Original Article

An economically efficient strategy for diagnosing atypia of undetermined significance or follicular lesion of undetermined significance thyroid nodules with ultrasound-based risk stratification systems and *BRAF*^{V600E} testing

Xiao Ma^, Yuanjing Huang, Runhui Shen, Yiyang Cui, Yaning Kang, Jing Yu, Kefei Cui

Department of Ultrasound, The First Affiliated Hospital of Zhengzhou University, Zhengzhou, China

Contributions: (I) Conception and design: X Ma, R Shen, K Cui; (II) Administrative support: K Cui, J Yu; (III) Provision of study materials or patients: K Cui, Y Huang, J Yu; (IV) Collection and assembly of data: X Ma, Y Huang, Y Cui; (V) Data analysis and interpretation: X Ma, Y Cui, R Shen, Y Kang; (VI) Manuscript writing: All authors; (VII) Final approval of manuscript: All authors.

Correspondence to: Jing Yu, MD; Kefei Cui, MD. Department of Ultrasound, The First Affiliated Hospital of Zhengzhou University, No. 1 Jianshe East Road, Erqi District, Zhengzhou 450052, China. Email: yujing822@163.com; cuikefei2010@126.com.

Background: The management of thyroid nodules classified as atypia of undetermined significance or follicular lesion of undetermined significance (AUS/FLUS) has been a subject of ongoing debate. Therefore, the aim of this study was to investigate a cost-effective approach for managing these nodules by combining *BRAF*^{V600E} mutation analysis with the guidelines provided by the American Thyroid Association (ATA) or the American College of Radiology (ACR) Thyroid Imaging Reporting and Data System (TIRADS).

Methods: This study included 762 AUS/FLUS nodules in 551 patients with a postoperative pathology. A preoperative *BRAF*^{V600E} gene test and an evaluation using the ATA guidelines and ACR-TIRADS were performed. Two combined diagnostic approaches were employed: In method 1, all nodules underwent *BRAF*^{V600E} gene testing, and nodules testing positive for *BRAF*^{V600E} or for risk stratification systems (RSSs) were diagnosed as malignant, while those with negative results in both tests were considered benign. In method 2 (modified combination method), nodules were reclassified into low-risk (category 2 and 3 in the ATA guidelines and ACR-TIRADS), medium-risk (category 4), and high-risk (category 5) groups based on the malignancy rate of the RSSs. *BRAF*^{V600E} gene testing was applied only with the medium-risk group. Nodules with positive *BRAF*^{V600E} mutation were upgraded to the high-risk group, while negative cases remained in the medium-risk group.

Results: Both malignancy rates and positive *BRAF*^{V600E} mutation rates increased with the increase in RSS category (P<0.001). The combination of ACR with *BRAF*^{V600E} gene testing significantly improved the area under the curve (AUC) compared to the use of ACR or *BRAF*^{V600E} alone (the AUCs for ACR combined with *BRAF*^{V600E}, modified ACR combined with *BRAF*^{V600E}, ACR alone, and *BRAF*^{V600E} alone were 0.875, 0.878, 0.832, and 0.839, respectively; P<0.05 for both combinations *vs.* ACR or *BRAF*^{V600E} alone). Similarly, ATA combined with *BRAF*^{V600E} showed significant improvements in AUC compared to ATA alone (the AUCs for ATA combined with *BRAF*^{V600E}, modified ATA combined with *BRAF*^{V600E}, and ATA alone were 0.851, 0.846, 0.809, respectively; P<0.001 for both combination methods *vs.* ATA alone), but there was no significant difference observed compared to using *BRAF*^{V600E} alone (P=0.450 and P=0.680 for both combination methods *vs.* *BRAF*^{V600E}). Notably, the AUC of ACR combined with *BRAF*^{V600E} was greater than that of ATA combined with *BRAF*^{V600E} (P=0.047 and P=0.007 for both combination methods, respectively). There were

^ ORCID: 0009-0000-3758-5039.

no significant differences in diagnostic performance between the two combination approaches (P=0.428 for ACR combined with *BRAF*^{V600E} and P=0.314 for ATA combined with *BRAF*^{V600E}). Performing *BRAF*^{V600E} gene testing only on the medium-risk groups (modified combination method) significantly reduced the rate of *BRAF*^{V600E} gene testing (P<0.001) without increasing the false-negative rate (P=0.818 and P=0.394 for ACR and ATA, respectively).

Conclusions: Incorporating the *BRAF*^{V600E} gene test exclusively for nodules in the medium-risk group significantly improved diagnostic efficacy, reduced the utilization of gene tests, and maintained a consistent false-negative rate.

Keywords: Atypia of undetermined significance or follicular lesion of undetermined significance (AUS/FLUS); thyroid risk stratification system; *BRAF*^{V600E}; thyroid imaging reporting and data system; thyroid nodule

Submitted Jul 28, 2023. Accepted for publication Nov 14, 2023. Published online Jan 02, 2024.
doi: 10.21037/qims-23-1066
View this article at: https://dx.doi.org/10.21037/qims-23-1066

Introduction

Fine needle aspiration (FNA) biopsy is the most accurate and cost-effective method for the preoperative evaluation of thyroid nodule properties (1). However, although FNA can obtain satisfactory specimens, there are still cases where nodules cannot be definitively diagnosed (2). The Bethesda System for Reporting Thyroid Cytopathology (BSRTC) refers to atypia of undetermined significance or follicular lesion of undetermined significance (AUS/FLUS), and this phenomenon is challenging for both pathologists and surgeons (3). The incidence and malignancy rates of AUS/FLUS vary widely in recent studies, ranging from 0.8% to 28% and from 6% to 48%, respectively (2,4,5). Currently, there is no consensus on a standardized clinical approach for the treatment of these nodules (6). The 2015 American Thyroid Association (ATA) guidelines recommend repeat FNA (7). However, rediagnosis of AUS/FLUS nodules occurs in 20% to 32% of patients (8), and repeated FNA does not necessarily improve the malignancy rate of these nodules (8,9). This has led to clinical controversy regarding the management of AUS/FLUS nodules.

