Thyroid imaging reporting and data system combined with Bethesda classification in qualitative thyroid nodule diagnosis

Huiwen Tan, MD[a,c], Zhihui Li, MD[b,*], Nong Li, MD[c,*], Jianrong Qian, MD[d], Fengchun Fan, MD[e], Huiling Zhong, MBBS[f], Jinquan Feng, MBBS[e], Huajun Xu, MBBS[f], Zhongxing Li, MBBS[d]

Abstract

Objective: We aimed to investigate the value of the combined use of high-resolution ultrasound thyroid imaging reporting and data system (TI-RADS) classification and thyroid fine needle aspiration cytology (Bethesda classification) for the qualitative diagnosis of benign and malignant thyroid nodules.

Methods: We enrolled 295 patients with 327 thyroid nodules who were scheduled to undergo thyroid nodule surgery. Before surgery, all the patients underwent ultrasound and scoring with the TI-RADS classification, along with thyroid fine needle biopsy cytology under ultrasound guidance (US-FNAC) and scoring with the Bethesda classification. After surgery, the TI-RADS and Bethesda classification scores, separately and in combination, were compared with the postoperative pathological results in terms of the differential diagnosis of thyroid nodules.

Results: TI-RADS classification score 4 exhibited the highest diagnostic value for thyroid cancer; the sensitivity, specificity, and accuracy were 92.7%, 70.7%, and 87.1%, respectively, whereas the Kappa and receiver-operating characteristics (ROC) values were 0.651 and 0.817, respectively. Moreover, Bethesda classification score 3 exhibited the highest diagnostic value for thyroid cancer; the sensitivity, specificity, and accuracy were 90.0%, 94.3%, and 91.1%, respectively, whereas the Kappa and ROC values were 0.78 and 0.914, respectively. With regard to the combined diagnostic method, a score of 7 exhibited the highest diagnostic value for thyroid cancer; the sensitivity, specificity, and accuracy were 97.3%, 92.0%, and 95.9%, respectively, whereas the Kappa and ROC values were 0.893 and 0.946, respectively.

Conclusion: The combination of high-resolution ultrasonography TI-RADS classification and US-FNAC (Bethesda classification) can improve the accuracy of malignant thyroid nodules diagnosis.

Abbreviations: BI-RADS = breast imaging reporting and data system, ROC = receiver-operating characteristics, TBSRTC = Bethesda grading diagnostic report system, TI-RADS = thyroid imaging reporting and data system, US-FNAC = ultrasound-guided thyroid fine-needle aspiration cytology.

Keywords: Bethesda grading, combined diagnosis, thyroid imaging reporting and data system (TI-RADS), thyroid nodules

1. Introduction

Thyroid nodules are commonly observed in the population (incidence: 4%–7%/10–18 million people), and may develop as a result of various thyroid disorders. They are characterized as the local growth of abnormal thyroid cells that are significantly different from normal thyroid tissue surrounding the scattered mass.[1] Most of the discovered nodules are benign, although an increase in the incidence of thyroid cancer has been reported. The detection rate of thyroid nodules through physical examination (cervical palpation) is approximately 5%,[2] whereas the rate of nodules found incidentally on ultrasonography indicates a prevalence of 20% to 76%.[3]

In a 2015 study in China, thyroid disease epidemiology was assessed in 10 cities, and included 15,008 urban adults.[4] The researchers found that the overall prevalence of thyroid nodules was 12.8%, whereas the prevalence of thyroid nodules in children aged 6 to 18 years was 10.59%, lower than that of Chinese adults. Moreover, the prevalence of thyroid nodules was 11.89% in females and 9.26% in males.[5] The prevalence rate of thyroid nodules and the multiple nodules rate were positively correlated with age and body mass index, and this trend was more apparent in women.[6,7]

Thyroid nodules can be classified as benign or malignant. The purpose of thyroid nodule detection is to identify cases with

thyroid cancer (malignant nodules). Thyroid cancer is the most common malignant tumor of the endocrine system, and accounts for 1.1% of all malignant tumors. In fact, thyroid cancer is considered one of the 10 most common malignant tumors; it is the eighth most common malignant tumor in women. The incidence of thyroid cancer varies with geographical location, age, and sex. The male to female morbidity rate is approximately 1:3. The factors associated with malignant nodules include a family history of thyroid cancer; history of neck irradiation after 15 years of age; male sex; fast-growing nodules, except for vocal cord lesions (such as inflammatory reaction or polyps) after persistent hoarseness; pronunciation difficulties; irregular nodular shape; poor mobility and adhesion to tissues surrounding the nodules; and increased serum thyroid-stimulating hormone (TSH) level.[8–11]

