

A Predictive Model for Lateral Cervical Lymph Node Metastasis in Thyroid Cancer

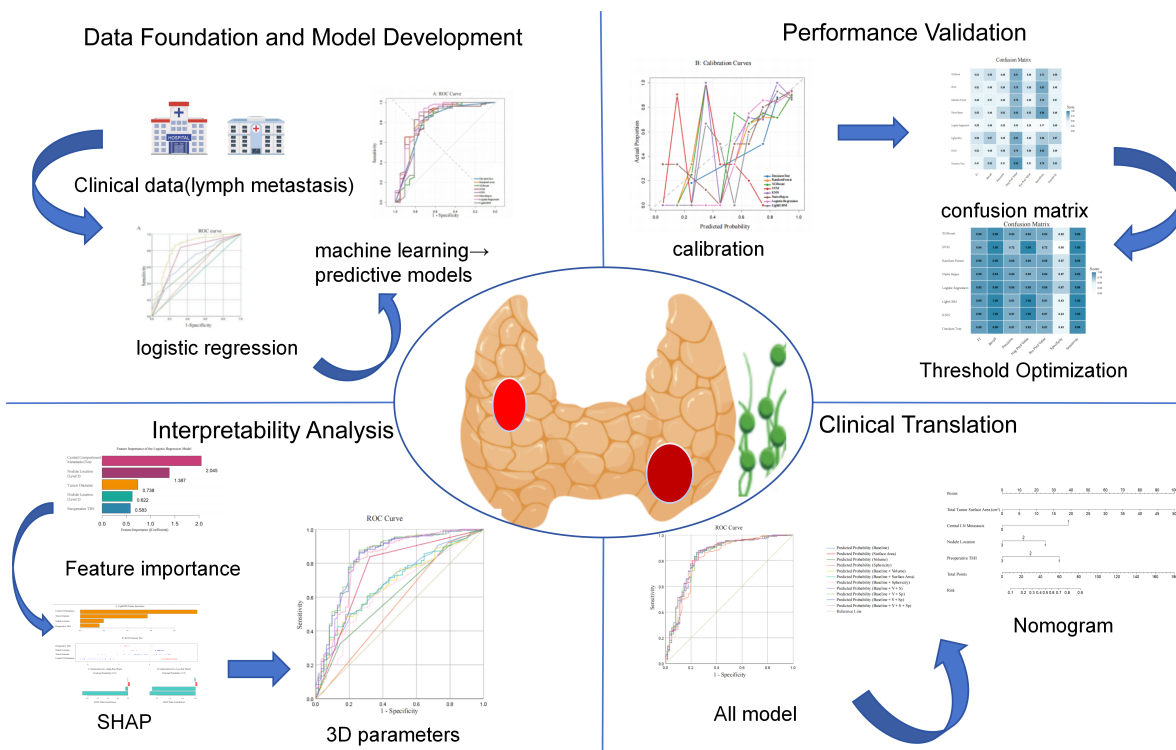
Authors

Shu Wang, Yanqiang Zhang, Ziyue Fu, Chuanlu Shen, Siyu Jia, Shugang Zhao, Kaile Wu

Correspondence

960092599@qq.com (S. Zhao), wukaile@fy.ahmu.edu.cn (K. Wu)

Graphical Abstract



<https://doi.org/10.71321/5r3ppj45>

© 2026 The Author(s). Published by Life Conflux Press Limited. This is an open access article distributed under the terms of the Creative Commons Attribution License (CC BY 4.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

A Predictive Model for Lateral Cervical Lymph Node Metastasis in Thyroid Cancer

Shu Wang¹, Yanqiang Zhang¹, Ziyue Fu¹, Chuanlu Shen¹, Siyu Jia², Shugang Zhao^{2*}, Kaile Wu^{1*}

Received: 2026-03-20 | Accepted: 2026-03-26 | Published online: 2026-04-03

Abstract

Objective: To investigate the predictive factors for lateral lymph node metastasis (LLNM) in patients with papillary thyroid carcinoma (PTC) and to develop an individualized prediction model.

Methods: Clinical data from 241 PTC patients who underwent lateral neck dissection were analyzed. Logistic regression and machine learning methods were employed to identify predictive factors and construct a model. The predictive value of three-dimensional morphological parameters (total tumor surface area and total tumor volume) was also evaluated.

Results: Maximum tumor diameter, central lymph node metastasis, preoperative Thyroid-Stimulating Hormone (TSH), and tumor location were identified as independent predictors of LLNM. A baseline combined model based on maximum tumor diameter showed good predictive performance (AUC 0.832). Furthermore, three-dimensional parameters (total surface area and total volume) demonstrated complementary predictive potential compared to the baseline model.

Conclusion: An effective clinical prediction model for assessing LLNM risk was successfully developed. Three-dimensional morphological parameters represent promising predictive indicators with potential complementary value.

Keywords: Papillary thyroid carcinoma; Lateral lymph node metastasis; Machine learning; Clinical prediction model; SHapley Additive exPlanations (SHAP)

Introduction

Thyroid cancer is a common malignant tumor of the endocrine system, with its global incidence continuing to rise [1-4]. Papillary thyroid carcinoma (PTC) is the most common pathological type. Patients with cervical lymph node metastasis have a higher risk of recurrence and distant metastasis, leading to a poorer prognosis [5-10]. Relevant guidelines recommend that PTC patients assessed preoperatively with lymph node metastasis should undergo surgical treatment [11]. Cervical lymph node dissection is a critical step in surgery, and the extent of dissection needs to balance complete tumor removal with postoperative quality of life. For lateral cervical lymph nodes, current guidelines recommend therapeutic dissection only when preoperative imaging suggests metastasis [3, 12]. However, existing imaging modalities have limitations in assessing lymph node metastasis [13-14]. Therefore, accurately predict-

ing lateral cervical lymph node metastasis is of great significance for formulating individualized surgical strategies. Current preoperative assessment primarily relies on ultrasound, CT, etc., but their sensitivity and specificity are limited [2, 15-18]. Preoperative lymph node biopsy is the standard method for evaluating LLNM [3]. However, ultrasound-guided fine-needle aspiration (FNA) has limitations in accuracy and carries the potential for sampling error [3, 19]. Constructing prediction models by combining clinicopathological features has become an important approach to improving preoperative assessment accuracy. This study aims to develop an individualized prediction model for lateral cervical lymph node metastasis using routine clinical indicators and to validate its performance through various statistical methods and machine learning algorithms, thereby providing a tool for clinical decision-making.

1 The First Affiliated Hospital of Anhui Medical University, Hefei, Anhui, 230022, China

2 Hefei Cancer Hospital, Chinese Academy of Sciences, Hefei, Anhui, 230031, China

* Corresponding Author.

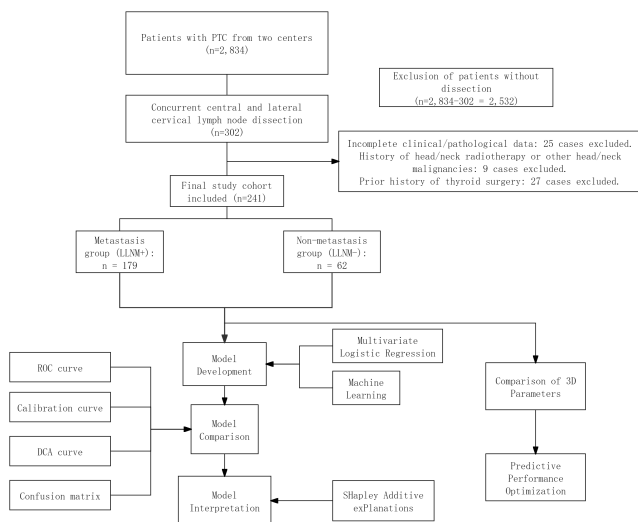
Methods

Clinical Study and Design

Study Design

This study included patients with PTC who were admitted to the First Affiliated Hospital of Anhui Medical University from March 2021 to September 2025 and to the Hefei Cancer Hospital of the Chinese Academy of Sciences from March 2019 to September 2025. All surgeries were performed by surgeons from the same treatment team, and postoperative pathological examination confirmed the diagnosis of PTC. Patients who underwent concurrent central and lateral neck lymph node dissection during surgery were selected. Inclusion criteria were: 1) preoperative conventional ultrasound examination; 2) no history of thyroid surgery. Exclusion criteria were: 1) history of head and neck radiotherapy; 2) concurrent other head and neck malignancies; 3) incomplete clinical, pathological, or ultrasound data (see Figure 1). The study protocol was approved by the Ethics Committees of The First Affiliated Hospital of Anhui Medical University and Hefei Cancer Hospital of Chinese Academy of Sciences.