Ultrasound (US)-based risk stratification systems (RSSs) have played a crucial role in standardizing the treatment of thyroid nodules and are widely recognized and applied in clinical practice. However, there is a lack of agreement regarding the significance of AUS/FLUS nodules (10,11). The *BRAF*^{V600E} gene mutation is commonly found in papillary thyroid carcinoma (PTC) and serves as a widely used molecular marker for detecting PTC in clinical practice (12). Studies have shown that the *BRAF*^{V600E} gene mutation exhibits high specificity and general sensitivity for diagnosing PTC (13,14). In some studies, the combination of *BRAF*^{V600E} and RSSs was used for the diagnosis of thyroid nodules (15,16). The general approach involved performing a *BRAF*^{V600E} test on all thyroid nodules and diagnosing those with positive *BRAF*^{V600E} or RSSs results as positive while diagnosing those with negative results for both markers as negative. However, since *BRAF*^{V600E} testing can be costly, it may impose a financial burden on patients, especially those from economically disadvantaged backgrounds. Hence, is not clear whether it is necessary to conduct *BRAF*^{V600E} testing on all nodules. To clarify this issue, we examined a substantial sample size to investigate the most effective approach for managing patients with AUS/FLUS nodules. We present this article in accordance with the STARD reporting checklist (available at https://qims.amegroups.com/article/view/10.21037/qims-23-1066/rc).

Methods

Patients

The study was conducted in accordance with the Declaration of Helsinki (as revised in 2013). This study was approved by the Scientific Research and Clinical Trials Ethics Committee of the First Affiliated Hospital of Zhengzhou University (No. 2022-KY-0974-001). The requirement for informed consent was waived due to the retrospective nature of the study, specifically with the sole purpose of conducting secondary analysis on existing data.

A total of 1,408 consecutive patients with 2,097 thyroid nodules initially diagnosed as AUS/FLUS after US-FNA from February 2019 to May 2023 at the First Affiliated

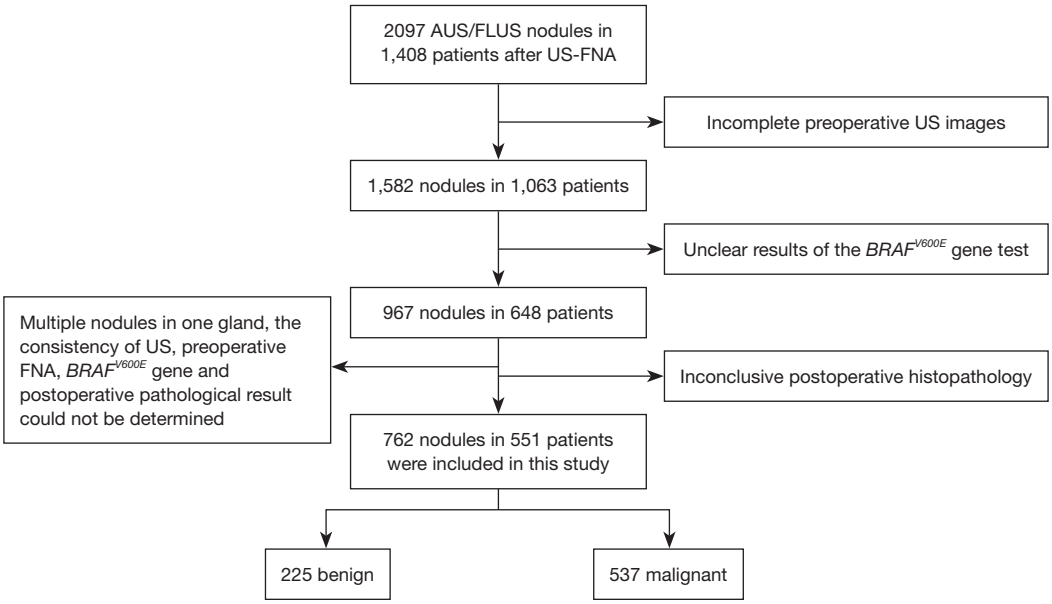

Figure 1 Flowchart of participant inclusion. AUS/FLUS, atypia of undetermined significance or follicular lesion of undetermined significance; US-FNA, ultrasound fine needle aspiration.

Hospital of Zhengzhou University were enrolled in this study. After the exclusion criteria were applied, 762 nodules in 551 patients were examined in the final analysis (*Figure 1*). The exclusion criteria were as follows: (I) incomplete preoperative US images; (II) unclear results of $BRAF^{V600E}$ gene testing; (III) inconclusive postoperative histopathology, including noninvasive follicular thyroid neoplasms with papillary-like nuclear features (NIFTP); and (IV) cases with multiple nodules in one gland, where the consistency of US images, preoperative FNA, and postoperative pathological results could not be determined.

US examination and image interpretation

Each examination was conducted by one of three highly experienced US specialists in thyroid imaging, each of whom had more than 10 years of expertise. Each US examination was conducted within half a month before the surgical procedure. The examinations were performed using Aplio 300 or 500 (Toshiba Corporation, Tokyo, Japan) US machines equipped with a 5 to 12-MHz linear array transducer. To assess the US features of the nodules, two additional sonographers with 8 and 11 years of expertise in analyzing thyroid US images were recruited. Their responsibilities included evaluating the nodules' characteristics, such as location (left, right, isthmus region), size, composition (solid or almost solid, mixed cystic and solid, cystic, spongiform), echogenicity (hyperechoic, isoechoic, hypoechoic, markedly hypoechoic), shape (taller-than-wide, wider-than-tall), margin (smooth, ill-defined, irregular or lobulated, extrathyroidal extension), and echogenic foci (punctate echogenic foci, macrocalcification, peripheral calcifications, comet-tail artifacts).

The nodules were then classified into different risk categories based on the ATA guidelines (*Figure 2*) and American College of Radiology (ACR) Thyroid Imaging Reporting and Data System (TIRADS) (*Figure 3*) (17). The ACR-TIRADS system encompasses categories 1 through 5, while the ATA guidelines, while not explicitly employing the term "TIRADS", classifies thyroid nodules into five groups: benign, very-low suspicion, low suspicion, intermediate suspicion, and high suspicion. For ease of expression, we refer to them as categories 1 to 5.

