

RESEARCH

Open Access



Preoperative prediction of extrathyroidal extension: radiomics signature based on multimodal ultrasound to papillary thyroid carcinoma

Fang Wan¹, Wen He^{1*}, Wei Zhang¹, Yukang Zhang¹, Hongxia Zhang¹ and Yang Guang^{1*}

Abstract

Background There is a recognized need for additional approaches to improve the accuracy of extrathyroidal extension (ETE) diagnosis in papillary thyroid carcinoma (PTC) before surgery. Up to now, multimodal ultrasound has been widely applied in disease diagnosis. We investigated the value of radiomic features extracted from multimodal ultrasound in the preoperative prediction of ETE.

Methods We retrospectively pathologically confirmed PTC lesions in 235 patients from January 2019 to April 2022 in our hospital, including 45 ETE lesions and 205 non-ETE lesions. MaZda software was employed to obtain radiomics parameters in multimodal sonography. The most valuable radiomics features were selected by the Fisher coefficient, mutual information, probability of classification error and average correlation coefficient methods (F + MI + PA) in combination with the least absolute shrinkage and selection operator (LASSO) method. Finally, the multimodal model was developed by incorporating the clinical records and radiomics features through fivefold cross-validation with a linear support vector machine algorithm. The predictive performance was evaluated by sensitivity, specificity, accuracy, F1 scores and the area under the receiver operating characteristic curve (AUC) in the training and test sets.

Results A total of 5972 radiomics features were extracted from multimodal sonography, and the 13 most valuable radiomics features were selected from the training set using the F + MI + PA method combined with LASSO regression. The multimodal prediction model yielded AUCs of 0.911 (95% CI 0.866–0.957) and 0.716 (95% CI 0.522–0.910) in the cross-validation and test sets, respectively. The multimodal model and radiomics model showed good discrimination between ETE and non-ETE lesions.

Conclusion Radiomics features based on multimodal ultrasonography could play a promising role in detecting ETE before surgery.

[†]Wen He, Yang Guang contributed equally to this work as co-corresponding authors.

*Correspondence:

Wen He

hewen@bjtth.org

Yang Guang

guangycmu@163.com

Full list of author information is available at the end of the article



© The Author(s) 2023. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

Keywords Multimodal ultrasound, Extrathyroidal extension, Radiomics signature

Introduction

Extrathyroidal extension (ETE) reflects the spread of a primary thyroid tumour beyond the thyroid gland capsule. It appears in 5–45% of papillary thyroid carcinomas (PTCs) according to previous research [1]. Depending on the extent of invasion, ETE is subclassified as gross ETE (gETE) and minimal ETE (mETE). gETE is regarded as a macroscopic event that was initially suspected or identified by intraoperative, radiologic, or clinical examination and is associated with disease recurrence and survival. mETE is defined as microscopically detected invasion into perithyroidal soft tissue [2]. However, the importance of mETE remains a much-debated topic over the years. However, the 8th American Joint Committee on Cancer (AJCC) no longer includes mETE as the protocol to define T3 for tumour staging of PTC [3]. Some studies have shown that mETE in patients with PTC presents aggressive biological behaviour and is closely related to the risk of recurrence and metastasis [4, 5].

Ultrasound (US) is the first choice for the diagnosis of thyroid cancer and suspicious cervical lymphadenopathy. When physicians evaluate the risk of malignancy through sonographic patterns, evidence of ETE is also the point of examination according to the 2015 American Thyroid Association (ATA) Management Guidelines [6]. Previous work has only focused on the visible characteristics of US while ignoring the limitations of human visual resolution. It has been reported that US estimation of minimal ETE is less sensitive (30.1%) [7]. Thus, there is a recognized need for additional approaches to improve the accuracy of ETE diagnosis on US examinations before surgery. Radiomics allows the rapid quantitative extraction of countless high-throughput features from digital images and is widely used to solve medical problems, such as predicting malignant disease and lymph metastasis [8]. Recent research on US estimation of ETE has been carried out with radiomics based on B-mode images. This approach has enabled more precise prediction of ETE [9].

Extensive research has shown that multimodal US, including B-mode US (BMUS), colour Doppler flow imaging (CDFI), shear-wave elastography (SWE), superb microvascular imaging (SMI) and contrast-enhanced US (CEUS), has been frequently applied to differentiate benign and malignant lesions in the thyroid, breast and liver [10–12]. Hard malignancy as measured with elastographic results was significantly associated with pathological extrathyroidal extension [13]. Quantitative CEUS analysis showed that the time from peak to one-half has good diagnostic value in detecting ETE [14]. To date, far too little attention has been given to predicting ETE by multimodal US radiomics. In this research, we aim to

develop and validate a state-of-the-art radiomics model based on clinical records and multimodal US for predicting ETE in PTC patients. Our findings should contribute to the field of noninvasive assessment of ETE and clinical decisions about PTC.

Methods

The study was approved by the Ethical Committee of the Beijing Tiantan Hospital of Capital Medical University and complied with the Declaration of Helsinki. All patients signed informed consent before the CEUS examination.

Patients

This retrospective trial assessed consecutive individuals with thyroid nodules first diagnosed from January 2019 to April 2022 at Beijing Tiantan Hospital. All patients were examined by multimodal US and subsequently administered thyroid surgery, subtotal or total thyroidectomy, within a month following the US examination.

The inclusion criteria were as follows: (1) age ≥ 18 years; (2) primary PTC confirmed after surgery; and (3) multimodal US performed one month before surgery. The exclusion criteria were as follows: (1) preoperative anticancer therapy (radiotherapy, chemotherapy, etc.); (2) no association of pathological results with US imaging findings; (3) poor US quality; and (4) insufficient pathological samples for the assessment of ETE.

Based on a previous study [9], with an ETE prevalence of 20%, a sensitivity of 70%, and a specificity of 80%, the necessary sample size was calculated to be 47 ETE cases and 188 non-ETE cases to detect such a difference at a two-sided α level of 0.05 with 80% power. We used all available data to maximize the power and generalizability of our results.

Clinical and laboratory information

The following clinical and laboratory variables were retrospectively considered: age, sex, tumour location and thyroid function laboratory test results. All data were recorded before the operation. Thyroid function tests, including thyroglobulin (Tg), thyroid peroxidase antibody (TPOAb), and thyroglobulin antibody (TgAb), were performed in our hospital within one month before surgery [15]. The laboratory results were classified as low/normal/high based on comparison with the normal range, and missing data were imputed using the median value.

Ultrasound examination

Two sonologists with more than 5 years of experience in thyroid US performed examinations in this study.

For each tumour, BMUS, SWE, SMI and CEUS images were acquired on an Aplio 900 or Aplio 500 US system (Toshiba, Tokyo, Japan) with a linear array transducer i18LX5 probe. BMUS scans were first performed to find the optimal scanning plane. Then, CDFI, SMI, SWE and CEUS examinations were performed. For CEUS imaging, 5 mL of the sulphur hexafluoride microbubble contrast agent SonoVue (Bracco SpA, Milan, Italy) was injected into the median cubital vein of patients. Data acquisition was started after injection and lasted for at least 120 s. After that, sonologists performed a neck ultrasound to report whether suspicious cervical lymph nodes existed. Images from BMUS, CDFI, SMI, SWE and CEUS were exported in BMP format for further analysis. The frame at the peak intensity in the CEUS video was used to analyse the enhancement pattern.

