

RESEARCH ARTICLE OPEN ACCESS

Impact of Diffusion-Weighted Magnetic Resonance Imaging Parameters on Diagnostic Accuracy for Thyroid Nodules: A Systematic Review and Meta-Analysis

Caterina Giannitto^{1,2}  | Angela Ammirabile^{1,2} | Giorgia Carnicelli^{1,2} | Ludovica Lofino¹ | Andrea Costantino^{2,3} | Fabrizio Natali⁴ | Andrea Alessandro Esposito⁵ | Armando De Virgilio^{2,3} | Antonio Lo Casto⁶ | Giovanni Savini^{1,2} | Lorenzo Ugga⁷ | Giuseppe Mercante^{2,3} | Letterio Salvatore Politi^{1,2}  | Steve Connor⁸

¹Department of Diagnostic and Interventional Radiology, IRCCS Humanitas Research Hospital, Milan, Italy | ²Department of Biomedical Sciences, Humanitas University, Milan, Italy | ³Otorhinolaryngology Unit, IRCCS Humanitas Research Hospital, Milan, Italy | ⁴School of Medicine and Surgery, University of Milano-Bicocca, Milan, Italy | ⁵Department of Diagnostic Radiology, ASST, Bergamo, Italy | ⁶Sezione di Scienze Radiologiche, DIBIMEF, Università Degli Studi di Palermo, Palermo, Italy | ⁷Department of Advanced Biomedical Sciences, University of Naples "Federico II", Naples, Italy | ⁸Department of Neuroradiology, King's College Hospital, London, UK

Correspondence: Caterina Giannitto (caterina.giannitto@gmail.com)

Received: 25 June 2025 | **Revised:** 16 September 2025 | **Accepted:** 22 September 2025

Keywords: data accuracy | diffusion-weighted MRI | meta-analysis | thyroid nodule

ABSTRACT

Background: Surgery is the gold standard to differentiate benign from malignant thyroid nodules, but it is invasive and often unnecessary in indeterminate cases. Diffusion-Weighted MRI (DW-MRI) has emerged as a promising, non-invasive tool, though its accuracy and the impact of acquisition parameters remain unclear.

Purpose: To evaluate the diagnostic accuracy of DW-MRI in distinguishing malignant from benign thyroid nodules and identifying influencing acquisition parameters.

Study Type: Systematic Review and Meta-analysis (researchregistry11482).

Population: 2073 patients, 2403 thyroid nodules (1067 malignant).

Field Strength/Sequence: DW-MRI at 3.0T or 1.5T.

Assessment: A systematic search of Pubmed, Embase, Cochrane Library, Scopus, and Web of Science was conducted through July 2025 following the Preferred Reporting Items for Systematic Reviews and Meta-Analyses of Diagnostic Test Accuracy guidelines. Studies on DW-MRI in thyroid nodules were included. Risk of bias and applicability were assessed using QUADAS-2.

Statistical Tests: A bivariate random-effects model estimated pooled sensitivity (SE), specificity (SP), and area under the curve (AUC). Univariable and multivariable meta-regressions explored the influence of DW-MRI parameters. Subgroup analyses and pooled Apparent Diffusion Coefficient (ADC) comparisons were also performed. Results were considered statistically significant at $p < 0.05$.

Results: Thirty-seven studies were included. DW-MRI showed pooled SE 0.84, SP 0.87, and AUC 0.91, confirmed in studies at low risk of bias (AUC 0.80). Univariable meta-regression revealed that 3.0T field strength reduced SP, while < 5 averages and acquisition matrix ≥ 130 increased SE. Nodule size ≥ 10 mm and circular region of interest improved SP. Multivariable analysis confirmed increased SE with b -values > 2 and improved SP with $b \geq 1000$, while 3.0T remained associated with reduced SP. Malignant nodules showed lower ADC (1.08 vs. 1.73×10^{-3} mm²/s vs. 1.70×10^{-3} mm²/s for benign), especially with $b \geq 1000$.

This is an open access article under the terms of the [Creative Commons Attribution-NonCommercial-NoDerivs](https://creativecommons.org/licenses/by-nc-nd/4.0/) License, which permits use and distribution in any medium, provided the original work is properly cited, the use is non-commercial and no modifications or adaptations are made.

© 2025 The Author(s). *Journal of Magnetic Resonance Imaging* published by Wiley Periodicals LLC on behalf of International Society for Magnetic Resonance in Medicine.

Data Conclusion: DW-MRI shows good diagnostic accuracy for thyroid nodule assessment, although it is influenced by both technical and methodological factors.

Level of Evidence: 3.

Technical Efficacy: Stage 3.

1 | Introduction

The initial diagnostic assessment of thyroid nodules aims to differentiate benign lesions, which may be managed conservatively, from those with malignant potential requiring surgical intervention [1]. Among imaging modalities, ultrasound (US) is currently the most widely used and validated first-line tool for thyroid nodule evaluation, owing to its accessibility, non-invasiveness, and cost-effectiveness [1]. Multiple risk stratification systems, Thyroid Imaging Reporting and Data System (TI-RADS), have been developed to standardize malignancy risk assessment based on sonographic features [2]. A recent network meta-analysis of 39 studies involving nearly 50,000 patients demonstrated that the American College of Radiology TI-RADS had the highest diagnostic performance, with sensitivities (SEs) and specificities (SPs) at optimal thresholds ranging from 64% to 77% and 82% to 90%, respectively [2]. For thyroid nodules of indeterminate risk, different TI-RADS demonstrate consistently high SE (97%–99%) but limited SP (3%–31%) [2]. Likewise, conventional US combined with elastography has shown pooled SE and SP of 0.88 and 0.86, with an area under the Receiver Operating Characteristic (ROC) curve of 0.92, confirming the strong, but not absolute, performance of US-based methods [3]. Fine-needle aspiration cytology (FNAC) is a diagnostic mainstay due to its accuracy, reproducibility, and cost-effectiveness [4]. However, in cases of indeterminate cytology, FNAC fails to reliably distinguish benign from malignant lesions, often leading to unnecessary surgery for definitive histopathological diagnosis in indeterminate nodules [5, 6]. Notably, up to 80% of these surgically excised nodules are ultimately benign, raising concerns regarding overtreatment and cost-effectiveness [7].

Diffusion-weighted-MRI (DW-MRI) has emerged as a promising non-invasive tool for tissue characterization. DW-MRI reflects the degree of Brownian motion of water molecules in tissues, with malignant lesions typically exhibiting restricted diffusion and lower apparent diffusion coefficient (ADC) values due to high cellularity. Multiple studies have shown significantly lower ADC values in malignant thyroid nodules compared to benign ones, suggesting a potential role for DW-MRI in preoperative diagnostic pathways [8–11].

However, heterogeneity in imaging protocols (*b*-values, ADC calculation methods), and study populations have limited the clinical applicability of DW-MRI for thyroid nodules assessment. Existing meta-analyses have not comprehensively addressed how technical factors influence diagnostic performance [8, 9].

Therefore, the aim of this study was to systematically evaluate the diagnostic accuracy of DW-MRI in differentiating benign from malignant thyroid nodules and to assess the impact of technical parameters on diagnostic performance across published studies.

2 | Materials and Methods

This systematic review and meta-analysis was conducted and reported in accordance with the PRISMA-DTA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses of Diagnostic Test Accuracy Studies) guidelines [12].

2.1 | Search Strategy and Information Sources

A comprehensive search was conducted in PubMed, EMBASE, Cochrane CENTRAL, Web of Science, and Scopus from inception to July 24, 2025. To ensure a systematic and focused literature review, the research question was framed according to the PIT (Population–Index test–Target condition) framework. The Population comprised patients with thyroid nodules. The Index test was DW-MRI using ADC cut-offs to dichotomize nodules as test-positive vs. test-negative. The target condition was thyroid malignancy, as determined by the reference standard of histopathology (preferred) or cytology. The full electronic search strategies for each database, including limits and filters, are reported in (Table S1). Additional studies were identified through backward citation searching and review of related articles.

2.2 | Eligibility Criteria

The inclusion and exclusion criteria were defined according to the PIT framework and the PRISMA recommendations.

2.3 | Inclusion Criteria

- Population (P): Patients with benign or malignant thyroid nodules, confirmed by histological or cytological examination.
- Index test (I): DW-MRI with field strength ≥ 1.5 Tesla, including studies reporting ADC values with defined cut-off thresholds for differentiating benign from malignant nodules. DW-MRI was considered positive when the reported ADC value was below the study-specific cut-off (or above, if defined accordingly), as prespecified by each primary study.
- Target condition (T): Malignant thyroid nodules.
- Outcome measures: Diagnostic performance indicators such as SE, SP, accuracy, area under the curve (AUC), or other relevant metrics.
- Language: Articles published in English.
- Availability: Full-text articles, published and peer-reviewed.

Plain Language Summary

This study investigated a potential tool to diagnose thyroid nodules. Thyroid nodules are lesions in the thyroid gland. Some are benign, while others may be cancerous. Doctors often use ultrasound and cell sampling, but many nodules remain unclear. In this study, researchers tested an advanced type of magnetic resonance imaging called diffusion-weighted MRI (DW-MRI). This method showed differences between benign and suspicious nodules, and some technical parameters influenced the accuracy. The results suggest that DW-MRI can support doctors in choosing the right treatment, reducing unnecessary surgery and improving care for people with thyroid nodules.

2.4 | Exclusion Criteria

- Conference abstracts, unpublished manuscripts, or non-peer-reviewed sources.
- Studies not written in English.
- Studies that did not report outcomes of interest (e.g., SE, SP, AUC) and ADC cut-off values or did not allow comparison between benign and malignant nodules.

