




Article

Sonographic Risk Stratification Systems for Thyroid Nodules as Rule-Out Tests in Older Adults

Giorgio Grani ¹, Gabriela Brenta ², Pierpaolo Trimboli ^{3,4}, Rosa Falcone ¹, Valeria Ramundo ¹, Marianna Maranghi ¹, Piernatale Lucia ¹, Sebastiano Filetti ^{1,*} and Cosimo Durante ^{1,*}

¹ Department of Translational and Precision Medicine, “Sapienza” University of Rome, Viale del Policlinico 155, 00161 Rome, Italy; giorgio.grani@uniroma1.it (G.G.); rosa.falcone@uniroma1.it (R.F.); valeria.ramundo@uniroma1.it (V.R.); marianna.maranghi@uniroma1.it (M.M.); piernatale.lucia@uniroma1.it (P.L.)

² Endocrinology Department, Cesar Milstein Hospital, Buenos Aires CABA C1221ACI, Argentina; gbrenta@gmail.com

³ Clinic of Endocrinology, Ente Ospedaliero Cantonale, 6900 Lugano, Switzerland; pierpaolo.trimboli@eoc.ch

⁴ Faculty of Biomedical Sciences, Università della Svizzera Italiana (USI), 6900 Lugano, Switzerland

* Correspondence: sebastiano.filetti@uniroma1.it (S.F.); cosimo.durante@uniroma1.it (C.D.)

Received: 31 July 2020; Accepted: 28 August 2020; Published: 30 August 2020



Simple Summary: The use of risk-stratification systems for thyroid nodules based on ultrasound features may reduce the number of biopsies to be performed. The aim of our study was to assess the diagnostic performance of these systems in different age groups. We confirmed that all systems had a significant discriminative performance in all age groups. The system proposed by the American College of Radiology was the best performing one, but all risk-stratification systems could avoid a sizable number of biopsies when applied as rule-out tests (to exclude malignancy) in elderly patients.

Abstract: Ultrasonographic risk-stratification systems (RSS), including various Thyroid Imaging Reporting and Data Systems (TIRADS), were proposed to improve reporting and reduce the number of fine-needle aspiration biopsies. However, age might be a confounder since some suspicious ultrasonographic features lack specificity in elderly patients. We aimed to investigate whether the diagnostic performance of the RSS varied between age groups. All patients consecutively referred for thyroid biopsy between November 1, 2015, and March 10, 2020, were included. The malignancy risk of each nodule was estimated according to five RSS: the American Association of Clinical Endocrinologists/American College of Endocrinology/Associazione Medici Endocrinologi guidelines, the American College of Radiology (ACR) TIRADS, the American Thyroid Association guidelines, the European TIRADS, and the Korean TIRADS. Overall, 818 nodules (57 malignant) were evaluated. The malignancy rate was higher in patients ≤ 65 years (8.1%) than in patients > 65 years (3.8%; $p = 0.02$). All RSS confirmed a significant discriminative performance in both age groups, with a negative predictive value of 100% in patients > 65 years, although specificity was lower in older patients. The ACR TIRADS was the best performing in both age groups. RSS could avoid a sizable number of biopsies when applied as rule-out tests in elderly patients.

Keywords: ultrasonography; ultrasound; thyroid nodule; reproducibility of results; sensitivity and specificity; aged adults; elderly

1. Introduction

Various published risk-stratification guidelines [1–5] provide recommendations for the evaluation of thyroid nodules based on the combination of nodule size and ultrasonographic (US) appearance [6],

with the aim of improving the standardization of thyroid ultrasound reporting and the identification of the small subset of nodules that warrant fine-needle aspiration biopsy (FNAB). The performance of these systems has been validated in retrospective [7–10] and prospective studies [11–13] and has also been confirmed by a recent meta-analysis [14]. Classification is usually based on the recognition of patterns of sonographic features, though the American College of Radiology (ACR) Thyroid Imaging Reporting and Data System (TIRADS) [4] assigns nodules points for each of five US categories, which are then added to determine a final class. The decision of whether to perform a biopsy or monitor the nodule is based on the maximum nodule diameter, with a different threshold for each risk class. For nodules in high-risk classes, FNAB is usually indicated if the maximum diameter is 1 cm or more. For nodules in lower risk classes, the size thresholds for FNAB range from 1.5 to 3 cm, depending on the risk-stratification system. It has been demonstrated that the various risk-stratification schemes vary in their ability to reduce the number of unnecessary FNABs. However, the ACR TIRADS has been found to outperform the other risk-stratification systems in its ability to decrease the number of biopsies while improving diagnostic accuracy [7,11,14].

Most recently, the ACR TIRADS and the sonographic risk-stratification systems proposed by the American Thyroid Association (ATA) [2] and the American Association of Clinical Endocrinologists/American College of Endocrinology/Associazione Medici Endocrinologi (AACE/ACE/AME) guidelines [1] have been validated in a geriatric population [15]. In that study, it is suggested that age might be a confounder since some suspicious US features of thyroid nodules lack specificity in elderly patients [15].

The aim of this study was to investigate whether the diagnostic performance (and the number of avoided biopsies) of the five most widely used sonographic risk-stratification systems (also including the EU-TIRADS of the European Thyroid Association [3] and the K-TIRADS of the Korean Society of Thyroid Radiology [5]) varied between age groups.