The clinical treatment of benign and malignant thyroid nodules differs, and there are also significant differences in the quality of life of patients and the medical costs involved.[12] Thus, the differentiation of benign and malignant thyroid nodules is vital in clinical evaluation. Diagnostic tests can be used to determine whether a thyroid nodule is benign or malignant (cancerous). High-resolution ultrasound is currently the preferred method to evaluate thyroid nodules.[13,14] In fact, ultrasound imaging plays a major role in the diagnosis, monitoring, and therapeutic decisions of thyroid disease. With regard to the assessment of benign and malignant thyroid nodules, ultrasound is considered to be superior to computed tomography and magnetic resonance imaging. Although new diagnostic imaging tools, such as elastosonography and thyroid contrast-enhanced ultrasonography, have been commonly used for the assessment of thyroid nodules in recent years, their clinical value remains unclear.[15,16] For nodules that cannot be diagnosed via ultrasound, ultrasound-guided thyroid fine-needle puncture cytology (US-FNAC) can be used for further diagnosis. However, same as ultrasound examination, the diagnosis of thyroid nodules using US-FNAC depends on the experience of the physician, and hence, in certain cases, clinical diagnosis can be incorrect. Therefore, researchers have proposed using TI-RADS grading and Bethesda grading (with thyroid fine-needle puncture cell pathology) for evaluating thyroid nodules via ultrasound, thus enabling an objective evaluation of thyroid nodules.

However, to our knowledge, no study has comprehensively assessed the sensitivity, specificity, and accuracy of the diagnosis of malignant thyroid nodules via combined high-resolution TI-RADS and Bethesda grading (thyroid fine-needle aspiration cytology).

In the present study, the pathological results after thyroid nodule surgery were taken as the criterion standard for evaluating thyroid nodules. High-resolution ultrasound TI-RADS grading and thyroid fine-needle puncture cytology grading (Bethesda classification), separately and in combination, were performed, and their accuracy, sensitivity, and specificity in the diagnosis of thyroid malignant nodules were compared. We aimed to explore the value of the combination of these 2 standardized grading methods in the differential diagnosis of thyroid nodules.

2. Subjects and methods

2.1. Subjects

A total of 295 patients hospitalized with thyroid nodules (n= 327) were enrolled in this study. Of these patients, 150 with thyroid nodules (162 nodules) were scheduled to undergo thyroid nodule surgery in Karamay People's Hospital and 145 patients with thyroid nodules (165 nodules) were scheduled to undergo thyroid surgery in the West China Hospital of Sichuan University from March 2015 to December 2016. The nodules were detected at an outpatient clinic or during medical check-ups at an endocrine metabolism clinic, and the patients were hospitalized and scheduled to undergo thyroid surgery.

All the enrolled subjects were aged >12 years, and were not contraindicated for thyroidectomy. Patients with blood system diseases (as those with blood system diseases are prone to bleeding) and those with severe coronary heart disease or long-term anticoagulant drug administration (warfarin) were excluded. None of the patients had heart, lung, liver, and renal dysfunction, and none of the patients exhibited significant hypertension (systolic blood pressure >180 mmHg and diastolic blood pressure >110 mmHg) or uncontrolled blood pressure with oral antihypertensive treatment. The average age of the 295 surgical patients was 45.33 ± 12.17 years (range, 12–76 years). The cohort consisted of 89 males and 206 females, and the maximum nodule diameter ranged from 0.40 to 5.5 cm.

The study followed the principles expressed in the World Medical Association Declaration of Helsinki and the International Ethical Guidelines for Biomedical Research Involving Subjects (GIOMS, Geneva, 1993) and Chinese clinical research management regulations. The study program was approved by the medical ethics committee of the Karamay People's Hospital and West China Hospital of Sichuan University. All subjects voluntarily provided written informed consent on each occasion of diagnostic examinations and therapeutic procedures and also for the publication of this study.

2.2. Methods

2.2.1. Nodule score. High-resolution ultrasonography and the TI-RADS hierarchical diagnostic criteria were used to evaluate the 295 patients who volunteered to participate in the study; the scoring criteria are demonstrated in Table 1. US-FNAC was also conducted, and the thyroid fine-needle aspiration cytology Bethesda grading diagnostic report system (TBSRTC)[16–19] was used to estimate scores; the scoring criteria are shown in Table 2. Moreover, the TI-RADS grading scores and Bethesda classification scores were combined using an accumulative method to determine a comprehensive score (joint score: TI-RADS classification score + Bethesda classification score), and the ultrasound TI-RADS classification scores and thyroid fine-needle puncture Bethesda scores were compared with the postoperative pathological results. Thus, we finally compared the TI-RADS grading, Bethesda classification, and combined scores for the diagnosis of thyroid cancer in terms of sensitivity, specificity, and accuracy.

2.2.2. Two-dimensional and color Doppler blood flow imaging. The devices used for ultrasound examination were the ATL HDI 5000 (Absolute Medical Equipment, Wesley Hills, NY) and IU22 (Philips Healthcare, Andover, MA) ultrasound diagnostic instrument with a 5 to 12 MHz probe.

The patient was maintained in a supine position, with a pillow under the neck to slightly tilt the head back. First, conventional 2-dimensional ultrasound (gray scale) was performed to confirm the presence of the nodules, and for detailed observation by enlarging the nodules. The clearest section was selected, and the following parameters of thyroid nodules were investigated:

Table 1

High-resolution ultrasound TI-RADS classification and scoring.