- Studies focused on non-nodular thyroid lesions or non-human populations.
- The included studies were grouped according to population type (benign vs. malignant nodules) and intervention (DW-MRI), relative to the reference standard of histological or cytological confirmation, in alignment with the review's objective. Restrictions such as English language and exclusion of abstracts were adopted to ensure methodological rigor and the reliability of the analyzed data.

2.5 | Study Selection

Two reviewers (A.A. and L.F.) independently screened titles and abstracts for relevance, retrieved full texts, and applied inclusion/exclusion criteria. Discrepancies were documented and addressed according to predefined eligibility criteria. After revision of discrepancies, full agreement was achieved. A PRISMA flow diagram details the selection process (Figure 1).

2.6 | Data Extraction

Data were extracted independently by two reviewers (A.A. and L.F.), radiologists with 2 years of experience in head and neck imaging, using a standardized and piloted data collection form.

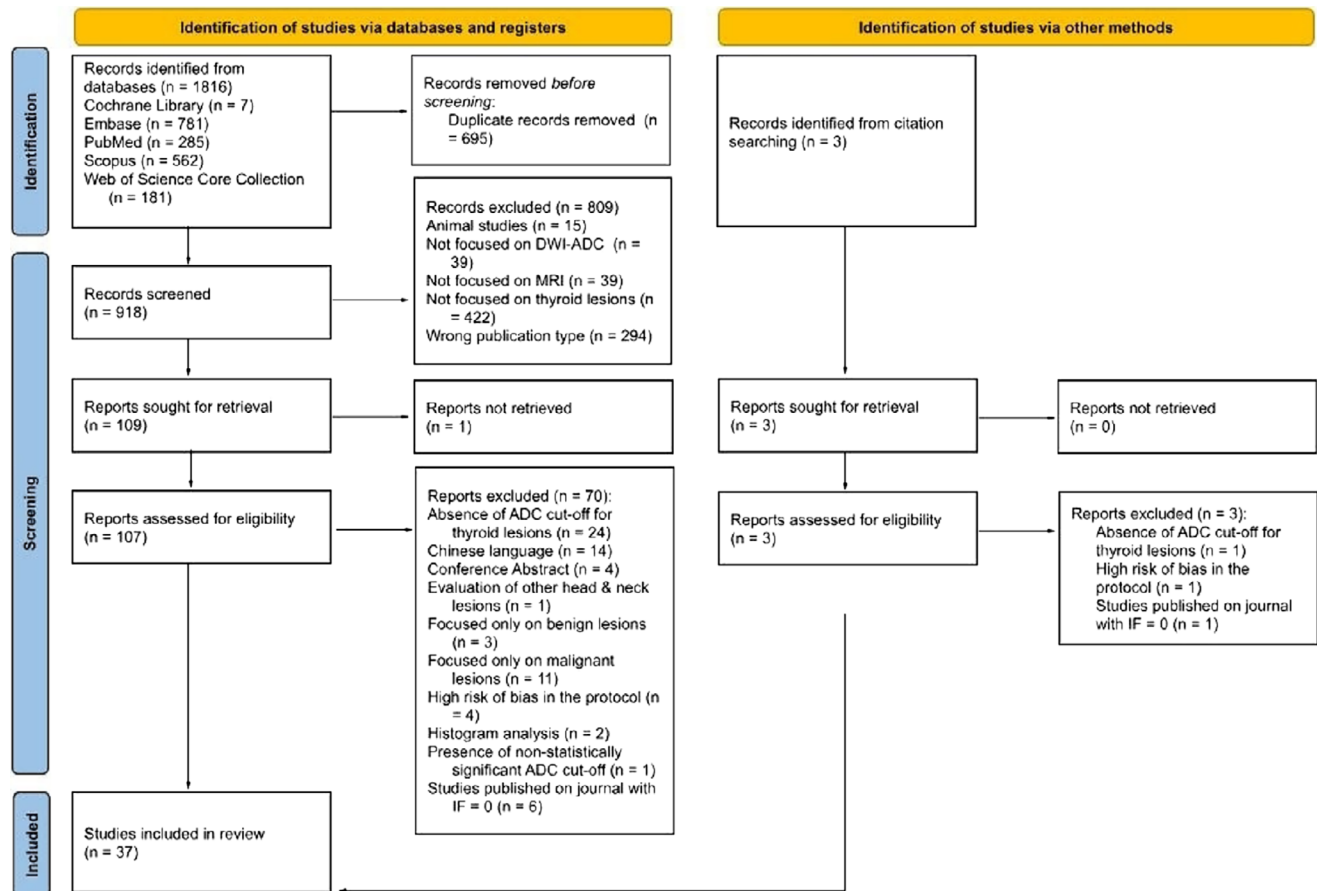


FIGURE 1 | Literature search and study selection process flowchart.

Extracted variables included:

- Study characteristics: design, publication year; impact factor of the journal.
- Lesion data: total number of nodules, benign/malignant ratio, histologic subtypes, minimum thyroid nodule size;
- Reference standard: histopathology, FNAC/fine needle biopsy (FNAB) or both.
- DW-MRI protocol details: field strength, b -values, number of b values, section thickness, acquisition matrix size, Time of Repetition (TR)/Time of Echo (TE), Field of View (FOV), number of averages, slice gap, coil type, and sequence.
- ROI characteristics for ADC evaluation: delineation method, inclusion/exclusion of cystic portions, segmentation of entire lesion (whole lesion volume); type of ROI used (free-hand or circular).
- ADC values: means and standard deviations for benign and malignant thyroid nodules.
- Diagnostic accuracy data: true positives, false positives, true negatives, false negatives, SE, SP.

When multiple b -values or cut-offs were used in the same population, each analysis was considered separately.

2.7 | Quality Assessment

Risk of bias and applicability concerns were assessed independently by two reviewers (A.A. and L.L.) using the QUADAS-2 tool. Disagreements were resolved by consensus.

2.8 | Statistical Analysis

All statistical analyses were performed using R software (version 4.4.2) with the mada and meta packages. A bivariate random-effects model was applied to jointly estimate pooled SE and SP, accounting for their natural correlation across studies. This model also provided the summary ROC curve, 95% confidence interval, and the AUC as a global indicator of diagnostic accuracy.

Between-study heterogeneity was assessed using I^2 statistics based on multiple methods Zhou and Dendukuri approach [13], Holling sample size-unadjusted method, and Holling sample size-adjusted method [14, 15].

A SE analysis restricted to studies judged at low risk of bias according to QUADAS-2 was also conducted to assess the robustness of the findings.

Univariable bivariate meta-regression was performed for the following covariates, which were most frequently reported across the included studies: b -values, number of b -values, field strength (1.5T vs. 3T), section thickness, acquisition matrix size, number of signal averages, slice gap, repetition time (TR), echo time (TE), field of view (FOV), thyroid nodule size cut-off, thyroid nodule segmentation method, and type of ROI used (free-hand or circular). A multivariable bivariate meta-regression including variables with minimal missing data (b values, number of b values, field

strength, section thickness) and QUADAS-2 evaluation (low and not low) was also performed using a bivariate random-effects approach. A p -value < 0.05 was considered statistically significant.

A single-group meta-analysis was conducted on studies that reported extractable mean ADC values to estimate pooled ADCs for benign and malignant thyroid nodules, respectively.

3 | Results

3.1 | Study Selection and Study Characteristics

The search resulted in 1816 articles; after removing duplicates, 918 studies were screened for titles and abstracts. Of the 109 retrieved full texts, 37 articles were included in the analysis [10, 11, 16–50] (Figure 1).

A general overview of the study characteristics is presented in Table 1. A total of 2073 patients and 2403 thyroid nodules, including 1067 malignant, were included. The study design was prospective in 22 studies, retrospective in 11, and not clearly specified in 4 studies.

In 20 studies DW-MRI was performed at 1.5T [10, 11, 16–18, 22, 27, 29, 31–37, 39–41, 46, 49], while 17 studies used a 3.0T scanner [19–21, 23–26, 28, 30, 38, 42–45, 47, 48, 50]. Technical parameters from each study are displayed in Tables S2 and S3.

3.1.1 | Risk of Bias and Applicability

Figure 2 shows the distribution of QUADAS-2 scores for the risk of bias and applicability concerns.

Twelve studies (32.4%) showed a low risk of bias in all items without applicability concerns [11, 19–21, 25, 31, 34, 38, 42, 46, 47, 50].

Concerning the reference standard bias, seven studies (18.9%) demonstrated an unclear risk of bias in classifying the target condition due to a lack of information on whether cytology or histology was performed [30, 35–37, 45, 49]; 15 studies (40.5%) showed a high risk of reference standard bias due to the use of different pathological techniques within the same cohort, for example, use of FNAC and surgery or FNAB, surgery, and follow-up [16, 18, 22, 24, 27–29, 32, 33, 39–41, 43, 44, 48].

In the flow and timing domain of the QUADAS-2 tool, which assesses whether all patients received the same reference standard, whether there was an appropriate interval between index test and reference standard, and whether all patients were included in the analysis, we reported an unclear risk of bias in 11 studies (29.7%) [23, 27–30, 32, 35, 37, 39, 45, 49] and a high risk in two studies (5.4%) [16, 24].

Applicability concerns were present in nine studies (24.3%) which utilized cytology as the reference standard [16, 18, 24, 27, 29, 30, 32, 36, 40]. In cases where insufficient details were provided, or FNAC/FNAB without surgery was used (33.3%), the risk of applicability concerns remained unclear [22, 28, 33, 35, 37, 39, 41, 43–45, 48, 49].

TABLE 1 | Baseline characteristics of the included studies.