2. Results

A total of 1349 thyroid nodule sonographic examinations before biopsy were evaluated. Some biopsies were performed multiple times on the same nodule during the study period ($n = 119$) due to cytology report suggestions, indeterminate cytology, non-diagnostic cytology, nodule growth, or the appearance of new suspicious features. In these cases, only the last examination was considered. The actual number of biopsied nodules was 1230 (1145 patients). Of these, 113 nodules were excluded because the maximum diameter was less than one centimeter, and 299 were excluded because of an inconclusive diagnosis (non-diagnostic or indeterminate cytology report without surgical pathology). To evaluate the potential impact of these exclusions on the age distribution of our final cohort, we compared the age distribution in the excluded and analyzed groups. Individuals with smaller nodules were younger (median 52 years (interquartile range, IQR 42–63) versus 57 years (IQR 47–67), $p = 0.003$), while patients with an inconclusive diagnosis were older (median 58 years (IQR 47–68 years) vs. 55 (IQR 46–66), $p = 0.005$). However, the age distribution was comparable between the group with excluded nodules and the final cohort (Figure 1).

The final cohort included 818 thyroid nodules, with a median maximum diameter of 20.7 (IQR 15–28.8) mm, of which 57 (7%) were classified as malignant. Seventy-five patients were submitted to surgery (23 benign nodules, and 52 of the malignant nodules), with a median maximum diameter of the biopsied nodule of 16.8 (IQR 13.1–27.7) mm, smaller than the not resected biopsied nodules (21.1 mm, IQR 15.4–29.1 mm; $p = 0.025$). The malignancy rate was higher in patients ≤ 65 years (8.1%) than in patients older than 65 years (3.8%; $p = 0.02$). The need for surgery was not significantly different between groups (13, 6.1% in the elderly group, and 62, 10.2% in the younger group; $p = 0.096$). We analyzed the distribution of single sonographic features (Table 1) and found no differences between the two age groups except for cystic nodules, which were more common in young patients, and calcifications, which were more frequent in the elderly.

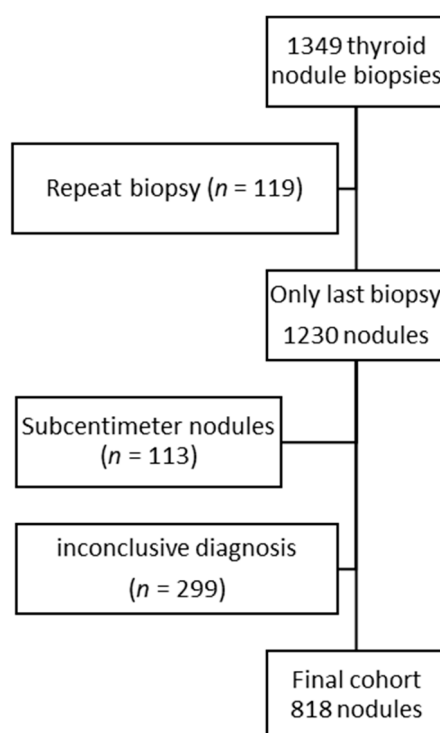


Figure 1. Flow chart of included nodules. The age distribution of the final cohort (median 56 years, interquartile range 46–66 years) was not significantly different from the age distribution of the excluded nodules (median 56 years, interquartile range 46–67 years; $p = 0.75$).

When using these features to classify nodules according to the five sonographic risk-stratification systems, we found no differences in the distribution of the two age groups with the AACE/ACE/AME, ACR TIRADS, and K-TIRADS systems (Table 2). However, elderly patients more commonly harbored EU-TIRADS 5 nodules and lesions that were non-classifiable in the ATA scheme (i.e., isoechoic nodules with other suspicious features like microcalcification, irregular margins, taller-than-wide shape, disrupted rim calcifications with a small extrusive hypoechoic soft tissue component, or evidence of extrathyroidal extension). However, if non-classifiable nodules were grouped with intermediate-suspicion nodules, the difference disappeared (Chi-square test; $p = 0.214$). The malignancy rate for each sonographic risk class is reported for each age group in Table 2.

Finally, we evaluated the diagnostic accuracy of the five systems by calculating sensitivity, specificity, positive and negative predictive values, and area under the receiver operating characteristic curve (AUROC) for patients younger and older than 65 years (Table 3). All systems confirmed a statistically significant discriminative performance in both age groups, with the specificity and positive predictive values of the systems being generally lower in older patients. However, all systems achieved a negative predictive value of 100% in patients > 65 years since no malignancy was missed by any of the systems. However, it is worth noting that for the ATA system, such a test performance was not confirmed if non-classifiable nodules were not submitted to biopsy. In fact, 16/172 (9.3%) of these nodules harbored a malignancy, and if they were not subjected to biopsy, the negative predictive value of the ATA system would decrease to 96.1% (95% CI 90.3–98.9%) in the > 65 group and to 94.1% (95% CI 90.4–96.6%) in patients ≤ 65 years. The application of these systems would avoid 13.2–45.3% of all FNABs in patients > 65 years. The ACR TIRADS was the best performing system as it was able to prevent the highest number of biopsies and achieve the best discriminative performance as estimated by the AUROC in both age groups.

Table 1. Sonographic features of thyroid nodules according to age group.