TI-RADSGrade	Evaluation	Description of ultrasound pattern	Score
1	Negative	Normal thyroid	
2	Benign	Liquid, mixed, regular shape, well-defined margin	1
3	Probably benign	No suspicious ultrasound features	2
4	Probably malignant	Solid, hypoechoic, microcalcifications, poor margins, defined/lobulated, aspect ratio >1 (taller-than-wide shape)	
4A	Low suspicion of malignancy	One suspicious ultrasound feature	3
4B	Intermediate suspicion of malignancy	Two suspicious ultrasound features	4
4C	Moderate concern but not classic for malignancy	Three or four suspicious ultrasound features	5
5	Highly suggestive of malignancy	More than 4 suspicious ultrasound features, including microcalcifications and. lobulated, particularly	6
6	Malignant	Biopsy-proven malignancy	

Suspicious ultrasound features of nodules: solid, hypoechoic, microcalcifications, irregular shape and margins (infringement, lobulated), taller-than-wide shape, uneven acoustic halo with rear attenuation, infringement of the thyroid envelope, rich blood flow within the nodule, abnormally shaped lymph nodes. TI-RADS = thyroid imaging reporting and data system.

internal echogenicity (to understand whether the nodules are hyperechoic or anechoic, hypoechoic or mixed echogenicity, and so on), acoustic halo (yes/no), microcalcification within the nodules (strong echo points <1.0 mm, no acoustic shadow), rhyme, and blood supply to the nodules.

The size of the left and right lobe of the thyroid gland and the thickness of the isthmus were measured and the location, size, boundary, internal echo, and presence of calcification around the nodules were recorded. Color and spectral Doppler were used to observe the peripheral and internal blood flow signals and the blood flow velocity of the tumor, and the images were stored for later use.

Real-time dynamic imaging of the operation process was performed, and the results were analyzed by 2 experienced senior ultrasound physicians using the double-blind method. Upon disagreement, consensus was reached through discussion.

2.2.3. US-FNAC. US-FNAC was jointly performed by specially trained interventional ultrasound physicians and pathologists engaged in pathological cytology. The patient was maintained in the supine position with a high neck pad and the thyroid fully exposed. The biopsy site skin was cleaned with 75% alcohol. US-FNAC was performed using a 20-mL empty syringe and a 22-gauge needle without local anesthesia. A thin needle was inserted into the thyroid and a sample of thyroid cells and fluid was collected. The nodes were first selected for puncture and biopsy, a needle was inserted into the nodules under ultrasound guidance, 2 mL air was drawn in, and the nodule was punctured 2 to 3 times back and forth in different directions. The needle was quickly pulled out after eliminating the negative pressure, and the suction

smear was fixed with 95% ethanol; the puncture was pressed with sterile gauze for 5 to 10 minutes. In cases with bleeding, including subcutaneous or subcapsular hemorrhage, the incidence of hematoma was low, and the hematoma could subside over time without any treatment. The cytopathological results were recorded using the thyroid cytopathic Bethesda reporting system of Cibas et al[16], and were classified into 6 categories: grade 1, the specimen was not satisfactory (insufficient number of cells or excessive amount of blood for diagnosis); grade 2, benign; grade 3, atypical lesions or vacuolar lesions with unclear significance; grade 4, follicular neoplasms or suspected follicular lesions; grade 5, suspicious malignancies; and grade 6, malignant tumor. The images were read by pathologists who were familiar with the pathological cytology of the thyroid and were analyzed using a blinded method.

2.3. Statistical analysis

Clinical observation data were collected in an Excel database, and were examined using the SPSS 17. 0 statistical software package, with the postoperative pathological examination data used as the criterion standard. Moreover, the comprehensive scores from the TI-RADS and Bethesda grading system in each patient, and the individual ultrasound TI-RADS and thyroid fine-needle puncture scores were compared with the criterion standard. Diagnostic tests were performed using a 4-grid table to calculate the sensitivity, specificity, accuracy, false-positive rate, false-negative rate, positive predictive value, negative predictive value, ROC curves, and Kappa value.

The Kappa test is a statistical test to evaluate the agreement between 2 classifications, and is a relatively effective method for testing reliability.[20] A larger Kappa value is associated with a greater consistency of the judgment result.[21] ROC curves were used to compare for sensitivity and specificity. A larger area under the ROC curves was associated with a higher diagnostic accuracy.

3. Results

With regard to the pathological results of the 327 thyroid nodules, 235 malignant nodules were determined as papillary thyroid carcinoma, whereas the other 92 nodules were benign. The comprehensive scores with the TI-RADS classification and Bethesda classification in the diagnosis of thyroid cancer

Table 2

Bethesda system for reporting thyroid cytopathology and scoring.

