REVIEW ARTICLE

Diagnostic performance of ultrasound elastography combined with ACR TI-RADS for thyroid nodule assessment: a systematic review and meta-analysis

Joud Alghamdi[1]*, Thana Algarni[1], Rama Alawi[1], Malath Alkathiri[1], Arwa Alariany[1], Basmah Sherbini[1], Hams Aljadani[1], Taif Alenezi[1], Dareen Bajamaan[1], Raghad Filimban[1], Sara Aljahdali[1], Amr M. Daous[2]

ABSTRACT

Introduction: Thyroid nodules are common, and most are benign. Risk stratification is important for detection of malignancy and avoidance of unnecessary fine-needle aspiration (FNA). The American College of Radiology Thyroid Imaging Reporting and Data System (ACR TI-RADS) provides standardized ultrasound-based assessment, and elastography has been suggested as an adjunct. This study aimed to evaluate the diagnostic performance of elastography combined with ACR TI-RADS and its impact on biopsy recommendation.

Methods: This systematic review and meta-analysis followed Preferred Reporting Items for Systematic Reviews and Meta-Analyses guidelines and was registered a priori in the International Prospective Register of Systematic Reviews (PROSPERO). Studies assessing the combined method of elastography plus ACR TI-RADS in correlation with histopathology and/or FNA cytology were included. Pooled sensitivity and specificity were calculated through a bivariate random-effect model, and diagnostic performance was assessed by hierarchical summary receiver operating characteristic analysis.

Results: Fifteen studies were selected. The pooled sensitivity, specificity, and area under the curve (AUC) of ACR TI-RADS alone were 84.1%, 76.0%, and 0.874, respectively. Adding elastography raised sensitivity to 90.5% but reduced specificity to 18.8%, with a lower AUC of 0.711. The pooled estimate indicated 41.1% of biopsies could be potentially avoided, but the estimate was compromised by substantial heterogeneity (I^2 = 98.4%) and limited generalizability. The combined false-negative rate was 3.0% with significant heterogeneity.

Conclusion: In our study, ACR TI-RADS alone showed more balanced performance than the combined systems. Adding elastography increased sensitivity but decreased specificity and overall diagnostic accuracy. Evidence on biopsy reduction is uncertain due to high heterogeneity. These results do not support routine elastography for reducing FNA, but selective adjunctive use warrants further prospective studies.

Keywords: Thyroid nodules, ultrasound elastography, ACR TI-RADS, fine-needle aspiration, diagnostic accuracy.

Introduction

Thyroid nodules are among the most frequently encountered abnormalities in clinical practice, particularly with the increasing use of high-resolution imaging [1]. Their reported prevalence varies according to the detection method and may reach approximately 68% when ultrasonography is used. Despite their high prevalence, most thyroid nodules are benign, making the accurate identification of the relatively small proportion of malignant lesions while avoiding unnecessary invasive procedures an important clinical challenge [1,2].

Correspondence to: Joud Alghamdi
*College of Medicine, University of Jeddah, Jeddah, Saudi Arabia.
Email: Joudmh0@gmail.com
Full list of author information is available at the end of the article.
Received: 25 August 2026 | **Revised (1):** 03 September 2026| **Accepted:** 03 October 2026

OPEN ACCESS
This is an open access article distributed in accordance with the Creative Commons Attribution (CC BY 4.0) license: https://creativecommons.org/licenses/by/4.0/) which permits any use, Share — copy and redistribute the material in any medium or format, Adapt — remix, transform, and build upon the material for any purpose, as long as the authors and the original source are properly cited. © The Author(s) 2026.
AMEM

Ultrasonography is the primary imaging modality for the evaluation of thyroid nodules and plays a major role in the decision-making process regarding the need for further assessment of lesions. To standardize the interpretation and management of nodules based on sonographic features and guidelines for follow-up or fine-needle aspiration (FNA) by size, the American College of Radiology developed the Thyroid Imaging Reporting and Data System (ACR TI-RADS) [2]. ACR TI-RADS has satisfactory diagnostic performance with a pooled sensitivity of 89% and a specificity of 70% [3]. However, the high prevalence of unnecessary biopsies is still a major concern. Meta-analyses comparing ultrasound risk-stratification systems have shown that ACR TI-RADS has an unnecessary biopsy rate of about 25%, which is lower than that of several commonly used systems [4].

Ultrasound elastography has therefore emerged as a promising complementary technique for better characterizing thyroid nodules. Elastography offers information regarding tissue stiffness, taking advantage of the fact that malignant nodules tend to demonstrate higher stiffness than benign lesions. Strain elastography (SE) and shear-wave elastography (SWE) have been investigated as an adjunct to conventional ultrasonography and TI-RADS-based assessment [5,6]. Previous studies have suggested that elastography parameters may improve discrimination of benign and malignant nodules when added to ultrasound risk stratification and potentially change biopsy recommendations. For example, shear-wave elastography, which is a promising tool in the reclassification of some ACR TI-RADS 4 nodules, and prospective data showed that the use of TI-RADS combined with elastography could reduce the number of FNA procedures [5–7].

However, the incremental diagnostic benefit of elastography remains uncertain. Studies vary considerably in elastography modality, stiffness thresholds, patient populations, reference standards, and the strategies used to combine elastography with TI-RADS [5–7]. Consequently, whether adding elastography provides sufficient diagnostic benefit to improve risk stratification and safely reduce unnecessary FNA remains unclear. Therefore, this systematic review and meta-analysis aimed to evaluate the diagnostic performance of the combined method of elastography plus ACR TI-RADS (COMB) for thyroid nodules and to assess its potential impact on unnecessary FNA while maintaining adequate detection of malignancy.

Methods

Review of literature

This systematic review was conducted and reported in accordance with the Preferred Reporting Items for Systematic Reviews and Meta-Analyses guidelines to minimize the risk of bias [8]. The review protocol was prospectively registered in PROSPERO under registration ID: CRD420261322223 [9]. Given the nature of this study, which involves analysis of previously published data, ethical approval was not required.

Search strategy

A comprehensive systematic search was conducted in February 2026 in four electronic databases: PubMed, Web of Science, Cochrane Library, and Google Scholar. The following search terms were applied: (Thyroid Elastography OR Shear Wave Elastography OR Strain Elastography) AND (TI-RADS OR ACR TI-RADS OR Thyroid Imaging Reporting and Data System) AND (Thyroid nodule OR Thyroid neoplasm OR Thyroid lesion) AND (Fine Needle Aspiration Biopsy OR FNA OR Needle Biopsy) AND (Diagnostic Accuracy OR Diagnostic Performance).

Eligibility criteria

Study eligibility was determined using the PICO framework (Population, Intervention, Comparison, Outcome) [10]. Inclusion criteria: Studies were eligible if they enrolled adult patients (≥18 years of age), used a prospective or retrospective cohort or diagnostic accuracy study design, assessed COMB, used histopathology and/or fine-needle aspiration cytology (FNAC) as the reference standard, reported diagnostic accuracy outcomes, and were published in English. Exclusion criteria: Studies were excluded if they included patients younger than 18 years, were case reports, narrative reviews, editorials, letters, or expert opinions, were conference abstracts lacking complete data, were animal or laboratory-based studies, lacked extractable outcome data, or evaluated elastography without comparison to ACR TI-RADS.