Feature	Descriptor	Age		Total	p-Value ¹
		≤65 years	>65 years		
Maximum diameter, mm (IQR)		21 (14.9–29.2)	20.4 (15.2–27.7)	20.7 (15–28.8)	0.798 ²
	Regular	245 40.4%	68 32.1%	313 38.3%	
Margins	Irregular/lobulated	92 15.2%	32 15.1%	124 15.2%	0.246
	Ill-defined	37 6.1%	14 6.6%	51 6.2%	
	Infiltrating	3 0.5%	1 0.5%	4 0.5%	
	Hypoechoic halo	229 37.8%	97 45.8%	326 39.9%	
Cystic composition		23 3.8%	1 0.5%	24 2.9%	0.006
Solid composition		180 29.7%	66 31.1%	246 30.1%	0.379
	Septa	20 3.3%	7 3.3%	27 3.3%	
Mixed composition	Non-nodular	364 60.1%	137 64.6%	501 61.2%	0.271
	Central nodular solid portion	6 1%	1 0.5%	7 1%	
	Eccentric nodular solid portion	13 2.1%	0 0%	13 1.6%	
Echogenicity	Anechoic	7 1.2%	1 0.5%	8 1.0%	0.107
	Hyperechoic	4 0.7%	3 1.4%	7 0.9%	
	Isoechoic	427 70.5%	167 78.8%	594 72.6%	
	Hypoechoic	147 24.3%	36 17.0%	183 22.4%	
	Markedly hypoechoic	21 3.5%	5 2.4%	26 3.2%	
Hyperechoic Foci	None	437 72.1%	162 76.4%	599 73.2%	0.277
	Comet-tail	43 7.1%	9 4.2%	52 6.4%	
	Indeterminate	126 20.8%	41 19.3%	167 20.4%	
Calcifications	None	486 80.2%	147 69.3%	633 77.4%	0.005
	Macrocalcifications	74 12.2%	42 19.8%	116 14.2%	
	Microcalcifications	46 7.6%	23 10.8%	69 8.4%	
Suspicious extrathyroidal extension		6 1.0%	1 0.5%	7 0.9%	0.422
Suspicious lymph nodes		10 1.7%	2 0.9%	12 1.5%	0.361
Taller-than-wide shape		101 16.7%	42 19.8%	143 17.5%	0.175
Total		606	212	818	

¹ Differences in the distribution of sonographic features between age groups were analyzed using the Chi-square test or the Fisher exact test. ² Differences in size between age groups were analyzed using the Mann–Whitney U test. p-values < 0.05 are highlighted in bold.

Table 2. Distribution of risk classes according to 5 sonographic classification systems and actual malignancy rate in the two age groups.

RSS	Category	Age		Total	p-Value ¹	Malignancy Rate		
		≤65 years	>65 years			≤65 years	>65 years	Overall
ATA guidelines	Benign	6 1.0%	0 0%	6 0.7%	0.041	0 0%	- -	0 0%
	Very low suspicion	301 49.7%	94 44.3%	395 48.3%		5 1.7%	0 0%	5 1.3%
		Low suspicion	78 12.9%	33 15.6%		111 13.6%	1 1.3%	1 3.0%
	Intermediate suspicion	31 5.1%	5 2.4%	36 4.4%		6 19.4%	0 0%	6 16.7%
		High suspicion	75 12.4%	23 10.8%		98 12.0%	25 33.3%	3 13.0%
	Not classifiable	115 19.0%	57 26.9%	172 21.0%		12 10.4%	4 7.0%	16 9.3%

Table 2. Cont.

RSS	Category	Age		Total	<i>p</i> -Value ¹	Malignancy Rate		
		≤65 years	>65 years			≤65 years	>65 years	Overall
K-TIRADS	K-TIRADS 2	11 1.8%	2 0.9%	13 1.6%	0.477	0 0%	0 0%	0 0%
	K-TIRADS 3	375 61.9%	125 59.0%	500 61.1%		7 1.9%	1 0.8%	8 1.6%
	K-TIRADS 4	179 29.5%	73 34.4%	252 30.8%		20 11.2%	6 8.2%	26 10.3%
	K-TIRADS 5	41 6.8%	12 5.7%	53 6.5%		22 53.7%	1 8.3%	23 43.4%
	AACE/ACE/AME	Low risk	48 7.9%	11 5.2%		59 7.2%	0 0%	0 0%
	Intermediate risk	358 59.1%	119 56.1%	477 58.3%	0.190	10 2.8%	1 0.8%	11 2.3%
	High risk	200 33.0%	82 38.7%	282 34.5%		39 19.5%	7 8.5%	46 16.3%
ACR TIRADS	TR1	24 4.0%	5 2.4%	29 3.5%	0.489	0 0%	0 0%	0 0%
	TR2	164 27.1%	48 22.6%	212 25.9%		2 1.2%	0 0%	2 0.9%
	TR3	106 17.5%	39 18.4%	145 17.7%		2 1.9%	0 0%	2 1.4%
	TR4	208 34.3%	83 39.2%	291 35.6%		13 6.3%	5 6.0%	18 6.2%
	TR5	104 17.2%	37 17.5%	141 17.2%		32 30.8%	3 8.1%	35 24.8%
EU-TIRADS	EU TIRADS 2	6 1.0%	1 0.5%	7 0.9%	0.035	0 0%	0 0%	0 0%
	EU TIRADS 3	318 52.5%	113 53.3%	431 52.7%		6 1.9%	1 0.9%	7 1.6%
	EU TIRADS 4	88 14.5%	16 7.5%	104 12.7%		6 6.8%	0 0%	6 5.8%
	EU TIRADS 5	194 32.0%	82 38.7%	276 33.7%		37 19.1%	7 8.5%	44 15.9%