The hired sonographers were blinded to the cytological results, postoperative pathology, and $BRAF^{V600E}$ gene results of the nodules. In cases where there were differences of opinion between the two sonographers, the final decision was made by the third specialist, who had 33 years of experience in thyroid imaging. Prior to evaluating the US features, a training session involving 20 representative thyroid nodules was conducted to improve expertise and ensure consistent and dependable classification by the

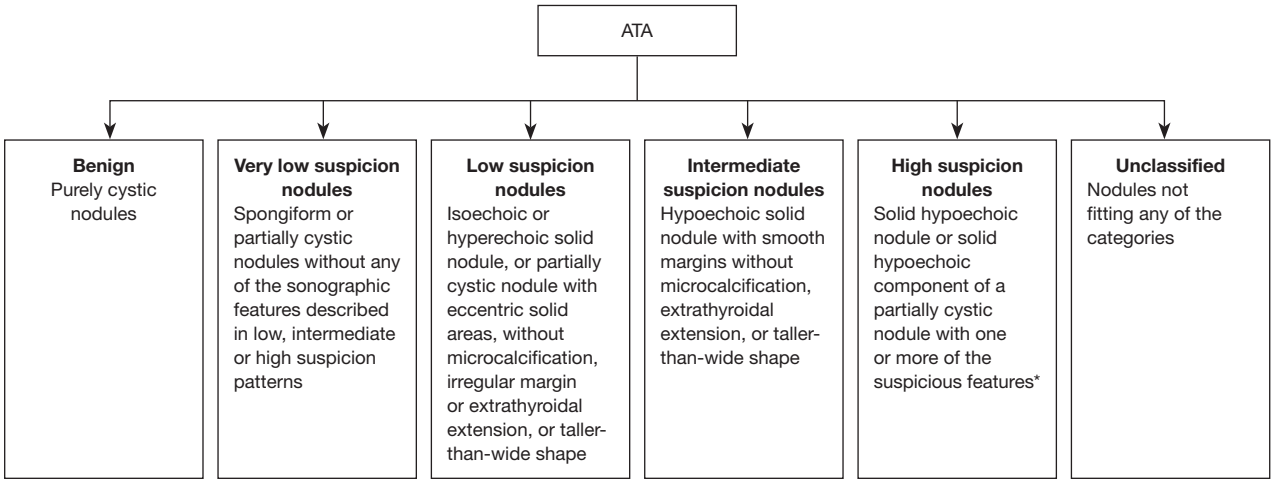

Figure 2 American Thyroid Association guidelines. *, suspicious features including irregular margins (infiltrative, microlobulated), microcalcifications, taller-than-wide shape, rim calcifications with small extrusive soft tissue component, and evidence of extrathyroidal extension. ATA, American Thyroid Association.

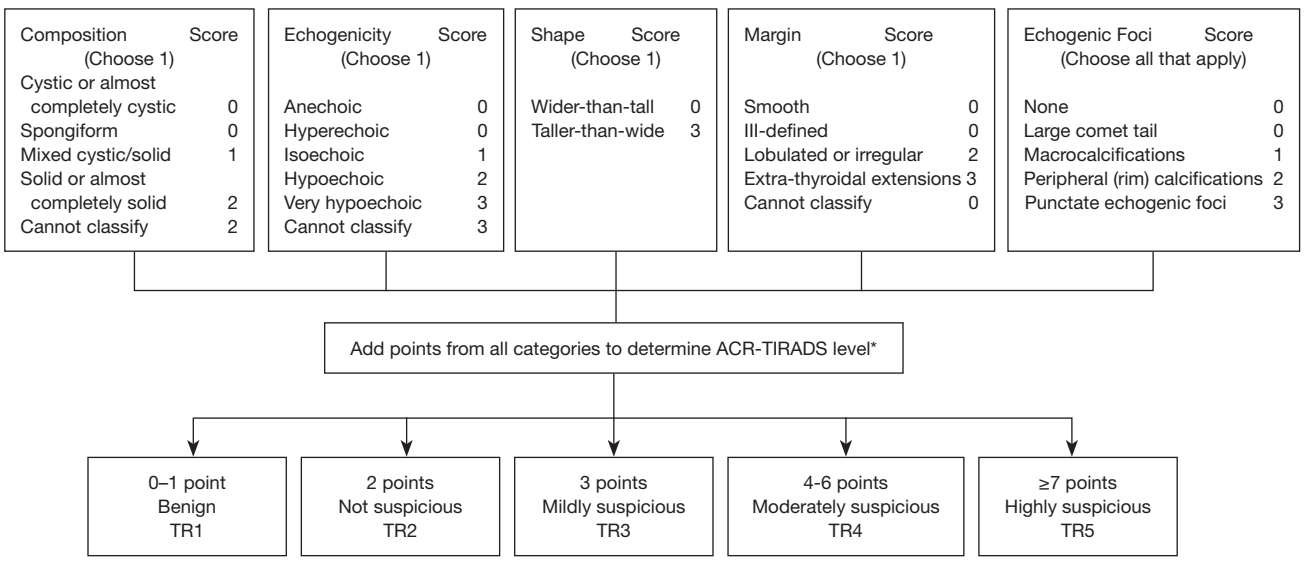

Figure 3 American College of Radiology Thyroid Imaging Reporting and Data System. *, cystic and spongiform nodules receive 0 points in total without addition of points for other categories. ACR, American College of Radiology; TIRADS, Thyroid Imaging Reporting and Data System; TR, Thyroid Imaging Reporting and Data System category.

participating sonographers. The unclassified (NC) nodules that did not fit into any of the categories outlined in the ATA guidelines were categorized as having intermediate suspicion (18,19).

