











Submitted:
30.11.2021
Accepted:
27.01.2022
Published:
13.04.2022

Summary of meta-analyses of studies involving TIRADS classifications (EU-TIRADS, ACR-TIRADS, and K-TIRADS) in evaluating the malignant potential of focal lesions of the thyroid gland

Katarzyna Dobruch-Sobczak*¹ , Zbigniew Adamczewski*² ,
Marek Dedejusz³ , Andrzej Lewiński^{4,5} , Bartosz Migda*⁶ ,
Marek Ruchała⁷ , Anna Skowrońska-Szcześniak*⁸ ,
Ewelina Szczepanek-Parulska*⁷ , Klaudia Zajkowska*³ , Agnieszka Żyłka*³ 

¹ Department of Radiology II, Maria Skłodowska-Curie National Research Institute of Oncology, Warsaw, Poland

² Department of Nuclear Medicine, Medical University of Lodz, Lodz, Poland

³ Department of Oncological Endocrinology and Nuclear Medicine, Maria Skłodowska-Curie National Research Institute of Oncology, Warsaw, Poland

⁴ Department of Endocrinology and Metabolic Diseases, Polish Mother's Memorial Hospital – Research Institute, Lodz, Poland

⁵ Department of Endocrinology and Metabolic Diseases, Medical University of Lodz, Lodz, Poland

⁶ Ultrasound Diagnostics Laboratory, Department of Paediatric Radiology, Faculty of Medicine, Medical University of Warsaw, Warsaw, Poland

⁷ Department of Endocrinology, Metabolism and Internal Diseases, Poznan University of Medical Sciences, Poznan, Poland

⁸ Department of Internal Medicine and Endocrinology, Medical University of Warsaw, Warsaw, Poland

* These authors contributed equally to this work

Correspondence: Katarzyna Dobruch-Sobczak; e-mail: katarzyna.dobruch-sobczak@pib-nio.pl

DOI: 10.15557/JoU.2022.0020

Keywords

thyroid;
ultrasound
classifications;
EU-TIRADS;
ACR-TIRADS;
K-TIRADS;
FNAB

Abstract

Numerous scientific societies around the world have published their TIRADS (Thyroid Imaging Reporting and Data System) classifications that evaluate the risk of malignancy of focal thyroid lesions, presenting different ultrasound features for each category and lesion size thresholds to determine eligibility for biopsy. The use of such risk estimation systems in focal thyroid lesions facilitates the reporting of thyroid ultrasound findings and improves the qualification of focal lesions for fine-needle aspiration biopsy (FNAB). In this publication, the three most popular TIRADS classifications, European – EU-TIRADS, Korean – K-TIRADS, and developed by the American Society of Radiology – ACR-TIRADS, are presented and discussed based on a literature review. The results of available head-to-head statistical analyses comparing the classifications are also presented. The advantage of the EU-TIRADS and K-TIRADS systems is that they include only the most important ultrasound features, so their application is not time-consuming, and the scores are easy to incorporate into clinical practice. ACR-TIRADS, unlike other scales, is based on a unique classification system and represents the most comprehensive classification. Each of the five categories of ultrasound features – morphology, echogenicity, shape, margins, microcalcifications – are evaluated and assigned a score from 0 to 3, with a higher score being associated with a higher risk of cancer. Based on the available data, the greatest benefit has been demonstrated for the ACR-TIRADS

classification, which also has implications for minimising the number of unnecessary FNABs. However, limitations related to the heterogeneity of the groups analysed in the study, including differences in the populations studied, inclusion criteria, proportions of patients of either sexes, and the number of malignant lesions analysed, should also be taken into account.

Introduction

Ultrasonography (USG) is a non-invasive and widely available diagnostic tool in the evaluation of focal thyroid lesions. The main limitation associated with the modality is the subjectivity of assessment of ultrasound features and the fact that it is impossible to visualise structures located retrosternally⁽¹⁾. It is characterised by relatively low diagnostic sensitivity for particular ultrasound features in estimating the malignant potential of focal lesions (27–63%), with high specificity reported for features including microcalcifications (87.8%), central vascularisation (78%), irregular margins (83.2%), shape higher than wider (96%), use of elastography (86.2%)⁽²⁾. In 2009, Horvath *et al.*, based on the Breast Imaging Reporting and Data System (BIRADS), was the first to propose a similar classification for thyroid imaging – Thyroid Imaging Reporting and Data System (TIRADS). The aim was to introduce a lexicon of sonographic features of focal thyroid lesions, to categorize findings, and to define further diagnostic and therapeutic management, including indications for fine-needle aspiration biopsy (FNAB)⁽³⁾. Targeted fine-needle aspiration biopsy is an important diagnostic tool for focal thyroid lesions, thus being a necessary complement to ultrasound diagnosis. Over the following years, numerous scientific societies worldwide have published classifications estimating the risk of focal thyroid lesions and presenting varied size thresholds and ultrasound features for each category, including those for biopsy.

The TIRADS classification is intended to have a high sensitivity and thus exclude low-risk lesions from biopsy and minimise the proportion of false-negative lesions.

However, it should be noted that in addition to sensitivity, it is important to obtain the highest possible level of specificity, in order to eliminate over-diagnosing and treating patients who do not actually need it. Unfortunately, there is no perfect tool with sensitivity and specificity levels equal to 100%, so it is essential to compare different classifications and adapt the best one to the diagnostic needs in a given population.

In this publication, the authors present a review of the literature based on meta-analyses on the most frequently cited TIRADS classifications.

EU-TIRADS classification – a literature review

The EU-TIRADS classification was introduced in 2017 and presented as a guideline by the European Thyroid Association (ETA)⁽⁴⁾. This scoring system is derived from the French classification, which was prospectively validated before its introduction, and its high diagnostic value was confirmed by Yoon *et al.* in a study of 4,696 thyroid nodules (TN), which showed a high level of sensitivity and negative predictive value (NPV)⁽⁵⁾. The lesions were divided into four categories – from 2 to 5 (EU-TIRADS 1 corresponds to the normal thyroid). A description of each category, tumour risk, and suggested indications for biopsy are shown in Tab. 1. A description of the lesions from each category is accompanied by sample ultrasound images (Fig. 1, Fig. 2, Fig. 3, Fig. 4, Fig. 5, Fig. 6).