Abbreviations: AACE/ACE/AME: American Association of Clinical Endocrinologists/American College of Endocrinology/Associazione Medici Endocrinologi; ACR TIRADS: American College of Radiology Thyroid Imaging Reporting and Data System; ATA: American Thyroid Association; EU-TIRADS: European Thyroid Imaging Reporting and Data System; K-TIRADS: Korean Thyroid Imaging Reporting and Data System; RSS: sonographic risk-stratification system. ¹ Differences in the distribution of RSS classes between age groups were analyzed using the Chi-square test. *p*-values <0.05 are highlighted in bold.

Table 3. Diagnostic performance of the 5 sonographic stratification systems, stratified according to age ≤ 65 years or > 65 years.

	RSS	Avoided Biopsies (%) ^b	Sensitivity (95% CI)	Specificity (95% CI)	PPV (95% CI)	NPV (95% CI)	AUROC (95% CI) ^c
≤65 years	ACR TIRADS	305 (50.3%)	83.67% (70.3–92.68%)	53.3% (49.1–57.5%)	13.6% (10.0–18.0%)	97.4% (94.9–98.9%)	0.68 (0.63–0.74)
	AACE/ACE/AME	197* (32.5%)	87.8% (75.2–95.4%)	34.3% (30.3–38.4%)	10.5% (7.7–13.9%)	96.95% (93.5–98.9%)	0.61 (0.56–0.66)
	ATA ^a	137* (22.6%)	93.9% (83.1–98.7%)	24.1% (20.6–27.8%)	9.8% (7.3–12.9%)	97.8% (93.7–99.5%)	0.59 (0.55–0.63)
	EU-TIRADS	154* (25.4%)	85.7% (72.8–94.1%)	26.4% (22.8–30.3%)	9.3% (6.8–12.3%)	95.4% (90.9–98.15%)	0.56 (0.51–0.61)
	K-TIRADS	79* (13.0%)	95.9% (86.0–99.5%)	13.8% (11.1–17%)	8.9% (6.6–11.7%)	97.47% (91.1–99.7%)	0.55 (0.52–0.58)
>65 years	ACR TIRADS	96 (45.3%)	100.0% (63.1–100.0%)	47.1% (40.0–54.1%)	6.9% (3.0–13.1%)	100.0% (96.2–100.0%)	0.73 (0.70–0.77)
	AACE/ACE/AME	61# (28.8%)	100.0% (63.1–100.0%)	29.9% (23.7–36.7%)	5.3% (2.3–10.2%)	100.0% (94.1–100.0%)	0.65 (0.62–0.68)
	ATA ^a	46# (21.7%)	100.0% (63.1–100.0%)	22.5% (17.0–28.9%)	4.8% (2.1–9.3%)	100.0% (92.3–100.0%)	0.61 (0.58–0.64)
	EU-TIRADS	52# (24.5%)	100.0% (63.1–100.0%)	25.5% (19.7–32.0%)	5.0% (2.2–9.6%)	100.0% (93.1–100.0%)	0.63 (0.60–0.66)
	K-TIRADS	28# (13.2%)	100.0% (63.1–100.0%)	13.7% (9.3–19.2%)	4.3% (1.9–8.4%)	100.0% (87.7–100.0%)	0.57 (0.54–0.59)

Abbreviations: AUROC: area under the receiver operating characteristic curve; CI: confidence interval; NPV: negative predictive value; PPV: positive predictive value. a) nodules not classifiable with the ATA system were considered intermediate-suspicion nodules. b) comparison with ACR TIRADS in the > 65 age group, McNemar test, * *p* < 0.001; comparison with ACR TIRADS in the ≤ 65 age group, McNemar test, # *p* < 0.001. c) comparison with ACR TIRADS in the ≤ 65 age group, DeLong approach, ACR vs. ATA: *p* = 0.006; ACR vs. AACE: *p* = 0.05; ACR vs. K-TIRADS: *p* < 0.001; ACR vs. EU-TIRADS: *p* = 0.002; AACE vs. K-TIRADS: *p* = 0.04; comparison with ACR TIRADS in the > 65 age group, DeLong approach, ACR vs. all other systems: *p* < 0.001; AACE vs. K-TIRADS: *p* < 0.001; EU-TIRADS vs. K-TIRADS: *p* = 0.002.

3. Discussion

While the prevalence of thyroid nodules increases with increasing age, the malignancy rate is reported to be lower [16]; thus, the proper identification of the small number of lesions requiring clinical attention is of paramount importance in elderly patients. The chances of diagnosing asymptomatic thyroid nodules are increased by the frequent use of high-frequency ultrasound and cross-sectional imaging in routine clinical care [17]. However, while confirmed cancers in elderly patients are more likely to be aggressive [16], the risks associated with overtreatment of benign or low-risk malignant diseases should be carefully avoided in frail patients since the benefits are uncertain [18]. It is now clear that less aggressive treatment approaches are safe for low-risk thyroid malignancies [19,20], even if these are still relatively uncommon in real-world practice [21]. In elderly patients, an active surveillance approach may be used to defer or even definitively avoid surgery [22]. However, clinicians may be concerned by the potential occurrence of more aggressive tumors in older patients if a long-term follow-up protocol is adopted instead of immediate thyroid nodule biopsy.