US-guided FNA and BRAF^{V600E} mutation analysis

US-FNA was performed by one of three physicians with more than 11 years of work experience using a EPIQ 7 (Philips, Amsterdam, The Netherlands) system equipped with a 5 to 12-MHz line array transducer. The procedure began with disinfection the neck area, which was then covered with a sterile towel to maintain a sterile field. The physician used US guidance to carefully insert a 23 G disposable puncture needle into the nodules and collected the specimens. The collected material was then injected

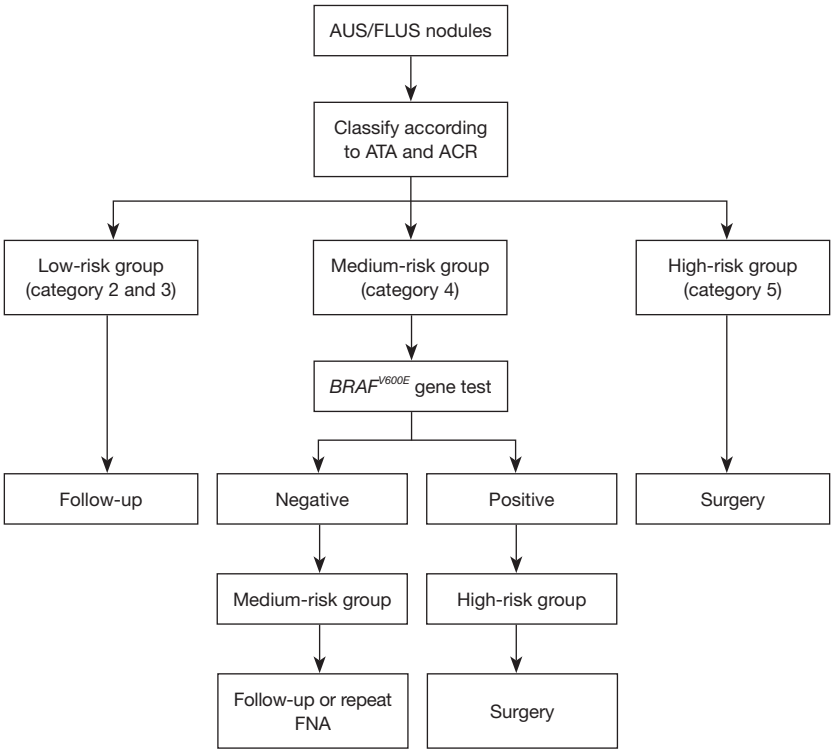

Figure 4 The new combined RSS and *BRAF*^{V600E} strategy for AUS/FLUS nodules. AUS/FLUS, atypia of undetermined significance or follicular lesion of undetermined significance; RSS, risk stratification system; ATA, American Thyroid Association; ACR, American College of Radiology; FNA, fine needle aspiration.

into a specimen tube with a liquid-based cell preservation solution. The needle was rinsed in the preservation solution to ensure all residual cells were transferred. To ensure accuracy and consistency, the diagnosis of nodules classified as AUS/FLUS was confirmed by two independent pathologists following the criteria outlined in the 2017 Bethesda system (3).

DNA samples were extracted from FNA samples using a DNA Kit (ADX-FF01 DNA Kit, AmoyDx, Xiamen, China). The *BRAF*^{V600E} mutation was assessed using the Human *BRAF*^{V600E} Gene Mutation Detection Kit (ADx-BR02, AmoyDx) according to the manufacturer's instructions. Real-time fluorescence polymerase chain reaction (PCR) was used as the detection method, and the ABI StepOne or ABI Sequence Analyzer (Applied Biosystems, Thermo Fisher Scientific, Waltham, MA, USA) was employed as the detection instrument.

Combined approaches

In method 1, all nodules underwent *BRAF*^{V600E} testing, and those found positive in the *BRAF*^{V600E} gene testing or RSSs were diagnosed as malignant. If both were negative, the diagnosis was benign.

In method 2, according to the malignancy rate of each category, the RSSs were reclassified into a low-risk group (category 2 and 3 in the ATA guidelines and ACR-TIRADS), medium-risk group (category 4), or high-risk group (category 5). *BRAF*^{V600E} gene testing was applied only in the medium-risk group. Nodules with positive *BRAF*^{V600E} mutation were upgraded to the high-risk group, while the negative ones in the medium-risk group remained unchanged (*Figure 4*).

Statistical analysis

The *BRAF*^{V600E} gene test rate (BGR) was considered to be the proportion of nodules that underwent *BRAF*^{V600E} gene testing. The false-negative rate (FNR) was considered to be the proportion of malignant nodules misdiagnosed as benign among all nodules. The false-positive rate (FPR) was considered to be the percentage of benign nodules

Table 1 Basic characteristics of patients with thyroid AUS/FLUS nodules

Variable	Total	Malignant	Benign	χ^2/t	P value
Sex, n					
Male	135	97	38	0.135	0.713
Female	416	292	124		
Age (years), mean ± SD	41.2±11.4	40.7±11.7	42.5±10.9	1.656	0.098
Nodule size (mm), mean ± SD	15.2±9.7	12.2±7.2	22.3±11.0	14.9	<0.001
Location, n				0.004	0.998
Left	334	235	99		
Right	370	261	109		
Isthmus region	58	41	17		

AUS/FLUS, atypia of undetermined significance or follicular lesion of undetermined significance; SD, standard deviation.

incorrectly diagnosed as malignant among all the nodules.

The statistical analysis was conducted using SPSS 26.0 software (IBM Corp., Armonk, NY, USA) and MedCalc 18.2.1 software (MedCalc Software Ltd, Ostend, Belgium). Continuous data were presented as the mean ± standard deviation (SD) and were compared using the independent two-sample *t*-test. Categorical data were compared using the Chi-squared test or Fisher exact test. Receiver operating characteristic (ROC) curves were constructed, and the area under the curve (AUC) was compared with the z test or DeLong method. Sensitivity, specificity, positive predictive value (PPV), negative predictive value (NPV), accuracy, BGR, FNR, and FPR, along with their 95% confidence intervals, were calculated and compared using the McNemar or Pearson Chi-squared test. A two-sided P value of <0.05 was considered statistically significant.

Results

Basic characteristics of AUS/FLUS nodules

Of the 762 AUS/FLUS nodules in the 551 patients included in this study, 225 were benign and 537 were malignant. Of the 225 benign nodules, 144 were nodular goiters, 31 were adenomatous goiters, 24 were follicular adenomas, 11 were Hashimoto thyroiditis nodules, 6 were subacute thyroiditis nodules, and 9 were others. Of the 537 malignant nodules, 505 were PTCs, 11 were follicular thyroid carcinomas, 5 were medullary thyroid carcinomas, and 16 were others. Among the patients included, 135 were male and 416 were female. The age range of the patients was 7–84 years old, with an average age of 41.2±11.4 years old. No significant differences in sex, age, or location were found between those with benign and those with malignant nodules (P>0.05; *Table 1*). Notably, the size of the benign nodules group was significantly larger than that of the malignant nodules group (P<0.001).

