

A Proposed Heterogeneous Ensemble Algorithm Model for Predicting Central Lymph Node Metastasis in Papillary Thyroid Cancer

Wenfei Liu*, Shoufei Wang*, Xiaotian Xia, Minggao Guo 

Department of Thyroid, Parathyroid, Breast and Hernia Surgery, Shanghai Jiao Tong University Affiliated Sixth People's Hospital, Shanghai, People's Republic of China

*These authors contributed equally to this work

Correspondence: Xiaotian Xia; Minggao Guo, Department of Thyroid, Parathyroid, Breast and Hernia Surgery, Shanghai Jiao Tong University Affiliated Sixth People's Hospital, No. 600 Yishan Road, Shanghai, People's Republic of China, Tel +8618930172917; +8618930172912, Email 18930172917@163.com; guominggao203@163.com

Purpose: To develop a heterogeneous ensemble algorithm model to precisely predict central lymph node metastasis (CLNM), which can provide a reference value on controversial topics of performing prophylactic central lymph node dissection for patients with papillary thyroid cancer (PTC).

Methods: The study included patients with PTC who underwent an initial thyroid resection in a single-center medical institution between January 2014 and December 2018. A total of 18 variables, including clinical features and ultrasound (US) features, were used in the univariate analysis, multivariate analysis, and feature selection and were also used to develop a heterogeneous ensemble model based on five basic machine learning models, including extreme gradient boosting, k-nearest neighbors, random forest, gradient boosting, and AdaBoost. Moreover, a partial dependent plot was used to explain the heterogeneous ensemble model.

Results: The area under the receiver operating characteristic curve of the heterogeneous ensemble algorithm model was 0.67, which is significantly better than that of the basic machine models in predicting CLNM. All machine learning models performed better than US. Based on multivariate analysis and receiver operating characteristic curve analysis, age ≤ 33 years, tumor size ≥ 0.8 cm, US-suspected CLNM, and microcalcification were risk factors for CLNM, and anti-thyroid peroxidase antibody and serum thyroglobulin levels were favorable factors for CLNM.

Conclusion: The proposed heterogeneous ensemble algorithm model may be optimal tool to predict CLNM by integrating clinical and US features.

Keywords: heterogeneous ensemble algorithm model, central lymph node metastasis, papillary thyroid cancer, ultrasound, machine learning model

Introduction

Incidence of thyroid cancer is rapidly increasing worldwide. Papillary thyroid cancer (PTC) is the most common pathological type, accounting for 80–85% of thyroid cancers.¹ In the United States, the overall incidence of thyroid cancer is increasing by 3% each year, and the incidence and mortality of advanced PTC have increased.^{2,3} The incidence-based mortality is increasing, mainly due to the changing biology of thyroid cancer, which in turn could be some unknown attributable reason, such as environmental, etiologic factors.⁴ Despite the growing molecular analysis in thyroid cancer prognostics, the actual prognostic value of molecular markers still needs to be validated.^{5,6} PTC has over 98% survival rate in patients aged <45 years. Nevertheless, PTC is prone to lymph node metastasis, which worsens the prognosis of patients.⁷ Central lymph node metastasis (CLNM) is the most common type of lymph node metastasis.⁸ Previous studies have shown that patients with CLNM are more likely to have distant metastases and a poor prognosis.^{9,10} Therefore, it is beneficial and necessary to perform lymph node dissection in patients with PTC who are identified with CLNM. However, it is

controversial whether to perform routine central lymph node dissection in clinically negative CLNM. Several researchers believed that performance of routine central neck dissection could achieve a less recurrence, better outcomes, decreased postoperative thyroglobulin levels and a lower morbidity rate related to initial operation.^{11–14} Nevertheless, others thought that a more aggressive procedure could increase the risk of surgical complication such as permanent hypocalcemia, hypoparathyroidism and recurrent laryngeal nerve paralysis.^{15,16} Cervical ultrasound (US) and computed tomography (CT) have poor performance in detecting CLNM in PTC patients.^{17–19}

Machine learning (ML) is a new and advanced processing method involving the study and development of computational models of learning processes and big data analysis. Additionally, this technique has been proven effective for analyzing medical problems.^{20,21} However, previous studies focused on utilizing the single-base ML model for medical research.^{22,23}

The ensemble learning algorithm, also known as multiple classifier systems, is a more effective ML model than the single-base model.²⁴ There are two types of ensemble algorithms. One ensemble algorithm consists of several identical basic learners, called homogeneous ensemble algorithms; the other algorithm consists of different basic learners, called heterogeneous ensemble algorithms. A previous study pointed out that the performance of the basic model is better than that of ordinary US in detecting CLNM in patients with PTC.²⁵ The current study proposes a heterogeneous ensemble algorithm based on clinical data from a single medical institution. To better present and interpret the heterogeneous ensemble algorithm for clinicians, we have used a partial dependence plot for visualization and to analyze how the significant variable affects CLNM in PTC patients.²⁶

The aim of the current study was to develop a new machine learning model to assist clinicians in decision-making. If a PTC patient with positive or high probability of CLNM according to the prediction of new machine learning model, the CLNM should be paid more attention and the patient should have more tendency to undergo central lymph node dissection.

Methods

Data Collection

This study was a retrospective clinical investigation approved by the Ethics Committee of Shanghai Jiao Tong University Affiliated Sixth People's Hospital. All patients provided informed consent to participate in the study. A total of consecutive 1046 patients who underwent initial thyroidectomy between January 2014 and December 2018 at the Shanghai Sixth People's Hospital were included. Patients with PTC were confirmed by postoperative pathology. We excluded PTC patients for the following reasons: Other subtypes of thyroid cancer; previous treatments including drug medication; or radiotherapy before initial surgery; relevant information was not detailed. The screening process of data collection is illustrated in Figure 1.

Surgical Method

Patients with unilateral PTC underwent unilateral thyroidectomy plus ipsilateral central lymph node dissection, and patients with bilateral PTC or isthmus tumors underwent total thyroidectomy with bilateral CLND.²⁷ The definition of the central neck (N1a) included both level VI (peritracheal, paratracheal, or Delphina) and level VII (upper mediastinal) lymph node compartments.²⁸

Clinical Features

The current study collected information on the patient's clinical characteristics, thyroid function tests, and US features. Clinical characteristics include basic information such as sex and age. We also collected data on the following items from thyroid function tests one week before surgery: free T3 (FT3), free T4, thyroid-stimulating hormone (TSH), thyroglobulin (TG), anti-thyroglobulin antibody, anti-thyroid peroxidase antibody (TPO-Ab), thyroid-stimulating hormone receptor antibody, and calcitonin. All patients underwent cervical US within one month before surgery at our hospital with GE Voluson E8 Color Doppler Ultrasound Diagnosis System. US characteristics of thyroid nodules included tumor size, unclear margin, hypo-echogenicity, the total number of lesions, microcalcification, irregular shape, US-suspected,