Figure 1. Study flow chart.



Data Collection

Data were retrieved from inpatient medical records. Based on guidelines and previous studies [20-23], 14 variables were selected as candidate predictors: age, sex, maximum tumor diameter, central compartment lymph node metastasis (assessed based on preoperative imaging and intraoperative frozen section pathology), number of metastatic central lymph nodes, number of tumor foci, capsular invasion, calcification, aspect ratio, punctate-strip blood flow signal, nodule margin, nodule location, preoperative TSH, and preoperative Parathyroid Hormone (PTH). Additionally, the three-dimensional measurements (long, intermediate, and short axes) of all tumor foci were collected.

Statistical Analysis

SPSS software (version 26.0) was used for statistical analysis. Categorical data are presented as n (%), and comparisons

between groups were performed using the χ^2 test. + Variables with a P value < 0.1 in univariate analysis were included in a subsequent multivariate logistic regression analysis to identify independent predictors. The predictive performance of individual factors and combined indicators was evaluated using Receiver Operating Characteristic (ROC) curves. A P value < 0.05 was considered statistically significant.

Further modeling was conducted using R software (version 4.5.1). The analysis was performed in two stages. Stage 1: Development and internal validation of the logistic regression model using the full dataset. A prediction model and a corresponding nomogram were constructed based on the independent predictors. Model performance was evaluated in terms of: 1) Discrimination ability (Receiver Operating Characteristic [ROC] curve and Area Under the Curve, AUC); 2) Calibration (calibration curve and Hosmer-Lemeshow test); 3) Clinical utility (Decision Curve Analysis and Clinical Impact Curve). Internal validation was performed using the bootstrap method with 1000 resamples. Stage 2: Comparative analysis using a training-test set split (70%/30%). The data were split into training and validation sets in a 7:3 ratio. Applied machine learning algorithms included: Decision Tree, Random Forest, XGBoost, Support Vector Machine, K-Nearest Neighbors, LightGBM, and Naïve Bayes. These algorithms were selected to represent a diverse range of machine learning paradigms, thereby enabling a comprehensive comparison with the traditional logistic regression model. Specifically, Decision Tree provides an intuitive, rule-based approach for risk classification; Random Forest and XGBoost (extreme gradient boosting) are ensemble methods that handle non-linear relationships and feature interactions effectively, with XGBoost known for its robust performance in structured data; LightGBM (light gradient boosting machine) offers high efficiency and accuracy in gradient boosting frameworks; Support Vector Machine (SVM) is suitable for high-dimensional feature spaces through kernel transformation; K-Nearest Neighbors (KNN) serves as a non-parametric, instance-based learning method; and Naïve Bayes provides a probabilistic approach based on conditional independence assumptions. Collectively, this set of algorithms covers key categories—including tree-based models, ensemble methods, kernel-based methods, distance-based methods, and probabilistic classifiers—allowing for a rigorous evaluation of model performance across different analytical frameworks. Models were comprehensively assessed using the Area Under the ROC Curve, calibration curves, and decision curves. The SHAP (SHapley Additive exPlanations) method was used to interpret the model decision mechanisms.

Using an ellipsoid geometric model, we calculated three refined metrics beyond the "maximum tumor diameter" for each patient: total volume sum of all tumor foci ($V = (\pi/6)abc$), total surface area sum ($S \approx \pi(((ab)^p + (ac)^p + (bc)^p)/3)^{1/p}$, $p = 1.6075$), and sphericity of the largest tumor focus ($\Psi = 3a/(a+b+c)$) (see Figure 1).

*The ellipsoid model was selected because it is the most commonly used geometric approximation for thyroid nodules in clinical ultrasound practice. In this setting, tumors are typically measured along three orthogonal axes (long, intermediate, and short). In our study population, the majority of tumor foci exhibited a shape consistent with the ellipsoid assumption (i.e., well-defined, roughly ovoid contours), as confirmed by preoperative ultrasound imaging. While more sophisticated three-di-

mensional reconstruction techniques exist, the ellipsoid model offers a practical and reproducible method for routine clinical application.

Results

Development of a Prediction Model Based on Logistic Regression

Univariate and Multivariate Analysis

Among the 241 patients, significant differences ($P < 0.1$) were found between the metastasis and non-metastasis groups regarding the following variables: age, sex, maximum tumor diameter (>1 cm), central compartment lymph node metastasis, number of metastatic central lymph nodes, ill-defined nodule margin, aspect ratio (>1), nodule location, preoperative

TSH, and preoperative PTH. These variables were included in a multivariate logistic regression analysis. The results identified the following as independent risk factors for predicting lateral lymph node metastasis (LLNM): maximum tumor diameter, central compartment lymph node metastasis status, nodule location, and preoperative TSH level (see [Tables 1, 2, and 3](#)).

Development and Predictive Performance of Model

Using the full dataset (241 patients), ROC curves were plotted for each independent predictor to evaluate their individual predictive value (see [Figure 2](#)). The prediction model was converted into a nomogram for direct clinical application (see [Figure 3](#)). The model exhibited good discriminatory ability with an AUC of 0.832. Internal validation using the bootstrap method (1000 repetitions) showed a high consistency between predicted and actual probabilities on the calibration curve (mean

Table 1. Univariate analysis of clinical characteristics in patients with PTC.

Factor	Non-metastasis Group (n=62), n (%)	Metastasis Group (n=179), n (%)	χ^2 Value	P Value
Age			44.780	0.000
≤ 35 years	52 (83.9%)	62 (34.6%)		
> 35 years	10 (16.1%)	117 (65.4%)		
Sex			7.340	0.007
Male	9 (14.5%)	58 (32.4%)		
Female	53 (85.5%)	121 (67.6%)		
Maximum Tumor Diameter			13.686	0.000
≤ 1 cm	40 (64.5%)	67 (37.4%)		
> 1 cm	22 (35.5%)	112 (62.6%)		
Central LN Metastasis			58.865	0.000
Yes	20 (32.3%)	150 (83.8%)		
No	42 (67.7%)	29 (16.2%)		
No. of Central LNs			30.920	0.000
≤ 3	56 (90.3%)	90 (50.3%)		
> 3	6 (9.7%)	89 (49.7%)		
No. of Tumor Foci			2.442	0.295
Single	34 (54.8%)	81 (45.3%)		
Double	15 (24.2%)	43 (24.0%)		
Multiple	13 (21.0%)	55 (30.7%)		

Table 2. Multivariate logistic regression analysis of factors associated with lateral cervical lymph node metastasis.