Study selection and data extraction

Three independent reviewers screened titles and abstracts using the Rayyan website and mobile app for systematic reviews [11], which was also used to identify and remove duplicate records. Full-text review of potentially eligible articles was then performed by two reviewers, with discrepancies resolved by a third reviewer.

Data were extracted by three reviewers into a standardized Microsoft Excel form capturing the following variables: total number of patients and nodules, study design, country, population type, mean patient age and range, number of female patients, number of malignant and benign nodules, inclusion and exclusion criteria, TI-RADS system used, definition of a positive TI-RADS result, TI-RADS biopsy threshold, elastography modality, elasticity cutoff values, definition of a positive elastography result, combination strategy, definition of a positive combined test, and reference standard used.

Quality assessment

The methodological quality and risk of bias of included studies were assessed using the Quality Assessment of Diagnostic Accuracy Studies 2 (QUADAS-2) tool [12]. Two reviewers independently evaluated each study. Disagreements were resolved through discussion; where consensus could not be reached, a third reviewer provided the final judgment.

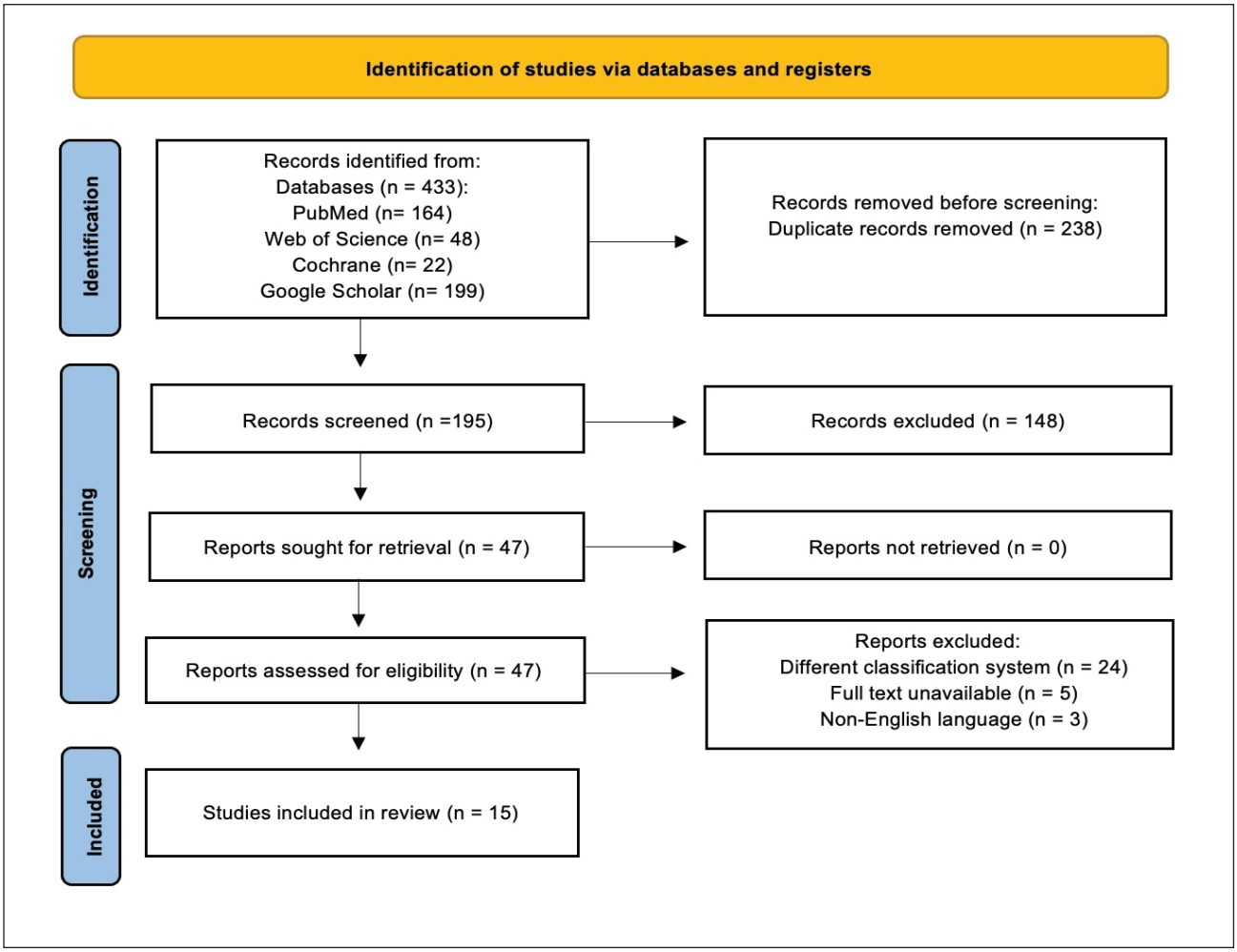


Figure 1. *PRISMA flow diagram for study selection.*

Statistical analysis

The data were extracted and subsequently converted into the four elements of a 2 × 2 contingency table (true positive (TP), false positive (FP), true negative (TN), false negative (FN)) based on the principles of Monaghan et al. 2021 [13] and a preconstructed calculation sheet. These were then used to generate confusion matrices for each study in the review. To provide pooled estimates of sensitivity and specificity and their 95% confidence intervals (CIs) whilst taking account of between-study heterogeneity and the inherent correlation between the two measures [14], a bivariate random-effects model was used. Forest plots were created with point estimates from individual studies, pooled estimates and 95% CIs for sensitivity and specificity that can be intuitively compared across the study samples. A Hierarchical Summary Receiver Operating Characteristic (HSROC) curve was used to plot the overall diagnostic performance and to illustrate the trade-off between sensitivity and specificity for the parameters generated by the bivariate model [15]. In the same model, the area under the HSROC curve (AUC) was computed as a summary measure of test performance. To assess heterogeneity between studies, the bivariate I^2 statistic developed by Zhou and Dendukuri was applied [16] along with the sample size-unadjusted and adjusted I^2 estimates developed by Holling et al. [17]. The R software program (version 4.4.2, 2024) was used to conduct all statistical analyses using the "mada" and "meta" packages [18].

Results

Search results

The search results across databases yielded a total of 433 potentially relevant records. After removal of 238 duplicated records, a total of 195 records underwent title and abstract screening. The initial screening resulted in the exclusion of 148 irrelevant records, with subsequent meticulous full-text assessment for the remaining 47 records based on eligibility criteria. Finally, 15 studies [6,7,19–31] were eligible for this systematic review and meta-analysis after the exclusion of 32 irrelevant articles (Figure 1).

