Article

Additional Value of Pertechnetate Scintigraphy to American College of Radiology Thyroid Imaging Reporting and Data Systems and European Thyroid Imaging Reporting and Data Systems for Thyroid Nodule Classification in Euthyroid Patients

Lea Sollmann [1], Maria Eveslage [2], Moritz Fabian Danzer [2], Michael Schäfers [1,3,4], Barbara Heitplatz [5], Elke Conrad [1], Daniel Hescheler [1,6], Burkhard Riemann [1,3] and Benjamin Noto [1,3,7,*]

1. Department of Nuclear Medicine, University of Münster and University Hospital Münster, 48149 Münster, Germany
2. Institute of Biostatistics and Clinical Research, University of Münster, 48149 Münster, Germany
3. West German Cancer Centre (WTZ), University Hospital Münster, 48149 Münster, Germany
4. European Institute for Molecular Imaging, University of Münster, 48149 Münster, Germany
5. Institute of Pathology, University Hospital of Münster, 48149 Münster, Germany
6. Department of Radiology and Nuclear Medicine, University Hospital Magdeburg, 39120 Magdeburg, Germany
7. Clinic for Radiology, University of Münster and University Hospital Münster, 48149 Münster, Germany
* Correspondence: benjamin.noto@ukmuenster.de; Tel.: +49-251-83-47302

Simple Summary: Thyroid nodules are a frequent finding yet remain a diagnostic challenge. While ultrasound and Thyroid Imaging Reporting and Data Systems (TIRADS) are widely accepted, the role of thyroid scintigraphy in euthyroid patients is debated. The European Association of Nuclear Medicine advocates its use, whereas the American Thyroid Association (ATA) and European Thyroid Association (ETA) do not include scintigraphy in their diagnostic algorithms. However, no study has systematically quantified whether scintigraphy adds diagnostic value to TIRADS in a multimodal approach. Our study addresses this gap, demonstrating a moderate improvement in diagnostic performance when incorporating scintigraphy into both ACR-TIRADS and EU-TIRADS, providing more accurate risk assessment, particularly in TIRADS categories with intermediate malignancy risk. This combined approach may benefit complex cases, such as multinodular goiter or indeterminate fine-needle aspiration results. Further research to determine if pertechnetate scintigraphy benefits the evaluation of thyroid nodules in selected cases of euthyroid patients seems warranted.

Abstract: Background: Thyroid nodules are common yet remain a diagnostic challenge. While ultrasound and Thyroid Imaging Reporting and Data Systems (TIRADS) are accepted as standard, the use of thyroid scintigraphy in euthyroid patients is debated. The European Association of Nuclear Medicine advocates it, whereas the American Thyroid Association and European Thyroid Association do not. However, it has not been evaluated whether scintigraphy adds value to TIRADS in a multimodal approach. Our study addresses this gap by assessing the impact of integrated pertechnetate scintigraphy on TIRADS accuracy. **Methods:** The diagnostic performance of ACR-TIRADS, EU-TIRADS, pertechnetate scintigraphy, and multimodal models were retrospectively analyzed for 322 nodules (231 benign, 91 malignant) in 208 euthyroid patients with histopathology as a reference. Generalized estimating equations were used for statistical analysis. **Results:** On scintigraphy, 210 nodules were hypofunctional, 99 isofunctional, and 13 hyperfunctional. The AUC for thyroid scintigraphy, ACR-TIRADS, and EU-TIRADS were 0.6 (95% CI: 0.55–0.66), 0.83 (95% CI: 0.78–0.88), and 0.78 (95% CI: 0.72–0.83). Integrating scintigraphy with ACR-TIRADS and EU-TIRADS slightly increased diagnostic accuracy (AUC 0.86 vs. 0.83, $p = 0.039$ and AUC 0.80 vs. 0.78, $p = 0.008$) and adjusted the malignancy probability for intermediate risk TIRADS categories, with iso- or hyperfunctioning nodules in ACR-TIRADS-TR4 or EU-TIRADS-4 showing comparable malignancy probabilities as hypofunctioning nodules in TR3 or EU-TIRADS-3, respectively. **Conclusions:** Integrating thyroid scintigraphy with ACR- or EU-TIRADS moderately improves diagnostic performance, potentially

benefiting management, especially in complex cases like multinodular goiter or indeterminate FNA. Further research is warranted to validate these findings and explore their clinical implications.

Keywords: thyroid cancer; thyroid nodules; TIRADS; scintigraphy; radionuclide scanning; ultrasound

1. Introduction

Thyroid nodules are a very common, often incidental clinical finding, detectable on high-resolution ultrasound in 25% to 75% of the general population [1,2]. While the majority of these nodules prove to be benign, up to 15% exhibit malignant characteristics, with thyroid cancer representing the most common endocrine malignancy [3–5].

Ultrasound is the primary imaging modality for the work-up of thyroid nodules in euthyroid patients. In recent years, several Thyroid Imaging Reporting and Data System (TIRADS), based on B-mode ultrasound characteristics, have been developed to this end [3,6–8]. Another imaging modality with a long history of use in thyroid evaluation is radionuclide scanning, mainly with [^{99m}Tc]Tc-pertechnetate. In contrast to ultrasound, thyroid scintigraphy primarily gives insight into the functional status. However, the use of thyroid scintigraphy for evaluating thyroid nodules in euthyroid patients is controversial.

The current German guidelines recommend the scintigraphic evaluation of every thyroid nodule larger than 1 cm [9]. The European Association of Nuclear Medicine recommends using thyroid scintigraphy to preselect only "cold" nodules for fine needle aspiration cytology (FNA), claiming that this strategy can greatly reduce the number of thyroid surgical procedures for benign entities, especially in populations with a high prevalence of iodine deficiency [10,11]. This implies that malignancy can be excluded with a high degree of certainty in isofunctional or hyperfunctional nodules. More so, hyperfunctioning nodules are considered to be virtually exclusively benign [12]. The American Thyroid Association (ATA) mentions the use of thyroid scintigraphy as a possibility for FNA-preselection in multinodular goiter but does not provide further details [3]. The European Thyroid Association mentions that thyroid scintigraphy may show hyperfunctioning nodules even in patients with normal TSH values and claims that hyperfunctioning nodules are rarely malignant. However, thyroid scintigraphy is not included in the proposed algorithm for the workup of thyroid nodules in euthyroid patients [13]. Finally, the British NICE-Guidelines recommend against the use of radionuclide scanning, basing their decision on insufficient and weak evidence concerning the use of thyroid scintigraphy in the evaluation of thyroid nodules in euthyroid patients [14].

Additionally, even the guidelines and societies that favor the use of thyroid scintigraphy for nodule evaluation do not elaborate on how to interpret the results of radionuclide scanning in conjunction with ultrasound in a multimodal approach.

Therefore, the aim of this study was to investigate whether thyroid scintigraphy adds diagnostic value to ACR- and EU-TIRADS in a multimodal approach.

2. Materials and Methods

2.1. Patients and Data Recruitment

Two separately recruited patient groups were retrospectively evaluated for this analysis. Inclusion criteria were the availability of preoperative thyroid ultrasound, pertechnetate scintigraphy, laboratory results, and pathology reports following thyroidectomy. Furthermore, only nodules ≥ 1 cm on ultrasound were included. Patients with Grave's disease or hyperthyroidism from other causes were excluded. The first group (Group 1) consisted of all patients presenting for thyroid cancer follow-up care at a tertiary referral hospital (*n* = 100) in 2022 and 2023, with nodal work-up performed at an outside facility. After informed consent was obtained, initial imaging was requested from external nuclear medicine specialists who performed the initial evaluation of the thyroid. The second group of patients (Group 2) comprised euthyroid patients who presented for thyroid nodule

work-up between 2011 and 2023 in the same department (*n* = 108). It should be noted that the original data set could be analyzed for each sonographic evaluation, as these were either stored in our system or the external facilites were able to provide us with their originals.

