



Ultrasound imaging classifications of thyroid nodules for malignancy risk stratification and clinical management: state of the art

Chiara Floridi¹, Michaela Cellina², Giorgio Buccimazza³, Antonio Arrichiello³, Andrea Sacrini¹, Francesco Arrigoni⁴, Giovanni Pompili¹, Antonio Barile⁴, Gianpaolo Carrafiello³

¹Unit of Diagnostic and Interventional Radiology, Department of Diagnostic Services, ASST Santi Paolo e Carlo, University of Milan, Milan, Italy;

²Department of Radiology, ASST Fatebenefratelli Sacco, Milan, Italy; ³Department of Health Sciences, Diagnostic and Interventional Radiology, San Paolo Hospital, University of Milan, Milan, Italy; ⁴Department of Biotechnological and Applied Clinical Sciences, University of L'Aquila, L'Aquila, Italy

Contributions: (I) Conception and design: C Floridi, M Cellina; (II) Administrative support: G Buccimazza; (III) Provision of study materials or patients: A Arrichiello; (IV) Collection and assembly of data: G Pompili, A Sacrini; (V) Data analysis and interpretation: F Arrigoni, A Barile; (VI) Manuscript writing: All authors; (VII) Final approval of manuscript: All authors.

Correspondence to: Chiara Floridi. Unit of Diagnostic and Interventional Radiology, Department of Diagnostic Services, ASST Santi Paolo e Carlo, via A. di Rudinì 8, 20142 Milan, Italy. Email: chiara.floridi@gmail.com.

Abstract: Assessing the risk of malignancy in the thyroid with ultrasound (US) is crucial in patients with nodules, as it can aid in selecting those who should have a fine-needle aspiration (FNA) biopsy performed. Many studies have examined whether the US characteristics of thyroid nodules are useful indicators of histological malignancy. Overall, these investigations have identified a few US features that are significantly more frequent in malignant thyroid nodules which can be coalesced into a defining set to be used as an indicator of a higher risk of malignancy. Despite these efforts, none of these classifications have been widely adopted worldwide, and there are still conflicting recommendations from different institutions. Understanding the role and appropriate utilization of these systems could facilitate the effective interpretation and communication of thyroid US findings among referring physicians and radiologists. In this comprehensive review, we outline the major US classification systems of thyroid nodules published in the last few years.

Keywords: Ultrasound (US); malignancy; thyroid nodule; classification

Submitted Jan 28, 2019. Accepted for publication Jun 27, 2019.

doi: 10.21037/gs.2019.07.01

View this article at: <http://dx.doi.org/10.21037/gs.2019.07.01>

Introduction

A thyroid nodule is defined by the American Thyroid Association (ATA) as “a discrete lesion within the thyroid gland that is radiologically distinct from the surrounding thyroid parenchyma” (1).

Thyroid nodules are very common in the general population, and their prevalence is dependent on the identification method used, with a high prevalence found in ultrasound (US) examination, ranging from 20% to 76% in the adult population (2). Non-palpable nodules detected on

US or on other imaging examinations are called “thyroid incidentalomas” or “incidentally discovered nodules”.

Thyroid nodules are more frequent in females, with an incidence 4 times higher in women than in men (3). This distribution can be explained by hormonal influences, as pregnancy is related with the increase in size and number of nodules (4). Thyroid nodules may determine gland dysfunction and compressive symptoms due to mass effect, but the main concern is to rule out their malignancy.

A US exam is a safe, non-invasive, and fast imaging

technique: it is sufficiently sensitive for detecting thyroid nodules and identifying suspicious features and can be used to plan further investigation and management decisions (5,6).

Considering that the average prevalence of thyroid nodule malignancy is variable and ranges from 4.0% to 6.5% (7–9), the accurate estimation of the risk of malignancy on US could help identify those nodules with a high risk of cancer. Moreover, many fine-needle aspirations (FNAs) may also be obviated by identifying those nodules with an acceptably low incidence of malignancy (10).

Many investigations have focused on the US characteristics of thyroid nodules as indicators of nodule malignancy, and several US systems have been developed by an array of authors and institutions to stratify the nodule malignancy risk and provide a standardized language for radiologists and endocrinologists.

Here, we offer an overview of the most recent and recognized US classification systems for thyroid nodules including the differing proposals of the Thyroid Imaging Reporting and Data System (TI-RADS) (9,11–14) along with other guidelines from renowned endocrinological and radiological associations (15–17).

TI-RADS

Many studies have investigated the feasibility of applying the breast imaging reporting and data system (BI-RADS) concept to the US evaluation of thyroid nodules for the purpose of developing a TI-RADS.

In 2009, Horvath (11) was the first to propose TI-RADS as a method to stratify the estimated risk of cancer in thyroid nodules and select those nodules needing to undergo FNA.

US features considered by this classification include echogenicity, microcalcifications, shape, irregular margins, peripheral halo, and presence of suspicious lymph nodes. TI-RADS includes 10 US patterns combined into categories with the increasing risk of malignancy and nodules classified as TI-RADS score 2 to 6 (11).

The rationale of this classification system is that the risk of malignancy rises in parallel with the increase of the number of suspicious US features and with the lack of benign findings.

TI-RADS classification has been validated through a study on 502 nodules belonging to 210 patients conducted by the same Horvath (18) who demonstrated a TI-RADS sensitivity of 99.6% (95% CI: 98.9–100.0%), a specificity of

74.35% (95% CI: 68.70–80.00%), a positive predictive value (PPV) of 82.1% (95% CI: 78.0–86.3%), and a negative predictive value (NPV) of 99.4% (95% CI: 98.3–100.0%) for malignancy. The estimated risk of malignancy was 0% in TI-RADS 2, 3.4% in TI-RADS 3, 10–80% in TI-RADS 4, and 87% in TI-RADS 5.