Study characteristics

This systematic review and meta-analysis included prospective and retrospective observational studies conducted mainly in China, Turkey, Greece, India, and Mexico, evaluating thyroid nodules with suspected malignancy or referred for ultrasound-based diagnostic evaluation. Across the included studies, the sample sizes

Table 1. *Summary of included studies and patient characteristics.*

Study ID	Study design	Location	Patients, n	Nodules, n	Mean age (SD), range, years	Malignant nodules, *n* (%)
Angelopoulos_2025	Prospective, single center	Greece	556	556	58 (12), (18-90)	55 (9.9)
SahinEdiz_2023	Prospective, single center	Turkey	91	91	49.8 (11.42)	12 (13.19)
Sekar_2025	Cross-sectional, single center	India	255	255	42.77 (13.7), (18-84)	51 (20)
Zhang_2021	Retrospective, multicenter	China	566	566	47.4 (13.4), (19-75)	341 (60.25)
Polat_2024	Prospective, single center	Turkey	308	370	50.6 (13.37), (18-70)	135 (36.49)
Yi_2024	Prospective, single center	China	237	270	NR, (>18)	198 (73.33)
Hao_2023	Retrospective, single center	China	109	112	malignant 39.9 (12.8), benign 46.8 (13.3), (22-77)	62 (55.36)
Qi_2023	Prospective, multicenter	China	543	593	44.96 (12.08)	385 (64.92)
Luo_2020	Retrospective, single center	China	152	152	45.2 (12.4), (19-74)	22 (14.47)
Paredes-Manjarrez_2021	Cross-sectional	Mexico	756	756	NR	NR
Zhang_2023	Retrospective, multicenter	China	548	548	49.39 (12.35), (22-80)	242 (44.16)
Hang_2018	Retrospective, multicenter	China	262	298	45.57 (12.11), (21-76)	177 (59.4)
Xu_2020	Retrospective, single center	China	864	922	45.81 (11.28), (28-62)	517 (56.07)
Pei_2020	Retrospective, multicenter	China	1498	1525	43.5 (12.4), (14-85)	799 (52.4)
Li_2019	Retrospective, single center	China	128	130	47.8 (10.5), (17-68)	73 (56.15)

NR = not reported, SD = standard deviation.

were variable between 91 and 1498 patients. Similarly, the mean age of participants was predominantly in the fifth decade of life, and sex proportions showed a strong female predominance. The proportions of malignant nodules varied widely across studies, from 9.9% to 73.33% (Table 1). Detailed population characteristics and eligibility criteria are provided in Supplementary Tables (S1–S2). For TI-RADS systems and elastography methods, the studies used different ultrasound-based systems, elastography techniques, cut-off values, and reference standards such as FNA, surgery, or pathology (Table 2). Detailed thresholds, positivity definitions, and combination rules are provided in Supplementary Tables (S3–S4).

Risk of bias assessment

The quality assessment using the QUADAS-2 tool showed that the included studies had overall acceptable methodological quality, as most studies were judged as low risk of bias in patient selection, index test, and reference standard domains. However, the flow and timing domain was rated as high risk across six studies [6,19,21,23,27,28] and unclear risk across three other studies [20,22,29]. Most studies had low applicability concerns in the patient selection and index test domains. However, the reference standard domain showed high variability, with six studies rated as unclear or high concern because of the diagnostic confirmation methods or incomplete pathological verification (Figure 2) [6,19,21,27,28,30].

Diagnostic accuracy of ACR TI-RADS

In bivariate analysis, ACR TI-RADS showed good diagnostic performance with a pooled sensitivity of 84.1% (95% CI [77.5%-89.1%]), specificity of 76.0% (95% CI [64.5%-84.7%]) (Figure 3). The false positive rate (FPR) was 24.0% (95% CI [15.3%-35.5%]), AUC and partial AUC were 0.874 and 0.880, respectively. Heterogeneity was moderate by Zhou and Dendukuri [16] I^2 (50.3%) and low or negligible with adjusted Holling I^2 (3.2%-5.7%) (Table S5). The univariate pooled estimates for specificity and sensitivity of ACR TI-RADS were 0.76 (95% CI [0.65-0.85]) and 0.85 (95% CI [0.78-0.89]), respectively (Figures S1 and 2).

Additionally, univariable meta-regression analysis demonstrated no significant association between elastography modality (SWE *vs.* SE) and pooled diagnostic performance. Compared with SWE, SE was associated with a non-significant reduction in sensitivity ($\beta = -0.41$, 95% CI: [−1.51 to 0.70], $p = 0.472$) and a non-significant reduction in FPR ($\beta = -0.59$, 95% CI: [−2.01 to 0.83], $p = 0.415$), indicating no significant difference in specificity between modalities.

Table 2. *Summary of TI-RADS, elastography, and combination methods.*

Study ID	TI-RADS system	Elastography modality	Main cutoff(s)	Combination strategy	Reference standard
Angelopoulos_2025	ACR TI-RADS	Strain elastography (SE / RTE)	ER: TR3 >1.60; TR4 >1.37; TR5 >0.95	Sequential; TI-RADS-guided FNA refined by ER	Cytology and histopathology
SahinEdiz_2023	ACR TI-RADS 2017	SWE	NR	NR	FNAC and/or histopathology
Sekar_2025	ACR TI-RADS + Kwak TI-RADS	ARFI elastography VTi –qualitative (VTq – quantitative, SWV)	SWV >3.4 m/s; VTi grade ≥3	No formal combined model	Histopathology
Zhang_2021	ACR TI-RADS + Kwak TI-RADS + ATA + KTA/KSThR	SWE, Emax in kPa	Emax >41.7 kPa	Sequential category upgrade by SWE	FNAC and/or histopathology
Polat_2024	ACR TI-RADS 2017	(SWE- VTIQ)	VTIQ ≥2.94 m/s	SWE-based reclassification	FNAC and histopathology
Yi_2024	ACR TI-RADS	SWE	E-min 18.25; E-mean 28.15; E-max 40.40; E-ratio 1.75; SD 4.25 kPa	Additive risk score (positive ≥8)	Histopathology
Hao_2023	ACR TI-RADS	3D-SWE	3D-C-Emax ≥51.5 kPa	SWE-guided upgrade of TR4; TR5 unchanged	Histopathology
Qi_2023	ACR TI-RADS	2D-SWE	E-whole-mean >11.35 kPa	Logistic regression model	FNAC and histopathology
Luo_2020	ACR TI-RADS	SWE	≤23.30 kPa downgrade; ≥52.14 kPa upgrade	SWE-guided category upgrade/downgrade	FNAC, histopathology, and follow-up
Paredes-Manjarrez_2021	ACR TI-RADS	SWE	TR5 >32.5 kPa; TR1-4 >21.5 kPa	NR	FNAC
Zhang_2023	ACR TI-RADS	Strain elastography and 2D SWE	VTIQ max 3.215-3.54 m/s; SE score ≥4	Predictive risk score	FNAC and histopathology
Hang_2018	ACR TI-RADS	SWE	Emax >52.7 kPa	Combined risk score (>8)	Histopathology
Xu_2020	ACR TI-RADS	Strain elastography + shear wave (VTIQ)	VTIQ 3.552 m/s; SR 3.371	Parallel combination (both tests positive)	FNAC and histopathology
Pei_2020	ACR TI-RADS	Strain elastography (RTE)	NR	Logistic model	FNAC and histopathology
Li_2019	ACR TI-RADS	Shear wave elastography (VTIQ)	SWV min 2.46 m/s	Sequential TR3/TR4 reclassification	FNAC and histopathology

ACR TI-RADS = American College of Radiology Thyroid Imaging Reporting and Data System, ARFI = Acoustic Radiation Force Impulse, ATA = American Thyroid Association ultrasound risk classification, ER = Elasticity Ratio, FNAC = fine needle aspiration cytology, KSThR = Korean Society of Thyroid Radiology, KTA = Korean thyroid ultrasound guidelines/classification system, NR = not reported, RTE = Real-Time Elastography, SE = Strain elastography, SR = Strain ratio, SWE = Shear Wave Elastography, SWV = Shear Wave Velocity, TI-RADS Risk Category, VTi = Virtual Touch Imaging, VTIQ = Virtual Touch Tissue Imaging and Quantification, VTq = Virtual Touch Quantification.