Eligible subjects who met inclusion criteria were identified. The complete preoperative data and pathology reports of the thyroidectomy specimens were available for 245 patients. For the evaluation of the diagnostic value of thyroid scintigraphy, TIRADS, and their combination, 21 patients were excluded due to hyperthyroidism, a further 5 because the pathology report could not be properly correlated with the nodules identified on imaging, 5 because of the insufficient image quality of the preoperative ultrasound or scintigraphy images, and 6 because the nodules were smaller than 1 cm. Preoperative imaging dated from February 2011 to July 2023.

The study protocol was approved by the ethics committee of the University of Münster (2022-428-f-S) and performed in accordance with the ethical standards as laid down in the 1964 Declaration of Helsinki and its later amendments.

2.2. Imaging and Image Analysis

In the 108 patients who had their thyroid nodule work-up in our department (Group 2), ultrasound of the thyroid gland was performed using either a Philips iU22 (Philips Healthcare, Eindhoven, The Netherlands), a Siemens Acuson NX3 Elite (Siemens Healthineers, Erlangen, Germany), or a Siemens Acuson S1000 (Siemens Healthineers, Erlangen, Germany) with a linear probe and a frequency of 5–10 MHz. Thyroid scintigraphy was acquired according to German guidelines 10–25 min post-injection of 70 MBq [^{99m}Tc]Tc-pertechnetate with a MIE Scintron (MiE Germany, Seth, Germany (LEHR collimation, matrix 128×128; acquisition time 10 min). One hundred patients (group 1) had their initial thyroid nodule work-up in various external institutions. Accordingly, different ultrasound and scintigraphy systems were used.

Ultrasound images were retrospectively classified according to the ACR-TIRADS [6] system by two experienced nuclear medicine specialists in a blinded fashion. The composition of nodules; echogenicity; shape; margin; and presence of echogenic foci were noted. On this basis, the corresponding EU-TIRADS categories of nodules were calculated. Nodule functional status was retrospectively assessed on preoperative [^{99m}Tc]Tc-pertechnetate thyroid scintigraphy scans as either hypofunctional (below the level of the surrounding thyroid tissue), isofunctional (at the level of the surrounding thyroid tissue), or hyperfunctional (above the level of the surrounding thyroid tissue) [15]. Only nodules that could be unequivocally correlated between ultrasound, scintigraphy, and pathology reports were included for further analysis.

2.3. Statistical Analyses

The distribution of quantitative parameters is summarized using the median and the interquartile range [lower–upper quartile]. The distribution of categorical variables is presented using absolute and relative frequencies, while 95% confidence intervals for proportions were calculated using the Wilson method. The association of categorical variables is visualized using bar charts and mosaic plots.

Multiple thyroid nodules per patient could be measured. To account for the resulting clustered nature of the data, further statistical analyses were based on generalized estimating equations (GEE) [16,17]. Measures of diagnostic performance and corresponding confidence intervals were calculated on a per nodule basis based on parameter estimates as well as the robust variance-covariance matrix resulting from the GEE approach [18]. The diagnostic performance of scintigraphy, the ACR-TIRADS, as well as the EU-TIRADS was quantified in terms of sensitivity, specificity, diagnostic odds ratio (DOR), as well as positive and negative diagnostic likelihood ratio (DLR+ and DLR−) [19]. The probability of malignancy was estimated depending on the result of the scintigraphy and the ACR-TIRADS or the EU-TIRADS. Confidence intervals for sensitivity, specificity, and the probability of malignancy were calculated using the back-transformation approach. Confidence intervals

for DLR+ and DLR− were calculated by applying the delta method to the log-transformed estimator followed by a back-transformation.

The diagnostic quality of different parameters was compared based on the area under the ROC curve (AUC) [20]. ROC curves facilitate the comparison of diagnostic methods, irrespective of the prevalence or the decision threshold. As sensitivity and specificity are independent of prevalence, the AUC is also independent of prevalence [19,21].

All analyses are intended as exploratory, as no adjustment for multiple testing was applied. The analyses were performed using R version 4.3.2 [22] and SAS ® software, version 9.4, for Windows (SAS Institute, Cary, NC, USA).

3. Results

3.1. Patient and Nodule Characteristics

Two hundred and eight patients met the inclusion criteria, with 68.9% being female and a median age of 47 years [37.3–58.0]. Chronic lymphocytic thyroiditis was identified in 28 patients based on pathology reports. Of the included patients, 129 were diagnosed with thyroid cancer, including papillary (98 cases), follicular (19 cases), anaplastic (1 case), medullary (4 cases), oncocytic (5 cases), and poorly differentiated (2 cases). Seventy-nine patients exhibited solely benign histological findings. The median number of nodules per patient was 1, with a maximum of 6. The median preoperative TSH level was 1.37 µU/mL [0.71–2.04 µU/mL].

The correlation of imaging and pathology reports was possible for 322 nodules (231 benign, 91 malignant). Note that, although 129 patients were diagnosed with thyroid cancer, only 91 malignant nodules were included in this analysis. This discrepancy arose because 44 patients with papillary microcarcinoma were included due to the presence of benign nodules larger than 1 cm, while microcarinomas were not included due to their size being less than 1 cm. One instance of non-invasive follicular thyroid neoplasm with papillary-like nuclear features (NIFTP) was classified as benign. Benign nodules had a median size of 20 mm [15–31 mm], and malignant nodules had a median size of 20 mm [15–32 mm].

On thyroid scintigraphy, 210 (65.22%) nodules were hypofunctional, 99 (30.75%) were isofunctional, and 13 (4.04%) were hyperfunctional, while 137, 83, and 11 out of 231 benign nodules and 73, 16, and 2 out of 91 malignant nodules were hypo-, iso-, or hyperfunctional, respectively. The malignancy rate of hypofunctional nodules was 34.8% (73/210), compared to 16.2% (16/99) for isofunctioning nodules and 15.4% (2/13) for hyperfunctioning nodules (see Figure 1).

The distribution of benign and malignant thyroid nodules across ACR-TIRADS scores/categories and EU-TIRADS categories, as well as the nodules' distribution among the categories and scores, is illustrated in Figure 1.

3.2. Diagnostic Performance of Thyroid Scintigraphy, ACR-TIRADS, and EU-TIRADS

As demonstrated, the malignancy rates for iso- and hyperfunctioning nodules were comparable—16.16% and 15.38%, respectively—while hypofunctioning nodules exhibited a higher malignancy rate of 34.76%. Based on these distinctions in malignancy rates, we analyzed the performance of thyroid scintigraphy by classifying it into two categories: hypofunctioning nodules and a combined group of isofunctioning and hyperfunctioning nodules. Our analysis revealed a sensitivity of 80% (95% CI: 0.71–0.88) and a specificity of 39% (95% CI: 0.32–0.47), with a positive likelihood ratio of 1.32 (95% CI: 1.13–1.56), a negative likelihood ratio of 0.50 (95% CI: 0.31–0.81), and a diagnostic odds ratio of 2.65 (95% CI: 1.41–4.98) (see Table 1).