Author (year)	Study design	Number of patients (male)	Age (mean)	Thyroid nodules	Malignant nodules	Reference standard
Abdel Rahman 2016 [16]	Prospective	60 (12)	n.s.	60	12	FNAC or FNAB
Abdel Razek 2008 [17]	Prospective	63 (19)	47	63	7	Surgery
Abdelgawad 2020 [18]	Prospective	73 (27)	36	73	15	FNAB
Aghaghazvini 2018 [19]	Cross-sectional	26 (7)	38	41	15	Surgery
Bayraktaroglu 2019 [20]	Prospective	32 (12)	52	76	13	Surgery
Brown 2016 [21]	Prospective	26	57.1	24	18	Surgery
El-Hariri 2012 [22]	Prospective	37 (9)	42	56	19	FNAB or Surgery
Hao 2016 [23]	Retrospective	93 (21)	41.7	101	66	Surgery
Ilica 2013 [24]	Retrospective	25 (11)	42.3	28	10	FNAB, Surgery, follow up
Jiang 2022 [25]	Prospective	48 (12)	n.s.	53	26	Surgery
Jiang 2024 [26]	n.s.	90 (22)	47.69	104	53	Surgery
Khizer 2015 [27]	Cross-sectional	35	n.s.	64	14	Histopathology
Kong 2019 [28]	Retrospective	100 (21)	49	137	86	Unclear
Latif 2021 [29]	Prospective	56 (20)	41.4	56	20	FNAC or Surgery
Li 2020 [30]	n.s.	44 (16)	42	44	24	Biopsy proven
Linh 2019 [31]	Prospective	93	n.s.	128	49	Histopathology
Monisha 2022 [32]	Observational	43 (8)	49.9	43	8	FNAC or FNAB
Mutlu 2012 [10]	Prospective	44 (17)	49	51	5	FNAB
Nakahira 2012 [11]	Retrospective	38 (13)	55.5	42	19	Surgery
Noda 2015 [33]	Retrospective	36 (10)	57.7	42	14	FNAB or Surgery
Özer 2024 [34]	Retrospective	26 (9)	48.62	46	9	Surgery
Sasaki 2013 [35]	Prospective	23 (8)	65	23	16	Histopathology
Shayganfar 2022 [36]	Prospective	33	50.1	37	22	FNAC
Shi 2013 [37]	Prospective	111 (36)	50	111	88	Surgery, Clinical examination
Shi 2017 [38]	Prospective	58 (11)	48.1	58	34	Surgery
Shi 2019 [39]	Retrospective	87 (27)	49.5	87	52	FNAC or FNAB
Shokry 2018 [40]	Retrospective	30 (10)	53.8	26	13	FNAC
Taha Ali 2017 [41]	Prospective	42 (15)	45.4	42	14	FNAB or Surgery
Tang 2022 [42]	Prospective	77 (22)	46	77	41	Surgery
Wang 2018 [43]	n.s.	85 (23)	47.4	114	52	FNAB or Surgery
Wang 2019 [44]	n.s.	92 (23)	47	123	65	FNAB or Surgery
Wang 2023 [45]	Retrospective	19 (5)	48.8	21	15	Histopathology
Wu 2013 [46]	Prospective	42 (10)	42.4	42	14	Surgery

(Continues)

TABLE 1 | (Continued)

Author (year)	Study design	Number of patients (male)	Age (mean)	Thyroid nodules	Malignant nodules	Reference standard
Yuan 2022 [47]	Prospective	95	n.s.	106	64	Surgery
Zhang 2021 [48]	Retrospective	80 (31)	46.5	80	43	FNAB or Surgery
Zhou 2019 [49]	Prospective	33 (7)	52.2	45	15	Histopathology
Zhu 2022 [50]	Prospective	78 (24)	53	79	17	Surgery

Abbreviations: FNAB: fine needle aspiration biopsy, FNAC: fine needle aspiration cytology, n.s: not specified.

3.1.2 | Diagnostic Accuracy

A total of 2403 nodules were used for the calculation of diagnostic accuracy parameters, with a prevalence of 44.4% for malignant lesions ($n = 1067$).

3.1.2.1 | Bivariate Meta-Analyses With Random Effects. A bivariate random-effects meta-analysis was performed to jointly estimate SE and false positive rate (FPR). The pooled SE was 0.84 (95% CI: 0.80–0.86), and the FPR was 0.129 (95% CI: 0.108–0.152). SP was calculated as $(1 - \text{FPR})$, resulting in a value of 0.87 (95% CI: 0.85–0.89). The AUC was 0.91, indicating excellent overall diagnostic performance (Figure 3). The between-study standard deviation was 0.291 for SE and 0.397 for FPR, with a weak negative correlation between them ($r = -0.169$). Heterogeneity was negligible, with $I^2 = 2.5\%$ using the Zhou and Dendukuri method and 1.2%–1.7% using the sample size-adjusted Holling approach, while the unadjusted Holling estimates ranged from 19.5% to 33.7%.

Forest plots for SE and SP were generated separately (Figure 4). Each plot displays the individual study point estimates along with their 95% confidence intervals. This visualization highlights the variability across studies and the associated uncertainty for each metric, without computing or displaying a pooled summary estimate, in line with the bivariate model used in the analysis.

The bivariate random-effects meta-analysis used to jointly estimate SE and FPR repeated in 11 studies [11, 19–21, 25, 31, 34, 38, 42, 46, 47, 50] judged at low risk of bias according to QUADAS-2 showed pooled SE of 0.81 (95% CI: 0.76–0.85), SP of 0.87 (95% CI: 0.81–0.92). The AUC was 0.80, indicating excellent overall diagnostic performance (Figure 3). The between-study standard deviation was 0.004 for SE and 0.721 for FPR, with a perfect negative correlation between them ($r = -1.000$). Heterogeneity was negligible, with $I^2 = 0.2\%$ using the Zhou and Dendukuri method and 0.8%–0.9% using the sample size-adjusted Holling approach, while the unadjusted Holling estimates ranged from 17% to 19.7%.

3.1.2.2 | Subgroup Analysis. Field strength was not significantly associated with SE ($p = 0.74$). By contrast, a significant association was observed with SP, with higher field-strength values (3T) being related to an increased FPR. These findings suggest that field strength may affect SP, while SE remains unaffected.

The number of averages was significantly associated with SE, but not with SP ($p = 0.97$); studies using fewer than 5 averages showed higher SE.

Acquisition matrix size was positively associated with SE, indicating that larger matrices (≥ 130) were related to increased SE, whereas no significant effect was observed on SP ($p = 0.82$).

The bivariate random-effects model revealed no significant association between the nodule cut-off and SE ($p = 0.57$). However, higher cut-off values were significantly associated with SP, leading to a reduction in false-positive cases.

A significant association was found between ROI delineation method and SP, with circular ROIs showing higher specificity compared with freehand ROIs. No significant effect on SE was observed ($p = 0.70$).

Overall, this univariate subgroup analysis did not identify a statistically meaningful influence of highest b value, number of b values, section thickness, gap, TE, TR, FOV, segmentation on either SE or SP (Tables 2 and 3 and bivariate meta-regression scatter plot with confidence ellipses for each covariate in the Supporting Information).

When the analysis was restricted to studies judged at low risk of bias according to QUADAS-2, subgroup analysis by covariates could not be performed because these studies showed no variability across the investigated parameters, precluding a meaningful univariable analysis.

The multivariable meta-regression showed that SE significantly increased when more than two b -values were used compared with only two b -values, holding other covariates constant. The FPR was significantly reduced when b -values ≥ 1000 were applied compared with b -values < 1000 , indicating a corresponding increase in SP. Conversely, field strength was significantly associated with a higher FPR at 3.0T compared with 1.5T, reflecting a reduction in SP, while other covariates showed no significant associations (Table 4).

Thirty-three studies were selected for meta-analysis on the overall mean ADC values of benign and malignant lesions, respectively [10, 11, 16–20, 22–31, 33–48]; a total of 2270 nodules of which benign and 1011 malignant nodules were evaluated.

STUDY	RISK OF BIAS				APPLICABILITY CONCERNS		
	PATIENT SELECTION	INDEX TEST	REFERENCE STANDARD	FLOW AND TIMING	PATIENT SELECTION	INDEX TEST	REFERENCE STANDARD
Abdel-Rahman 2016	+	?	-	-	+	+	-
Abdel-Razek 2008	+	?	+	+	+	+	+
Abdelgawad 2020	+	+	-	+	+	+	-
Aghaghazvini 2018	+	+	+	+	+	+	+
Bayraktaroglu 2019	+	+	+	+	+	+	+
Brown 2016	+	+	+	+	+	+	+
El-Hariri 2012	+	+	-	+	+	+	?
Hao 2016	+	+	+	?	+	+	+
Ilica 2013	+	+	-	-	+	+	-
Jiang 2022	+	+	+	+	+	+	+
Jiang 2024	+	?	+	+	+	+	+
Khizer 2015	+	+	-	?	+	+	-
Kong 2019	+	+	-	?	+	+	?
Latif 2021	+	+	-	?	+	+	-
Li 2020	?	-	?	?	+	+	-
Linh 2019	+	+	+	+	+	+	+
Monisha 2022	+	?	-	?	+	+	-
Mutlu 2012	?	-	+	+	+	+	+
Nakahira 2012	+	+	+	+	+	+	+
Noda 2015	+	+	-	+	+	+	?
Ozer 2024	+	+	+	+	+	+	+
Sasaki 2013	+	?	?	?	+	+	?
Shayganfar 2022	+	+	?	+	+	+	-
Shi 2013	+	+	?	?	+	+	?
Shi 2017	+	+	+	+	+	+	+
Shi 2019	+	+	-	?	+	+	?
Shokry 2018	+	+	-	+	+	+	-
Taha Ali 2017	+	?	-	+	+	+	?
Tang 2022	+	+	+	+	+	+	+
Wang 2018	+	+	-	+	+	+	?
Wang 2019	+	+	-	+	+	+	?
Wang 2023	+	+	?	?	+	+	?
Wu 2013	+	+	+	+	+	+	+
Yuan 2022	+	+	+	+	+	+	+
Zhang 2021	+	+	-	+	+	+	?
Zhou 2019	+	+	?	?	+	+	?
Zhu 2022	+	+	+	+	+	+	+