Malignant nodule proportions and $BRAF^{V600E}$ mutation rates in the RSS categories

Significant differences were observed in the malignancy rate across the ATA guidelines and ACR-TIRADS categories (P<0.001). Similarly, the $BRAF^{V600E}$ mutation rate also exhibited significant differences among these categories (P<0.001, except for ATA category 2). Both malignancy rates and $BRAF^{V600E}$ mutation rates increased with a higher category (*Table 2*).

Diagnostic value of $BRAF^{V600E}$ and the two RSSs

The positive $BRAF^{V600E}$ mutation rate was found to be 50.9% in all AUS/FLUS nodules and 70.9% in malignant nodules. $BRAF^{V600E}$ exhibited high specificity (96.8%) with moderate sensitivity (70.9%).

According to the ROC curves, the best cutoff was category 5 for ATA guidelines and ACR-TIRADS (*Figure 5*). There were no statistically significant differences in sensitivity, specificity, PPV, NPV, or accuracy between the ATA guidelines and ACR-TIRADS (P>0.05). Additionally, although the AUC of ACR was higher than that of ATA, this difference did not reach statistical significance (P=0.069) (*Table 3*).

Table 2 The proportion of malignant nodules and positive *BRAF*^{V600E} mutation in each category of RSSs

RSS	*BRAF*^{V600E} gene test (n)		Positive mutation rate (%)	P value	Pathological (n)		Calculated malignancy rate (%)	Recommended malignancy rate (%)	P value
	Negative	Positive			Benign	Malignant			
ACR-TIRADS				<0.001					<0.001
TR1	–	–			–	–		0	
TR2	32	0	0		31	1	3.1	<2	
TR3	41	4	8.9		35	10	22.2	5	
TR4	150	85	36.1		123	112	47.7	5–20	
TR5	151	299	66.4		36	414	92.0	>20	
ATA guideline				<0.001					<0.001
Benign	–	–			–	–		<1	
Very low suspicion	3	0	0		3	0	0	<3	
Low suspicion	85	7	7.6		71	21	22.8	5–10	
Intermediate suspicion (including unclassified)	131	76	36.7		106	101	48.8	10–20	
High suspicion	155	305	66.3		45	415	90.2	70–90	

RSS, risk stratification system; ACR, American College of Radiology; TIRADS, Thyroid Imaging Reporting and Data System; TR, TIRADS category; ATA, American Thyroid Association.

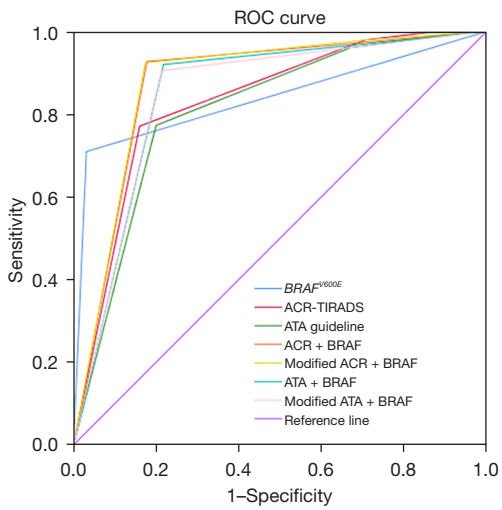

Figure 5 ROC curves of the two RSSs, *BRAF*^{V600E} mutation, and the combination approaches. ROC, receiver operating characteristic; RSS, risk stratification system; ATA, American Thyroid Association; ACR American College of Radiology; TIRADS, Thyroid Imaging Reporting and Data System.

Diagnostic value of the different combinations

Combining the RSSs with *BRAF*^{V600E} testing significantly improved the sensitivity, NPV, and accuracy compared to the RSSs or *BRAF*^{V600E} testing alone (all P values <0.05). However, the specificity and PPV were lower when compared to *BRAF*^{V600E} alone (all P values <0.05), but there were no significant differences compared to using a single RSS (all P>0.05). ACR combined with *BRAF*^{V600E} demonstrated a significantly higher AUC compared to using ACR or *BRAF*^{V600E} alone (the AUCs for ACR combined with *BRAF*^{V600E}, modified ACR combined with *BRAF*^{V600E}, ACR alone, and *BRAF*^{V600E} alone were 0.875, 0.878, 0.832, and 0.839, respectively; P<0.05 for both combinations *vs.* ACR or *BRAF*^{V600E}alone). Meanwhile, ATA combined with *BRAF*^{V600E} had a significantly higher AUC compared to using ATA alone (the AUCs for ATA combined with *BRAF*^{V600E}, modified ATA combined with *BRAF*^{V600E}, and ATA alone were 0.851, 0.846, and 0.809, respectively; P<0.001 for both combinations *vs.* ATA), but no difference was observed compared to using *BRAF*^{V600E} alone (P=0.450

Table 3 Diagnostic performance for AUS/FLUS nodules of the thyroid risk stratification systems, *BRAF*^{V600E}, and their combinations

Diagnostic method	Sensitivity (%) (95% CI)	Specificity (%) (95% CI)	PPV (%) (95% CI)	NPV (%) (95% CI)	Accuracy (%) (95% CI)	AUC (95% CI)
BRAF^{V600E} gene test	70.9 (67.3–75.0)	96.8 (94.6–99.2)	98.2 (96.9–99.5)	58.2 (53.4–63.5)	78.7 (75.8–81.7)	0.839 (0.811–0.867)
ACR-TIRADS	77.1 (73.5–80.7)	84.0 (79.2–88.8)	92.0 (89.5–94.5)	60.6 (55.1–66.0)	79.1 (76.2–82.0)	0.832 (0.798–0.865)
ATA guideline	77.3 (73.7–80.8)	80.0 (74.7–85.3)	90.2 (87.5–92.9)	59.6 (54.0–65.2)	78.1 (75.1–81.0)	0.809 (0.773–0.844)
ACR + *BRAF*^{V600E} (Method 1)	92.7 (90.5–94.9)	82.2 (77.2–87.3)	92.6 (90.3–94.8)	82.6 (77.6–87.6)	89.6 (87.5–91.8)	0.875 (0.843–0.907)
Modified ACR + *BRAF*^{V600E} (Method 2)	92.4 (90.1–94.6)	82.7 (77.7–87.7)	92.7 (90.5–94.9)	81.9 (76.9–87.0)	89.5 (87.3–91.7)	0.878 (0.846–0.909)
ATA + *BRAF*^{V600E} (Method 1)	92.0 (89.7–94.3)	78.2 (72.8–83.7)	91.0 (88.6–93.4)	80.4 (75.1–85.7)	87.9 (85.6–90.2)	0.851 (0.817–0.885)
Modified ATA + *BRAF*^{V600E} (Method 2)	90.5 (88.0–93.0)	78.7 (73.3–84.1)	91.0 (88.6–93.4)	77.6 (72.2–83.1)	92.5 (90.6–94.4)	0.846 (0.812–0.881)