This system has been part of standardizing thyroid US examination reports and consolidating the disparate terminologies used by radiologists and endocrinologists all over the world.

A modified version of TI-RADS that predicts malignancy risks only according to the number of suspicious US features was proposed by Kwak *et al.* in 2011 (9). These features include solid nodules, hypoechogenicity, microlobulated or irregular margins, microcalcifications or mixed calcification, and taller-than-wide shape.

The risk of malignancy increases as the number of suspicious US signs increases. According to this system, nodules with no suspicious US features have a 0.02–0.028 fitted probability of malignancy, implying that it is safe to follow up with the lesion rather than perform a biopsy. Nodules in categories TI-RADS 4 or 5 with at least 1 suspicious US feature have more than 0.036 fitted probability of malignancy, indicating that a biopsy is warranted.

In a study including 167 nodules from 150 patients (19), the TI-RADS by Kwak *et al.* showed a sensitivity of 94%, a specificity of 28.2%, a PPV of 37.3%, and an NPV of 91.2%; meanwhile, a study on 1,293 nodules (20) showed a sensitivity of 97.4%, a specificity of 29.3%, a PPV of 23.3%, and an NPV of 98.1%.

American College of Radiology (ACR)-TI-RADS

Another version of TI-RADS has been proposed by the ACR (12) in order to identify the most clinically significant malignancies and decrease the number of FNAs on benign nodules.

This classification is not a pattern-based approach but is based on the assessment of different US features of thyroid nodules: composition, echogenicity, shape, margin, and echogenic foci; each of these features is associated with a score ranging from 0 to 3 points.

The sum of the assigned points defines the risk of malignancies according to 5 grades, with each grade corresponding to benign, minimally suspicious, moderately suspicious, or highly suspicious for malignancy. This system does not include subcategories, nor a TR 0 group to

indicate the normal thyroid.

Indications for FNA execution or US follow-up are based on ACR-TI-RADS level and on the maximum diameter of the nodules: for risk grades TR3–TR5, a size threshold at or above which FNA should be performed is indicated. Also, the lower size limits to perform follow-up US for TR3, TR4, and TR5 nodules were established in order to reduce the number of repeat examinations for nodules that are likely to be benign or not clinically significant. Highly suspect nodules are submitted to biopsy only if they are 1 cm or larger, while nodules that have a low risk for malignancy should be further investigated only when they measure 2.5 cm or more.

The main goal of ACR TI-RADS is to balance the benefit of detecting clinically significant malignancies against the risk and cost of submitting benign nodules or indolent and non-aggressive tumours to invasive investigations and treatment. The indications for US follow-up aim at reducing the eventuality that significant lesions remain undetected over time.

In a study which included 1,112 thyroid nodules, this classification system had an area under the receiver operating characteristic curve (AUC) in discrimination of malignant nodules of 0.875 (21), while in another study on a series of 74 nodules from paediatric patients, the AUC was 0.75 (22).

European (EU)-TI-RADS

European (EU)-TI-RADS is a classification system first proposed by Russ *et al.* in 2011 (23), and has been subsequently modified into an easier-to-use version, validated in a large prospective study in 2013 (24) and finally published as a European guideline in 2017 (13).

The latest version of EU-TI-RADS consists of five categories, each which is scored in correspondence to features from the US examination: EU-TI-RADS 1 refers to a US examination where no thyroid nodule is found, while the other 4 categories—nodules benign and with low, intermediate, and high risk—correspond in an increasing risk of malignancy.

EU-TI-RADS 2, with a risk of malignancy close to 0%, includes two patterns: pure/anechoic cysts and entirely spongiform nodules; these nodules should be considered as benign and FNA is not indicated. EU-TI-RADS 3, with a risk of malignancy around 2–4%, consists of nodules with oval shape, smooth margins, isoechoic or hyperechoic features, and a lack of any high-risk features; these nodules

should be submitted to FNA only if >20 mm. EU TI-RADS 3, with a risk of malignancy around 6–17%, consists of nodules with oval shape, smooth margins, mildly hypoechoic features, and a lack of any high-risk features; the evidence of a thin halo, cystic components, comet-tail artefacts, peripheral vascularization, and low stiffness, lower the malignancy risk in this group, while the presence of discontinuous peripheral or rim macrocalcifications, a thick halo, predominantly central vascularization, and high stiffness, increase the risk of malignancy; the threshold for FNA is 15 mm.

EU-TI-RADS 5 involves nodules presenting at least 1 of the following high-risk signs: non-oval shape, irregular margins, microcalcifications, marked hypoechogenicity; the malignancy rate is 26–87% and nodules must be submitted to FNA when >1 cm.

In addition, the authors of EU-TI-RADS noted that extrathyroidal extension with disruption of the capsular margin increases the malignancy risk of a nodule and should be described in the report. On the other hand, hyperechoic spots associated with comet-tail artefacts, along with thin halo, are suggestive of benignity and reduce the suspicion of malignancy.

This classification system also considers extra thyroid US features in stratifying the malignancy risk and the indication to FNA.

Notably, EU-TI-RADS accounts for suspicious lymphadenopathies: in case of a suspicious lymph node of thyroid origin, FNA of the lymph node itself and of the most suspicious thyroid nodule should be implemented.