Image interpretation and analysis

The multimodal US characteristics of nodules were evaluated during real-time and multiview scanning before surgery. According to Lamartina L et al. [7] and Zhang Y et al. [16], the BMUS signs of ETE diagnosis were classified as follows: the angle of contact with the thyroid capsule (absent capsule contact, 0; acute, $<90^\circ$; straight or obtuse, = or $>90^\circ$), the degree of contact ($<25\%$, 25–50% or $>50\%$ circumference of thyroid nodule contact with the capsule), bulging (nodule that deforms thyroid contour and bulges out), and capsule echogenic loss. The CDFI and SMI characteristics of suspicious ETE were vascularity extending beyond the capsule. The CEUS characteristics of suspicious ETE were

classified as follows: (1) discontinuous capsular enhancement (the enhancement of the anterior and/or posterior hyperechoic thyroid capsular was discontinued) and (2) enhancement extending beyond the capsule (enhancement extended out of the capsule). The ultrasonographic representation for the diagnosis of extrathyroidal extension on US is shown in Fig. 1. The 2015 ATA guidelines were used to guide reporting cervical lymph node metastasis on US (US-LNM) [6].

Radiomics feature extraction

MaZda software (version 4.6, the Institute of Electronics, Technical University of Lodz, <http://www.eletel.p.lodz.pl/programy/mazda/>) was used to extract radiomics features from each lesion [17]. We applied image normalization to $\mu \pm 3\delta$ (μ : mean grey-level value, δ : standard deviation) to decrease intensity bias from different images [18, 19]. Radiomics features were automatically extracted from the region of interest (ROI) on the US image of the largest cross-section with MaZda software, including shape features and six common texture feature groups (histogram, absolute gradient, grey-level co-occurrence matrix, run length matrix, autoregressive model, and wavelet transform). Shape features were extracted from the actual ROI, which was manually delineated along the border of each tumour on the BMUS images. Considering that the periphery of the tumour also contained helpful information, other texture features were extracted from the disk structure dilated from the original segmented ROIs (the dilated ROI had a diameter approximately 10% larger than the contour of the lesion)

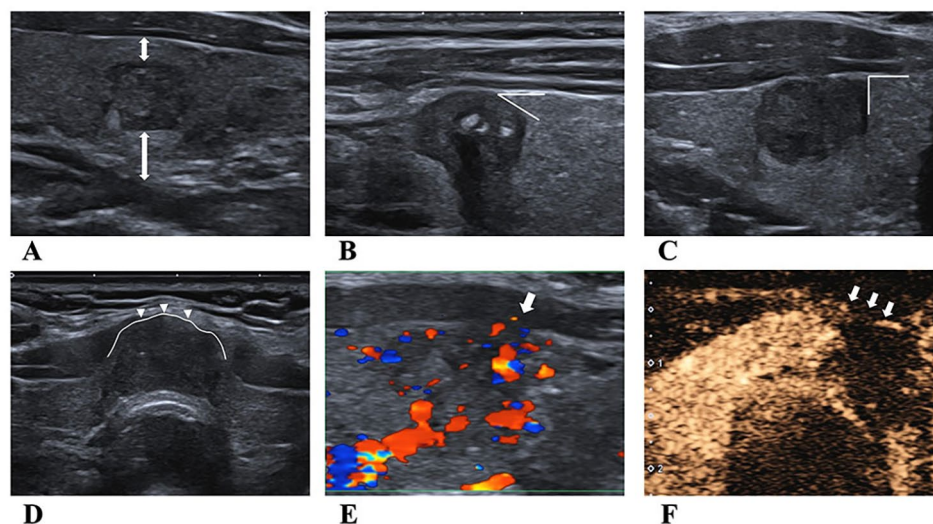


Fig. 1 Ultrasonography for diagnosis of extrathyroidal extension

(A) The lesion is not in contact with the capsule. According to the classification criteria, the contact angle with the thyroid capsule is 0° . (B) The lesion formed an acute angle with the thyroid capsule. (C) The lesion formed a straight angle with the thyroid capsule. (D) Thyroid capsule deformation for the bulging nodule. (E) The arrow indicates that the CDFI vascularity of suspicious extrathyroidal extension nodules extended beyond the thyroid capsule. (F) The arrows pointing at contrast-enhanced ultrasound imaging showed discontinuous enhancement of the anterior thyroid capsule

[20]. The reference point to create all the dilated ROIs was the gravity centre of the original ROI [21]. Because the CDFI and SWE images are three channels (RGB), we performed the colour analysis module function to extract features from each channel separately. The overall study process is shown in Fig. 2.

Intraobserver and interobserver agreement

Given the importance of radiological signature reproducibility, 20 samples were randomly selected to measure the interobserver agreement. The multimodal radiomics features of these samples were extracted by two independent radiologists (Guang Yang and Yukang Zhang) who had at least five years of experience performing US examinations and were blinded to patients' clinicopathological records. The first radiologist delineated the ROIs of these samples again after two weeks. Then, the delineation of the remaining images was completed by the first radiologist.

Dimensionality reduction and radiomics feature selection

To avoid the problem of dimensionality and reduce the bias in feature modelling, 30 optimal features were selected from shape features for BMUS, and 150 texture features were selected from texture features for BMUS, CDFI, SWE, SMI, and CEUS based on the Fisher coefficient, mutual information, probability of classification error and average correlation coefficient methods (F+MI+PA). After the normalization of radiomics data, least absolute shrinkage and selection operator (LASSO) regression was performed for dimensionality reduction.

The selected features were used to further develop prediction models.

Development and validation of ETE-predictive models

We randomly selected 74% of the samples as the training set, and the remaining 26% constituted the test set. The models were trained separately in the training dataset and were likewise validated independently in the testing dataset. Candidate variables included demographics, laboratory tests, selected radiomics features and observed multimodal characteristics that were risk factors for ETE ($P < 0.05$). To uncover the incremental value of the radiomics signatures to the risk for ETE estimation, both radiomics models and clinical models were developed. The multimodal model incorporated radiomics features and clinical risk factors based on a support vector machine with a linear kernel (linear SVM). A fivefold cross-validation protocol with randomly split training data was used to adjust the optimal weight parameters, and the accuracy for each fold was estimated to prevent overfitting. The performance of the models was evaluated and compared with respect to sensitivity, specificity, accuracy, and F1 scores. Receiver operating curves (ROCs) and precision-recall curves (PR curves) were plotted and quantified with the area under the curve (AUC) to evaluate the discriminatory performance between different models in the cross-validation and test cohorts.

