Journal of Clinical & Translational Endocrinology 40 (2025) 100390

Contents lists available at ScienceDirect

Journal of Clinical & Translational Endocrinology

journal homepage: www.elsevier.com/locate/jcte

Original research

Prediction of BRAF and TERT status in PTCs by machine learning-based ultrasound radiomics methods: A multicenter study

Hui Shi [a,b,1], Ke Ding [c,1], Xue Ting Yang [d], Ting Fan Wu [e], Jia Yi Zheng [f], Li Fan Wang [g], Bo Yang Zhou [g], Li Ping Sun [a,b], Yi Feng Zhang [a,d], Chong Ke Zhao [g], Hui Xiong Xu [a,b,g,*]

[a] Department of Medical Ultrasound, Center of Minimally Invasive Treatment for Tumor, Shanghai Tenth People's Hospital, School of Medicine, Tongji University, China
[b] Shanghai Engineering Research Center of Ultrasound Diagnosis and Treatment, Shanghai, China
[c] Department of Ultrasound, The First Affiliated Hospital of Zhengzhou University, Zhengzhou, China
[d] Department of Ultrasound, Shanghai First People's Hospital Affiliated to Shanghai Jiaotong University School of Medicine, Shanghai, China
[e] Bayer Healthcare, Radiology, Shanghai, China
[f] Department of Pathology, Shanghai Tenth People's Hospital, Tongji University School of Medicine, Shanghai, China
[g] Department of Ultrasound, Zhongshan Hospital, Institute of Ultrasound in Medicine and Engineering, Fudan University, Shanghai, China

ARTICLE INFO

Keywords:
Papillary thyroid carcinoma
BRAF V600E
TERT promoter
Machine learning
Radiomics
Prediction

ABSTRACT

Background: Preoperative identification of genetic mutations is conducive to individualized treatment and management of papillary thyroid carcinoma (PTC) patients. *Purpose*: To investigate the predictive value of the machine learning (ML)-based ultrasound (US) radiomics approaches for BRAF V600E and TERT promoter status (individually and coexistence) in PTC.

Methods: This multicenter study retrospectively collected data of 1076 PTC patients underwent genetic testing detection for BRAF V600E and TERT promoter between March 2016 and December 2021. Radiomics features were extracted from routine grayscale ultrasound images, and gene status-related features were selected. Then these features were included to nine different ML models to predicting different mutations, and optimal models plus statistically significant clinical information were also conducted. The models underwent training and testing, and comparisons were performed.

Results: The Decision Tree-based US radiomics approach had superior prediction performance for the BRAF V600E mutation compared to the other eight ML models, with an area under the curve (AUC) of 0.767 versus 0.547–0.675 ($p < 0.05$). The US radiomics methodology employing Logistic Regression exhibited the highest accuracy in predicting TERT promoter mutations (AUC, 0.802 vs. 0.525–0.701, $p < 0.001$) and coexisting BRAF V600E and TERT promoter mutations (0.805 vs. 0.678–0.743, $p < 0.001$) within the test set. The incorporation of clinical factors enhanced predictive performances to 0.810 for BRAF V600E mutant, 0.897 for TERT promoter mutations, and 0.900 for dual mutations in PTCs.

Conclusion: The machine learning-based US radiomics methods, integrated with clinical characteristics, demonstrated effectiveness in predicting the BRAF V600E and TERT promoter mutations in PTCs.

Introduction

The incidence of thyroid cancer (TC) has markedly risen in recent decades [1]. The predominant pathological variant of TC is thyroid papillary carcinoma (PTC), accounting for roughly 90 % of cases, characterized by indolent biological behavior [2]. Nonetheless, around 5 % to 10 % of all PTCs exhibit a tendency for recurrence, resulting in diminished disease-free survival, reduced overall survival, and elevated disease-related mortality [3,4]. Accumulating evidence underscores the critical prognostic value of B-Raf proto-onco-gene serine/threonine kinase (BRAF) and V600E and/or telomerase reverse transcriptase (TERT) promoter mutations in PTC [5]. The BRAF V600E mutation initially characterized as an independent prognostic indicator for extrathyroidal extension and lymphatic metastasis [6,7], While TERT promoter

* Correspondence author at: Department of Ultrasound, Zhongshan Hospital, Institute of Ultrasound in Medicine and Engineering, Fudan University, 180 Fenglin Street, Xuhui District, Shanghai, China.
E-mail address: xuhuixiong@126.com (H.X. Xu).
[1] These two authors contributed equally.

https://doi.org/10.1016/j.jcte.2025.100390
Received 12 October 2024; Received in revised form 17 March 2025; Accepted 28 March 2025
Available online 30 March 2025

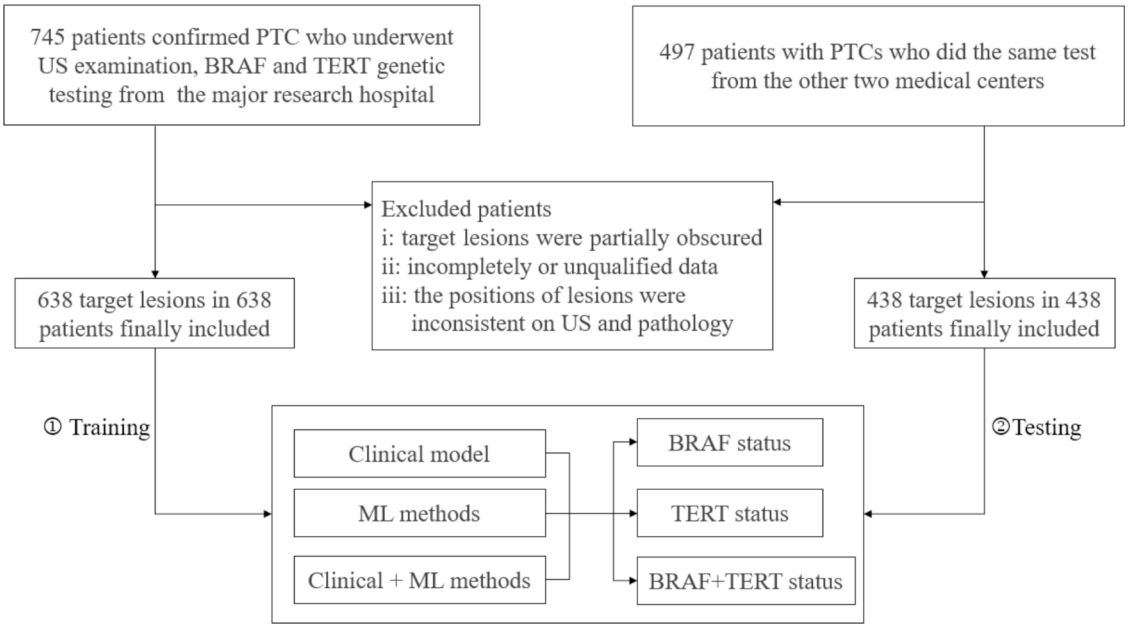

Fig. 1. Flowchart of patient enrolment and model development procedures for predicting BRAF V600E, TERT, and BRAF V600E + TERT mutations for patients with PTC in the training cohort and independent testing cohort.

mutations are an independent risk factor for disease-specific mortality and distant metastasis [8,9]. This genetic co-occurrence amplified oncogenic effects, including a greater risk of recurrence [10] and metastasis [9]. Notably, Xing et al. discovered that this genetic combination markedly elevated the mortality risk in patients with PTC [11]. Therefore, the identification of BRAF V600E and TERT promoter mutations (individually and coexistence) would promote risk-stratified management and individualized therapy for PTC patients [12,13]. Generally, the detection of BRAF V600E and TERT promoter mutations in PTC primarily relies on invasive diagnostic procedures, including preoperative fine-needle aspiration (FNA) or core needle biopsy (CNB), as well as postoperative histopathological analysis. [14,15]. Undifferentiated genetic testing may adversely affect individuals at minimal risk who do not require this intrusive procedure. Moreover, the availability of molecular testing remains limited in resource-constrained settings, where prohibitive costs of genetic analyses and the scarcity of physicians trained in performing thyroid biopsies pose significant implementation challenges.

