European Journal of Radiology Open 6 (2019) 343–349

Contents lists available at ScienceDirect

European Journal of Radiology Open

journal homepage: www.elsevier.com/locate/ejro

Classification and Regression Tree (CART) model of sonographic signs in predicting thyroid nodules malignancy

Tarek Smayra[a,*], Zahra Charara[a], Ghassan Sleilaty[b], Gaelle Boustany[a], Lina Menassa-Moussa[a], Georges Halaby[c]

[a] Radiology Department, School of Medicine, Hotel-Dieu Hospital, Saint Joseph University, Alfred Naccache Street, PO Box: 16-6830, Beirut, Lebanon
[b] Clinical Research Center, School of Medicine, Hotel-Dieu Hospital, Saint Joseph University, Alfred Naccache Street, PO Box: 16-6830, Beirut, Lebanon
[c] Endocrinology Department, School of Medicine, Hotel-Dieu Hospital, Saint Joseph University, Alfred Naccache Street, PO Box: 16-6830, Beirut, Lebanon

ARTICLE INFO

Keywords:
Thyroid nodule
Thyroid neoplasms
Biopsy
Fine-needle
Ultrasonography

ABSTRACT

Purpose: To develop a Classification and Regression Tree (CART) model in order to recognize the most suspicious sonographic features of thyroid nodules and efficiently guide their management.

Methods: 791 thyroid fine needle aspiration cytology (FNAC) performed under ultrasound guidance between January 2015 and January 2017 were reviewed. Retrieved data consisted in qualitative (patient's gender, composition, echogenicity, shape, margins and echogenic foci of the nodule) and quantitative (patient's age and maximal diameter of the nodule) variables as well as the Bethesda score.

Results: Patients were 48.5 ± 13.7 years old with female to male ratio of 8:2. The nodules had median size of 2.3 (1.5–3.5) cm with a majority of solid (62.5 %) and isoechoic (50.8 %) features. 700 nodules (88.5 %) had a wider-than-tall shape, 600 (75.9 %) smooth margins and 113 (14.3 %) ill-defined ones. Echogenic foci were absent in 388 nodules (49.1 %) and, when present, largely dominated by punctate foci (32.5 %). Bethesda classes 3, 4 and 5, which require surgery, represented only 10.6 % of cases. They were significantly correlated with the taller-than-wide shape and with solid or predominantly solid features. There was no significant correlation between echostructure and Bethesda scores but we did find more nodules classified Bethesda 4 and 5 in the categories hypoechoic and severely hypoechoic. In the CART model we developed, the sequence leading to most nodules classified Bethesda 4 and 5 is: taller-than-wide shape, solid composition and hypoechoic or severely hypoechoic feature.

Conclusions: Taller-than-wide, solid or predominantly solid, hypoechoic or severely hypoechoic nodules are likely to require surgery and might benefit from FNAC.

1. Introduction

Thyroid ultrasound is nowadays an essential tool for the management of thyroid nodules because of its high sensitivity and specificity, as well as its wide availability. The use of ultrasound in thyroid disease dates back to the late 1960s, with the apparition of B mode scanning and low-resolution images [1]. Currently, thyroid ultrasound is performed using high frequency linear probes, allowing high-resolution imaging and accurate guidance of fine needle aspiration cytology (FNAC) when needed. In parallel, machine-based data mining is allowing more and more decision trees extraction and clinical algorithms to be established. In this study, we statistically analyze different sonographic characteristics of nodules and develop the most complete Classification and Regression Tree (CART) model yet in order to recognize the most suspicious features and efficiently guide nodules management.

2. Materials and methods

2.1. Patient's selection

Institutional Review Board approved this retrospective study, performed in a single center, and informed consent was waived. Our patients are all those referred for ultrasound guided FNAC of thyroid nodules found on clinical examination or on a previous ultrasound. A new comprehensive ultrasound is always performed and nodules are selected as FNAC targets on the basis of consensus recommendations [2]. The selected route is the one that goes directly to the nodule,

* Corresponding author.
E-mail address: tarek_smayra@icloud.com (T. Smayra).

https://doi.org/10.1016/j.ejro.2019.11.003
Received 1 September 2019; Received in revised form 8 November 2019; Accepted 18 November 2019
Available online 28 November 2019

preferably through a trans-isthmic path while avoiding vascular structures. The patient is informed of the FNAC technique, its potential complications and his consent is obtained according to our institution policy. Thyroid ultrasound examination and FNAC were performed by a specialized radiologist with more than 15 years of experience.