Low Risk High Risk Unclear Risk

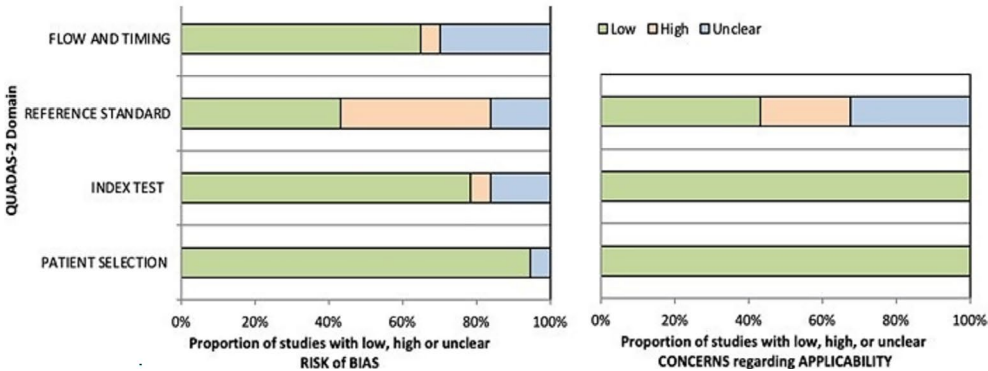


FIGURE 2 | Methodological quality of the included studies assessed according to the Quality. Assessment of Diagnostic Accuracy Studies 2 tool for risk of bias and applicability concerns. The green circle represents the low risk of bias, the yellow circle the unclear risk of bias, and the red circle the high risk of bias.

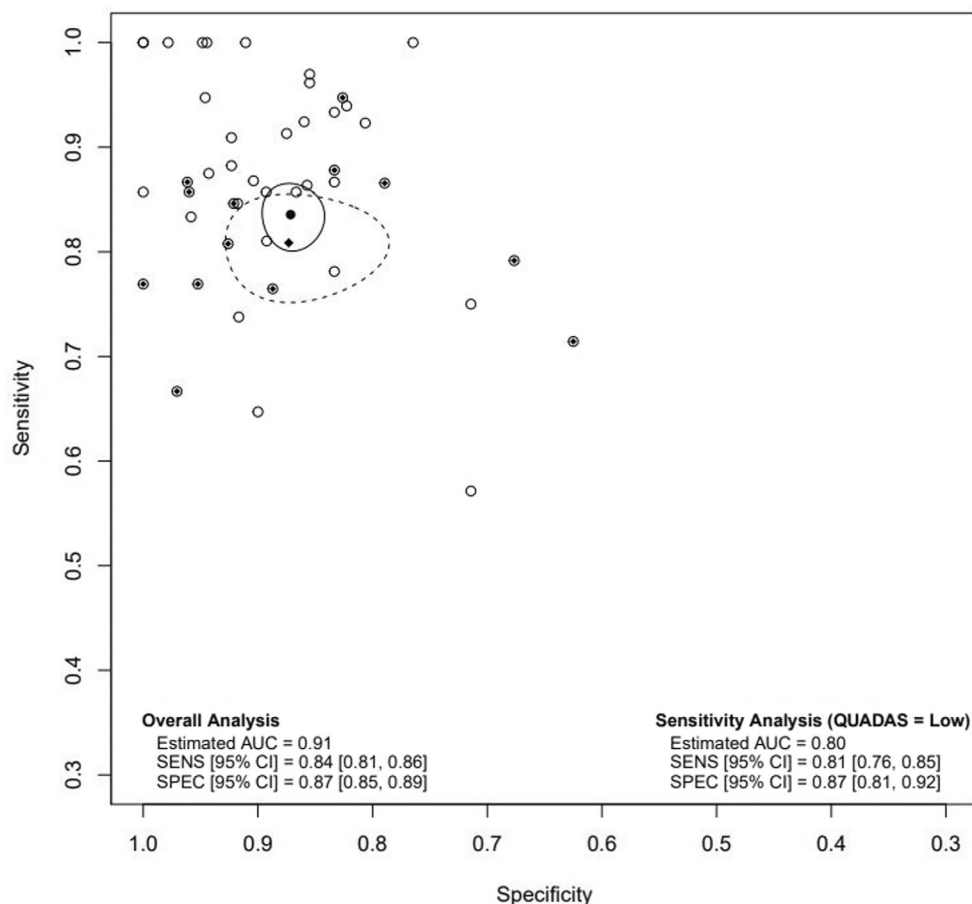


FIGURE 3 | Summary receiver operating characteristic (SROC) plot of diagnostic accuracy. Open circles represent sensitivity and specificity estimates for individual studies. Black diamonds indicate studies at low risk of bias (QUADAS-2 = Low). The solid black circle denotes the overall meta-analytic estimate with its 95% confidence region (solid line). The black diamond indicates the meta-analytic estimate restricted to low-risk-of-bias studies, with the corresponding 95% confidence region shown by the dashed line. Overall analysis: AUC = 0.91; SE = 0.84 (95% CI, 0.81–0.86); SP = 0.87 (95% CI, 0.85–0.89). Sensitivity analysis (QUADAS-2 = Low): AUC = 0.80; SE = 0.81 (95% CI, 0.76–0.85); SP = 0.87 (95% CI, 0.81–0.92).

A single-group meta-analysis using the inverse variance method reported an overall mean ADC of $1.08 \times 10^{-3} \text{ mm}^2/\text{s}$ for malignant lesions [95% CI 0.96–1.20; random-effects model] (Figure 5a) with significant heterogeneity ($I^2 = 98.2\%$ [97.9%; 98.4%]). Pooled analysis of the ADC values in benign lesions revealed an overall mean ADC value of $1.73 \times 10^{-3} \text{ mm}^2/\text{s}$ [95% CI 1.6116–1.8462; random-effects model] with significant heterogeneity ($I^2 = 97.7\%$ [97.3%; 98.1%]) (Figure 5b).

The overall mean ADC value of malignant lesions at $b \geq 1000$ was $0.94 \times 10^{-3} \text{ mm}^2/\text{s}$ (95% CI 0.8363–1.0220; REM) with heterogeneity expressed as $I^2 = 97.2\%$ (CI 96.5%; 97.8%) (Figure 6a), while the mean ADC for benign lesions was $1.63 \times 10^{-3} \text{ mm}^2/\text{s}$ (95% CI 1.52–1.74; REM) with heterogeneity expressed as $I^2 = 97.0\%$ [CI 96.3%; 97.6%] (Figure 6b).

4 | Discussion

DW-MRI is increasingly recognized as an additional tool in the characterization of thyroid nodules. Our systematic review and meta-analysis demonstrated optimal diagnostic performance with a cumulative AUC of 0.91. When the analysis was restricted to studies judged at low risk of bias according to QUADAS-2, the

pooled estimates confirmed excellent diagnostic performance, with an AUC of 0.80, only slightly lower than the overall value of 0.91.

In our analysis, DWI-MRI demonstrated a SP of 0.87 and a SE of 0.84, which is comparable to the pooled diagnostic performance of conventional ultrasound combined with elastography (SE 0.88, SP 0.86, AUC 0.92) [3] and ACR-TIRADS (SE ranging from 64% to 77% and SP from 82% to 90%) [2]. It should be kept in mind that different TIRADS and guideline-based classifications demonstrate consistently high SE (97%–99%) but limited SP (3%–31%), resulting in modest overall accuracy ranging from 32% to 61% [2]; unfortunately, in the available studies, the accuracy of DW-MRI in indeterminate nodules has not been specifically assessed.

The ADC value, obtained by DW-MRI, could serve as a complementary biomarker to ultrasound in the assessment of indeterminate thyroid nodules, potentially improving risk stratification. Evaluating the combined diagnostic performance of both modalities may lead to a more accurate approach, reducing uncertainty and better guiding clinical decision-making. DWI-MRI may be also helpful in patients undergoing neck MRI for other indications, where incidental