In our cohort, we found that nodules submitted to biopsy in individuals > 65 years had more calcifications, even if the overall rate of malignancy was lower than in younger patients [16]. US-detected microcalcifications are associated with the presence of psammoma bodies [23] in papillary thyroid cancer. However, dystrophic or stromal calcifications and eosinophilic colloid may also appear as punctate hyperechogenic foci [24], similar to microcalcifications.

The distribution of risk categories was comparable between age groups in the different sonographic risk-stratification systems, with the exception of the EU-TIRADS and ATA guideline systems. Due to the higher rate of microcalcifications, the number of EU-TIRADS five and ATA non-classifiable nodules was higher in older patients. The ATA non-classifiable nodules are a significant proportion of the whole cohort and have a non-negligible malignancy rate, as previously reported by other authors [25]. This is due to the presence of key suspicious features in the context of isoechogenic nodules. For this reason, as suggested in the recent literature [25], non-classifiable nodules were counted as intermediate-suspicion nodules: in this way, the difference in ATA risk category distribution between age groups disappeared.

When analyzing the diagnostic performance of sonographic stratification systems, their discriminative ability was confirmed in people > 65 years, even if the low positive predictive values and specificities suggested the need to revise the definition or the relative weight of some features in the > 65 age group (microcalcifications seemed to be the most critical). These points should be taken into consideration when current guidelines are updated or in the development of new systems. These results were consistent with data reported by Di Fermo et al. [15], which supported the validity of sonographic stratification systems in elderly patients, even if the specificity of suspicious features, in this setting, was lower than expected. Conversely, in our cohort, there was no significant difference in diagnostic performance between the ATA and AACE/ACE/AME systems. The ACR TIRADS achieved the best discriminative performance. It is important to note that this system weighs microcalcifications and other punctate echogenic foci equally. Our results might be due to the low overall malignancy rate in our cohort. In settings with a higher pretest probability of malignancy (e.g., 18F-fluorodeoxyglucose positron emission-positive nodules), TIRADS with a higher propensity to indicate FNAB may be preferred [26].

This study had some limitations. First of all, the sample size might be limited. Most malignancies were confirmed by surgical histology, but false positives could not be excluded for patients with cytological diagnoses of malignancy who opted for conservative management. For cytologically-benign nodules, false negatives may occur, and a false negative rate of 3.7% has been reported [27]. Furthermore, we excluded subcentimeter nodules (9.2%) and lesions with inconclusive cytology (24.3%), although these exclusions did not alter the age distribution of our final cohort. The exclusion of indeterminate cytology nodules from the analysis might have reduced the amount of follicular thyroid cancers. However, scoring systems have also been found to correctly classify these cancers [28], mainly due to the suggestion to biopsy nodules greater than 20–25 mm, regardless of their sonographic pattern.

4. Materials and Methods

All patients consecutively referred to our center for FNAB of a thyroid nodule between November 1, 2015, and March 10, 2020, were included in the study. The study was conducted with institutional review board approval (Sapienza University Ethics Committee, study number 806/16) and written consent.

Patients were referred by our thyroid nodule clinic and by other specialists, including hospitalists, endocrinologists, nuclear medicine physicians, and surgeons, based on clinical risk factors, sonographic risk features, or patient preference.

Prior to FNAB, each nodule was examined with a HI-VISION Avius[®] system (Hitachi Medical Corporation, Inc., Tokyo, Japan) and a 13-MHz linear-array transducer. During this re-examination, two clinicians experienced in thyroid imaging recorded their joint evaluation of the sonographic features of each nodule on a standardized form. Full details on the enrollment criteria and procedures used for sonographic assessment, risk stratification, and FNAB examination of the nodules have previously been published [11,29,30]. We previously used a subset of this cohort in previous studies we conducted to compare the diagnostic performance of the systems, evaluate the impact of intrathyroidal location, and propose a better definition of the taller-than-wide shape, the results of which have already been reported [11,31,32]. In summary, all nodule sonographic features were collected, and the malignancy risk of each nodule was estimated automatically according to five sonographic risk-stratification systems by applying an algorithmic approach: the AACE/ACE/AME guidelines, the ACR TIRADS, the ATA guidelines, the EU-TIRADS, and the K-TIRADS. Nodules that could not be classified with the ATA guidelines were considered intermediate-suspicion nodules (i.e., iso or hyperechoic nodules with high-suspicion features, including irregular margins, microcalcifications, taller-than-wide shape, disrupted rim calcifications with a small extrusive hypoechoic soft tissue component, or evidence of extrathyroidal extension) [25]. Nodules with a maximum diameter of less than 1 cm were excluded from this study since none of the risk-stratification systems routinely recommend FNAB for subcentimeter thyroid nodules.

4.1. Reference Standard

Cytology was classified according to the criteria published in the Italian consensus for thyroid cytopathology [33,34], a six-tiered system comparable to the Bethesda System for Reporting Thyroid Cytopathology. If surgery had been performed, the reference standard diagnosis (malignant vs. benign) was based on histological examination of the resected nodule. If the nodule was not resected, a cytology-based reference standard was applied. Nodules were considered malignant if they were classified as TIR4 or TIR5 (corresponding to Bethesda classes V and VI), and benign if they were classified as TIR2, corresponding to Bethesda class II. Unresected nodules that were cytologically classified as TIR1 (non-diagnostic), 3A (low-risk indeterminate), or 3B (high-risk indeterminate) were excluded.

