

Unnecessary fine needle aspiration cytological biopsy of thyroid nodules according to two ultrasound classification systems

Biopsia citológica por aspiración con aguja fina innecesaria de nódulos tiroideos según dos sistemas de clasificación ecográfica

SM Batallés¹, JL Novelli¹, L Caciarelli¹, I Baumlis¹, C Argutti¹, R Asenjo¹, C Blanco¹, P Goldstein¹, L Martinessi¹, S Ahn¹

¹Universidad Abierta Interamericana. Rosario, Argentina.

Submitted: 06-01-2023

Revised: 13-04-2023

Accepted: 09-07-2024

Published: 10-07-2024

How to Cite: Batallés SM, Novelli JL, Caciarelli L, Baumlis I, Argutti C, Asenjo R, et al. Unnecessary fine needle aspiration cytological biopsy of thyroid nodules according to two ultrasound classification systems. Interamerican Journal of Health Sciences. 2023; 3:18. <https://doi.org/10.59471/ijhsc202318>

ABSTRACT

Objective: to identify cytological FNA of thyroid nodules that would be unnecessary to perform according to US classification systems applied in our environment (ATA and ACR TI-RADS).

Method: n=346 thyroid nodules were US and cytological evaluated in 293 patients (238 women and 55 men). Mean age without statistical significance between sexes (48,4/47,6 years old for males/women).

Results: when the nodule characteristics are suspicious, both systems were similar in suggesting an FNA. But to discern which nodules not require FNA, ACR TI-RADS was significantly better than ATA classification, in terms of lower amount of Positive False and Specificity ($P<0,05$). Complications: 5/293 (1,7 %) superficial hematomas (early complications); no late complications were observed. Why were there FNA in nodules that had FNA even when both classifications suggested not to do so? doctor's request (43,4 %) and/or patient preference (32,4 %), nodules in a context of a chronic thyroiditis parenchyma (12,4 %) and purely cystic nodules >35 mm, to evacuate the liquid content (11,8 %).

Conclusion: nodules cytologically identified as suspicious was low. 94 % of the FNA performed only required clinical follow-up. ACR TI-RADS classification induces fewer unnecessary FNA than ATA classification. Both classifications are wrong in equal proportion in suggesting no FNA suspicious nodules.

KEYWORDS

Thyroid, Nodule, Cytology, Needle, Aspiration.

RESUMEN

Objetivo: identificar las PAAF citológicas de nódulos tiroideos que sería innecesario realizar según los sistemas de clasificación US aplicados en nuestro medio (ATA y ACR TI-RADS).

Método: n=346 nódulos tiroideos fueron evaluados US y citológicamente en 293 pacientes (238 mujeres y 55 hombres). Edad media sin significación estadística entre sexos (48,4/47,6 años varones/mujeres).

Resultados: cuando las características del nódulo son sospechosas, ambos sistemas fueron similares en sugerir una PAAF. Pero para discernir qué nódulos no requieren FNA, ACR TI-RADS fue significativamente mejor que la clasificación ATA, en términos de menor cantidad de Falso Positivo y Especificidad ($P<0,05$). Complicaciones:

5/293 (1,7 %) hematomas superficiales (complicaciones precoces); no se observaron complicaciones tardías. ¿Por qué se realizaron PAAF en nódulos en los que se practicó incluso cuando ambas clasificaciones sugerían no hacerlo? A petición del médico (43,4 %) y/o preferencia del paciente (32,4 %), nódulos en un contexto de parénquima de tiroiditis crónica (12,4 %) y nódulos puramente quísticos >35 mm, para evacuar el contenido líquido (11,8 %).

Conclusiones: los nódulos identificados citológicamente como sospechosos fueron escasos. El 94 % de las PAAF realizadas sólo requirieron seguimiento clínico. La clasificación ACR TI-RADS induce menos PAAF innecesarias que la clasificación ATA. Ambas clasificaciones se equivocan en igual proporción al sugerir la no realización de BAAF de nódulos sospechosos.

PALABRAS CLAVE

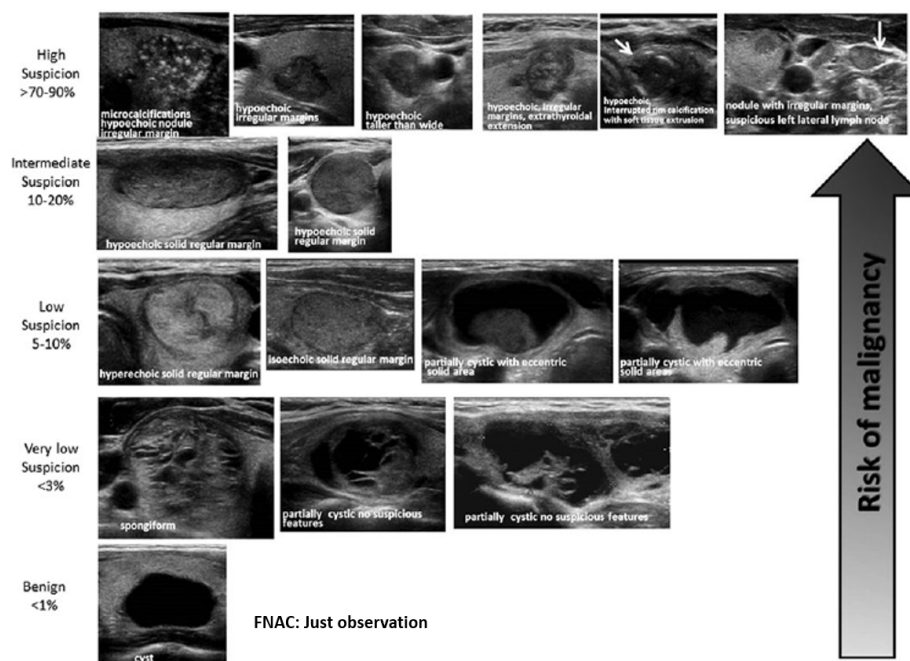
Tiroides, Nódulo, Citología, Aguja, Aspiración.