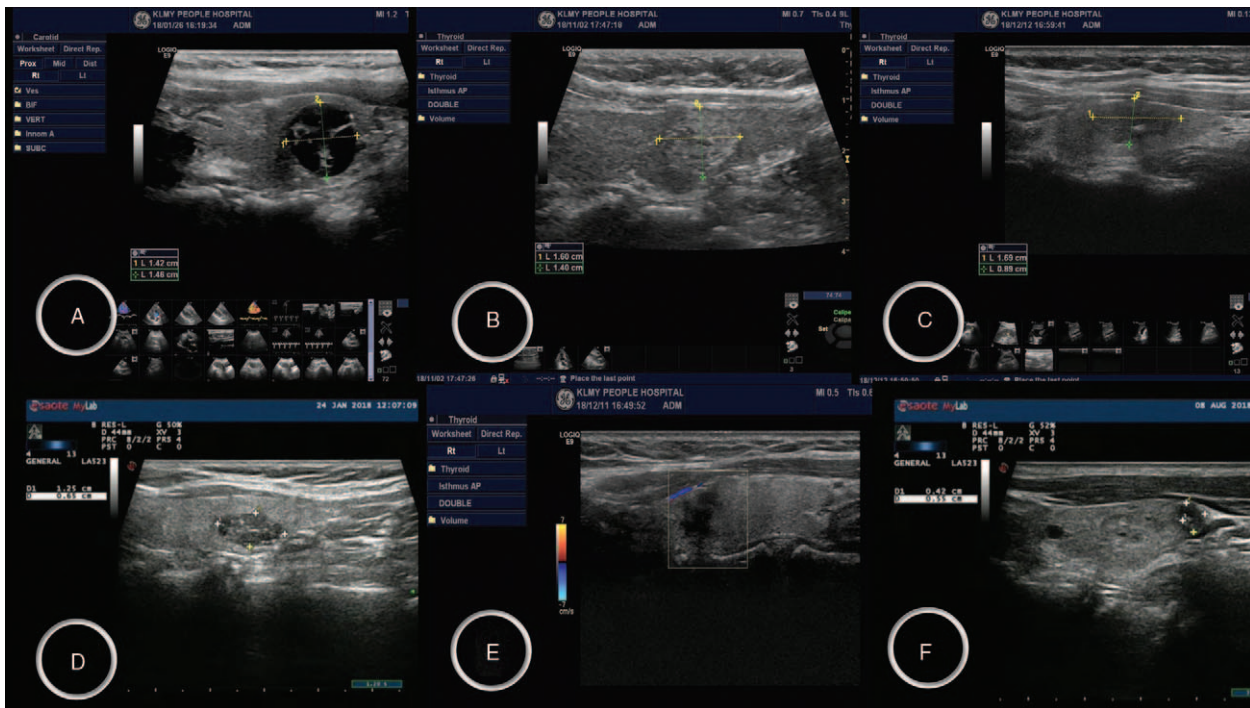| Diagnostic category	Score
Category 1: Nondiagnostic or unsatisfactory	1
Category 2: Benign	0
Category 3: Atypia of undetermined significance or follicular lesion of undetermined significance	2
Category4: Follicular neoplasm or suspicious for a follicular neoplasm	3
Category 5: Suspicious for malignancy	4
Category 6: Malignant	6

Figure 1. Thyroid ultrasound images (thyroid imaging reporting and data system, A–F).

indicated that, of 327 thyroid nodules, 1 was of grade 2, 12 were of grade 3, 66 were of grade 4A, 71 were of grade 4B, 87 were of grade 4C, and 90 were of grade 5. Figure 1 shows the thyroid ultrasound image, containing ultrasound image from TI-RADS 2–5 grades. ① TI-RADS 2 means typical and well-defined benign nodules, such as adenomas and cystic nodules; ② TI-RAD 3: atypical benign nodules, such as complex nodular goiter, with a malignant risk of less than 5%;TI-RAD 4: Suspicious malignant nodules, which subdivided into ③ 4a, ④ 4band ⑤ 4c subtypes,

with a malignant risk of 5% to 85%; ⑥ TI-RAD 5: is a typical thyroid cancer, malignant risk of 85% to 100%, suspected thyroid malignant nodules with cervical lymph node metastasis (Fig. 1). Table 3 shows that, as the TI-RADS grading scores increased from 3 to 6 points, the sensitivity decreased, but its specificity increased. Thus, TI-RADS4 had the highest accuracy (87.1%), TI-RADS4 and TI-RADS5 had the highest sensitivity + specificity (163.4%), and TI-RADS4 had the highest Kappa value (0.651). Among the various TI-RADS classifications in the

Table 3

TI-RADS and Bethesda classification scores and the combination of the 2 comprehensive scores for the diagnosis of thyroid carcinoma in terms of sensitivity, specific, and accuracy.

Score	Accuracy	Sensitivity	Specificity	False-positive (%)	False-negative (%)	Positive predicitive (%)	Negative predicitive (%)	Kappa value
TI-RADS 3	78.60%	100.00%	14.30%	85.70%	0.00%	77.90%	100.00%	0.200
TI-RADS 4	87.10%	92.70%	70.70%	29.30%	7.30%	90.30%	76.81%	0.651
TI-RADS 5	76.60%	71.40%	92.00%	8.00%	28.60%	96.30%	52.27%	0.507
TI-RADS 6	54.60%	39.50%	98.70%	1.30%	60.50%	98.90%	35.75%	0.242
Bethesda 2	90.50%	89.50%	93.30%	6.70%	10.50%	97.50%	75.27%	0.768
Bethesda 3	91.10%	90.00%	94.30%	5.70%	10.00%	97.90%	75.86%	0.780
Bethesda 4	81.70%	75.50%	100.00%	0.00%	24.50%	100.00%	58.14%	0.610
Bethesda 6	78.60%	71.40%	100.00%	0.00%	28.60%	100.00%	54.35%	0.559
Combination 4	78.30%	100.0%	14.7%	85.3%	0.0%	77.5%	100.00%	0.204
Combination 5	82.4%	100.0%	30.7%	69.3%	0.0%	80.9%	100.00%	0.397
Combination 6	94.30%	99.10%	80.00%	20.00%	0.90%	93.70%	96.55%	0.839
Combination 7	95.90%	97.30%	92.00%	8.00%	2.70%	97.30%	92.00%	0.893
Combination 8	91.20%	90.90%	92.00%	8.00%	9.10%	97.10%	77.53%	0.781
Combination 9	87.10%	84.50%	94.70%	5.30%	15.50%	97.90%	67.62%	0.700
Combination 10	82.70%	76.80%	100.00%	0.00%	23.20%	100.00%	59.52%	0.628
Combination 11	68.50%	57.70%	100.00%	0.00%	42.30%	100.00%	44.64%	0.410
Combination 12	48.80%	31.40%	100.00%	0.00%	68.60%	100.00%	33.19%	0.189