AUS/FLUS, atypia of undetermined significance or follicular lesion of undetermined significance; CI, confidence interval; PPV, positive predictive value; NPV, negative predictive value; AUC, area under the curve; ACR, American College of Radiology; TIRADS, Thyroid Imaging Reporting and Data System; ATA, American Thyroid Association.

Table 4 Comparison of the two combined approaches

Guidelines	BGR (%) (95% CI)	FNR (%) (95% CI)	FPR (%) (95% CI)
ACR + *BRAF*^{V600E}	100	5.1 (3.5–6.7)	5.2 (3.7–6.8)
Modified ACR + *BRAF*^{V600E}	30.8 (27.6–34.1)	5.4 (3.8–7.0)	5.1 (3.5–6.7)
ATA + *BRAF*^{V600E}	100	5.6 (4.0–7.3)	6.4 (4.7–8.2)
Modified ATA + *BRAF*^{V600E}	27.2 (24.0–30.3)	6.7 (4.9–8.5)	6.3 (4.6–8.0)

BGR, *BRAF*^{V600E} gene test rate; CI, confidence interval; FNR, false-negative rate; FPR, false-positive rate; ACR, American College of Radiology; ATA, American Thyroid Association.

and P=0.680 for both combinations *vs.* *BRAF*^{V600E}). Additionally, the AUC of ACR combined with *BRAF*^{V600E} was significantly higher than that of ATA combined with *BRAF*^{V600E} (P=0.047 and P=0.007) (*Table 3*).

The two combined methods showed no statistically significant difference in diagnostic performance (P=0.428 for ACR combined with *BRAF*^{V600E} and P=0.314 for ATA combined with *BRAF*^{V600E}). Method 2 (modified combination method) significantly reduced the BGR compared to method 1 (both P values <0.001), but there was no statistically significant difference in FNR (P=0.818 and P=0.394 for ACR and ATA, respectively) or FPR (P=0.908 and P=0.916 for ACR and ATA, respectively) between the two methods (*Table 4*).

Discussion

Thyroid nodules classified as AUS/FLUS have historically presented a significant diagnostic and therapeutic challenge in clinical practice. Our findings revealed that the malignancy rates for category 2 nodules in the ATA guidelines and ACR-TIRADS ranged from 0% to 3.1%, suggesting that additional *BRAF*^{V600E} testing would be unnecessary. In fact, no positive *BRAF*^{V600E} mutations were detected in this category. For category 3 nodules in the ATA guidelines and ACR-TIRADS, the malignancy rates ranged from 22.2% to 22.8%, with positive mutation rates ranging only from 7.6% to 8.9%. It is worth noting that the majority of low-risk nodules underwent FNA due to

their larger size. However, in cases where nodules lack suspicious US features, it is advisable to refrain from using *BRAF*^{V600E} testing, as it does not aid in identifying non-PTC malignancies, such as follicular thyroid carcinomas and poorly differentiated thyroid carcinoma (20). Therefore, we recommend a more cost-effective follow-up approach for managing low-risk nodules.

In the high-risk group, the malignancy rates were found to be as high as 90.2–92.0%. Due to the limited sensitivity of *BRAF*^{V600E} alone in AUS/FLUS nodules (21), the utility of *BRAF*^{V600E} was considered less valuable compared to that of the RSSs in such cases. It is worth mentioning that a negative *BRAF*^{V600E} result could not definitively exclude a malignant diagnosis. However, combining *BRAF*^{V600E} with the RSSs for the high-risk group yielded significant benefits. For these nodules, it is recommended to consider more aggressive treatments, such as surgical intervention.

The diagnostic capabilities of the RSSs have been found to be limited for medium-risk nodules. To address this, *BRAF*^{V600E} testing has been used to aid in the identification of malignant nodules. Given the high specificity of *BRAF*^{V600E}, the detection of a positive mutation indicates a high likelihood of PTC (21). Therefore, in this study, combining RSSs with *BRAF*^{V600E} allowed for the screening of malignant nodules. Nodules with a positive *BRAF*^{V600E} mutation were upgraded to the high-risk group, while those with a negative mutation remained in the medium-risk group. For this subset of nodules, follow-up or repeat FNA could be advised based on the patient's preference. Compared to testing all nodules with *BRAF*^{V600E}, performing *BRAF*^{V600E} testing only on the medium-risk group resulted in a significant 69.2–72.8% reduction in BGR in AUS/FLUS nodules. Importantly, although there was a slight increase in FNR, it was not statistically significant (P=0.818 and P=0.394 for ACR and ATA, respectively). This reduction in BGR alleviated the financial burden on patients and avoided wasting medical resources. The study findings suggested that combining ACR with *BRAF*^{V600E} can provide superior diagnostic performance compared to the use of ATA alone. The observed superiority could be attributed to ACR's higher specificity, which led to a lower FPR when there was a direct classification of category 5 in the ATA guidelines and ACR-TIRADS for the high-risk group. Therefore, combining ACR with *BRAF*^{V600E} appears to be a more appropriate approach for AUS/FLUS nodules.