In other studies, the authors of EU-TI-RADS suggest that stiffness on elastography should be considered as complementary information in nodule assessment (25), even if there is yet no reliable threshold for delineating benign from malignant nodules. They also suggest that FNA should not be indicated on a nodule if thyroid scintigraphy shows this to be warm/hot.

According to the EU-TI-RADS guidelines, the blood flow within the thyroid nodules can be investigated with different US imaging modalities, and the type of vascularization pattern can be assessed and scored from I to III: malignant nodules are usually associated with type III vascularization, while benign nodules are usually associated with type I and II signal patterns. However, EU-TI-RADS emphasizes the fact that vascularity also rises in benign nodules and increases in size, while the presence of type III vascularization shows low sensitivity, specificity, and PPV for malignancy and should be associated with other

suspicious US findings (23). Therefore, the reliability of Color Doppler in the malignancy assessment of thyroid nodules remains controversial: the sensitivity of Doppler depends markedly on the US equipment and settings, the establishment of central vascularity is highly observer-dependent, and a quantitative analysis of the Color signals is still unclear. Given this, the Color Doppler evaluation is not included in the EU-TI-RADS score, but only recommended to be used in differentiating solid tissue from thick colloid.

In a study to assess the reliability of the first version of EU-TI-RADS (26), 430 nodules were examined, and the malignancy risks of the EU-TI-RADS categories were 0% for TI-RADS 2, 2.2% for TI-RADS 3, 5.9–57.9% for TI-RADS 4, and 100% for TI-RADS 5.

Recently, Xu *et al.* (27) observed that the percentage of unnecessary FNAs based on EU-TI-RADS was 25.2% while another study by Grani *et al.* (28) reported a EU-TI-RADS sensitivity of 86.1%, a specificity of 32%, a PPV of 8.9%, and an NPV of 96.7%.

Korean (K)-TI-RADS

K-TI-RADS was proposed in 2016 by Shin *et al.* (14). In this classification system, the malignancy risk estimated by US examination is not determined by a single US sign, but by a combination of them; the rationale is that any single US predictor is not sensitive and specific enough to determine the suspicion of malignancy. Based on US patterns including solidity, echogenicity, and suspicious US characteristics, thyroid nodules in the K-TI-RADS system are classified as high suspicion, intermediate suspicion, low suspicion, and benign.

Evidence of microcalcification, speculated or microlobulated margins, and taller-than-wide shape are considered strong predictors of malignancy, with a high specificity (>80%), while US appearance of solid internal content and hypoechogenicity is considered predictive of malignancy, with an intermediate specificity. The decision to perform FNA should be based on the calculated risk of malignancy according to the US features and nodule size, except in case of the presence of poor prognostic factors, such as suspected lymph node metastasis, extrathyroidal tumour extension, or distant metastasis from thyroid cancer.

It should be noted that nodular vascularization at Color Doppler evaluation and elastography appearance are described as potentially useful in the distinction between

benign and malignant nodules, but these are not included in the classification as further investigations are needed to assess their complementary function in the risk categories.

According to a study on 2000 nodules (29), FNA criteria of K-TI-RADS categories 4 or 5 for nodules ≥ 1 cm and K-TI-RADS 3 for nodules ≥ 1.5 cm, proved to have a sensitivity, specificity, PPV, NPV, and accuracy for overall malignancy of 94.5%, 26.8%, 27.5%, 94.3%, and 42.2%, respectively (30). K-TI-RADS has also been validated by a multicentre prospective study of four institutions comprising 902 thyroid nodules, which proved the significantly higher malignancy risk of solid hypoechoic nodules (73.4%) compared to partially cystic or iso/hyper echoic nodules (4.3–38.5%; $P < 0.001$); this study also demonstrated that the malignancy risk calculated in K-TI-RADS categories 5, 4, 3, and 2 were 73.4%, 19.0%, 3.5%, and 0.0%, respectively; meanwhile, the sensitivity, specificity, PPV, NPV, and accuracy of K TI-RADS for identification of malignant nodules were 95.5%, 58.6%, 44.5%, 96.9%, and 69.5%, respectively.

American Association of Clinical Endocrinologists (AACE), American College of Endocrinology (ACE), and Associazione Medici Endocrinologi (AME) classification

AACE/ACE/AME (15) used information provided by US examination to stratify nodules into a 3-class system based on their risk of malignancy (low, intermediate, and high risk) combining US thyroid nodule features and nodule dimension with patient clinical history. FNA execution is indicated for “high US risk thyroid lesions ≥ 10 mm”; “intermediate US risk thyroid lesions > 20 mm”; “low US risk thyroid lesions only when > 20 mm and increasing in size or associated with a risk history and before thyroid surgery or minimally invasive ablation therapy”; “subcapsular or paratracheal lesions”; “suspicious lymph nodes or extra thyroidal spread”; “positive personal or family history of thyroid cancer”; and “coexistent suspicious clinical findings (e.g., dysphonia)”.

In addition, “nodules with a major diameter < 5 mm should be monitored, rather than biopsied” and “in nodules with a major diameter of 5–10 mm that are associated with suspicious US signs, either FNA or watchful waiting may be considered”.