TI-RADS = thyroid imaging reporting and data system.

Variable	AUC	SE	95% CI	
			Lower limit	Upper limit
T-RADS 4	0.817	0.033	0.752	0.882
T-RADS5	0.817	0.027	0.765	0.869
Bethesda3	0.914	0.021	0.874	0.955
Combination7	0.946	0.019	0.909	0.984

AUC = area under the curve, CI = confidence interval, ROC = receiver-operating characteristic, SE = standard error.

diagnosis of thyroid cancer, TI-RADS4 (TI-RADS3B level) exhibited the maximum values for predictive qualitative diagnosis.

Bethesda classification grade 4 exhibited the highest accuracy and sensitivity for the diagnosis of thyroid carcinoma (90.1% and 90.0%, respectively), Bethesda classification grade 5 and 6 exhibited the highest specificity (100%), Bethesda classification grade 4 showed the highest sensitivity + specificity (Bethesda classification score of 3) and highest Kappa value (0.780). Among the Bethesda classification grades, a joint score of 7 indicated the highest accuracy and sensitivity, a joint score of 11 and 12 indicated the highest specificity, a score of 10 indicated high sensitivity (100%), score of 7 indicated high sensitivity + specificity (189.3%), and a score of 7 indicated the greatest Kappa value (0.893). In Table 4, the TI-RADS 4 score, TI-RADS 5 score, Bethesda classification 3 score, and combined score of 7 under ROC curve had an area under the ROC curve of 0.817, 0.914, and 0.946, respectively (Table 4); the combined score of 7 indicated the largest area under the ROC curve (Fig. 1).

4. Discussion

The differentiation of benign and malignant thyroid nodules is of great clinical significance. At present, although high-resolution ultrasound remains the preferred method for evaluating thyroid nodules, the pathology of thyroid nodules is complex, the ultrasound findings of benign and malignant nodules often overlap, and the diagnostic results are closely related to the clinical experience of the ultrasound specialists. Due to the differences in the outcomes between hospitals and doctors, it is vital to establish a unified standard for the classification and evaluation of the severity of thyroid nodules, so that each physician can perform an objective and consistent evaluation of the nature (benign and malignant) of thyroid nodules.

In 2009, Horvath et al[22] proposed a thyroid imaging reporting and data system (TI-RADS) based on the breast imaging reporting and data system of the American Society of Radiology (breast imaging reporting and data system; BI-RADS). In recent years, studies have attempted to evaluate the reliability and diagnostic performance of TI-RADS,[12,23,24,25] and have found that TI-RADS has high sensitivity, specificity, and accuracy. The TI-RADS diagnostic system is useful for differentiating benign from malignant thyroid nodules; however, the sensitivity, specificity, and accuracy of the differential diagnosis of benign and malignant thyroid nodules by TI-RADS are affected by nodule size. Moreover, the clinical application of TI-RADS is complicated, and different sonographers may have different classification results for the same image. Hence, FNAC can be used to improve the diagnostic accuracy of benign and malignant thyroid nodules. Preoperative detection of thyroid cancer by FNAC showed a sensitivity of 83% (range, 65%–98%), specificity of 92% (range, 72%–100%), positive predictive value of 75% (range, 50%–96%), false-negative rate of 5% (range, 1%–11%), and false-positive rate of 5% (range, 0–7%).[26] Currently, US-FNAC has been used to improve the success rate of puncture and the accuracy of the puncture results. Some studies reported that the false-negative rate of US-FNAC was <3%.[27,28]

As the methods and diagnostic terminology of FNAC widely differ among hospitals, uniform diagnostic terminology and criteria are needed. The establishment of the thyroid cytopathic Bethesda reporting system led to the standardized reporting of thyroid fine-needle aspiration results. In fact, the Bethesda reporting system is widely used worldwide, and pathologists use this system to efficiently communicate with clinicians, thus offering uniform reporting template for thyroid fine-needle puncture, as recommended by the American Thyroid Association.[29]

However, US-FNAC is susceptible to specimen puncture bleeding, pathological cytological diagnosis errors, and relatively difficult nodular puncture, which could lead to substandard puncture specimen quality and difficulty in qualitative diagnosis. Despite repeated puncture, a certain proportion of pathological cytological results remain unclear.