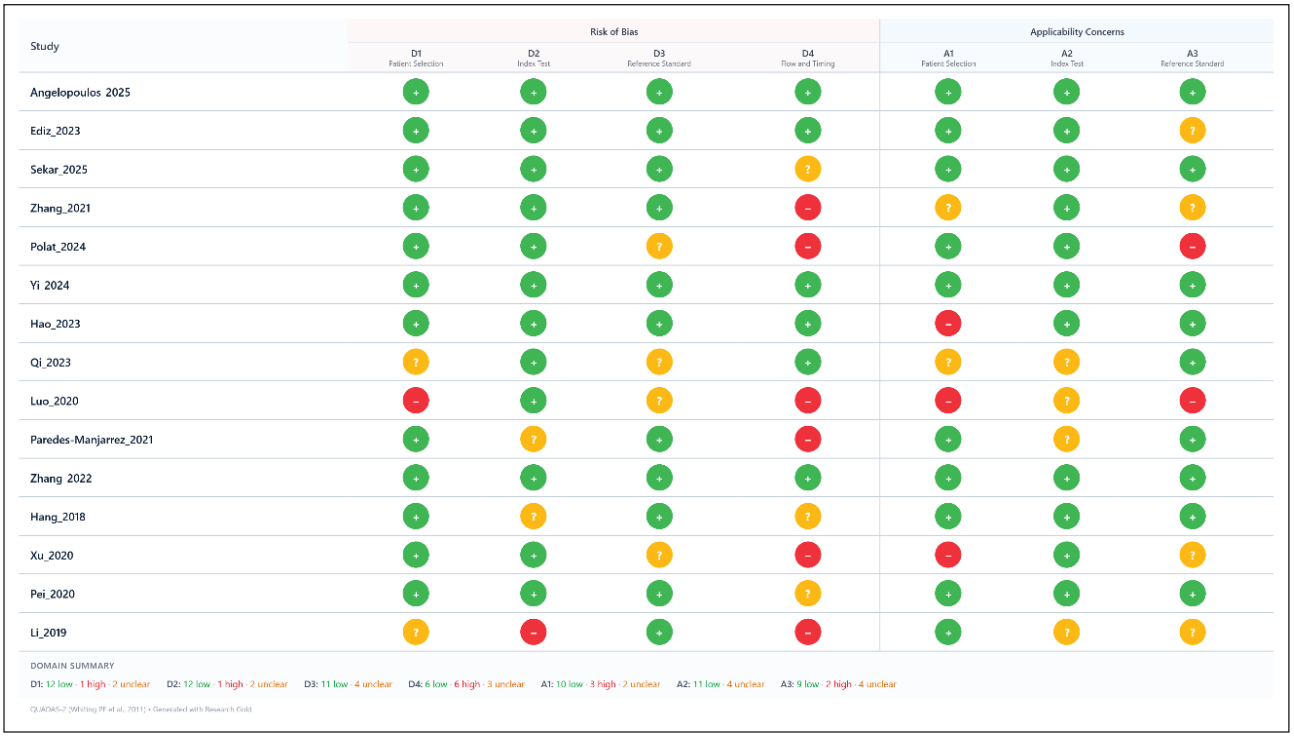


Figure 2. *Summary of risk of bias and applicability concerns assessed using QUADAS-2 for the included studies.*

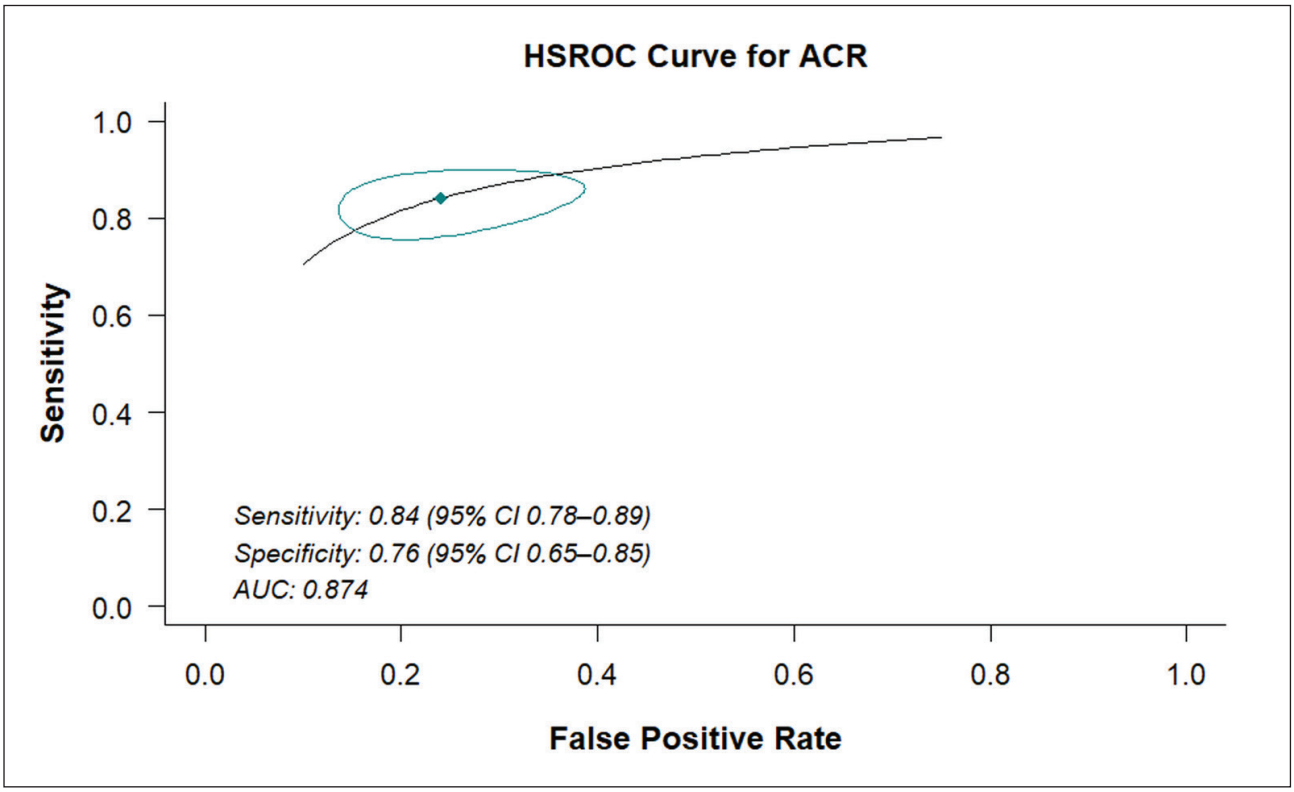


Figure 3. *Shows Hierarchical Summary Receiver Operating Characteristic (HSROC) curve for ACR TI-RADS.*

Diagnostic accuracy of combined method

In bivariate analysis, COMB demonstrated high sensitivity but very low specificity, with a pooled sensitivity of 90.5% (95% CI [85.7%-93.8%]) and specificity of 18.8% (95% CI [11.3%-29.6%]), yielding an AUC of 0.711 and a partial AUC of 0.855 (Figure 4).