The actual performance of the EU-TIRADS classification has also been tested on the Polish population. It was first

Tab. 1. EU-TIRADS classification according to the recommendations of the European Thyroid Association (2017)

| Category EU-TIRADS | Type of change | Risk of malignancy | Indications for biopsy |
|--------------------|--|--------------------|---|
| 1 | without focal changes | close to 0% | not recommended |
| 2 | anechogenic mixed solid-cystic with spongiform structure | close to 0% | not recommended (exception: therapeutic biopsy in symptomatic patients, e.g. cyst emptying) |
| 3 | oval shape regular margins isoechogenic or hyperechogenic without high-risk ultrasound features | 2–4% | >20 mm |
| 4 | oval shape regular margins slightly hypoechoic without high-risk ultrasound features | 6–17% | >15 mm |
| 5 | deeply hypoechoic shape other than oval irregular margins microcalcifications | 26–87% | >10 mm |

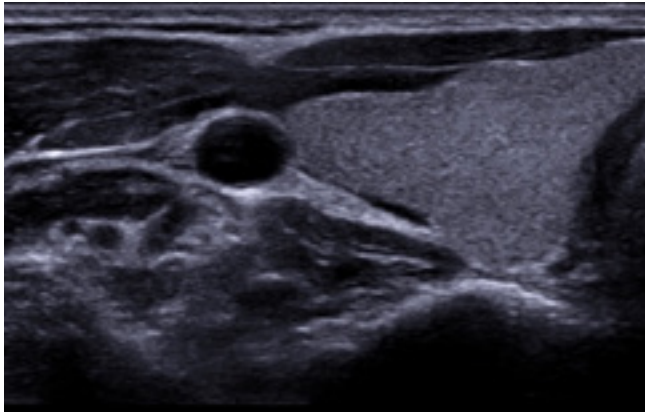


Fig. 1 EU-TIRADS 1, normal parenchyma (cross-section of the right lobe)

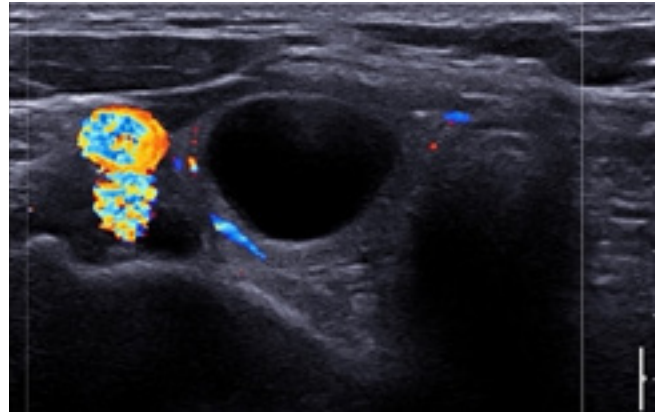


Fig. 2 EU-TIRADS 2, simple cyst

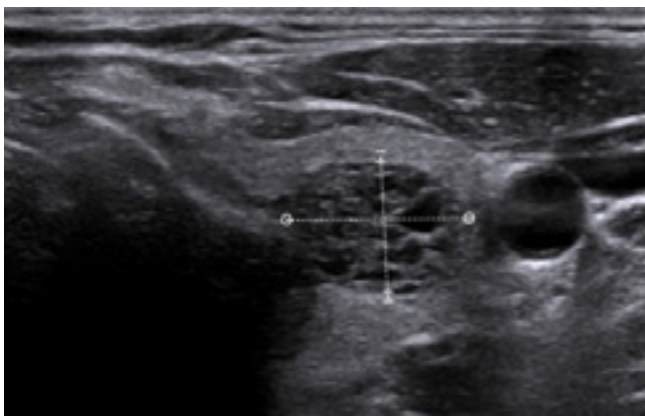


Fig. 3 EU-TIRADS 2, spongiform lesion

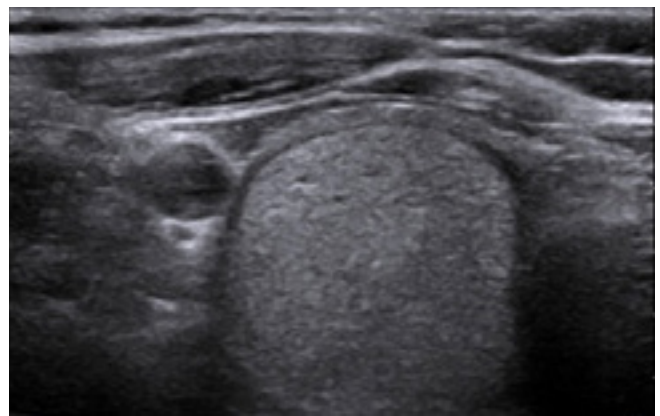


Fig. 4 EU-TIRADS 3, solid isoechogenic lesion with hypoechoogenic halo

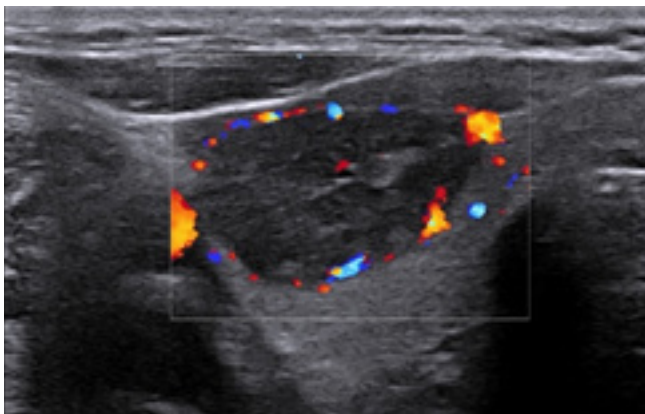


Fig. 5 EU-TIRADS 4, solid hypoechoogenic lesion. CD imaging shows marginal and central vascularisation