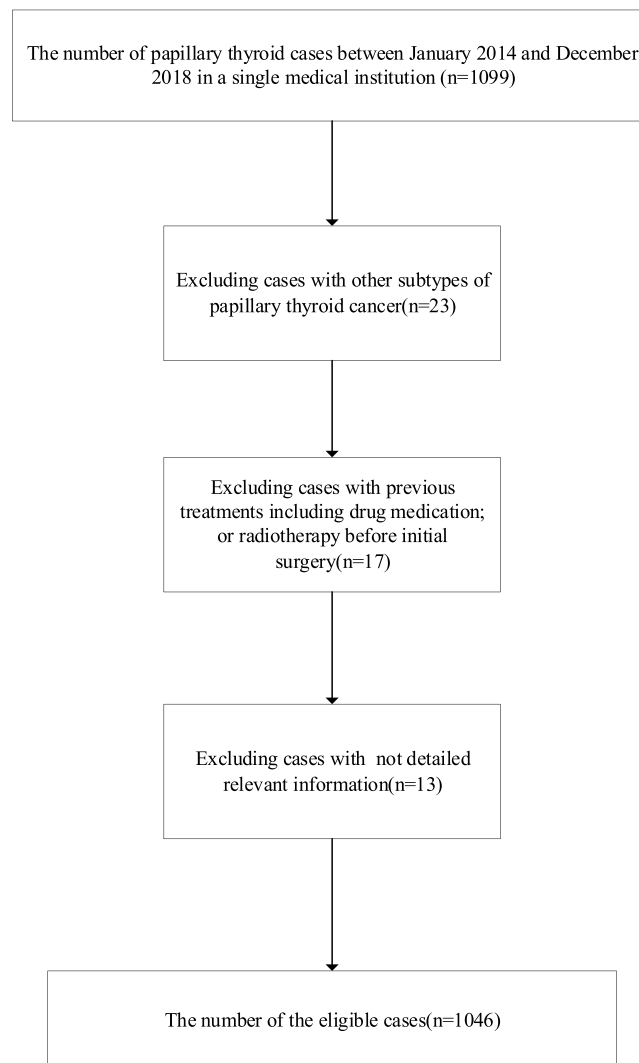


Figure 1 Detailed screening process of data collection.

and blood supply.²⁹ The US feature of hypo-echogenicity of a thyroid nodule is strongly related to an increased risk of malignancy.³⁰ Therefore, the mass was divided into hypo-echogenicity and none hypo-echogenicity. Tumor size was defined as the sum of the maximum diameter of a single-lesion mass and the total tumor diameter of a multi-lesion mass from US examination.³¹ An unclear boundary indicated that the boundary of the mass is difficult to determine using US detection. The total number of lesions was defined as the number of thyroid gland tumors. Blood supply was defined as obvious blood flow signals in the tumor using the US with color Doppler flow imaging. Microcalcification was defined as a spot-like strong echogenicity with an inner diameter <1 mm. Irregular shape was defined as a ratio of >1 in the anteroposterior diameter to the horizontal diameter when measured in the transverse plane. US-suspected was defined as the high possibility of tumor metastasis in level VI lymph nodes.

Feature Selection

Univariate and multivariate analyses were utilized for feature selection to avoid the overfitting state of the constructed machine learning model due to many variables. Variables with $P < 0.05$ in univariate analysis were included in the multivariate analysis. Variables with $P < 0.05$ in the multivariate analysis were important variables for ML model construction.

Development of Basic Machine Learning Model

Using the ratio of 8:2, the overall data set was divided into a training set (n=837) and a test set (n=209) using a random stratified sampling method.³² One-hot encoding is adopted to achieve data transformation of categorical features. For example, sex (male, female) with two values can be described as [(1,0), (0,1)]. Based on the training set, five basic ML models were developed: extreme gradient boosting, k-nearest neighbors (KNN), random forest, gradient boosting, and AdaBoost. The test set was used to assess the performances of the above models. Several previous studies confirmed that the above models perform well in category classification processing.^{23,33} Extreme gradient boosting is an effective algorithm, which has been applied for regression and classification. KNN is a typical classification supervised algorithm, which can solve complex classification problems. Random forest is a homogeneous ensemble algorithm of decision trees, which is used to deal with the classification issues. Gradient boosting is machine model that integrated several same weak classifiers. AdaBoost is the combination of some multiple classifiers. In our research, the above five machine models were described as the basic model elements that were used to create the new proposed machine learning model.

Development and Explanation of the Heterogeneous Ensemble Algorithm Model

In this study, the heterogenous ensemble algorithm model (HEM) was defined as the advanced machine learning model, which is composed of five basic learning algorithms. We adopted an ensemble learning to develop HEM. Ensemble learning is a method to determine the final prediction target by combining the prediction of several different basic ML models.²⁴ The final prediction of the HEM depended on the prediction of the five basic ML models. When predicting, we obtain the five basic algorithms' prediction results and use the prediction result with the most votes as the result. The algorithm formula was as following:

$$M(X) = \begin{cases} C_j & \text{if } \sum_{i=1}^T B_i^j(X) > 0.5 \\ \text{reject, otherwise} & \end{cases}$$

Where X is n dimensional feature vector inputs. C_j is the j class with more than half of the votes.

T means the number of basic ML models. B_j i(x) is the outputs of basic ML on the j class. M(x) is the prediction of the HEM model. For example, the prediction of three ML algorithms is a PTC patient with CLNM, and predictions of two ML algorithms is a PTC patient with NCLNM. The final prediction of HEM is a PTC patient with CLNM. The specific process of the heterogeneous integration algorithm is shown in Figure 2. Furthermore, the partial dependent plot, which is a tool for graphic visualization, is utilized to show how the response of the predicted CLNM is affected by dependent variables for explanation of HEM.²⁶ Moreover, optimal cut-off of tumor size and age by the ROC curve analysis.

The programming software version R version 4.0.5 (<https://www.r-project.org>) and Python version 3.8.1 (<https://www.python.org>) were used for the statistical analysis and development of models.

Results

Clinical Characteristics

A total of 1046 consecutive PTC patients consisted of 821 PTC patients with NCLNM (78.5%) and 225 PTC patients with CLNM (21.5%), including 207 patients with ipsilateral CLNM (19.8%) and 18 PTC patients with bilateral CLNM (1.7%). Our sample included 256 males (24.5%) and 790 females (75.5%) and the mean age was 43.9±12.92 years. The mean tumor size was 1.02±0.76 cm and the mean number of tumor lesions was 2.33. The mean number of removed and positive lymph nodes was 5.02 and 1.56, respectively. A total of 307 patients with PTC (29.3%) underwent total thyroidectomy plus bilateral central lymph node dissection, and 739 (70.7%) underwent unilateral lobe thyroidectomy plus unilateral central lymph node dissection. Additionally, there were 69 PTC patients (6.6%) with US-suspected CLNM. More detailed clinical information regarding PTC patients is shown in Table 1. Additionally, younger PTC patients were more likely to have CLNM compared to those with older age, as illustrated in Table 2. Moreover, FT3, TG, TPO-Ab, size, sex, irregular shape, microcalcification, US-suspected CLNM and surgical resection showed significant difference between groups.