The FPR was 81.2% (95% CI [70.4%-88.7%]). The heterogeneity test showed low to moderate inconsistency, with I^2 values of 45.6% Zhou and Dendukuri [16], and

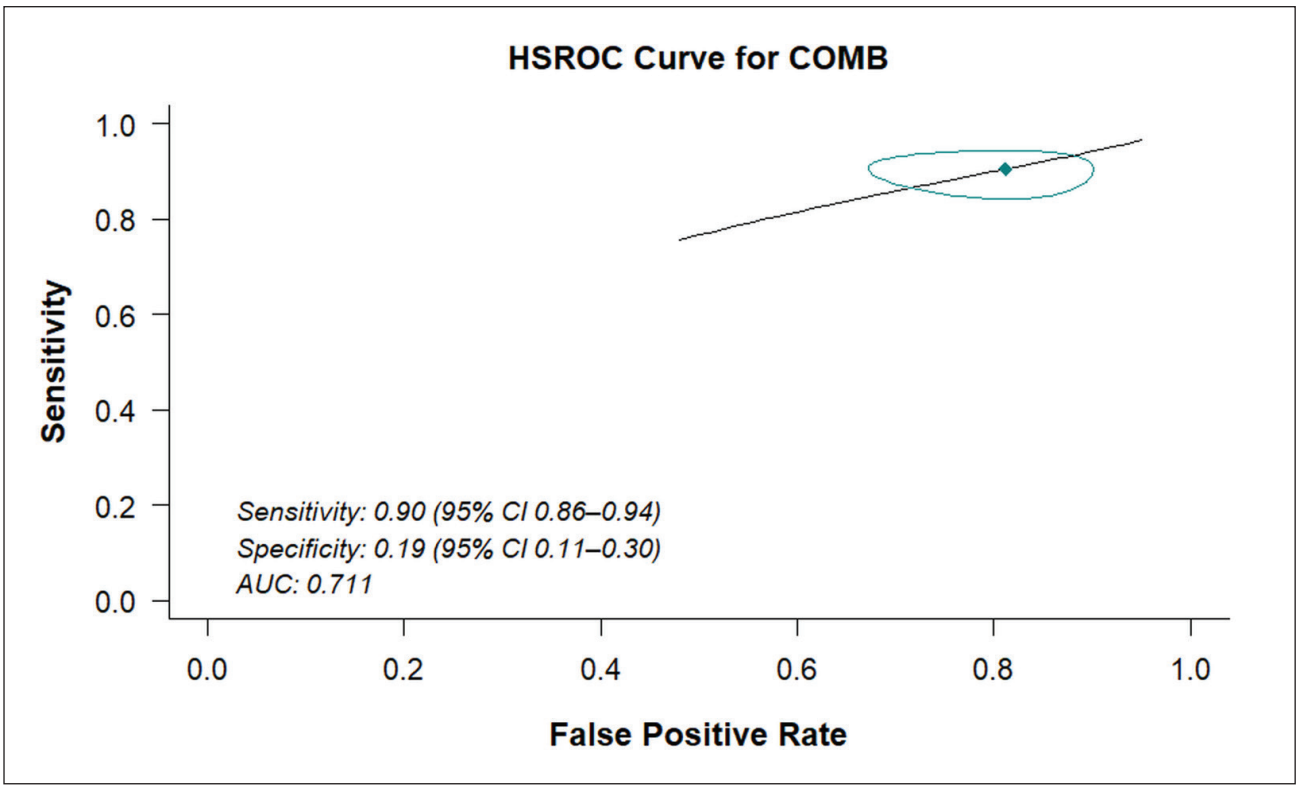


Figure 4. *Shows the Hierarchical Summary Receiver Operating Characteristic (HSROC) curve for the combined method of elastography plus ACR TI-RADS (COMB).*

1.9%-2.8% (adjusted Holling) (Table S6). The univariate analysis showed that COMB had high pooled sensitivity of 0.91 (95% CI [0.86-0.94]) and low pooled specificity of 0.19 (95% CI [0.11-0.29]), with substantial heterogeneity across studies (Figures S3 and 4).

Meta-regression showed no statistically significant difference in sensitivity between SWE and SE (β = −0.135, 95% CI: [−1.291 to 1.021]; p = 0.819). However, SE showed a non-significant trend toward a higher FPR compared with SWE (β = 1.137, 95% CI: [−0.122 to 2.397]; p = 0.077), suggesting a possible reduction in specificity, although this finding did not reach statistical significance.

Avoided biopsies and missed cases

The pooled proportion of false-negative cases after biopsy was 3% (25/769), 95% CI: [0%–13.2%], indicating a low FN rate across the included studies, but with significant between-study heterogeneity (I^2 = 76.7%) (Figure 5). In addition, biopsy was avoided in 41.1% of cases, suggesting a substantial proportion of patients could potentially avoid biopsy as recommended by COMB, despite considerable heterogeneity across studies (I^2 = 98.4%) (Figure 6).

Discussion

Clinical interpretation

The pooled analysis showed that the ACR TI-RADS system had a strong diagnostic performance with a pooled sensitivity of 84.1% (95% CI: 77.5%-89.1%) and a pooled specificity of 76.0% (95% CI: 64.5%-84.7%), resulting in an AUC of 0.874. These results align with the performance criteria of the ACR TI-RADS framework originally described by Tessler et al. [2] and validate the results reported within the individual studies contributing to this analysis. The acknowledged limitation of ACR TI-RADS, that some benign nodules are inappropriately referred for FNA biopsy, which is an acknowledged problem in the contemporary guidelines for the management of thyroid nodules [32], is reflected in the observed false-positive rate of 24.0%. The moderate heterogeneity in the bivariate analysis (Zhou and Dendukuri [16] I^2 = 50.3%) could be due to differences in study design, prevalence of malignant tumors, and differences in interpretation of the TI-RADS system between centers.