Factor	β	SE	Wald χ^2	P Value	OR (95% CI)
Factor	-0.761	0.478	2.532	0.112	0.467 (0.183–1.191)
Sex (Male vs. Female)	-0.010	0.015	0.440	0.507	0.990 (0.961–1.020)
Age	0.689	0.259	7.062	0.008	1.992 (1.199–3.310)
Maximum Tumor Diameter	1.727	0.419	17.008	0.000	5.623 (2.471–12.794)
Central LN Metastasis	1.058	0.554	3.644	0.056	2.882 (0.973–8.534)
No. of Central LNs	0.085	0.495	0.030	0.864	1.089 (0.413–2.873)
Ill-defined Nodule Margin	-0.739	0.315	5.490	0.019	0.478 (0.258–0.885)
Nodule Location	0.055	0.420	0.017	0.896	1.056 (0.463–2.409)
Aspect Ratio (>1)	-1.056	0.448	5.562	0.018	0.348 (0.145–0.837)
Preoperative TSH	0.664	0.560	1.406	0.236	1.943 (0.648–5.823)
Preoperative PTH	1.762	2.123	0.689	0.406	5.827

Note: Statistically significant factors ($P < 0.05$) are highlighted in bold.

absolute deviation [MAD] = 0.023). The Hosmer-Lemeshow test yielded a P-value of 0.137, suggesting good model calibration. Decision curve analysis demonstrated that within the threshold probability range of 0.1 to 0.5, the net benefit of applying this model was higher than both the "intervention for all" and "intervention for none" strategies, indicating its potential clinical utility.

Construction and Comparison

To compare the performance of different modeling approaches, a training-test set split strategy (70%/30%) was employed. Based on the preliminary analysis, four predictive variables were included: central compartment lymph node metastasis status, maximum tumor diameter, nodule location, and preoperative TSH. All models were developed on the training set (70% of the data) and evaluated on the validation set (30% of the data). Multivariate logistic regression was used as the primary method to construct the prediction model. For supplementary comparison, seven machine learning algorithms were also applied, resulting in a total of eight prediction models. The confusion matrices for each model in the validation set are shown in Figure 4. The LightGBM model achieved the highest F1 score (0.56), indicating the best balance between sensitivity and specificity. However, the logistic regression model outperformed all other models in terms of discriminative ability (highest AUC), calibration (lowest Brier score [a measure of calibra-

tion accuracy], 0.124), and clinical utility (broadest range of net benefit on the decision curve analysis), suggesting it holds the greatest potential for broad clinical applicability (see Figure 5).

Figure 2. ROC curve, calibration curve, and decision curve analysis of the prediction model.

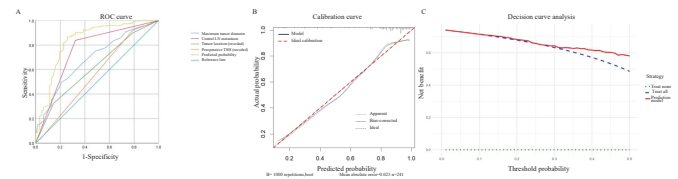


Figure 3. Nomogram model constructed based on multivariate variables.

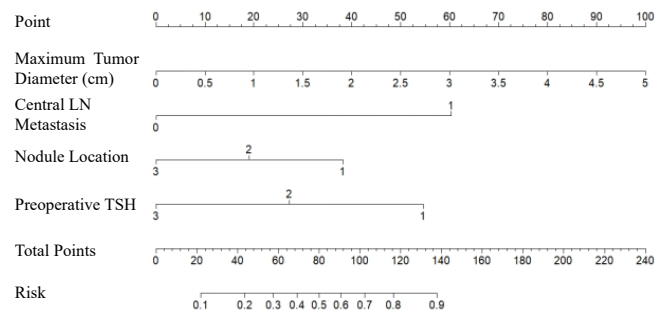


Table 3. Univariate analysis of ultrasound characteristics and serological indicators in patients with PTC.

Factor	Non-metastasis Group (n=62), n (%)	Metastasis Group (n=179), n (%)	χ ² Value	P Value
Capsular Invasion			1.773	0.183
Yes	12 (19.4%)	50 (27.9%)		
No	50 (80.6%)	129 (72.1%)		
Ultrasound Calcification			0.417	0.518
Yes	31 (50.0%)	98 (54.7%)		
No	31 (50.0%)	81 (45.3%)		
Aspect Ratio (>1)			3.436	0.064
Yes	24 (38.7%)	47 (26.3%)		
No	38 (61.3%)	132 (73.7%)		
Punctate-Strip Blood Flow			0.039	0.843
Yes	32 (51.6%)	95 (53.1%)		
No	30 (48.4%)	84 (46.9%)		
Ill-defined Nodule Margin			25.513	0.000
Yes	32 (51.6%)	148 (82.7%)		
No	30 (48.4%)	31 (17.3%)		
Nodule Location			9.9	0.007
Upper	9 (14.5%)	58 (32.4%)		
Middle	40 (64.5%)	103 (57.5%)		
Lower	13 (21.0%)	18 (10.1%)		
Preoperative TSH			8.116	0.017
Low	3 (4.8%)	8 (4.5%)		
Normal	46 (74.2%)	157 (87.7%)		
High	13 (21.0%)	14 (7.8%)		
Preoperative PTH			5.218	0.074
Low	1 (1.6%)	7 (3.9%)		
Normal	59 (95.2%)	150 (83.8%)		
High	2 (3.2%)	22 (12.3%)		

Figure 4. Ensemble of confusion matrices for each model.

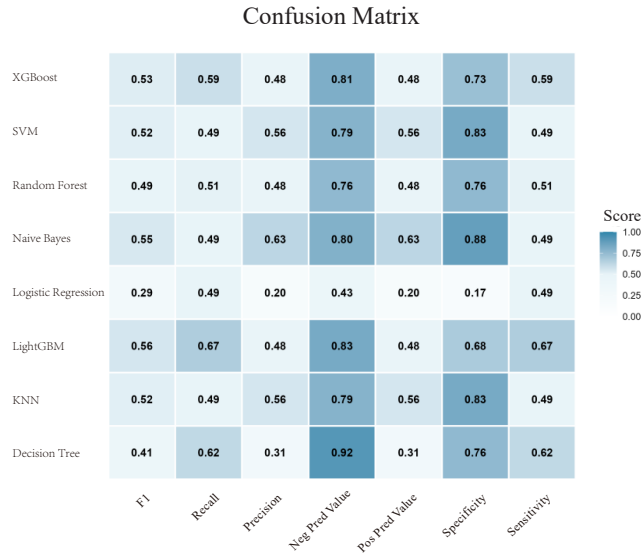
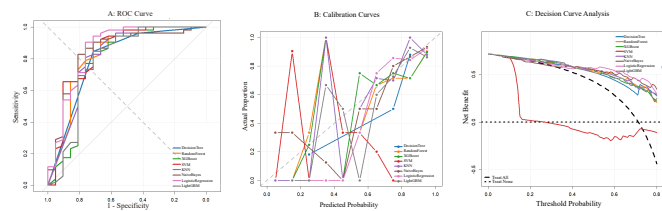


Figure 5. Comparison of discrimination, calibration, and clinical utility among the models.

A. The Logistic Regression model achieved the highest area under the curve (AUC) among all models, demonstrating the best overall discriminative ability for lateral cervical lymph node metastasis status. **B.** This model also had the lowest Brier score (0.124), and its calibration curve was closest to the ideal diagonal, indicating the best agreement between predicted probabilities and actual risk. **C.** In the decision curve analysis, the Logistic Regression model provided a net clinical benefit across the threshold probability range of 0.05–0.75, with a wider effective range than other models, suggesting broader clinical applicability.