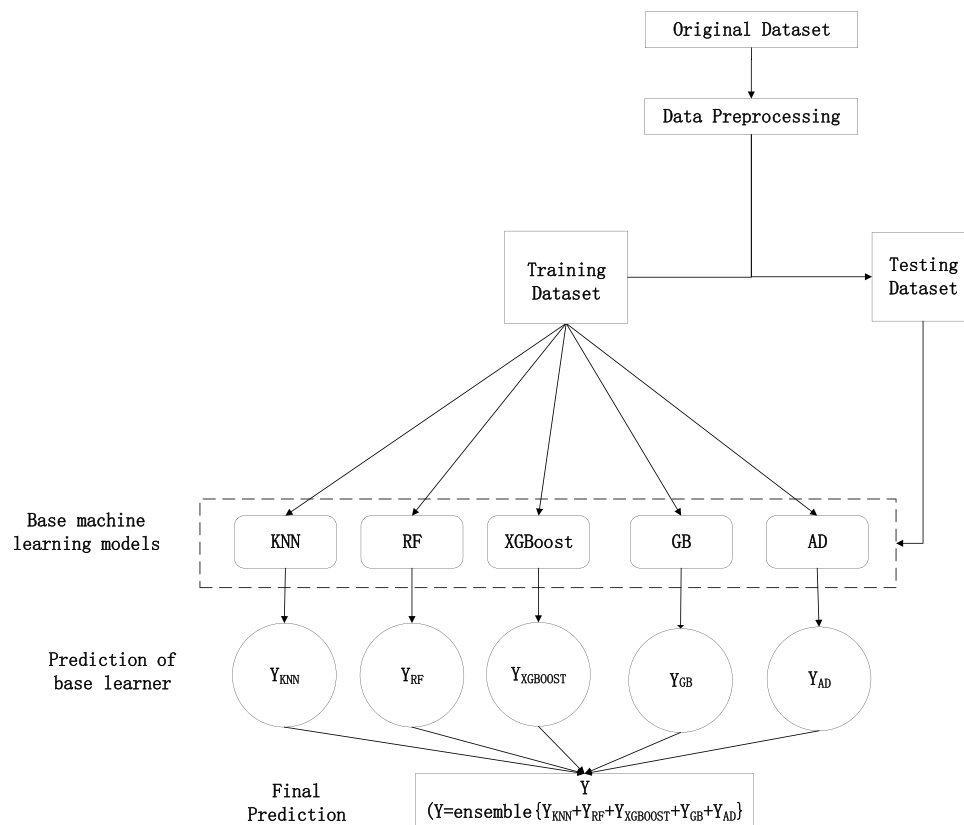


Figure 2 Detailed procedure of development of heterogeneous ensemble algorithm model.

Feature Selection

Univariate logistic analysis showed that CLNM was significantly associated with clinical characteristics such as sex, age, FT3, TSH, TG, TPO-Ab, and the US features such as microcalcifications, irregular shape, US-suspected CLNM, and tumor size ($p < 0.05$) (Table 3). Variables with a P value < 0.05 in the univariate logistic analysis were selected for the multivariate logistic analysis. The multivariate logistic analysis showed that age (OR=0.974, 95% CI=0.961–0.986, $P < 0.001$), TG (OR=0.997, 95% CI=0.994–0.999, $P=0.011$), TPO-Ab (OR=0.996, 95% CI=0.993–0.998, $P=0.001$), US-suspected CLNM (OR=2.202, 95% CI=1.251–3.387, $P=0.006$), size (OR=1.261, 95% CI=1.021–1.558, $P=0.032$), and microcalcification (OR=1.559, 95% CI=1.13–2.15, $P=0.007$) were significantly related with the CLNM in PTC patients and selected for development of ML model. More information is provided in Table 4.

Performance and Visualization of HEM Model

Six variables significantly associated with CLNM were selected to develop the five basic predictive ML models, which were used to create HEM. Compared with US (area under the receiver-operating characteristic curve [AUC]=0.525), basic predictive ML models, including extreme gradient boosting (AUC=0.558), KNN (AUC=0.629), random forest (AUC=0.663), gradient boosting (0.609), and AdaBoost (AUC=0.614) performed better in detecting CLNM in PTC patients, as shown in Figure 3 and Table 5. The HEM model performed better (AUC=0.670) in predicting CLNM compared to the basic ML model. Additionally, brier score of the HEM model was 0.169 as shown in calibration curve (Figure 4). To better understand the HEM model, the response of predicted CLNM affected by age and TPO-Ab is shown in the partial dependence plot. As illustrated in Figure 5A, the predicted probability of CLNM shows an overall decreasing tendency as the age increases. Thus, for PTC patients, the earlier age diagnosed, the prediction of HEM is more likely to be CLNM. Additionally, the predicted probability of CLNM decreases sharply when the value of age increases from 20 to 40 years. The phenomenon indicates the PTC patients with an earlier age of diagnosis are more

Table I Detailed Clinical Information of the PTC Patients

Features	Overall (n=1046)
Age, years*	43.90 (12.92)
FT3, pmol/L*	4.99 (0.64)
FT4, pmol/L*	17.08 (2.53)
TSH, mIU/L*	2.55 (1.94)
TG, ng/mL*	60.76 (32.31)
TG-Ab, KIU/L*	149.85 (43.48)
TPO-Ab, KIU/L*	50.69 (10.41)
TR-Ab, IU/L*	0.47 (0.08)
Calcitonin, ng/L*	4.18 (0.44)
No. of nodules*	2.33 (0.71)
No. of HCLN*	5.02 (4.92)
No. of PCLN*	1.56 (0.56)
Size, cm*	1.02 (0.76)
Sex, n (%)	
Male	256 (24.5)
Female	790 (75.5)
CLNM	
Unilateral	207 (19.8)
Bilateral	18 (1.7)
Negative	821 (78.5)
Hypo-echogenicity, n (%)	
YES	906 (86.6)
NO	140 (13.4)
Irregular shape, n (%)	
YES	465 (44.5)
NO	581 (55.5)
Unclear margin, n (%)	
YES	606 (57.9)
NO	440 (42.1)
Blood flow signals in nodules, n (%)	
YES	664 (63.5)
NO	382 (36.5)

(Continued)

Table 1 (Continued).

Features	Overall (n=1046)
Microcalcification, n (%)	
YES	567 (54.2)
NO	479 (45.8)
Suspected CLNM, n (%)	
Present	69 (6.6)
Absent	977 (93.4)
Surgical resection, n (%)	
Total thyroidectomy with bilateral CLND	307 (29.3)
Lobectomy plus with ipsilateral CLND	739 (70.7)

Note: *values are expressed as mean $\pm$ standard deviation.