For all 322 nodules

(A)
Scintigraphy

(B)
ACR–TIRADS Score

(C)
ACR–TIRADS Category

(D)
EU–TIRADS Category
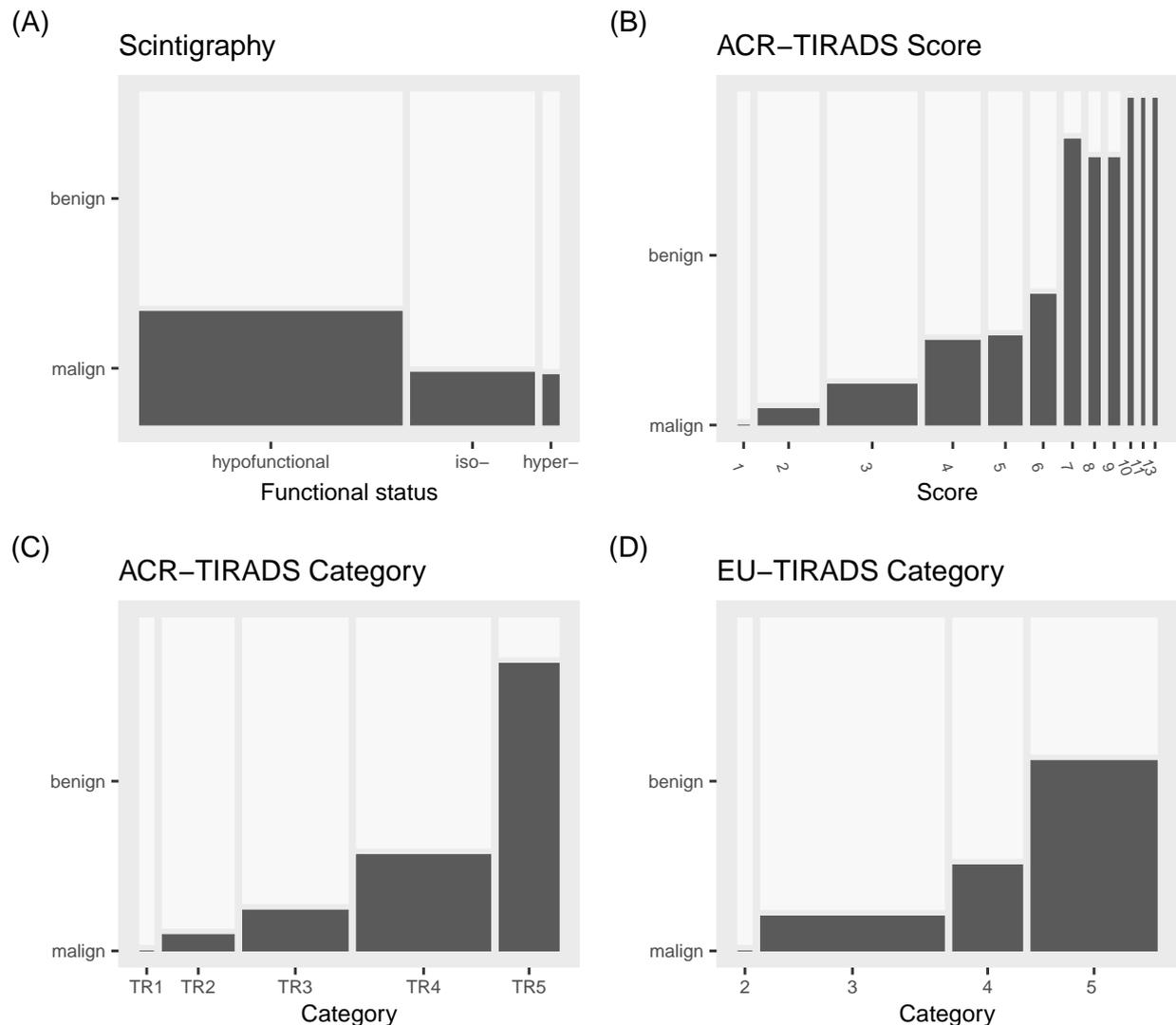
Figure 1. Mosaic plot illustrating the relationship between nodule status and (**A**) functional status assessed by scintigraphy, (**B**) ACR-TIRADS score, (**C**) ACR-TIRADS category, and (**D**) EU-TIRADS category. The width of each vertical segment represents the proportion of nodules that are hypofunctional, isofunctional, or hyperfunctional on thyroid scintigraphy (**A**), and the proportion of nodules in each ACR-TIRADS score (**B**), ACR-TIRADS category (**C**), and EU-TIRADS category (**D**). The height of each colored section corresponds to the proportion of malignant nodules, with full height representing 100%.

Receiver operating characteristic (ROC) analysis revealed virtually identical areas under the curve (AUC) for thyroid scintigraphy when categorized as a two- or three-level variable. Specifically, the AUC was 0.60 (95% CI: 0.55–0.66) on the two-level basis and 0.60 (95% CI: 0.55–0.66) on the three-level basis. Notably, the AUC of thyroid scintigraphy, whether categorized as a two- or three-level variable, was significantly lower than that of ACR-TIRADS at both the score level (0.60 versus 0.84, $p < 0.001$) and the category level (0.60 versus 0.83, $p < 0.001$), as well as compared to EU-TIRADS (0.60 versus 0.78, $p < 0.001$).

Table 1. Diagnostic performance metrics of scintigraphy, ACR-TIRADS, and EU-TIRADS using different cutoffs. Analyses based on 322 nodules in 208 patients. Numbers in brackets are 95% confidence intervals. * AUC for scintigraphy with the two levels "cold" and "indeterminate" or "warm". AUC for ACR-TIRADS using scores. ** AUC for scintigraphy with the three levels "cold", "indeterminate", and "warm". AUC for ACR-TIRADS using categories. [†] All 12 nodules classified as ACR-TIRADS TR1 or EU-TIRADS 2 were benign. Hence, using ACR-TIRADS ≥ 2 or EU-TIRADS ≥ 3 as a cutoff, the sensitivity was 100%, and the specificity 0.05%. For this combination, 95% confidence intervals and diagnostic likelihood ratios are not computable with the GEE-Model.

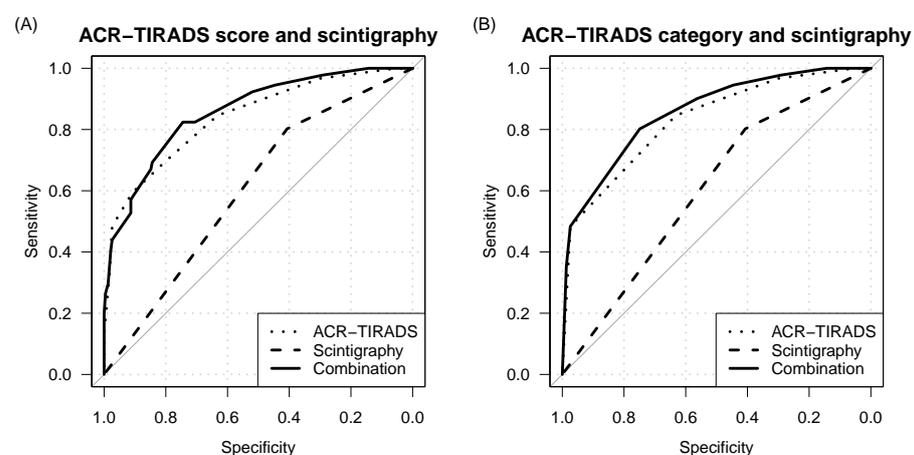|	Scintigraphy	ACR-TIRADS ≥ 2	ACR-TIRADS ≥ 3	ACR-TIRADS ≥ 4	ACR-TIRADS = 5	EU-TIRADS ≥ 3	EU-TIRADS ≥ 4	EU-TIRADS = 5
Sensitivity	0.80 (0.71–0.88)	1 [†]	0.97 (0.90–0.99)	0.85 (0.76–0.91)	0.48 (0.38–0.59)	1 [†]	0.83 (0.74–0.90)	0.67 (0.56–0.76)
Specificity	0.39 (0.32–0.47)	0.05 [†]	0.29 (0.23–0.36)	0.62 (0.55–0.69)	0.97 (0.94–0.99)	0.05 [†]	0.62 (0.54–0.68)	0.81 (0.74–0.86)
DLR+	1.32 (1.13–1.56)	-	1.36 (1.23–1.50)	2.26 (1.84–2.77)	18.56 (8.23–41.82)	-	2.17 (1.77–2.67)	3.46 (2.52–4.77)
DLR−	0.50 (0.31–0.81)	-	0.11 (0.04–0.36)	0.24 (0.14–0.40)	0.53 (0.43–0.65)	-	0.27 (0.17–0.43)	0.41 (0.30–0.56)
DOR	2.65 (1.41–4.98)	-	12.08 (3.55–41.10)	9.44 (4.90–18.19)	35.04 (14.11–86.98)	-	8.01 (4.31–14.91)	8.48 (4.82–14.92)
AUC *	0.6 (0.55–0.66)	0.84 (0.79–0.89)				0.78 (0.72–0.83)		
AUC **	0.6 (0.55–0.66)	0.83 (0.78–0.88)						