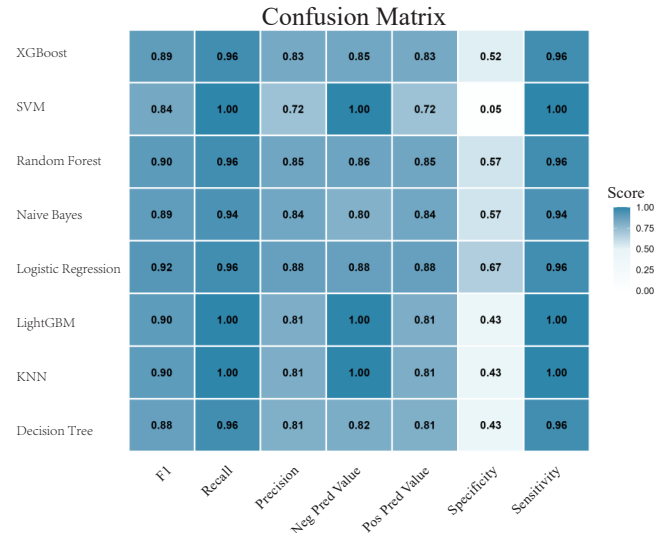
Threshold Optimization

Systematic optimization of classification thresholds revealed that the default threshold (0.5) was not optimal. Logistic Regression achieved its best F1 score (0.917) at a threshold of 0.6, making it suitable for scenarios emphasizing specificity. LightGBM reached its optimal F1 score (0.889) at a threshold of 0.3, rendering it more appropriate for screening contexts requiring high sensitivity (see Table 4). After threshold optimization, Logistic Regression, with the highest AUC and the optimal F1 score, emerged as the prediction model with the best overall performance (see Figure 6).

Table 4. Comparison of model performance after threshold optimization.

Model	Optimal Threshold	AUC	F1 Score	Accuracy	Sensitivity	Specificity
Logistic Regression	0.6	0.859	0.917	0.877	0.877	0.875
LightGBM	0.3	0.782	0.889	0.822	0.800	1.000

Figure 6. Ensemble of confusion matrices for each model under the optimal threshold range.



Comprehensive Evaluation and Discussion of Algorithmic Models

Based on the comprehensive performance evaluation, the Logistic Regression model demonstrated the best performance in terms of predictive accuracy and clinical utility. This model exhibited the highest discriminative ability (AUC=0.859), the most accurate predicted probabilities (Brier score=0.124), and its calibration curve aligned most closely with the ideal diagonal line. Within the clinical decision threshold range of 0.05 to 0.75, this model could provide the greatest net benefit (0.579). It ranked first across all three core evaluation dimensions: the ROC curve, calibration curve, and decision curve analysis. Its performance improved significantly after optimizing the classification threshold to 0.6. This model was thus identified as the optimal prediction model in this study.

Model Interpretability Analysis Based on SHAP

Model interpretability analysis was conducted using the SHAP (SHapley Additive exPlanations) method. In the Logistic Regression model, the order of variable importance was as follows: Central Compartment Metastasis > Nodule Location > Maximum Tumor Diameter > Preoperative TSH (see Figure 7). In the LightGBM model, the order of feature importance was: Central Compartment Metastasis > Maximum Tumor Diameter > Nodule Location > Preoperative TSH (see Figure 8). The SHAP summary plot illustrates the contribution of each feature's value to the model's predictive output, visually explaining the process of transforming feature inputs into predicted probabilities.

Figure 7. Feature importance analysis of the Logistic Regression model.

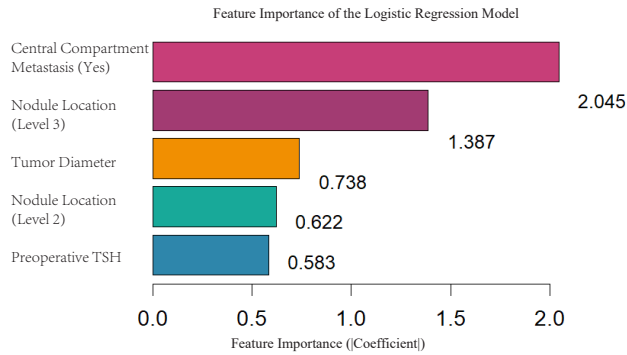
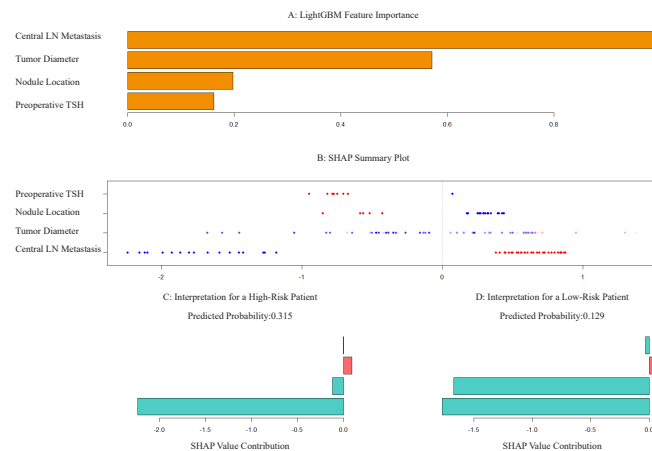


Figure 8: SHAP plot for the LightGBM model.

A. Feature Importance: The mean absolute SHAP value was used to evaluate each feature's contribution to the model's predictions. In the LightGBM model, the order of feature importance (highest to lowest) was: central compartment metastasis, tumor diameter, nodule location, preoperative TSH. **B. SHAP Summary Plot:** This plot visualizes the relationship between individual feature values and their corresponding SHAP values, providing a visual representation of each feature's contribution to the final predicted probability. **C. High-Risk Case Example:** For a patient with actual lateral lymph node metastasis, both the Logistic Regression and LightGBM models attributed increased risk prediction to features such as larger tumor diameter (and preoperative TSH in LightGBM). The final predicted probability was 0.609, correctly identifying this high-risk individual. **D. Low-Risk Case Example:** For a patient without metastasis, protective features like preoperative TSH (in Logistic Regression) and the absence of central metastasis/smaller tumor diameter (in both models) contributed negatively to the risk score. The final predicted probability was 0.194, correctly classifying this low-risk individual.



Exploration of Multivariate Prediction Models Using Alternatives to Maximum Tumor Diameter

Consistent with the comparative analysis described above, the training-test set split strategy (70%/30%) was used in this section as well. The variable "maximum tumor diameter" in the original model was sequentially replaced with the total tumor volume, total tumor surface area, and sphericity, following the same modeling process. The results showed that the AUCs for models based on total tumor volume, total tumor surface area, and sphericity were 0.835, 0.839, and 0.816, respectively. Notably, the AUCs for models using total tumor volume and

total surface area were 0.835 and 0.839, respectively, which were slightly higher compared to that of the original model (AUC=0.832) (see Table 5).

ROC curves and nomograms were plotted for each alternative indicator (see Figures 9 and 10). The optimal cut-off value of total tumor surface area for predicting LLNM was 4.6295 cm², yielding a sensitivity of 49.7% and a specificity of 82.3%, with an AUC of 0.689 (95% CI: 0.616-0.762). Patients with a total tumor surface area exceeding this threshold should be alerted to a higher risk of LLNM.

Figure 9. ROC curves of individual factors and the prediction model for predicting lateral cervical lymph node metastasis.

