



Developing and validating a multivariable machine learning model for the preoperative prediction of lateral lymph node metastasis of papillary thyroid cancer

Junwei Huang¹, Zufe Li^{1,2}, Qi Zhong¹, Jugao Fang¹, Xiaohong Chen¹, Yang Zhang¹, Zhigang Huang¹

¹Department of Otorhinolaryngology Head and Neck Surgery, Beijing Tongren Hospital, Capital Medical University, Key Laboratory of Otolaryngology Head and Neck Surgery (Capital Medical University), Ministry of Education, Beijing, China; ²Department of Otorhinolaryngology, Head and Neck Surgery, Beijing Chaoyang Hospital, Capital Medical University, Beijing, China

Contributions: (I) Conception and design: J Huang, Z Li, Y Zhang, Z Huang; (II) Administrative support: Y Zhang, Z Huang; (III) Provision of study materials or patients: J Huang, Y Zhang, Z Huang, Q Zhong, X Chen, J Fang; (IV) Collection and assembly of data: Z Li, J Huang; (V) Data analysis and interpretation: Z Li; (VI) Manuscript writing: All authors; (VII) Final approval of manuscript: All authors.

Correspondence to: Zhigang Huang, MD. Department of Otorhinolaryngology Head and Neck Surgery, Beijing Tongren Hospital, Capital Medical University, 1, Dongjiaominxiang Street, Dongcheng District, Beijing 100730, China. Email: enthzg@trhos.com. Yang Zhang, MD. Department of Otorhinolaryngology Head and Neck Surgery, Beijing Tongren Hospital, Capital Medical University, 1, Dongjiaominxiang Street, Dongcheng District, Beijing 100730, China. Email: zhangyangent@163.com.

Background: At present, preoperative diagnosis of lateral cervical lymph node metastasis (LLNM) in patients with papillary thyroid carcinoma (PTC) mostly depends on the training and expertise of ultrasound doctors. A machine-learning model for predicting LLNM accurately before PTC surgery may help to determine the scope of surgery and reduce unnecessary surgical trauma.

Methods: The data of patients with primary PTC who underwent thyroidectomy with lateral cervical lymph node surgery at Beijing Tongren Hospital between July 2009 and June 2021 were retrospectively analyzed. All patients had complete ultrasonic examination, clinical data, and definite pathology diagnosis of lymph nodes. LLNM was confirmed by postoperative pathology. The patients were randomly divided into a training set (155 cases) and a test set (98 cases) at a ratio of 6:4. Eleven parameters, including patient demographics, ultrasound results, and tumor-related conditions, were collected, and a prediction model was established using the support vector machine (SVM) algorithm. Several other machine-learning algorithms were also used to establish models for comparison. The accuracy, precision, recall, F1-score, sensitivity, specificity, Cohen's kappa value, and area under the receiver operating characteristic curve (AUC) were used to evaluate model performance.

Results: A total of 87 males and 156 females were included in the study, aged 14–80 years. One hundred and four patients of them had LLNM and 139 did not have LLNM. The pandas Python library was used for the statistical analysis, and the Spearman coefficient was used to analyze the correlation between each parameter and the prediction index. The SVM model performed the best among all the models. Its accuracy, precision, recall, F1-score, sensitivity, specificity, Cohen's kappa value, and AUC were 90.8%, 91.0%, 90.8%, 90.8%, 87.5%, 94.0%, 81.6%, and 91.0%, respectively.

Conclusions: This model can enable surgeons to improve the accuracy of ultrasonography in predicting LLNM without additional examination, thus avoiding missing positive lateral cervical lymph nodes and reducing the sequelae caused by unnecessary lateral neck dissection.

Keywords: Support vector machine (SVM); papillary thyroid carcinoma (PTC); neck lymph node dissection; lateral cervical lymph node metastasis (LLNM); machine learning

Submitted Nov 23, 2022. Accepted for publication Jan 06, 2023. Published online Jan 15, 2023

doi: 10.21037/gs-22-741

View this article at: <https://dx.doi.org/10.21037/gs-22-741>

Introduction

Papillary thyroid carcinoma (PTC) is the most common thyroid malignancy. Its annual incidence has been increasing year by year (1). In the United States, the incidence rate increased from 3.4/100,000 in 1975 to 15/100,000 in 2014 (2,3). PTC has the best prognosis among all types of thyroid cancer. The 10-year survival rate of PTC is >90% (1-3). However, the cervical lymph node metastasis rate of PTC is as high as 30–80%, especially in the central area. In addition, lateral cervical lymph node metastasis (LLNM) may also occur (4).