The diagnostic performance characteristics of ACR-TIRADS and EU-TIRADS for various cutoffs are detailed in Table 1. The AUC of ACR-TIRADS outperformed that of EU-TIRADS, both at the category and score levels (0.83 and 0.84 vs. 0.78, respectively, both $p < 0.001$).

### 3.3. Diagnostic Value of Combined TIRADS and Pertechnetate Scintigraphy

To evaluate whether incorporating radionuclide scanning alongside ACR- or EU-TIRADS enhances diagnostic efficacy, we conducted a comparative analysis of the area under the curve (AUC) between standalone TIRADS and multimodal TIRADS/scintigraphy models. One of the evaluated models combined ACR-TIRADS at the category level and scintigraphy as a two-level variable (hypofunctional vs. isofunctional and hyperfunctional). The integrated model demonstrated a superior AUC of 0.86 in contrast to 0.83 for ACR-TIRADS alone ($p = 0.039$). Similarly, the composite model pairing EU-TIRADS with scintigraphy as a two-level variable exhibited enhanced performance over EU-TIRADS alone in terms of AUC (0.80 versus 0.78, $p = 0.008$). Although the model combining ACR-TIRADS on a score level with scintigraphy showed a higher AUC than ACR-TIRADS on a score level alone, the difference did not reach statistical significance (0.86 vs. 0.84, $p = 0.064$). Further details regarding performance and differences between other tested models can be found in Table 2. Receiver operating characteristic (ROC) curves for standalone and combined models are shown in Figures 2 and 3.

Table 2. AUC comparison of TIRADS alone versus combined multimodal models. Analysis based on 322 nodules in 208 patients.

|	AUC	*p* Value
ACR-TIRADS (category)		
TIRADS alone	0.83 (0.78–0.88)	
+ scintigraphy (two-levels)	0.86 (0.81–0.90)	0.039
+ scintigraphy (three-levels)	0.86 (0.81–0.90)	0.039
ACR-TIRADS (score)		
TIRADS alone	0.84 (0.79–0.89)	
+ scintigraphy (two-levels)	0.86 (0.81–0.90)	0.064
+ scintigraphy (three-levels)	0.86 (0.81–0.91)	0.063
EU-TIRADS		
TIRADS alone	0.78 (0.72–0.83)	
+ scintigraphy (two-levels)	0.80 (0.75–0.85)	0.008
+ scintigraphy (three-levels)	0.80 (0.75–0.85)	0.011

Figure 2. ROC curves for: (**A**) thyroid scintigraphy (dashed line, AUC: 0.6), ACR-TIRADS score (dotted line, AUC: 0.84), and the combined model (solid line, AUC: 0.86). (**B**) Thyroid scintigraphy (dashed line, AUC: 0.6), ACR-TIRADS categories (dotted line; AUC: 0.83), and the combined model (solid line, AUC: 0.86).
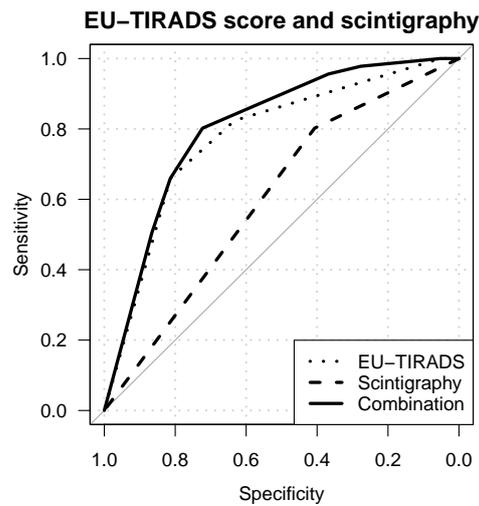
Figure 3. ROC curves for: thyroid scintigraphy (dashed line, AUC: 0.6), EU-TIRADS (dotted line; 0.78), and the combined model (solid line, AUC: 0.80).

As shown in Figure 4, malignant nodules were more frequently hypofunctional than benign nodules across all ACR-TIRADS categories except TR1, which had no malignant nodules. Similarly, malignant nodules were more frequently hypofunctional than benign nodules across all EU-TIRADS categories (except EU-TIRADS orange2, which had no malignant nodules), underscoring the potential of radionuclide scanning to further stratify nodules based on the probability of malignancy (Figure 5).
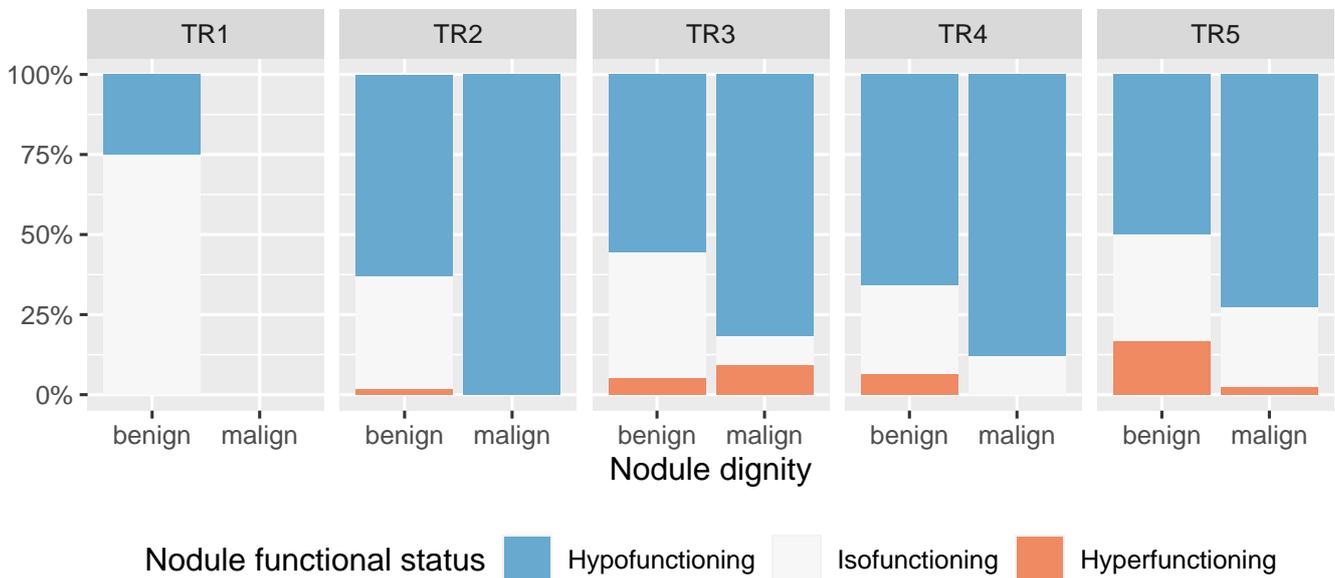