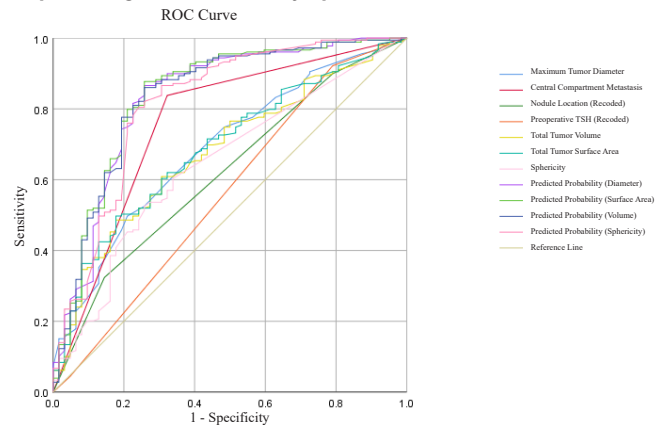
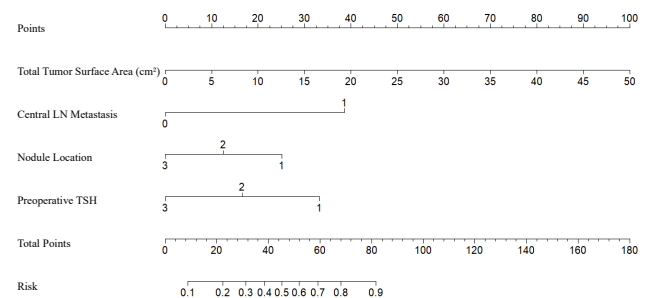


Figure 10: Nomogram model constructed based on significant new variables in the multivariate analysis.

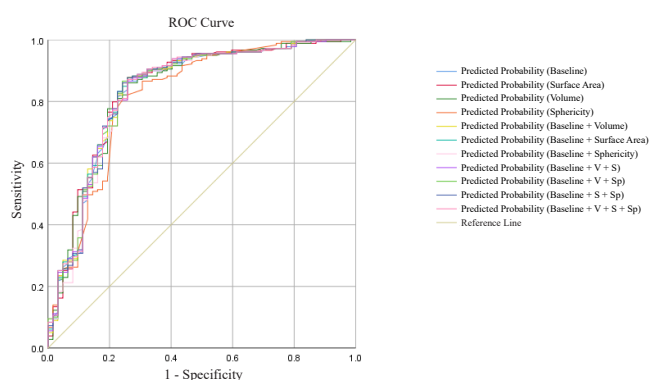


Note: The coding for categorical variables is as follows: Central Compartment Metastasis (0 = No, 1 = Yes); Nodule Location (1 = Upper, 2 = Middle, 3 = Lower); Preoperative TSH (1 = Low, 2 = Normal, 3 = High).

Maximum tumor diameter, total tumor volume, total tumor surface area, and sphericity provide complementary multidimensional information: maximum diameter serves as the classic one-dimensional indicator of linear growth; total tumor volume quantifies the overall tumor burden, which may be more directly related to metastatic potential; total tumor surface area characterizes the contact interface between the tumor and its microenvironment (blood vessels, lymphatic vessels) and is theoretically more closely associated with invasion and metastasis; sphericity reflects the morphological irregularity of the dominant tumor focus, potentially indicating more aggressive growth. Although these metrics are computationally related, they individually characterize the tumor's size, burden, interface, and shape features, holding independent biological significance. We regard them as a set of geometric candidate predictors to verify their incremental predictive value relative to maximum diameter indicators.

A baseline model was constructed based on maximum tumor diameter, central compartment metastasis, nodule location, and preoperative TSH, with each three-dimensional parameter subsequently added for comparison. The results revealed that the AUC of all combined models did not exceed that of the "Predicted probability (surface area sum)" model (AUC=0.839) (Figure 11, Table 6). This model incorporated only the core three-dimensional parameter. This model achieved the highest sensitivity (87.7%) and accuracy (84.0%) while maintaining stable specificity (74.2%), and its AUC was also superior to that of the baseline model (0.839 vs. 0.832). The performance of other models containing three-dimensional parameters was comparable but did not surpass it, suggesting that three-dimensional parameters—especially surface area sum—already possess promising independent predictive potential.

Figure 11. ROC Curves of the New Predictive Model for Lateral Lymph Node Metastasis.



To further validate the incremental predictive value of three-dimensional geometric parameters, this study adopted a training-test set split strategy (70%/30%) to redevelop and systematically compare 11 prediction models. It should be noted that while the specific numerical results obtained from this

validation approach differ from those derived from the Stage 1 full-dataset analysis, the core conclusions remain consistent, reflecting the generalization performance of the models in an independent test set.

In the test set, the model with the highest AUC was the "Basic + Volume" model (AUC = 0.8475, 95% CI: 0.7521–0.9428). Although its point estimate of AUC was marginally higher than that of the basic model (AUC = 0.8455), the difference was not statistically significant (Delong test, adjusted P = 1.000) (Table 7). This finding suggests that adding total tumor volume to the maximum diameter may provide modest complementary information in quantifying tumor burden, thereby achieving favorable discriminative performance within this dataset. From a clinical perspective, the lack of statistical significance indicates that the basic model—which relies on the more readily available maximum tumor diameter—achieves predictive performance comparable to the more complex model incorporating total tumor volume. However, the slightly higher AUC of the "Basic + Volume" model suggests that total tumor volume may capture additional information related to overall tumor burden in multifocal cases, which could be clinically valuable when preoperative assessment aims to maximize predictive accuracy. Thus, model selection can be tailored to clinical priorities: the basic model offers simplicity and ease of use, making it a more practical alternative for general application, while the "Basic + Volume" model may be considered when a marginally higher discriminative performance is desired, particularly in patients with multifocal disease. By incorporating only four routine variables, the basic model achieves nearly equivalent discriminative efficacy (AUC difference: 0.002) while substantially enhancing generalizability and operational feasibility.

Discussion

Surgery is the primary treatment for papillary thyroid carci-

Table 5. Predictive value of individual factors and combined model indicators for lateral cervical lymph node metastasis.

Factor / Model Indicator	Sensitivity (%)	Specificity (%)	Accuracy (%)	AUC	95% CI	P Value
Maximum Tumor Diameter	83.8	67.7	79.6	0.687	0.613–0.761	<0.001
Central LN Metastasis	49.7	79.0	63.0	0.758	0.682–0.833	<0.001
Nodule Location (Recoded)	32.4	85.4	46.0	0.617	0.538–0.697	0.006
Preoperative TSH (Recoded)	92.2	21.0	73.8	0.561	0.474–0.648	0.152
Total Tumor Volume	48.6	82.3	58.3	0.678	0.604–0.752	<0.001
Total Tumor Surface Area	49.7	82.3	58.9	0.689	0.616–0.762	<0.001
Sphericity	52.5	74.2	58.8	0.646	0.567–0.724	0.001
Model (Based on Diameter)	86.6	74.2	83.4	0.832	0.766–0.898	<0.001
Model (Based on Surface Area)	87.7	74.2	81.0	0.839	0.774–0.903	<0.001
Model (Based on Volume)	86.0	74.2	83.0	0.835	0.770–0.899	<0.001
Model (Based on Sphericity)	82.1	74.2	80.1	0.816	0.748–0.884	<0.001

Note: Key indicators are highlighted in bold for emphasis.

noma (PTC). Accurate preoperative assessment of lateral cervical lymph node metastasis (LLNM) status is crucial for determining the extent of lymph node dissection, which directly impacts surgical complications (such as recurrent laryngeal nerve injury [24], hypoparathyroidism [25], chylous leakage [26], etc.) and patients' quality of life.

This study collected data from PTC patients across two centers. Based on three indicators—maximum tumor diameter, central compartment lymph node metastasis status, and nodule location—an individualized prediction model for LLNM was developed. The model demonstrated good predictive performance (AUC=0.832) and calibration. Decision curve analysis confirmed its clinical utility.