According to the literature, the false-negative rate is higher for the cytological examination results of thyroid nodules with highly suspicious ultrasound manifestations. An analysis of 1343 cytological examination results of benign nodules indicated that the malignant rate could be as high as 29% if the ultrasonography was suspicious, but only 0.6% if ultrasound result was normal.[27] Therefore, the combination of US-FNAC and high-resolution ultrasound could reduce the rate of missed diagnosis. However, if the nature of the thyroid nodules still remains unclear, certain molecular thyroid cancer markers (such as mutations in RAS, RET/PTC, PAX8/PPARgamma, v-raf, mouse sarcoma filter viral oncogene homolog B1 [BRAF]) can be detected in the biopsy specimens. The detection of these gene mutations in clinical practice has been found to improve the diagnosis rate.[30] However, such genetic testing needs to be performed in specialized medical testing institutions with special methods, which can be expensive[31] and difficult to implement, particularly for primary medical institutions. Therefore, the detection of molecular markers of thyroid nodule lesions is not a cost-effective solution.

The British Thyroid Association guidelines recommend a multidisciplinary discussion when the clinical, cytological, and ultrasound findings are inconsistent.[32] Therefore, in the present study, we combined the high-resolution ultrasound TI-RADS grade and Bethesda classification grade in patients with thyroid nodules for a more comprehensive evaluation, and compared the outcomes with the pathology results after surgery. We found that the combined method of diagnosis has better accuracy than the individual Bethesda classification or T-RADS classification systems.

We observed that a score of 4 with the ultrasound TI-RADS classification exhibited the highest Kappa value and accuracy (0.651 and 87.1%, respectively). With regard to the Bethesda classification system, a score of 3 exhibited the highest Kappa value and accuracy (0.78 and 91.10%, respectively). A score of 7 with the combined system exhibited the highest Kappa value and accuracy (0.893 and 95.90%, respectively; Table 3). By
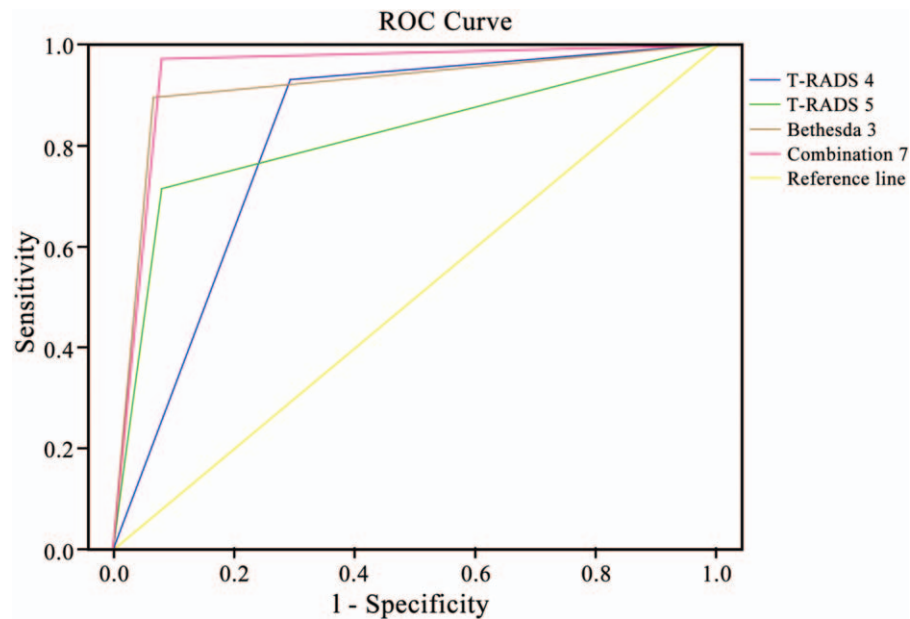
Figure 2. Receiver-operating curves of ultrasound in the evaluation of Bethesda class thyroid nodules. ROC = receiver-operating characteristic.

comparing the Kappa value and accuracy of these 3 methods, we concluded that the combined method is more accurate than the individual Bethesda classification or T-RADS classification systems.

Furthermore, the ROC curve was used to compare the strengths and weaknesses of the different methods for the diagnosis of the same disease. The area under the ROC curve reflects the overall accuracy of a test. A larger area is associated with a higher accuracy. The area under the ROC curve of the 3 diagnostic methods with the highest Kappa value and accuracy was also compared, and the combined score of 7 indicated the largest area under the ROC curve and the highest diagnostic accuracy (Table 4 and Fig. 2).

Thus, a major finding of this study is that if thyroid nodules are strongly suspected of malignancy (ie, TI-RADS grade 5) on high-resolution ultrasound, the potential for malignancy of approximately 94% should be carefully considered even if the US-FNAC biopsy specimen indicates benign disease or cannot be evaluated; in fact, the accuracy of qualitative diagnosis for malignant thyroid nodules can be as high as 94% to 96%. Moreover, when the thyroid nodules of the patient exhibit 2 to 3 malignant signs on high-resolution ultrasound examination, the potential for malignancy should be carefully considered, even if US-FNAC examination with the Bethesda system indicates grades 3 to 4, and there is no indication of malignancy or suspected malignancy. The results of present study indicated that the combination of the 2 methods is helpful to give clinicians a better reference for the diagnosis and treatment of thyroid nodules. According to our results, if the combination score ≥7, the false-positive rate is about 8.00%. However, it does not mean that patients with thyroid nodules who have low incidence of malignancy will receive unnecessary surgery because clinicians will make clinical decisions based on the reports of radiologists and the actual situation. Due to the variety of US-FNAC performance of the thyroid nodules and the possibility of its false-negative results, current America Thyroid Association guidelines recommend ultrasonographic follow-up instead of direct surgery in patients with suspected benign thyroid nodules.[32]