2.2. FNAC procedure

The patient is placed on the examination table, with a slight hyperextension of the head. Ultrasound examination is performed using one of two machines upon availability. The first is an ACUSON Antares (Siemens Healthcare, Erlangen, Germany) ultrasound machine and the second a LOGIQ E9 (GE Healthcare, Little Chalfont, United Kingdom). In both cases, a high frequency linear array probe with high resolution is used (VFX13-5 and ML6-15 respectively). First, we systematically take the dimensions of the two thyroid lobes (length, width and thickness) and the maximum thickness of the isthmus. Then, all identified anomalies such as abnormal echo structure, nodules and calcifications are reported. For each nodule, we note its sonographic characteristics listed in 6 categories according to the lexicon published by Grant et al. [3]: composition, echogenicity, shape, size, margins and echogenic foci. The study is concluded with the analysis of cervical lymph nodes areas in search of suspicious lymph nodes. FNAC is performed under ultrasound guidance, using the "in-plane" technique that permits to see and control the progression of the needle and the precise location of the sampling area in real-time (Fig. 1). After cleaning the skin with alcohol solution and without the use of local anesthesia, the patient is instructed to swallow two or three times before inserting the needle. A simple 22 G needle is used without a syringe, keeping the Luer extremity in the open air. When the target is penetrated, the Luer is caught between the thumb and index finger and small movements back and forth are performed within the nodule, which, capillarity helping, allow the needle to fill with cytological material. Two to three passages through each nodule are made based on the collected material, which is immediately spread over several glass slides, half of which will be placed in absolute alcohol for Papanicolaou staining and the others left to dry for Diff-Quick staining technique. A small bandage is applied to the puncture entry points and instructions are given to the patient for possible pain or the apparition of a hematoma.

2.3. Studied variables

791 thyroid FNAC were performed under ultrasound guidance between January 2015 and January 2017. For each case, the ultrasound images were retrospectively reviewed and various patient and nodules data were selected for analysis (Table 1). Qualitative variables are patient gender, composition, echogenicity, shape, margins and echogenic foci of the nodule. Quantitative variables are patient's age and maximal diameter of the nodule. Also, the Bethesda System of Classification was used in the pathology reports and the Bethesda score of each nodule was obtained (Fig. 2).

2.4. Statistical analysis

Departure from normality was assessed based on Quartile-Quartile plots along with Kolmogorov-Smirnov and Shapiro-Wilk tests. Continuous variables not departing from normality assumptions were expressed as mean ± standard deviation. Variables that departed significantly from normality (e.g. nodule size) were expressed as median with its interquartile range. In this latter case, a Box-Cox power transformation was applied, to optimize subsequent use of these variables in the multivariate analysis. In univariate analysis, the Mann-Whitney test, the Kruskall-Wallis test, the Jonckheere-Tespstra test, and Chi square tests were used for ordinal and categorical variables. ANOVA with linear contrasts was used to compare continuous variables (or their Box-Cox transform) among Bethesda classes. 95 % confidence intervals were derived by bootstrapping, based on 1000 samples. Before proceeding to multivariable analysis, the variables were classified by decreased order of their predictive potential: shape, echogenicity, margins, composition, size, and age. The independent variables were then optimized: No change was made for the variable 'shape', with a high predictive power. Variable 'echogenicity' was recoded into two classes, the first one including hyperechoic and isoechoic nodules and the second one hypoechoic and severely hypoechoic nodules. Variable 'margins' was recoded into 3 classes. In the first one we kept the smooth margins. In the second one we included both the irregular and lobulated margins, and in the third one we put the remainder: ill-defined, halo and extra-thyroidal extension. Variable 'composition' was also recoded into 3 classes: solid, predominantly solid, and the remaining

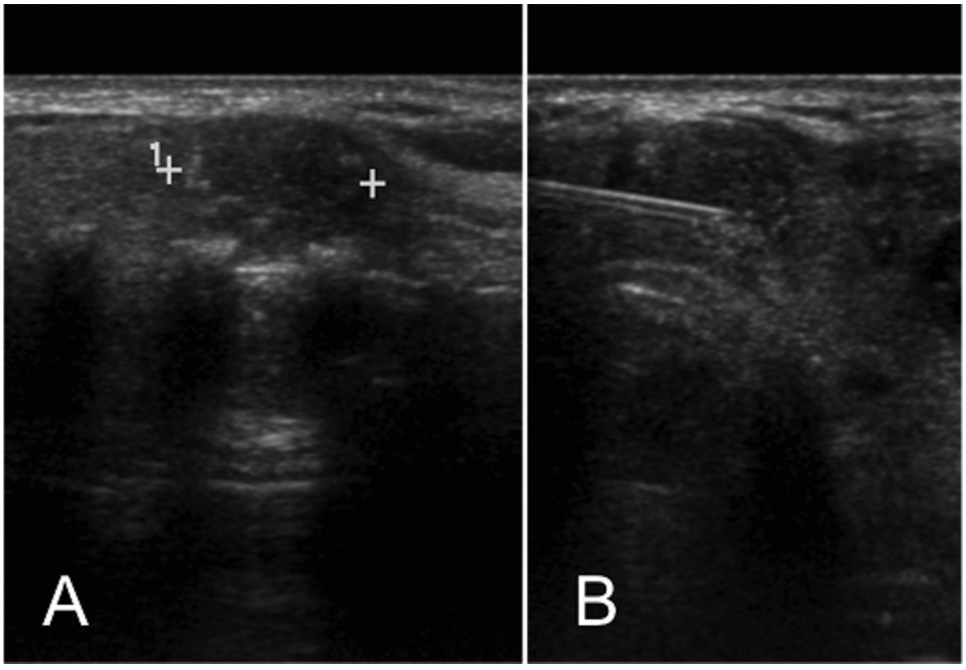

Fig. 1. Thyroid nodule with suspicious sonographic features such as hypoechogenicity and microcalcifications (A). FNAC of the same nodule (B).