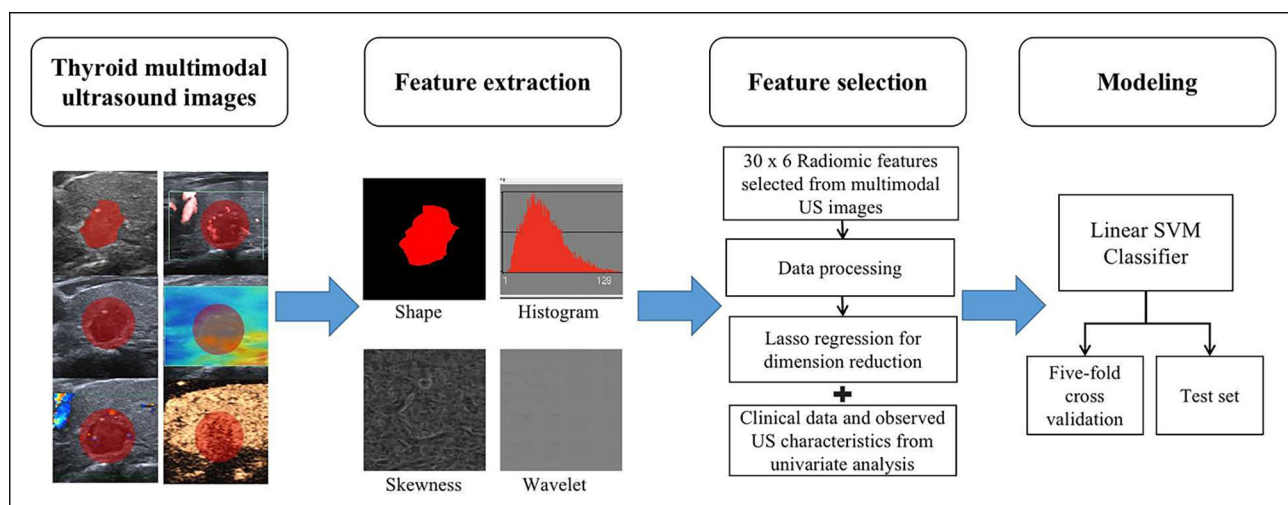


Fig. 2 The overall study process

The most representative image of each tumour on the thyroid multimodal ultrasound image was selected. Radiomics features, including shape, histogram, absolute gradient, grey-level co-occurrence matrix, run length matrix, autoregressive model, and wavelet transform, were extracted. Radiomics features were generated using the Fisher coefficient, mutual information, probability of classification error and average correlation coefficient methods (F+MI+PA) and LASSO. These selected features were used to train the linear SVM in fivefold cross-validation and test in an test set. Univariate analysis was performed to determine the association between the clinical variables and ETE. Another SVM classifier was built using radiomics features plus clinical variables and observed ultrasound characteristics. SVM: support vector machine

Pathology

All patients underwent hemithyroidectomy or total thyroidectomy. Tissue slices from all patients were independently reviewed by our institutional pathologist (with more than 15 years of experience) to confirm whether the tumour had ETE, which was defined as a tumour extending through the capsule only. mETE was defined as microscopic invasion of adjacent connective tissue, and gETE was defined as gross invasion involving perithyroidal strap muscles or beyond subcutaneous soft tissues, larynx, trachea, oesophagus, or recurrent laryngeal nerve [22].

Statistical analysis

Statistical analysis was conducted with SPSS 26.0 (SPSS, Chicago, United States) and MedCalc version 20.0.22 (MedCalc Software, Mariakerke, Belgium). Categorical variables were compared by the χ^2 test. Continuous variables were compared by the Mann-Whitney U test for abnormally distributed variables and t test for normally distributed variables. The level of significant difference reported was two-tailed, and p values of less than 0.05 were considered statistically significant. The inter/intra-class correlation coefficient (ICC) was used to evaluate interobserver and intraobserver agreement. ICC values were considered excellent for $ICC \geq 0.75$, satisfactory for $0.4 \leq ICC < 0.75$, and poor for $ICC < 0.4$. Z score normalization data processing was performed using R software version 3.4.1 with the caret package, and the glmnet package was used for LASSO regression. The classification learner toolbox in MATLAB R2021b (MathWorks, Natick, MA) was used to build linear SVM models for predicting ETE status. The discrimination metrics of established models, including AUC, classification accuracy, sensitivity, specificity and F1 scores, were calculated using MedCalc.

Results

Clinicopathological characteristics

Among the 1009 thyroid patients examined by multimodal US, 235 patients with 250 suspicious thyroid lesions on multimodal ultrasonography who underwent surgery between 01 January 2019 and 30 April 2022 were included (shown in Fig. 3). The 250 lesions were randomly allocated across two sets, namely, a training set and a test set, containing 184 lesions (40 males and 144 females; median age 41.0 years) and 66 lesions (15 males and 51 females; median age 41.5 years), respectively. The clinicopathological characteristics of the training and test sets are shown in Table 1. Positive ETE accounted for 17.9% (33/184) and 18.2% (12/66) of cases in the training and test sets, respectively. There was no significant difference between the two cohorts in the presence of ETE ($P=0.964$). In addition, there were no significant

differences between the two cohorts in other clinicopathological characteristics. These results justified the use of the training and test cohorts.

Clinical and multimodal US characteristics

The multimodal US characteristics of the study population are presented in Table 2. Univariate analysis was used to identify potential variables associated with ETE. The Tg level, maximum size, US-LNM, degree of contact with the capsule, angle of contact, bulging, capsule echogenic loss on BMUS, SMI vascularity extending beyond the capsule and discontinuous capsular enhancement were associated with ETE in PTCs ($P < 0.05$).

ETE prediction

Based on the training cohort, a total of 5972 radiomics features were extracted from multimodal sonography. These features were reduced to 180 using F+MI+PA, and LASSO regression was used to select the predictors with the greatest potential (Fig. 4). One shape feature, two BMUS features, three SMI features, three SWE features and four CEUS features were selected to further develop prediction models. The proportion of texture features derived from CEUS was the highest (30.8%). All 13 radiomics features showed highly significant differences between ETE and non-ETE masses with a t test ($P < 0.05$, Table 3). The radiomic score calculation formula was as follows:

$$\begin{aligned} \text{Rad-score} = & -1.842561 - 0.03930704 \times \text{Shape-GeoX} + 0.23981317 \times \text{BMUS-S}_{(5,5)} \\ & + \text{InvDf-Mom} + 0.29167245 \times \text{BMUS-GrKurtosis} + 0.08002657 \\ & \times \text{CEUS-Vertl_GLevNonU} + 0.14981715 \times \text{CEUS-Vertl_RLNonUni} \\ & - 0.03253232 \times \text{CEUS-WavEnHL}_{s4} - 0.56894438 \times \text{CEUS-Perc.99\%} \\ & - 0.15750629 \times \text{SWE-H}_{(0,1)} + \text{SumOfSqs} + 0.03310617 \times \text{SWE-Teta1} \\ & + 0.26388574 \times \text{SWE-R}_{\text{Horzl_GLevNonU}} + 0.02669084 \times \text{SMI-Vertl_GLevNonU} \\ & - 0.23046584 \times \text{SMI-WavEnHH}_{s5} + 0.12426557 \times \text{SMI-WavEnHL}_{s6}. \end{aligned}$$

The interobserver reproducibility and intraobserver reproducibility of feature extraction are shown in Table 4. All the ICC values were reported to be excellent or satisfactory. Therefore, all the selected radiomics features were input into the radiomics model.

Figure 5 shows ROC curves for the radiomics model in distinguishing ETE from non-ETE masses in the cross-validation and test cohorts. The multimodal prediction model yielded AUCs of 0.911 (95% CI 0.866–0.957) and 0.716 (95% CI 0.522–0.910) in the cross-validation (Fig. 5A) and test (Fig. 5B) sets, respectively. The clinical decision curve of the multimodal and clinical model is depicted in Fig. 6A and B. Table 5 shows the diagnostic performance of the three models. The sensitivity, specificity, accuracy, AUC and F1 score of the clinical SVM