Figure 4. Barplot showing the relation of nodule dignity and functional status as assesed by scintigraphy, stratified by ACR-TIRADS category (*n* = 322 nodules in 208 patients). The proportion of hypofunctioning nodules is higher in malignant nodules in all ACR-TIRADS categories. Note that there are malignant hyperfunctional nodules.
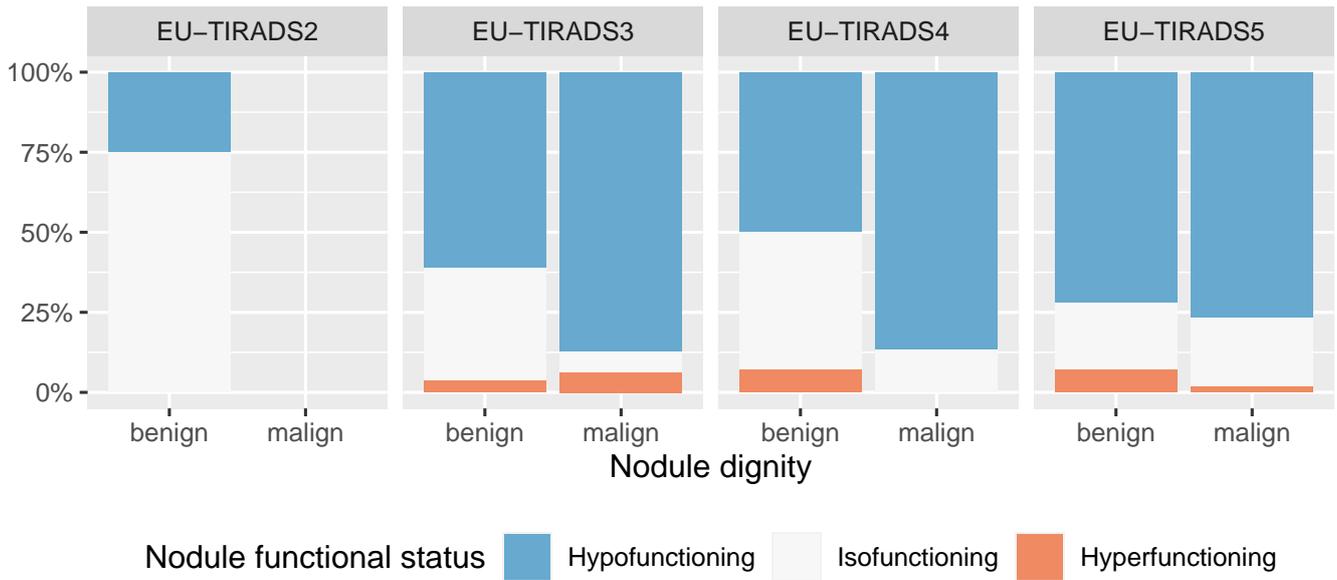
Figure 5. Barplot showing the relation of nodule dignity and functional status as assessed by scintigraphy, stratified by EU-TIRADS category (*n* = 322 nodules in 208 patients).

Indeed, integrating scintigraphy into the two TIRADS schemes led to a reordering of the estimated probability for malignancy across TIRADS categories (Figure 6). Notably, for ACR-TIRADS, the probability of malignancy for a hyper- or isofunctioning nodule in the TR4 category was comparable to that of a hypofunctional nodule in the TR3 category. Similarly, the probability of malignancy for an isofunctioning or hyperfunctioning nodule in the TR3 category was comparable to that of a hypofunctional nodule in the TR2 category. For EU-TIRADS, the probability of malignancy for a hyper- or isofunctioning nodule in the EU-TIRADS 4 category was comparable to that of a hypofunctional nodule in the EU-TIRADS 3 category (Figure 6).
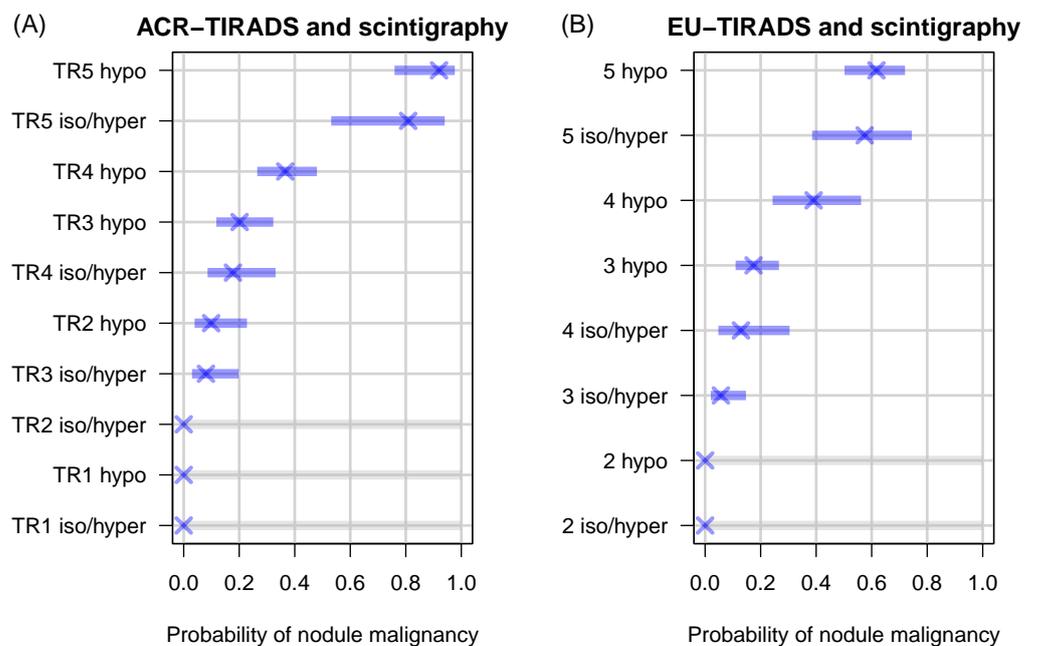
Figure 6. Estimated probability for nodule malignancy (purple cross) and corresponding 95% confidence intervals (bars, *n* = 322. (**A**) Combinations of ACR-TIRADS categories and thyroid

scintigraphy. Note that the probability of malignancy for a hyper- or isofunctioning nodule in the category TR4 is comparable to a hypofunctional nodule in the category TR3. The probability of malignancy for a iso-/or hyperfunctioning nodule in the category TR3 is comparable to a hypofunctioning nodule in the category TR2 (Confidence intervals for hyper- or isofunctioning nodules in categories TR1 and TR2 could not be estimated; hence, their error bars are greyed out for visual distinction). (**B**) Combinations of EU-TIRADS categories and thyroid scintigraphy. Note that the probability of malignancy for a hyper- or isofunctioning nodule in the category EU-TIRADS 4 is comparable to a hypofunctional nodule in the category EU-TIRADS 3 (confidence intervals cannot be estimated for category EU-TIRADS 2; hence, their error bars are greyed out for visual distinction).