INTRODUCTION

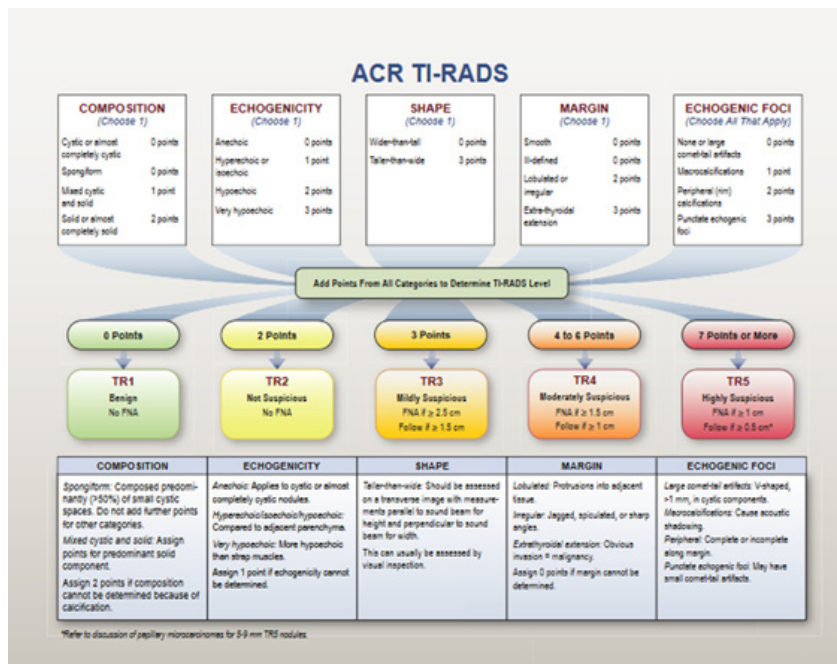
Thyroid nodules are common; large-scale studies have reported a prevalence of palpable nodules of 5 % in women and 1 % in men in iodine-sufficient areas.⁽¹⁾ Ultrasound (US) findings raise this prevalence to 68 %.^(2,3)

Although the prevalence of the thyroid nodule is high, the prevalence of cancer in them is about 5 % of the nodules. However, when patients are selected based on US characteristics, the prevalence of malignancy rises to close to 13 %.^(1,2,3)

When a thyroid nodule shows US characteristics suspected of malignancy, guidelines agree that the diagnosis should be completed with a fine needle aspiration cytological biopsy (FNA) since the sum of both diagnostic methods (US + FNA) increases the certainty of the nature of nodule and the need for surgery. But when US shows features considered benign, the indication for FNA is questionable and should be justified. To identify suspected cases of malignancy based on US characteristics, risk stratifications have been proposed. The two most widely used are: the guidelines of the American Thyroid Association (ATA), in its version of the year 2015, which stratified the risks of malignancy of the nodules according to US patterns (figure 1)⁽⁷⁾ and the classification proposed by the American College of Radiology (ACR TI-RADS) in 2017, based on a score achieved by the different US characteristics of nodule (figure 2).⁽⁸⁾ Both classification systems suggest in which nodules a FNA is recommended.



Source: Haugen BR, Alexander EK, Bible KC et al. Thyroid. 26(1):1-133, 2016
Figure 1. US characteristics of thyroid nodule and malignancy risk proposed by ATA (2015)



Source: Tessler FN, Middleton WD, Grant EG. Et al. J Am Coll Radiol 2017; 14:587-595

Figure 2. Chart showing five categories on the basis of the ACR Thyroid Imaging, Reporting and Data System (TI-RADS) lexicon, TR levels, and criteria for fine-needle aspiration or follow-up ultrasound. Explanatory notes appear at the bottom

Worldwide, there are other US classifications like ATA and ACR TI-RADS, dedicated to establishing malignancy risk and suggestion of FNA^(1,2,3,4,5,6,7,8) but in our environment, the most used are those included in this analysis. FNAs are interpreted and classified with the Bethesda system (figure 3).⁽⁹⁾ In that classification, Bethesda II nodules are compatible with benignity (<3 % of malignancy risk) and clinical follow-up of them is suggested.

Bethesda	Diagnostic Category	Risk of malignancy Usual management
I	Nondiagnostic Unsatisfactory cyst fluid only virtually acellular specimen other (obscuring blood, clotting artifact, etc.)	- Repeat FNAC with ultrasound
II	Benign consistent with a benign follicular nodule (includes adenomatoid nodule, colloid nodule, etc.) consistent with lymphocytic (Hashimoto) thyroiditis in the proper clinical context consistent with granulomatous (subacute) thyroiditis	0 – 3% Clinical follow-up
IIIa	Atypia of undetermined significance (AUS)	~5 – 15% Repeat FNAC with ultrasound
IIIb	Follicular lesion of undetermined significance (FLUS)	
IV	Follicular neoplasm/"suspicious" for follicular neoplasm Specify in Hürthle cell type	15 – 30% Surgical lobectomy
V	Suspicious for malignancy suspicious for papillary carcinoma suspicious for medullary carcinoma suspicious for metastatic carcinoma suspicious for lymphoma	60 – 75% Near-total thyroidectomy or surgical lobectomy
VI	Malignant papillary thyroid carcinoma poorly differentiated carcinoma medullary thyroid carcinoma undifferentiated (anaplastic) carcinoma squamous cell carcinoma carcinoma with mixed features metastatic	97 – 99% Near-total thyroidectomy