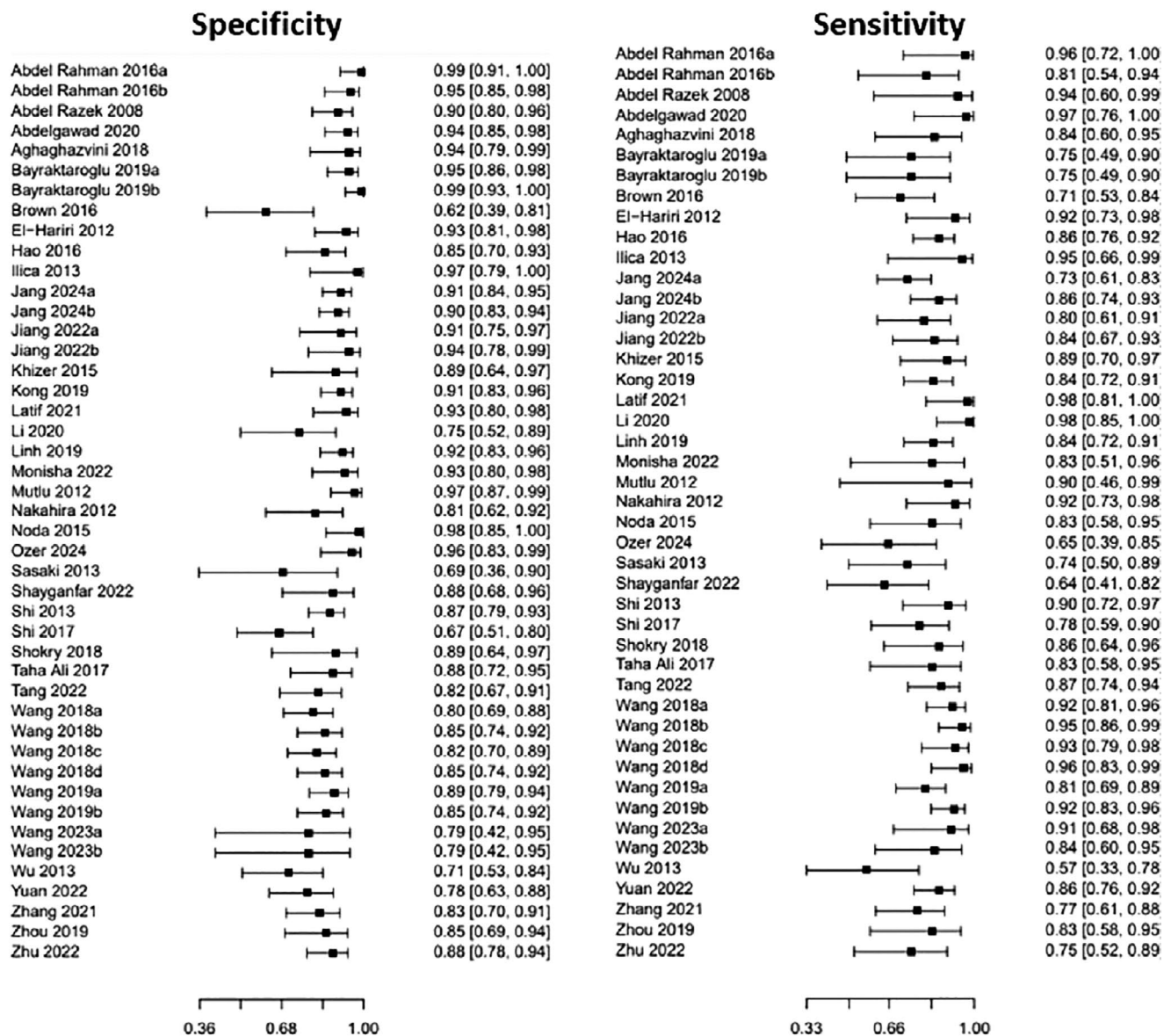


FIGURE 4 | Forest plots of the individual study sensitivity (a) and specificity (b) for thyroid nodule malignancy.

thyroid nodules are identified. Its ability to non-invasively characterize tissue could enhance diagnostic confidence and help refine surgical selection, potentially reducing unnecessary procedures. Improved characterization of thyroid nodules would have an impact on non-diagnostic FNAC rates—currently 34% of all procedures, on indeterminate Bethesda III–IV categories and consequent diagnostic lobectomies, reducing risks for patients [51, 52].

Our meta-analysis, which includes 37 studies and 2403 thyroid nodules, confirms the diagnostic value of DWI-MRI reported in previous research. Compared to an earlier meta-analysis based on 15 studies and 765 nodules (SE = 0.90, SP = 0.95) [8], and a more recent one involving 24 studies and 1714 lesions (pooled SE = 90.07%, SP = 87.97%) [9], our findings showed slightly lower values (SE = 0.84, SP = 0.86), yet still excellent overall performance. These results underscore the robust diagnostic accuracy of DWI-MRI and support its potential role

as an adjunct imaging modality in the characterization of thyroid nodules.

Technical acquisition parameters have the greatest impact on DW-MRI readability, as diagnostic efficacy is strongly influenced by image quality and the presence of artifacts [53]. The univariable meta-regression highlighted that some technical and methodological factors influenced diagnostic performance. Field strength appeared to affect SP but not SE. The number of averages and the acquisition matrix size were mainly associated with SE, while the cut-off definition of nodules and the method of ROI delineation showed a significant impact on SP. These findings underline how both acquisition parameters and methodological choices may contribute differently to test accuracy.

Despite the theoretical advantage of increased signal-to-noise ratio at 3T, several technical drawbacks inherent to higher field strength may explain the increased FPR observed in clinical

TABLE 2 | Results of subgroup analysis of MRI parameters.

Covariate	Number of studies (nodules)		Level	Estimate	Confidence interval lower	Confidence interval upper	p
Highest <i>b</i> value (mm ² /s)	36 (2403 nodules)	SE	< 1000	0.83	0.79	0.86	0.47
			≥ 1000	0.85	0.80	0.88	
		SP	< 1000	0.86	0.82	0.89	0.16
			≥ 1000	0.89	0.86	0.92	
Number of <i>b</i> values	36 (2403 nodules)	SE	2	0.83	0.79	0.86	0.29
			> 2	0.86	0.81	0.89	
		SP	2	0.85	0.82	0.88	0.13
			> 2	0.89	0.85	0.91	
Field strength (T)	36 (2403 nodules)	SE	1.5	0.83	0.78	0.88	0.74
			3.0	0.84	0.81	0.87	
		SP	1.5	0.91	0.88	0.93	0.00
			3.0	0.85	0.82	0.87	
Slice thickness (mm)	30 (1907 nodules)	SE	< 4	0.86	0.81	0.89	0.22
			≥ 4	0.82	0.78	0.86	
		SP	< 4	0.87	0.83	0.90	0.55
			≥ 4	0.85	0.82	0.88	
Number of averages	20 (1312 nodules)	SE	< 5	0.87	0.83	0.91	0.04
			≥ 5	0.79	0.71	0.85	
		SP	< 5	0.87	0.82	0.90	0.97
			≥ 5	0.87	0.80	0.91	
Slice gap (mm)	27 (1628 nodules)	SE	< 1	0.84	0.80	0.88	0.97
			≥ 1	0.85	0.78	0.89	
		SP	< 1	0.86	0.83	0.89	0.13
			≥ 1	0.89	0.86	0.92	
Time of echo (ms)	28 (1558 nodules)	SE	< 75	0.83	0.79	0.86	0.37
			≥ 75	0.80	0.74	0.85	
		SP	< 75	0.87	0.82	0.91	0.86
			≥ 75	0.87	0.81	0.91	
Time of repetition (ms)	30 (1823 nodules)	SE	< 4000	0.83	0.77	0.87	0.62
			≥ 4000	0.84	0.80	0.87	
		SP	< 4000	0.83	0.78	0.87	0.06
			≥ 4000	0.88	0.85	0.90	
FOV (mm)	25 (1675 nodules)	SE	< 200×200	0.85	0.82	0.88	0.11
			≥ 200×200	0.80	0.75	0.85	
		SP	< 200×200	0.86	0.82	0.89	0.92
			≥ 200×200	0.86	0.82	0.90	

(Continues)

TABLE 2 | (Continued)

Covariate	Number of studies (nodules)		Level	Estimate	Confidence interval lower	Confidence interval upper	p
Acquisition matrix	24 (1602 nodules)	SE	< 130×130	0.80	0.75	0.84	0.00
			≥ 130×130	0.89	0.85	0.92	
		SP	< 130×130	0.87	0.83	0.90	0.82
			≥ 130×130	0.87	0.83	0.91	

Note: For each covariate of MRI parameters and pre-specified cut-offs, the values of sensitivity (SE) and specificity (SP) and *p*-value are displayed covariate^a.

^aSix studies [22, 25, 31, 46, 47], and Jang 2024 reported two thresholds (according to different *b*-values), and one study [45] reported four thresholds (according to different *b*-values); all thresholds were considered in the analysis.

TABLE 3 | Results of subgroup analysis for thyroid nodule size, segmentation and ROI type.

Covariate	Number of studies		Level	Estimate	Confidence interval lower	Confidence interval upper	p
Thyroid nodule size (mm)	14 (1009 nodules)	SE	< 10	0.86	0.80	0.90	0.57
			≥ 10	0.84	0.77	0.88	
		SP	< 10	0.84	0.76	0.89	0.04
			≥ 10	0.91	0.87	0.94	
Thyroid nodule segmentation	30 (2031 nodules)	SE	Not whole	0.84	0.78	0.89	0.84
			Whole	0.85	0.81	0.88	
		SP	Not whole	0.88	0.84	0.91	0.19
			Whole	0.85	0.82	0.88	
ROI type	31 (1.985 nodules)	SE	Freehand	0.85	0.81	0.89	0.70
			Circular	0.84	0.79	0.88	
		SP	Freehand	0.84	0.80	0.87	0.03
			Circular	0.88	0.85	0.91	

Note: For each methodological covariate and pre-specified cut-offs, the values of sensitivity (SE) and specificity (SP) and *p*-value are displayed covariate.

DWI, especially in anatomically complex regions like the neck. As Graves et al. [54] outlined, 3T systems are significantly more sensitive to magnetic susceptibility differences between tissues such as air, bone, and soft tissue. This leads to increased spatial misregistration, geometric distortions, and signal pile-up, particularly evident in single-shot echo planar imaging based DW-MRI sequences. These artifacts can mimic or exaggerate diffusion restriction, potentially being misinterpreted as malignant lesions, thereby increasing false positive cases. Additionally, 3T systems suffer from greater field inhomogeneities, which result in non-uniform signal intensities across the FOV. This variability reduces the reproducibility of ADC measurements and compromises quantitative accuracy.