Research has shown that cervical lymph node metastasis is an important factor in the local recurrence of PTC after surgery (4,5). Preventive central lymph node (CLN) dissection has been suggested for PTC patients without clinical metastatic lymph nodes (4,5); however, controversy remains as to whether a neck dissection should be performed, especially a lateral neck dissection. An additional lateral neck dissection requires an expansion of the incision in the neck, which directly affects the patients' physical appearance, and this may have a great effect on the patient's quality of life. It also requires a modification to the originally planned endoscopic surgical procedure (4,5). Conversely, any occult metastatic lymph nodes present in the lateral cervical region are difficult to treat with postoperative iodine-131. In postoperative lymphadenopathy or neck recurrence, secondary surgery is often required (4,5). Thus, accurate preoperative knowledge of LLNM is helpful in formulating an appropriate surgical plan.

Currently, as a non-invasive, economical, and repeatable examination, ultrasound is the preferred procedure for

assessing thyroid tumors. However, the accuracy of an ultrasonic diagnosis greatly depends on the training and expertise of the doctor assessing the ultrasound images. An earlier study reported that LLNM assessments only had an accuracy of 82% (6), a sensitivity of 70%, and a specificity of 84% (7). Other studies have reported that the number of CLN metastases and tumor size are related to the extent of LLNM (4,8). Heng *et al.* reported that when the number of CLN metastases was ≥ 5 , the patient had a high risk (83.0%) of LLNM involvement (9). However, these studies relied on intraoperative pathological results obtained during CLN dissection. The interpretation of frozen sections is not as accurate as that of paraffin pathology, and not all hospitals are equipped to undertake intraoperative pathological examinations of frozen sections. The relatively accurate prediction of LLNM before surgery would be of great benefit in the selection of a surgical strategy.

Most guidelines do not recommend prophylactic lateral neck dissection for PTC. Surgeons usually rely on the accuracy of ultrasound reports and their own experience to decide whether to perform lateral neck dissection. In many cases, there is a lack of objective and effective evaluation indicators. Recent advances in computing power and big data handling has enabled some artificial intelligence applications began to outperform human intelligence in solving tasks that require complex decision-making (10). Common machine-learning algorithms include the support vector machine (SVM), random forest (RF), and decision tree (DT) algorithms. Currently, these algorithms are widely used in medical fields in the areas of disease diagnosis, curative effects prediction, and prognosis prediction (10). There are also many related studies in the field of otolaryngology head and neck surgery (11). Because different algorithms use different principles, it is impossible to know which algorithm will perform best before trying different algorithms. Researchers need to find the optimal algorithm according to different purposes and conditions. Lee *et al.* established an automatic prediction model for lymph node metastasis of PTC using an artificial intelligence deep-learning method for the interpretation of thyroid computed tomography (CT) images and found that the area under the receiver operating characteristic (ROC) curve (AUC) was 88.4% (12). CT is not the preferred examination for thyroid tumors, thus the use of this model is limited.

Ultrasound is a commonly used and accurate technique for thyroid examination. Yu *et al.* established a random forest predictive model of central lymph node metastasis

Highlight box

Key findings

- This study established and validated a machine-learning model that can accurately predict lymph node metastasis (LLNM) of primary papillary thyroid carcinoma (PTC) before surgery.

What is known and what is new?

- The judgment of lateral cervical LLNM is of the utmost importance for the surgery of PTC.
- This study proposed a machine-learning model to predict LLNM of PTC before surgery.

What is the implication, and what should change now?