Source: Baloch ZW, Alexander EK, Gharib H y col. Chapter 1. En: SZ Ali y ES Cibas (eds) The Bethesda System for Reporting Thyroid Cytopathology. New York, NY: Springer, pp:1–4, 2010

Figure 3. Cytological classification of Bethesda System, malignancy risk and therapeutic management suggestions

Nodules classified with Bethesda III have indeterminate behavior and, in all cases, repeat FNA is suggested because within Bethesda category two levels are distinguished: III-FLUS (indeterminate follicular lesion) -which have a malignancy risk similar to Bethesda II level- and III-AUS (atypia of indeterminate meaning) -which resemble the risk of malignancy of the Bethesda IV (suspicion of follicular neoplasia). Bethesda IV has a malignancy risk of 15-30 % and where it is suggested to perform lobectomy of the nodule. Bethesda V (suspected malignancy) and VI (malignant) levels cast no doubt on the nodule cytology; their malignancy risk exceeds 60 % and surgical resection is the indicated therapeutic option (lobectomy or thyroidectomy less than total and total thyroidectomy or less than total, respectively).

Which nodules are punctured unnecessarily, according to the guidelines? In theory, those whose US characteristics suggest benignity and, therefore, there would be no need to perform an FNA, since these nodules would not require surgery, but would only be subject to evolutionary control over time.

It is hypothesized that many professionals prescribe FNA in nodules despite not being indicated by the guidelines. If FNA were performed even without being indicated by the guidelines and were reported as benign, those indications are correct, they should not be performed and those are unnecessary FNA. If the guidelines indicated that they should not be performed a FNA and for some reason, were done and biopsied informed as malignant, then it could be the mistake of the guidelines.

The general objective of this study was to identify cytological FNA of thyroid nodules that would be unnecessary to perform according to US classification systems applied in our environment (ATA and ACR TI-RADS). As specific objectives, the motives of FNA indication contrarian to those and to determine which of the two classifications was more effective in their suggestions settled.

METHOD

A prospective observational study was conducted. Between January 2020 and January 2022, US information of with one or more thyroid nodules was prospectively collected from patients who were referred to the US Service to perform a FNA under US guidance. The practice prescription was indicated by the patient's endocrinologist/clinician/primary care surgeon.

Patients of both sexes, older than 18 years, with FNA indication of one or more thyroid nodules were included. Patients with previous thyroid surgeries, a personal history of neck radiation and a family history of papillary thyroid carcinoma were excluded from the study.

All patients were evaluated with thyroid color-doppler US and the nodules were categorized with two classification systems used in our environment (ATA and ACR TI-RADS). All nodules were punctured and reported with the Bethesda classification system; the presence of the pathologist at the time of the FNA ensured the obtaining of sufficient material for cytological diagnosis and preservation of it until the definitive diagnosis.

To reduce the risk of loss information, the study was limited to US and FNA images under US guidance performed at the US and Cytological FNA Service of the Thyroid Unit at the Oroño Group, Rosario (Argentina). The two US classifications and the extraction of FNA material were performed by the same specialist operator (SMB in diagnostic by imaging and OB in pathology), thus avoiding the bias of inter-observer variability as in other published works.

It was originally planned to collect approximately 400 nodules in the aforementioned period. The effects of mandatory emergence confinement due to a new strain of severe acute respiratory syndrome virus type-2 (SARS-CoV-2) which caused COVID-19 infection, caused many patients delayed medical consultations and studies. The final sample included 346 nodules.

For each punctured nodule, information was recorded on the patient's age and sex, the context (nodular/multinodular goiter), US evidence of chronic thyroiditis, the respective US variables included in the two classification systems (ATA and ACR TI-RADS) and the Bethesda cytological level. We considered an US pattern of chronic thyroiditis when a heterogeneous thyroid parenchyma accompanied or not by rounded or oval pseudonodular focal areas, with normal, increased or decreased intralobular vascularization were identified. In addition, in those patients classified as low or very low malignancy suspicion according ATA and ACR TI-RADS, the reasons that indicated the FNA were consigned.

For the purposes of this study, in accordance with ATA and ACR TI-RADS systems that recommend "FNA suggested" or "FNA not suggested" according to the different levels of malignancy, the nodules classified with Bethesda II and III-FLUS system were categorized together, in the recommendation "surgery not required" and Bethesda III-FLUS, IV, V and VI as "surgery required".

The project obtained the approval of the Ethic Committee of the Universidad Abierta Interamericana (UAI), Argentina, for its implementation. All patients included signed an informed consent before performing the FNA. At the end of the procedure, all patients assessed perceived pain by completing a Verbal Numerical Scale (VNS), choosing a number between 0 and 10 reflecting the level of pain perceived during the FNA, where 0 represented

“no pain” and 10 “excruciating pain”. A record of early and late complications of FNA was also kept. The statistical analysis included the respective individual evaluation of the successes and errors of ATA and ACR TI-RADS US classifications with respect to the Bethesda system, analyzing them in terms of Sensitivity, Specificity, False (+) and (-) and concordance. Subsequently, those values were compared between both methods, evaluating the statistical significance with the chi-square or Fisher test, when appropriate; in all cases, a value of $P < 0,05$ was considered statistically significant. All statistical analysis was performed with IBM-SPSS®V.21 software.