In a prospective study on 987 thyroid nodules consecutively referred for FNA, whose aim was determining

the diagnostic accuracy of different US classification systems and their reliability for FNA indication (31), the AACE/ACE/AME malignancy rate was 1.1% in low-risk, 4.4% in intermediate-risk, and 54.9% in high-risk US categories of nodules. The sensitivity, specificity, PPV, NPV, and accuracy were 82%, 87%, 54%, 96%, and 86% respectively, with a kappa coefficient of inter-observer agreement of 82%, which was higher than those of other classifications. In another study including 859 FNAs (32), according to AACE/ACE/AME guidelines, 341 (39.7%) were low-, 489 (56.9%) intermediate-, and 29 (3.4%) were high-risk nodules. Of these, 88.5% and 74.9% of low- and intermediate-risk nodules, respectively, were cytologically benign, whereas 84.6% of high-risk nodules had a moderate-to-elevated risk of malignancy or were malignant. They suggested that the biopsy decision be patient-oriented (considering the clinical setting, operator expertise, patient's preference) rather than simply based on nodule size. A further study on 2,000 nodules with analysis of the diagnostic performance of different US guidelines for thyroid nodules (33), AACE/ACE/AME classification showed a specificity of 80.4%.

Despite this research, the management of nodules with intermediate risk at AACE/ACE/AME classification is still controversial. Deng *et al.* (34), in a study on 1,000 patients with 1,000 nodules tried to propose a solution with a risk stratification system for intermediate-risk nodules of AACE/ACE/AME classification, elaborating a predictive model: in univariate analysis, they observed that “younger patient age, smaller nodule maximum diameter, slight hypo-echogenicity, slightly ill-defined margin, no macro-calcification and indeterminate hyper-echoic spot were significantly associated with malignancy” ($P<0.05$).

With multivariate logistic regression analysis, the authors found that indeterminate hyper-echoic spot was the most significant independent risk predictor for intermediate risk, followed by slightly ill-defined margin, slight hyper-echo, and no macro-calcification; the probability of malignancy increased with the number of risk factors increasing (34).

ATA classification

The ATA classification (16) of thyroid nodules is based on echostructure, echogenicity, margins, presence/absence of microcalcifications, and shape. It provides 5-class scoring with a risk for malignancy estimated to each category in the following manner: <1% for benign class, <3% for very-low suspicion, 5–10% for low suspicion, 10–20% for

intermediate suspicion, and >70–90% for high suspicion.

In ATA, classifications completely exclude Color Doppler and elastosonography parameters. A previous study on 1,293 nodules by Yoon *et al.* (20) considered very-low-suspicion nodules as negative, and low- to high-suspicion nodules as positive. The results demonstrated that ATA had a sensitivity of 95.3%, a specificity of 37.4%, a PPV of 25.2%, an NPV of 97.3%, and an accuracy of 47.9%, with a higher specificity, PPV, and accuracy in comparison to Kwak's TI-RADS ($P<0.001$ for all), albeit a lower specificity ($P=0.024$). In another study (19), ATA showed a sensitivity of 98.0%, a specificity of 17.3%, a PPV of 35.0%, and an NPV of 95.0%.

In the study by Persichetti *et al.* (31), ATA achieved a sensitivity of 81%, a specificity of 87%, a PPV of 54%, an NPV of 96%, and an accuracy of 86%; the interobserver agreement was 76.9%.

In the study by Grani *et al.* (28), the ATA classification system achieved a sensitivity of 75%, a specificity of 45.3%, a PPV of 9.6%, and an NPV of 95.9%.

In a further study including 112 children totalling 145 thyroid nodules (31), which compared radiologists' impressions and ATA risk stratification to histology and cytology results, the sensitivity in identifying malignancy was comparable between ATA guidelines (91%) and the radiologists' impression (90%), while the specificity of the radiologists' malignant impression (80%) was better than the ATA guidelines (54%).

Another study by Pandya *et al.*, which examined 1,947 thyroid nodules (35), was performed to identify the limitations of ATA guidelines and demonstrated that PPV for cancer increased by ATA category 1 to 5 with measures of 0%, 2%, 5%, 12%, and 28%, respectively; this study stated that the ATA classification system was optimistic about cancer detection rates for the higher-risk nodules and that the inter-rater agreement among the 5 readers was fair (intraclass correlation coefficient = 0.460).

British Thyroid Association (BTA) US classification of thyroid nodules

Another existing classification for the US of thyroid nodules is that of the BTA (17), which consists of five categories: normal (U1), benign (U2), equivocal/indeterminate (U3), suspicious (U4), and malignant (U5). It considers the following US features as predictors of malignancy: eccentric location of the solid portion in partially cystic nodules, non-

smooth margins, hypoechogenicity of the solid portion, microcalcification, and taller-than-wide shape.

In a study on 104 nodules, with only U2 being classified as a negative test and only U5 being classified as a positive test, the sensitivity, specificity, positive likelihood ratio, negative likelihood ratio, PPV, NPV, and accuracy of BTA guidelines were 100%, 91.4%, 11.7%, 0.0%, 78.6%, 100% and 93.5%, respectively (36). In the previously cited study by Chang *et al.* (19), BTA demonstrated a sensitivity of 90%, a specificity of 50.9%, a PPV of 45.5%, and an NPV of 91.8%, while in the study by Persichetti *et al.* (31), sensitivity was 74%, specificity was 92%, PPV was 62%, NPV was 95%, and accuracy was 89%. In a study on 47 patients for correctly identifying nodules requiring further investigation, BTA guidelines showed a sensitivity of 56%, a specificity of 81%, a PPV of 81%, and an NPV of 56% (37).

Comparison between the different US classifications for thyroid nodules

It is critical to assess the risk of malignancy in the thyroid with US in order to properly identify those patients for whom FNA biopsy is a necessary priority. Many studies have investigated whether the US characteristics of thyroid nodules are useful indicators of histological malignancy. Overall, these investigations have identified a few US features that are significantly more frequent in malignant than in benign thyroid nodules, thus defining a set of US features at a higher risk of malignancy. The main features of these classifications are summarized on *Table 1*.