- This model can benefit clinicians by assisting with preoperative planning in patients with primary PTC.

in cN0 papillary thyroid microcarcinoma patients with an AUC of 79.4% (13). This result showed the possibility and advantages of machine-learning model in predicting PTC lymph node metastasis. This model based on ultrasonic examination can also avoid radiation damage or iodine intake caused by additional CT scan. However, LLNM usually does not occur in PMTC patients, and this prediction only applied to lymph nodes in the central region. To the best of our knowledge, no machine-learning model has been developed to predict LLNM of PTC using only thyroid ultrasound scans and patient clinical data. This study aimed to use a machine-learning algorithm to predict LLNM of PTC before surgery to provide better preoperative reference information. We present the following article in accordance with the TRIPOD reporting checklist (available at <https://gs.amegroups.com/article/view/10.21037/gs-22-741/rc>).

Methods

Demographic data

This retrospective, cross-sectional study was approved by the Beijing Tongren Hospital Ethics Committee (approval number: TRECKY18-007). The study was conducted in accordance with the Declaration of Helsinki (as revised in 2013). Individual consent for this retrospective analysis was waived. The routinely collected data of 243 patients who underwent surgery for PTC at Beijing Tongren Hospital between July 2009 and June 2021 were retrospectively analyzed. To be eligible for inclusion in this study, the patients had to meet the following inclusion criteria: (I) be aged >14 years and have primary PTC; (II) have ultrasound scan and clinical data available; and (III) have undergone first-time thyroid and lateral cervical lymph node surgery and have a diagnosis confirmed by pathology. Patients were excluded from the study if they met any of the following exclusion criteria: (I) had incomplete data; (II) had pathology suggestive of non-thyroid papillary carcinoma; and/or (III) had other head and neck cancers. The patients were randomly divided into an independent training set (155 cases) and a test set (98 cases) at a ratio of 6:4.

All the primary PTCs were routinely excised via an ipsilateral neck dissection at Level VI. With bilateral lesions, bilateral neck lymph node dissection was performed at Level VI. If LLNM was suspected on preoperative imaging, a selective cervical lymph node dissection was performed during surgery. Generally, a selective neck dissection was performed rather than a lymph node biopsy or removal. If

the frozen sections returned a positive result, a Level II–IV or Level II–V dissection was performed depending on the number and location of positive lymph nodes in the frozen sections. A selective neck dissection was performed in some patients who had no lymph node metastases in the lateral neck according to the imaging findings at our hospital if: (I) the intraoperative frozen section showed that there were a large number of lymph node metastases at Level VI (unilateral >5); (II) the imaging findings from other hospitals suggested that the lymph nodes were swollen and the patients strongly requested investigation; or (III) the clinician, based on their experience, strongly suspected that the patient might have LLNM (e.g., neoplasms were observed in the upper pole involving the capsule).

Some PTC patients were diagnosed using fine-needle aspiration (FNA) before surgery. If there was no pathological diagnosis by FNA before surgery, thyroidectomy was first performed for the intraoperative frozen sections to confirm PTC. Lymph node dissection was performed after the diagnosis of a malignant tumor. According to the 2017 American Joint Committee on Cancer (AJCC) TNM system, N1b stage confirmed by postoperative pathology was diagnosed as LLNM.

Clinical data collection and feature selection

Eleven parameters, comprising 5 clinical parameters and 6 ultrasound scan characteristics, were used as input features for the model. The clinical data collected were as follows: (I) age at admission; (II) sex; (III) tumor location (upper, middle, or lower thyroid); (IV) tumor size; and (V) history of preoperative FNA biopsy of the thyroid nodule.

If an ultrasound scan showed that the lymph node size was >5 mm in the central neck or >8 mm in the lateral neck, it was classified as lymphadenopathy. If the lymph node size did not meet the above criteria, but there was microcalcification, cystic degeneration, a round shape, an irregular or fuzzy boundary, cortical thickening, or other abnormalities in the lymph node, it was classified as lymphadenopathy. Lymphadenopathy was recorded on ultrasonography.