RESULTS

Total patients

N=346 thyroid nodules were US and cytological evaluated in 293 patients (238 women and 55 men). The mean age of men (48,4 years) was slightly higher than that of women (47,6 years), without statistical significance ($P=0,633$ from t-Student test). The nodule diameters (axial, cephalocaudal and anteroposterior) were evaluated; the largest of them was used to determine the nodular size. Thus, 346 FNA in thyroid nodules ranged between 6 and 60 mm, averaging 17,4 mm were performed (FNA was not performed in nodules ≤ 5 mm). US and cytological characteristics of nodules evaluated are shown in figure 4.

Characteristics		n= 346 nodules	Characteristics		n= 346 nodules
Age (years)	Mean	47.7	ATA Classification		
	Median (range)	47.5 (16-78)	High suspicion	21 (6%)	
Sex	Female	238 (68.8%)	Intermediate	46 (13.3%)	
	Male	55 (31.2%)	Low suspicion	180 (52%)	
In multinodular context		273 (78.9%)	Very low suspicion	89 (25.7%)	
Nodule site	Right lobe	178 (51.4%)	ACR TI-RADS classification		
	Left lobe	168 (48.6%)	1	12 (3.5%)	
Chronic thyroiditis		117 (33.8%)	2	98 (28.3%)	
Nodule diameter (mm)	<10	27 (7.8%)	3	122 (35.3%)	
	10-14	91 (26.3%)	4	97 (28.0%)	
	15-19	97 (28%)	5	17 (4.9%)	
	≥ 20	131 (37.9%)	Bethesda System		
			II	289 (83.5%)	
			III-FLUS	36 (10.4%)	
			III-AUS	8 (2.3%)	
			IV	2 (0.6%)	
			V-VI	11 (3.2%)	

Figure 4. Characteristics of evaluated thyroid nodules

Relationship between ATA classification and Bethesda System

The relationship of the indication of FNA between ATA classification and Bethesda system is shown in figure 5.

ATA Malignancy suspicion	Suggestion according nodule diameter	Bethesda System					Total
		II	III		IV	V-VI	
			FLUS	AUS			
High	<10 mm No FNA	6	-	1	-	1	8
	≥ 10 mm	12	5	2	-	4	23
Intermediate	<10 mm No FNA	3	-	-	-	1	4
	≥ 10 mm	31	7	1	1	2	42
Low	<15 mm No FNA	60	3	1	-	-	64
	≥ 15 mm	98	13	3	-	2	116
Very low	<20 mm No FNA	52	3	-	-	1	56
	≥ 20 mm	27	5	-	1	-	33
Total		289	36	8	2	11	346

Figure 5. Relationship between malignancy suspicion ATA classification and Bethesda system

ATA classification indicated FNA in 214/346 nodules (61,8 %); 6/214 (2,8 %) were Bethesda III-AUS, 2/214 (0,9 %) were Bethesda IV and 8/214 (3,7 %) were Bethesda V-VI, ratifying the suggestion of FNA in 7,4 % of these cases. However, among nodules that ATA indicated FNA, 168 were Bethesda II and another 30 were Bethesda III-FLUS, adding 198/214 (92,5 %) of diagnostic error suggesting FNA for those nodules. Moreover, according to ATA classification, only nodules ≥ 20 mm of very low risk should perform an FNA. According figure 5, only 1/33 (3 %) was Bethesda IV; and those that should not perform FNA, 1/56 (1,7 %) was a carcinoma (V-VI). According to the grouping proposed in Material and Methods, figure 6 relates the suggestion of FNA according to ATA classification and the indication of thyroid surgery according to Bethesda system.

		Bethesda System		Total	Diagnostic performance
		Surgery required	Surgery not required		
ATA Classification	FNA suggested	16	198	214	Sensibility: 16/21= 76.2% Specificity: 127/325= 39.1% Positive False: 198/325= 60.9% Negative False: 5/21= 23.8% Positive Predicted Value: 16/214= 7.5%* Negative Predicted Value: 127/132= 96.2%* Observed Concordance: (16+127) / 346= 41.3% Prevalence of nodules that required surgery: 21/346= 6%
	FNA not suggested	5	127	132	
	Total	21	325	346	

Figure 6. Diagnostic performance of ATA through FNA suggestion to identify nodules that surgery is required according to Bethesda system

Note: although predictive values were estimated, they were not considered due to the low prevalence of thyroid nodules that surgery required

Based on a low prevalence of nodules that would require surgery (6 %), ATA classification showed a good ability to detect nodules that require FNA (76 %), but a poor ability to identify those that do not (39 %), causing a high amount of Positive False.

Responding to the objectives of this work, using the ATA classification to suggest FNA, the following were settled:

- Unnecessary FNA: 61 % of nodules where surgery was not required.
- FNA performed despite not being indicated: 24 % of nodules that surgery was required.
- Concordance diagnostic between ATA and Bethesda: 41 % of nodules evaluated.

Relationship between ACR TI-RADS classification and Bethesda System

The relationship of the indication of FNA between ACR TI-RADS classification and Bethesda system is shown in figure 7.