Some radiologists have attempted to investigate the potential of utilizing ultrasound (US) images to predict gene mutation [16]. Several studies manifested that hypoechogenicity, microlobulated margins, non-parallel orientation/taller-than-wide shape, and the presence of microcalcifications are independent risk factors of visual interpretation in predicting genetic mutations [17,18]. But these reports are inconsistent in the correlation factor of US image features due to inter- and intra-observer variability [16].

ML has recently become a hot topic in the medical imaging field, which contains crucial information of medical images. Radiomics, a subset of ML, can provide quantitative features extracted through computerized algorithms beyond the visual interpretation of human eyes [19,20]. Previous studies have found that US-based radiomics features of thyroid lesions could provide some information about the risk of BRAF V600E mutation [21] and lymph node metastasis in PTC [22]. The proliferation of more ML models may enhance the precision of disease diagnosis.

The integration of radiomics derived from ultrasound imaging and various machine learning models may play a significant role in preoperatively predicting the status of BRAF V600E and TERT promoter in PTC. Thus, this study sought to evaluate the predictive efficacy of nine ML-based ultrasound radiomics models for BRAF V600E and TERT promoter status, both alone and in combination, in PTCs.

Materials and methods

Study design

This multicenter and retrospective study was approved by the institutional ethics committees of the major research hospital (approval number: SHSY-IEC-4.1/21-51/01), and informed consent was allowed to be waived. Clinical research protocol was registered at https://www.chictr.org.cn (ChiCTR2100044902).

Patients

Between March 2016 to July 2021, 638 consecutive patients with PTCs were enrolled from the primary research hospital as the training cohort. From July 2021 to December 2021, 438 cases from the additional two medical centers were designated as an independent testing cohort.

The inclusion criteria were as follows: (a) patients underwent thyroid US examination before surgery; (b) patients never received preoperative ablation or other therapies; (c) lesions were confirmed by postoperative pathology as PTCs with maximum diameters of ≥10 mm; (d) genetic testing of target lesions only checked for BRAF V600E and TERT promoter mutation status. The exclusion criteria included: (a) partial obscurity of the target lesions in the ultrasound images; (b) incomplete or untrained data, including absent ultrasound images and compromised image quality; (c) discrepancies in the nodule placements between ultrasound and pathology findings. In cases involving individuals with multifocal PTCs, the biggest lesion was designated as the target lesion. aforementioned patient selection criteria were applied to all three involved medical institutions. Finally, a total of 1076 Chinese patients (mean age, 45.4 ± 13.6 years; range, 16–87 years) with 1076 PTCs were enrolled in this study (Fig. 1). The training cohort consists of 638 patients (mean age, 45.1 ± 13.7 years; range, 18–84 years) with 638 PTCs, while an additional 438 patients (mean age, 45.9 ± 13.5 years; range, 16–87 years) with 438 PTCs constituted the testing cohort.

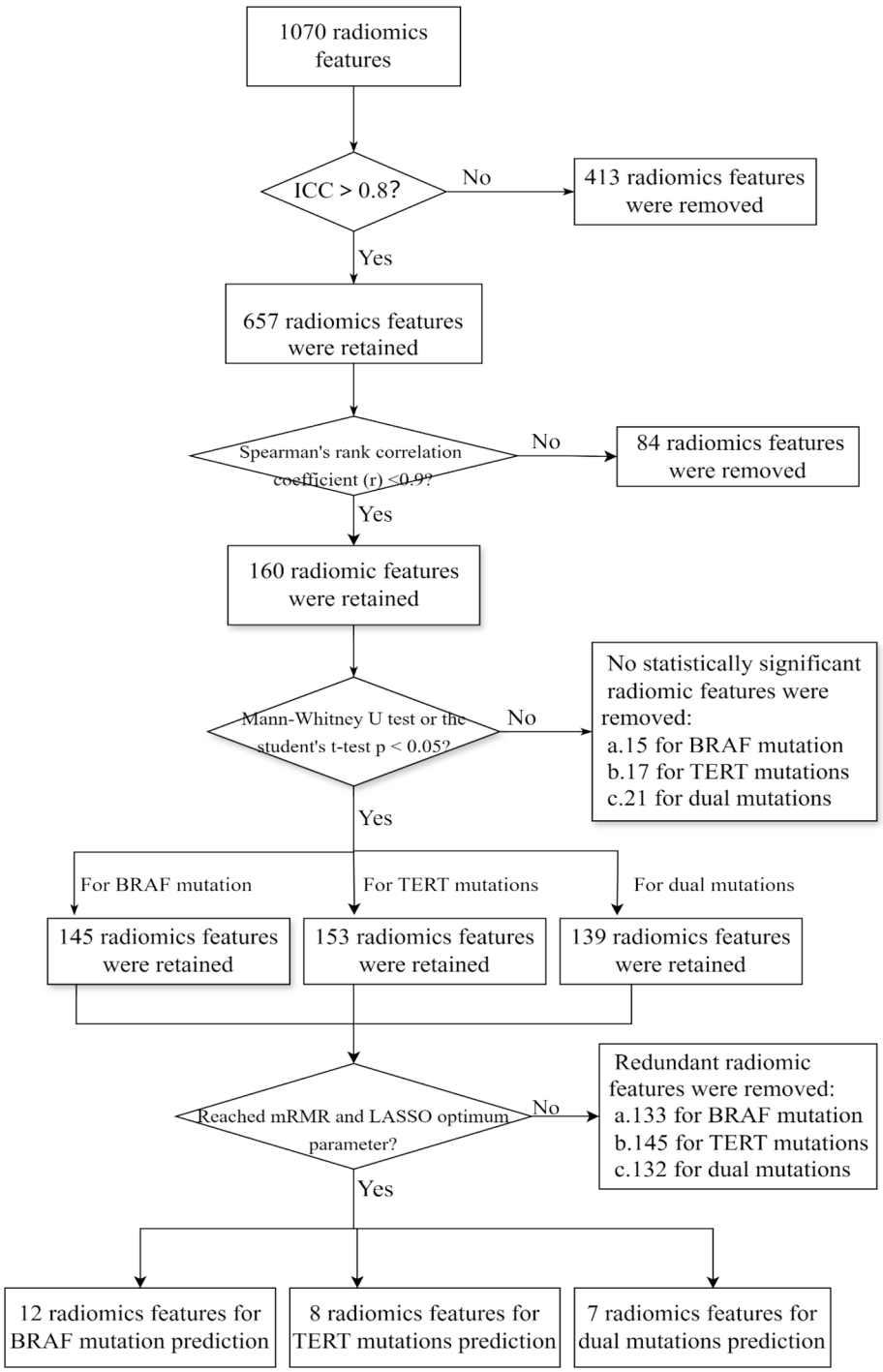

Fig. 2. The Flow chart for radiomics features dimension reduction and selection.

Data acquisition and genetic testing analysis

All US images were collected from the respective picture archiving and communication system workstation (PACS). Twelve US instruments and five experienced sonographers were joined in the images collection (Supplementary Table 1). The clinical data (age and gender) and the pathological data (the status of BRAF V600E and TERT promoter, the lesion size, number, and lymphatic metastasis) were retrieved from the respective clinical record system.

Original genomic DNA isolated from PTCs, polymerase chain reaction (PCR) amplification of primers, and standard detecting procedures of BRAF V600E and TERT promoter mutations were detailed and described previously [14,18]. According to gene mutations status, these enrolled patients were assigned into three mutation groups: (a) BRAF V600E mutation, (b) TERT promoter mutations, and (c) dual mutations group.