Despite these efforts, none of these classifications have been widely adopted worldwide, and there are still conflicting recommendations from several different societies. However, in the last few years, several studies have compared the different US risk stratifications systems in large cohorts of patients with thyroid nodules in order to evaluate the most accurate and useful system. The major results of studies published in the last year are listed below.

In the study of Xu *et al.*, which included a total of 1,460 benign and 1,005 malignant nodules (27), the distribution of malignancies among ACR-TI-RADS levels was 0%, 2.3%, 7.5%, 40.1%, and 81.4%, respectively. ACR-TI-RADS showed the best sensitivity (96.6%) in comparison to EU-TI-RADS and K-TI-RADS, but the lowest specificity (52.9%) and AUC (0.846). K-TI-RADS had good specificity (87.4%) but had the lowest (71.4%) sensitivity.

In a study which included two observers and 1,055 images of thyroid nodules (38), EU-TI-RADS proved to have the higher rate of inter-observer agreement in the first set of nodules and for indication to perform FNA in comparison to the AACE/ACE/AME and ATA guidelines, in addition to ACR-TI-RADS and K-TI-RADS.

Middleton *et al.* (39) evaluated 3,422 thyroid nodules for which pathologic findings were available according to the ACR-TI-RADS, the K-TI-RADS, and the ATA guidelines. The rate of malignant nodules was 89.2% according to the ACR-TI-RADS, which was the best performance in comparison to the other classification systems. The percentage of benign nodules that would be biopsied was 47.1% for the ACR-TI-RADS, 79.7% for the K-TI-RADS, and 78.1% for the ATA guidelines.

In a study including 902 nodules, whose aim was to compare the diagnostic performance of ATA guidelines with K-TI-RADS and ACR-TI-RADS (40), the calculated malignancy risks for nodules of categories 5, 4, 3, 2, and 1 were 71.7%, 21.5%, 2.6%, 3.8%, and 0%, respectively. The ATA guidelines in this study demonstrated significantly higher diagnostic sensitivity (95.0%) than the ACR-TI-RADS (80.2%; $P=0.001$), but a lower specificity (38.1% versus 68.9%; $P<0.001$); meanwhile, when compared to K-TI-RADS, the ATA guidelines showed a lower diagnostic sensitivity (100.0%; $P=0.07$) but a higher specificity (28.2%; $P<0.001$). The unnecessary FNA rate was the lowest according to the ACR system (25.8%), followed by the ATA (51.2%) and K (59.4%) guidelines.

Finally, in the most recent study on 477 patients (28), Grani *et al.* observed that ACR-TI-RADS proved to be the most effective system when compared to other systems considered in the analysis (EU-TI-RADS, AACE/ACE/AME, K-TI-RADS), and allowed for the largest reduction of biopsy numbers (268 of 502) with the lowest false negative rate of 2.2% (NPV, 97.8%; 95% CI: 95.2–99.2%), a sensitivity of 83.3%, a specificity of 56.2%, and a PPV of 12.8%.

Figures 1–3 summarizes some features of the different US classifications.

In conclusion, according to different studies, all US malignancy risk stratification systems demonstrated high sensitivity and specificity, but unfortunately, the results obtained are not yet sufficient to establish the best system. Prospective studies and larger cohorts are needed to standardize a US classification for application in routine clinical practice.

Table 1 Thyroid nodule scores summary

Classification	US features				Size	ECD features	Lymph node evaluation	Categories	FNA recommendations
	Echogenicity	Shape	Margins	Echogenic foci (calcification)					
TIRADS (11)	√	√	√	√	x	Included	x	TI-RADS 1: normal thyroid gland TI-RADS 2: benign conditions TI-RADS 3: probably benign nodules TI-RADS 4: suspicious nodules TI-RADS 5: probably malignant nodules TI-RADS 6: malignant nodules	FNA recommended from TI-RADS 3
ACR-TIRADS (12)	√	√	√	√	√	Included	x	ACR-TI-RADS 1: benign ACR-TI-RADS 2: not suspicious ACR-TI-RADS 3: mildly suspicious ACR-TI-RADS 4: moderately suspicious ACR-TI-RADS 5: high suspicious	FNA recommended from mildly suspicious, according on nodule size
EU-TIRADS (13)	√	√	√	√	x	Not included	√	EU-TI-RADS 1: normal EU-TI-RADS 2: benign EU-TI-RADS 3: low risk EU-TI-RADS 4: intermediate risk EU-TI-RADS 5: high risk	FNA recommended considering US scoring + clinical setting

Table 1 (continued)

Table 1 (continued)

Classification	US features				Size	ECD features	Lymph node evaluation	Categories	FNA recommendations
	Echogenicity	Shape	Margins	Echogenic foci (calcification)					
K-TI-RADS (14)	√	√	√	√	√	Not included	×	K-TI-RADS 1: no nodule K-TI-RADS 2: benign K-TI-RADS 3: low suspicion K-TI-RADS 4: intermediate suspicion K-TIRADS 5: high suspicion	FNA recommended from benign lesion, according on nodule size
AACE/ACE/AME (15)	√	√	√	√	×	Included	×	1. Low risk lesion 2. Intermediate risk lesion 3. High risk lesion	FNA recommended from low risk lesion, according on nodule size and patient history
ATA (16)	√	√	√	√	×	Not included	×	1. Benign 2. Very low suspicion 3. Low suspicion 4. Intermediate suspicion 5. High suspicion	N/A
BTA (17)	√	√	√	√	×	Included	√	U1: normal U2: benign U3: intermediate/equivocal U4. suspicious U5 malignant	N/A

TI-RADS, Thyroid Imaging Reporting and Data System; ACR-TI-RADS, American College of Radiologists-TI-RADS; EU-TI-RADS, European-TI-RADS; K-TI-RADS, Korean-TI-RADS; AACE/ACE/AME, American Association of Clinical Endocrinologists/American College of Endocrinology/Associazione Medici Endocrinologi; ATA, American Thyroid Association; BTA, British Thyroid Association.