ACR TI-RADS Malignancy suspicion	Suggestion according nodule diameter	Bethesda System					Total
		II	III		IV	V-VI	
			FLUS	AUS			
5	<10 mm No FNA	2	-	1	-	-	3
	≥ 10 mm	4	1	1	-	3	9
4	<15 mm No FNA	30	2	1	-	4	37
	≥ 15 mm	44	10	2	2	3	61
3	<25 mm No FNA	82	6	3	-	-	91
	≥ 25 mm	23	8	-	-	-	31
2	No FNA	87	9	-	-	1	97
1	No FNA	17	-	-	-	-	17
Total		289	36	8	2	11	346

Figure 7. Relationship between malignancy suspicion ACR TI-RADS classification and Bethesda system

ACR TI-RADS classification suggested that FNA 245/346 nodules (70,8 %) should not performed, because 218/245 (88,9 %) were Bethesda II and 17/245 (6,9 %) were III-FLUS, ratifying that FNA was not necessary in 95,9 % of those nodules. However, among those that ACR TI-RADS indicated no FNA, 5 nodules with Bethesda III-AUS and another 5 nodules with Bethesda V-VI were found, adding up to a total of 10/245 (4 %) of diagnostic error suggesting not FNA for those nodules.

ACR TI-RADS classification indicated FNA in 101/346 nodules (29,2 %); 3/101 (2,9 %) were Bethesda III-AUS, 2/101 (1,9 %) were Bethesda IV and 6/101 (5,9 %) were Bethesda V-VI, ratifying the suggestion of FNA in almost 11 % of those cases. However, among nodules that ACR TI-RADS suggested FNA, 71 were Bethesda II and 19 were III-FLUS, adding up 90/101 (89,1 %) of diagnostic error suggesting FNA for those nodules.

According to the grouping proposed in Material and Methods, figure 8 relates the suggestion of FNA according to ACR TI-RADS classification and the indication of thyroid surgery according to Bethesda system.

		Bethesda System		Total	Diagnostic performance
		Surgery required	Surgery not required		
ACR TI-RADS Classification	FNA suggested	11	90	101	Sensibility: 11/21= 52,3% Specificity: 235/325= 72,3% Positive False: 90/325= 27,7% Negative False: 10/21= 47,7% Positive Predicted Value: 11/101= 10,9%* Negative Predicted Value: 235/245= 95,9%* Observed Concordance: (11+235) / 346= 71,1% Prevalence of nodules that required surgery: 21/346= 6%
	FNA not suggested	10	235	245	
	Total	21	325	346	

Figure 8. Diagnostic performance of ACR TI-RADS through FNA suggestion to identify nodules that surgery is required according to Bethesda system

Note: although predictive values were estimated, they were not considered due to the low prevalence of thyroid nodules that surgery required

Based on a low prevalence of nodules that would require surgery (6 %), ATA classification showed a regular ability to detect nodules that require FNA (52,3 %), but an appropriate ability to identify those that do not (72,3 %), causing a low amount of Positive False but twice of Negative False.

Responding to the objectives of this work, using the ATA classification to suggest FNA, the following were settled:

- Unnecessary FNA: 28 % of nodules where surgery was not required.
- FNA performed despite not being indicated: 48 % of nodules that surgery was required.
- Concordance diagnostic between ATA and Bethesda: 71 % of nodules evaluated.

Comparison of diagnostic performance estimations

Diagnostic performance estimators	ATA classification	ACR TI-RADS classification	P*
Sensibility	16/21 (76.2%)	11/21 (52.3%)	0.11 (NS)
Specificity	127/325 (39.1%)	235/325 (72.3%)	0.0001
Positive False	198/325 (60.9%)	90/325 (27.7%)	0.0001
Negative False	5/21 (23.8%)	10/21 (47.7%)	0.11 (NS)
Concordance	143/346 (41.3%)	246/346 (71.1%)	0.0001

Figure 9. Comparison of diagnostic performance estimators between ATA and ACR TI-RADS classification to suggest which nodules requires an FNA

Note: Chi-squared test and NS: without statistical significance

The ability of ATA and ACR TI-RADS classifications to determine which of the two systems was the best one to suggest which nodules required FNA and which should not be exposed in figure 9.

The diagnostic success rate was significantly higher in ACR TI-RADS than in ATA classification. When the nodule characteristics are suspicious, both systems were similar in suggesting an FNA. But to discern which nodules not require FNA, ACR TI-RADS was significantly better than ATA classification, in terms of lower amount of Positive False and Specificity.

Regarding the complications secondary to perform FNA, only 5/293 (1,7 %) patients presented superficial hematomas (early complications) that did not required special interventions, but they were treated with analgesics and anti-inflammatories, resolving completely between 3 to 5 days. No late complications were observed.

The results of the VNS evaluating the perceived pain kept infer that FNA was a well-tolerated procedure, since 66,2 % of patients (n= 194) perceived the pain as “mild”, 32,4 % (n = 95) considered it “moderate” and only 1,4 % (n = 4) qualified such pain as “intense”.

Why were there FNA in nodules that were punctured even when the classifications suggested not to do so (low or very low suspicion according to ATA and level II according to ACR)? The choice had to do with the preference of the requesting physician (43,4 %), patient preference (insecurity/fear because the existence of the nodule) (32,4 %), nodules in a context of a chronic thyroiditis parenchyma (12,4 %) and purely cystic nodules >35 mm in diameter greater in which FNA was performed in order to evacuate the liquid content (11,8 %).

DISCUSSION

The technological advancement of recent years in diagnostic imaging is staggering. In thyroid pathology, the evolution in US scanners has made it possible to detect nodules with characteristics and/or size that before, with less powerful equipment, were not detected. Thyroid nodules are extremely common to diagnose; reports that up to 68 % of adults have nodules, leading to problems with overdiagnosis and overtreatment.^(2,11)

Several US classification systems have been developed to help professionals to identify in which thyroid nodules a FNA should be suggested; those classifications are based on US characteristics to stratify the malignancy risk.^(7,9,12,11) Although there are many US guidelines to recommend FNA in a thyroid nodule, there is currently no single, collective classification system that applies worldwide.