Radiomics features extraction and prediction models construction

Using the ITK-SNAP software (version 3.8.0, https://www.itksnap.org) to outline the region of interest (ROI), that is, to manually track the nodule boundary on the maximum long-axis cross section of the grayscale US image. The intra-operator reproducibility was

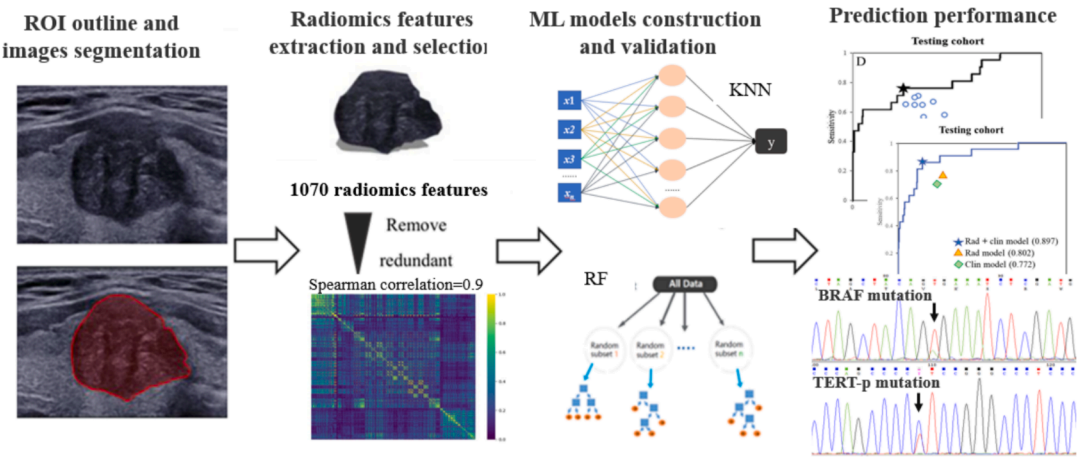

Fig. 3. Diagram depicting overall workflow of US-based radiomics ML models construction and validation.

implemented by the same radiologist with five years of experience in performing thyroid US examinations delineated the ROIs in grayscale US images of 100 randomly chosen images twice with a two-week interval. The inter-operator reproducibility was evaluated by the previous readers and another radiologist with four years of experience in performing thyroid US examinations, who independently segmented the same grayscale US images in the same time span. Then image segmentation of the retaining data was completed by the former radiologist. Two radiologists who performed the segmentations were able to access all the images captured during the diagnostic sonograms but were blinded to the gene status of the target lesions. After drawing all the ROIs of US images, those processed images were imported into the IFoundry software (Intelligence Foundry 1.2; GE Healthcare, China) to extract radiomics features. A total of 1070 radiomics features were automatically extracted from the region of interest (ROI) of each grayscale US image. (a) 18 first-order statistical intensity features; (b) 73 texture features consisted of gray-level co-occurrence matrix (GLCM), gray-level run-length matrix (GLRLM), gray-level size zone matrix (GLSZM), neighborhood Graystone difference matrix (NGTDM), gray-level dependence matrix (GLDM), and neighboring gray tone difference matrix (NGTDM); (c) contour features (n = 5), shape features (n = 13), and textural phenotype features (n = 13); (d) 948 filter-based features contained wavelet filters (n = 432), Gabor filter (n = 468), and local ternary patterns filters (n = 48).

Then radiomics features extraction, dimension reduction, and selection to identify the meaningful radiomics features are present in Fig. 2. Finally, 12 non-zero coefficient radiomics features were reserved as the most significant features for predicting BRAF V600E mutation, 8 for TERT-p mutations, and 7 for dual mutations were as meaningful impact factors brought into later predictive models (Supplementary Table 2).

Nine machine learning methods utilizing the aforementioned ultrasonic radiomics features were employed to predict three mutation groups. The classifiers employed were support vector machine (SVM), artificial neural network (ANN), logistic regression (LR), decision tree (DT), gradient boosting tree (GBT), K-nearest neighbor (KNN), naïve Bayes (NB), random forest (RF), and adaptive boosting (AB). L1 regularization was employed to mitigate model overfitting. The overall workflow of radiomics model construction and validation was depicted in Fig. 3. Univariate analysis and multivariate logistic regression analysis were used to identify clinical risk factors and construct logistic regression models for predicting target mutations. Ultimately, the optimal algorithm was employed to develop the ML-based clinical and US radiomics models for predicting target mutations. The stabilities of all prediction models were validated and compared.

Table 1
Baseline characteristics of the 1076 patients.

Characters	Total (n = 1076)	Training cohort (n = 638)	Independent validationcohort (n = 438)	P value
BRAF V600E mutation (%)	869 (80.8)	514 (80.6)	355 (81.1)	0.843
TERT-p mutations (%)	51 (4.7)	30 (4.7)	21 (4.8)	0.944
dual mutations (%)	48 (4.4)	28 (4.3)	20 (4.6)	0.754
Age (y)a	45.4 ± 13.6 (16–87)	45.1 ± 13.7 (18–84)	45.9 ± 13.5 (16–87)	0.316
Male (%)	292 (27.1)	190 (29.8)	118 (26.9)	0.311
Size (mm)a	15.2 ± 8.7 (10–95)	17.6 ± 8.2 (10–95)	17.2 ± 7.7 (10–56)	0.984
Multifocality (%)	198 (31.0)	198 (31.0)	144 (32.8)	0.688
Lymphatic metastasis (%)				
No	504 (46.8)	294 (46.1)	210 (47.9)	0.927
Central	444 (41.2)	266 (41.7)	178 (40.6)	
Lateral	78 (11.9)	78 (12.2)	50 (11.4)	
TNM stage				0.548
I + II (%)	769 (71.5)	451 (70.7)	318 (72.6)	
III + IV (%)	307 (28.5)	187 (29.3)	120 (27.4)	

Statistical analysis

Statistical analyses were performed using R software (version 4.0.3; https://www.Rproject.org) and IBM SPSS 25.0 Statistics for Windows (IBM, Armonk, NY, USA). Using independent-samples T tests or Mann-Whitney U tests (continuous variables) and chi-square tests (categorical variables) to compare the differences between gene mutations and non-mutations. Using ICC to evaluate the intra-operator and inter-operator reproducibility and the Spearman's rank correlation coefficient (r) to assess the correlations for the extracted radiomics features. The diagnostic performances of training and validation sets were evaluated according to the area under the curve (AUC) with a 95 % confidence interval (CI), sensitivity, specificity, positive predictive value (PPV), negative predictive value (NPV), and accuracy. Delong's test was used to compare the clinical model or US radiomics methods with the US radiomics plus clinical information method. The reported statistical

Table 2
The clinical characters univariate analysis of three mutation group in the training cohort.

	BRAF V600E status			TERT status			BRAF V600E + TERT status		
	Wild type (n = 124)	Mutant type (n = 514)	P Value	Wild type (n = 608)	Mutant type (n = 30)	P Value	Nobi-mutation (n = 610)	Bi-mutations (n = 28)	P Value
Age (y)[a]	40 ± 13.2	46 ± 13.6	<0.001*	44 ± 13.3	61 ± 10.6	<0.001*	40 ± 12.5	61 ± 10.7	<0.001*
Male (%)	24(29.6)	166(29.8)	0.444	171(28.1)	19(63.2)	<0.001*	23(29.1)	18(64.3)	0.001*
Tumor size (mm)[a]	17.5 ± 9.9	17.7 ± 7.8	0.849	17.2 ± 7.3	25.6 ± 17.6	0.014*	18.2 ± 7.6	23.2 ± 12.2	0.039*
Multifocality (%)	23(28.4)	175(31.4)	0.279	179(29.4)	19(63.3)	<0.001*	22(27.8)	18(64.3)	0.001*
Lymphatic metastasis (%)			0.374			0.001*			0.024*
No	39(48.2)	255(45.8)		287(47.2)	7(23.3)		38(48.1)	6(21.4)	
Central	29(35.8)	237(42.5)		253(41.6)	13(43.3)		28(35.4)	12(37.7)	
Lateral	13(16.0)	65(11.7)		68(11.2)	10(33.3)		13(16.5)	10(20.8)	
TNM stage			0.043*			<0.001*			<0.001*
I + II (%)	98 (79.0)	353 (68.6)		448 (73.7)	3 (10.0)		448 (72.1)	3 (10.7)	
III + IV (%)	26 (21.0)	161 (31.4)		160 (26.3)	27 (90.0)		162 (27.9)	25 (89.3)	

[a] Mean ± standard deviation (range).
* Statistically significant difference.

Table 3
Multivariate analysis of clinical features in predicting gene mutation.