To further enhance and compare performance, this study constructed multiple models, including multivariate logistic regression and seven machine learning algorithms. A comprehensive comparison revealed that after threshold optimization, the Logistic Regression model performed optimally at a threshold of 0.6, making it suitable for decision-making scenarios emphasizing specificity. The LightGBM model performed best at a threshold of 0.3, rendering it more appropriate for high-sensitivity screening. Considering all evaluation metrics, Logistic Regression was identified as the optimal prediction model.

Furthermore, this study innovatively introduced three-dimensional morphological parameters. For multifocal PTC, the total tumor volume and total tumor surface area were calculated.

Table 6. Predictive value of various prediction model indicators for lateral cervical lymph node metastasis.

Model Indicator	Sensitivity (%)	Specificity (%)	Accuracy (%)	AUC	95% CI	P Value
Predicted Probability (Baseline)	86.6	74.2	83.3	0.832	0.766–0.898	<0.001
Predicted Probability (Surface Area)	87.7	74.2	84.0	0.839	0.774–0.903	<0.001
Predicted Probability (Volume)	86.0	74.2	82.9	0.835	0.770–0.899	<0.001
Predicted Probability (Sphericity)	82.1	74.2	80.4	0.816	0.748–0.884	<0.001
Predicted Probability (Baseline + Volume)	87.7	74.2	84.0	0.833	0.768–0.899	<0.001
Predicted Probability (Baseline + Surface Area)	86.6	74.2	83.3	0.835	0.769–0.901	<0.001
Predicted Probability (Baseline + Sphericity)	86.0	75.8	83.5	0.829	0.762–0.896	<0.001
Predicted Probability (Baseline + V + S)	86.6	74.2	83.3	0.834	0.768–0.900	<0.001
Predicted Probability (Baseline + V + Sp)	86.6	75.8	83.9	0.833	0.767–0.899	<0.001
Predicted Probability (Baseline + S + Sp)	87.7	74.2	84.0	0.835	0.769–0.900	<0.001
Predicted Probability (Baseline + V + S + Sp)	87.2	74.2	84.0	0.835	0.769–0.900	<0.001

Notes: Baseline model predictors: Maximum tumor diameter, central compartment metastasis, nodule location, preoperative TSH.

Table 7. Delong Test Results for Key Model Comparisons.

Comparison Group	Model A AUC	Model B AUC	AUC Difference	Raw P-value	Adjusted P-value	Conclusion
Basic model vs. Surface area substitution	0.8455	0.8426	0.0029	0.807	1.000	No significant difference
Basic model vs. Volume substitution	0.8455	0.8397	0.0058	0.622	1.000	No significant difference
Basic model vs. Basic + Volume	0.8455	0.8475	0.0019	0.818	1.000	No significant difference
Basic model vs. Basic + Surface area	0.8455	0.8397	0.0058	0.714	1.000	No significant difference
Surface area substitution vs. Volume substitution	0.8426	0.8397	0.0029	0.478	1.000	No significant difference
Surface area substitution vs. Basic + All 3D parameters	0.8426	0.8124	0.0302	0.305	1.000	No significant difference
Basic model vs. Basic + All 3D parameters	0.8455	0.8124	0.0331	0.297	1.000	No significant difference

*Note: Adjusted P-values were calculated using the Bonferroni correction for multiple comparisons; the significance threshold was set at P < 0.05.

Models based on total tumor surface area (AUC=0.839) or total tumor volume (AUC=0.835) showed slightly higher predictive performance than the model based on maximum diameter (AUC=0.832). When the total tumor surface area exceeds 4.6295 cm², a high risk of LLNM should be strongly suspected. This finding validates the hypothesis that three-dimensional morphological parameters possess complementary discriminatory power, providing a new basis for clinical assessment. Maximum tumor diameter reflects only unidimensional linear growth and does not accurately quantify total tumor burden or the true spatial configuration of multifocal lesions. Total tumor volume directly represents the overall tumor cell load, and its association with metastatic potential may be stronger than that of maximum diameter alone in multifocal PTC. Total surface area more closely approximates the interface between the tumor and surrounding lymphatic and vascular pathways. A larger surface area may provide greater opportunity for tumor cells to breach the basement membrane and invade the vasculature, thus serving as a more direct indicator of invasive potential. Sphericity, as a morphological descriptor, offers a novel perspective for evaluating tumor growth patterns, with irregular morphology often suggesting enhanced proliferative activity and aggressive behavior.

Although the inclusion of three-dimensional parameters did not significantly improve the AUC of the basic model, both the "Surface area substitution" and "Volume substitution" models demonstrated predictive performance comparable to that of the basic model, with no statistically significant differences. These findings indicate that the predictive information embedded in three-dimensional parameters—particularly total surface area—largely encompasses that conveyed by unidimensional maximum diameter, conferring independent and equivalent discriminative capacity. Accordingly, total surface area derived from 3D imaging reconstruction holds promise as a preoperative predictive biomarker. It is at least equivalent to, and potentially provides more information than, maximum tumor diameter. It may facilitate risk assessment and guide individualized surgical decision-making, underscoring its clear potential for clinical translation.

In summary, this study advances the exploration from "unidimensional metrics" to the refinement of "three-dimensional geometric characterization." The results demonstrate that three-dimensional geometric parameters—especially total surface area—more comprehensively elucidate metastatic risk across dimensions such as tumor burden, interfacial effects, and growth morphology. These findings establish them as valuable complementary predictors of lateral cervical lymph node metastasis in PTC. Future research should validate their robustness in prospective, multicenter cohorts and investigate their integrated application with radiomic features and molecular markers.

Analysis of Predictive Factors

This study confirms that the maximum tumor diameter is an independent predictive factor for LLNM, consistent with previous research [27-30]. Larger tumor diameters are typically associated with greater invasiveness and a higher risk of metastasis. Central compartment lymph node metastasis was also confirmed as a robust predictive indicator [31-32]. As the primary lymphatic drainage site, its status effectively reflects the tumor's metastatic propensity. However, in some cases,

patients may exhibit skip metastasis: negative central lymph nodes with positive lateral lymph nodes [33-34], which can be easily missed during preoperative evaluation and surgery [35]. The influence of tumor location on metastasis risk may be related to anatomical differences in lymphatic drainage [27, 36-37]. The results of this study indicate that when the total tumor surface area exceeds 4.6295 cm², patients face a significantly increased risk of LLNM. This threshold can serve as an effective early-warning indicator.

Model Comparison and Strengths

Compared to previous studies, the strengths of our model include: 1) It integrates multiple independent predictive factors to provide individualized risk assessment; 2) Visualization through a nomogram facilitates direct clinical application; 3) Comprehensive validation of model performance using various statistical methods; and 4) Systematic comparison of multiple machine learning algorithms, revealing their respective application potentials.

Clinical Significance and Application

The prediction model and nomogram tool developed herein can assist clinicians in more accurately assessing the risk of lateral cervical lymph node metastasis preoperatively, thereby facilitating the formulation of individualized surgical plans. For patients identified as high-risk by the model, therapeutic or prophylactic lateral neck dissection may be considered. For low-risk patients, overtreatment can be potentially avoided.

Study Limitations and Future Directions

This study is a dual-center investigation with a limited sample size; therefore, its conclusions require further validation through multicenter, prospective studies with larger cohorts. Future research could integrate multi-dimensional information such as radiomics and circulating tumor markers [29, 38-40] to enhance predictive accuracy. Despite these limitations, this study successfully developed and validated a prediction model based on readily available clinical indicators. It demonstrates good predictive performance and clinical utility, providing a valuable reference for individualized preoperative lymph node assessment in PTC patients.