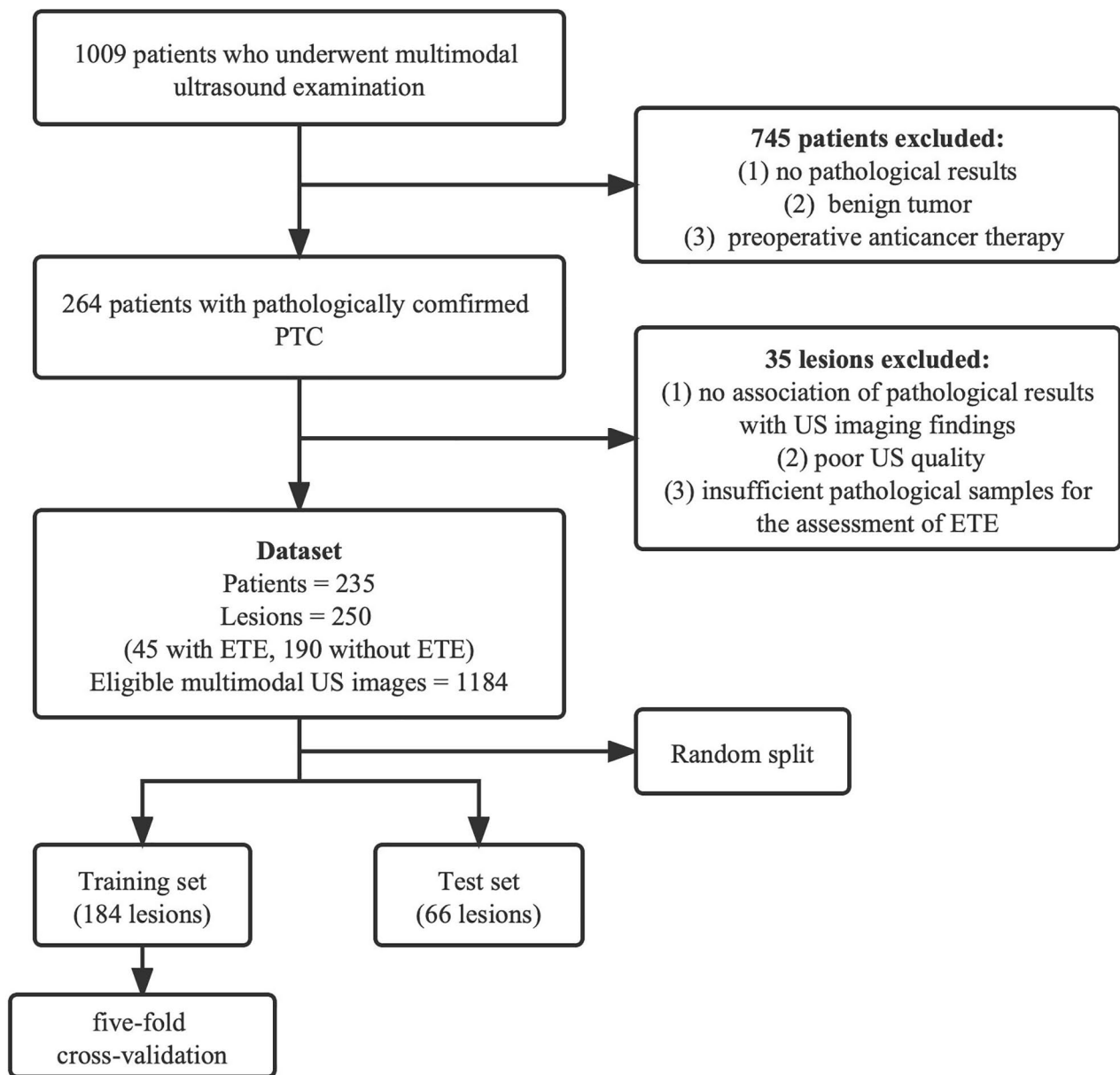


Fig. 3 Study flowchart

US: ultrasound; PTC: papillary thyroid carcinoma; ETE: extrathyroidal extension

model were 12.1%, 96.7%, 81.5%, 0.700, and 0.472 in the cross-validation set and 8.3%, 96.3%, 80.3%, 0.676, and 0.421 in the test set. The sensitivity, specificity and accuracy, AUC and F1 scores of the radiomics SVM model were 57.6%, 94.7%, 88.0%, 0.908, and 0.691 in the cross-validation set and 58.3%, 94.4%, 87.9%, 0.698, and 0.636 in the test set. Using combined clinical data and radiomics parameters, we developed a multimodal SVM model, which yielded 54.5%, 94.0%, 87.0%, 0.911, and 0.691 in the cross-validation set and 50.0%, 92.6%, 84.8%, 0.716, and 0.571 in the test set, respectively. In the cross-validation cohort, there was a significant difference between

the AUC of the multimodal model and the clinical model. However, there was no significant AUC difference among the three models in the test set (shown in Table S1).

Discussion

ETE suggests the aggressive behaviour of thyroid cancer. Previous studies have shown insufficient sensitivity and accuracy in ETE assessment before surgery. Another question is that the US evaluation seems more subjective and relies on expert opinion, which may cause poor interobserver agreement in capsule invasion diagnosis [23]. It is very urgent to utilize modern techniques to

Table 1 Clinicopathological characteristics of PTCs in the training and test sets

Clinicopathological characteristics	Training set (n=184)	Test set (n=66)	P Value
Gender			0.868
Male	40 (16.0%)	15 (6.0%)	
Female	144 (57.6%)	51 (20.4%)	
Age (years)	41.0 (33.0, 51.0)	41.5 (32.0, 53.0)	0.939
Maximum Size (cm)	0.8 (0.6, 1.2)	0.9 (0.6, 1.2)	0.915
Location			0.425
Isthmus	7 (2.8%)	4 (1.6%)	
Left lobe	91 (36.4%)	27 (10.8%)	
Right lobe	86 (34.4%)	35 (14.0%)	
US-LNM			0.177
Normal	146 (58.4%)	47 (18.8%)	
Suspicious	38 (15.2%)	19 (7.6%)	
Pathological LNM			0.227
No	94 (37.6%)	28 (11.2%)	
Yes	90 (36.0%)	38 (15.2%)	
ETE			0.964
No	151 (60.4%)	54 (21.6%)	
Yes	33 (13.2%)	12 (4.8%)	

Data are presented as median with interquartile range and number where applicable

PTC: papillary thyroid carcinoma; LNM: cervical lymph node metastasis; US-LNM: cervical lymph node metastasis on ultrasound; ETE: extrathyroidal extension

improve ETE diagnostic accuracy in preoperative thyroid US.

The present study was designed to determine the effect of multimodal US on ETE prediction before surgery. Researchers have investigated ETE by BMUS in prior studies, defining ETE as contact of more than 25% of the lesion with the thyroid capsule or loss of capsular hyperechogenicity, consistent with our study [24]. It has also been reported that vascularity beyond the capsule on CDFI images showed high specificity but low sensitivity [25]. Conversely, we did not find a relationship between CDFI vascularity and ETE. SMI is a mature method to display tiny blood flow [26]. It reached higher sensitivity in detecting microvessels of thyroid tumours, possibly being more valuable in predicting ETE. In the present study, with a larger sample size, we found that SMI vascularity performed better in diagnosing ETE than CDFI. CEUS has shown a solid ability to distinguish malignant and benign lesions by dynamically exhibiting blood supply patterns. The normal thyroid capsule manifests high enhancement during CEUS because of the vessels in the intrinsic capsule. Like Yan Zhang et al., we found that discontinuous enhancement of the thyroid capsule was highly suggestive of ETE [16]. However, lesion enhancement beyond the capsule could not predict ETE. This may be because few PTCs present hyper or iso-enhancement. Moreover, several studies have demonstrated the

correlation between LNM and ETE. Our study found that suspicious lymph nodes in cervical US examination also indicate the possibility of ETE.

Human vision has a limitation in distinguishing details in medical imaging, which causes limited applicability in ETE diagnosis. The clinical model established by linear SVM for ETE prediction showed insufficient discriminatory ability. To overcome this problem, we employed radiomics. Radiomics is known to rapidly extract numerable quantitative features from digital images through high-throughput computing for analysis [27]. It can be widely applied in medical imaging diagnosis and treatment decision-making. The experimental results show that radiomics features extracted from multimodal US images were independently associated with ETE. Adding these radiomics features to the clinical model increased the accuracy, AUC and F1 score in the cross-validation and test sets, which is also better than the previously reported clinical model based on US [7]. Therefore, our study suggested that radiomics has a favourable discriminatory ability to solve medical imaging classification problems. We also investigated cases with inconsistent predictive results from the clinical and multimodal models. Figure 7 provides a series of multimodal US images for two misdiagnosed cases. We speculated that the multimodal model enhances the sensitivity of ETE prediction, which may lead to overestimation or misdiagnosis of minor nodules with a low risk of ETE. On the other hand, the multimodal model still showed better performance than the clinical model, while the thyroidal capsule seemed to be continuous on BMUS imaging (Fig. 7E).