US Features	B	SE	OR	95 % CIs	P value
BRAF V600E mutation group					
Age	0.046	0.010	1.037	1.017–1.058	<0.001*
TNM III + IV stage	0.548	0.340	1.730	1.889–3.368	0.107
TERT-p mutations group					
Age	0.074	0.025	1.077	1.025–1.132	0.003*
Male	1.451	0.489	0.234	0.090–0.611	0.003*
Tumor size	0.068	0.025	1.071	1.019–1.125	0.006*
Multifocality	0.855	0.458	2.352	0.959–5.771	0.062
Lymphatic metastasis (%)					
No					
Central	1.014	0.569	2.757	0.904–8.907	0.075
Lateral	0.957	0.652	2.604	0.725–9.352	0.142
TNM III + IV	1.743	0.824	5.712	1.136–28.712	0.034*
BRAF V600E + TERT-p mutations group					
Age	0.072	0.025	1.075	1.023–1.130	0.008*
Male	1.363	0.495	0.256	0.097–0.675	0.006*
Tumor size	0.042	0.018	1.034	1.008–1.080	0.017*
Multifocality	0.904	0.458	2.470	0.530–6.390	0.049*
Lymphatic metastasis (%)					
No					0.136
Central	0.999	0.584	2.715	0.865–8.524	0.087
Lateral	1.207	0.642	3.342	0.949–11.772	0.060
TNM III + IV	1.474	0.819	4.368	0.878–21.737	0.045*

US = Ultrasound, B = regression coefficient, SE = standard error, OR= Odds Ratios.
* Statistically significant difference.

significance levels were all two-sided, and p-values of less than 0.05 were considered statistically significant.

Results

Characteristics of PTC patients

The frequency of BRAF V600E mutation in PTCs accounted for 80.6 % (514/638), and TERT promoter mutations for 4.7 % (30/638), and coexisting BRAF V600E and TERT mutations for 4.3 % (28/638) in the training cohort. Similarly, the testing cohort included 81.1 % (355/438) BRAF V600E mutation, 4.8 % (21/438) TERT promoter mutations, and 4.6 % (20/438) coexisting BRAF V600E and TERT mutations in patients with PTCs.

The TNM III/IV stage of PTCs was 187 (29.3 %) and 120 (27.4 %) in the training and testing cohorts, respectively. Besides, no demographic differences were found between the training cohort and testing cohort in other clinicopathological characteristics such as sex, age, PTC size, multifocal lesion, and lymphatic metastasis (Table 1, all $p > 0.05$). The similar clinicopathologic characteristics of patients with PTCs justified the use of the training and testing cohorts in this study.

Through comparisons of clinicopathologic characteristics analysis among different mutation groups by univariate and multivariate logistic regression analysis, it was found that age, gender, tumor size, lymphatic metastasis, and TNM III/IV stage showed statistically significant differences in the TERT promoter mutations group (all $p < 0.05$). Additionally, the same clinicopathological features, plus multifocality, were found to be statistically significant in patients with both BRAF V600E and TERT promoter dual mutations (all $p < 0.05$). But in predicting BRAF V600E mutation, only age and TNM III/IV stage were proved as the valid predictive factors ($p < 0.001$), while other characters were not statistically significant predictors in this study (all $p > 0.05$) (Table 2, 3).

Construction and validation of the ML-based US radiomics models

The previously screened target mutation-related features were input into nine machine learning algorithms for modeling. The difference in prediction performance due to algorithm differences is shown in Fig. 4.

In comparison with 8 other ML algorithms, the DT-based US radiomics model presented the optimal performance in predicting BRAF V600E mutation (Supplementary Fig. 1). That AUCs were achieved 0.813 in the training cohort and 0.767 in the testing cohort. In the other two groups of mutation prediction, LR-based US radiomics model exhibited superior performance. This model achieved a AUCs with 0.812 in the training cohort and 0.802 in the testing cohort for predicting TERT promoter mutations, 0.839 in the training cohort and 0.805 in the testing cohort for predicting dual promoter mutations (Table 4). The sensitivity, specificity, and accuracy of nine ML algorithm were detailed in the Supplementary Table 3.

Construction and validation of the clinical models

After multivariate logistic regression analysis, independent predictors were used to calculate AUC to evaluate the predictive performance of different mutation groups. In predicting BRAF V600E mutation group, age and TNM III/IV stages as predictive factors and the clinical model obtained an AUC of 0.691 in the training cohort and 0.643 in the

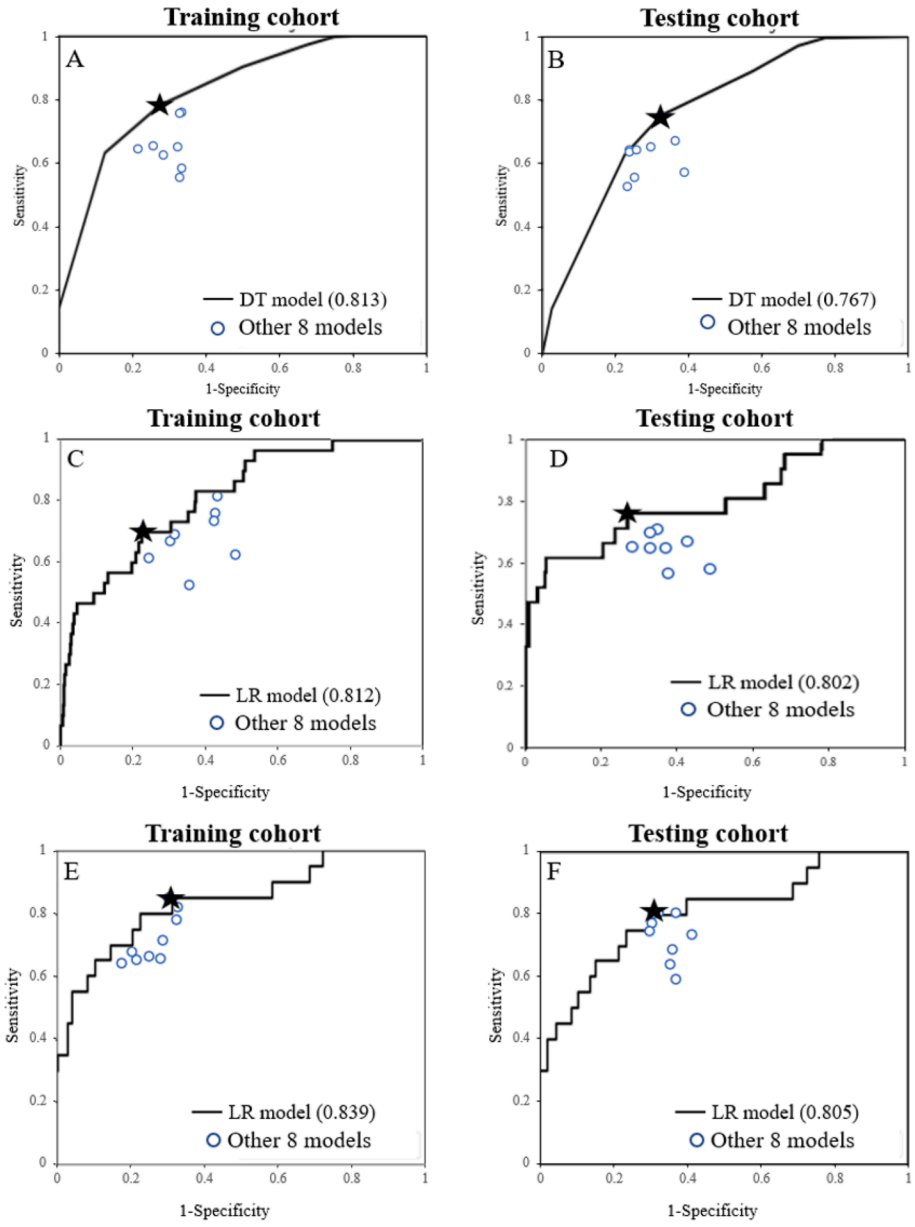

Fig. 4. Graph shows receiver operating characteristic curves (AUCs) of the optimal model (black pentagram) and other eight machine learning models (blue circle) for predicting BRAF V600E mutation (A-B), TERT promoter mutations (C-D) and the genetic duet (E-F) in training and validation cohorts.

testing cohort. In predicting TERT promoter mutation group, the clinical model with age, male, size, lymphatic metastasis, and TNM III/IV stages achieved an AUC of 0.784 in the training cohort and 0.772 in the testing cohort. In predicting coexisting dual mutations group, six clinically significant features were included in the clinical model, which showed an AUC of 0.788 in the training cohort and 0.781 in the testing cohort (Table 5).