Although the operator-dependent factor must be considered in those systems, in this study it was avoided, since all thyroid US were performed and classified with two systems (ATA and ACR TI-RADS) by the same specialist in diagnostic imaging (SMB); the same happened with the pathologist-cytologist who analyzed and reported the Bethesda system (OB).

To reach the objective, this study compared the ability of two US classification systems to appropriately select nodules that should require a FNA. A total of 346 nodules were evaluated by the two systems. Importantly, 94 % (325 nodules) were classified with Bethesda II and III-FLUS, which was interpreted as benign cytology and, therefore, FNA in such nodules would be unnecessary. Only 6 % (21 nodules) required FNA because they were suspicious or malignant nodules (Bethesa III-AUS and IV-VI).

Many studies showed that ACR TI-RADS had better diagnosis compared to other guidelines and avoiding, with their suggestions, a higher number of unnecessary FNA.^(9,11,12,13,14) In terms of unnecessary FNA -Positive False in technical terms-, in the study of Wu et al.⁽¹⁰⁾ compared and discussed the reasons for the differences in diagnosis between ACR TI-RADS and ATA 2015. The rates of Positive False (34 %) and concordance with the FNA (75,5 %) for ACR TI-RADS, leded to conclusion that ACR TI-RADS suggested less unnecessary FNA than ATA classification (P = 0,007).

Yoon et al.⁽¹³⁾ and Xu et al.⁽¹⁴⁾ evaluated the similarities and differences between three TI-RADS systems: the Korean Society of Radiology, ACR and the European Thyroid Association, comparing the diagnostic performance of US with the results of cytology in the detection of malignant nodules. Both studies concluded that, although the three systems have similarities in US characteristics, they did find significant differences in diagnosing nodules suspected of malignancy: they demonstrated that ACR TI-RADS, compared to other TI-RADS systems of US classification, had the lowest rate of unnecessary FNA (28 % ACR vs 52,7 % -Korean- and 66,3 % -European-). Middleton et al.⁽⁹⁾ reported that Positive False rates were 78,1 % according to ATA classification and 47,1 % according to ACR TI-RADS. The authors concluded that those results significantly kept reduce the number of unnecessary FNA in benign nodules. In our study, considering ATA system, 198/325 (60,9 %) unnecessary FNA were performed reducing the rate to 90/325 (27,7 %) with ACR TI-RADS system. TIRADS ACR system significantly reduced more than twice as much unnecessary FNA as ATA would suggest (P~0).

Although the objective of this work was to identify unnecessary FNA, both US classification systems have also been wrong to suggest not FNA that were later suspicious or malignant -Negative False in technical terms-. In this study, 21 nodules were included in this group. The Negative False rate was 24 % for ATA and 48 % for ACR TI-RADS systems, without statistical significance (P=0,11).

Other recent studies also evaluated the concordance of three TI-RADS systems: ACR, European Thyroid Association and that from of Siriraj Hospital (Bangkok)^(15,16,17,18,19) highlighting both the optimal inter-observer agreement and the measures of sensitivity and negative predictive value for the diagnosis of malignant thyroid nodules. From our point of view, we do not consider predictive values to evaluate diagnostic capability because those are calculated on a total of FNA results (as has been demonstrated, many of them unnecessary) and not based on the result of cytology.^(20,21,22)

A heterogeneous parenchyma (e.g., chronic thyroiditis) could be considered a “confounding variable”. Many times, a nodule is categorized as such when in fact it is a pseudonodule or focal area of chronic thyroiditis. Despite this, it is reported that some endocrinologists/surgeons/clinicians prefer to indicate FNA in nodules present in the context of chronic thyroiditis due to the probable increased malignancy risk of nodules associated with this condition⁽¹¹⁾ especially in areas with high prevalence of chronic thyroiditis.⁽¹²⁾ It is also necessary to consider the patient’s preference (insecurity / fear of knowing the existence of the nodule) to indicate. In our studies, requesting physician and patient’s preference added the 75 % of motives of performed FNA.^(23,24)

Why do ACR TI-RADS suggestions mean making fewer unnecessary FNA than ATA’s? From our point of view, ATA classification system is based on US patterns to determine the malignancy risk; instead ACR TI-RADS is based on quite similar US characteristics that are measured in ATA but quantified through a numerical score (and not all characteristics are weighted equally). Evidence showed that the quantified ACR TI-RADS score better identifies those FNA that will be unnecessary to perform.^(25,26)

It could be thought as a potential limitation of this study to suppose that the specialist in charge of US classifying nodules (SMB) had a certain tendency to suggest FNA more nodules unnecessarily, in order not to leave without FNA those nodules that should do it so. It could also be thought of similarly with the cytology specialist (OB); after all, both ATA and Bethesda classifications are based on US and cytological patterns respectively, and they are interpretative, not quantifiable numerically.

CONCLUSIONS

- The rate of nodules cytologically identified as suspicious was low. The 94 % of the FNA performed only required clinical follow-up.
- ACR TI-RADS classification induces fewer unnecessary FNA than ATA classification.
- Both classifications are wrong in equal proportion in suggesting no FNA suspicious nodules.
- Reasons of requesting physician and/or patient’s preference to perform a FNA should be respected.