Previous studies showed that the computed tomography (CT) radiomics nomogram for preoperative prediction of ETE in PTC patients was slightly better than that in our research, with an AUC of 0.80–0.84 in the test set [28–30]. Wei R et al. used multimodal magnetic resonance imaging (MRI) radiomics scores to calculate the risk of ETE in PTCs with tumour diameters ≥ 5 mm, which reached an AUC of 0.87 [31]. Given the radiation of CT and renal unsafety of iodine-based contrast media in patients with chronic renal insufficiency and the cost and consuming time of MRI, we prefer to explore a predictive model based on US. However, there needs to be more studies investigating ETE by US imaging radiomics analysis. Although Wang X et al. developed a radiomics nomogram based on the BMUS radiomics score and clinical factors (tumour location and radiologist diagnosis) that reached good discrimination in the validation cohort (AUC=0.824) [9], an external test set is needed to confirm the generalization ability. In the present study, the SVM-supervised classifier was chosen rather than the nomogram for automatic computational analysis of the possibility of ETE. It performed with high accuracy

Table 2 Univariate analysis of clinical and multimodal US characteristics for ETE in the training set

Characteristics	Non-ETE (n = 151)	ETE (n = 33)	P Value
Gender			0.075
Male	29 (15.8%)	11 (6.0%)	
Female	122 (66.3%)	22 (12.0%)	
Age (years)	41.0 (33.0, 51.0)	44.0 (36.0, 52.5)	0.514
Tg			0.030
Low	34 (18.5%)	3 (1.6%)	
Normal	112 (60.9%)	26 (14.1%)	
High	5 (2.7%)	4 (2.2%)	
TgAb			0.640
Normal	118 (64.1%)	27 (14.7%)	
Abnormal	33 (17.9%)	6 (3.3%)	
TPOAb			0.206
Normal	118 (64.1%)	29 (15.8%)	
Abnormal	33 (17.9%)	4 (2.2%)	
Maximum size (cm)	0.70 (0.60,1.00)	1.20 (0.85,1.55)	< 0.001
Location			0.119
Isthmus	4 (2.2%)	3 (1.6%)	
Left Lobe	73 (39.7%)	18 (9.8%)	
Right Lobe	74 (40.2%)	12 (6.5%)	
US-LNM			0.047
No	124 (67.4%)	22 (12.0%)	
Yes	27 (14.7%)	11 (6.0%)	
Degree of contact			< 0.001
< 25%	93 (50.5%)	7 (3.8%)	
25-50%	46 (25.0%)	22 (12.0%)	
>50%	12 (6.5%)	4 (2.2%)	
Angle of contact			< 0.001
None	70 (38.0%)	3 (1.6%)	
Acute	63 (34.2%)	26 (14.1%)	
Straight or obtuse	18 (9.8%)	4 (2.2%)	
Bulging			0.014
No	118 (64.1%)	19 (10.3%)	
Yes	33 (17.9%)	14 (7.6%)	
Capsule echogenic loss			< 0.001
No	125 (67.9%)	16 (8.7%)	
Yes	26 (14.1%)	17 (9.2%)	
CDFI vascularity extending			0.419
No	143 (77.7%)	30 (16.3%)	
Yes	8 (4.4%)	3 (1.6%)	
SMI vascularity extending			0.009
No	145 (78.8%)	27 (14.7%)	
Yes	6 (3.3%)	6 (3.3%)	
Enhancement extending			0.082
No	145 (78.8%)	29 (15.8%)	
Yes	6 (3.3%)	4 (2.2%)	
Discontinuous capsular enhancement			< 0.001
No	127 (69.0%)	18 (9.8%)	
Yes	24 (13.0%)	15 (8.2%)	

Data are presented as median with interquartile range and number where applicable

US, ultrasound; ETE, extrathyroidal extension; Tg, thyroglobulin; TPOAb, thyroid peroxidase antibody; TgAb, thyroglobulin antibody; US, ultrasound; CDFI, colour Doppler flow imaging; SMI, superb microvascular imaging; US-LNM, cervical lymph node metastasis on ultrasound

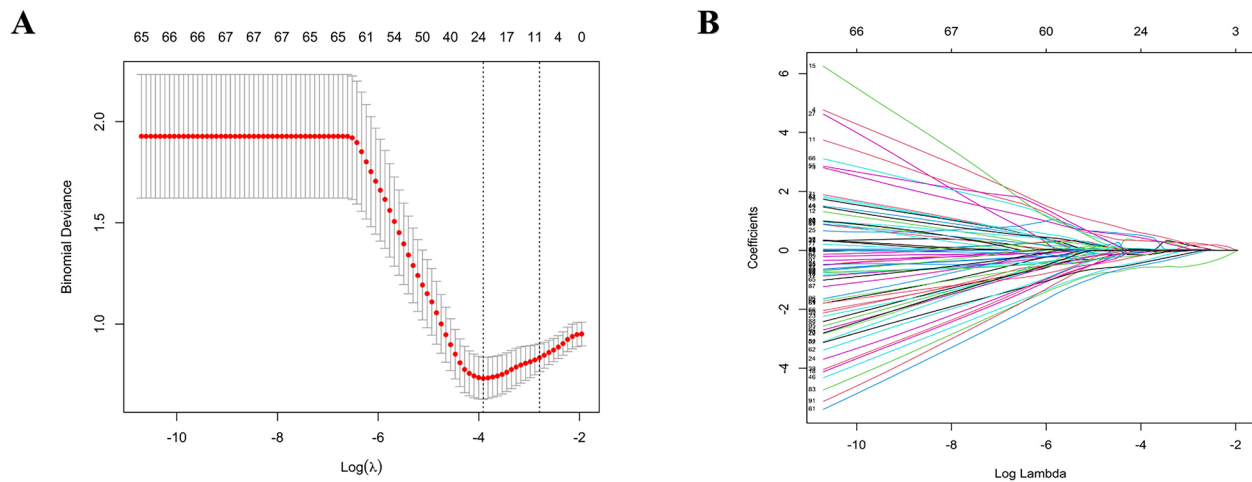


Fig. 4 Radiomics feature selection using LASSO regression in the training set (A) The 10-fold cross-validation process was used to generate the 13 selected features in LASSO regression for further modelling. (B) The feature coefficient convergence graph

Table 3 Comparison of radiomics values of ETE and non-ETE groups in the training set