4.2. Age Groups

Patients were grouped according to their chronological age, a younger group (≤ 65 years), and an elderly group (> 65 years). It is the classical, conventional threshold, that we adopted, even if it is subject to changes based on comprehensive evidence in various aspects of social, cultural, and medical sciences [35].

4.3. Statistical Analysis

The nodules for which FNAB was indicated in each system were flagged as test positive. The sensitivity, specificity, positive and negative predictive values (PPV and NPV), and the AUROC, each with 95% confidence intervals, were computed for each system. Differences in categorical variables between groups were analyzed using the Chi-square test or the Fisher exact test.

The proportion of biopsies that would not have been indicated by the various systems were compared using the McNemar test. Data were analyzed with IBM SPSS Statistics, version 25.0 (IBM Corp., Armonk, NY, US). AUROC was compared with the DeLong approach [36] using the easyROC package [37].

5. Conclusions

In conclusion, when current risk-stratification systems were applied in clinical practice as rule-out tests for older patients, all were able to avoid a sizable number of biopsies, with a negative predictive value of 100%. Indeed, no malignancy was missed in any of the systems, though this result required that non-classifiable nodules in the ATA guidelines be considered intermediate-suspicion lesions. As previously reported in the general population, the ACR TIRADS outperformed the other systems as it avoided the highest number of biopsies and had the best discriminative power in the > 65 age group.

Author Contributions: Conceptualization, G.G. and C.D.; methodology, G.G. and C.D.; formal analysis, G.G.; investigation, G.G., R.F., V.R., M.M., and P.L.; data curation, G.G., R.F., V.R., M.M., and P.L.; writing—original draft preparation, G.G. and C.D.; writing—review and editing, all authors.; supervision, S.F. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Conflicts of Interest: The authors declare no conflict of interest.

References

- Gharib, H.; Papini, E.; Garber, J.R.; Duick, D.S.; Harrell, R.M.; Hegedüs, L.; Paschke, R.; Valcavi, R.; Vitti, P.; AACE/ACE/AME Task Force on Thyroid Nodules. American Association of Clinical Endocrinologists, American College of Endocrinology, and Associazione Medici Endocrinologi medical guidelines for clinical practice for the diagnosis and management of thyroid nodules 2016 update. *Endocr. Pr.* **2016**, *22*, 622–639. [[CrossRef](#)] [[PubMed](#)]
- Haugen, B.R.; Alexander, E.K.; Bible, K.C.; Doherty, G.M.; Mandel, S.J.; Nikiforov, Y.E.; Pacini, F.; Randolph, G.W.; Sawka, A.M.; Schlumberger, M.; et al. 2015 American Thyroid Association Management Guidelines for Adult Patients with Thyroid Nodules and Differentiated Thyroid Cancer: The American Thyroid Association Guidelines Task Force on Thyroid Nodules and Differentiated Thyroid Cancer. *Thyroid* **2016**, *26*, 1–133. [[CrossRef](#)] [[PubMed](#)]
- 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](#)] [[PubMed](#)]
- 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](#)]
- Shin, J.H.; Baek, J.H.; Chung, J.; Ha, E.J.; Kim, J.H.; Lee, Y.H.; Lim, H.K.; Moon, W.J.; Na, D.G.; Park, J.S.; et al. Ultrasonography Diagnosis and Imaging-Based Management of Thyroid Nodules: Revised Korean Society of Thyroid Radiology Consensus Statement and Recommendations. *Korean J. Radiol.* **2016**, *17*, 370–395. [[CrossRef](#)] [[PubMed](#)]
- Grani, G.; Sponziello, M.; Pecce, V.; Ramundo, V.; Durante, C. Contemporary Thyroid Nodule Evaluation and Management. *J. Clin. Endocrinol. Metab.* **2020**, *105*. [[CrossRef](#)] [[PubMed](#)]
- Hoang, J.K.; Middleton, W.D.; Farjat, A.E.; Langer, J.E.; Reading, C.C.; Teefey, S.A.; Abinanti, N.; Boschini, F.J.; Bronner, A.J.; Dahiya, N.; et al. Reduction in Thyroid Nodule Biopsies and Improved Accuracy with American College of Radiology Thyroid Imaging Reporting and Data System. *Radiology* **2018**, *287*, 185–193. [[CrossRef](#)]
- Ahmadi, S.; Oyekunle, T.; Jiang, X.; Scheri, R.; Perkins, J.; Stang, M.; Roman, S.; Sosa, J.A. A direct comparison of the ata and TI-Rads ultrasound scoring systems. *Endocr. Pr.* **2019**, *25*, 413–422. [[CrossRef](#)]
- Middleton, W.D.; Teefey, S.A.; Reading, C.C.; Langer, J.E.; Beland, M.D.; Szabunio, M.M.; Desser, T.S. Multiinstitutional Analysis of Thyroid Nodule Risk Stratification Using the American College of Radiology Thyroid Imaging Reporting and Data System. *AJR Am. J. Roentgenol.* **2017**, *208*, 1331–1341. [[CrossRef](#)]