The large reduction in specificity seen with COMB can be partly explained by the heterogeneity of the combination strategies used in the included studies. Like the parallel or cumulative designs used by Xu et al. [21] and Qi et al. [24], a positive result on either test was sufficient to label the nodule suspicious, which by nature increases sensitivity and reduces specificity. Sequential reclassification strategies, however, such as those employed by Angelopoulos et al. [31], Hao et al. [25], and Li et al. [19] used elastography to up- or down-classify ACR TI-RADS categories instead of referencing nodules independently. These approaches are theoretically more conservative and clinically appealing, as they incorporate elastography into the existing risk-stratification framework. However, the pooled analysis

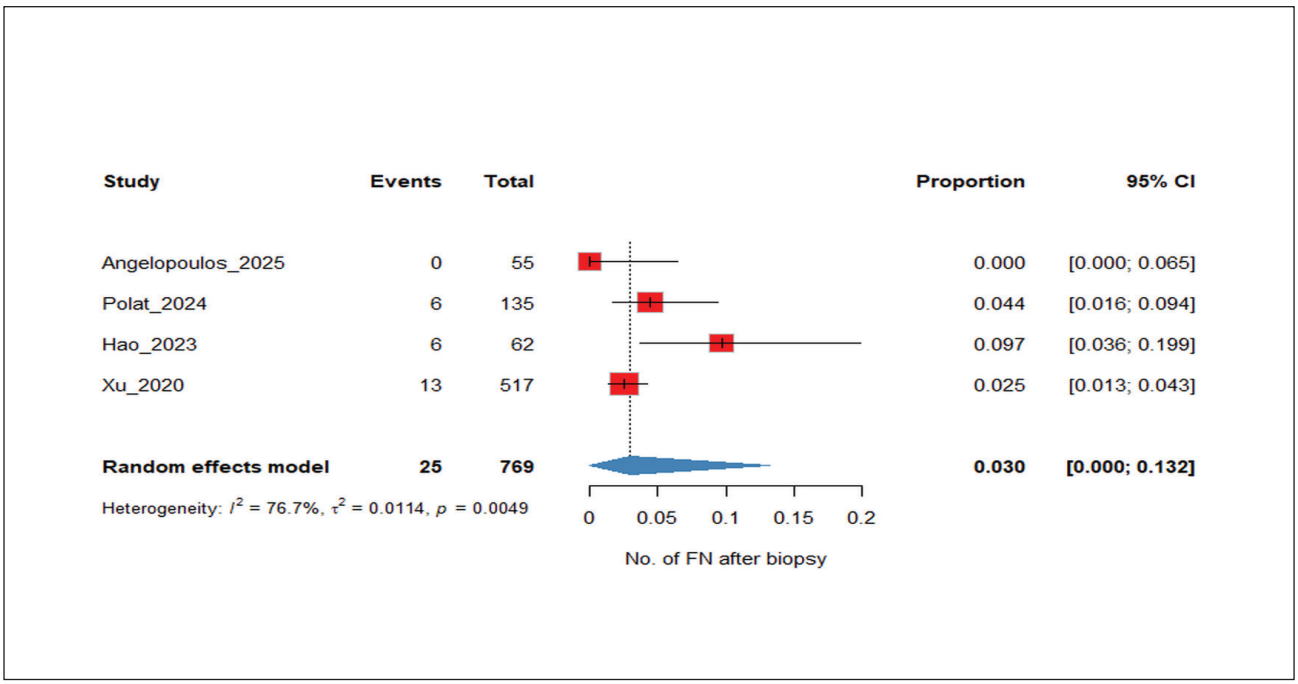


Figure 5. *Forest plot showing the pooled proportion of false-negative results after biopsy.*

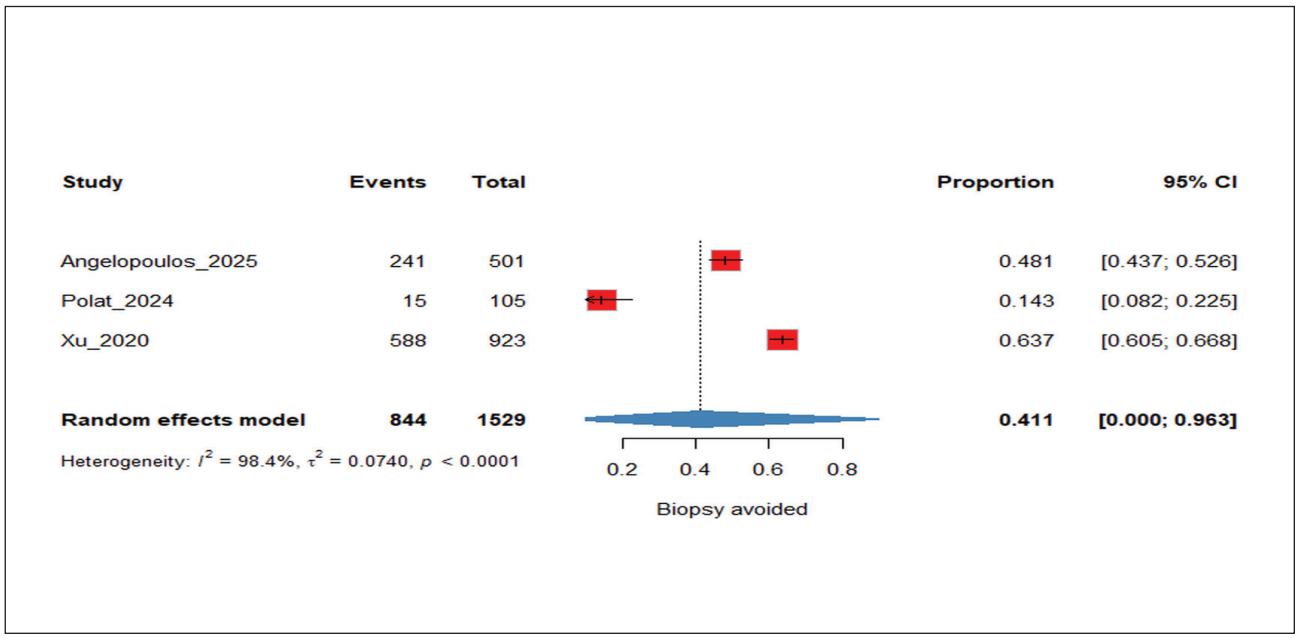


Figure 6. *Forest plot showing the pooled proportion of biopsy avoided as recommended by the combined method of elastography plus ACR TI-RADS (COMB) method.*

still demonstrated significantly low specificity, suggesting that elastography is more reliable for increasing suspicion of malignancy than for excluding malignancy across heterogeneous TI-RADS categories.

One of the most interesting findings of this meta-analysis was the recommendation against biopsy in about 41.1% of the cases using COMB. However, this estimate should be interpreted with caution due to the large heterogeneity across studies (I^2 = 98.4%). The variability may be due to differences in combination strategies, biopsy thresholds, patient selection, and prevalence of underlying malignancy. The combined false-negative rate in our study is generally consistent with acceptable cut-offs suggested by the major thyroid guidelines [32]. However, the high heterogeneity observed (I^2 = 76.7%), and the broad CI limited this estimate. FN rates may also vary between populations because of differences in the incidence of baseline malignancy. For example, some of the Chinese cohorts included in this review reported relatively high prevalence of malignancy [19–22,24,28], which may partially explain the differences in diagnostic performance when compared to studies conducted in lower-risk populations.

The observed no significant difference in diagnostic performance between SWE and SE plus ACR TI-RADS in the meta-regression analysis aligns with the World Federation for Ultrasound in Medicine and Biology (WFUMB) recommendations that both SE and SWE are beneficial adjunctive techniques for evaluating thyroid nodules [33].

Heterogeneity sources

The main source of heterogeneity in the included studies was variability in elastography parameters and cut-off values. The thresholds of SWE ranged from 11.35 kPa in the study of Qi et al. [24] to 52.7 kPa in the study of Hang et al. [22]. Likewise, studies that used SE used different semi-quantitative scoring systems. The absence of a universally accepted protocol for elastography assessment of thyroid nodules still represents a major challenge and a factor hampering the establishment of standardized thresholds for routine clinical practice [33]. Therefore, the current findings support that elastography is most useful when incorporated into pre-defined reclassification algorithms rather than as a solitary indication for biopsy.