Abbreviations: CLNM, central lymph node metastasis; CLND, central lymph node dissection; PTC, papillary thyroid cancer; HCLN, harvested central lymph nodes; PCLN, positive central lymph nodes.

likely to have CLNM, and clinicians should pay more attention to younger patients and plan appropriate surgery for them. As illustrated in [Figure 5B](#), the predicted probability of CLNM shows an overall decreasing tendency as the TPO-Ab increases. In other words, for PTC patients, the larger the value of the TPO-Ab, the prediction of HEM is more likely to be NCLNM. Several changes in the probability of CLNM should be noticed in the range of the TPO-Ab. Additionally, the predicted probability of CLNM increases sharply when the value of TPO-Ab increases from 25 KIU/L to 75 KIU/L. However, the predicted probability of CLNM decreases sharply when the value of age increases from 110 KIU/L to 240 KIU/L. The phenomenon indicates a non-linear relationship between the TPO-Ab and the CLNM. The cut-off age and size were 33 years and 0.8 cm, respectively, using ROC analysis.

Discussion

PTC has a better overall prognosis than other pathological types of thyroid cancer, with a 5-year survival rate of >98%.³⁴ However, PTC is prone to CLNM, which is associated with a higher risk of tumor recurrence, lateral cervical lymph node metastasis, and locally advanced thyroid cancer.³⁵ Clinically, it is necessary to perform central lymph node dissection in PTC patients with CLNM. However, it is controversial whether performing central lymph node dissection is recommended for PTC patients with unidentified CLNM due to surgical complications, such as permanent hypocalcemia and laryngeal nerve paralysis.^{36–38} Several researches about detecting cervical lymph node metastasis in PTC patients have been reported.^{39,40} High success results of artificial intelligence assisted ultrasound in detecting central lymph node metastasis have been discussed previously.²⁵ Therefore, to further improve the efficiency of diagnosing CLNM, we have proposed using HEM, consisting of a combination of basic ML models.

In this study, we found that the US sensitivity and specificity were only 10.6% and 40.7%, respectively, and the AUC was only 0.525, which was consistent with the idea of low efficacy of US in previous studies.^{17–19} Additionally, all basic ML models perform better than US, demonstrating that the ML model may be used to predict CLNM. Wu et al confirmed that ML algorithms might predict lymph node metastasis in the central area.²⁵ A possible reason is that the ML algorithm analyzes other important variables about CLNM, ignored by US detection. The innovation in this research is to develop a HEM to predict CLNM by integrating basic learning algorithms. We found that HEM performed better than basic models, demonstrating that HEM is an optimal and novel model for predicting CLNM. The HEM offers an effective technique for obtaining increased levels of predictive accuracy by combining the predictions of different ML models.⁴¹ The probable reason is that HEM uses different base algorithms for training, which can identify different types of errors

Table 2 Comparison of Information Between PTC Patients with NCLNM and PTC with CLNM

Factors	NCLNM (N=821)	CLNM (N=225)	P value
Age, years*	44.87±13.01	40.35±11.94	<0.001
FT3, pmol/L*	4.96 ±0.64	5.11±0.64	0.003
FT4, pmol/L*	17.05±2.58	17.20±2.36	0.441
TSH, mIU/L*	2.61±2.09	2.32±1.27	0.05
TG, ng/mL*	67.41±36.28	36.53±6.56	0.02
TG-Ab, KIU/L*	157.30±43.29	122.69±44.14	0.29
TPO-Ab, KIU/L*	56.35±10.85	30.05±6.62	0.001
TR-Ab, IU/L*	0.48±0.08	0.43±0.20	0.447
Calcitonin, ng/L*	4.43±0.53	3.30±1.66	0.751
No. of nodules*	2.32±0.72	2.35±0.69	0.595
No. of HCLN*	4.44±2.99	5.15±3.62	0.09
Size, cm*	0.99±0.76	1.11±0.77	0.04
Sex, n (%)			
Male	187 (22.8)	69 (30.7)	0.019
Female	634 (77.2)	156 (69.3)	
Hypo-echogenicity, n (%)			
YES	705 (85.9)	201 (89.3)	0.215
NO	116 (14.1)	24 (10.7)	
Irregular shape, n (%)			
YES	347 (42.3)	118 (52.4)	0.008
NO	474 (57.7)	107 (47.6)	
Unclear margin, n (%)			
YES	464 (56.5)	142 (63.1)	0.089
NO	357 (43.5)	83 (36.9)	
Blood flow signals in nodules, n (%)			
YES	532 (64.8)	132 (58.7)	0.106
NO	289 (35.2)	93 (41.3)	
Microcalcification, n (%)			
YES	424 (51.6)	143 (63.6)	0.002
NO	397 (48.4)	82 (36.4)	
US-suspected CLNM, n (%)			
Present	45 (5.5)	24 (10.7)	0.009
Absent	776 (94.5)	201 (89.3)	

(Continued)

Table 2 (Continued).

Factors	NCLNM (N=821)	CLNM (N=225)	P value
Surgical resection, n (%)			
Total thyroidectomy with bilateral CLND	218 (26.6)	89 (39.6)	<0.001
Lobectomy plus with ipsilateral CLND	603 (73.4)	136 (60.4)	

Note: *values are expressed as mean $\pm$ standard deviation.

Abbreviations: CLNM, central lymph node metastasis; CLND, central lymph node dissection; PTC, papillary thyroid cancer; HCLN, harvested central lymph nodes.

during training procedure and correct false prediction in time, which is beneficial for improving the accuracy of integration during testing. Herein, five different algorithms were used; the principles of each model differed increasing the number of errors that could be eliminated and the accuracy rate.

In addition to developing a novel ML model, we found that several risk factors were significantly associated with CLNM. In the multivariate analysis, we found that age was strongly associated with CLNM. Younger patients with PTC have a higher risk of CLNM. Receiver operating characteristic (ROC) analysis suggested the optimal cut-off age was 33 years (AUC=0.520; sensitivity = 0.639; specificity=0.533). Moreover, the predicted probability of CLNM with the change of age indicates that younger PTC patients were more prone to CLNM as shown in Figure 3A. Liu et al have previously demonstrated that age ≤ 45 years is an independent risk factor for CLNM.⁴² It is reasonable that younger patients with PTC should be paid more attention. However, more clinical research is needed to find the optimal cut-off age.

Additionally, TPO-Ab may be a protective factor against CLNM according to the multivariate logistic analysis. Aydoğan et al identified that TPO-Ab was negatively associated with CLNM in differentiated thyroid cancer patients with chronic lymphocytic thyroiditis (CLT), which is consistent with our research result.^{43,44}

However, Babli et al demonstrated that the coexistence of CLT and PTC appeared to be associated with unfavorable pathological features in patients aged <45 years.⁴⁵ Currently, it is still controversial whether PTC patients with CLT have favorable clinicopathological characteristics compared with PTCs without CLT. As shown in Figure 3B, we found an unclear relationship between CLNM and TPO-Ab levels. To clarify the relationship between outcomes of PTC and CLT, more research is required to determine the association between the serum TG and CLNM.