Modality	Method	Radiomics parameters	Non-ETE (n=151)	ETE (n=33)	P
BMUS	Geometry	GeoX	$(6.26 \pm 0.88) \times 10^2$	$(5.88 \pm 1.02) \times 10^2$	
	Grey-level co-occurrence matrix	S (5,5) InvDfMom	$(1.60 \pm 0.30) \times 10^{-1}$	$(1.90 \pm 0.40) \times 10^{-1}$	
	Absolute gradient	GrKurtosis	2.47 ± 1.14	3.20 ± 1.01	
SMI	Grey-level run length matrix	Vertl_GLevNonU	$(6.85 \pm 5.20) \times 10^2$	$(12.58 \pm 9.07) \times 10^2$	
	Wavelets transform	WavEnHH_s5	$(7.63 \pm 3.18) \times 10^1$	$(5.78 \pm 1.89) \times 10^1$	
	Wavelets transform	WavEnHL_s6	$(1.11 \pm 0.6) \times 10^2$	$(1.36 \pm 0.77) \times 10^2$	
SWE	Grey-level co-occurrence matrix	H_S (0,1) SumOfSqs	$(1.08 \pm 0.70) \times 10^2$	$(1.03 \pm 0.88) \times 10^2$	
	Autoregressive model	Teta1	$(8.8 \pm 0.30) \times 10^{-1}$	$(9.0 \pm 0.30) \times 10^{-1}$	
	Grey-level run length matrix	R_Horzl_GLevNonU	$(5.26 \pm 4.21) \times 10^2$	$(9.64 \pm 6.54) \times 10^2$	
CEUS	Grey-level run length matrix	Vertl_GLevNonU	$(8.14 \pm 5.87) \times 10^2$	$(15.31 \pm 12.1) \times 10^2$	
	Grey-level run length matrix	Vertl_RLNonUni	$(1.98 \pm 1.46) \times 10^4$	$(3.68 \pm 2.83) \times 10^4$	
	Wavelets transform	WavEnHL_s4	$(1.04 \pm 0.30) \times 10^2$	$(0.85 \pm 0.20) \times 10^2$	
	Histogram	Perc.99%	$(1.96 \pm 0.15) \times 10^2$	$(1.78 \pm 0.25) \times 10^2$	

BMUS, B-mode US; CDFI, colour Doppler flow imaging; SWE, shear-wave elastography; SMI, superb microvascular imaging; CEUS, contrast-enhanced ultrasound. S(x,y), grey level co-occurrence matrix for inter-pixel distance x along rows and y along columns; InvDfMom, inverse difference moment; Gr Kurtosis, absolute gradient kurtosis; Horzl, Horizontal; Vertl, vertical; GLevNonU, gray-level non-uniformity; WavEnHH (HL), energies of wavelet transform coefficients in frequency channels HH (HL); SumOfSqs, Sum of squares; RLNonUni, run-length non-uniformity; Perc.99%, percentile 99%

and a good guarantee against overfitting. The multimodal model based on multimodal US radiomics features and clinical records by linear SVM yielded a cross-validation AUC of 0.911 and a test AUC of 0.716, indicating that multimodal US contained valuable information about ETE. TableS2 provides a detailed comparison of the data that have been reported previously for ETE prediction.

However, the number of mETE samples in this study was much larger than that of gETE samples. There were 41 nodules with mETE and 4 with gETE. Thus, this research used the same model to predict both of them. We utilized the subgroup accuracy to assess the predictive ability in different subgroups [subgroup accuracy=the number of correct classifications/total number of samples in the subgroup (%)] (Table 6). The model

could predict mETE as well as gETE. Fisher’s exact test showed no difference between the subgroup accuracy of the multimodal model (P value of the cross-validation: 0.489; P value of the test set: 1.00).

A major clinical contribution of this paper is providing the possibility of improving individualized treatment in PTC patients. However, the 8th AJCC no longer includes mETE as the protocol to define T3 for tumour staging of PTC. ETE is still a prognostic factor, according to previous studies. Evidence shows that active surveillance is an alternative to surgery in low-risk PTC patients. However, the therapeutic strategies for those with ETE may need to be stricter. The pathological standard for ETE in PTC lesions remains controversial [1, 2]. The incidence of pathological ETE in our study is reasonable due to the

Table 4 The inter-observer reproducibility and intra-observer reproducibility of selected feature

Selected features	Interclass coefficient	Intraclass coefficient
Shape-GeoX	0.993	0.986
BMUS-S (5,5) InvDfMom	0.895	0.947
BMUS-GrKurtosis	0.705	0.772
SMI-Vertl_GLevNonU	0.858	0.790
SMI-WavEnHH_s5	0.802	0.862
SMI-WavEnHL_s6	0.668	0.798
SWE-H_S (0,1) SumOfSqs	0.721	0.966
SWE-Teta1	0.402	0.652
SWE-R_Horzl_GLevNonU	0.740	0.774
CEUS-Vertl_GLevNonU	0.816	0.930
CEUS-Vertl_RLNonUni	0.840	0.900
CEUS-WavEnHL_s4	0.858	0.942
CEUS-Perc.99%	0.423	0.908

Inter/intra class correlation coefficient values were considered excellent for $ICC \geq 0.75$, satisfactory for $0.4 \leq ICC < 0.75$, poor for $ICC < 0.4$

prevalence of micro-PTC. Thus, the established model may provide better application in the real world.

Our study has several limitations. First, none of the qualitative multimodal features (e.g., peak intensity, time to peak, SMI vascularity index, and shear wave speed) were included in our study. Further research should use these parameters as input in predictive models. Second, it was a retrospective, one-centre study with unavoidable selection bias. More prospective and multicentre studies are required to explore the possible relationship between multimodal US features and the presence of ETE. Third, radiomics features extracted from static images lead to missing some of the information, especially for CEUS images. Evaluating the thyroid nodule with dynamic real-time US assessment could provide more accurate ETE results. More importantly, a major decrease in AUC in the test set indicated that the model was overfitted. Although we use 5-fold cross-validation to minimize overfitting when modelling, there is still a risk of

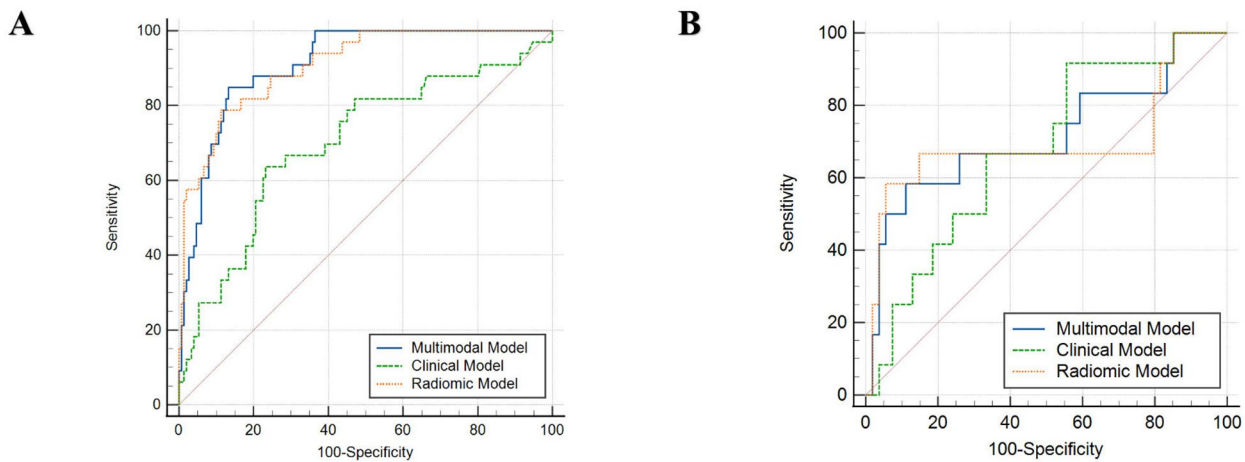


Fig. 5 (A) ROC curves of the three models in the cross-validation set. (B) ROC curves of the three models in the test set