Construction and validation of the ML-based clinical plus radiomics models

According to the previous analysis, the age and TNM III/IV stage in clinical predictors and 12 radiomics predictors were input together to predict BRAF V600E mutation, the result provided an AUC of 0.839 in the training cohort and 0.810 in the testing cohort. Likewise, using 5 significantly clinical features and 8 radiomics predictors to constructed TERT promoter mutations prediction models which got an AUC of 0.914 for the training set and 0.897 for the testing set, and using 6 clinical predictive factors and 7 radiomics predictors to establish dual mutations prediction, which obtained an AUC of 0.913 for the training cohort and 0.900 for the independent testing cohort. The average accuracy, sensitivity, specificity, PPV, and NPV were present in Table 5.

Comparison of the three prediction models in different mutation groups

There were statistically significant differences between the clinical model and the ML-based US radiomics incorporated clinical model when it came to predicting the BRAF V600E mutation (AUC: 0.691 vs. 0.810, $p < 0.001$), the TERT promoter mutations (AUC: 0.784 vs. 0.914, $p < 0.001$), and the dual mutations group (AUC: 0.788 vs. 0.913, $p < 0.001$) in the training cohort. The similar results were also reflected in the testing cohort (AUC: 0.643 vs. 0.810 for the BRAF mutation group, 0.772 vs. 0.897 for the TERT promoter mutations group, and 0.775 vs. 0.900 for the dual mutations group). Moreover, the US radiomics-incorporated clinical characteristics model also showed superiority

Table 4
Predictive performance of the nine Machine learning models based on radiomics features for the training and validation cohorts.

		SVM	ANN	LR	DT	GBT	KNN	NB	RF	AB
BRAF V600E group										
Training set	AUC	0.683	0.608	0.648	0.813	0.646	0.651	0.650	0.735	0.635
	95 % CI	0.628–0.736	0.533–0.662	0.536–0.673	0.744–0.876	0.565–0.751	0.561–0.652	0.529–0.727	0.682–0.779	0.572–0.736
Validation set	AUC	0.619	0.547	0.654	0.767	0.615	0.653	0.639	0.675	0.604
	95 % CI	0.576–0.681	0.521–0.612	0.592–0.687	0.726—–0.815	0.579–0.664	0.526–0.663	0.597–0.689	0.586–0.789	0.569–0.677
TERT-p group										
Training set	AUC	0.651	0.701	0.812	0.652	0.677	0.525	0.573	0.676	0.683
	95 % CI	0.612–0.764	0.592–0.722	0.779—0.841	0.688–0.812	0.676–0.789	0.612–0.732	0.521–0.621	0.611–0.754	0.556–0.776
Validation set	AUC	0.651	0.701	0.802	0.652	0.677	0.525	0.573	0.676	0.683
	95 % CI	0.553–0.751	0.632–0.783	0.761—0.838	0.513–0.717	0.552–0.741	0.489–0.662	0.512–0.691	0.562–0.752	0.538–0.756
BRAF V600E + TERT-p group										
Training set	AUC	0.774	0.767	0.839	0.736	0.767	0.652	0.701	0.761	0.737
	95 % CI	0.613–0.887	0.653–0.863	0.712–0.912	0.656–0.813	0.627–0.876	0.628–0.756	0.621–0.823	0.617–0.839	0.587–0.883
Validation set	AUC	0.743	0.737	0.805	0.703	0.714	0.678	0.695	0.711	0.721
	95 % CI	0.612–0.812	0.629–0.816	0.719–0.879	0.581–0.812	0.632–0.816	0.582–0.705	0.598–0.793	0.611–0.812	0.625–0.802

AUC = the area under the receiver operating characteristic curve, SVM = support vector machine, ANN = artificial neural network, LR = logistic regression, DT = decision tree, GBT = gradient boosting tree, KNN = K-nearest neighbor classification, NB = naïve bayes, RF = random forest, AB = adaptive boosting.

over the ML-based US radiomics method in training (AUC: 0.813 vs. 0.839 for the BRAF group, 0.812 vs. 0.914 for the TERT promoter group, and 0.818 vs. 0.913 for the coexisting BRAF V600E and TERT promoter group) and in the testing cohort (AUC: 0.767 vs. 0.810 for the BRAF group, 0.802 vs. 0.897 for the TERT promoter group, and 0.811 vs. 0.900 for the coexisting BRAF V600E and TERT promoter group). And the prediction performance of the three mutation groups did not differ statistically between the training and testing sets (Table 5, Fig. 5).

Discussion

In this study, we present an ML-based US radiomics approach capable of predicting the mutation status of BRAF V600E, TERT promoter, or the coexistence of BRAF V600E and TERT promoter mutations in PTC patients. The results demonstrate that ML-based US radiomics features achieved moderate predictive performance for the three mutation groups, with AUC values ranging from 0.547 to 0.767 for BRAF V600E mutation, 0.525 to 0.802 for TERT promoter mutations, and 0.695 to 0.805 for the dual mutations. Furthermore, after incorporation of clinical characteristics, the predictive performance improved from 0.767 to 0.810 for predicting BRAF V600E mutation, from 0.802 to 0.897 for predicting TERT promoter mutations, and from 0.818 to 0.900 for predicting the dual promoter mutations. These promising results could be attributed to the advantages of a standardized lesion outline, a large-scale sample size from a multicenter database, and the utilization of nine ML algorithms.

In the current cohort, the prevalence of BRAF mutation (80.8 %) was higher and the TERT promoter mutations (4.7 %) was lower compared with western countries [23,24]. These findings could be explained by population-specific variations or methodological differences: (i) In Chinese cohorts, BRAF V600E rates in PTC have been reported as high as 76–86 % [25–27], potentially indicating regional genetic. (ii) The use of Next-Generation Sequencing enhances sensitivity compared to conventional Sanger sequencing approaches [28]. In thyroid cancer, genetic mutations have played a useful role in determining the appropriate initial treatment, risk-adapted management, and postoperative follow-up [3,4,29]. For instance, patients with 1–4 cm intrathyroidal PTC without TERT promoter mutations typically only require lateral lobectomy for favorable outcomes [30]. The dual mutations of PTC have higher recurrence and mortality [31], which affects the postoperative treatment and follow-up time of patients [30]. Moreover, the dual mutations can further strengthen the risk assessment based on the ATA or TNM staging system [32]. Our study demonstrates a strong predictive performance for dual mutations (AUC: 0.900) and TERT promoter mutations (AUC: 0.897), may providing valuable insights for clinicians in risk stratification, treatment decisions, and patient management, particularly in environments where genetic testing may not be readily available.

Previous studies have reported that some grayscale US features are associated with molecular mutations in PTC [16,21,33]. Hypoechogenicity, spiculated/microlobulated margins, microcalcifications, non-parallel orientation, and mixed-type non-increased vascularity have been shown to be predictive factors of the BRAF V600E mutation [34]. Tumors with typically multifocal, microlobulated margin, microcalcifications, taller-than-wide shape, and capsule contact or involvement are frequently associated with TERT promoter mutations [35]. Furthermore, Watutantrige et al. reported that the coexistence of BRAF V600E and TERT mutations was more likely to exhibit bilateral, multifocal, angio-invasive, extrathyroid tissue involvement and intratumoral necrosis [36]. These findings suggested a strong link between molecular mutations and US imaging features. However, it is difficult to identify molecular mutations by visual identification of ultrasonic features in actual clinical work.