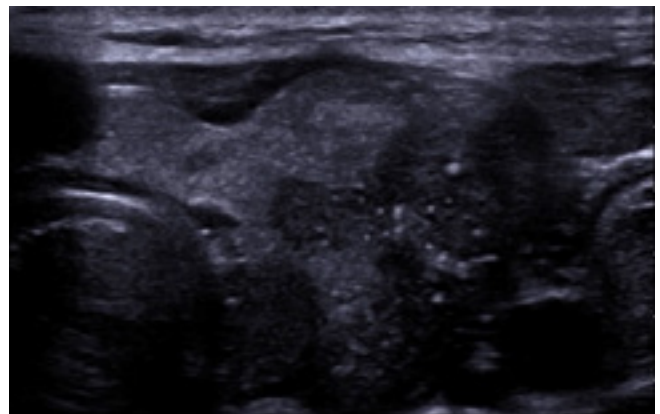


Fig. 6 EU-TIRADS 5, solid hypoechoogenic lesion with bright echoes (microcalcifications), uneven margins, irregular in shape

evaluated by Skowrońska *et al.* on a group of 52 patients with the presence of 40 TNs; the evaluation of the nature of the lesions was performed in relation to the histopathological examination⁽⁶⁾. The percentage of thyroid cancer was 0% in EU-TIRADS 2; 0% in EU-TIRADS 3; 5.9% in EU-TIRADS 4, and 75% in EU-TIRADS 5. When EU-TIRADS ≥ 4 was used as the cut-off value, the sensitivity, specificity, PPV (positive predictive value) and NPV were 75%, 94.1%, 75% and 94.1%, respectively.

Another study conducted to verify the diagnostic efficacy of the EU-TIRADS RSS (Risk Stratification Scale) in the Polish setting was a multi-centre study published by Dobruch-Sobczak *et al.*, in which a total of 842 TNs (613 benign, 229 malignant) were identified in 428 patients, and their nature was confirmed by histopathological verification⁽⁷⁾. The malignancy rate per EU-TIRADS category was as follows: 0% (EU-TIRADS 2), 3.2% (EU-TIRADS 3), 11.7% (EU-TIRADS 4), and 43.5% (EU-TIRADS 5).

The authors showed that if the cut-off points for FNAB proposed in the guidelines were used, 36% of malignant TNs would be missed. The cut-off value at EU-TIRADS 4 was characterised by high sensitivity (100%) and low specificity (25.1%); at the cut-off for EU-TIRADS 5 the values were 93.4% and 54.6%, respectively. Another study seeking to evaluate the benefits of the EU-TIRADS classification in the Polish population was the prospective study by Szczepanek-Parulska *et al.* in which a total of 133 TNs from 88 patients were analysed, bringing the following results: sensitivity 90.9%, specificity 61.2%, PPV 69.8%, NPV 87.2%, and accuracy 75.9%⁽⁸⁾. Another study conducted in Poland assessed the usefulness of the EU-TIRADS classification in the diagnosis of TNs from Hürthle cells (HC) with ambiguous cytological findings. The aim of the study was to compare the diagnostic efficacy of EU-TIRADS in two groups of TNs with equivocal cytology (categories III–V of the Bethesda system), with and without Hürthle cells. The study included a total of 162 HC and 378 non-HC TNs with an established histopathological diagnosis (17.9% and 15.6% of tumours). The authors concluded that the EU-TIRADS classification did not support clinical decision-making in patients with cytologically equivocal HC TNs, especially those classified as category IV of the Bethesda System for Reporting Thyroid Cytopathology (BSRTC)⁽⁹⁾. The EU-TIRADS score was also found to correlate with a positive BRAF mutation in patients with early-stage papillary thyroid carcinoma (PTC). However, since the conclusions were drawn based on the analysis of just 49 lesions, they need to be confirmed in large-scale studies⁽¹⁰⁾.

A recently published paper including 80 patients supported the idea of active surveillance of TN <1 cm classified as EU-TIRADS 5⁽¹¹⁾. Sixteen (20.0%) of these patients underwent surgery after a median follow-up of 57.2 months, which confirmed the diagnosis of PTC in 15 out of 16 cases, and all were in remission after a follow-up of 6–12 months. The results suggest good accuracy of the EU-TIRADS 5 category in the classification of malignant TNs even <1 cm in diameter, and the possibility of careful monitoring for TNs classified as EU-TIRADS 5⁽¹¹⁾.

The largest published multi-centre study undertaken to date with a view to evaluating the diagnostic performance of EU-TIRADS with respect to histopathology was that conducted by Trimboli *et al.*⁽¹²⁾. This retrospective study included a total of 1,058 TNs, of which 24.3% proved to be malignant. The incidence of malignancy was 1.4%, 3.5%, 17%, and 87.7% in grades 2–5. When EU-TIRADS categories 4 and 5 were combined, 93% sensitivity and 97% NPV were achieved. Using the recommended criteria for FNAB, a decrease in NPV to 90.9% was achieved. The largest meta-analysis on this topic was published by Castellana *et al.*⁽¹³⁾ Seven studies evaluating a total of 5,672 TNs with histopathological evaluation as the reference standard were included. The incidence of malignancy in each EU-TIRADS category was: 0.5%, 5.9%, 21.4%, and 76.1%. The sensitivity, specificity, PPV, NPV of EU-TIRADS 5 were 83.5%, 84.3%, 76.1%, and 85.4%, respectively. Further improvement in results was found

after excluding two studies with limited sample size and low incidence of malignancies in class 5⁽¹³⁾. The efficacy of the EU-TIRADS score was also tested in a paediatric population of 24 patients (31 TNs), 14 of whom underwent thyroidectomy; all malignant TNs were classified as EU-TIRADS grade 4 or 5. The sensitivity of the EU-TIRADS classification in detecting malignant TNs was 100%, specificity was 25%, PPV was 44%, and NPV was 100%⁽¹⁴⁾.