4. Discussion

This retrospective study assessed the diagnostic efficacy of [^{99m}Tc]Tc-pertechnetate radionuclide scanning and two ultrasound-based Thyroid Imaging Reporting and Data Systems (TIRADS)—the American College of Radiology (ACR-TIRADS) and the European (EU-TIRADS)—for classifying thyroid nodules in euthyroid patients. Our investigation evaluated these methodologies both independently and within a multimodal framework, demonstrating that the integration of pertechnetate thyroid scintigraphy with ACR-TIRADS or EU-TIRADS enhances predictive accuracy beyond the capabilities of either TIRADS alone moderately.

The area under the receiver operating characteristic curve (AUC) was used to quantify and compare the diagnostic accuracy of different methods. The AUC calculated in our study can be interpreted as the probability that a randomly selected malignant nodule has a test result indicating greater suspicion than a randomly chosen non-malignant nodule [19,23]. Unlike both the negative and positive predictive values, the AUC assesses the diagnostic effectiveness of a test independently of disease prevalence.

Concerning the diagnostic performance of the investigated TIRADS, our findings revealed an AUC of 0.83 (95% CI: 0.78–0.88) for ACR-TIRADS. This aligns with the AUC ranges of 0.78 to 0.88 reported in prior studies [24–28]. For EU-TIRADS, we observed an AUC of 0.78 (95% CI: 0.72–0.83), consistent with previously reported ranges of 0.73 to 0.80 [24–26,28]. Comparatively, ACR-TIRADS outperformed EU-TIRADS in terms of AUC, with a statistically notable margin (0.83 vs. 0.78, $p < 0.001$), a finding echoed by Zhe Jin et al. in their analysis of 3438 thyroid nodules, primarily evaluated against surgical outcomes and histopathology [29]. Research by Sun Huh et al. and Yi-Xin Shi et al. similarly identified an advantage in AUC for ACR-TIRADS over EU-TIRADS, a distinction not observed in Seifert et al.'s study [24,30,31]. Unlike many previous studies, ours employed histopathological examination of each thyroid nodule as the reference standard, in contrast to relying on follow-up evaluations or fine-needle aspiration (FNA) for validation, providing a robust basis for comparing the diagnostic value of these methodologies.

The diagnostic performance of thyroid scintigraphy was limited compared to both ACR-TIRADS and EU-TIRADS, with an AUC of 0.60 versus 0.83 and 0.78, respectively (both $p < 0.001$). To enable comparisons with prior studies evaluating the diagnostic capacity of thyroid scintigraphy, which did not report AUC values, we conducted supplementary analyses of sensitivity and specificity. For these analyses, thyroid scintigraphy outcomes were grouped into two categories: hypofunctioning nodules in one group and iso- or hyperfunctioning nodules in the other, based on their similar malignancy rates. We observed a sensitivity of 80% (95%CI: 0.71–0.88) and a specificity of 39% (95%CI: 0.32–0.47), reflecting a pattern of relatively high sensitivity but modest specificity, aligning with existing literature. In close agreement with our results, Jones et al. reported a sensitivity of 82% and a specificity of 34% in their study of 175 euthyroid patients [32]. Similarly, Paneerselvan et al. observed a sensitivity of 60% and a specificity of 13% in a cohort of 109 euthyroid patients [33]. The studies of Kountakis et al. (*n* = 243) and Lumachi et al. (*n* = 496) reported higher sensitivities for scintigraphy of 0.91 and 0.95, respectively, but lower specificities of 0.19 and 0.21 [34,35]. Of note, TSH-values are were not reported in those two studies, which may limit the comparability of results [34,35].

Hyperfunctioning nodules are widely regarded as almost exclusively benign [10]. However, in our study, 2 of 13 hyperfunctioning nodules were found to be malignant (15.4%, 95% CI: 4.3–42.2%). While our study was not designed to assess the prevalence of malignancy in hyperfunctioning nodules, the observed proportion of malignancies is notable if hyperfunctional nodules were indeed virtually always benign. This assumption is predominantly founded upon research conducted during the nascent stages of radioiodine diagnostics and therapy [36]. Subsequent research has reported higher malignancy rates in hyperfunctioning nodules, more in line with our findings. For example, Ashcraft et al. (1981) identified malignancy in 4% (10/283) of nodules deemed hyperfunctional on radioiodine scans, and up to 29% (6/21) on [^{99m}Tc]Tc-pertechnetate scans [37]. Ikekubo et al. (1989) identified seven carcinomas in 16 resected hot nodules, and more recently, Dirikoc et al. (2015) reported that among 227 toxic and 31 nontoxic autonomous nodules operated on between 2008 and 2014, 20 (8.8%) and 2 (6.5%), respectively, were malignant [38,39]. Given the contrast between these findings and the prevailing assumption in the scientific literature that hyperfunctional nodules are almost exclusively benign, further research into the frequency of malignancy in hyperfunctioning nodules seems warranted.

Beyond comparing the individual diagnostic values of ultrasound-based TIRADS and thyroid scintigraphy, this study introduces an innovative perspective by systematically exploring the added diagnostic value of thyroid scintigraphy to TIRADS in a multimodal approach. Our findings suggest that integrated models combining radionuclide scanning with TIRADS improve diagnostic accuracy compared to TIRADS alone. Particularly, the integration of thyroid scintigraphy adjusted the estimated malignancy probability among TIRADS categories with intermediate risk. Comparable malignancy probabilities were observed between iso- or hyperfunctioning nodules in ACR-TIRADS TR4 and hypofunctioning nodules in TR3. Similarly, comparable malignancy probabilities were observed between the iso- and hyperfunctioning nodules in EU-TIRADS 4 and the hypofunctional nodule in the category EU-TIRADS 3, albeit with broad confidence intervals. This novel insight, pending validation by future research, holds potential clinical implications. In a study investigating 1257 nodules in 566 euthyroid patients presenting for thyroid nodule work-up, 54% of nodules were classified as TR3 or TR4 according to ACR-TIRADS and 75% as EU-TIRADS 3 or 4 according to EU-TIRADS [15]. Hence, the adjustment of estimated malignancy probability among TIRADS categories with intermediate risk, achieved by integrating thyroid scintigarphy, might have impacted a significant proportion of cases in this population. Our findings might be especially relevant in the management of multinodular goiter, potentially reducing the number of nodules necessitating fine-needle aspiration cytology (FNA). Furthermore, incorporating thyroid scintigraphy as a complementary diagnostic modality could be beneficial in up to 30% of cases where fine-needle aspiration cytology (FNA) yields indeterminate results [40]. However, the incremental diagnostic value of scintigraphy to TIRADS found in our study was only modest. Therefore, further research is essential to validate these findings and determine whether this moderate enhancement in diagnostic performance translates into meaningful clinical benefits.

Our study is subject to several limitations. Firstly, its retrospective design, which merges two distinct patient groups—Group 1, consisting of patients evaluated for thyroid nodules in our department with subsequent histopathology results, and Group 2, comprising patients undergoing thyroid cancer follow-up with available preoperative imaging—may have introduced bias. The prevalence of malignant nodules in our study population is high. Nevertheless, the comparison of the diagnostic potential of different methods remains valid as the AUC is independent of the prevalence [19,21]. Also, we consider the type of nodules a representative sample of malignant and non-malignant nodules. Therefore, our findings are transferable to patients undergoing initial nodule evaluation, where a lower prevalence of malignancy is typically expected [41]. Furthermore, approximately half of the patients underwent preoperative imaging at various institutions outside our department, leading to potential variability in imaging quality and execution. Also, all thyroid scans included in this study were conducted using [^{99m}Tc]Tc-pertechnetate.