In recent studies, artificial intelligence (AI) methods based on US images have been utilized to assess the genetic mutations in PTC. Kwon et al. and Yoon et al. attempted to predict the BRAF V600E mutation utilizing US-based radiomics and machine learning techniques, attaining modest performance with AUC values between 0.516 and 0.742 [21,37]. Similarly, our research also yielded unsatisfactory results despite applying more machine learning algorithms (AUC, 0.547–0.767). This may result from the predictive performance being undermined by data biases that fit more closely with real-world events or from a weak connection between the radiomic characteristics and BRAF mutations. The optimal predictive outcomes were attained for TERT promoter mutations (AUC: 0.802) and dual mutations (AUC: 0.805). To date, studies using radiomics features based US images to predict TERT promoter mutations and dual mutations have not been found. These results indicate that ML-based US radiomics approaches hold promise for predicting genetic mutations of PTCs.

Clinical features are also considered as factors associated with the

Table 5
The diagnostic performance of three different methods for predicting three groups mutation status.

Methods	Cohorts	AUC	Sensitivity (%)	Specificity (%)	PPV (%)	NPV (%)	P_1	P_2
BRAF V600E group								
Clinical model	Training set	0.691 (0.653–0.726)	57.2 (52.8–61.5)	75.0 (66.4–82.3)	90.5 (86.7–93.4)	29.7 (24.7–35.1)	/	<0.001*
	Testing set	0.643 (0.596–0.688)	56.3 (51.0–61.6)	67.5 (56.3–77.4)	88.1 (83.2–92.0)	26.5 (20.7–33.0)	0.128	<0.001*
US radiomics method	Training set	0.813 (0.744–0.876)	78.0 (51.6–61.6)	72.5 (65.3–84.6)	91.0 (86.4–94.4)	44.4 (37.3–55.8)	/	0.032#
	Testing set	0.767 (0.726–0.815)	75.1 (50.2–60.7)	67.5 (64.0–83.6)	90.4 (85.7–93.9)	38.3 (30.1–45.7)	0.099	0.025#
US radiomics + Clinical information method	Training set	0.839 (0.775–0.913)	81.4 (77.0–85.3)	72.4 (61.8–81.5)	92.4 (88.9–95.1)	45.5 (39.6–58.4)	/	/
	Testing set	0.810 (0.768–0.857)	76.9 (72.2–81.2)	72.3 (61.6–81.5)	92.0 (88.4–94.8)	43.2 (35.0–51.6)	0.309	/
TERT-p group								
Clinical model	Training set	0.784 (0.750–0.815)	63.3 (43.9–80.1)	82.7 (79.5–85.7)	15.3 (9.5–22.9)	97.9 (96.2–98.9)	/	<0.001*
	Testing set	0.772 (0.729–0.810)	76.2 (52.8–91.8)	73.1 (68.6–77.3)	12.5 (7.3–19.5)	98.4 (96.3–99.5)	0.913	<0.001*
US radiomics method	Training set	0.812 (0.779–0.841)	70.0 (50.6–85.3)	77.5 (73.9–80.7)	13.3 (8.4–19.6)	98.1 (96.5–99.1)	/	<0.001#
	Testing set	0.802 (0.761–0.838)	71.4 (47.8–88.7)	76.5 (72.1–80.5)	13.3 (7.6–20.9)	98.2 (96.0–99.3)	0.317	<0.001#
US radiomics + Clinical information method	Training set	0.914 (0.890–0.935)	83.33 (65.3–94.4)	87.34 (84.4–89.9)	24.5 (16.5–34.0)	99.1 (97.8–99.7)	/	/
	Testing set	0.897 (0.865–0.924)	85.71 (63.7–97.0)	85.61 (81.9–88.8)	23.1 (14.3–34.0)	99.2 (97.6–99.8)	0.098	/
BRAF V600E + TERT-p group								
Clinical model	Training set	0.788 (0.704–0.873)	64.3 (44.1–81.4)	84.6 (81.5–87.4)	16.1 (9.8–24.2)	98.1 (96.5–99.1)	/	<0.001*
	Testing set	0.775 (0.733–0.813)	57.1 (34.0–78.2)	81.1 (77.0–84.7)	13.2 (7.0–21.9)	97.4 (95.1–98.8)	0.896	<0.001*
US radiomics method	Training set	0.839 (0.721–0.911)	76.1 (52.8–91.8)	78.4 (74.2–82.3)	15.1 (8.9–23.4)	98.5 (96.5–99.5)	/	<0.001#
	Testing set	0.805 (0.717–0.900)	79.6 (76.3–82.8)	77.8 (74.4–81.1)	13.4 (5.1–26.8)	98.9 (97.4–99.6)	0.425	<0.001#
US radiomics + Clinical information method	Training set	0.913 (0.857–0.970)	85.7 (67.3–96.0)	83.44 (80.3–86.3)	19.2 (12.7–27.2)	99.2 (98.0–99.8)	/	/
	Testing set	0.900 (0.868–0.927)	85.7 (63.7–97.0)	81.3 (77.2–84.9)	18.8 (11.5–28.0)	99.1 (97.5–99.8)	0.105	/

US = Ultrasound, AUC = the area under the receiver operating characteristic curve,
P_1 = Comparison the AUC between training group and independent testing group.
P_2 = The training and testing cohorts of the clinical model or ultrasound radiomics model were respectively compared with the corresponding cohorts of the integrated US radiomics + Clinica linformation method.
* Statistically significant difference between the clinical model and US radiomics + Clinical information methods.
# Statistically significant difference in US radiomics methods compared with US radiomics + Clinical information methods.

prediction of molecular mutations in PTC [5,38,39]. Older age, male patients, larger tumor size, extrathyroidal extension, multifocality, lymph node metastasis, and TNM III/IV stage are associated with TERT promoter mutations [40,41]. The age at diagnosis, advanced TNM stage, and extrathyroidal extension are somewhat associated with the BRAF V600E mutation. Dual mutations are substantially correlated with age, female, advanced TNM stage, extrathyroidal extension, lymph node metastases, and distant metastasis [5]. This work employed mutation-related clinical data to develop predictive models, which yielded sub-optimal performance (AUC: 0.663 for BRAF V600E mutant, 0.784 for TERT promoter mutations, and 0.788 for dual mutations). By integrating statistically significant clinical characteristics and radiomic features to develop machine learning-based clinical and ultrasound radiomic models, the predictive capability across three mutation groups has been markedly enhanced (AUC: 0.810 for the BRAF V600E mutation group, 0.897 for the TERT promoter mutations group, and 0.900 for the dual mutations group). The results indicate that although clinical features provide limited predictive capability for genetic mutations, their integration with imaging radiomics features can substantially enhance prediction accuracy.

There are several limitations in this study. Firstly, this study did not include non-PTC controls, so the potential scope of application of this model would be only in patients with confirmed PTC and not in patients with non-PTC thyroid cancers or benign thyroid nodules. Secondly, this retrospective study may have selection bias, as it only included PTC patients who underwent simultaneous testing for both BRAF V600E and TERT promoter status. Thirdly, the use of a limited number of US equipment for preoperative US examinations in the main cohort may have impacted the texture information of the US images and the generalizability of our findings. Therefore, future research may benefit from the inclusion of images from various ultrasonic machines. Lastly, it is essential to consider the potential bias in the data due to variations in the prevalence of BRAF V600E and TERT promoter mutations across different countries and ethnic groups. Larger prospective studies are needed to validate our results and assess their potential for clinical application in diverse populations in the future.