We also found a positive correlation between tumor size and CLNM in patients with PTC. Liu et al confirmed that tumor size >1.0 cm was an independent risk factor for CLNM.⁴² Moreover, we also identified that the optimal cut-off of tumor size was 0.8 cm (AUC=0.539; sensitivity=0.435; specificity=0.528). Zhou et al believe that 0.7 cm is a better cut-off point for tumor size for CLNM in his research.⁴⁶ Therefore, tumor size is closely connected to CLNM, and more clinical research is required to investigate the optimal cut-off of the tumor size.

Additionally, we found that the serum TG level was negatively associated with the risk for CLNM. Nevertheless, a previous study has demonstrated that patients with TG >5000 ng/mL have a significant disease burden and preoperative TG levels are associated with the metastatic disease.⁴⁵ Moreover, Scappaticcio et al thought that Sustained de novo TG-Ab appearance is rare and may predict structural recurrences and similar disease-free survival was noticed in patients with continued de novo TG-Ab and negative TG-Ab DTC patients.⁴⁷ Therefore, more investigation is required to understand the associations between TG and CLNM. Currently, the US has relatively low accuracy for the diagnosis of CLNM, but we found that US-suspected CLNM is a significant predictive factor for CLNM. Therefore, it is necessary to reconsider the surgical method for a patient with suspected US-CLNM. Moreover, we found that microcalcification was significantly associated with the CLNM. Microcalcification implies deposition of calcium salt in the thyroid nodules owing to the proliferation of blood vessels and fibrous, reflecting the rapid growth of cancer cells. Therefore, if microcalcification is present in the nodules, the lymph node status in the central region should be evaluated more cautiously.⁴⁸

Table 3 Univariate Analysis of Clinical Features Related to Central Lymph Node Metastasis in the PTC Patients

Factors	OR	95% CI	P value
Age	0.972	0.96–0.984	<0.001
FT3	1.398	1.119–1.746	0.003
FT4	1.023	0.966–1.083	0.441
TSH	0.896	0.805–0.997	0.043
TG	0.997	0.995–1	0.034
TG-Ab	1.000	0.999–1	0.297
TPO-Ab	0.996	0.993–0.998	0.001
TR-Ab	0.884	0.622–1.254	0.489
Calcitonin	0.999	0.996–1.003	0.754
No. of nodules	1.056	0.863–1.293	0.595
Size	1.208	1.007–1.448	0.041
Sex			
Male	1.500	1.081–2.08	0.015
Female	Reference		
Hypo-echogenicity			
YES	1.378	0.864–2.197	0.178
NO	Reference		
Irregular shape			
YES	1.506	1.12–2.025	0.007
NO	Reference		
Unclear margin			
YES	1.316	0.971–1.784	0.076
NO	Reference		
Blood flow signals in nodules			
YES	0.771	0.570–1.042	0.091
NO	Reference		
Microcalcification			
YES	1.633	1.205–2.213	0.002
NO	Reference		
US-suspected CLNM			
Present	2.059	1.225–3.460	0.006
Absent	Reference		

Abbreviations: CLNM, central lymph node metastasis; CLND, central lymph node dissection; PTC, papillary thyroid cancer; HCLN, harvested central lymph nodes.

Table 4 Multivariate Analysis of Clinical Features Related to Central Lymph Node Metastasis in the PTC Patients

Features	OR	95% CI	P value
Age	0.974	0.961–0.986	<0.001
FT3	1.235	0.958–1.591	0.103
TSH	0.929	0.83–1.039	0.198
TG	0.997	0.994–0.999	0.011
TPO-Ab	0.996	0.993–0.998	0.001
Size	1.261	1.021–1.558	0.032
Sex			
Male	1.232	0.855–1.775	0.263
Female	Reference		
Irregular shape			
YES	1.304	0.907–1.876	0.152
NO	Reference		
Microcalcification			
YES	1.559	1.13–2.15	0.007
NO	Reference		
US-suspected CLNM			
Present	2.202	1.251–3.878	0.006
Absent	Reference		
Unclear margin			
YES	0.965	0.665–1.4	0.851
NO	Reference		

Abbreviations: CLNM, central lymph node metastasis; CLND, central lymph node dissection; PTC, papillary thyroid cancer; HCLN, harvested central lymph nodes.

The innovation in this research is a combination of traditional and new statistical methods. First, univariate and multivariate analyses were used to select relatively important factors associated with CLNM. We then developed a HEM by combining basic ML models to predict CLNM in PTC patients with better performance. Second, to help clinicians better understand the novel model, partial dependency graphs were used to explain the model and visualize the change in tendency of risk of CLNM.

Limitations

The main purpose of this research was to develop a HEM model to predict the CLNM of PTC patients. However, there were some limitations to this research. Firstly, the number of cases studied was relatively small, and the model should be optimized based on the dataset with more PTC cases. Secondly, some important variables may have been ignored during data collection in the development of models after feature selection to avoid overfitting of models, such as TSH, thyroid imaging reporting and data system (TIRADS) classification.⁴⁹ Thirdly, potential selection bias, survivorship bias and study scope limitations cannot be avoided because of the retrospective study. Finally, prospective research is needed to