Interestingly, studies using fewer than five signal averages reported higher SE. This apparent paradox can be explained by the reduced acquisition time, which minimizes swallowing and motion artifacts in the head and neck region, thereby preserving lesion conspicuity despite the lower theoretical Signal-to-Noise Ratio (SNR). Conversely, higher averages, while improving SNR, may increase motion-related image degradation and reduce diagnostic SE.

In DW-MRI, the acquisition matrix size directly affects the in-plane spatial resolution. A higher matrix (≥ 130×130) results in smaller voxels and finer anatomical detail, allowing better delineation of lesion borders and improved detection of small or subtle abnormalities. This leads to an increase in SE, as more true positive lesions are correctly identified. However, this improvement in resolution does not necessarily reduce the FPR.

Using a cut-off of ≥ 10 mm for the thyroid nodules resulted in a lower FPR and therefore improved SP, without a significant drop in SE. This supports the idea that applying a stricter threshold for lesion size (excluding subcentimetric nodules) helps reduce overdiagnosis and may be useful in optimizing the diagnostic accuracy of DW-MRI. This supports the notion that DW-MRI should be selectively applied to nodules that are indeterminate or suspicious, particularly when their size exceeds 10 mm. This aligns with current clinical guidelines, which recommend further characterization of nodules > 10 mm that remain equivocal after ultrasound evaluation [4].

The higher SP observed with circular ROIs may be explained by their greater reproducibility and reduced inclusion of peripheral or heterogeneous areas that can bias ADC measurements, compared with freehand delineation. By standardizing the placement within the solid portion of the nodule, circular ROIs provide more reliable values and reduce false positives. Importantly, all included studies reported excluding cystic, necrotic, and hemorrhagic components from ROI placement, further minimizing sources of variability.

In multivariable meta-regression, the number of *b*-values and the highest *b*-value reached statistical significance, while the unfavorable effect of 3T on SP remained significant. The use of multiple *b*-values likely improves SE because it allows a more

accurate fitting of the diffusion signal decay, reducing misclassification of malignant lesions [55]. Conversely, higher maximum *b*-values enhance SP by minimizing perfusion effects and T2 shine-through, thereby improving discrimination of benign lesions and reducing false positives [56].

The fact that these associations were not significant in univariable analysis but became significant in the multivariable model may reflect the confounding effect of correlated acquisition parameters. By adjusting for other covariates, the multivariable analysis isolates the independent contribution of each factor, revealing their true influence on diagnostic performance [57].

As an added value to the existing body of literature, we performed a single-group meta-analysis to estimate the mean ADC values of malignant thyroid lesions in comparison with benign nodules, providing a cut-off suggestive of malignancy. Overall, malignant lesions showed lower ADC values than benign ones, consistent with the expected higher tissue cellularity and restricted diffusion. These findings are in line with those reported in a previous meta-analysis, which similarly demonstrated lower ADC values for malignant nodules compared with benign lesions. Moreover, in line with the results of our meta-regression, ADC values were influenced by the choice of *b*-values, with higher thresholds (≥ 1000) leading to more specific measurements. Also, the *b*-value at which an ADC sampling is performed plays a role and has often been heterogeneous across studies [58].

4.1 | Limitations

This meta-analysis has some limitations. First, not all technical parameters known to influence DW-MRI acquisition were consistently reported across the included studies. As a result, we

TABLE 4 | Results of meta-regression analysis.

Covariate (reference)	Association with SE (p)	Association with SP via FPR (p)
Highest <i>b</i> value (≥ 1000 vs < 1000)	$\uparrow$ SE ($p = 0.031$)	NS ($p = 0.759$)
Number of <i>b</i> -values (> 2 vs ≤ 2)	NS ($p = 0.187$)	$\uparrow$ SP ($p = 0.02$)
Field strength (3.0T vs 1.5T)	NS ($p = 0.190$)	$\downarrow$ SP ($p = 0.00$)
Slice thickness (≤ 3 mm vs > 3 mm)	NS ($p = 0.727$)	NS ($p = 0.362$)
QUADAS-2 (low vs high/unclear risk)	NS ($p = 0.102$)	NS ($p = 0.768$)

Note: For each methodological covariate, the values of association with sensitivity (SE), specificity (SP) and *p*-value are displayed covariate. Abbreviation: NS, not significant.

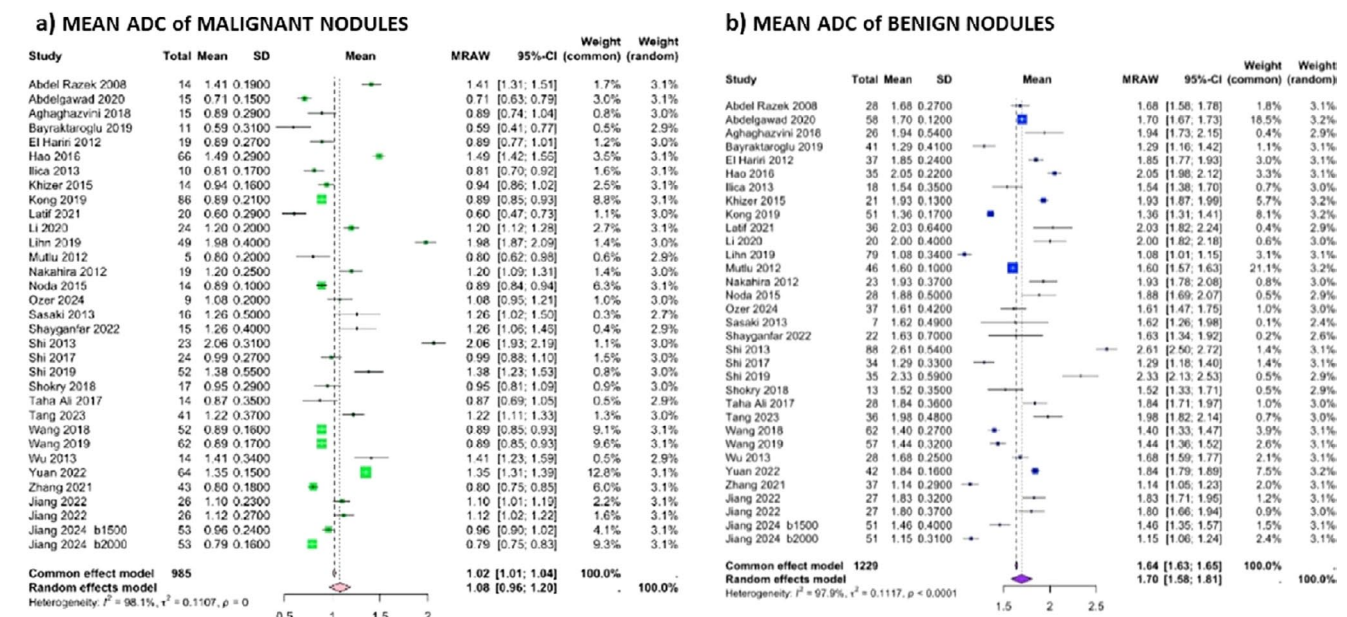


FIGURE 5 | Forest plots of apparent diffusion coefficient values of diffusion-weighted magnetic resonance imaging in differentiation of thyroid nodule (a) Mean ADC values of malignant nodules; (b) Mean ADC values of benign nodules.

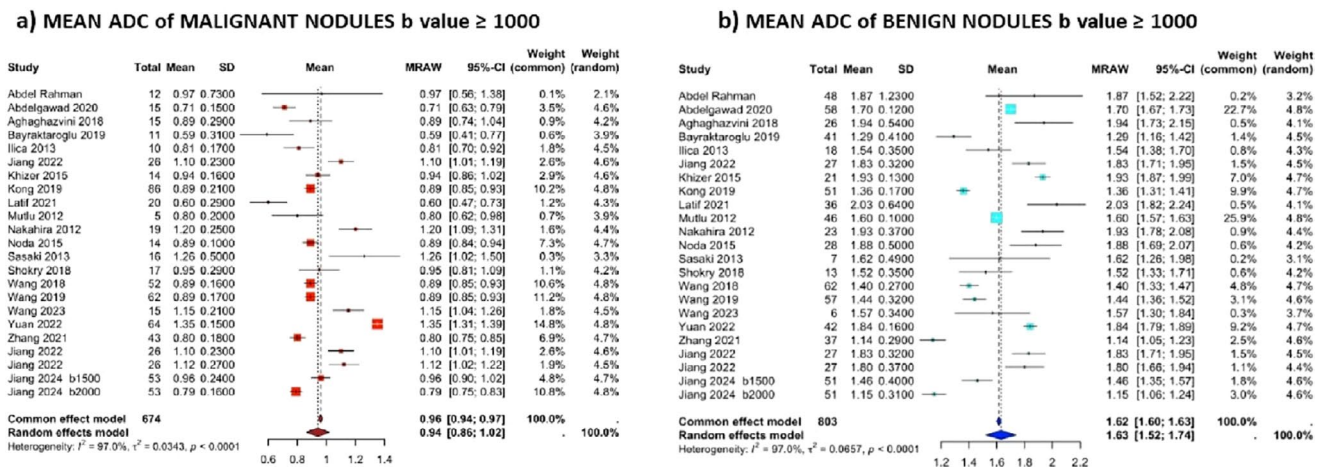


FIGURE 6 | Forest plots of apparent diffusion coefficient values of diffusion-weighted magnetic resonance imaging in differentiation of thyroid nodule of subgroup analysis performed on selected studies with b max value ≥ 1000 (a) Mean ADC values of malignant nodules; (b) Mean ADC values of benign nodules.

were unable to perform a comprehensive multivariable meta-regression. Specifically, we could not investigate the potential impact of surface coil usage, sequence types, scanner models, or manufacturer-related differences, all of which may affect image quality and diagnostic accuracy. Furthermore, since some variables were not available in all datasets, their effects could only be explored descriptively, and any observed trends should be interpreted with caution. Therefore, the findings of this meta-analysis should not be generalized to parameters that were not consistently included, and future prospective studies with standardized imaging protocols are warranted to validate these results. The majority of the studies included had a low or unclear risk of bias, represented by the method of nodule characterization (a clear distinction between nodules characterized by cytological analysis and those confirmed by histological examination is often lacking), variability in DW-MRI protocols and methods for ADC sampling. It is worth noting the significant heterogeneity observed in ADC analysis outcomes, which may reflect differences in MRI protocols, ROI placement, or patient populations.