In conclusion, our study demonstrates the potential of ML-based US radiomics plus clinical characteristics as a non-invasive diagnostic method to predict genetic mutations, particularly for coexisting BRAF V600E and TERT mutations in PTCs. This integrated method may provide valuable insights for preoperative screening of patients necessitating genetic testing, personalized therapies, and risk-stratified individualized management in patients with PTCs. In addition, this approach may be less time-consuming and more cost-effective than

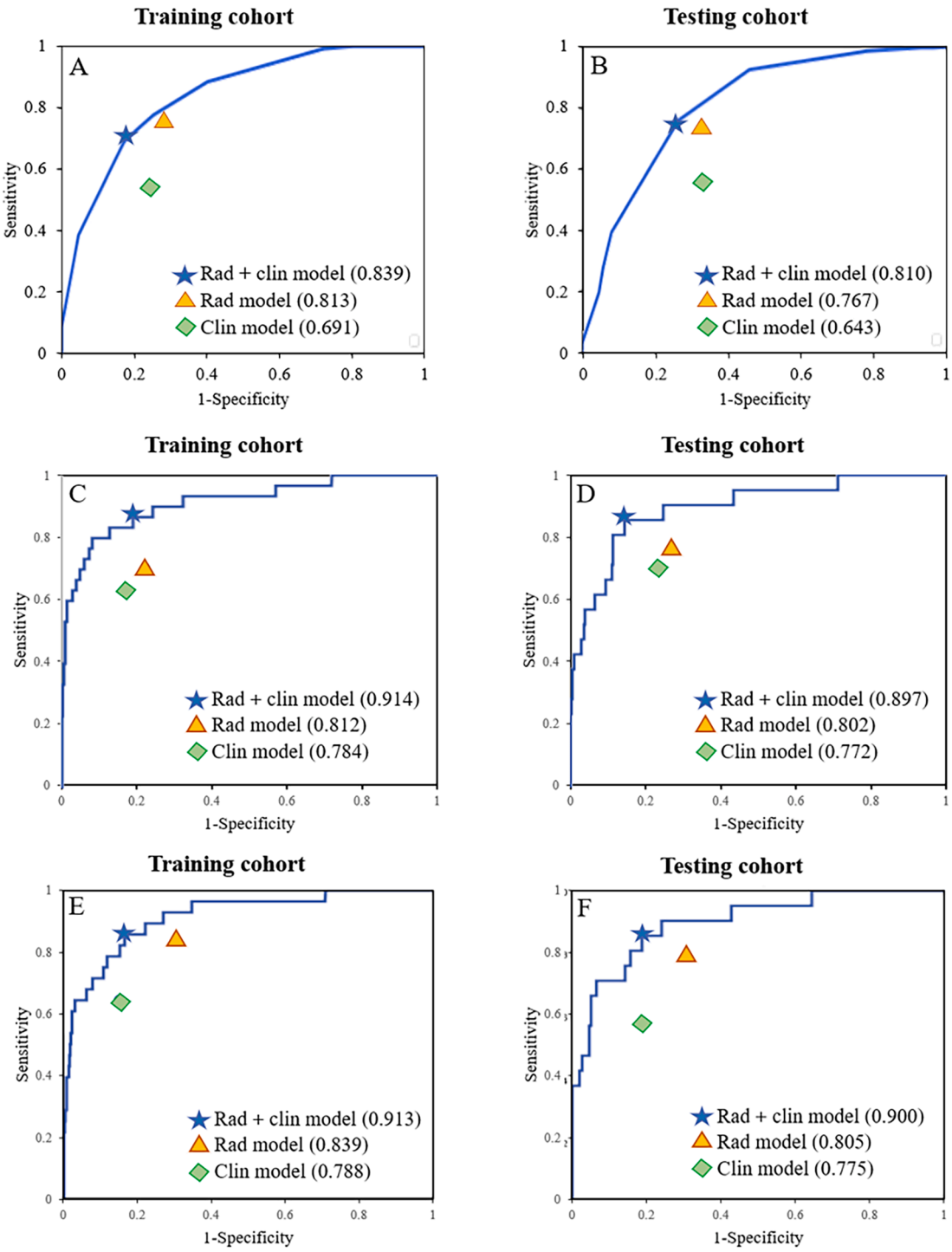

Fig. 5. Comparison of ROC curves between the Clinical model (rhombus), the machine learning model (triangle), and the combined model (pentagram) for predicting BRAF V600E mutation (A-B), TERT promoter mutations (C-D) and the genetic duet (E-F) in training and validation cohorts.

routine genetic testing as it leverages existing imaging data and could be performed without additional invasive procedures. However, integrating these US radiomics into PACS software faces challenges now. Image preprocessing, feature extraction and selection, compatibility with the PACS system, as well as the performance and stability of the PACS system after the integration of image auto-recognition algorithms, are all issues that need to be addressed. We believe that with later efforts, the incorporation of ML-based US radiomics algorithms into PACS software could facilitate the prediction of BRAF and/or TERT mutations, thereby improving disease diagnosis and treatment in the future.

CRediT authorship contribution statement

Hui Shi: Writing – original draft. **Ke Ding:** Validation, Data curation. **Xue Ting Yang:** Data curation. **Ting Fan Wu:** Software. **Jia Yi Zheng:** Resources. **Li Fan Wang:** Investigation. **Bo Yang Zhou:** Methodology. **Li Ping Sun:** Project administration. **Yi Feng Zhang:** Supervision. **Chong Ke Zhao:** Conceptualization. **Hui Xiong Xu:** Writing – original draft.

Funding

This study was supported by the National Natural Science Foundation of China (Grant 82202174), the Science and Technology Commission of Shanghai Municipality (Grants 18441905500, and 19DZ2251100), Shanghai Municipal Health Commission (Grants 2019LJ21 and SHSLCZDZK03502), Shanghai Science and Technology Innovation Action Plan (21Y11911200), and Fundamental Research Funds for the Central Universities (ZD11202151), Scientific Research and Development Fund of Zhongshan Hospital of Fudan University (Grant 2022ZSQD07).

Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: [Hui Xiong Xu reports financial support was provided by Shanghai Tenth People's Hospital. Hui Xiong Xu reports a relationship with Shanghai Tenth People's Hospital that includes: employment. No has patent no pending to no. The authors do not hold any positions with the journal and therefore have no other conflicts of interest If there are other authors, they declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper].

Acknowledgments

The authors wish to thank the patients enrolled in our study of Thyroid papillary carcinoma involved in genetic testing

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jcte.2025.100390.