[^{99m}Tc]Tc-pertechnetate is taken up into thyrocytes but not organified. Previous studies have shown that about 5% of nodules hyperfunctional in 99mTc-pertechnetate are so-called trapping only nodules, i.e., they are cold in ^{123}I scintigraphy [15,42]. Another limitation is that histopathological analyses were not revisited for this study, and some of the utilized reports precede the 2016 introduction of the "non-invasive follicular thyroid neoplasm with papillary-like nuclear features" (NIFTP) [43]. Consequently, certain nodules identified as malignant due to the diagnosis of the follicular variant of papillary thyroid carcinoma might actually correspond to NIFTP, which is regarded as benign in our analysis.

5. Conclusions

In conclusion, our study demonstrates an enhancement of diagnostic performance achieved by integrating thyroid scintigraphy with ACR- and EU-TIRADS for classifying thyroid nodules in euthyroid patients. Such a multimodal approach could improve risk stratification and management decisions, particularly in complex scenarios like multinodular goiter or with indeterminate FNA. Further research is warranted to validate these findings and explore their clinical implications.

Abbreviations

The following abbreviations are used in this manuscript:

AUC	Area under the receiver operating characteristic curve
ATA	American Thyroid Association
DOR	Diagnostic odds ratio
DLR+	Positive diagnostic likelihood ratio
DLR-	Negative diagnostic likelihood ratio
FNA	Fine-needle aspiration cytology
GEE	Generalized estimating equations
ROC	Receiver operating characteristic
TIRADS	Thyroid Imaging Reporting and Data System

References

1. Fisher, S.B.; Perrier, N.D. The incidental thyroid nodule. *CA A Cancer J. Clin.* **2018**, *68*, 97–105. [CrossRef] [PubMed]
2. Guth, S.; Theune, U.; Aberle, J.; Galach, A.; Bamberger, C. Very high prevalence of thyroid nodules detected by high frequency (13 MHz) ultrasound examination. *Eur. J. Clin. Investig.* **2009**, *39*, 699–706. [CrossRef]
3. HaugenBryan, R.; AlexanderErik, K.; BibleKeith, C.; DohertyGerard, M.; MandelSusan, J.; NikiforovYuri, E.; RandolphGregory, W.; SawkaAnna, M.; SchuffKathryn, G.; ShermanSteven, I.; 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**, *25*, 567–610 .

4. Wong, R.; Farrell, S.G.; Grossmann, M. Thyroid nodules: Diagnosis and management. *Med. J. Aust.* **2018**, *209*, 92–98. [CrossRef] [PubMed]

5. Orlandella, F.M.; Imperlini, E.; Pane, K.; Luciano, N.; Braile, M.; De Stefano, A.E.; Iervolino, P.L.C.; Ruocco, A.; Orrù, S.; Franzese, M.; et al. miR-331-5p Affects Motility of Thyroid Cancer Cell Lines and Regulates BID Expression. *Biomedicines* **2024**, *12*, 658. [CrossRef] [PubMed]

6. Tessler, F.N.; Middleton, W.D.; Grant, E.G.; Hoang, J.K.; Berland, L.L.; Teefey, S.A.; Cronan, J.J.; Beland, M.D.; Desser, T.S.; Frates, M.C.; et al. ACR thyroid imaging, reporting and data system (TI-RADS): White paper of the ACR TI-RADS committee. *J. Am. Coll. Radiol.* **2017**, *14*, 587–595. [CrossRef] [PubMed]

7. Russ, G.; Bonnema, S.J.; Erdogan, M.F.; Durante, C.; Ngu, R.; Leenhardt, L. European Thyroid Association guidelines for ultrasound malignancy risk stratification of thyroid nodules in adults: The EU-TIRADS. *Eur. Thyroid J.* **2017**, *6*, 225–237. [CrossRef]

8. Kwak, J.Y.; Han, K.H.; Yoon, J.H.; Moon, H.J.; Son, E.J.; Park, S.H.; Jung, H.K.; Choi, J.S.; Kim, B.M.; Kim, E.K. Thyroid imaging reporting and data system for US features of nodules: A step in establishing better stratification of cancer risk. *Radiology* **2011**, *260*, 892–899. [CrossRef]

9. Dietlein, M.; Eschner, W.; Lassmann, M.; Verburg, F.; Luster, M. *DGN-Handlungsempfehlung (S1-Leitlinie) Schilddrüsenszintigraphie (Version 4) Stand: 10/2014–AWMF-Registernummer: 031-011*; Deutsche Gesellschaft für Nuklearmedizin: Leipzig, Germany , 2019.

10. Verburg, F.A.; Aktolun, C.; Chiti, A.; Frangos, S.; Giovanella, L.; Hoffmann, M.; Iakovou, I.; Mihailovic, J.; Krause, B.J.; Langsteger, W.; et al. Why the European Association of Nuclear Medicine has declined to endorse the 2015 American Thyroid Association management guidelines for adult patients with thyroid nodules and differentiated thyroid cancer. *Eur. J. Nucl. Med. Mol. Imaging* **2016**, *43*, 1001–1005. [CrossRef] [PubMed]

11. Schicha, H.; Hellmich, M.; Lehmacher, W.; Eschner, W.; Schmidt, M.; Kobe, C.; Schober, O.; Dietlein, M. Should all patients with thyroid nodules ≥1 cm undergo fine-needle aspiration biopsy? *Nuklearmedizin-NuclearMedicine* **2009**, *48*, 79–83.

12. Giovanella, L.; Avram, A.M.; Iakovou, I.; Kwak, J.; Lawson, S.A.; Lulaj, E.; Luster, M.; Piccardo, A.; Schmidt, M.; Tulchinsky, M.; et al. EANM practice guideline/SNMMI procedure standard for RAIU and thyroid scintigraphy. *Eur. J. Nucl. Med. Mol. Imaging* **2019**, *46*, 2514–2525. [CrossRef] [PubMed]

13. Durante, C.; Hegedüs, L.; Czarniecka, A.; Paschke, R.; Russ, G.; Schmitt, F.; Soares, P.; Solymosi, T.; Papini, E. 2023 European Thyroid Association clinical practice guidelines for thyroid nodule management. *Eur. Thyroid J.* **2023**, *12*, e230067. [CrossRef] [PubMed]

14. NICE guideline [NG230]. *Thyroid Cancer: Assessment and Management*; National Institute for Health and Care Excellence (NICE): London, UK, 2022.

15. Noto, B.; Eveslage, M.; Pixberg, M.; Gonzalez Carvalho, J.M.; Schäfers, M.; Riemann, B.; Kies, P. Prevalence of hyperfunctioning thyroid nodules among those in need of fine needle aspiration cytology according to ATA 2015, EU-TIRADS, and ACR-TIRADS. *Eur. J. Nucl. Med. Mol. Imaging* **2020**, *47*, 1518–1526. [CrossRef]

16. Liang, K.Y.; Zeger, S.L. Longitudinal data analysis using generalized linear models. *Biometrika* **1986**, *73*, 13–22. [CrossRef]

17. Zeger, S.L.; Liang, K.Y. Longitudinal data analysis for discrete and continuous outcomes. *Biometrics* **1986**, *42*, 121–130. [CrossRef]

18. Sternberg, M.R.; Hadgu, A. A GEE approach to estimating sensitivity and specificity and coverage properties of the confidence intervals. *Stat. Med.* **2001**, *20*, 1529–1539. [CrossRef]

19. Zhou, X.H.; Obuchowski, N.A.; McClish, D.K. *Statistical Methods in Diagnostic Medicine*; John Wiley & Sons: Hoboken, NJ, USA, 2011; Volume 712.