ACKNOWLEDGEMENTS

The authors thank Marta Alarcón for her collaboration in the statistical analyzes.

REFERENCES

1. Yassa, L., Cibas, E. S., Benson, C. B., Frates, M. C., Doubilet, P. M., Gawande, A. A., Moore, F. D., Jr, Kim, B. W., Nosé, V., Marqusee, E., Larsen, P. R., & Alexander, E. K. (2007). Long-term assessment of a multidisciplinary approach to thyroid nodule diagnostic evaluation. *Cancer*, 111(6), 508–516. <https://doi.org/10.1002/cncr.23116>
2. Guth, S., Theune, U., Aberle, J., Galach, A., & Bamberger, C. M. (2009). Very high prevalence of thyroid nodules detected by high frequency (13 MHz) ultrasound examination. *European journal of clinical investigation*, 39(8), 699–706. <https://doi.org/10.1111/j.1365-2362.2009.02162.x>
3. 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 (2016). 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. *Endocrine practice: official journal of the American College of Endocrinology and the American Association of Clinical Endocrinologists*, 22(5), 622–639. <https://doi.org/10.4158/EP161208.GL>
4. Ezzat, S., Sarti, D. A., Cain, D. R., & Braunstein, G. D. (1994). Thyroid incidentalomas. Prevalence by palpation and ultrasonography. *Archives of internal medicine*, 154(16), 1838–1840. <https://doi.org/10.1001/archinte.154.16.1838>
5. Papini, E., Guglielmi, R., Bianchini, A., Crescenzi, A., Taccogna, S., Nardi, F., Panunzi, C., Rinaldi, R., Toscano, V., & Pacella, C. M. (2002). Risk of malignancy in nonpalpable thyroid nodules: predictive value of ultrasound

- and color-Doppler features. *The Journal of clinical endocrinology and metabolism*, 87(5), 1941–1946. <https://doi.org/10.1210/jcem.87.5.8504>
6. Harach, H. R., Franssila, K. O., & Wasenius, V. M. (1985). Occult papillary carcinoma of the thyroid. A “normal” finding in Finland. A systematic autopsy study. *Cancer*, 56(3), 531–538. [https://doi.org/10.1002/1097-0142\(19850801\)56:3<531:aid-cnrcr2820560321>3.0.co;2-3](https://doi.org/10.1002/1097-0142(19850801)56:3<531:aid-cnrcr2820560321>3.0.co;2-3)
 7. 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., Schuff, K. G., Sherman, S. I., Sosa, J. A., Steward, D. L., Tuttle, R. M., & Wartofsky, L. (2016). 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: official journal of the American Thyroid Association*, 26(1), 1–133. <https://doi.org/10.1089/thy.2015.0020>
 8. American College of Radiologist. ACRTI-RADS Steering Committee. <https://doi.org/10.1016/j.jacr.2017.01.046>
Available from: <https://www.acr.org/Clinical-Resources/Reporting-and-Data-Systems/TI-RADS>
 9. Middleton, W. D., Teefey, S. A., Reading, C. C., Langer, J. E., Beland, M. D., Szabunio, M. M., & Desser, T. S. (2018). Comparison of Performance Characteristics of American College of Radiology TI-RADS, Korean Society of Thyroid Radiology TIRADS, and American Thyroid Association Guidelines. *AJR. American journal of roentgenology*, 210(5), 1148–1154. <https://doi.org/10.2214/AJR.17.18822>
 10. Wu, X. L., Du, J. R., Wang, H., Jin, C. X., Sui, G. Q., Yang, D. Y., Lin, Y. Q., Luo, Q., Fu, P., Li, H. Q., & Teng, D. K. (2019). Comparison and preliminary discussion of the reasons for the differences in diagnostic performance and unnecessary FNA biopsies between the ACR TIRADS and 2015 ATA guidelines. *Endocrine*, 65(1), 121–131. <https://doi.org/10.1007/s12020-019-01886-0>
 11. Ahmadi, S., Herbst, R., Oyekunle, T., Jiang, X., Strickland, K., Roman, S., & Sosa, J. A. (2019). Using the ATA and ACR TI-RADS sonographic classifications as adjunctive predictors of malignancy for indeterminate thyroid nodules. *Endocrine practice: official journal of the American College of Endocrinology and the American Association of Clinical Endocrinologists*, 25(9), 908–917. <https://doi.org/10.4158/EP-2018-0559>
 12. Persichetti, A., Di Stasio, E., Guglielmi, R., Bizzarri, G., Taccogna, S., Misischi, I., Graziano, F., Petrucci, L., Bianchini, A., & Papini, E. (2018). Predictive Value of Malignancy of Thyroid Nodule Ultrasound Classification Systems: A Prospective Study. *The Journal of clinical endocrinology and metabolism*, 103(4), 1359–1368. <https://doi.org/10.1210/jc.2017-01708>
 13. Yoon, S. J., Na, D. G., Gwon, H. Y., Paik, W., Kim, W. J., Song, J. S., & Shim, M. S. (2019). Similarities and Differences Between Thyroid Imaging Reporting and Data Systems. *AJR. American journal of roentgenology*, 213(2), W76–W84. <https://doi.org/10.2214/AJR.18.20510>
 14. 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. (2019). Validation and comparison of three newly-released Thyroid Imaging Reporting and Data Systems for cancer risk determination. *Endocrine*, 64(2), 299–307. <https://doi.org/10.1007/s12020-018-1817-8>
 15. 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. (2019). Validation and comparison of three newly-released Thyroid Imaging Reporting and Data Systems for cancer risk determination. *Endocrine*, 64(2), 299–307. <https://doi.org/10.1007/s12020-018-1817-8>
 16. Shen, Y., Liu, M., He, J., Wu, S., Chen, M., Wan, Y., Gao, L., Cai, X., Ding, J., & Fu, X. (2019). Comparison of Different Risk-Stratification Systems for the Diagnosis of Benign and Malignant Thyroid Nodules. *Frontiers in oncology*, 9, 378. <https://doi.org/10.3389/onc.2019.00378>
 17. Baloch ZW, Alexander EK, Gharib H et al. 2010. In: SZ Ali & ES Cibas (eds) *The Bethesda System for Reporting Thyroid Cytopathology*. New York, NY: Springer. Chapter 1, pp:1-4.
 18. Grani, G., Lamartina, L., Cantisani, V., Maranghi, M., Lucia, P., & Durante, C. (2018). Interobserver agreement of various thyroid imaging reporting and data systems. *Endocrine connections*, 7(1), 1–7. <https://doi.org/10.1530/EC-17-0336>