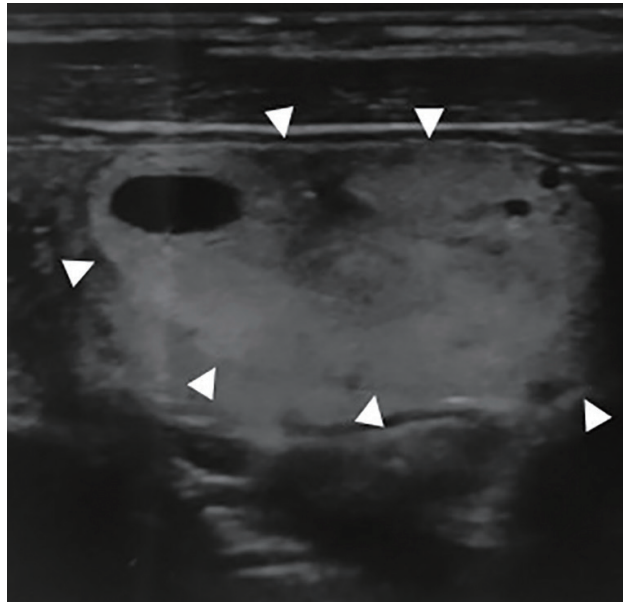


Figure 1 Transverse ultrasonography shows an isoechoic and hyperechoic mixed solid-cystic nodule with well-defined margins (arrowheads). In AACE/ACE/AME classification, this nodule is grade 1 (low-risk lesion); in TI-RADS by Horvath, the lesion is Colloid Type 1 (TI-RADS 2); in TI-RADS by Kwak, it is TI-RADS 3 (not suspicious US features); in ACR-TI-RADS, it is TR 2 (not suspicious); in ATA classification, it is a nodule with low suspicion of malignancy; in EU-TI-RADS, it is grade 3 (low risk); in K-TI-RADS, it is a lesion with low suspicion of malignancy; in BTA classification, the lesion is grade U2 (benign). According to the US classifications, the patient did not undergo to fine-needle biopsy. AACE, American Association of Clinical Endocrinologists; ACE, American College of Endocrinology; AME, Associazione Medici Endocrinologi; TI-RADS, Thyroid Imaging Reporting and Data System; ACR-TI-RADS, American College of Radiology (ACR)-TI-RADS; EU-TI-RADS, European (EU)-TI-RADS; K-TI-RADS, Korean (K)-TI-RADS; BTA, British Thyroid Association; ATA, American Thyroid Association; US, ultrasound.

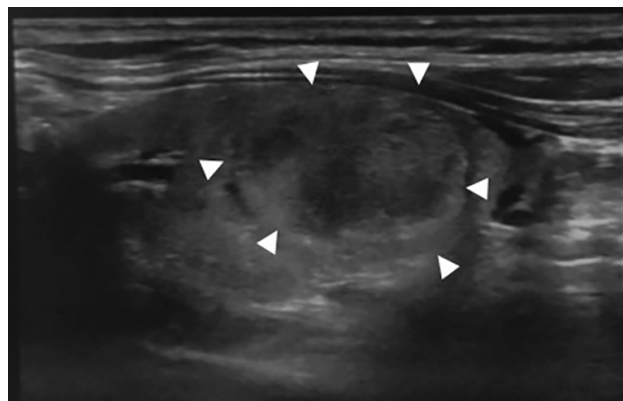


Figure 2 A longitudinal ultrasonography image shows a single, iso-hypoechoic, thyroid nodule with regular hypoechoic margins (arrowheads) and a taller-than-wide shape. In AACE/ACE/AME classification, this nodule is grade 2 (intermediate risk lesion); in TI-RADS, it is TI-RADS 4a (undetermined); in TI-RADS by Kwak, it is TI-RADS 4b (2 suspicious US features); in ACR-TI-RADS, it is TR 4 (moderately suspicious); in ATA classification, the nodule has intermediate suspicion of malignancy; in EU-TI-RADS, it is grade 4 (intermediate risk); in K-TI-RADS, it is a nodule with intermediate suspicion of malignancy; in BTA classification, it is grade U3 (intermediate/equivocal). The fine-needle aspiration cytology confirmed a benign nodule. AACE, American Association of Clinical Endocrinologists; ACE, American College of Endocrinology; AME, Associazione Medici Endocrinologi; TI-RADS, Thyroid Imaging Reporting and Data System; ACR-TI-RADS, American College of Radiology (ACR)-TI-RADS; EU-TI-RADS, European (EU)-TI-RADS; K-TI-RADS, Korean (K)-TI-RADS; BTA, British Thyroid Association; ATA, American Thyroid Association; US, ultrasound.