10. Xu, T.; Wu, Y.; Wu, R.X.; Zhang, Y.Z.; Gu, J.Y.; Ye, X.H.; Tang, W.; Xu, S.H.; Liu, C.; Wu, X.H. Validation and comparison of three newly-released Thyroid Imaging Reporting and Data Systems for cancer risk determination. *Endocrine* **2019**, *64*, 299–307. [[CrossRef](#)]
11. Grani, G.; Lamartina, L.; Ascoli, V.; Bosco, D.; Biffoni, M.; Giacomelli, L.; Maranghi, M.; Falcone, R.; Ramundo, V.; Cantisani, V.; et al. Reducing the Number of Unnecessary Thyroid Biopsies While Improving Diagnostic Accuracy: Toward the “Right” TIRADS. *J. Clin. Endocrinol. Metab.* **2019**, *104*, 95–102. [[CrossRef](#)] [[PubMed](#)]
12. Basha, M.A.A.; Alnaggar, A.A.; Refaat, R.; El-Maghraby, A.M.; Refaat, M.M.; Abd Elhamed, M.E.; Abdalla, A.; Aly, S.A.; Hanafy, A.S.; Mohamed, A.E.M.; et al. The validity and reproducibility of the thyroid imaging reporting and data system (TI-RADS) in categorization of thyroid nodules: Multicentre prospective study. *Eur. J. Radiol.* **2019**, *117*, 184–192. [[CrossRef](#)] [[PubMed](#)]
13. Persichetti, A.; Di Stasio, E.; Guglielmi, R.; Bizzarri, G.; Taccogna, S.; Misischi, I.; Graziano, F.; Petrucci, L.; Bianchini, A.; Papini, E. Predictive Value of Malignancy of Thyroid Nodule Ultrasound Classification Systems: A Prospective Study. *J. Clin. Endocrinol. Metab.* **2018**, *103*, 1359–1368. [[CrossRef](#)] [[PubMed](#)]
14. Castellana, M.; Castellana, C.; Treglia, G.; Giorgino, F.; Giovanella, L.; Russ, G.; Trimboli, P. Performance of five ultrasound risk stratification systems in selecting thyroid nodules for FNA. A meta-analysis. *J. Clin. Endocrinol. Metab.* **2019**, *105*, 1659–1669. [[CrossRef](#)]
15. Di Fermo, F.; Sforza, N.; Rosmarin, M.; Morosan Allo, Y.; Parisi, C.; Santamaria, J.; Pacenza, N.; Zuk, C.; Faingold, C.; Ferraro, F.; et al. Comparison of different systems of ultrasound (US) risk stratification for malignancy in elderly patients with thyroid nodules. Real world experience. *Endocrine* **2020**, *69*, 331–338. [[CrossRef](#)] [[PubMed](#)]
16. Kwong, N.; Medici, M.; Angell, T.E.; Liu, X.; Marqusee, E.; Cibas, E.S.; Krane, J.F.; Barletta, J.A.; Kim, M.I.; Larsen, P.R.; et al. The Influence of Patient Age on Thyroid Nodule Formation, Multinodularity, and Thyroid Cancer Risk. *J. Clin. Endocrinol. Metab.* **2015**, *100*, 4434–4440. [[CrossRef](#)]
17. Russ, G.; Leboulleux, S.; Leenhardt, L.; Hegedüs, L. Thyroid incidentalomas: Epidemiology, risk stratification with ultrasound and workup. *Eur. Thyroid J.* **2014**, *3*, 154–163. [[CrossRef](#)]
18. Wang, Z.; Vyas, C.M.; Van Benschoten, O.; Nehs, M.A.; Moore, F.D.; Marqusee, E.; Krane, J.F.; Kim, M.I.; Heller, H.T.; Gawande, A.A.; et al. Quantitative Analysis of the Benefits and Risk of Thyroid Nodule Evaluation in Patients ≥ 70 Years Old. *Thyroid* **2018**, *28*, 465–471. [[CrossRef](#)]
19. Ramundo, V.; Sponziello, M.; Falcone, R.; Verrienti, A.; Filetti, S.; Durante, C.; Grani, G. Low-risk papillary thyroid microcarcinoma: Optimal management toward a more conservative approach. *J. Surg. Oncol.* **2020**, *121*, 958–963. [[CrossRef](#)]
20. Lamartina, L.; Grani, G.; Durante, C.; Filetti, S. Recent advances in managing differentiated thyroid cancer. *F1000Res* **2018**, *7*, 86. [[CrossRef](#)]
21. Holoubeck, S.A.; Yan, H.; Kuchta, K.M.; Winchester, D.J.; Moo-Young, T.A.; Prinz, R.A. Trends in nonoperative management of papillary thyroid microcarcinoma. *J. Surg. Oncol.* **2020**. [[CrossRef](#)]
22. Ito, Y.; Miyauchi, A.; Kihara, M.; Higashiyama, T.; Kobayashi, K.; Miya, A. Patient age is significantly related to the progression of papillary microcarcinoma of the thyroid under observation. *Thyroid* **2014**, *24*, 27–34. [[CrossRef](#)] [[PubMed](#)]
23. Kim, B.K.; Lee, E.M.; Kim, J.H.; Oak, S.Y.; Kwon, S.K.; Choi, Y.S.; Kim, Y.O. Relationship between ultrasonographic and pathologic calcification patterns in papillary thyroid cancer. *Medicine* **2018**, *97*, e12675. [[CrossRef](#)] [[PubMed](#)]
24. Tahvildari, A.M.; Pan, L.; Kong, C.S.; Desser, T. Sonographic-Pathologic Correlation for Punctate Echogenic Reflectors in Papillary Thyroid Carcinoma: What Are They? *J. Ultrasound Med.* **2016**, *35*, 1645–1652. [[CrossRef](#)] [[PubMed](#)]
25. Rosario, P.W.; da Silva, A.L.; Nunes, M.S.; Ribeiro Borges, M.A.; Mourão, G.F.; Calsolari, M.R. Risk of malignancy in 1502 solid thyroid nodules >1 cm using the new ultrasonographic classification of the American Thyroid Association. *Endocrine* **2017**, *56*, 442–445. [[CrossRef](#)] [[PubMed](#)]
26. Trimboli, P.; Knappe, L.; Treglia, G.; Ruberto, T.; Piccardo, A.; Ceriani, L.; Paone, G.; Giovanella, L. FNA indication according to ACR-TIRADS, EU-TIRADS and K-TIRADS in thyroid incidentalomas at 18F-FDG PET/CT. *J. Endocrinol. Investig.* **2020**. [[CrossRef](#)]
27. Bongiovanni, M.; Spitalo, A.; Faquin, W.C.; Mazzucchelli, L.; Baloch, Z.W. The Bethesda System for Reporting Thyroid Cytopathology: A meta-analysis. *Acta Cytol.* **2012**, *56*, 333–339. [[CrossRef](#)]