Limitations

There are a few limitations to note. First, significant methodological heterogeneity was found in the studies regarding TI-RADS interpretation, elastography modality, cut-off values, and combination strategies, which may limit the reliability of pooled estimates. Second, the prevalence of malignancy at baseline ranged widely (9.9%-73.3%), raising the possibility that spectrum effects might be influencing estimates of sensitivity and specificity. Third, most of the included studies were performed in China, which may limit the generalizability of the results because of regional differences in referral pathways, iodine intake, prevalence of malignancy, and healthcare infrastructure. Fourth, several studies were found to have a high or unclear risk of bias in the flow and timing domain by the QUADAS-2 assessment [12]. Moreover, the bivariate meta-analytic model adequately accounts for the correlation between sensitivity and specificity across studies, but complex and heterogeneous integration strategies between ACR TI-RADS and elastography may not be fully captured by a single pooled diagnostic model, which may contribute to residual unexplained heterogeneity.

Future implications

Future prospective multicenter studies with standardized elastography protocols and internationally validated cut-off values are needed to account for the large heterogeneity seen across current studies. Moreover, the overall elastography utilization in prior studies across all ACR TI-RADS categories limits the clinical interpretability of reported biopsy reduction rates, as the diagnostic utility of elastography may differ by risk category. Therefore, future studies should prospectively evaluate elastography-based reclassification strategies that are selectively applied to intermediate-risk thyroid nodules (TR3–TR4) to maximize biopsy reduction rates and maintain acceptable malignancy detection rates. The current findings also require further validation in different geographic populations.

Conclusion

This systematic review and meta-analysis showed that the ACR TI-RADS technique may have a good and balanced diagnostic performance for evaluating thyroid nodules with reliable sensitivity and specificity. In our study, ultrasound elastography improved sensitivity and decreased the likelihood of missing malignant tumors, but at the cost of a significant drop in specificity and a marked increase in FPs. COMB showed potential to reduce the number of biopsies performed. However, the large heterogeneity observed across studies and the reduction in overall diagnostic accuracy limit the routine application of COMB. Based on current evidence, elastography could be used as a selective adjunct in particular clinical situations, especially when incorporated with structured reclassification approaches. Multicenter prospective studies with standardized elastography protocols are needed to define the optimal role of elastography in the evaluation of thyroid nodules.

Acknowledgments

None.

List of Abbreviations

ACR TI-RADS	American College of Radiology Thyroid Imaging Reporting and Data System
AUC	Area Under the Curve
CI	Confidence interval
COMB	The combined method of elastography plus ACR TI-RADS
ER	Elasticity ratio
FN	False Negative
FNA	Fine-Needle Aspiration
FNAC	Fine-Needle Aspiration Cytology
FP	False Positive
FPR	False Positive Rate
HSROC	Hierarchical Summary Receiver Operating Characteristic
PROSPERO	International Prospective Register of Systematic Reviews
QUADAS-2	Quality Assessment of Diagnostic Accuracy Studies 2
SD	Standard Deviation
SE	Strain Elastography
SWE	Shear Wave Elastography
TN	True Negative
TP	True Positive
TR	TI-RADS Risk Category

Conflict of interest

The authors declare that there is no conflict of interest regarding the publication of this article.

Funding

None.

Consent for publication

Not applicable.

Ethical approval

Not applicable.

Data availability

The data are available upon request from the corresponding author.

Consent to participate

Not applicable.

Author details

Joud Alghamdi[1], Thana Algarni[1], Rama Alawi[1], Malath Alkathiri[1], Arwa Alariany[1], Basmah Sherbini[1], Hams Aljadani[1], Taif Alenezi[1], Dareen Bajamaan[1], Raghad Filimban[1], Sara Aljahdali[1], Amr M. Daous[2]

1. College of Medicine, University of Jeddah, Jeddah, Saudi Arabia
2. Department of Radiology, College of Medicine, University of Jeddah, Jeddah, Saudi Arabia

Supplementary content (If any) is available online.

References

1. Abou Shaar B, Meteb M, Awad El-Karim G, Almalki Y. Reducing the number of unnecessary thyroid nodule biopsies with the American College of Radiology (ACR) Thyroid Imaging Reporting and Data System (TI-RADS). Cureus. 2022;14(3):e23118. https://doi.org/10.7759/cureus.23118
2. 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(5):587–95. https://doi.org/10.1016/j.jacr.2017.01.046
3. 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(1):38–47. https://doi.org/10.2214/AJR.19.22691
4. 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(5):2877–85. https://doi.org/10.1007/s00330-020-07384-6
5. Russ G, Royer B, Bigorgne C, Rouxel A, Bienvenu-Perrard M, Leenhardt L. Prospective evaluation of thyroid imaging reporting and data system on 4550 nodules with and without elastography. Eur J Endocrinol. 2013;168(5):649–55. https://doi.org/10.1530/EJE-12-0936
6. Luo J, Chen J, Sun Y, Xu F, Wu L, Huang P. A retrospective study of reducing unnecessary thyroid biopsy for American College of Radiology Thyroid Imaging Reporting and Data Systems 4 assessment through applying shear wave elastography. Arch Endocrinol Metab. 2020;64(4):349–55. https://doi.org/10.20945/2359-3997000000267
7. Zhang Y, Huang QY, Wu CJ, Chen Q, Xia CJ, Liu BJ, et al. Predicting malignancy in thyroid nodules based on conventional ultrasound and elastography: the value of predictive models in a multi-center study. Endocrine. 2023;80(1):111–23. https://doi.org/10.1007/s12020-022-03271-w
8. Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. BMJ. 2021;372(71):n71. https://doi.org/10.1136/bmj.n71
9. PROSPERO 2026 CRD420261322223: Diagnostic value of adding thyroid elastography to TI-RADS in reducing unnecessary fine-needle aspiration biopsies of thyroid nodules. York (UK): Centre for Reviews and Dissemination, University of York; 2026. Available from: https://www.crd.york.ac.uk/PROSPERO/view/CRD420261322223
10. Richardson WS, Wilson MC, Nishikawa J, Hayward RS. The well-built clinical question: a key to evidence-based decisions. ACP J Club. 1995;123(3):A12–3. https://doi.org/10.7326/ACPJC-1995-123-3-A12
11. Ouzzani M, Hammady H, Fedorowicz Z, Elmagarmid A. Rayyan-a web and mobile app for systematic reviews. Syst Rev. 2016;5(1):210. https://doi.org/10.1186/s13643-016-0384-4
12. Whiting PF, Rutjes AW, Westwood ME, Mallett S, Deeks JJ, Reitsma JB, et al.; QUADAS-2 Group. QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. Ann Intern Med. 2011;155(8):529–36. https://doi.org/10.7326/0003-4819-155-8-201110180-00009
13. Monaghan TF, Rahman SN, Agudelo CW, Wein AJ, Lazar JM, Everaert K, et al. Foundational statistical principles in medical research: sensitivity, specificity, positive predictive value, and negative predictive value. Medicina (Kaunas). 2021;57(5):503. https://doi.org/10.3390/medicina57050503
14. Negeri ZF. A bivariate finite mixture random effects model for identifying and accommodating outliers in diagnostic test accuracy meta-analyses. Biom J. 2025 Jun;67(3):e70062. https://doi.org/10.1002/bimj.70062
15. Rutter CM, Gatsonis CA. A hierarchical regression approach to meta-analysis of diagnostic test accuracy evaluations. Stat Med. 2001;20(19):2865–84. https://doi.org/10.1002/sim.942
16. Zhou Y, Dendukuri N. Statistics for quantifying heterogeneity in univariate and bivariate meta-analyses of binary data: the case of meta-analyses of diagnostic accuracy. Stat Med. 2014;33(16):2701–17. https://doi.org/10.1002/sim.6115
17. Holling H, Böhning W, Masoudi E, Böhning D, Sangnawakij P. Evaluation of a new version of I2 with emphasis on diagnostic problems. Commun Stat Simul Comput. 2020;49(4):942–72. https://doi.org/10.1080/03610918.2018.1489553
18. R: The R Project for Statistical Computing [Internet]. [cited 2026 Sep 2]. Available from: https://www.r-project.org/
19. Li X, Hou XJ, Du LY, Wu JQ, Wang L, Wang H, et al. Virtual Touch Tissue Imaging and Quantification (VTIQ) combined with the American College of Radiology Thyroid Imaging Reporting and Data System (ACR TI-RADS) for malignancy risk stratification of thyroid nodules. Clin Hemorheol Microcirc. 2019;72(3):279–91. https://doi.org/10.3233/CH-180477
20. Pei S, Zhang B, Cong S, Liu J, Wu S, Dong Y, et al. Ultrasound real-time tissue elastography improves the diagnostic performance of the ACR Thyroid Imaging Reporting and Data System in Differentiating Malignant from Benign Thyroid Nodules: a summary of 1525 Thyroid Nodules. Int J Endocrinol. 2020;2020:1749351. https://doi.org/10.1155/2020/1749351
21. Xu L, Zhou Y, Li Y, Lu B, Liu T. reducing unnecessary biopsy of American College of Radiology Thyroid Imaging Reporting and Data System Category 4 Nodules. J Ultrasound Med. 2021;40(2):227–36. https://doi.org/10.1002/jum.15391