20. Obuchowski, N.A. Nonparametric analysis of clustered ROC curve data. *Biometrics* **1997**, *53*, 567–578. [CrossRef] [PubMed]

21. Obuchowski, N.A. ROC Analysis. *Am. J. Roentgenol.* **2005**, *184*, 364–372. PMID: 15671347. [CrossRef] [PubMed]

22. R Core Team. *R: A Language and Environment for Statistical Computing*; R Foundation for Statistical Computing: Vienna, Austria, 2023.

23. Hanley, J.A.; McNeil, B.J. The meaning and use of the area under a receiver operating characteristic (ROC) curve. *Radiology* **1982**, *143*, 29–36. [CrossRef] [PubMed]

24. Shi, Y.X.; Chen, L.; Liu, Y.C.; Zhan, J.; Diao, X.H.; Fang, L.; Chen, Y. Differences among the Thyroid Imaging Reporting and Data System proposed by Korean, the American College of Radiology and the European Thyroid Association in the diagnostic performance of thyroid nodules. *Transl. Cancer Res.* **2020**, *9*, 4958–4967. [CrossRef] [PubMed]

25. Torshizian, A.; Hashemi, F.; Khoshhal, N.; Ghodsi, A.; Rastegar, H.; Mousavi, Z.; Dadgar Moghadam, M.; Mohebbi, M. Diagnostic Performance of ACR TI-RADS and ATA Guidelines in the Prediction of Thyroid Malignancy: A Prospective Single Tertiary Center Study and Literature Review. *Diagnostics* **2023**, *13*, 2972. [CrossRef]

26. Kim, J.S.; Kim, B.G.; Stybayeva, G.; Hwang, S.H. Diagnostic Performance of Various Ultrasound Risk Stratification Systems for Benign and Malignant Thyroid Nodules: A Meta-Analysis. *Cancers* **2023**, *15*, 424. [CrossRef] [PubMed]

27. 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. *Am. J. Roentgenol.* **2021**, *216*, 38–47. [CrossRef]

28. Hekimsoy, I.; Öztürk, E.; Ertan, Y.; Orman, M.N.; Kavukçu, G.; Özgen, A.G.; Özdemir, M.; Özbek, S.S. Diagnostic performance rates of the ACR-TIRADS and EU-TIRADS based on histopathological evidence. *Diagn. Interv. Radiol.* **2021**, *27*, 511. [CrossRef]

29. Jin, Z.; Pei, S.; Shen, H.; Ouyang, L.; Zhang, L.; Mo, X.; Chen, Q.; You, J.; Zhang, S.; Zhang, B. Comparative study of C-TIRADS, ACR-TIRADS, and EU-TIRADS for diagnosis and management of thyroid nodules. *Acad. Radiol.* **2023**, *30*, 2181–2191. [CrossRef] [PubMed]

30. Huh, S.; Lee, H.S.; Yoon, J.; Kim, E.K.; Moon, H.J.; Yoon, J.H.; Park, V.Y.; Kwak, J.Y. Diagnostic performances and unnecessary US-FNA rates of various TIRADS after application of equal size thresholds. *Sci. Rep.* **2020**, *10*, 10632. [CrossRef]

31. Seifert, P.; Schenke, S.; Zimny, M.; Stahl, A.; Grunert, M.; Klemenz, B.; Freesmeyer, M.; Kreissl, M.C.; Herrmann, K.; Görges, R. Diagnostic performance of Kwak, EU, ACR, and Korean TIRADS as well as ATA guidelines for the ultrasound risk stratification of non-autonomously functioning thyroid nodules in a region with long history of iodine deficiency: A German multicenter trial. *Cancers* **2021**, *13*, 4467. [CrossRef] [PubMed]

32. Jones, A.; Aitman, T.; Edmonds, C.; Burke, M.; Hudson, E.; Tellez, M. Comparison of fine needle aspiration cytology, radioisotopic and ultrasound scanning in the management of thyroid nodules. *Postgrad. Med. J.* **1990**, *66*, 914–917. [CrossRef] [PubMed]

33. Panneerselvan, R.; Schneider, D.F.; Sippel, R.S.; Chen, H. Radioactive iodine scanning is not beneficial but its use persists for euthyroid patients. *J. Surg. Res.* **2013**, *184*, 269–273. [CrossRef]

34. Kountakis, S.E.; Skoulas, I.G.; Maillard, A. The radiologic work-up in thyroid surgery: Fine-needle biopsy versus scintigraphy and ultrasound. *Ear. Nose Throat J.* **2002**, *81*, 151–154. [CrossRef]

35. Lumachi, F.; Varotto, L.; Borsato, S.; Tregnaghi, A.; Zucchetta, P.; Marzola, M.C.; Cecchin, D.; Bui, F. Usefulness of 99mTc-pertechnetate scintigraphy and fine-needle aspiration cytology in patients with solitary thyroid nodules and thyroid cancer. *Anticancer Res.* **2004**, *24*, 2531–2534. [PubMed]

36. Lau, L.W.; Ghaznavi, S.; Frolkis, A.D.; Stephenson, A.; Robertson, H.L.; Rabi, D.M.; Paschke, R. Malignancy risk of hyperfunctioning thyroid nodules compared with non-toxic nodules: Systematic review and a meta-analysis. *Thyroid Res.* **2021**, *14*, 1–16. [CrossRef] [PubMed]

37. Ashcraft, M.W.; Van Herle, A.J. Management of thyroid nodules. II: Scanning techniques, thyroid suppressive therapy, and fine needle aspiration. *Head Neck Surg.* **1981**, *3*, 297–322. [CrossRef]

38. Ikekubo, K.; Hino, M.; Ito, H.; Otani, M.; Yamaguchi, H.; Saiki, Y.; Ui, K.; Habuchi, Y.; Ishihara, T.; Mori, T. Thyroid carcinoma m solitary hot thyroid lesions on Tc-99m sodium pertechnetate scans. *Ann. Nucl. Med.* **1989**, *3*, 31–36. [CrossRef] [PubMed]

39. Dirikoc, A.; Polat, S.B.; Kandemir, Z.; Aydin, C.; Ozdemir, D.; Dellal, F.D.; Ersoy, R.; Cakir, B. Comparison of ultrasonography features and malignancy rate of toxic and nontoxic autonomous nodules: A preliminary study. *Ann. Nucl. Med.* **2015**, *29*, 883–889. [CrossRef]

40. Rossi, E.D.; Pantanowitz, L.; Faquin, W.C. The role of molecular testing for the indeterminate thyroid FNA. *Genes* **2019**, *10*, 736. [CrossRef]

41. Grussendorf, M.; Ruschenburg, I.; Brabant, G. Malignancy rates in thyroid nodules: A long-term cohort study of 17,592 patients. *Eur. Thyroid J.* **2022**, *11*, e220027. [CrossRef]

42. Reschini, E.; Ferrari, C.; Castellani, M.; Matheoud, R.; Paracchi, A.; Marotta, G.; Gerundini, P. The trapping-only nodules of the thyroid gland: Prevalence study. *Thyroid* **2006**, *16*, 757–762. [CrossRef]

43. Nikiforov, Y.E.; Seethala, R.R.; Tallini, G.; Baloch, Z.W.; Basolo, F.; Thompson, L.D.; Barletta, J.A.; Wenig, B.M.; Al Ghuzlan, A.; Kakudo, K.; et al. Nomenclature revision for encapsulated follicular variant of papillary thyroid carcinoma: A paradigm shift to reduce overtreatment of indolent tumors. *JAMA Oncol.* **2016**, *2*, 1023–1029. [CrossRef] [PubMed]