Thus, the following ultrasound scan-related parameters were selected: (I) if an enlarged CLN/lymphadenopathy was identified on the ultrasound scan, regardless of whether metastasis or reactive hyperplasia was suspected, then this parameter was defined as “CLN reported” and set as “1”, and was otherwise set as “0”; and (II) if an enlarged lateral lymph node (LLN)/lymphadenopathy was identified on the

Table 1 Demographic data of the enrolled patients

Characteristics	Data
Age (years), mean \pm SD	42.5 \pm 12.3
Sex (male/female)	87/156
Location (upper/middle/lower)	54/116/73
FNA history (yes/no)	175/68
Capsule invasion (yes/no)	83/160
No. of CLNs in specimen	8.4 \pm 6.5
No. of positive CLNs, mean \pm SD	2.9 \pm 3.7
Tumor size, mean \pm SD	1.2 \pm 0.9
Dissection/pathology ratio, mean \pm SD	0.32 \pm 0.34
Enlarged CLN by ultrasound (yes/no)	73/170
Enlarged LLN by ultrasound (yes/no)	150/93
Unclear hilar-medullary structure (yes/no)*	81/162
Cortical thickening (yes/no)*	61/182
Punctate calcification (yes/no)*	105/138
Cystic change (yes/no)*	22/221

*, the state of LLN by ultrasound examination. FNA, fine-needle aspiration; CLN, central lymph node; LLN, lateral cervical lymph node.

ultrasound scan, regardless of whether metastasis or reactive hyperplasia was suspected, then this parameter was defined as “LLN reported” and set as “1”, and was otherwise set as “0”.

The description of “LLN reported” was further refined as follows: (I) unclear hilar-medullary structure of LLN; (II) punctate calcification of LLN; (III) cortical thickening of LLN; and (IV) cystic change of LLN. If there was no visible LLN, then these 4 ultrasound scan features were set as “0”. A Spearman’s correlation analysis was conducted to examine the correlation between these clinical parameters and the presence or absence of LLNM.

Machine-learning algorithm

In this study, a SVM algorithm was mainly used for model construction. The kernel function was “linear,” the penalty coefficient was set as 0.1, and other parameters were set as the default values. In addition, we evaluated and compared the performance of the models established using several commonly used machine-learning algorithms, including DT, RF, linear discriminant analysis, k-nearest

neighbor (KNN), polynomial naive Bayesian classifier, multilayer perceptron, XGBoost, CATBoost, AdaBoost, and LightGBM. All the machine-learning algorithms were implemented using the “sklearn” machine-learning library of Python programming software.

Model performance evaluation and statistical analysis

The accuracy, precision, recall, and F1-score of the sklearn machine-learning library were used to evaluate model performance with the test set. In addition, we used the AUC, specificity, sensitivity, and Cohen’s kappa coefficient to evaluate the models further. A confusion matrix was used to visually display the prediction results of the models. A prediction model with accuracy greater than 82%, sensitivity greater than 70% and specificity greater than 84% was considered to be a good model. The higher the values of all parameters, the better the prediction performance of the model. When kappa value was less than 60% or AUC was less than 50%, the model may not be considered as sufficient enough for clinical practice. The pandas Python library was used for the statistical analysis, and the Spearman coefficient was used to analyze the correlation between each parameter and the prediction index. A two-sided P value <0.05 was considered statistically significant.

Results

Description of demographic data of the enrolled patients

A total of 243 patients with PTC who underwent neck lymph node dissection were included in this study. The basic demographic data of the patients are shown in *Table 1*. Among the patients, 87 were male and 156 were female, their ages ranged from 14 to 80 years (42.5 \pm 12.3 years), and 104 had LLNM and 139 did not have LLNM as confirmed by pathology.

Feature importance

Feature importance was determined by a Spearman correlation analysis of all the features and LLNM states (*Figure 1*). The positive values in the *Figure 1* indicated a positive correlation coefficient, r . The numbers in the *Figure 1* are the correlation coefficients, and positive and negative values represent positive and negative correlations, respectively. The following characteristics were negatively