Through a review of the literature on the assessment and treatment of thyroid nodules during the past 3 years in PubMed and Scopus, Cosimo et al concluded that the risk models for malignant thyroid nodules were primarily based on the ultrasound manifestations of thyroid nodules. Thus, the initial indication with FNAC can be used for better guidance and during follow-up.[33] Accordingly, based on our findings, the follow-up of thyroid nodules should be performed as follows: if thyroid ultrasonography is performed and the TI-RADS grade is ≥4 (score ≥3), fine-needle puncture (US-FNAC) should be performed under ultrasound guidance; if the Bethesda classification grade is grade 1 to 2 (score ≤1), outpatient follow-up should be performed; if the Bethesda classification grade is ≥3 (score ≥2), surgery should be performed. Our findings are consistent with the recommendations of Grani et al.[34]

The present study has certain limitations. First, due to acceptable ethical reasons, we did not surgically confirm all nodules classified as TI-RADS scores ≤3. Most of the patients undergoing surgery and US-FNAC included those with TI-RADS grades ≥4B (75.8% of all cases), and only a few cases (24.2% of all cases) had TI-RADS grades <4B (score ≤3). Hence, the limited number of cases may affect the sensitivity, false-negative rate, and negative predictive value. This bias is partially compensated by the sample size of the present study. Second, stratified analysis was not performed according to nodule size. Third, most of the cases in this study involved single nodules and only few cases involved double nodules (263 single nodules and 32 double nodules). The risk of malignant tumors was not estimated in both single and double nodules. Finally, only 2 medical centers were involved in the present study, and hence, future studies should prospectively include nonspecialized members from multiple centers, if possible. Moreover, without violating any ethics, more patients with TI-RADS classification score <4B (≤3 points) should be included in the study to obtain more meaningful results.

In the differential diagnosis of benign and malignant thyroid nodules, qualitative diagnosis can lead to misdiagnosis and missed diagnosis, due to the use of a single standard based on ultrasound TI-RADS grading. Moreover, there is a great amount of uncertainty with Bethesda classification diagnosis. Therefore, a combined diagnosis system could improve the accuracy of diagnosing malignant thyroid nodules.

Acknowledgments

Author contributions

Conceptualization: Huiwen Tan, Zhihui Li, Nong Li, Huajun Xu, Zhongxing Li.

Data curation: Huiwen Tan, Zhihui Li, Jianrong Qian, Fengchun Fan, Huiling Zhong, Huajun Xu, Zhongxing Li.

Formal analysis: Huiwen Tan, Zhihui Li, Jianrong Qian, Fengchun Fan, Huiling Zhong, Jinquan Feng.

Funding acquisition: Nong Li.

Investigation: Huiwen Tan, Zhihui Li, Nong Li, Jianrong Qian, Fengchun Fan, Huajun Xu, Zhongxing Li.

Methodology: Huiwen Tan, Zhihui Li, Nong Li, Jianrong Qian, Fengchun Fan, Huiling Zhong, Jinquan Feng, Huajun Xu, Zhongxing Li.

Project administration: Nong Li, Jianrong Qian.

Resources: Zhihui Li, Nong Li, Jinquan Feng, Zhongxing Li.

Software: Huiwen Tan, Fengchun Fan, Huiling Zhong, Jinquan Feng, Huajun Xu, Zhihui Li.

Supervision: Nong Li, Zhihui Li.

Validation: Nong Li, Fengchun Fan, Jinquan Feng.

Visualization: Nong Li, Huiling Zhong, Jinquan Feng, Huajun Xu, Zhongxing Li.

Writing – original draft: Huiwen Tan, Zhihui Li.

Writing – review & editing: Huiwen Tan, Nong Li.

Huiwen Tan orcid: 0000-0002-0451-8283.

References

[1] Haugen BR, Alexander EK, Bible KC, et al. 2015 American Thyroid Association Management Guidelines for adult patients with thyroid nodules and differentiated thyroid cancer. Thyroid 2016;26:1–33.

[2] Mazzaferri EL. Management of a solitary thyroid nodule. N Engl J Med 1993;328:553–9.

[3] Gharib H, Papini E, Paschke R, et al. AACE/AME/ETA Task Force on Thyroid Nodules. American Association of Clinical Endocrinologists, Associazione Medici Endocrinologi, and European Thyroid Association Medical Guidelines for Clinical Practice for the Diagnosis and Management of Thyroid Nodules. Endocr Pract 2010;16:1–43.

[4] Shan Z, Chen L, Lian X, et al. Iodine status and prevalence of thyroid disorders after introduction of mandatory universal salt iodization for 16 years in China: a cross-sectional study in 10 cities. Thyroid 2016;26:1125–30.