The reproducibility metrics (e.g., inter-reader agreement or Bland-Altman analysis) could not be extracted from the included articles; therefore, we only considered differences related to ROI placement methods (free-hand vs. circular).

A potential limitation of our meta-analysis is the lack of patient-level data, which prevented us from accounting for clustering effects due to multiple nodules arising from the same patient. As most of the included studies reported diagnostic performance on a per-nodule basis, we could not adjust for within-patient correlation, and this may have led to a slight overestimation of SE. Additionally, in our dataset, most studies pooled malignant and benign nodules without per-subtype stratification, so a histology-based meta-analysis was not feasible. However, Abdelgawad et al. [18] found significant ADC differences between benign and malignant nodules but no consistent differences among histologic subtypes.

4.2 | Conclusion

This meta-analysis confirms that DW-MRI and ADC values provide substantial diagnostic accuracy in differentiating benign from malignant thyroid nodules. Subgroup and univariable meta-regression analyses highlighted that some technical and methodological factors influenced diagnostic performance. Field strength mainly affected SP, while the number of averages and the acquisition matrix size were associated with SE. In addition, the nodule cut-off definition and the method of ROI delineation significantly impacted SP. Multivariable meta-regression further demonstrated that the use of multiple b-values improved SE, higher maximum b-values enhanced SP, and imaging at 3T was associated with reduced SP. Taken together, these findings emphasize that both acquisition parameters and methodological choices contribute differently to diagnostic accuracy, underlining the need for standardized protocols to ensure reproducible and clinically reliable use of DW-MRI in thyroid nodule characterization.

Acknowledgments

The authors would like to thank Giorgia Carnicelli and Andrea Costantino for their valuable support to Fabrizio Natali in the statistical analysis, which significantly contributed to the quality of this work. Caterina Giannitto and Angela Ammirabile share first authorship.

References

1. L. Giovannella, A. Campenni, M. Tuncel, and O. P. Petranović, "Integrated Diagnostics of Thyroid Nodules," *Cancers* 16 (2024): 2.
2. D. H. Kim, S. W. Kim, M. A. Basurrah, J. Lee, and S. H. Hwang, "Diagnostic Performance of Six Ultrasound Risk Stratification Systems for Thyroid Nodules: A Systematic Review and Network Meta-Analysis," *American Journal of Roentgenology* 220, no. 6 (2023): 791–803.
3. M. Shi, D. Nong, M. Xin, and L. Lin, "Accuracy of Ultrasound Diagnosis of Benign and Malignant Thyroid Nodules: A Systematic Review and Meta-Analysis," *International Journal of Clinical Practice* 2022 (2022): 5056082.

4. C. Durante, L. Hegedüs, A. Czarniecka, et al., "European Thyroid Association Clinical Practice Guidelines for Thyroid Nodule Management," *European Thyroid Journal* 2023 (2023): 12.
5. F. N. Tessler, W. D. Middleton, E. G. Grant, et al., "ACR Thyroid Imaging, Reporting and Data System (TI-RADS): White Paper of the ACR TI-RADS Committee," *Journal of the American College of Radiology* 14, no. 5 (2017): 587–595.
6. E. Abelleira and F. Jerkovich, "Dynamic Risk Assessment in Patients With Differentiated Thyroid Cancer," *Reviews in Endocrine and Metabolic Disorders* 25, no. 1 (2024): 79–93.
7. G. Grani, M. Sponziello, V. Pecce, V. Ramundo, and C. Durante, "Contemporary Thyroid Nodule Evaluation and Management," *Journal of Clinical Endocrinology and Metabolism* 105, no. 9 (2020): 2869–2883.
8. L. Chen, J. Xu, J. Bao, et al., "Diffusion-Weighted MRI in Differentiating Malignant From Benign Thyroid Nodules: A Meta-Analysis," *BMJ Open* 6, no. 1 (2016): e008413.
9. H. J. Meyer, A. Wienke, and A. Surov, "Discrimination Between Malignant and Benign Thyroid Tumors by Diffusion-Weighted Imaging – A Systematic Review and Meta Analysis," *Magnetic Resonance Imaging* 84 (2021): 41–57.
10. H. Mutlu, A. K. Sivrioglu, G. Sonmez, et al., "Role of Apparent Diffusion Coefficient Values and Diffusion-Weighted Magnetic Resonance Imaging in Differentiation Between Benign and Malignant Thyroid Nodules," *Clinical Imaging* 36, no. 1 (2012): 1–7.
11. M. Nakahira, N. Saito, S. Murata, et al., "Quantitative Diffusion-Weighted Magnetic Resonance Imaging as a Powerful Adjunct to Fine Needle Aspiration Cytology for Assessment of Thyroid Nodules," *American Journal of Otolaryngology* 33, no. 4 (2012): 408–416.
12. M. D. F. McInnes, D. Moher, B. D. Thombs, et al., "Preferred Reporting Items for a Systematic Review and Meta-Analysis of Diagnostic Test Accuracy Studies: The PRISMA-DTA Statement," *Journal of the American Medical Association* 319, no. 4 (2018): 388–396.
13. Y. Zhou and N. Dendukuri, "Statistics for Quantifying Heterogeneity in Univariate and Bivariate Meta-Analyses of Binary Data: The Case of Meta-Analyses of Diagnostic Accuracy," *Statistics in Medicine* 33, no. 16 (2014): 2701–2717.
14. P. Doebler, H. Holling, and D. Böhning, "A Mixed Model Approach to Meta-Analysis of Diagnostic Studies With Binary Test Outcome," *Psychological Methods* 17, no. 3 (2012): 418–436.
15. H. Holling, W. Böhning, and D. Böhning, "Meta-Analysis of Diagnostic Studies Based Upon SROC-Curves: A Mixed Model Approach Using the Lehmann Family," *Statistical Modelling* 12, no. 4 (2012): 347–375.
16. H. M. Abdel-Rahman, M. H. Abowarda, and S. M. Abdel-Aal, "Diffusion-Weighted MRI and Apparent Diffusion Coefficient in Differentiation of Benign From Malignant Solitary Thyroid Nodule," *Egyptian Journal of Radiology and Nuclear Medicine* 47, no. 4 (2016): 1385–1390.
17. A. A. Abdel Razek, A. G. Sadek, O. R. Kombar, T. E. Elmahdy, and N. Nada, "Role of Apparent Diffusion Coefficient Values in Differentiation Between Malignant and Benign Solitary Thyroid Nodules," *American Journal of Neuroradiology* 29, no. 3 (2008): 563–568.
18. E. A. Abdelgawad, E. A. AbdelGawad, O. AbuElCebaa, and A. M. Atiya, "Can Quantitative Diffusion-Weighted MR Imaging Differentiate Between Different Subtypes of Benign and Malignant Solitary Thyroid Nodules?," *Egyptian Journal of Radiology and Nuclear Medicine* 51, no. 1 (2020): 6.
19. L. Aghaghazvini, H. Sharifian, N. Yazdani, et al., "Differentiation Between Benign and Malignant Thyroid Nodules Using Diffusion-Weighted Imaging, a 3-T MRI Study," *Indian Journal of Radiology and Imaging* 28, no. 4 (2018): 460–464.
20. S. Bayraktaroglu, P. K. Öztürk, N. Ceylan, Ö. Makay, G. İçöz, and Y. Ertan, "3 Tesla Apparent Diffusion Coefficient (ADC) Values of Thyroid Nodules: Prediction of Benignancy and Malignancy," *Iranian Journal of Radiology* 16, no. 3 (2019): e63974.
21. A. M. Brown, S. Nagala, M. A. McLean, et al., "Multi-Institutional Validation of a Novel Textural Analysis Tool for Preoperative Stratification of Suspected Thyroid Tumors on Diffusion-Weighted MRI," *Magnetic Resonance in Medicine* 75, no. 4 (2016): 1708–1716.
22. M. A. El-Hariri, G. K. Gouhar, N. S. Said, and M. M. Riad, "Role of Diffusion-Weighted Imaging With ADC Mapping and In Vivo 1H-MR Spectroscopy in Thyroid Nodules," *Egyptian Journal of Radiology and Nuclear Medicine* 43, no. 2 (2012): 183–192.
23. Y. Hao, C. Pan, W. Chen, T. Li, W. Zhu, and J. Qi, "Differentiation Between Malignant and Benign Thyroid Nodules and Stratification of Papillary Thyroid Cancer With Aggressive Histological Features: Whole-Lesion Diffusion-Weighted Imaging Histogram Analysis," *Journal of Magnetic Resonance Imaging* 44, no. 6 (2016): 1546–1555.
24. A. T. Ilica, H. Artas, A. Ayan, et al., "Initial Experience of 3 Tesla Apparent Diffusion Coefficient Values in Differentiating Benign and Malignant Thyroid Nodules," *Journal of Magnetic Resonance Imaging* 37, no. 5 (2013): 1077–1082.
25. L. Jiang, J. Zhang, J. Chen, et al., "rFOV-DWI and SMS-RESOLVE-DWI in Patients With Thyroid Nodules: Comparison of Image Quality and Apparent Diffusion Coefficient Measurements," *Magnetic Resonance Imaging* 91 (2022): 62–68.
26. L. Jiang, J. Chen, Y. Tan, et al., "Comparative Analysis of the Image Quality and Diagnostic Performance of the Zooming Technique With Diffusion-Weighted Imaging Using Different b-Values for Thyroid Papillary Carcinomas and Benign Nodules," *Frontiers in Oncology* 14 (2024): 1241776.
27. A. T. Khizer, S. Raza, and A. U. Slehra, "Diffusion-Weighted MR Imaging and ADC Mapping in Differentiating Benign From Malignant Thyroid Nodules," *Journal of the College of Physicians and Surgeons–Pakistan* 25, no. 11 (2015): 785–788.
28. W. Kong, X. Yue, J. Ren, and X. Tao, "A Comparative Analysis of Diffusion-Weighted Imaging and Ultrasound in Thyroid Nodules," *BMC Medical Imaging* 19, no. 1 (2019): 92.
29. M. A. Latif, M. M. El Rakhawy, and M. F. Saleh, "Diagnostic Accuracy of B-Mode Ultrasound, Ultrasound Elastography and Diffusion Weighted MRI in Differentiation of Thyroid Nodules (Prospective Study)," *Egyptian Journal of Radiology and Nuclear Medicine* 52, no. 1 (2021): 256.
30. G. Li, G. Jiang, Y. Mei, et al., "Applying Amide Proton Transfer-Weighted Imaging (APTWI) to Distinguish Papillary Thyroid Carcinomas and Predominantly Solid Adenomatous Nodules: Comparison With Diffusion-Weighted Imaging," *Frontiers in Oncology* 10 (2020): 918.
31. L. T. Linh, N. N. Cuong, T. V. Hung, et al., "Value of Diffusion Weighted MRI With Quantitative ADC Map in Diagnosis of Malignant Thyroid Disease," *Diagnostics* 9, no. 4 (2019): 129.
32. V. Monisha, N. Rache Gowda, and S. Anil Kumar, "The Role of Diffusion-Weighted Magnetic Resonance Imaging in Differentiating Benign From Malignant Thyroid Nodules: A Descriptive Observational Study," *Cureus* 14, no. 10 (2022): e30493.
33. Y. Noda, M. Kanematsu, S. Goshima, et al., "MRI of the Thyroid for Differential Diagnosis of Benign Thyroid Nodules and Papillary Carcinomas," *American Journal of Roentgenology* 204, no. 3 (2015): W332–W335.
34. B. M. Özer, Y. Pabuşçu, S. Tarhan, et al., "Effectiveness of Diffusion-Weighted Magnetic Resonance Imaging (DW-MRI) in the Differentiation of Thyroid Nodules," *Thyroid Research* 17, no. 1 (2024): 24.
35. M. Sasaki, M. Sumi, K. Kaneko, K. Ishimaru, H. Takahashi, and T. Nakamura, "Multiparametric MR Imaging for Differentiating Between Benign and Malignant Thyroid Nodules: Initial Experience in 23 Patients," *Journal of Magnetic Resonance Imaging* 38, no. 1 (2013): 64–71.