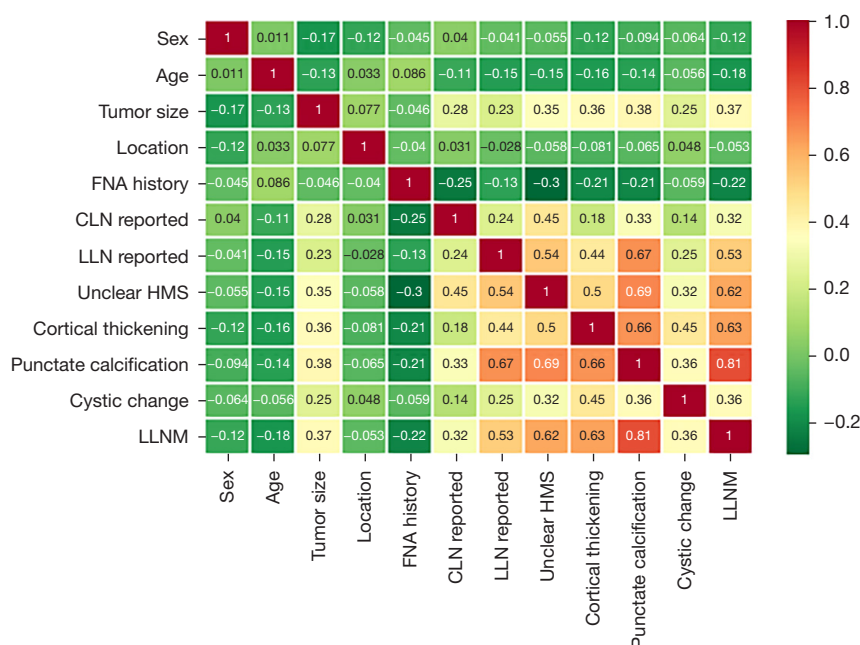


Figure 1 Correlation heatmap of the 11 features and LLNM characteristics. FNA, fine-needle aspiration; CLN, central lymph node; LLN, lateral cervical lymph node; HMS, hilar-medullary structure of the LLN reported by ultrasound scan; LLNM, lateral cervical lymph node metastasis.

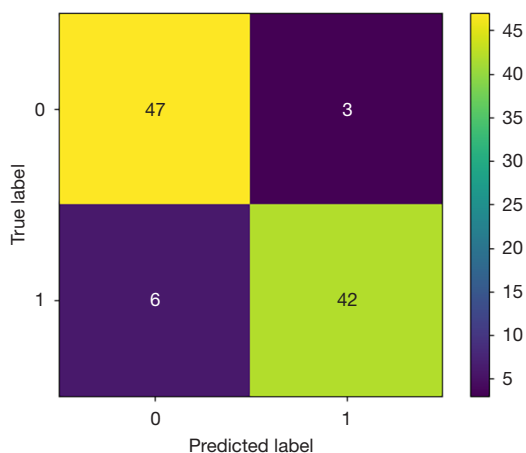


Figure 2 Confusion matrix of the proposed model with the SVM algorithm; “0” represents “no LLNM”, and “1” represents “LLNM”. SVM, support vector machine; LLNM, lateral cervical lymph node metastasis.

correlated with the presence or absence of cervical metastasis: age ($r=-0.18$), FNA history ($r=-0.22$) (both $P<0.05$). The following characteristics were positively correlated with the presence or absence of cervical

metastasis: tumor size ($r=0.37$), enlarged CLN reported by ultrasound scan ($r=0.32$), enlarged LLN reported by ultrasound scan ($r=0.53$), unclear hilar-medullary structure ($r=0.62$), cortical thickening ($r=0.63$), punctate calcification ($r=0.81$), and cystic change ($r=0.36$) (all $P<0.05$). Sex and tumor location were not significantly correlated with LLNM. However, as these 2 features represent key clinical data, they were also included in the training characteristics.

SVM model performance

The accuracy, precision, recall, F1-score, sensitivity, specificity, Cohen’s kappa value, and AUC of the LLNM prediction model established using the SVM algorithm were 90.8%, 91.0%, 90.8%, 90.8%, 87.5%, 94.0%, 81.6%, and 91.0%, respectively. The confusion matrix and AUC of the model are shown in Figures 2,3, respectively. As the confusion matrix shows, in the test set, 47 of 50 patients (94.0%) were correctly predicted to have no LLNM, and only 3 patients were incorrectly predicted to have no LLNM. Among the 48 cases of LLNM confirmed by pathology, 42 cases (87.5%) were correctly predicted, and 6 cases were incorrectly predicted.

Comparison of the different algorithm models

The results of the comparison of all the machine-learning algorithms (i.e., RF, LR, KNN, naive Bayes algorithm, DT, AdaBoost, XGBoost, CATBoost, LightGBM, and SVM) are shown in *Table 2*. Notably, the results showed that the SVM algorithm model was the most accurate model.