[5] Wang N, Zhou Y, Fu C, et al. The association of thyroid nodule with Non-iodized salt among Chinese children. PLoS One 2014;9:e102726.

[6] Guth S, Theune U, Aberle J, et al. Very high prevalence of thyroid nodules detected by high frequency (13 MHz) ultrasound examination. Eur J Clin Invest 2009;39:699–706.

[7] Kwong N, Medici M, Angell TE, et al. The influence of patient age on thyroid nodule formation, multinodularity, and thyroid cancer risk. J Clin Endocrinol Metab 2015;100:4434–40.

[8] Pacini F, Vorontsova T, Demidchik EP, et al. Post-Chernobyl thyroid carcinoma in Belarus children and adolescents: comparison with naturally occurring thyroid carcinoma in Italy and France. J Clin Endocrinol Metab 1997;82:3563–9.

[9] Lai S, Page JB, Lai H. Solid cancers after bone marrow transplantation. N Engl J Med 1997;336:897–904.

[10] Hegedfis L. Clinical practice. The thyroid nodule. N Engl J Med 2004;351:1764–71.

[11] McLeod DS, Watters KF, Carpenter AD, et al. Thyrotropin and thyroid cancer diagnosis: a systematic review and dose. response meta analysis. J Clin Endocrinol Metab 2012;97:2682–92.

[12] Kwak JY, Han KH, Yoon JH, et al. Thyroid imaging reporting and data system for US features of nodules: a step in establishing better stratification of cancer risk. Radiology 2011;260:892–9.

[13] Baskin HJ. Ultrasound of Thyroid Nodules. Boston, MA: Kluwer Academic Publishers; 2000. 12(4):71–86.

[14] Shweel M, nansour E. Diagnostic performance of combined elastosonographyscoring and high-resolution uhrasonography for the differentiation of benign and malignant thyroid nodules. Eur J Radiol 2013;82:995–1001.

[15] Rago T, Santini F, Scutari M, et al. Elastography: new developments in ultrasound for predicting malignancy in thyroid nodules. J Clin Endocrinol Metab 2007;92:2917–22.

[16] Cibas ES, Ali SZ. The Bethesda System for reporting thyroid cytopathology. Thyroid 2009;19:1159–65.

[17] Ali SZ, Cibas ES. The Bethesda System for reporting thyroid cytopathology II. Acta Cytol 2016;60:397–8.

[18] Cibas ES, Ali SZ. The 2017 Bethesda System for reporting thyroid cytopathology. Thyroid 2017;27:1341–6.

[19] Kundel HL, Polansky M. Measurement of observer agreement. Radiology 2003;228:303–8.

[20] Landis JR, Koch GG. The measurement of observer agreement for categorical data. Biometrics 1977;33:157–74.

[21] Horvath E, Majlis S, Rossi R, et al. An ultrasonogram reportin system for thyroid nodules stratifying cancer risk for clinical management. J Clin Endocrinal Metab 2009;94:1748–51.

[22] Cheng SP, Lee JJ, Lin JL, et al. Characterization of thyroid nodules using the proposed thyroid imaging reporting and data system (TI-RADS). Head Neck 2013;35:541–7.

[23] Park JY, Lee HJ, Jang HW, et al. A proposal for a thyroid imaging reporting and data system for ultrasound features of thyroid carcinoma. Thyroid 2009;19:1257–64.

[24] Tessler FN, Middleton WD, Grant EG, et al. ACR Thyroid Imaging, Reporting and Data System(TI-RADS): white paper of the ACR TI-RADS Committee. J AmColl Radiol 2017;14:587–95.

[25] Liu Z, Huang T. Papillary thyroid microcarcinoma: an over-treated malignancy? World J Surg 2016;40:764–5.

[26] Durante C, Costante G, Lucisano G, et al. The natural history of benign thyroid nodules. JAMA 2015;313:926–35. 28.

[27] Kwak JY, Koo H, Youk JH, et al. Value of US correlation of a thyroid nodule with initially benign cytologic results. Radiology 2010;254:292–300.

[28] Haugen BR, Alexander EK, Bible KC, 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;26:1–33.

[29] Nikiforov YE, Steward DL, Robinson-Smith TM, et al. Molecular testing for mutations in improving the fine-needle aspiration diagnosis of thyroid nodules[J]. J Clin Endocrinol Metab 2009;94:2092–8.

[30] Nishino M. Molecular cytopathology forthyroid nodules: a review of methodology and test performance. Cancer Cytopathol 2016;124:14–27.

[31] Perros P, Boelaert K, Colley S, et al. British Thyroid AssociationGuidelines for the management of thyroid cancer. Clin Endocrinol (Oxf) 2014;81(suppl 1):1–22.

[32] Durante C, Grani G, Lamartina L, et al. The diagnosis and management of thyroid nodules: a review. JAMA 2018;319:914–24.

[33] Gao LY, Wang Y, Jiang YX, et al. Ultrasound is helpful to differentiate Bethesda class III thyroid nodules: A PRISMA-compliant systematic review and meta-analysis. Medicine (Baltimore) 2017;96:e6564.

[34] Grani G, Lamartina L, Ascoli V, et al. ltrasonography scoring systems can rule out malignancy in cytologically indeterminatethyroid nodules. Endocrine 2017;57:256–61.