36. A. Shayganfar, N. Azin, P. Hashemi, A. M. Ghanei, and S. Hajjah-madi, "Diagnostic Accuracy of Multiple MRI Parameters in Dealing With Incidental Thyroid Nodules," *SN Comprehensive Clinical Medicine* 4, no. 1 (2022): 228.
37. H. F. Shi, Q. Feng, J. W. Qiang, R. K. Li, L. Wang, and J. P. Yu, "Utility of Diffusion-Weighted Imaging in Differentiating Malignant From Benign Thyroid Nodules With Magnetic Resonance Imaging and Pathologic Correlation," *Journal of Computer Assisted Tomography* 37, no. 4 (2013): 505–510.
38. R. Y. Shi, Q. Y. Yao, Q. Y. Zhou, et al., "Preliminary Study of Diffusion Kurtosis Imaging in Thyroid Nodules and Its Histopathologic Correlation," *European Radiology* 27, no. 11 (2017): 4710–4720.
39. H. Shi, Z. Yuan, C. Yang, et al., "Role of Multi-Modality Functional Imaging in Differentiation Between Benign and Malignant Thyroid (18) F-Fluorodeoxyglucose Incidentaloma," *Clinical & Translational Oncology* 21, no. 11 (2019): 1561–1567.
40. A. M. Shokry, T. A. Hassan, A. A. Baz, A. S. O. Ahmed, and M. H. Zedan, "Role of Diffusion Weighted Magnetic Resonance Imaging in Differentiation of Benign and Malignant Thyroid Nodules," *Egyptian Journal of Radiology and Nuclear Medicine* 49, no. 4 (2018): 1014–1021.
41. T. F. Taha Ali, "Solitary Thyroid Nodule: Diagnostic Yield of Combined Diffusion Weighted Imaging and Magnetic Resonance Spectroscopy," *Egyptian Journal of Radiology and Nuclear Medicine* 48, no. 3 (2017): 593–601.
42. Q. Tang, X. Liu, Q. Jiang, et al., "Unenhanced Magnetic Resonance Imaging of Papillary Thyroid Carcinoma With Emphasis on Diffusion Kurtosis Imaging," *Quantitative Imaging in Medicine and Surgery* 13, no. 4 (2023): 2697–2707.
43. Q. Wang, Y. Guo, J. Zhang, et al., "Utility of High b-Value (2000 Sec/mm²) DWI With RESOLVE in Differentiating Papillary Thyroid Carcinomas and Papillary Thyroid Microcarcinomas From Benign Thyroid Nodules," *PLoS One* 13, no. 7 (2018): e0200270.
44. Q. Wang, Y. Guo, J. Zhang, et al., "Diagnostic Value of High b-Value (2000 s/mm²) DWI for Thyroid Micronodules," *Medicine* 98, no. 10 (2019): e14298.
45. X. Wang, P. Wang, H. Zhang, X. Wang, J. Shi, and S. Hu, "Multiplexed Sensitivity-Encoding Versus Single-Shot Echo-Planar Imaging: A Comparative Study for Diffusion-Weighted Imaging of the Thyroid Lesions," *Japanese Journal of Radiology* 42, no. 3 (2024): 268–275.
46. Y. Wu, X. Yue, W. Shen, et al., "Diagnostic Value of Diffusion-Weighted MR Imaging in Thyroid Disease: Application in Differentiating Benign From Malignant Disease," *BMC Medical Imaging* 13 (2013): 23.
47. L. Yuan, P. Zhao, X. Lin, T. Yu, R. Diao, and G. Ning, "T1 Mapping and Reduced Field-Of-View DWI at 3.0T MRI for Differentiation of Thyroid Papillary Carcinoma From Nodular Goiter," *Clinical Physiology and Functional Imaging* 43, no. 3 (2023): 137–145.
48. J. Zhang, W. X. Fu, W. P. Li, et al., "Diagnostic Value of Thyroid Micronodules With High b-Value Diffusion Weighted Imaging: Comparative Study With High-Resolution Ultrasound," *European Journal of Radiology* 143 (2021): 109912.
49. X. Zhou, C. Ma, Z. Wang, et al., "Effect of Region of Interest on ADC and Interobserver Variability in Thyroid Nodules," *BMC Medical Imaging* 19, no. 1 (2019): 55.
50. X. Zhu, J. Wang, Y. C. Wang, et al., "Quantitative Differentiation of Malignant and Benign Thyroid Nodules With Multi-Parameter Diffusion-Weighted Imaging," *World Journal of Clinical Cases* 10, no. 24 (2022): 8587–8598.
51. E. D. Rossi, A. J. Adeniran, and W. C. Faquin, "Pitfalls in Thyroid Cytopathology," *Surgical Pathology Clinics* 12, no. 4 (2019): 865–881.
52. R. Stewart, Y. J. Leang, C. R. Bhatt, S. Grodski, J. Serpell, and J. C. Lee, "Quantifying the Differences in Surgical Management of Patients With Definitive and Indeterminate Thyroid Nodule Cytology," *European Journal of Surgical Oncology* 46, no. 2 (2020): 252–257.
53. C. D. Norris, S. E. Quick, J. G. Parker, and N. A. Koontz, "Diffusion MR Imaging in the Head and Neck: Principles and Applications," *Neuroimaging Clinics of North America* 30, no. 3 (2020): 261–282.
54. M. J. Graves, "3 T: The Good, the Bad and the Ugly," *British Journal of Radiology* 95, no. 1130 (2022): 20210708.
55. R. L. Magin, M. M. Karaman, M. G. Hall, W. Zhu, and X. J. Zhou, "Capturing Complexity of the Diffusion-Weighted MR Signal Decay," *Magnetic Resonance Imaging* 56 (2019): 110–118.
56. S. Chawla, S. Kim, S. Wang, and H. Poptani, "Diffusion-Weighted Imaging in Head and Neck Cancers," *Future Oncology* 5, no. 7 (2009): 959–975.
57. R. D. Riley, D. Jackson, G. Salanti, et al., "Multivariate and Network Meta-Analysis of Multiple Outcomes and Multiple Treatments: Rationale, Concepts, and Examples," *British Medical Journal* 358 (2017): j3932.
58. A. H. El Beltagi, A. H. Elsotouhy, A. M. Own, W. Abdelfattah, K. Nair, and S. Vattoth, "Functional Magnetic Resonance Imaging of Head and Neck Cancer: Performance and Potential," *Neuroradiology Journal* 32, no. 1 (2019): 36–52.

Supporting Information

Additional supporting information can be found online in the Supporting Information section. **Data S1:** Supporting Information. **Data S2:** Supporting Information.