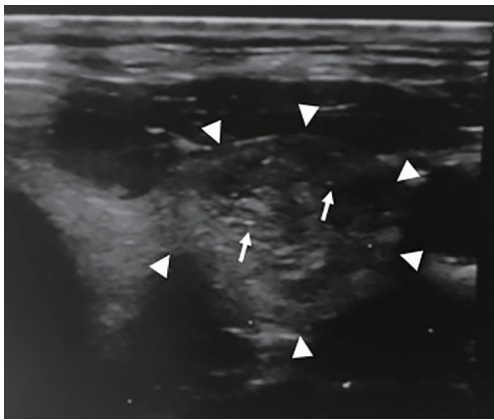


Figure 3 A transverse ultrasonography shows a single, hypoechoic, solid nodule with irregular margins (arrowheads), containing some fine areas of hyperechogenicity (arrows) without any comet-tail. In AACE/ACE/AME classification, this nodule is grade 3 (high-risk lesion); in TI-RADS by Horvath, it is TI-RADS 4a (undetermined) because there are no other nodules; in TI-RADS by Kwak, it is TI-RADS 4c (3–4 suspicious US features); in ACR-TI-RADS, it is TI-RADS 3 (mildly suspicious); in ATA classification, it is a nodule with high suspicion of malignancy; in EU-TI-RADS, it is grade 5 (high risk); in K-TIRADS, the nodule has high suspicion of malignancy; in BTA classification, it is U4 (suspicious). The fine-needle aspiration cytology confirmed a papillary carcinoma. AACE, American Association of Clinical Endocrinologists; ACE, American College of Endocrinology; AME, Associazione Medici Endocrinologi; TI-RADS, Thyroid Imaging Reporting and Data System; ACR-TI-RADS, American College of Radiology (ACR)-TI-RADS; EU-TI-RADS, European (EU)-TI-RADS; K-TI-RADS, Korean (K)-TI-RADS; BTA, British Thyroid Association; ATA, American Thyroid Association; US, ultrasound.

Acknowledgments

None.

Footnote

Conflicts of Interest: 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.

References

1. American Thyroid Association (ATA) Guidelines Taskforce on Thyroid Nodules and Differentiated Thyroid Cancer. Cooper DS, Doherty GM, et al. Revised American Thyroid Association management guidelines for patients with thyroid nodules and differentiated thyroid cancer. *Thyroid* 2009;19:1167-214.
2. Periakaruppan G, Seshadri KG, Vignesh Krishna GM, et al. Correlation between Ultrasound-based TIRADS and Bethesda System for Reporting Thyroid-cytopathology: 2-year Experience at a Tertiary Care Center in India. *Indian J Endocrinol Metab* 2018;22:651-5.
3. Mazzaferri EL. Management of a solitary thyroid nodule. *N Engl J Med* 1993;328:553-9.
4. Kung AW, Chau MT, Lao TT, et al. The effect of pregnancy on thyroid nodule formation. *J Clin Endocrinol Metab* 2002;87:1010-4.
5. De Bernardi IC, Floridi C, Muollo A, et al. Vascular and interventional radiology radiofrequency ablation of benign thyroid nodules and recurrent thyroid cancers: literature review. *Radiol Med* 2014;119:512-20.
6. Barile A, Quarchioni S, Bruno F, et al. Interventional radiology of the thyroid gland: critical review and state of the art. *Gland Surg* 2018;7:132-46.
7. de Matos PS, Ferreira AP, Ward LS. Prevalence of papillary microcarcinoma of the thyroid in Brazilian autopsy and surgical series. *Endocr Pathol* 2006;17:165-73.
8. Kovács GL, Gonda G, Vadász G, et al. Epidemiology of thyroid microcarcinoma found in autopsy series conducted in areas of different iodine intake. *Thyroid* 2005;15:152-7.
9. 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.
10. Pompili GG, Tresoldi S, Ravelli A, et al. Use of the ultrasound-based total malignancy score in the management of thyroid nodules. *Ultrasonography* 2018;37:315-22.
11. Horvath E, Majlis S, Rossi R, et al. An ultrasonogram reporting system for thyroid nodules stratifying cancer risk for clinical management. *J Clin Endocrinol Metab* 2009;94:1748-51.
12. 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 Am Coll*