One recent study by Kovatcheva *et al.* aimed to evaluate the utility of the EU-TIRADS score in the assessment of TNs for FNAB and its ability to reduce unnecessary FNABs⁽¹⁵⁾. The study was a prospective single-centre analysis and included 741 patients with a total of 942 TNs. The malignancy rates in the categories 2 to 5 were 0, 0, 3.8, and 30.6%, respectively. The sensitivity, specificity, PPV, NPV and accuracy of EU-TIRADS with a cut-off point in category 5 were 91.3, 30.6, 98.6, and 76.4%, respectively. The diagnostic performance in aspects other than sensitivity and NPV was better in nodules ≥ 10 mm in size. The authors noted that the number of FNABs would decrease by 53.4% if the FNAB criteria were strictly followed. When FNAB indications were followed, the estimated sensitivity, specificity, PPV and NPV of EU-TIRADS were 69.9, 56.3, 16.4 and 93.8%, respectively⁽¹⁵⁾. The usefulness of EU-TIRADS was also demonstrated in a group of 75 TNs showing increased radioisotope uptake on 18F-FDG PET/CT, which was further evaluated on thyroid ultrasound. The cancer rate in this group was 0% in EU-TIRADS 2, 2.9% in EU-TIRADS 3, 4.2% in EU-TIRADS 4 and 78.6% in EU-TIRADS 5 ($p < 0.001$). The sensitivity, specificity, PPV, NPV and scale accuracy were found to be 5%, 95%, 79%, 97% and 93%, respectively. Based on these findings, it was concluded that EU-TIRADS might be helpful in predicting the nature of TNs detected by 18F-FDG PET/CT⁽¹⁶⁾.

In conclusion, the EU-TIRADS system is a simple and valuable RSS with a very good level of sensitivity, easy to incorporate into clinical practice. Its main strength lies in very high NPV, which contributes to reducing the number of FNABs.

ACR-TIRADS classification – a literature review

In 2017, the American College of Radiology (ACR), in collaboration with the American Association of Clinical Endocrinologists (AAEC) and the American Thyroid Association (ATA), developed a practical ultrasound system for assessing the malignancy risk of thyroid nodules, called the ACR Thyroid Imaging, Reporting and Data System (ACR-TIRADS)⁽¹⁷⁾. The concept is based on the commonly and widely used BIRADS system for the assessment of focal breast lesions⁽¹⁸⁾. The primary goal of the ACR-TIRADS classification, as well as other ultrasound RSSs, is to identify among TNs those which, based on the ultrasound findings, raise suspicion of neoplastic lesion and thus require deeper diagnostics and additional

Tab. 2. ACR TI-RADS classification – ultrasound features and their score values

| Ultrasound features | Ultrasound scoring characteristics <i>Details of US characteristic with point values</i> |
|----------------------|---|
| Morphology | cystic = 0 spongiform = 0 mixed cystic and solid = 1 solid = 2 indeterminate = 2 |
| Echogenicity | anechogenic = 0 isoechogenic = 1 hyperechogenic = 1 heterogeneous = 1 hypoechoic = 2 deeply hypoechoic = 3 |
| Shape | wider than tall = 0 taller and wide = 3 |
| Margins/borders | smooth = 0 ill-defined = 0 cannot be determined = 0 lobulated or irregular = 2 extra-thyroidal extension = 3 |
| Hyperechogenic areas | none = 0 comet tail artefacts = 0 macrocalcifications = 1 peripheral (rim) calcifications = 2 punctate echogenic foci (microcalcifications) = 3 |

procedures, most often using FNAB. Additionally, the classification helps to standardise the conclusions arising from the interpretation of ultrasound images, thus improving and streamlining communication between ultrasound technologists and clinicians, while contributing to a reduction in over-performed FNAB.

ACR-TIRADS, unlike other scales, is based on a unique scoring classification. Of the five categories of ultrasound features, individual elements – morphology, echogenicity, shape, margins, and microcalcifications – are evaluated and assigned a value of 0 to 3 points, with a higher score being associated with a higher risk of malignancy (Tab. 2). The total score determines the TIRADS malignancy risk (TR) divided into the following categories: TR1 – benign (0 point), TR2 – not suspicious (2 points), TR3 – slightly suspicious (3 points), TR4 – moderately suspicious (4–6 points) and TR5 – highly suspicious (≥ 7 points). The combined assessment of TR grade and maximum lesion diameter determines further management, which may be invasive (FNAB) or observation. The estimated risk of malignancy in each group is: $<2\%$ for TR1 and TR2, $<5\%$ for TR3, ranging from 5.1% to 20% for TR4 and $>20\%$ for TR5. In addition, ACR-TIRADS is part of a new interactive online algorithm, called TNAPP (Thyroid Nodule App), which was presented at the American Association for Clinical Endocrinology annual conference in 2021⁽¹⁹⁾. The innovation of TNAPP is based on combining clinical factors, ultrasound and FNAB features and provide clear suggestions for further management of TNs.

The diagnostic value of ACR-TIRADS has been confirmed in a number of recent studies. In a meta-analysis published by Li *et al.*, including 16 studies and evaluating 21,882 TNs, the authors showed that the meta-analytic overall sensitivity and specificity for risk stratification were 89% (95% CI, 81–93) and 70% (95% CI, 60–78)⁽²⁰⁾. The diagnostic odds ratio (DOR) was 18.46 (95% CI 9.77–34.88). Among these studies, 10/16 articles compared the direct diagnostic utility of ACR with ATA guidelines, while 6 studies compared ACR with Korean TIRADS, observing similar sensitivity (83% vs 87% and 85% vs 91%), but higher specificity of ACR compared with the other two (65% vs 50%, 57% vs 24%).

In conclusion, ACR-TIRADS is a well-established, valuable tool used in the assessment of TN malignancy risk that provides extensive and detailed lesion characterisation with clear management recommendations, thus avoiding unnecessary FNABs. The main disadvantage of this system when used in daily clinical practice is that it is quite time-consuming to use, particularly for novice clinicians.

K-TIRADS classification – a literature review

In 2016, the Korean Society of Thyroid Radiology (KSThR) presented the K-TIRADS classification and recommendations, describing the terminology and symptomatology of ultrasound imaging of focal thyroid lesions⁽²¹⁾. These ultrasound image features were used to develop the K-TIRADS classification and recommendations for each category. In addition to the B-mode image, the authors also detailed the vascularisation types of focal lesions and the potential role of sonoelastography in the differential diagnosis of TNs, but did not include them in the classification itself due to the divergent results presented in the literature. The classification itself is based on the ultrasonographic features of the B-mode image. The following were considered suspicious features: microcalcifications, vertical shape of the lesion, spiculated/microlobulated margins⁽²¹⁾. K-TIRADS 1 corresponds to healthy thyroid parenchyma without focal lesions. K-TIRADS 2 should be assigned to cysts, partially cystic lesions with comet tail artefacts present, and spongiform nodules. Partially cystic lesions that are iso-, hyperechogenic or of mixed echogenicity (iso-/hyper-) without any of the ultrasound suspicious features are assigned to category “3”, while lesions with any of the suspicious features present are assigned to category “4”. Similarly, category 4 is assigned to solid, hypoechoic lesions without suspicious features. K-TIRADS 5 refers to solid hypoechoic lesions with at least one suspicious feature present. Further management of focal lesions depends on their size (referring to the largest dimension)⁽²¹⁾:

- Category 2:
 - spongiform lesions, FNAB ≥ 2 cm (risk of malignancy, RA $<3\%$),
 - partially cystic lesions with comet tail artefacts present, FNAB not recommended (RA $<1\%$),

- Category 3: FNAB ≥ 1.5 cm (RA 3–15%),
- Category 4: FNAB ≥ 1 cm (RA 15–30%),
 - solid hypoechoic focal lesion without suspicious features,
 - partially cystic or iso-/hyperechoic lesion with any suspicious feature present,
- Category 5: FNAB ≥ 1 cm or selectively lesions > 5 mm (RA $> 60\%$).

Compared to ACR-TIRADS, the K-TIRADS classification seems more intuitive, however it should be remembered that ACR-TIRADS allows (and also requires) a more objective assessment (analysis) of each individual lesion (in contrast to K- and EU-TIRADS). Moreover, it should be noted that focal lesions with suspicious features, with mixed or solid structure and iso-/hyperechoic in the K-TIRADS classification are in category 4⁽²¹⁾ as in ACR-TIRADS⁽¹⁷⁾ classification, whereas in the EU-TIRADS classification they are assigned to 5⁽⁴⁾. K-TIRADS, like EU-TIRADS, do not consider macro-calcifications, annular calcifications and infiltration beyond the thyroid capsule as higher risk features in contrast to ACR-TIRADS, thus omitting quite important TN features considered suspicious in the literature. Another discrepancy among the classification is seen in the cut-off thresholds for the size of nodules referred for FNAB. The authors of K-TIRADS recommend cytological verification in categories 4 and 5 for lesions measuring ≥ 10 mm, noting the possibility of selective FNAB in category 5 for lesions > 5 mm in size. For EU-TIRADS and ACR-TIRADS category 4, FNAB is recommended if the lesion is > 15 mm in size and > 10 mm in category 5^(4,17,21). In the meta-analysis by Kim *et al.*, comparing the diagnostic ability of K, ACR- and EU-TIRADS classification against the cut-off points for category 5 (Tab. 3) and 4 or 5 (Tab. 4), no statistically significant differences were found. However, there was a trend towards higher sensitivity values for EU-TIRADS and higher specificity for K-TIRADS, and the sensitivity and specificity results are presented in Tab. 3 and Tab. 4⁽²²⁾. In a second meta-analysis comparing the diagnostic ability of these classifications, a similar conclusion was reached – the results for sensitivity and specificity for these classifications against category 5 (Tab. 5) and 4 or 5 (Tab. 6) are similar⁽²³⁾.

The discussion on the TIRADS classification should also include the problem of fine-needle biopsy and, more specifically, the percentage of unnecessary procedures generated by each classification. This aspect is further discussed in the following meta-analysis⁽²⁴⁾. In this study, the authors indicate that the K-TIRADS classification is characterised by a high percentage of unnecessary biopsies: 55%. The lowest percentage characterises the ACR-TIRADS classification (25%), significantly lower than for K-TIRADS ($p < 0.05$) and lower compared to EU-TIRADS (38%; $p = 0.087$). Clearly, this percentage is closely related to both the stratification of focal lesions (assignment to categories) and the nodule size cut-off point. Thus, borrowing the cut-off points from the ACR-TIRADS

Tab. 3. Comparison of the diagnostic parameters of the K-, ACR- and EU-TIRADS classifications against the cut-off point for category 5⁽²²⁾

| Category 5 | Sensitivity (%) | Specificity (%) |
|------------|-----------------|-----------------|
| K-TIRADS | 64 | 93 |
| ACR-TIRADS | 70 | 89 |
| EU-TIRADS | 78 | 89 |

Tab. 4. Comparison of diagnostic parameters of the K-, ACR- and EU-TIRADS classification in relation to the cut-off point for category 4 or 5⁽²²⁾

| Category 4 or 5 | Sensitivity (%) | Specificity (%) |
|-----------------|-----------------|-----------------|
| K-TIRADS | 92 | 61 |
| ACR-TIRADS | 95 | 49 |
| EU-TIRADS | 96 | 48 |

Tab. 5. Comparison of diagnostic parameters of K-, ACR- and EU-TIRADS classification in relation to the cut-off point for category 5⁽²³⁾

| Category 5 | Sensitivity (%) | Specificity (%) |
|------------|-----------------|-----------------|
| K-TIRADS | 55 | 95 |
| ACR-TIRADS | 66 | 91 |
| EU-TIRADS | 82 | 90 |

Tab. 6. Comparison of diagnostic parameters of the K-, ACR- and EU-TIRADS classification in relation to the cut-off point for category 4 or 5⁽²³⁾

| Category 4 or 5 | Sensitivity (%) | Specificity (%) |
|-----------------|-----------------|-----------------|
| K-TIRADS | 89 | 64 |
| ACR-TIRADS | 95 | 55 |
| EU-TIRADS | 96 | 52 |

classification and applying them to the K-TIRADS and EU-TIRADS classifications will reduce the proportion of unnecessary FNAB⁽²⁵⁾. However, it is important to note that a reduction in so-called unnecessary biopsies is not necessarily a desirable solution. It may entail an increase in the percentage of undiagnosed tumours, which may negatively affect the survival curve of patients with thyroid cancer. Based on meta-analyses alone, it is difficult to draw a definitive conclusion on the rate of unnecessary biopsies. The economic element, i.e. the percentage

of unnecessary biopsies and thus the possible rate of false positives with their consequences that researchers are willing to accept, may be a factor that facilitates the decision.

Summary of meta-analyses of studies including head-to-head comparative assessment of the use of TIRADS classification in estimating the malignancy potential of focal lesions of the thyroid gland: a review of the literature

The aim of this analysis is to evaluate the application of five major classifications: ACR-TIRADS (American College of Radiology guidelines), ATA (American Thyroid Association guidelines), Kwak-TIRADS, K-TIRADS (Korean Thyroid Association/Korean Society of Thyroid Radiology (KTA/KSThR) guidelines) (K-TIRADS is a development of the Kwak-TIRADS classification), and EU-TIRADS (European Thyroid Association (ETA) guidelines) in estimating the malignancy potential of focal thyroid lesions. Over the last few years, a number of studies have been published comparing the use of the TIRADS classification in the assessment of the malignancy risk of focal thyroid lesions, selected elements of which are presented in the following summary.