Conclusion

In this study, an effective clinical prediction model for assessing lateral lymph node metastasis risk was successfully developed using readily available clinical indicators, demonstrating good predictive performance and clinical utility. This model, based on four routine variables (maximum tumor diameter, central lymph node metastasis, nodule location, and preoperative TSH), is well-suited for widespread application across various clinical settings—from primary hospitals where advanced imaging resources may be limited to tertiary referral centers seeking rapid, standardized risk stratification. The incorporation of three-dimensional morphological parameters—particularly total tumor surface area and total tumor volume—shows promising potential for enhancing risk assessment in multifocal papillary thyroid carcinoma. However, these findings require further validation in prospective, multicenter cohorts before clinical implementation. Additionally, the relatively com-

plex measurement procedures and the need for automated tools currently limit the widespread applicability of three-dimensional parameters. Future efforts should focus on developing efficient, automated measurement tools to facilitate clinical translation.

Abbreviations

AUC: Area Under the Curve; CI: Confidence Interval; CT: Computed Tomography; FNA: Fine-Needle Aspiration; LightGBM: Light Gradient Boosting Machine; LLNM: Lateral Lymph Node Metastasis; MAD: Mean Absolute Deviation; PTC: Papillary Thyroid Carcinoma; PTH: Parathyroid Hormone; ROC: Receiver Operating Characteristic; SHAP: SHapley Additive exPlanations; TSH: Thyroid-Stimulating Hormone; SVM: Support Vector Machine; KNN: K-Nearest Neighbors; XGBoost: Extreme Gradient Boosting.

Acknowledgements

Not applicable.

Author Contributions

The manuscript was conceived by all authors. WS,ZYQ and WKL designed the article, WS drafted the manuscript which was subsequently edited by FZY,SCL,JSY,ZSG,WKL. All authors reviewed the manuscript.

Funding Information

Not applicable.

Ethical Approval and Consent to Participate

This study was conducted in accordance with the principles of the Declaration of Helsinki. With approval from the Ethics Committees of The First Affiliated Hospital of Anhui Medical University (Approval No. PJ 2024-12-91) and Hefei Cancer Hospital of Chinese Academy of Sciences (Approval No. PJ-KY2025-008), the need for informed consent was waived due to the non-interventional, retrospective study design.

Competing Interests

The authors declare no competing interests.

Data Availability

The datasets obtained and/or analyzed during the current study are available from the corresponding author on reasonable request.

References

- [1] Miranda-Filho A, Lortet-Tieulent J, Bray F, Cao B, Franceschi S, Vaccarella S, et al. Thyroid cancer incidence trends by histology in 25 countries: a population-based study. *Lancet Diabetes Endocrinol.* 2021;9:225–34. [https://doi.org/10.1016/S2213-8587\(21\)00027-9](https://doi.org/10.1016/S2213-8587(21)00027-9).
- [2] Bray F, Laversanne M, Sung H, Ferlay J, Siegel RL, Soerjomataram I, et al. Global cancer statistics 2022: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin.* 2024;74:229–63. <https://doi.org/10.3322/caac.21834>.
- [3] Haugen BR, Alexander EK, Bible KC, Doherty GM, Mandel SJ, Nikiforov YE, et al. 2015 American Thyroid Association Management Guidelines for Adult Patients with Thyroid Nodules and Differentiated Thyroid Cancer: The American Thyroid Association Guidelines Task Force on Thyroid Nodules and Differentiated Thyroid Cancer. *Thyroid Off J Am Thyroid Assoc.* 2016;26:1–133. <https://doi.org/10.1089/thy.2015.0020>.
- [4] Siegel RL, Miller KD, Wagle NS, Jemal A. Cancer statistics, 2023. *CA Cancer J Clin.* 2023;73:17–48. <https://doi.org/10.3322/caac.21763>.
- [5] Zaydfudim V, Feuer ID, Griffin MR, Phay JE. The impact of lymph node involvement on survival in patients with papillary and follicular thyroid carcinoma. *Surgery.* 2008;144:1070–7; discussion 1077–1078. <https://doi.org/10.1016/j.surg.2008.08.034>.
- [6] Basolo F, Macerola E, Poma AM, Torregrossa L. The 5th edition of WHO classification of tumors of endocrine organs: changes in the diagnosis of follicular-derived thyroid carcinoma. *Endocrine.* 2023;80:470–6. <https://doi.org/10.1007/s12020-023-03336-4>.
- [7] Shen G, Ma H, Huang R, Kuang A. Predicting large-volume lymph node metastasis in the clinically node-negative papillary thyroid microcarcinoma: a retrospective study. *Nucl Med Commun.* 2020;41:5–10. <https://doi.org/10.1097/MNM.0000000000001119>.
- [8] Smith VA, Sessions RB, Lentsch EJ. Cervical lymph node metastasis and papillary thyroid carcinoma: does the compartment involved affect survival? Experience from the SEER database. *J Surg Oncol.* 2012;106:357–62. <https://doi.org/10.1002/jso.23090>.
- [9] de Meer SGA, Dauwan M, de Keizer B, Valk GD, Borel Rinkes IHM, Vriens MR. Not the number but the location of lymph nodes matters for recurrence rate and disease-free survival in patients with differentiated thyroid cancer. *World J Surg.* 2012;36:1262–7. <https://doi.org/10.1007/s00268-012-1427-1>.
- [10] Ruan J, Chen Z, Chen S, Xu Z, Wen L, Mao Z, et al. Lateral lymph node metastasis in papillary thyroid microcarcinoma: a study of 5241 follow-up patients. *Endocrine.* 2024;83:414–21. <https://doi.org/10.1007/s12020-023-03486-5>.
- [11] Haddad RI, Bischoff L, Ball D, Bernet V, Blomain E, Busaidy NL, et al. Thyroid Carcinoma, Version 2.2022, NCCN Clinical Practice Guidelines in Oncology. *J Natl Compr Cancer Netw JNCCN.* 2022;20:925–51. <https://doi.org/10.6004/jnccn.2022.0040>.
- [12] Fei Y, Wang B, Yao X, Wu J. Factors associated with