19. Vaccarella, S., Franceschi, S., Bray, F., Wild, C. P., Plummer, M., & Dal Maso, L. (2016). Worldwide Thyroid-Cancer Epidemic? The Increasing Impact of Overdiagnosis. *The New England journal of medicine*, 375(7), 614–617. <https://doi.org/10.1056/NEJMp1604412>
20. 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., Hammers, L. W., Hamper, U. M., Langer, J. E., Reading, C. C., Scoutt, L. M., & Stavros, A. T. (2017). ACR Thyroid Imaging, Reporting and Data System (TI-RADS): White Paper of the ACR TI-RADS Committee. *Journal of the American College of Radiology: JACR*, 14(5), 587–595. <https://doi.org/10.1016/j.jacr.2017.01.046>
21. Grani, G., Lamartina, L., Ascoli, V., Bosco, D., Biffoni, M., Giacomelli, L., Maranghi, M., Falcone, R., Ramundo, V., Cantisani, V., Filetti, S., & Durante, C. (2019). Reducing the Number of Unnecessary Thyroid Biopsies While Improving Diagnostic Accuracy: Toward the “Right” TIRADS. *The Journal of clinical endocrinology and metabolism*, 104(1), 95–102. <https://doi.org/10.1210/jc.2018-01674>
22. Ruan, J. L., Yang, H. Y., Liu, R. B., Liang, M., Han, P., Xu, X. L., & Luo, B. M. (2019). Fine needle aspiration biopsy indications for thyroid nodules: compare a point-based risk stratification system with a pattern-based risk stratification system. *European radiology*, 29(9), 4871–4878. <https://doi.org/10.1007/s00330-018-5992-z>
23. Ha, S. M., Baek, J. H., Na, D. G., Suh, C. H., Chung, S. R., Choi, Y. J., & Lee, J. H. (2019). Diagnostic Performance of Practice Guidelines for Thyroid Nodules: Thyroid Nodule Size versus Biopsy Rates. *Radiology*, 291(1), 92–99. <https://doi.org/10.1148/radiol.2019181723>
24. Ha, E. J., Na, D. G., Baek, J. H., Sung, J. Y., Kim, J. H., & Kang, S. Y. (2018). US Fine-Needle Aspiration Biopsy for Thyroid Malignancy: Diagnostic Performance of Seven Society Guidelines Applied to 2000 Thyroid Nodules. *Radiology*, 287(3), 893–900. <https://doi.org/10.1148/radiol.2018171074>
25. Gómez Sáez J. M. (2014). Chronic autoimmune thyroiditis and thyroid cancer. *Endocrinología y nutrición: organo de la Sociedad Espanola de Endocrinología y Nutricion*, 61(6), 299–301. <https://doi.org/10.1016/j.endonu.2014.03.006>
26. Matsubayashi, S., Kawai, K., Matsumoto, Y., Mukuta, T., Morita, T., Hirai, K., Matsuzuka, F., Kakudoh, K., Kuma, K., & Tamai, H. (1995). The correlation between papillary thyroid carcinoma and lymphocytic infiltration in the thyroid gland. *The Journal of clinical endocrinology and metabolism*, 80(12), 3421–3424. <https://doi.org/10.1210/jcem.80.12.8530576>

FINANCING

The authors did not receive financing for the development of this research.

CONFLICT OF INTEREST

The authors declare that there is no conflict of interest.

AUTHORSHIP CONTRIBUTION

Conceptualization: SM Batallés, JL Novelli, L Caciarelli, I Baumlis, C Argutti, R Asenjo, C Blanco, P Goldstein, L Martinessi, S Ahn.

Formal analysis: SM Batallés, JL Novelli, L Caciarelli, I Baumlis, C Argutti, R Asenjo, C Blanco, P Goldstein, L Martinessi, S Ahn.

Research: SM Batallés, JL Novelli, L Caciarelli, I Baumlis, C Argutti, R Asenjo, C Blanco, P Goldstein, L Martinessi, S Ahn.

Methodology: SM Batallés, JL Novelli, L Caciarelli, I Baumlis, C Argutti, R Asenjo, C Blanco, P Goldstein, L Martinessi, S Ahn.

Drafting - original draft: SM Batallés, JL Novelli, L Caciarelli, I Baumlis, C Argutti, R Asenjo, C Blanco, P Goldstein, L Martinessi, S Ahn.

Writing - proofreading and editing: SM Batallés, JL Novelli, L Caciarelli, I Baumlis, C Argutti, R Asenjo, C Blanco, P Goldstein, L Martinessi, S Ahn.