Yang *et al.* conducted a meta-analysis of 19 studies evaluating the clinical benefit of using the five classifications enumerated above. The meta-analysis included the results of research presented in articles published between 2015 and 2020. Between 100 and 4,696 focal lesions of the thyroid gland (24,325 lesions in total) diagnosed in between

92 and 4,585 patients were evaluated. All focal thyroid lesions underwent cyto- and/or histopathological evaluation. Twelve of the analysed studies included ACR-TIRADS assessment, 10 – included ATA classification, 6 included Kwak-TIRADS, 4 – EU-TIRADS, and 4 – K-TIRADS. Based on the statistical analysis performed, the high sensitivity of the individual classifications was obtained in the range of 0.84–0.96 (95% CI), while the specificity was presented sequentially for ACR-TIRADS, ATA, Kwak-TIRADS, KTA, and EU-TIRADS, respectively: 0.68, 0.44, 0.62, 0.47, and 0.61 (95% CI) – Tab. 7. An area under the ROC curve (AUC) value >0.8 was obtained for all the classifications, with a distribution of values for the above qualifications amounting to, respectively, 0.8553, 0.8976, 0.9101, 0.9022, and 0.8810, which supports the very good diagnostic accuracy of their use⁽²⁶⁾ (Tab. 7).

In contrast, the head-to-head comparison method yielded relative diagnostic odds ratio RDOR values of 1.57 (ACR vs ATA), 1.37 (ACR vs EU), 1.8 (ACR vs Kwak), 1.74 (ACR vs K), which distinguishes ACR-TIRADS in the analysis (Tab. 8). The authors of the study further highlighted the differences in the TIRADS classification system. ACR and Kwak-TIRADS are based on point classification, whereas the others categorise focal lesions based on sonographic patterns, which in clinical practice seems to be a more intuitive procedure, but associated with lower test accuracy. For example, the EU-TIRADS category 5 or 4 may correspond to ACR-TIRADS TR4/3 or K-TIRADS TR4/3, whereas focal lesions classified as K-TIRADS TR3 and EU-TIRADS category 3 (low risk of malignancy) may be categorised as ACR-TIRADS TR2, meaning no suspicion of malignancy. These differences in eligibility criteria,

Tab. 7. Sensitivity, specificity, LR(+), LR(-), DOR, AUC for individual TIRADS qualifications

| TIRADS classification | Number of tests | Sensitivity (95% CI) | Specificity (95% CI) | LR(+) (95% CI) | LR(-) (95% CI) | DOR (95% CI) | AUC |
|-----------------------|-----------------|----------------------|----------------------|------------------|------------------|--------------|--------|
| ACR-TIRADS | 13 | 0.85 (0.84–0.86) | 0.68 (0.6–0.69) | 2.98 (2.37–3.75) | 0.22 (0.16–0.29) | 15.23 | 0.8553 |
| EU-TIRADS | 4 | 0.85 (0.83–0.87) | 0.61 (0.59–0.62) | 2.84 (1.43–5.64) | 0.21 (0.13–0.34) | 13.18 | 0.8810 |
| K-TIRADS | 4 | 0.85 (0.83–0.86) | 0.47 (0.46–0.48) | 2.60 (1.2–5.57) | 0.18 (0.08–0.39) | 14.57 | 0.9022 |
| Kwak-TIRADS | 6 | 0.94 (0.94–0.95) | 0.62 (0.6–0.63) | 3.23(0.90–11.61) | 0.08 (0.04–0.16) | 43.15 | 0.9101 |

CI – confidence interval; LR(+) – positive likelihood ratio; LR(-) – negative likelihood ratio; DOR – diagnostic odds ratio; AUC – area under the ROC curve

Tab. 8. Head-to-head comparison of relative diagnostic odds ratio (RDOR) with CI 95%

| | ACR-TIRADS | EU-TIRADS | K-TIRADS | Kwak-TIRADS |
|-------------|------------------------|------------------------|------------------------|------------------------|
| ACR-TIRADS | - | 0.7308 (0.3000–1.7803) | 0.5734 (0.2759–1.1919) | 0.5564 (0.2552–1.2131) |
| EU-TIRADS | 1.3683 (0.5617–3.3332) | - | 0.7846 (0.3075–2.0020) | 0.7614 (0.2498–2.3208) |
| K-TIRADS | 1.7439 (0.8390–3.6247) | 1.2745 (0.4995–3.2518) | - | 0.9703 (0.3697–2.5466) |
| Kwak-TIRADS | 1.7972 (0.8243–3.9183) | 1.3138 (0.4309–4.0035) | 1.0306 (0.3927–2.7048) | - |

RDOR – relative diagnostic odds ratio; CI – confidence interval

which also take into account the size of the lesion, lead to differences in the specificity of the methods used. On the basis of the meta-analysis performed for ACR-TIRADS, the highest diagnostic accuracy in estimating the risk of malignancy and related limitation of indications to perform FNAB were noted⁽²⁶⁾. The authors point out certain limitations of the meta-analysis, including the lack of histopathological verification for each analysed focal lesion. Another limitation of the study was the numerical difference of lesions with malignant potential qualified for the analysis, moreover, there were no sufficient data related to the comparative analysis using K-TIRADS and EU-TIRADS. Also noteworthy is the limitation due to the incompatibility of ultrasound results obtained by different investigators and by the same investigator (inter-observer and intra-observer variability)⁽²⁷⁾. Conclusions from the presented meta-analysis emphasize the clinical usefulness and confirm the validity of recommending the use of TIRADS classification in estimating the malignancy potential of focal thyroid lesions, with an emphasis on the highest accuracy of the method for ACR-TIRADS.