- Radiol 2017;14:587-95.
13. Russ G, Bonnema SJ, Erdogan ME, et al. European Thyroid Association Guidelines for Ultrasound Malignancy Risk Stratification of Thyroid Nodules in Adults: The EU-TIRADS. *Eur Thyroid J* 2017;6:225-37.
 14. Shin JH, Baek JH, Chung J, et al. Ultrasonography diagnosis and imaging-based management of thyroid nodules: Revised Korean Society of Thyroid Radiology consensus statement and recommendations. *Korean J Radiol* 2016;17:370-95.
 15. Gharib H, Papini E, Garber JR, et al. 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.
 16. 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;26:1-133.
 17. Perros P, Boelaert K, Colley S, et al. Guidelines for the management of thyroid cancer. *Clin Endocrinol (Oxf)* 2014;81 Suppl 1:1-122.
 18. Horvath E, Silva CF, Majlis S, et al. Prospective validation of the ultrasound based TIRADS (Thyroid Imaging Reporting And Data System) classification: results in surgically resected thyroid nodules. *Eur Radiol* 2017;27:2619-28.
 19. Chng CL, Tan HC, Too CW, et al. Diagnostic performance of ATA, BTA and TIRADS sonographic patterns in the prediction of malignancy in histologically proven thyroid nodules. *Singapore Med J* 2018;59:578-83.
 20. Yoon JH, Lee HS, Kim EK, et al. Malignancy Risk Stratification of Thyroid Nodules: Comparison between the Thyroid Imaging Reporting and Data System and the 2014 American Thyroid Association Management Guidelines. *Radiology* 2016;278:917-24.
 21. Ha SM, Ahn HS, Baek JH, et al. Validation of Three Scoring Risk-Stratification Models for Thyroid Nodules. *Thyroid* 2017;27:1550-7.
 22. Lim-Dunham JE, Toslak IE, Reiter MP, et al. Assessment of the American College of Radiology Thyroid Imaging Reporting and Data System for Thyroid Nodule Malignancy Risk Stratification in a Pediatric Population. *AJR Am J Roentgenol* 2019;212:188-94.
 23. Russ G, Bigorgne C, Royer B, et al. The Thyroid Imaging Reporting and Data System (TIRADS) for ultrasound of the thyroid. *J Radiol* 2011;92:701-13.
 24. Russ G, Royer B, Bigorgne C, et al. Prospective evaluation of thyroid imaging reporting and data system on 4,550 nodules with and without elastography. *Eur J Endocrinol* 2013;168:649-55.
 25. Cantisani V, Grazhdani H, Drakonaki E, et al. Strain US Elastography for the Characterization of Thyroid Nodules: Advantages and Limitation. *Int J Endocrinol* 2015;2015:908575.
 26. Moifo B, Takoeta EO, Tambe J, et al. Reliability of thyroid imaging reporting and data system (TIRADS) classification in differentiating benign from malignant thyroid nodules. *Open J Radiol* 2013;3:103-7.
 27. Xu T, Wu Y, Wu RX, et al. Validation and comparison of three newly-released Thyroid Imaging Reporting and Data Systems for cancer risk determination. *Endocrine* 2019;64:299-307.
 28. Grani G, Lamartina L, Ascoli V, et al. Reducing the Number of Unnecessary Thyroid Biopsies While Improving Diagnostic Accuracy: Toward the "Right" TIRADS. *J Clin Endocrinol Metab* 2019;104:95-102.
 29. Ha EJ, Moon WJ, Na DG, et al. A Multicenter Prospective Validation Study for the Korean Thyroid Imaging Reporting and Data System in Patients with Thyroid Nodules. *Korean J Radiol* 2016;17:811-21.
 30. Creo A, Alahdab F, Al Nofal A, et al. Ultrasonography and the American Thyroid Association Ultrasound-Based Risk Stratification Tool: Utility in Pediatric and Adolescent Thyroid Nodules. *Horm Res Paediatr* 2018;90:93-101.
 31. Persichetti A, Di Stasio E, Guglielmi R, et al. Predictive Value of Malignancy of Thyroid Nodule Ultrasound Classification Systems: A Prospective Study. *J Clin Endocrinol Metab* 2018;103:1359-68.
 32. Negro R, Greco G, Colosimo E. Ultrasound Risk Categories for Thyroid Nodules and Cytology Results: A Single Institution's Experience after the Adoption of the 2016 Update of Medical Guidelines by the American Association of Clinical Endocrinologists and Associazione Medici Endocrinologi. *J Thyroid Res* 2017;2017:8135415.
 33. Ha EJ, Na DG, Baek JH, et al. US Fine-Needle Aspiration Biopsy for Thyroid Malignancy: Diagnostic Performance of Seven Society Guidelines Applied to 2000 Thyroid Nodules. *Radiology* 2018;287:893-900.
 34. Deng XH, Tang LN, Liu SQ, et al. A Proposal to Stratify the Intermediate-Risk Thyroid Nodules According to the AACE/ACE/AME Guidelines with Ultrasound Features.

- Sci Rep 2017;7:17901.
35. Pandya A, Caoili EM, Jawad-Makki F, et al. Limitations of the 2015 ATA Guidelines for Prediction of Thyroid Cancer: A Review of 1947 Consecutive Aspirations. *J Clin Endocrinol Metab* 2018;103:3496-502.
 36. Farihah AG, Nurismah MI, Husyairi H, et al. Reliability of the ultrasound classification system of thyroid nodules in predicting malignancy. *Med J Malaysia* 2018;73:9-15
 37. Brimiouille M, Al-Lami A, Marzouk S, et al. Ensuring patient safety when implementing a new diagnostic pathway for thyroid nodules. *Ann R Coll Surg Engl* 2018;100:366-70.
 38. Grani G, Lamartina L, Cantisani V, et al. Interobserver agreement of various thyroid imaging reporting and data systems. *Endocr Connect* 2018;7:1-7.
 39. Middleton WD, Teefey SA, Reading CC, et al. Comparison of Performance Characteristics of American College of Radiology TI-RADS, Korean Society of Thyroid Radiology TIRADS, and American Thyroid Association Guidelines. *AJR Am J Roentgenol* 2018;210:1148-54.
 40. Ha EJ, Na DG, Moon WJ, et al. Diagnostic Performance of Ultrasound-Based Risk-Stratification Systems for Thyroid Nodules: Comparison of the 2015 American Thyroid Association Guidelines with the 2016 Korean Thyroid Association/Korean Society of Thyroid Radiology and 2017 American Congress of Radiology Guidelines. *Thyroid* 2018;28:1532-7. Erratum in: Correction to: Diagnostic Performance of Ultrasound-Based Risk-Stratification Systems for Thyroid Nodules: Comparison of the 2015 American Thyroid Association Guidelines with the 2016 Korean Thyroid Association/Korean Society of Thyroid Radiology and 2017 American College of Radiology Guidelines Ha EJ, Na DG, Moon W-J, Lee YH, and Choi N *Thyroid* 2018;28:1532-7. DOI: 10.1089/thy.2018.0094. [Thyroid. 2019].

Cite this article as: Floridi C, Cellina M, Buccimazza G, Arrichiello A, Sacrini A, Arrigoni F, Pompili G, Barile A, Carrafiello G. Ultrasound imaging classifications of thyroid nodules for malignancy risk stratification and clinical management: state of the art. *Gland Surg* 2019;8(Suppl 3):S233-S244. doi: 10.21037/gs.2019.07.01