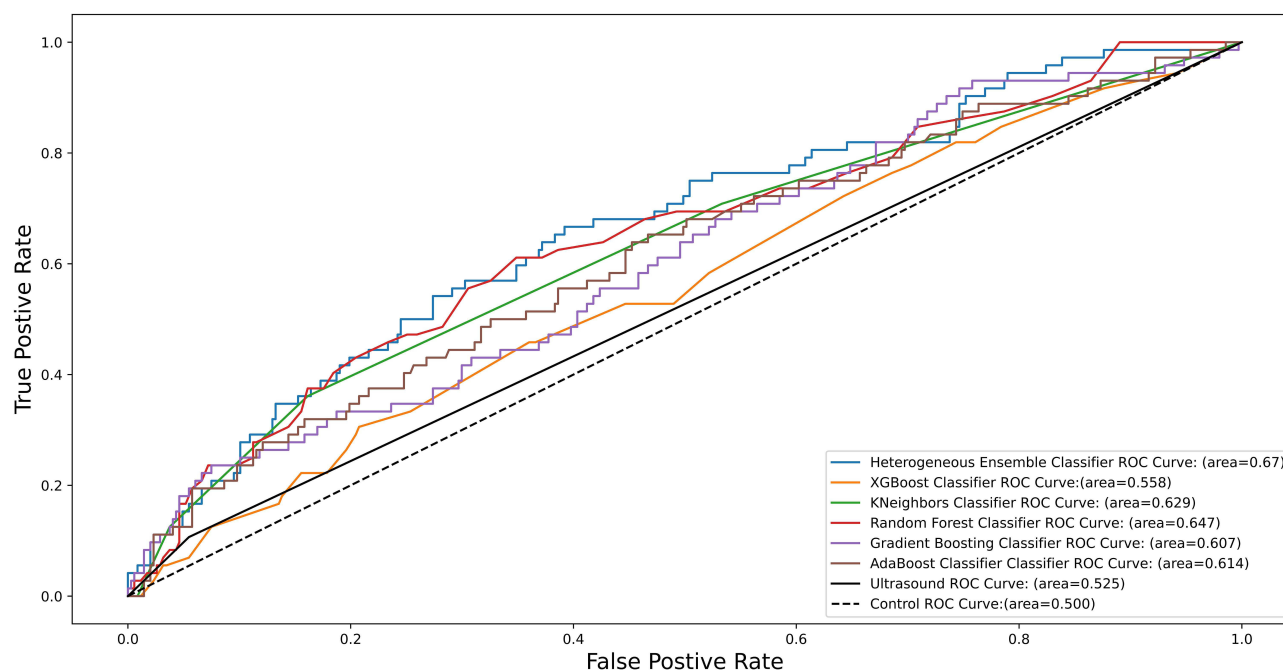


Figure 3 Receiver operating characteristic curves of all machine learning models.

verify the efficacy of the novel HEM model. As for other subtype thyroid cancer, the novel model should be designed based on the collection of data about patients with other subtypes of thyroid cancer to achieve implementation and validation of this algorithm.

Conclusion

Age ≤ 33 years, tumor size ≥ 0.8 cm, US-suspected CLNM, and microcalcification were risk factors for CLNM. Anti-thyroid peroxidase antibody and serum thyroglobulin levels were favorable factors for CLNM. The HEM has better performance than the US and other basic ML models, indicating that ML models could improve the function for

Table 5 The Performance of the Machine Learning Models and Ultrasonography

Model	AUC Value	Sensitivity	Specificity	Positive Value
HEM	0.67	44.30%	73.60%	61.30%
RF	0.647	38.50%	62.70%	57.40%
AD	0.614	34.50%	61.70%	46.60%
KNN	0.629	31.70%	62.70%	52.40%
XGBoost	0.558	15.60%	66.50%	45.30%
GB	0.607	22.20%	68.30%	43.20%
US	0.525	10.60%	40.70%	37.70%

Abbreviations: AUC, area under the ROC curve; HEM, Heterogeneous Ensemble Algorithm Model; RF, random forest classifier; KNN, k-nearest neighbors; XGBoost, extreme gradient boosting; GB, gradient boosting; US, ultrasonography; AD, Adaboost.

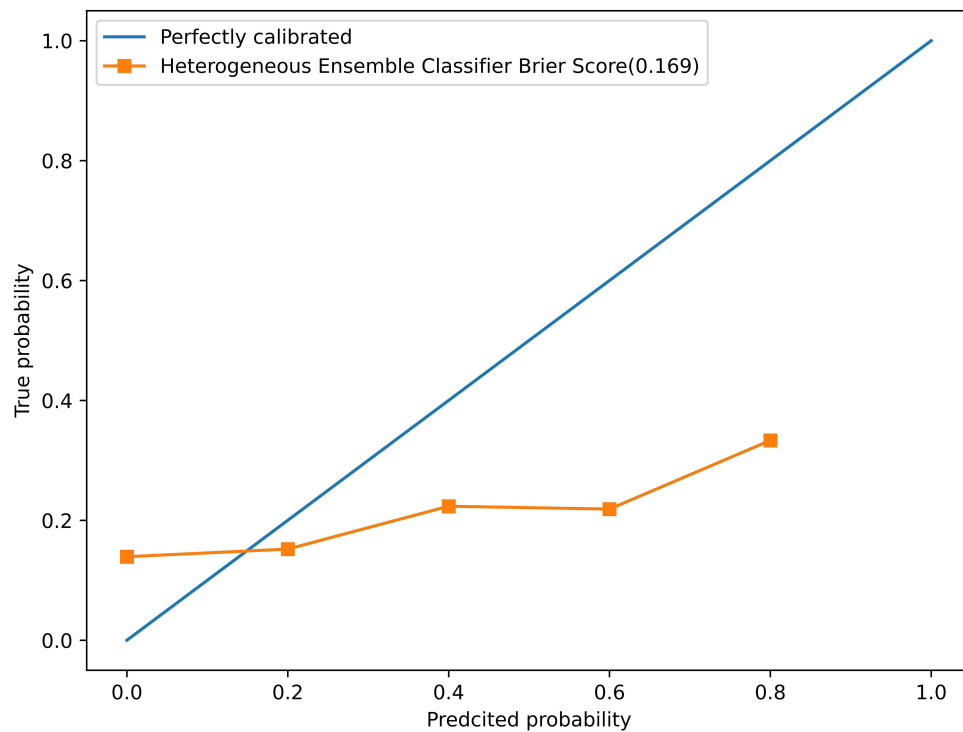


Figure 4 Calibration curve of the heterogeneous ensemble algorithm model.

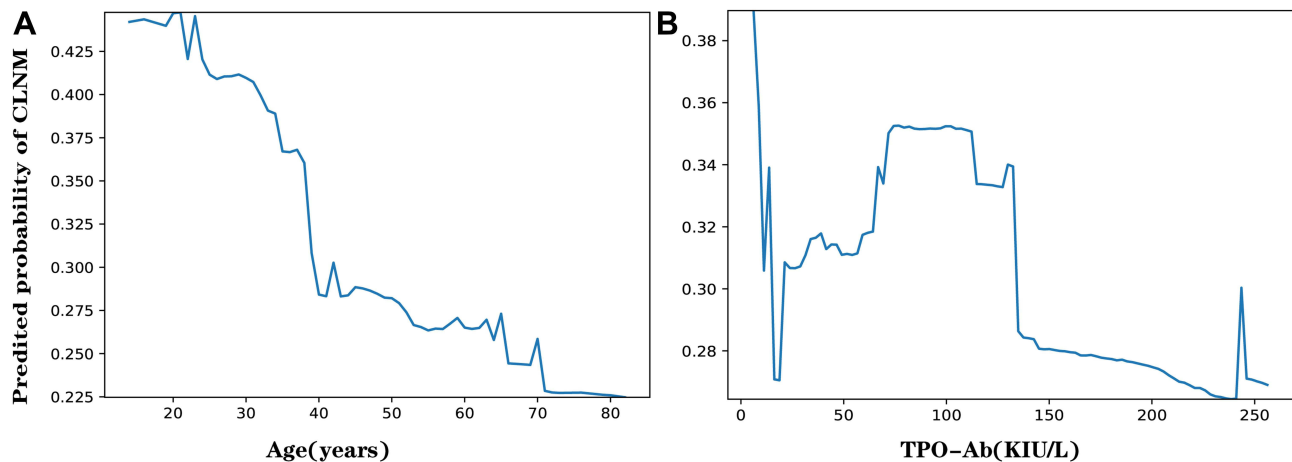


Figure 5 Partial dependent plot of predicted probability of CLNM against age (panel **(A)**) and partial dependent plot of predicted probability of CLNM against TPO-Ab (panel **(B)**).

predicting the CLNM. We believe that there will be more refined algorithms for clinical disease prediction based on this research.