Similar results on the benefits of using the ACR-TIRADS classification were obtained by Castellana *et al.* who conducted a meta-analysis of 12 studies (six of which were included in the previous meta-analysis) including five classifications (AACE/ACE/AME, ACR-TIRADS, ATA, EU-TIRADS, and K-TIRADS) to evaluate their use for TN typing for FNAB⁽²⁸⁾. A total of 18,750 TNs were analysed (including 4,378 histopathologically verified malignant lesions and 14,372 cyto- and/or histopathologically verified benign lesions). A DOR value ranging from 2.2 to 4.9 was obtained, with the highest RDOR value in the head-to-head comparative evaluation for ACR-TIRADS vs ATA ($p = 0.02$) and vs K-TIRADS ($p = 0.002$), which was due to the highest reliability quotient for a positive result (LR+ (95% CI)). The above results confirm the advantage of using for ACR-TIRADS in evaluating lesions for FNAB and thus reducing the number of unnecessary procedures.

In a subsequent study by Kim *et al.*, a total of 34 studies (analysing a total of 37,585 focal thyroid lesions) were analysed to assess the diagnostic utility of TIRADS in estimating the malignancy risk of focal thyroid lesions by classifying them into TIRADS categories 4 and 5 (TR4, TR5) with three classifiers: EU-TIRADS, ACR-TIRADS, and K-TIRADS⁽²²⁾. For TR5, the highest sensitivity was reported for EU-TIRADS: 78% (95% CI, 64–88%), followed by ACR-TIRADS: 70% (95% CI, 61–79%), and K-TIRADS: 64% (95% CI, 58–70%). The highest specificity for TR5 was observed for K-TIRADS: 93% (95% CI, 91–95%), ACR-TIRADS: 89% (95% CI, 85–92%), and EU-TIRADS: 89% (95% CI, 77–95%). Categorisation up to TR4/5 in all three classifications has a sensitivity of >90%, while the highest specificity was reported for K-TIRADS: 61% (95% CI, 50–72%), though no statistically significant difference was observed between the classifications. However, the results of the meta-analysis indicate a trend towards the highest sensitivity for the use of EU-TIRADS and the highest specificity for K-TIRADS as diagnostic tools in estimating the risk of malignancy associated with focal thyroid lesions.

In a subsequent meta-analysis of 29 papers evaluating a total of 33,748 focal thyroid lesions, Kim *et al.* reported comparable diagnostic accuracy of the four presented TIRADS classifications in estimating the risk of malignancy. The sensitivity and specificity of the method were considered in relation to ACR-TIRADS, being respectively: 66% and 91% for TR5, and 95% and 55% for TR4/5, K-TIRADS classification: 55% and 95% for TR5, and 89% and 64% for TR4/5, and EU-TIRADS: 82% and 90% for TR5, and 96% and 52% for TR4/5⁽²³⁾. There was no significant advantage for either classification as a prognostic factor for the risk of focal lesion malignancy. Attention was drawn to factors affecting the limitations of the meta-analysis performed; these include differences in the populations studied, inclusion criteria, proportions of patients of either sex, and numbers of malignant lesions analysed.

Kim *et al.* in their meta-analysis of 8 studies with an analysis of 13,092 focal thyroid lesions reported the lowest rate of unnecessary FNABs compared to ACR-TIRADS classification of 25% (95% CI, 22–29%), statistically significantly lower compared to ATA: 51% (95% CI, 44–58% ($p < 0.001$)) and K-TIRADS: 55% (95% CI, 42–67% ($p < 0.001$)). In contrast, there was no significant relationship compared with EU-TIRADS: 38% (95% CI, 16–66%, ($p = 0.087$))⁽²⁴⁾. The above analysis demonstrates the greatest benefit of using the ACR-TIRADS and EU-TIRADS classifications in typing lesions for FNAB, which translates into minimising the number of unnecessary invasive medical procedures.

Conclusions

The presented findings of the meta-analyses prove the diagnostic utility of the TIRADS classification in estimating the risk of malignancy of focal thyroid lesions and confirm the validity of recommending its applications in clinical practice. A head-to-head statistical analysis of the available data provides evidence for the superiority of the ACR-TIRADS classification, which also translates into minimising the number of unnecessary FNAB. However, in order to draw more far-reaching conclusions, it is necessary to conduct more studies using individual TIRADS classifications (including EU-TIRADS, K-TIRADS), taking into account the limitations presented above related to the heterogeneity of the groups subjected to the analysis. An advantage of the EU-TIRADS scale is that it includes only the most important ultrasound features, so its use is not time-consuming, and the scale is easy to apply in clinical practice. The application of the system would facilitate reporting of the thyroid ultrasound result and could improve TN qualification for FNAB.

Conflict of interest

The authors do not report any financial or personal connections with other persons or organizations which might negatively affect the contents of this publication and/or claim authorship rights to this publication.

Author contributions

Original concept of study: KD-S. *Writing of manuscript:* KD-S, AZ, BM, AS-S, ES-P, KZ, AŻ. *Analysis and interpretation of data:* KD-S, AZ, MD, BM, AS-S, ES-P, KZ, AŻ. *Final acceptance of manuscript:* KD-S, AZ, MD, AL, BM, MR, AS-S, ES-P, KZ, AŻ. *Collection, recording and/or compilation of data:* KD-S, AZ, BM, AS-S, ES-P, KZ, AŻ. *Critical review of manuscript:* MD, AL, MR.