- occult lateral lymph node metastases in patients with clinically lymph node negative papillary thyroid carcinoma: a systematic review and meta-analysis. *Front Endocrinol.* 2024;15:1353923. <https://doi.org/10.3389/fendo.2024.1353923>.
- [13] Zhang S, Liu R, Wang Y, Zhang Y, Li M, Wang Y, et al. Ultrasound-Base Radiomics for Discerning Lymph Node Metastasis in Thyroid Cancer: A Systematic Review and Meta-analysis. *Acad Radiol.* 2024;31:3118–30. <https://doi.org/10.1016/j.acra.2024.03.012>.
- [14] Mulla MG, Knoefel WT, Gilbert J, McGregor A, Schulte K-M. Lateral cervical lymph node metastases in papillary thyroid cancer: a systematic review of imaging-guided and prophylactic removal of the lateral compartment. *Clin Endocrinol (Oxf).* 2012;77:126–31. <https://doi.org/10.1111/j.1365-2265.2012.04336.x>.
- [15] Alabousi M, Alabousi A, Adham S, Pozdnyakov A, Ramadan S, Chaudhari H, et al. Diagnostic Test Accuracy of Ultrasonography vs Computed Tomography for Papillary Thyroid Cancer Cervical Lymph Node Metastasis: A Systematic Review and Meta-analysis. *JAMA Otolaryngol– Head Neck Surg.* 2022;148:107–18. <https://doi.org/10.1001/jamaoto.2021.3387>.
- [16] Zhang Y, Lu Y-Y, Li W, Zhao J-H, Zhang Y, He H-Y, et al. Lymphatic Contrast-enhanced US to Improve the Diagnosis of Cervical Lymph Node Metastasis from Thyroid Cancer. *Radiology.* 2023;307:e221265. <https://doi.org/10.1148/radiol.221265>.
- [17] Lin S, Zhong Y, Lin Y, Liu G. Prediction model for lateral lymph node metastasis of papillary thyroid carcinoma in children and adolescents based on ultrasound imaging and clinical features: a retrospective study. *BMC Med Imaging.* 2024;24:228. <https://doi.org/10.1186/s12880-024-01384-4>.
- [18] Yang Z, Wang X, Tao T, Zou J, Qiu Z, Wang L, et al. Diagnostic value of contrast-enhanced ultrasonography in the preoperative evaluation of lymph node metastasis in papillary thyroid carcinoma: a single-center retrospective study. *BMC Surg.* 2023;23:325. <https://doi.org/10.1186/s12893-023-02199-w>.
- [19] Tee YY, Lowe AJ, Brand CA, Judson RT. Fine-needle aspiration may miss a third of all malignancy in palpable thyroid nodules: a comprehensive literature review. *Ann Surg.* 2007;246:714–20. <https://doi.org/10.1097/SLA.0b013e-3180f61adc>.
- [20] Shao L, Wang Z, Dong W, Sun W, Zhang H. Risk factors associated with preferential lateral lymph node metastasis in papillary thyroid carcinoma. *Cancer Med.* 2023;12:20670–6. <https://doi.org/10.1002/cam4.6567>.
- [21] So YK, Kim M-J, Kim S, Son Y-I. Lateral lymph node metastasis in papillary thyroid carcinoma: A systematic review and meta-analysis for prevalence, risk factors, and location. *Int J Surg Lond Engl.* 2018;50:94–103. <https://doi.org/10.1016/j.ijssu.2017.12.029>.
- [22] Zhao L, Wu F, Zhou T, Lu K, Jiang K, Zhang Y, et al. Risk factors of skip lateral cervical lymph node metastasis in papillary thyroid carcinoma: a systematic review and meta-analysis. *Endocrine.* 2022;75:351–9. <https://doi.org/10.1007/s12020-021-02967-9>.
- [23] Zhao H, Huang T, Li H. Risk factors for skip metastasis and lateral lymph node metastasis of papillary thyroid cancer. *Surgery.* 2019;166:55–60. <https://doi.org/10.1016/j.surg.2019.01.025>.
- [24] Mizuno K, Takeuchi M, Kanazawa Y, Kitamura M, Ide K, Omori K, et al. Recurrent laryngeal nerve paralysis after thyroid cancer surgery and intraoperative nerve monitoring. *The Laryngoscope.* 2019;129:1954–60. <https://doi.org/10.1002/lary.27698>.
- [25] Yazıcıoğlu MÖ, Yılmaz A, Kocaöz S, Özçağlayan R, Parlak Ö. Risks and prediction of postoperative hypoparathyroidism due to thyroid surgery. *Sci Rep.* 2021;11:11876. <https://doi.org/10.1038/s41598-021-91277-1>.
- [26] Hassan I, Hassan L, Bacha F, Alsalameh M, Qatee O, Hassan W. Clinico-pathological initial outcome of a newly adopted novel surgical technique for nodal metastatic thyroid cancer at a large-volume centre in a high-income developing country. *Front Surg.* 2023;10:1204230. <https://doi.org/10.3389/fsurg.2023.1204230>.
- [27] Zhang X, Chen W, Fang Q, Fan J, Feng L, Guo L, et al. Lateral Lymph Node Metastases in T1a Papillary Thyroid Carcinoma: Stratification by Tumor Location and Size. *Front Endocrinol.* 2021;12:716082. <https://doi.org/10.3389/fendo.2021.716082>.
- [28] Bae SY, Jung SP, Choe J-H, Kim JS, Kim JH. Prediction of lateral neck lymph node metastasis according to preoperative calcitonin level and tumor size for medullary thyroid carcinoma. *Kaohsiung J Med Sci.* 2019;35:772–7. <https://doi.org/10.1002/kjm2.12122>.
- [29] Yu Q, Hao W, He Y, Ruan X, Liu L, Yun X, et al. Multi-omics analysis unveils dysregulation of the tumor immune microenvironment and development of a machine learning-based multi-gene classifier for predicting lateral lymph node metastasis in papillary thyroid carcinoma. *Endocrine.* 2025;90:172–87. <https://doi.org/10.1007/s12020-025-04308-6>.
- [30] Ywata de Carvalho A, Kohler HF, Gomes CC, Vartanian JG, Kowalski LP. Predictive factors for recurrence of papillary thyroid carcinoma: analysis of 4,085 patients. *Acta Otorhinolaryngol Ital Organo Uff Della Soc Ital Otorinolaringol E Chir Cerv-facc.* 2021;41:236–42. <https://doi.org/10.14639/0392-100X-N1412>.
- [31] Wada N, Duh Q-Y, Sugino K, Iwasaki H, Kameyama K, Mimura T, et al. Lymph Node Metastasis From 259 Papillary Thyroid Microcarcinomas. *Ann Surg.* 2003;237:399–407. <https://doi.org/10.1097/01.SLA.0000055273.58908.19>.
- [32] Zhan S, Luo D, Ge W, Zhang B, Wang T. Clinicopathological predictors of occult lateral neck lymph node metastasis in papillary thyroid cancer: A meta-analysis. *Head Neck.* 2019;41:2441–9. <https://doi.org/10.1002/hed.25762>.
- [33] Hu D, Lin H, Zeng X, Wang T, Deng J, Su X. Risk Factors for and Prediction Model of Skip Metastasis to Lateral Lymph Nodes in Papillary Thyroid Carcinoma. *World J Surg.* 2020;44:1498–505. <https://doi.org/10.1007/s00268-019-05332-0>.
- [34] Feng J-W, Qin A-C, Ye J, Pan H, Jiang Y, Qu Z. Predictive Factors for Lateral Lymph Node Metastasis and Skip Metastasis in Papillary Thyroid Carcinoma. *Endocr Pathol.* 2020;31:67–76. <https://doi.org/10.1007/s12022-019-09599-w>.
- [35] Wu X, Li B, Zheng C, He X. Risk factors for skip metastasis in patients with papillary thyroid microcarcinoma. *Cancer Med.* 2023;12:7560–6. <https://doi.org/10.1002/>

- cam4.5507.
- [36] Qubain SW, Nakano S, Baba M, Takao S, Aikou T. Distribution of lymph node micrometastasis in pN0 well-differentiated thyroid carcinoma. *Surgery*. 2002;131:249–56. <https://doi.org/10.1067/msy.2002.120657>.
- [37] Back K, Kim JS, Kim J-H, Choe J-H. Superior Located Papillary Thyroid Microcarcinoma is a Risk Factor for Lateral Lymph Node Metastasis. *Ann Surg Oncol*. 2019;26:3992–4001. <https://doi.org/10.1245/s10434-019-07587-2>.
- [38] Yu Y, Ouyang W, Huang Y, Huang H, Wang Z, Jia X, et al. Artificial intelligence-based multi-modal multi-tasks analysis reveals tumor molecular heterogeneity, predicts pre-operative lymph node metastasis and prognosis in papillary thyroid carcinoma: a retrospective study. *Int J Surg Lond Engl*. 2025;111:839–56. <https://doi.org/10.1097/JS9.0000000000001875>.
- [39] Yu S, Liu C, Hou Y, Li J, Guo Z, Chen X, et al. Integrative metabolomic characterization identifies plasma metabolomic signature in the diagnosis of papillary thyroid cancer. *Oncogene*. 2022;41:2422–30. <https://doi.org/10.1038/s41388-022-02254-5>.
- [40] Zhang L, Zhang J, Fan S, Zhong Y, Li J, Zhao Y, et al. A case-control study of urinary concentrations of bisphenol A, bisphenol F, and bisphenol S and the risk of papillary thyroid cancer. *Chemosphere*. 2023;312 Pt 1:137162. <https://doi.org/10.1016/j.chemosphere.2022.137162>.