There were 45 NC nodules according to the ATA guidelines. However, their malignancy rate of 51.1% was similar to that of the intermediate suspicion group (48.1%). As a result, we classified them as intermediate suspicion nodules. Subsequently, the sensitivity, specificity, PPV, NPV, accuracy, and AUC were determined to be 77.3%, 80.0%, 90.2%, 59.6%, 78.1%, and 0.809, respectively. When the NC nodules were excluded, the corresponding values became 80.7%, 77.8%, 90.2%, 61.5%, 75.2%, and 0.812, respectively. There was no significant difference observed between the two definition methods (P>0.05). This is consistent with the study by Kuru *et al.* (19), in which a similar risk of malignancy was found in the NC group (63%) compared to the intermediate suspicion group (57%). Consequently, they suggested that the NC group should be included in the intermediate category of the ATA guidelines.

The sensitivity and specificity of *BRAF*^{V600E} in this study were consistent with those reported previously (13,14). It should be pointed out that *BRAF*^{V600E} gene mutations were detected in seven benign AUS/FLUS nodules in our study. Among these, two nodules were diagnosed as atypical hyperplasia, while the remaining five were nodular goiters. Similar findings were reported by Chung *et al.* (22), who described a benign nodule with a positive *BRAF*^{V600E} mutation. This particular nodule was initially classified as indeterminate for PTC based on FNA, but subsequent postoperative pathology revealed atypical hyperplasia. This suggests that atypical hyperplasia may represent a precancerous thyroid lesion, potentially leading to false-positive *BRAF*^{V600E} results. Among the five nodular goiters in our study, two surgical specimens were negative for the *BRAF*^{V600E} gene mutation, while the remaining three cases were not evaluated for the *BRAF*^{V600E} mutation after surgery. A similar case was reported by Yin *et al.* (23), where cytology indicated a positive result, but histology did not detect the *BRAF*^{V600E} mutation. In such cases, it is speculated that there may be small and indolent papillary carcinomas present, which could be missed even with careful examination. Additionally, other studies have reported *BRAF*^{V600E} mutations in benign nodules, and the reason for this could be attributed to oversensitive detection methods or procedural errors (24).

In this study, benign nodules were generally larger than malignant ones. This observation can be explained by the fact that benign nodules are typically managed through follow-up, and surgery is only considered when they cause compression-related symptoms or aesthetic concerns. The combination of *BRAF*^{V600E} and RSS assessments outperformed individual evaluations, highlighting the improved diagnostic efficacy of incorporating *BRAF*^{V600E}

gene testing as compared to using US alone. Of particular note is the combination of ACR with *BRAF*^{V600E}, which exhibited superior performance compared to ATA combined with *BRAF*^{V600E}. Furthermore, the restriction of *BRAF*^{V600E} gene testing to medium-risk nodules significantly reduced the rate of BGR while maintaining consistent FNR, supporting it as a preferable strategy over testing all nodules.

This study also had several limitations. First, the use of postoperative pathology as a reference led to a selection bias, which resulted in a relatively higher malignancy rate. Second, this study employed a retrospective design, thus limiting to the use of static images, which might have introduced some variability in the assessment of US features. Future prospective research is needed for further validation. Third, this study encountered a subset of nodules that were misdiagnosed as benign by all tests (both RSSs and *BRAF*^{V600E}), with the majority of these cases being non-PTC malignancies, primarily follicular-patterned lesions. This underscores the limited effectiveness of the methods employed in this study for non-PTC malignancies. To enhance diagnostic accuracy in such cases in the future, it is imperative to develop more suitable diagnostic approaches, such as elasticity imaging, contrast-enhanced US, *RAS* gene analysis, and other potential modalities.

In conclusion, the combination of RSSs with *BRAF*^{V600E} significantly enhances the efficacy of diagnosing AUS/FLUS nodules. Moreover, compared to testing all nodules with *BRAF*^{V600E}, only testing those nodules considered to be medium risk can significantly reduce the BGR without increasing the FNR.

Acknowledgments

Funding: None.

Footnote

Reporting Checklist: The authors have completed the STARD reporting checklist. Available at https://qims.amegroups.com/article/view/10.21037/qims-23-1066/rc

Conflicts of Interest: All authors have completed the ICMJE uniform disclosure form (available at https://qims.amegroups.com/article/view/10.21037/qims-23-1066/coif). The authors have no conflicts of interest to declare.

Ethical Statement: The authors are accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved. The study was conducted in accordance with the Declaration of Helsinki (as revised in 2013) and was approved by the Scientific Research and Clinical Trials Ethics Committee of the First Affiliated Hospital of Zhengzhou University (No. 2022-KY-0974-001). The requirement for informed consent was waived due to the retrospective nature of the study.

References

1. Ke J, Jianyong L, Ying L, Genpeng L, Linlin S, Zhihui L, Jinnan L, Xueying S, Yong J, Jingqiang Z. The use of The Bethesda System for Reporting Thyroid Cytopathology in a Chinese population: An analysis of 13 351 specimens. Diagn Cytopathol 2019;47:876-80.
2. Sauter JL, Lehrke H, Zhang X, Al Badri OT, Rodriguez-Gutierrez R, Delivanis DA, Singh Ospina N, Donegan D, Hamidi O, Iñiguez-Ariza N, Sharma A, Kittah NEN, Tamhane SU, Hurtado Andrade MD, Kotwal A, Jenkins SM, Spears G, Rivera M, Dean DS, Henry MR. Assessment of The Bethesda System for Reporting Thyroid Cytopathology. Am J Clin Pathol 2019;152:502-11.
3. Kakudo K, Higuchi M, Hirokawa M, Satoh S, Jung CK, Bychkov A. Thyroid FNA cytology in Asian practice-Active surveillance for indeterminate thyroid nodules reduces overtreatment of thyroid carcinomas. Cytopathology 2017;28:455-66.
4. Mileva M, Stoilovska B, Jovanovska A, Ugrinska A, Petrushevska G, Kostadinova-Kunovska S, Miladinova D, Majstorov V. Thyroid cancer detection rate and associated risk factors in patients with thyroid nodules classified as Bethesda category III. Radiol Oncol 2018;52:370-6.
5. Crescenzi A, Palermo A, Trimboli P. Cancer prevalence in the subcategories of the indeterminate class III (AUS/FLUS) of the Bethesda system for thyroid cytology: a