References

1. Ginat DT, Butani D, Giampoli EJ, Patel N, Dogra V: Pearls and pitfalls of thyroid nodule sonography and fine-needle aspiration. *Ultrasound Q* 2010; 26: 171–178.
2. Remonti LR, Kramer CK, Leitão CB, Pinto LCF, Gross JL: Thyroid ultrasound features and risk of carcinoma: a systematic review and meta-analysis of observational studies. *Thyroid* 2015; 25: 538–550.
3. Horvath E, Majlis S, Rossi R, Franco C, Niedmann JP, Castro A *et al.*: An ultrasonogram reporting system for thyroid nodules stratifying cancer risk for clinical management. *J Clin Endocrinol Metab* 2009; 94: 1748–1751.
4. Russ G, Bonnema SJ, Erdogan MF, Durante C, Ngu R, Leenhardt L: European Thyroid Association guidelines for ultrasound malignancy risk stratification of thyroid nodules in adults: the EU-TIRADS. *Eur Thyroid J* 2017; 6: 225–237.
5. Yoon JH, Han K, Kim E-K, Moon HJ, Kwak JY: Diagnosis and management of small thyroid nodules: a comparative study with six guidelines for thyroid nodules. *Radiology* 2017; 283: 560–569.
6. Skowrońska A, Milczarek-Banach J, Wiechno W, Chudziński W, Żach M, Mazurkiewicz M *et al.*: Accuracy of the European Thyroid Imaging Reporting and Data System (EU-TIRADS) in the valuation of thyroid nodule malignancy in reference to the post-surgery histological results. *Pol J Radiol* 2018; 83: e579–586.
7. Dobruch-Sobczak K, Adamczewski Z, Szczepanek-Parulska E, Mígda B, Woliński K, Krauze A *et al.*: Histopathological verification of the diagnostic performance of the EU-TIRADS classification of thyroid nodules—results of a multicenter study performed in a previously iodine-deficient region. *J Clin Med* 2019; 8: E1781.
8. Szczepanek-Parulska E, Wolinski K, Dobruch-Sobczak K, Antosik P, Ostalowska A, Krauze A *et al.*: S-Detect Software vs. EU-TIRADS Classification: a dual-center validation of diagnostic performance in differentiation of thyroid nodules. *J Clin Med* 2020; 9: E2495.
9. Słowińska-Klencka D, Wysocka-Konieczna K, Klencki M, Popowicz B: Usability of EU-TIRADS in the diagnostics of Hürthle cell thyroid nodules with equivocal cytology. *J Clin Med* 2020; 9: E3410.
10. Skubisz K, Januszkiewicz-Caulier J, Cybula P, Bakula-Zalewska E, Goryca K, Paziewska A *et al.*: Higher EU-TIRADS-score correlated with BRAF V600E positivity in the early stage of papillary thyroid carcinoma. *J Clin Med* 2021; 10: 2304.
11. Rozenbaum A, Buffet C, Bigorgne C, Royer B, Rouxel A, Bienvenu M *et al.*: Outcomes of active surveillance of EU-TIRADS 5 thyroid nodules. *Eur J Endocrinol* 2021; 184: 677–686.
12. Trimboli P, Ngu R, Royer B, Giovanella L, Bigorgne C, Simo R *et al.*: A multicentre validation study for the EU-TIRADS using histological diagnosis as a gold standard. *Clin Endocrinol* 2019; 91: 340–347.
13. Castellana M, Grani G, Radzina M, Guerra V, Giovanella L, Deandrea M *et al.*: Performance of EU-TIRADS in malignancy risk stratification of thyroid nodules: a meta-analysis. *Eur J Endocrinol* 2020; 183: 255–264.
14. Yeste Fernández D, Vega Amenabar E, Coma Muñoz A, Arciniegas Vallejo L, Clemente León M, Planes-Conangla M *et al.*: Ultrasound criteria (EU-TIRADS) to identify thyroid nodule malignancy risk in adolescents. Correlation with cyto-histological findings. *Endocrinol Diabetes Nutr (Engl Ed)* 2021; S2530-0164(21)00078-1.
15. Kovatcheva RD, Shinkov AD, Dimitrova ID, Ivanova RB, Vidinov KN, Ivanova RS: Evaluation of the diagnostic performance of EU-TIRADS in discriminating benign from malignant thyroid nodules: a prospective study in one referral center. *Eur Thyroid J* 2021; 9: 304–312.
16. Trimboli P, Paone G, Treglia G, Virili C, Ruberto T, Ceriani L *et al.*: Fine-needle aspiration in all thyroid incidentalomas at 18 F-FDG PET/CT: Can EU-TIRADS revise the dogma? *Clin Endocrinol* 2018; 89: 642–648.
17. Tessler FN, Middleton WD, Grant EG, Hoang JK, Berland LL, Teefey SA *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: 587–595.
18. Vanel D: The American College of Radiology (ACR) Breast Imaging and Reporting Data System (BI-RADS): a step towards a universal radiological language? *Eur J Radiol* 2007; 61: 183.
19. Garber JR, Papini E, Frasoldati A, Lupo MA, Harrell RM, Parangi S *et al.*: American Association of Clinical Endocrinology and Associazione Medici Endocrinologi Thyroid Nodule Algorithmic Tool. *Endocr Pract* 2021; 27: 649–660.
20. Li W, Wang Y, Wen J, Zhang L, Sun Y: Diagnostic performance of American College of Radiology TI-RADS: a systematic review and meta-analysis. *AJR Am J Roentgenol* 2021; 216: 38–47.
21. Shin JH, Baek JH, Chung J, Ha EJ, Kim J-H, Lee YH *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: 370–395.
22. Kim PH, Chung SR, Choi SH, Kim KW: Accuracy of thyroid imaging reporting and data system category 4 or 5 for diagnosing malignancy: a systematic review and meta-analysis. *Eur Radiol* 2020; 30: 5611–5624.
23. Kim PH, Suh CH, Baek JH, Chung SR, Choi YJ, Lee JH: Diagnostic performance of four ultrasound risk stratification systems: a systematic review and meta-analysis. *Thyroid* 2020; 30: 1159–1168.
24. Kim PH, Suh CH, Baek JH, Chung SR, Choi YJ, Lee JH: Unnecessary thyroid nodule biopsy rates under four ultrasound risk stratification systems: a systematic review and meta-analysis. *Eur Radiol* 2021; 31: 2877–2885.
25. Ha SM, Baek JH, Na DG, Suh CH, Chung SR, Choi YJ *et al.*: Diagnostic performance of practice guidelines for thyroid nodules: thyroid nodule size versus biopsy rates. *Radiology* 2019; 291: 92–99.
26. Yang R, Zou X, Zeng H, Zhao Y, Ma X: Comparison of diagnostic performance of five different ultrasound TI-RADS classification guidelines for thyroid nodules. *Front Oncol* 2020; 10: 598225.
27. Liu H, Ma A-L, Zhou Y-S, Yang D-H, Ruan J-L, Liu X-D *et al.*: Variability in the interpretation of grey-scale ultrasound features in assessing thyroid nodules: a systematic review and meta-analysis. *Eur J Radiol* 2020; 129: 109050.
28. Castellana M, Castellana C, Treglia G, Giorgino F, Giovanella L, Russ G *et al.*: Performance of five ultrasound risk stratification systems in selecting thyroid nodules for FNA. *J Clin Endocrinol Metab* 2020; 105: dgz170.