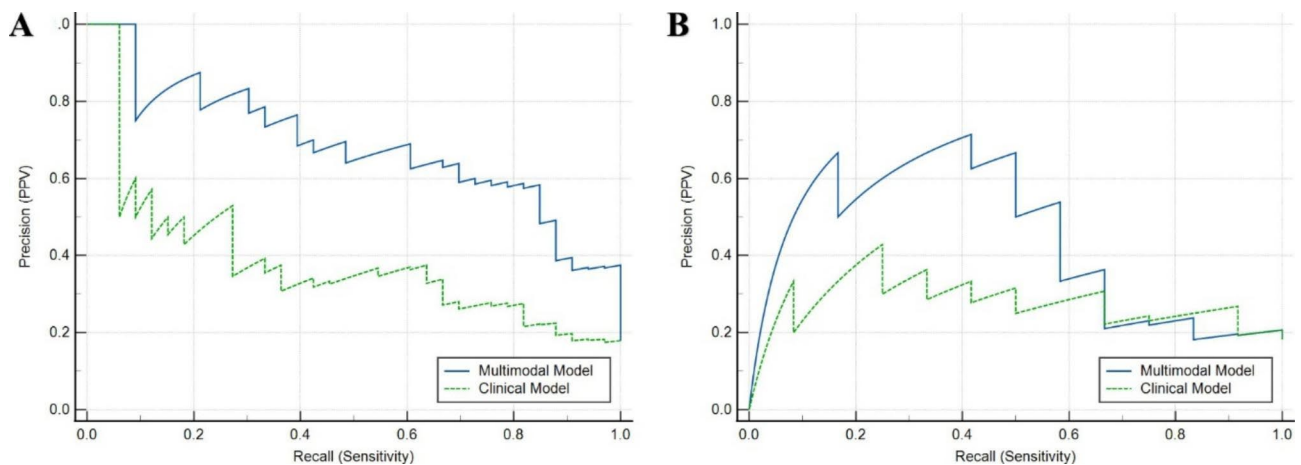


Fig. 6 (A) The precision-recall curve in the cross-validation set. (B) The precision-recall curve in the test set

Table 5 Diagnosis performance of the three models

Group	Model	Sen	Spe	Acc	AUC (95% CI)	F1 score
Cross-validation	Multimodal Model	54.5%	94.0%	87.0%	0.911 (0.866–0.957) *	0.691
	Clinical Model	12.1%	96.7%	81.5%	0.700 (0.593–0.807)	0.472
	Radiomic Model	57.6%	94.7%	88.0%	0.908 (0.857–0.960) *	0.691
Test set	Multimodal Model	50.0%	92.6%	84.8%	0.716 (0.522–0.910)	0.571
	Clinical Model	8.3%	96.3%	80.3%	0.676 (0.513–0.839)	0.421
	Radiomic Model	58.3%	94.4%	87.9%	0.698(0.477–0.918)	0.636

*Compared to AUC of Clinical Model in cross-validation cohort, there is a significant different ($P < 0.05$). Sen, sensitivity; Spe, specificity; Acc, accuracy; AUC, area under the curve; CI, confidence interval

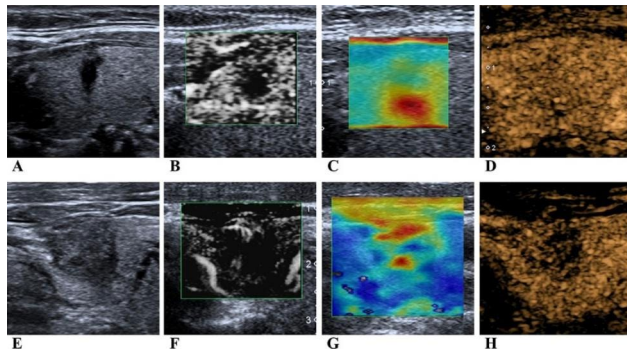


Fig. 7 A~D, multimodal ultrasound images for a non-ETE tumour misdiagnosed by the multimodal model but correctly diagnosed by the clinical model; E~H, an ETE case misdiagnosed by clinical models but diagnosed correctly by radiomics

overfitting in machine learning. This is more likely if the sample size is small and unbalanced. Thus, if sufficient training examples are provided, a deep learning system capable of video image processing will be able to address this issue in the future.

Conclusions

The diagnostic power of the clinical model in ETE prediction was inferior to the diagnostic power of the radiomics model based on multimodal US. The superiority of multimodal US radiomics features in predicting ETE provides further evidence of the potential clinical value of radiomics analyses in assessing ETE. The linear SVM predictive model based on information supplied by multimodal US radiomics and clinical features demonstrates a promising approach for predicting preoperative ETE in PTC patients.

Abbreviations

PTC	Primary papillary thyroid carcinoma
ETE	Extrathyroidal extension

mETE	Micro extrathyroidal extension
gETE	Gross extrathyroidal extension
US	Ultrasound
LNM	Cervical lymph node metastasis
BMUS	B-mode ultrasound
CDFI	Color doppler flow imaging
CEUS	Contrast-enhanced ultrasound
SWE	Shear wave elastography
SMI	Superb microvascular imaging
LASSO	Least absolute shrinkage and selection operator
SVM	Support vector machine
ROC	Receiver operating characteristic
ACC	Accuracy
AUC	Area under the curve
95% CI	95% confidence interval

Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s12880-023-01049-8>.

Supplementary Material 1

Supplementary Material 2

Acknowledgements

We appreciate the assistance provided by Fuying Zeng during the preparation of this manuscript.

Author contributions

YG, WH and FW – conceive the idea for the work. WZ and HXZ – design the analysis. FW and YKZ – analyze the data. FW - wrote the manuscript. YG and YKZ – interpret data and review the manuscript. WH – provide overall supervision for the work. All authors have approved the submitted version.

Funding

This work was supported by the National Natural Science Foundation of China (grant no. 81901744) and Natural Science Foundation of Beijing (grant no. 7204255).

Data availability

The datasets generated and/or analyzed during the current study are not publicly available due to potential compromise of patient's privacy but are available from the corresponding author on reasonable request.

Table 6 The predictive accuracy of gETE and mETE group in different models

Models	gETE		mETE	
	Cross-validation	Test set	Cross-validation	Test set
Clinical model	0/2 (0%)	1/2 (50%)	4/31 (12.9%)	0/10 (0%)
Radiomic model	2/2 (100%)	2/2 (100%)	17/31 (54.8%)	5/10 (50%)
Multimodal model	2/2 (100%)	1/2 (50%)	16/31 (51.6%)	5/10 (50%)

Number of correct classification/ total number of samples in the subgroup (subgroup accuracy, %)

Declarations

Competing interests

The authors declare no competing interests.

Ethics approval and consent to participate

The studies involving human participants were reviewed and approved by Beijing Tian Tan Hospital Ethics Committee (ethics approval number KY 2020-007-02) in accordance with the Declaration of Helsinki. Authors confirm that all methods were in accordance with the relevant guidelines and regulations. The patients/participants provided their written informed consent to participate in this study.

Consent for publication

Not applicable.

Author details

¹Department of Ultrasound, Beijing Tiantan Hospital, Capital Medical University, No. 119 West Road of South 4th Ring Road, Fengtai District, 100160 Beijing, China