meta-analysis. J Endocrinol Invest 2021;44:1343-51.
6. Abelardo AD, Sotalbo KCJ. Clinical management of thyroid aspirates diagnosed as atypia of undetermined significance in the Philippines. Gland Surg 2020;9:1788-96.
7. Haugen BR, Alexander EK, Bible KC, Doherty GM, Mandel SJ, Nikiforov YE, Pacini F, Randolph GW, Sawka AM, Schlumberger M, Schuff KG, Sherman SI, Sosa JA, Steward DL, Tuttle RM, Wartofsky L. 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;26:1-133.
8. Bayona A, Benavent P, Muriel A, Abuchaibe C, Sharpe SC, Tarasova V, McIver B, Valderrabano P. Outcomes of repeat fineneedle aspiration biopsy for AUS/FLUS thyroid nodules. Eur J Endocrinol 2021;185:497-506.
9. Evranos Ogmen B, Aydin C, Kilinc I, Aksoy Altinboga A, Ersoy R, Cakir B. Can Repeat Biopsies Change the Prognoses of AUS/FLUS Nodule? Eur Thyroid J 2020;9:92-8.
10. Li Q, Yang L, Yang L, Jiang X, Li S. Utility of Six Ultrasound-Based Risk Stratification Systems in the Diagnosis of AUS/FLUS Thyroid Nodules. Acad Radiol 2023. [Epub ahead of print]. doi: 10.1016/j.acra.2023.04.029.
11. Sahli ZT, Karipineni F, Hang JF, Canner JK, Mathur A, Prescott JD, Sheth S, Ali SZ, Zeiger MA. The association between the ultrasonography TIRADS classification system and surgical pathology among indeterminate thyroid nodules. Surgery 2019;165:69-74.
12. Choden S, Keelawat S, Jung CK, Bychkov A. VE1 Immunohistochemistry Improves the Limit of Genotyping for Detecting BRAFV600E Mutation in Papillary Thyroid Cancer. Cancers (Basel) 2020;12:596.
13. Zheng B, Zarka MA, Chen C, You J, Sun L, Chen L. The largest CAP-certified Chinese reference laboratory experience with the Bethesda system for reporting thyroid cytopathology: correlation with histologic and BRAF data. J Am Soc Cytopathol 2018;7:16-21.
14. Yoo SK, Lee S, Kim SJ, Jee HG, Kim BA, Cho H, Song YS, Cho SW, Won JK, Shin JY, Park do J, Kim JI, Lee KE, Park YJ, Seo JS. Comprehensive Analysis of the Transcriptional and Mutational Landscape of Follicular and Papillary Thyroid Cancers. PLoS Genet 2016;12:e1006239.
15. Zhou YF, Zhang YF, Fu HJ, Yang WP, Zhao CK, Xu HX. Improving the diagnosis of AUS/FLUS thyroid nodules using an algorithm with combination of BRAFV600E mutation analysis and ultrasound pattern-based risk stratification. Clin Hemorheol Microcirc 2021;77:273-85.
16. Li Q, Yang L, Lv J, Xu L, Zhang M, Li S. The combination of BRAF(V600E) mutation and Chinese Thyroid Imaging Reporting and Data System is helpful in the management of AUS/FLUS thyroid nodules. Endocrine 2022;78:507-16.
17. Tessler FN, Middleton WD, Grant EG, Hoang JK, Berland LL, Teefey SA, Cronan JJ, Beland MD, Desser TS, Frates MC, Hammers LW, Hamper UM, Langer JE, Reading CC, Scoutt LM, Stavros AT. ACR Thyroid Imaging, Reporting and Data System (TI-RADS): White Paper of the ACR TI-RADS Committee. J Am Coll Radiol 2017;14:587-95.
18. Yim Y, Na DG, Ha EJ, Baek JH, Sung JY, Kim JH, Moon WJ. Concordance of Three International Guidelines for Thyroid Nodules Classified by Ultrasonography and Diagnostic Performance of Biopsy Criteria. Korean J Radiol 2020;21:108-16.
19. Kuru B, Kefeli M, Danaci M. Comparison of 5 Thyroid Ultrasound Stratification Systems for Differentiation of Benign and Malignant Nodules and to Avoid Biopsy Using Histology as Reference Standard. Endocr Pract 2021;27:1093-9.
20. Chung JH. BRAF and TERT promoter mutations: clinical application in thyroid cancer. Endocr J 2020;67:577-84.
21. Nikiforov YE, Carty SE, Chiosea SI, Coyne C, Duvvuri U, Ferris RL, Gooding WE, LeBeau SO, Ohori NP, Seethala RR, Tublin ME, Yip L, Nikiforova MN. Impact of the Multi-Gene ThyroSeq Next-Generation Sequencing Assay on Cancer Diagnosis in Thyroid Nodules with Atypia of Undetermined Significance/Follicular Lesion of Undetermined Significance Cytology. Thyroid 2015;25:1217-23.
22. Chung KW, Yang SK, Lee GK, Kim EY, Kwon S, Lee SH, Park DJ, Lee HS, Cho BY, Lee ES, Kim SW. Detection of BRAFV600E mutation on fine needle aspiration specimens of thyroid nodule refines cyto-pathology diagnosis, especially in BRAF600E mutation-prevalent area. Clin Endocrinol (Oxf) 2006;65:660-6.
23. Yin L, Tang Y, Yu S, Wang C, Xiao M, Wang Y, Liu SJ, Gao L, Huang K, Jin L. The Role of BRAF V600E in Reducing AUS/FLUS Diagnosis in Thyroid Fine Needle Aspiration. Endocr Pathol 2019;30:312-7.
24. Gharib H, Papini E, Garber JR, Duick DS, Harrell RM, Hegedüs L, Paschke R, Valcavi R, Vitti P; AACE/

ACE/AME Task Force on Thyroid Nodules. American Association of Clinical Endocrinologists, American College of Endocrinology, and Associazione Medici Endocrinologi Medical Guidelines for Clinical Practice for The Diagnosis and Management of Thyroid Nodules--2016 Update. Endocr Pract 2016;22:622-39.

Cite this article as: Ma X, Huang Y, Shen R, Cui Y, Kang Y, Yu J, Cui K. An economically efficient strategy for diagnosing atypia of undetermined significance or follicular lesion of undetermined significance thyroid nodules with ultrasound-based risk stratification systems and *BRAF*^{V600E} testing. Quant Imaging Med Surg 2024;14(1):920-931. doi: 10.21037/qims-23-1066