Discussion

In this study, 11 parameters were used as features to

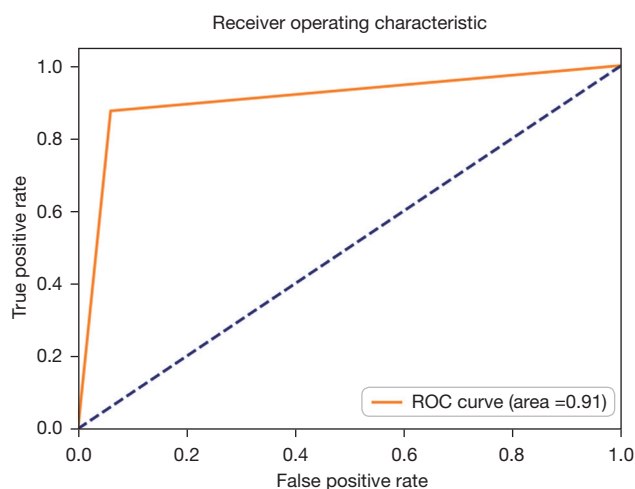


Figure 3 The AUC of the proposed model was 91.0%. ROC, receiver operating characteristic; AUC, area under the receiver operating characteristic curve.

construct a model that could preoperatively predict LLNM in patients with PTC. The parameters included were age, sex, tumor size, location, FNA history, CLN, and LLN reported by ultrasound scan, as well as the 4 characteristics of LLN (i.e., cortical thickening, punctate calcification, an unclear hilar-medullary structure, and cystic changes). Using artificial intelligence machine-learning algorithms, a model was constructed for preoperative LLNM prediction in patients with PTC. Its accuracy reached 90.8%, and precision, recall, F1-score, sensitivity, specificity, Cohen's kappa value, and AUC were 91.0%, 90.8%, 90.8%, 87.5%, 94.0%, 81.6%, and 91.0%, respectively, suggesting that the performance of the model was satisfactory.

Research has shown that age is negatively correlated with CLN metastasis in patients with PTC, such that young patients appear to be more prone to cervical metastasis (14). Conversely, other study has shown that cervical lymph node metastasis is not correlated with age (15). The relationship between sex and cervical metastasis is also controversial (14,15). Our study showed that LLNM was not correlated with sex and was only weakly negatively correlated with age ($r=-0.18$). Regardless of whether or not there is a correlation between age and sex and cervical metastasis, given that age and sex represent the most basic clinical data of patients, their inclusion as input features in the model should have helped to improve its generalization.

The rate of LLNM in patients with preoperative puncture may have been lower than that in patients without preoperative puncture ($r=-0.22$) for 2 possible reasons. First, if an ultrasound scan strongly suggests cervical

Table 2 Comparison of the results for different machine-learning algorithm models

Algorithm	Accuracy (%)	Precision (%)	recall (%)	Sensitivity (%)	Specificity (%)	F1-score (%)	Kappa value (%)	AUC (%)
RF	86.7	87.1	86.7	81.2	92.0	86.7	73.4	87.0
LR	85.7	86.7	85.7	77.1	94	85.6	71.3	86.0
KNN	76.5	81.1	76.5	56.2	96.0	75.5	52.7	76
Naive Bayes	87.7	88.3	87.8	81.2	94	87.7	75.4	88.0
DT	80.6	80.7	80.6	77.1	84.0	80.6	61.2	81.0
AdaBoost	82.6	83.0	82.7	77.1	88.0	82.6	65.2	93.0
XGBoost	84.6	85.0	84.7	79.2	90	84.6	69.3	95.0
CATBoost	86.7	86.8	86.7	83.3	90.0	86.7	73.4	87.0
LightGBM	72.0	79.6	72.0	64.7	87.5	72.9	44.8	76.0
SVM	90.8	91.0	90.8	87.5	94.0	90.8	81.6	91.0

AUC, area under the receiver operating characteristic curve; RF, random forest; LR, logistic regression; KNN, k-nearest neighbor; DT, decision tree; SVM, support vector machine.

lymph node metastasis, patients may choose direct surgical treatment without FNA. Second, we have observed that after puncture, some patients have inflammatory nodes in the lateral neck area, which may be misdiagnosed by ultrasound scan, which in turn would reduce the positive LLNM rate of lateral neck dissection. This reduction may be due to the inflammatory response; thus, further studies should be conducted to investigate the differences between this inflammatory reaction and metastasis.