Abbreviation

HEM, heterogeneous ensemble algorithm model; CLNM, central lymph node metastasis; US, ultrasound; PTC, papillary thyroid cancer; FT3, free T3; FT4, free T4; TSH, thyroid-stimulating hormone; TG, thyroglobulin; TG-Ab, anti-thyroglobulin antibody; TPO-Ab, anti-thyroid peroxidase antibody; TR-Ab, thyroid-stimulating hormone receptor antibody; ML, machine learning; AUC, area under the curve; HEM, heterogenous ensemble algorithm model; RF, random

forest classifier; KNN, k-nearest neighbors; XGBoost, extreme gradient boosting; GB, gradient boosting; US, ultrasonography; AD, Adaboost; OR, odds ratio; CI, confidence interval.

Data Sharing Statement

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

Ethics Approval

The study was conducted with the approval of the Ethics Committee of Shanghai Jiao Tong University Affiliated Sixth People's Hospital and in accordance with the declaration of Helsinki. Written informed consent was obtained from all participants.

Code Availability

The code used or analyzed during the current study is available from the corresponding author on reasonable request.

Acknowledgments

The authors would like to thank all the reviewers who participated in the review. We would like to thank Editage (www.editage.com) for English language editing.

Funding

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Disclosure

The authors report no conflicts of interest.

References

1. Carling T, Udelsman R. Thyroid Cancer. *Annu Rev Med*. 2014;65(1):125–137. doi:10.1146/annurev-med-061512-105739
2. La Vecchia C, Malvezzi M, Bosetti C, et al. Thyroid cancer mortality and incidence: a global overview. *Int J Cancer*. 2015;136(9):2187–2195. doi:10.1002/ijc.29251
3. Lim H, Devesa SS, Sosa JA, Check D, Kitahara CM. Trends in thyroid cancer incidence and mortality in the United States, 1974–2013. *JAMA*. 2017;317(13):1338–1348. doi:10.1001/jama.2017.2719
4. Marotta V, Malandrino P, Russo M, et al. Fathoming the link between anthropogenic chemical contamination and thyroid cancer. *Crit Rev Oncol Hematol*. 2020;150:102950. doi:10.1016/j.critrevonc.2020.102950
5. Marotta V, Sciammarella C, Colao A, Faggiano A. Application of molecular biology of differentiated thyroid cancer for clinical prognostication. *Endocr Relat Cancer*. 2016;23(11):R499–R515. doi:10.1530/ERC-16-0372
6. Marotta V, Sciammarella C, Capasso M, et al. Germline polymorphisms of the VEGF pathway predict recurrence in nonadvanced differentiated thyroid cancer. *J Clin Endocrinol Metab*. 2017;102(2):661–671. doi:10.1210/jc.2016-2555
7. Cabanillas ME, McFadden DG, Durante C. Thyroid cancer. *Lancet*. 2016;388(10061):2783–2795. doi:10.1016/S0140-6736(16)30172-6
8. Al Afif A, Williams BA, Rigby MH, et al. Multifocal papillary thyroid cancer increases the risk of central lymph node metastasis. *Thyroid*. 2015;25(9):1008–1012. doi:10.1089/thy.2015.0130
9. Xiao GZ, Gao L. Central lymph node metastasis: is it a reliable indicator of lateral node involvement in papillary thyroid carcinoma? *World J Surg*. 2010;34(2):237–241. doi:10.1007/s00268-009-0347-1
10. Cheema Y, Repplinger D, Elson D, Chen H. Is tumor size the best predictor of outcome for papillary thyroid cancer? *Ann Surg Oncol*. 2006;13(11):1524–1528. doi:10.1245/s10434-006-9176-8
11. Calò PG, Pisano G, Medas F, et al. Total thyroidectomy without prophylactic central neck dissection in clinically node-negative papillary thyroid cancer: is it an adequate treatment? *World J Surg Oncol*. 2014;12(1):152. doi:10.1186/1477-7819-12-152
12. Shan CX, Zhang W, Jiang DZ, Zheng XM, Liu S, Qiu M. Routine central neck dissection in differentiated thyroid carcinoma: a systematic review and meta-analysis. *Laryngoscope*. 2012;122(4):797–804. doi:10.1002/lary.22162
13. Docimo G, Tolone S, Ruggiero R, et al. Tiroidectomia totale senza dissezione profilattica del compartimento centrale del collo combinata con somministrazione di routine di calcio e vitamina D per via orale: È una scelta giusta per ottenere un basso tasso di ricorrenza evitando l'ipocalcemia? Uno studio retrospettivo. *Minerva Chir*. 2013;68(3):321–328.
14. White ML, Gauger PG, Doherty GM. Central lymph node dissection in differentiated thyroid cancer. *World J Surg*. 2007;31(5):895–904. doi:10.1007/s00268-006-0907-6
15. Barczyński M, Konturek A, Stopa M, Nowak W. Prophylactic central neck dissection for papillary thyroid cancer. *Br J Surg*. 2013;100(3):410–418. doi:10.1002/bjs.8985
16. Conzo G, Pasquali D, Bellastella G, et al. Total thyroidectomy, without prophylactic central lymph node dissection, in the treatment of differentiated thyroid cancer. Clinical retrospective study on 221 cases. *Endocrine*. 2013;44(2):419–425. doi:10.1007/s12020-013-9877-2