References

[1] Bao WQ, Zi H, Yuan QQ, Li LY, Deng T. Global burden of thyroid cancer and its attributable risk factors in 204 countries and territories from 1990 to 2019. Thorac Cancer 2021;12(18):2494–503.
[2] Li M, Brito JP, Vaccarella S. Long-term declines of thyroid cancer mortality: An international age-period-cohort analysis. Thyroid 2020;30(6):838–46.
[3] Haugen BR, Alexander EK, Bible KC, Doherty GM, Mandel SJ, Nikiforov YE, et al. 2015 American thyroid association management guidelines for adult patients with thyroid nodules and differentiated thyroid cancer: The american thyroid association guidelines task force on Thyroid nodules and differentiated thyroid cancer. Thyroid 2016;26(1):1–133.
[4] Patel KN, Yip L, Lubitz CC, Grubbs EG, Miller BS, Shen W, et al. The American association of endocrine surgeons guidelines for the definitive surgical management of thyroid disease in adults. Ann Surg 2020;271(3):e21–93.
[5] Moon S, Song YS, Kim YA, Lim JA, Cho SW, Moon JH, et al. Effects of coexistent BRAF(V600E) and TERT promoter mutations on poor clinical outcomes in papillary thyroid cancer: A meta-analysis. Thyroid 2017;27(5):651–60.
[6] Kim TH, Park YJ, Lim JA, Ahn HY, Lee EK, Lee YJ, et al. The association of the BRAF(V600E) mutation with prognostic factors and poor clinical outcome in papillary thyroid cancer: a meta-analysis. Cancer 2012;118(7):1764–73.
[7] Xing M, Alzahrani AS, Carson KA, Viola D, Elisei R, Bendlova B, et al. Association between BRAF V600E mutation and mortality in patients with papillary thyroid cancer. JAMA 2013;309(14):1493–501.
[8] Melo M, da Rocha AG, Vinagre J, Batista R, Peixoto J, Tavares C, et al. TERT promoter mutations are a major indicator of poor outcome in differentiated thyroid carcinomas. J Clin Endocrinol Metab 2014;99(5):E754–65.
[9] Melo M, Gaspar da Rocha A, Batista R, Vinagre J, Martins MJ, Costa G, et al. TERT, BRAF, and NRAS in Primary thyroid cancer and metastatic disease. J Clin Endocrinol Metab 2017;102(6):1898–907.
[10] Xing M, Liu R, Liu X, Murugan AK, Zhu G, Zeiger MA, et al. BRAF V600E and TERT promoter mutations cooperatively identify the most aggressive papillary thyroid cancer with highest recurrence. J Clin Oncol 2014;32(25):2718–26.
[11] Liu R, Bishop J, Zhu G, Zhang T, Ladenson PW, Xing M. Mortality risk stratification by combining BRAF V600E and TERT promoter mutations in papillary thyroid cancer: Genetic duet of BRAF and TERT promoter mutations in thyroid cancer mortality. JAMA Oncol 2017;3(2):202–8.
[12] Raue F, Frank-Raue K. Thyroid cancer: Risk-stratified management and individualized therapy. Clin Cancer Res 2016;22(20):5012–21.
[13] Ngeow J, Eng C. TERT and BRAF in thyroid cancer: teaming up for trouble. J Clin Oncol 2014;32(25):2683–4.
[14] Zhao CK, Zheng JY, Sun LP, Xu RY, Wei Q, Xu HX. BRAF(V600E) mutation analysis in fine-needle aspiration cytology specimens for diagnosis of thyroid nodules: The influence of false-positive and false-negative results. Cancer Med 2019;8(12):5577–89.
[15] Liu R, Xing M. Diagnostic and prognostic TERT promoter mutations in thyroid fine-needle aspiration biopsy. Endocr Relat Cancer 2014;21(5):825–30.
[16] Lewiński A, Adamczewski Z, Zygmunt A, Markuszewski L, Karbownik-Lewińska M, Stasiak M. Correlations between molecular landscape and sonographic image of different variants of papillary thyroid carcinoma. J Clin Med 2019;8(11).
[17] Kabaker AS, Tublin ME, Nikiforov YE, Armstrong MJ, Hodak SP, Stang MT, et al. Suspicious ultrasound characteristics predict BRAF V600E-positive papillary thyroid carcinoma. Thyroid 2012;22(6):585–9.
[18] Shi H, Guo LH, Zhang YF, Fu HJ, Zheng JY, Wang HX, et al. Suspicious ultrasound and clinicopathological features of papillary thyroid carcinoma predict the status of TERT promoter. Endocrine 2020;68(2):349–57.
[19] Shan J, Alam SK, Garra B, Zhang Y, Ahmed T. Computer-aided diagnosis for breast ultrasound using computerized BI-RADS features and machine learning methods. Ultrasound Med Biol 2016;42(4):980–8.
[20] Choi YJ, Baek JH, Park HS, Shim WH, Kim TY, Shong YK, et al. Computer-aided diagnosis system using artificial intelligence for the diagnosis and characterization of thyroid nodules on ultrasound: initial clinical assessment. Thyroid 2017;27(4):546–52.
[21] Kwon MR, Shin JH, Park H, Cho H, Hahn SY, Park KW. Radiomics study of thyroid ultrasound for predicting BRAF Mutation in papillary thyroid carcinoma: Preliminary results. AJNR Am J Neuroradiol 2020;41(4):700–5.
[22] Yu J, Deng Y, Liu T, Zhou J, Jia X, Xiao T, et al. Lymph node metastasis prediction of papillary thyroid carcinoma based on transfer learning radiomics. Nat Commun 2020;11(1):4807.
[23] Kimura ET, Nikiforova MN, Zhu Z, Knauf JA, Nikiforov YE, Fagin JA. High prevalence of BRAF mutations in thyroid cancer: genetic evidence for constitutive activation of the RET/PTC-RAS-BRAF signaling pathway in papillary thyroid carcinoma. Cancer Res 2003;63(7):1454–7.
[24] Soares P, Trovisco V, Rocha AS, Lima J, Castro P, Preto A, et al. BRAF mutations and RET/PTC rearrangements are alternative events in the etiopathogenesis of PTC. Oncogene 2003;22(29):4578–80.
[25] Cong R, Ouyang H, Zhou D, Li X, Xia F. BRAF V600E mutation in thyroid carcinoma: a large-scale study in Han Chinese population. World J Surg Oncol 2024;22(1):259. https://doi.org/10.1186/s12957-024-03539-7.
[26] Wang Z, Tang P, Hua S, Gao J, Zhang B, Wan H, et al. Genetic and clinicopathologic characteristics of papillary thyroid carcinoma in the chinese population: High BRAF mutation allele frequency, multiple driver gene mutations, and RET fusion may indicate more advanced TN stage. Onco Targets Ther 2022;15:147–57.
[27] Rashid FA, Munkhdelger J, Fukuoka J, Bychkov A. Prevalence of BRAF(V600E) mutation in Asian series of papillary thyroid carcinoma-a contemporary systematic review. Gland Surg 2020;9(5):1878–900.
[28] Pant S, Weiner R, Marton MJ. Navigating the rapids: the development of regulated next-generation sequencing-based clinical trial assays and companion diagnostics. Front. Oncol 2014;4:78.
[29] Xing M, Haugen BR, Schlumberger M. Progress in molecular-based management of differentiated thyroid cancer. Lancet 2013;381(9871):1058–69.
[30] Ebina A, Togashi Y, Baba S, Sato Y, Sakata S, Ishikawa M, et al. TERT promoter mutation and extent of thyroidectomy in patients with 1-4 cm intrathyroidal papillary carcinoma. Cancers (Basel) 2020;12(8).
[31] Bullock M, Ren Y, O'Neill C, Gill A, Aniss A, Sywak M, et al. TERT promoter mutations are a major indicator of recurrence and death due to papillary thyroid carcinomas. Clin Endocrinol (Oxf) 2016;85(2):283–90.
[32] Song YS, Lim JA, Choi H, Won JK, Moon JH, Cho SW, et al. Prognostic effects of TERT promoter mutations are enhanced by coexistence with BRAF or RAS mutations and strengthen the risk prediction by the ATA or TNM staging system in differentiated thyroid cancer patients. Cancer 2016;122(9):1370–9.
[33] Hahn SY, Kim TH, Ki CS, Kim SW, Ahn S, Shin JH, et al. Ultrasound and clinicopathological features of papillary thyroid carcinomas with BRAF and TERT promoter mutations. Oncotarget 2017;8(65):108946–57.
[34] Baek H, Kim D, Shin G, Heo Y, Baek J, Lee Y, et al. Ultrasonographic Features of Papillary Thyroid Carcinomas According to Their Subtypes 2018;9:223.
[35] Kim T, Ki C, Hahn S, Oh Y, Jang H, Kim S, et al. Ultrasonographic prediction of highly aggressive telomerase reverse transcriptase (TERT) promoter-mutated papillary thyroid cancer 2017;57(2):234–40.
[36] Watutantrige-Fernando S, Vianello F, Barollo S, Bertazza L, Galuppini F, Cavedon E, et al. The Hobnail Variant of Papillary Thyroid Carcinoma: Clinical/Molecular Characteristics of a Large Monocentric Series and Comparison with Conventional Histotypes 2018;28(1):96–103.
[37] Yoon JH, Han K, Lee E, Lee J, Kim EK, Moon HJ, et al. Radiomics in predicting mutation status for thyroid cancer: A preliminary study using radiomics features for predicting BRAFV600E mutations in papillary thyroid carcinoma. PLoS One 2020;15(2):e0228968.
[38] Huang M, Yan C, Xiao J, Wang T, Ling R. Relevance and clinicopathologic relationship of BRAF V600E, TERT and NRAS mutations for papillary thyroid carcinoma patients in Northwest China. Diagn Pathol 2019;14(1):74. https://doi.org/10.1186/s13000-019-0849-6.

[39] Yang J, Gong Y, Yan S, Chen H, Qin S, Gong R. Association between TERT promoter mutations and clinical behaviors in differentiated thyroid carcinoma: a systematic review and meta-analysis. Endocrine 2020;67(1):44–57.
[40] Yin DT, Yu K, Lu RQ, Li X, Xu J, Lei M, et al. Clinicopathological significance of TERT promoter mutation in papillary thyroid carcinomas: a systematic review and meta-analysis. Clin Endocrinol (Oxf) 2016;85(2):299–305.
[41] Nasirden A, Saito T, Fukumura Y, Hara K, Akaike K, Kurisaki-Arakawa A, et al. In Japanese patients with papillary thyroid carcinoma, TERT promoter mutation is associated with poor prognosis, in contrast to BRAF (V600E) mutation. Virchows Arch 2016;469(6):687–96.