It has been reported that tumor size in PTC is positively correlated with cervical lymph node metastasis (16). Our study showed that the correlation coefficient between tumor size and LLNM was 0.37, and the correlation coefficient with CLN metastasis was 0.43 (data not shown). Previous study has set tumor size cut-off values of 1 cm (16). Our research directly input the value of tumor size into the model as a feature without using cut-off values, which is one of the advantages of using a machine-learning algorithm.

Tumor location in the thyroid is often divided into the upper third, middle third, and lower third of the thyroid. One study showed that the location of the tumor is related to cervical metastasis, and the upper third of the PTC is more prone to metastasis (17). Other study reported that tumor location is not correlated with LLNM (16). Our findings indicated that tumor location was not correlated with the presence of LLNM. However, given the results in the literature, we used this parameter as one of the input features in the machine-learning model.

Some studies have suggested that the number of lymph node metastases and tumor size in the central region should be used as indicators of whether to perform cervical dissection (18,19). However, this study aimed to construct a preoperative prediction model. As a result, our prediction model did not consider the number of lymph node dissections in the central region, the number of lymph node-confirmed metastases in the central region, or the ratio of these 2 parameters.

The following ultrasound scan-related parameters were positively correlated with LLNM characteristics: (I) lymphadenopathy in the central region; (II) lymphadenopathy in the lateral neck; (III) unclear hilar-medullary structure of LLN; (IV) punctate calcification of LLN; (V) cortical thickening of LLN; and (VI) cystic change of LLN. With a correlation coefficient of 0.81, the presence of punctate calcification had the strongest correlation with LLNM. Thus, these 6 parameters were all taken as input features for the prediction model in this study.

Our study had several strengths. To the best of our knowledge, this is the first study to construct an artificial intelligence model to predict LLNM in patients with PTC before surgery. Second, the results of this study were satisfactory with an accuracy of 90.8%, sensitivity of 87.5%, and specificity of 94.0%, which were better than those reported for ultrasonography (which had an accuracy of 82%, a sensitivity of 70%, and a specificity of 84%) (6,7). Third, the model can predict LLNM characteristics before surgery, which reduces the need to wait for the intraoperative pathology results. This model can be used to create software packages to predict LLN metastasis in PTC. The user can input the above characteristics into the model to determine whether the patient has lateral cervical metastasis, and a lateral neck dissection should be performed.

This study had some limitations. First, the number of cases was small, and all were obtained from a single center. Second, there was no prospective evaluation of the current model. We tested the model using an independent test data set; however, the generalizability of the current model should be further investigated. In the future, we will apply this model to more centers to further clarify its applicability.

Conclusions

In summary, based on data from thyroid B ultrasound and clinical data from patients, we established a prediction model for LLNM in patients with PTC. This model could benefit clinicians by assisting with preoperative planning for patients with primary PTC.

Acknowledgments

The authors would like to thank the numerous individuals who participated in this study.

Funding: This study was supported by the National Natural Science Foundation of China (No. 82072997).

Footnote

Reporting Checklist: The authors have completed the TRIPOD reporting checklist. Available at <https://gs.amegroups.com/article/view/10.21037/gS-22-741/rc>

Data Sharing Statement: Available at <https://gs.amegroups.com/article/view/10.21037/gS-22-741/dss>

Conflicts of Interest: All authors have completed the ICMJE uniform disclosure form (available at <https://gs.amegroups.com/article/view/10.21037/ggs-22-741/coif>). The authors have no conflicts of interest to declare.

Ethical Statement: The authors are accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved. The study was conducted in accordance with the Declaration of Helsinki (as revised in 2013). The study was approved by the Beijing Tongren Hospital Ethics Committee (approval number: TRECKY18-007). Individual consent for this retrospective analysis was waived.