28. Castellana, M.; Piccardo, A.; Virili, C.; Scappaticcio, L.; Grani, G.; Durante, C.; Giovanella, L.; Trimboli, P. Can ultrasound systems for risk stratification of thyroid nodules identify follicular carcinoma? *Cancer Cytopathol.* **2020**, *128*, 250–259. [[CrossRef](#)]
29. Grani, G.; Lamartina, L.; Ascoli, V.; Bosco, D.; Nardi, F.; D'Ambrosio, F.; Rubini, A.; Giacomelli, L.; Biffoni, M.; Filetti, S.; et al. Ultrasonography scoring systems can rule out malignancy in cytologically indeterminate thyroid nodules. *Endocrine* **2017**, *57*, 256–261. [[CrossRef](#)]
30. Grani, G.; Lamartina, L.; Biffoni, M.; Giacomelli, L.; Maranghi, M.; Falcone, R.; Ramundo, V.; Cantisani, V.; Filetti, S.; Durante, C. Sonographically Estimated Risks of Malignancy for Thyroid Nodules Computed with Five Standard Classification Systems: Changes over Time and Their Relation to Malignancy. *Thyroid* **2018**, *28*, 1190–1197. [[CrossRef](#)]
31. Ramundo, V.; Lamartina, L.; Falcone, R.; Ciotti, L.; Lomonaco, C.; Biffoni, M.; Giacomelli, L.; Maranghi, M.; Durante, C.; Grani, G. Is thyroid nodule location associated with malignancy risk? *Ultrasonography* **2019**, *38*, 231–235. [[CrossRef](#)]
32. Grani, G.; Lamartina, L.; Ramundo, V.; Falcone, R.; Lomonaco, C.; Ciotti, L.; Barone, M.; Maranghi, M.; Cantisani, V.; Filetti, S.; et al. Taller-Than-Wide Shape: A New Definition Improves the Specificity of TIRADS Systems. *Eur. Thyroid J.* **2020**, *9*, 85–91. [[CrossRef](#)] [[PubMed](#)]
33. Nardi, F.; Basolo, F.; Crescenzi, A.; Fadda, G.; Frasoldati, A.; Orlandi, F.; Palombini, L.; Papini, E.; Zini, M.; Pontecorvi, A.; et al. Italian consensus for the classification and reporting of thyroid cytology. *J. Endocrinol. Investig.* **2014**, *37*, 593–599. [[CrossRef](#)] [[PubMed](#)]
34. Fadda, G.; Basolo, F.; Bondi, A.; Bussolati, G.; Crescenzi, A.; Nappi, O.; Nardi, F.; Papotti, M.; Taddei, G.; Palombini, L. Cytological classification of thyroid nodules. Proposal of the SIAPEC-IAP Italian Consensus Working Group. *Pathologica* **2010**, *102*, 405–408. [[PubMed](#)]
35. Orimo, H.; Ito, H.; Suzuki, T.; Araki, A.; Hosoi, T.; Sawabe, M. Reviewing the definition of “elderly”. *Geriatr. Gerontol. Int.* **2006**, *6*, 149–158. [[CrossRef](#)]
36. DeLong, E.R.; DeLong, D.M.; Clarke-Pearson, D.L. Comparing the Areas under Two or More Correlated Receiver Operating Characteristic Curves: A Nonparametric Approach. *Biometrics* **1988**, *44*, 837. [[CrossRef](#)] [[PubMed](#)]
37. Goksuluk, D.; Korkmaz, S.; Zararsiz, G.; Karaagaoglu, A.E. easyROC: An Interactive Web-tool for ROC Curve Analysis Using R Language Environment. *R. J.* **2016**, *8*, 213–230. [[CrossRef](#)]



© 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).