22. Hang J, Li F, Qiao XH, Ye XH, Li A, Du LF. combination of maximum shear wave elasticity modulus and TIRADS improves the diagnostic specificity in characterizing thyroid nodules: a retrospective study. Int J Endocrinol. 2018;2018:4923050. https://doi.org/10.1155/2018/4923050

23. Paredes-Manjarrez C, Arreola-Cháidez D, Magdalena-Buitrago A, Ferreira-Hermosillo A, Avelar-Garnica JF, Arreola-Rosales R. Shear-wave elastography as a tool in the assessment of thyroid nodules. Gac Med Mex. 2021;157(1):18–23. https://doi.org/10.24875/GMM.20005656

24. Qi WH, Jin K, Cao LL, Peng M, He NA, Zhan XL, et al. Diagnostic performance of a new two-dimensional shear wave elastography expression using siemens ultrasound system combined with ACR TI-RADS for classification of benign and malignant thyroid nodules: A prospective multi-center study. Heliyon. 2023;9(10):e20472. https://doi.org/10.1016/j.heliyon.2023.e20472

25. Hao L, Liu P, Ding C, Li J, Zhang Y. Diagnostic value of ACR TI-RADS combined with three-dimensional shear wave elastography in ACR TI-RADS 4 and 5 thyroid nodules. Chin Med J (Engl). 2023;136(10):1225–30. https://doi.org/10.1097/CM9.0000000000002655

26. Yi AJ, Yang WW, Cui XW, Dietrich CF, Wang B. The value of quantitative and a new qualitative color pattern shear wave elastography for the differentiation of ACR TI-RADS 4 or 5 category thyroid nodules measuring ≤10 mm. Front Endocrinol (Lausanne). 2024;14:1275256. https://doi.org/10.3389/fendo.2023.1275256

27. Polat Z, Elmalı M, Tanrivermis Sayit A, Kalkan C, Danacı M, Kefeli M. Comparative evaluation of shear wave elastography elasticity values in thyroid nodules with cytology results and TI-RADS scoring in differentiation of benign-malignant nodules. Eur Arch Oto-Rhino-Laryngol Off J Eur Fed Oto-Rhino-Laryngol Soc EUFOS Affil Ger Soc Oto-Rhino-Laryngol -. Head Neck Surg. 2024;281(5):2609–17. https://doi.org/10.1007/s00405-024-08516-0

28. Zhang WB, Xu W, Fu WJ, He BL, Liu H, Deng WF. Comparison of ACR TI-RADS, Kwak TI-RADS, ATA guidelines and KTA/KSThR guidelines in combination with SWE in the diagnosis of thyroid nodules. Clin Hemorheol Microcirc. 2021;78(2):163–74. https://doi.org/10.3233/CH-201021

29. Sekar S, Adithan S, Gantait S, Siddaraju N, Shankar G. Diagnostic Performance of Thyroid Imaging Reporting and Data system (TIRADS) and Acoustic Radiation Force Impulse (ARFI) imaging in Detection of Malignant Thyroid Nodules: A Cross-sectional Study. J Clin Diagn Res. 2025;19(3):1. https://doi.org/10.7860/JCDR/2025/73120.20708

30. Ediz SŞ, Tarhan M. Kabaalioğlu A. A prospective study; diagnostic role of shear-wave elastography to differentiate benign and non-benign thyroid nodules categorized to according to the ACR 2017 TI-RADS. Bezmialem Sci [Internet]. 2023 [cited 2026 Sep 2]. Available from: https://bezmialemscience.org/articles/a-prospective-study-diagnostic-role-of-shear-wave-elastography-to-differentiate-benign-and-non-benign-thyroid-nodules-categorized-to-according-to-the-acr-2017-ti-rads/bas.galenos.2023.72623

31. Angelopoulos N, Goulis DG, Chrisogonidis I, Iakovou I, Tessler FN. Elastography enhances diagnostic accuracy of ACR TI-RADS in thyroid nodule evaluation. J Clin Endocrinol Metab. 2026;111(6):e1616–25. https://doi.org/10.1210/clinem/dgaf670

32. Haugen BR, Alexander EK, Bible KC, Doherty GM, Mandel SJ, Nikiforov YE, et al. 2015 American Thyroid Association Management guidelines for adult patients with thyroid nodules and differentiated thyroid cancer: the American Thyroid Association guidelines task force on thyroid nodules and differentiated thyroid cancer. Thyroid. 2016;26(1):1–133. https://doi.org/10.1089/thy.2015.0020

33. Cosgrove D, Barr R, Bojunga J, Cantisani V, Chammas MC, Dighe M, et al. WFUMB Guidelines and Recommendations on the Clinical Use of Ultrasound Elastography: Part 4. Thyroid. Ultrasound Med Biol. 2017;43(1):4–26. https://doi.org/10.1016/j.ultrasmedbio.2016.06.022