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

References

- Miccoli P, Bakkar S. Surgical management of papillary thyroid carcinoma: an overview. *Updates Surg* 2017;69:145-50.
- Davies L, Welch HG. Current thyroid cancer trends in the United States. *JAMA Otolaryngol Head Neck Surg* 2014;140:317-22.
- Roman BR, Morris LG, Davies L. The thyroid cancer epidemic, 2017 perspective. *Curr Opin Endocrinol Diabetes Obes* 2017;24:332-6.
- Feng JW, Yang XH, Wu BQ, et al. Predictive factors for central lymph node and lateral cervical lymph node metastases in papillary thyroid carcinoma. *Clin Transl Oncol* 2019;21:1482-91.
- Pisanu A, Reccia I, Nardello O, et al. Risk factors for nodal metastasis and recurrence among patients with papillary thyroid microcarcinoma: differences in clinical relevance between nonincidental and incidental tumors. *World J Surg* 2009;33:460-8.
- Sugitani I, Fujimoto Y, Yamada K, et al. Prospective outcomes of selective lymph node dissection for papillary thyroid carcinoma based on preoperative ultrasonography. *World J Surg* 2008;32:2494-502.
- Zhao H, Li H. Meta-analysis of ultrasound for cervical lymph nodes in papillary thyroid cancer: Diagnosis of central and lateral compartment nodal metastases. *Eur J Radiol* 2019;112:14-21.
- Jeon MJ, Chung MS, Kwon H, et al. Features of papillary thyroid microcarcinoma associated with lateral cervical lymph node metastasis. *Clin Endocrinol (Oxf)* 2017;86:845-51.
- Heng Y, Yang Z, Zhou L, et al. Risk stratification for lateral involvement in papillary thyroid carcinoma patients with central lymph node metastasis. *Endocrine* 2020;68:320-8.
- Spicer J, Sanborn AN. What does the mind learn? A comparison of human and machine learning representations. *Curr Opin Neurobiol* 2019;55:97-102.
- Tama BA, Kim DH, Kim G, et al. Recent Advances in the Application of Artificial Intelligence in Otorhinolaryngology-Head and Neck Surgery. *Clin Exp Otorhinolaryngol* 2020;13:326-39.
- Lee JH, Ha EJ, Kim D, et al. Application of deep learning to the diagnosis of cervical lymph node metastasis from thyroid cancer with CT: external validation and clinical utility for resident training. *Eur Radiol* 2020;30:3066-72.
- Yu Y, Yu Z, Li M, et al. Model development to predict central lymph node metastasis in cN0 papillary thyroid microcarcinoma by machine learning. *Ann Transl Med* 2022;10:892.
- Lin DZ, Qu N, Shi RL, et al. Risk prediction and clinical model building for lymph node metastasis in papillary thyroid microcarcinoma. *Oncotargets Ther* 2016;9:5307-16.
- Liu Z, Wang L, Yi P, et al. Risk factors for central lymph node metastasis of patients with papillary thyroid microcarcinoma: a meta-analysis. *Int J Clin Exp Pathol* 2014;7:932-7.
- Liu C, Xiao C, Chen J, et al. Risk factor analysis for predicting cervical lymph node metastasis in papillary thyroid carcinoma: a study of 966 patients. *BMC Cancer* 2019;19:622.
- Kwak JY, Kim EK, Kim MJ, et al. Papillary microcarcinoma of the thyroid: predicting factors of lateral neck node metastasis. *Ann Surg Oncol* 2009;16:1348-55.
- Hu D, Zhou J, He W, et al. Risk factors of lateral lymph node metastasis in cN0 papillary thyroid carcinoma. *World J Surg Oncol* 2018;16:30.

19. Wang Y, Deng C, Shu X, et al. Risk Factors and a Prediction Model of Lateral Lymph Node Metastasis in CN0 Papillary Thyroid Carcinoma Patients With 1-2 Central Lymph Node Metastases. *Front Endocrinol*

(Lausanne) 2021;12:716728.

(English Language Editor: L. Huleatt)

Cite this article as: Huang J, Li Z, Zhong Q, Fang J, Chen X, Zhang Y, Huang Z. Developing and validating a multivariable machine learning model for the preoperative prediction of lateral lymph node metastasis of papillary thyroid cancer. *Gland Surg* 2023;12(1):101-109. doi: 10.21037/gs-22-741