoma. Int J Med Sci. 2024;21(9):1701–9. https://doi.org/10.7150/ijms.95022.
22. Shi Y, Zou Y, Liu J, Wang Y, Chen Y, Sun F, Yang Z, Cui G, Zhu X, Cui X, Liu F. Ultrasound-based radiomics XGBoost model to assess the risk of central cervical lymph node metastasis in patients with papillary thyroid carcinoma: Individual application of SHAP. Front Oncol. 2022;26(12): 897596. https://doi.org/10.3389/fonc.2022.897596.
23. Agyekum EA, Ren YZ, Wang X, Cranston SS, Wang YG, Wang J, Akortia D, Xu FJ, Gomashie L, Zhang Q, Zhang D, Qian X. Evaluation of cervical lymph node metastasis in papillary thyroid carcinoma using clinical-ultrasound radiomic machine learning-based model. Cancers (Basel). 2022;14(21):5266. https://doi.org/10.3390/cancers14215266.
24. Tong Y, Li J, Huang Y, Zhou J, Liu T, Guo Y, Yu J, Zhou S, Wang Y, Chang C. Ultrasound-based radiomic nomogram for predicting lateral cervical lymph node metastasis in papillary thyroid carcinoma. Acad Radiol. 2021;28(12):1675–84. https://doi.org/10.1016/j.acra.2020.07.017.
25. Zhou SC, Liu TT, Zhou J, Huang YX, Guo Y, Yu JH, Wang YY, Chang C. An ultrasound radiomics nomogram for preoperative prediction of central neck lymph node metastasis in papillary thyroid carcinoma. Front Oncol. 2020;4(10):1591. https://doi.org/10.3389/fonc.2020.01591.
26. Yan X, Mou X, Yang Y, Ren J, Zhou X, Huang Y, Yuan H. Predicting central lymph node metastasis in patients with papillary thyroid carcinoma based on ultrasound radiomic and morphological features analysis. BMC Med Imaging. 2023;23(1):111. https://doi.org/10.1186/s12880-023-01085-4.
27. Shen B, Zhou C, Xu C, Yang B, Wu X, Fu X, Liu S, Sun J, Xie Y, Zhu Z. Ultrasound-based radiomics for predicting metastasis in the lymph nodes posterior to the right recurrent laryngeal nerve in patients with papillary thyroid cancer. Curr Med Imaging. 2023. https://doi.org/10.2174/0115734056257332231024112410.
28. Liu Z, Zhang X, Zhao X, Guo Q, Li Z, Wei M, Niu L, An C. Combining radiomics with thyroid imaging reporting and data system to predict lateral cervical lymph node metastases in medullary thyroid cancer. BMC Med Imaging. 2024;24(1):64. https://doi.org/10.1186/s12880-024-01222-7.
29. Xue J, Li S, Qu N, Wang G, Chen H, Wu Z, Cao X. Value of clinical features combined with multimodal ultrasound in predicting lymph node metastasis in cervical central area of papillary thyroid carcinoma. J Clin Ultrasound. 2023;51(5):908–18. https://doi.org/10.1002/jcu.23465.
30. Zhang H, Li Y, Wu S, Peng Y, Liu Y, Gao S. Machine learning-based radiomics for guiding lymph node dissection in clinical stage I lung adenocarcinoma: a multicenter retrospective study. Transl Lung Cancer Res. 2024;13(12):3579–89. https://doi.org/10.21037/tlcr-24-668.
31. Daibo S, Homma Y, Ohya H, et al. Novel machine-learning model for predicting lymph node metastasis in resectable pancreatic ductal adenocarcinoma. Ann Gastroenterol Surg. 2024;9(1):161–8. https://doi.org/10.1002/ags3.12836.
32. Ke Z, Shen L, Shao J. Early warning of axillary lymph node metastasis in breast cancer patients using multi-omics signature: a machine learning-based retrospective study. Int J Gen Med. 2024;17:6101–14. https://doi.org/10.2147/IJGM.S499238.
33. Harmon SA, Sanford TH, Xu S, Turkbey EB, Roth H, Xu Z, Yang D, Myronenko A, Anderson V, Amalou A, Blain M, Kassin M, Long D, Varble N, Walker SM, Bagci U, Ierardi AM, Stellato E, Plensich GG, Franceschelli G, Girlando C, Irmici G, Labella D, Hammoud D, Malayeri A, Jones E, Summers RM, Choyke PL, Xu D, Flores M, Tamura K, Obinata H, Mori H, Patella F, Cariati M, Carrafiello G, An P, Wood BJ, Turkbey B. Artificial intelligence for the detection of COVID-19 pneumonia on chest CT using multinational datasets. Nat Commun. 2020;11(1):4080. https://doi.org/10.1038/s41467-020-17971-2.
34. Yang D, Xu Z, Li W, Myronenko A, Roth HR, Harmon S, Xu S, Turkbey B, Turkbey E, Wang X, Zhu W, Carrafiello G, Patella F, Cariati M, Obinata H, Mori H, Tamura K, An P, Wood BJ, Xu D. Federated semi-supervised learning for COVID region segmentation in chest CT using multi-national data from China Italy, Japan. Med Image Anal. 2021;70:101992. https://doi.org/10.1016/j.media.2021.101992.

Publisher's Note