17. 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. doi:10.1016/j.ejrad.2019.01.006
18. Hwang HS, Orloff LA. Efficacy of preoperative neck ultrasound in the detection of cervical lymph node metastasis from thyroid cancer. *Laryngoscope.* **2011**;121(3):487–491. doi:10.1002/lary.21227
19. Xing Z, Qiu Y, Yang Q, et al. Thyroid cancer neck lymph nodes metastasis: meta-analysis of US and CT diagnosis. *Eur J Radiol.* **2020**;129:109103. doi:10.1016/j.ejrad.2020.109103
20. Mitchell TM, Carbonell JG, Michalski RS. Q-learning. *Mach Learn.* **1986**;12. doi:10.1007/978-1-4613-2279-5
21. Deo RC. Machine learning in medicine. *Circulation.* **2015**;132(20):1920–1930. doi:10.1161/CIRCULATIONAHA.115.001593
22. Yang L, Wu H, Jin X, et al. Study of cardiovascular disease prediction model based on random forest in eastern China. *Sci Rep.* **2020**;10(1). doi:10.1038/s41598-020-62133-5
23. Murugan A, Nair SAH, Kumar KPS. Detection of skin cancer using SVM, random forest and kNN classifiers. *J Med Syst.* **2019**;43(8). doi:10.1007/s10916-019-1400-8
24. Polikar R. Ensemble learning. *Ensemble Mach Learn.* **2012**;1–34. doi:10.1007/978-1-4419-9326-7_1
25. Wu Y, Rao K, Liu J, et al. Machine learning algorithms for the prediction of central lymph node metastasis in patients with papillary thyroid cancer. *Front Endocrinol.* **2020**;11. doi:10.3389/fendo.2020.577537
26. Friedman JH. Greedy function approximation: a gradient boosting machine. *Ann Stat.* **2001**;29(5):1189–1232. doi:10.1214/aos/1013203451
27. Gao M, Ge M, Ji Q, et al. 2016 Chinese expert consensus and guidelines for the diagnosis and treatment of papillary thyroid microcarcinoma. *Cancer Biol Med.* **2017**;14(3):203–211. doi:10.20892/j.issn.2095-3941.2017.0051
28. Tuttle RM, Morris LF, Haugen BR, et al. *AJCC Cancer Staging Manual.* Springer; **2017**. doi:10.1007/978-3-319-40618-3
29. Lew JI, Rodgers SE, Solorzano CC. Developments in the use of ultrasound for thyroid cancer. *Curr Opin Oncol.* **2010**;22(1):11–16. doi:10.1097/CCO.0b013e3283337f16
30. Lee JY, Na DG, Yoon SJ, et al. Ultrasound malignancy risk stratification of thyroid nodules based on the degree of hypoechogenicity and echotexture. *Eur Radiol.* **2020**;30(3):1653–1663. doi:10.1007/s00330-019-06527-8
31. Feng JW, Pan H, Wang L, Ye J, Jiang Y, Qu Z. Total tumor diameter: the neglected value in papillary thyroid microcarcinoma. *J Endocrinol Invest.* **2020**;43(5):601–613. doi:10.1007/s40618-019-01147-x
32. Liu WC, Li ZQ, Luo ZW, Liao WJ, Liu ZL, Liu JM. Machine learning for the prediction of bone metastasis in patients with newly diagnosed thyroid cancer. *Cancer Med.* **2021**;10(8):2802–2811. doi:10.1002/cam4.3776
33. Ogunleye A, Wang QG. XGBoost model for chronic kidney disease diagnosis. *IEEE.* **2020**;17(6):2131–2140. doi:10.1109/TCBB.2019.2911071
34. Wang K, Xu J, Li S, Liu S, Zhang L. Population-based study evaluating and predicting the probability of death resulting from thyroid cancer among patients with papillary thyroid microcarcinoma. *Cancer Med.* **2019**;8(16):6977–6985. doi:10.1002/cam4.2597
35. Rajeev P, Ahmed S, Ezzat TM, Sadler GP, Mihai R. The number of positive lymph nodes in the central compartment has prognostic impact in papillary thyroid cancer. *Langenbeck's Arch Surg.* **2013**;398(3):377–382. doi:10.1007/s00423-012-1041-6
36. So YK, Seo MY, Son YI. Prophylactic central lymph node dissection for clinically node-negative papillary thyroid microcarcinoma: influence on serum thyroglobulin level, recurrence rate, and postoperative complications. *Surgery.* **2012**;151(2):192–198. doi:10.1016/j.surg.2011.02.004
37. Salem FA, Bergenfelz A, Nordenström E, Almquist M. Central lymph node dissection and permanent hypoparathyroidism after total thyroidectomy for papillary thyroid cancer: population-based study. *Br J Surg.* **2020**. doi:10.1002/bjs.12028
38. Prim MP, De Diego JI, Hardisson D, Madero R, Gavilan J. Factors related to nerve injury and hypocalcemia in thyroid gland surgery. *Otolaryngol Head Neck Surg.* **2001**;124(1):111–114. doi:10.1067/mhn.2001.112305
39. Ardakani AA, Reiaz R, Mohammadi A. A clinical decision support system using ultrasound textures and radiologic features to distinguish metastasis from tumor-free cervical lymph nodes in patients with papillary thyroid carcinoma. *J Ultrasound Med.* **2018**;37(11):2527–2535. doi:10.1002/JUM.14610
40. Kuzan TY, Canbey Goret C. Comparison of number of passes and cytopathological specimen adequacy for thyroid fine-needle aspiration biopsy in the absence of an on-site pathologist. *Eur Thyroid J.* **2020**;9(1):49–54. doi:10.1159/000504094
41. Gashler M, Giraud-Carrier C, Martinez T. Decision tree ensemble: small heterogeneous is better than large homogeneous. *IEEE.* **2009**;900–905. doi:10.1109/icmla.2008.154
42. 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(1). doi:10.1186/s12885-019-5835-6
43. Aydoğan Bİ, Mutlu ABB, Yüksel S, et al. The association of histologically proven chronic lymphocytic thyroiditis with clinicopathological features, lymph node metastasis, and recurrence rates of differentiated thyroid cancer. *Endocr Pathol.* **2021**;32(2):280–287. doi:10.1007/s12022-020-09653-y
44. Lee I, Kim HK, Soh EY, Lee J. The association between chronic lymphocytic thyroiditis and the progress of papillary thyroid cancer. *World J Surg.* **2020**;44(5):1506–1513. doi:10.1007/s00268-019-05337-9
45. Babli S, Payne RJ, Mitmaker E, Rivera J. Effects of chronic lymphocytic thyroiditis on the clinicopathological features of papillary thyroid cancer. *Eur Thyroid J.* **2018**;7(2):95–101. doi:10.1159/000486367
46. Zhou YL, Gao EL, Zhang W, et al. Factors predictive of papillary thyroid micro-carcinoma with bilateral involvement and central lymph node metastasis: a retrospective study. *World J Surg Oncol.* **2012**;10(1). doi:10.1186/1477-7819-10-67
47. Scappaticcio L, Trimboli P, Verburg FA, Giovanella L. Significance of “de novo” appearance of thyroglobulin antibodies in patients with differentiated thyroid cancer. *Int J Biol Markers.* **2020**;35(3):41–49. doi:10.1177/1724600820931517
48. Iannuccilli JD, Cronan JJ, Monchik JM. Risk for malignancy of thyroid nodules as assessed by sonographic criteria: the need for biopsy. *J Ultrasound Med.* **2004**;23(11):1455–1464. doi:10.7863/jum.2004.23.11.1455
49. Jeon S, Kwon SY, Ryu YJ, et al. Combined role of lymph node ratio and serum thyroglobulin levels in predicting prognosis of papillary thyroid carcinoma. *Nucl Med Commun.* **2020**;41(8):733–739. doi:10.1097/MNM.0000000000001214

International Journal of General Medicine**Dovepress****Publish your work in this journal**

The International Journal of General Medicine is an international, peer-reviewed open-access journal that focuses on general and internal medicine, pathogenesis, epidemiology, diagnosis, monitoring and treatment protocols. The journal is characterized by the rapid reporting of reviews, original research and clinical studies across all disease areas. The manuscript management system is completely online and includes a very quick and fair peer-review system, which is all easy to use. Visit <http://www.dovepress.com/testimonials.php> to read real quotes from published authors.

Submit your manuscript here: <https://www.dovepress.com/international-journal-of-general-medicine-journal>