Received: 13 October 2022 / Accepted: 16 June 2023

Published online: 20 July 2023

References

1. Turk AT, Asa SL, Baloch ZW, Faquin WC, Fellegara G, Ghossein RA, Giordano TJ, LiVolsi VA, Lloyd R, Mete O, et al. Interobserver Variability in the histopathologic Assessment of Extrathyroidal Extension of Well differentiated thyroid carcinoma supports the New American Joint Committee on Cancer Eighth Edition Criteria for Tumor Staging. *Thyroid*. 2019;29(5):619–24.
2. Nishino M, Jacob J. Invasion in thyroid cancer: controversies and best practices. *Semin Diagn Pathol*. 2020;37(5):219–27.
3. Perrier ND, Brierley JD, Tuttle RM. Differentiated and anaplastic thyroid carcinoma: major changes in the American Joint Committee on Cancer eighth edition cancer staging manual. *CA Cancer J Clin*. 2018;68(1):55–63.
4. Danilovic DLS, Castroneves LA, Suemoto CK, Elias LO, Soares IC, Camargo RY, Correa FA, Hoff AO, Marui S. Is there a difference between minimal and gross extension into the strap muscles for the risk of recurrence in papillary thyroid carcinomas? *Thyroid* 2020, 30(7):1008–16.
5. Bortz MD, Kuchta K, Winchester DJ, Prinz RA, Moo-Young TA. Extrathyroidal extension predicts negative clinical outcomes in papillary thyroid cancer. *Surgery*. 2021;169(1):2–6.
6. Haugen BR, Alexander EK, Bible KC, Doherty GM, Mandel SJ, Nikiforov YE, Pacini F, Randolph GW, Sawka AM, Schlumberger M, 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*. 2016;26(1):1–133.
7. Lamartina L, Bidault S, Hadoux J, Guerlain J, Girard E, Breuskin I, Attard M, Suciù V, Baudin E, Al Ghuzlan A, et al. Can preoperative ultrasound predict extrathyroidal extension of differentiated thyroid cancer? *Eur J Endocrinol*. 2021;185(1):13–22.
8. Cao Y, Zhong X, Diao W, Mu J, Cheng Y, Jia Z. Radiomics in differentiated thyroid Cancer and nodules: explorations, application, and Limitations. *Cancers (Basel)* 2021, 13(10).
9. Wang X, Agyekum EA, Ren Y, Zhang J, Zhang Q, Sun H, Zhang G, Xu F, Bo X, Lv W, et al. A Radiomic Nomogram for the Ultrasound-Based evaluation of Extrathyroidal Extension in Papillary thyroid carcinoma. *Front Oncol*. 2021;11:625646.
10. Li XL, Lu F, Zhu AQ, Du D, Zhang YF, Guo LH, Sun LP, Xu HX. Multimodal Ultrasound imaging in breast imaging-reporting and data system 4 breast lesions: a prediction model for Malignancy. *Ultrasound Med Biol*. 2020;46(12):3188–99.
11. Yao Z, Dong Y, Wu G, Zhang Q, Yang D, Yu JH, Wang WP. Preoperative diagnosis and prediction of hepatocellular carcinoma: Radiomics analysis based on multi-modal ultrasound images. *BMC Cancer*. 2018;18(1):1089.
12. Pei S, Cong S, Zhang B, Liang C, Zhang L, Liu J, Guo Y, Zhang S. Diagnostic value of multimodal ultrasound imaging in differentiating benign and malignant TI-RADS category 4 nodules. *Int J Clin Oncol*. 2019;24(6):632–9.
13. Jin ZQ, Lin MY, Hu WH, Li WY, Bai SJ. Gray-scale ultrasonography combined with elastography imaging for the evaluation of papillary thyroid microcarcinoma: as a prognostic clinicopathology factor. *Ultrasound Med Biol*. 2014;40(8):1769–77.
14. Liu Y, Liu H, Qian CL, Lin MS, Li FH. Utility of quantitative contrast-enhanced ultrasound for the prediction of extracapsular extension in papillary thyroid carcinoma. *Sci Rep*. 2017;7(1):1472.
15. Wu Y, Rao K, Liu J, Han C, Gong L, Chong Y, Liu Z, Xu X. Machine learning algorithms for the prediction of Central Lymph Node Metastasis in patients with papillary thyroid Cancer. *Front Endocrinol (Lausanne)*. 2020;11:577537.
16. Zhang Y, Zhang X, Li J, Cai Q, Qiao Z, Luo YK. Contrast-enhanced ultrasound: a valuable modality for extracapsular extension assessment in papillary thyroid cancer. *Eur Radiol*. 2021;31(7):4568–75.
17. Szczypinski PM, Strzelecki M, Materka A, Klepaczko A. MaZda—a software package for image texture analysis. *Comput Methods Programs Biomed*. 2009;94(1):66–76.
18. Huang X, Shu J, Yan Y, Chen X, Yang C, Zhou T, Li M. Feasibility of magnetic resonance imaging-based radiomics features for preoperative prediction of extrahepatic cholangiocarcinoma stage. *Eur J Cancer (Oxford England: 1990)*. 2021;155:227–35.
19. Gu Q, Feng Z, Liang Q, Li M, Deng J, Ma M, Wang W, Liu J, Liu P, Rong P. Machine learning-based radiomics strategy for prediction of cell proliferation in non-small cell lung cancer. *Eur J Radiol*. 2019;118:32–7.
20. Li Y, Liu Y, Zhang M, Zhang G, Wang Z, Luo J. Radiomics with Attribute bagging for breast tumor classification using Multimodal Ultrasound images. *J Ultrasound Med*. 2020;39(2):361–71.
21. Mohammadi A, Mirza-Aghazadeh-Attari M, Faeghi F, Homayoun H, Abolghasemi J, Vogl TJ, Bureau NJ, Bakhshandeh M, Acharya RU, Abbasian Ardakani A. Tumor Microenvironment, Radiology, and Artificial Intelligence: should we consider Tumor Periphery? *J Ultrasound Med*. 2022;41(12):3079–90.
22. Christiansen F, Epstein EL, Smedberg E, Åkerlund M, Smith K, Epstein E. Ultrasound image analysis using deep neural networks for discriminating between benign and malignant ovarian tumors: comparison with expert subjective assessment. *Ultrasound Obstet Gynecol*. 2021;57(1):155–63.
23. Grani G, Lamartina L, Cantisani V, Maranghi M, Lucia P, Durante C. Interobserver agreement of various thyroid imaging reporting and data systems. *Endocr Connect*. 2018;7(1):1–7.
24. Hu S, Zhang H, Sun Z, Ge Y, Li J, Yu C, Deng Z, Dou W, Wang X. Preoperative assessment of extrathyroidal extension of papillary thyroid carcinomas by ultrasound and magnetic resonance imaging: a comparative study. *Radiol Med*. 2020;125(9):870–6.
25. Kamaya A, Tahvildari AM, Patel BN, Willmann JK, Jeffrey RB, Desser TS. Sonographic Detection of Extracapsular extension in papillary thyroid Cancer. *J Ultrasound Med*. 2015;34(12):2225–30.
26. Fu Z, Zhang J, Lu Y, Wang S, Mo X, He Y, Wang C, Chen H. Clinical applications of superb microvascular imaging in the superficial tissues and Organs: a systematic review. *Acad Radiol*. 2021;28(5):694–703.
27. Gillies RJ, Kinahan PE, Hricak H. Radiomics: images are more than pictures, they are data. *Radiology*. 2016;278(2):563–77.
28. Yu P, Wu X, Li J, Mao N, Zhang H, Zheng G, Han X, Dong L, Che K, Wang Q, et al. Extrathyroidal extension prediction of papillary thyroid Cancer with computed tomography based Radiomics Nomogram: a Multicenter Study. *Front Endocrinol (Lausanne)*. 2022;13:874396.
29. Xu XQ, Zhou Y, Su GY, Tao XW, Ge YQ, Si Y, Shen MP, Wu FY. Iodine maps from dual-energy CT to Predict Extrathyroidal extension and recurrence in papillary thyroid Cancer based on a Radiomics Approach. *AJNR Am J Neuroradiol*. 2022;43(5):748–55.
30. Chen B, Zhong L, Dong D, Zheng J, Fang M, Yu C, Dai Q, Zhang L, Tian J, Lu W, et al. Computed Tomography Radiomic Nomogram for Preoperative Prediction of Extrathyroidal Extension in Papillary thyroid carcinoma. *Front Oncol*. 2019;9:829.
31. Wei R, Wang H, Wang L, Hu W, Sun X, Dai Z, Zhu J, Li H, Ge Y, Song B. Radiomics based on multiparametric MRI for extrathyroidal extension feature prediction in papillary thyroid cancer. *BMC Med Imaging*. 2021;21(1):20.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.