Table 1
Characteristics of the 791 nodules.

Variable	Categories	Frequency	Valid Percent	Cumulative Percent	95 % CI for percent[a]
Gender	Female	635	80.3	80.3	77.7–82.8
	Male	156	19.7	100.0	17.1–22.6
Composition	Solid	494	62.5	62.5	58.9–65.6
	Predominantly solid	157	19.8	82.3	17.3–22.8
	Predominantly cystic	69	8.7	91.0	7.1–10.6
	Cystic	12	1.5	92.5	.9–2.1
	Spongiform	59	7.5	100.0	5.6–9.4
Echogenicity	Hyperechoic	63	8.0	8.0	6.2–9.9
	Isoechoic	402	50.8	58.8	47.4–54.1
	Hypoechoic	254	32.1	90.9	29.1–35.0
	Severely hypoechoic	72	9.1	100.0	7.2–11.1
Shape	Height > Width	91	11.5	11.5	9.6–13.5
	Width > Height	700	88.5	100.0	86.3–90.6
Margins	Regular	600	75.9	75.9	72.9–78.6
	Irregular	43	5.4	81.3	4.0–7.0
	Lobulated	29	3.7	85.0	2.7–4.9
	Ill-defined	113	14.3	99.2	12.1–16.4
	Halo	5	.6	99.9	.1–1.3
	Extra-thyroid extension	1	.1	100.0	.0–.4
Echogenic foci	Absence	388	49.1	49.1	46.0–52.2
	Punctiform	257	32.5	81.5	29.5–35.8
	Macrocalcifications	112	14.2	95.7	12.0–16.3
	Peripheral	19	2.4	98.1	1.5–3.3
	Comet tail	15	1.9	100.0	1.1–2.6
Bethesda class	0: Non-diagnostic/unsatisfactory	71	9.0	9.0	7.1–11.0
	1: Benign/Non-cancerous	533	67.4	76.4	64.1–70.5
	2: Indeterminate	103	13.0	89.4	11.0–15.2
	3: Suspicious for follicular neoplasm	38	4.8	94.2	3.5–6.3
	4: Suspicious for cancer	28	3.5	97.7	2.4–4.8
	5: Positive for cancer	18	2.3	100.0	1.4–3.2
Variable	**Statistic**	**Result**			
Age (years)	Mean ± Std Deviation	48.5 ± 13.7			
Nodule size (cm)	Median (Quartile 1 – Quartile 3)	2.3 (1.5–3.5)			
BC Nodule size	Mean ± Std Deviation	.945 ± .691			

[a] Unless otherwise noted, bootstrap results are based on 1000 bootstrap samples; BC stands for Box-Cox transformation, in this case $5(\sqrt[5]{size} - 1)$, size denoting the nodule size.

SCORE	CATEGORY	MANAGEMENT	MALIGNANCY
0	Non diagnostic	Repeat FNAC	-
1	Benign	Follow-up by ultrasound	0 - 3%
2	Atypia of undetermined significance or follicular lesion of undetermined significance	Repeat FNAC	5 - 15%
3	Suspicious for follicular neoplasm	Surgery	15 - 30%
4	Suspicious for malignancy	Surgery	60 - 75%
5	Malignant	Surgery	97 - 99%

Fig. 2. The Bethesda Classification System.

predominantly cystic, cystic and spongiform were merged in one class. Variables 'age' and 'Box-Cox of nodule size' were already normalized. They were standardized (that is, standard normal distribution) and then recoded each in two classes: for 'age', below and above 48 years old, and for nodule 'size', smaller and larger than 2.4 cm. Multivariate analysis relied on CART (Classification And Regression Trees). Target categories were defined as Bethesda classes 3, 4 and 5. Iterative segregation used Twoing schema and cross validation was performed on 16 groups. Minimal parent and child node sizes were set to 40 and 20 respectively. Maximal tree depth was set to 3 levels. No pruning was done on the model. Statistical analysis was performed using SPSS software (IBM Corp. Released 2013. IBM SPSS Statistics for Windows, Version 22.0. Armonk, NY: IBM Corp.).

3. Results

3.1. Descriptive part

Between January 2015 and January 2017, 791 patients with a mean age of 48.5 ± 13.7 years had undergone FNAC of thyroid nodules, which were sent for cytological analysis. The nodules had median size of 2.3 (1.5–3.5) cm. The Box-Cox transformation for 'age' yielded lambda = 0.2, with the transform of size $5(\sqrt[5]{size} - 1)$ having a normal distribution. In total, there were 635 women (80.3 %) and 156 men (19.7 %). The largely predominant type of nodule composition was the solid one (62.5 %), followed by isoechoic nodules (50.8 %), and hypoechoic ones (32.1 %). There was a clear majority of wider-than-tall nodules, 700 cases (88.5 %). Analysis of the margins showed that 600 nodules (75.9 %) had smooth margins and 113 nodules (14.3 %) had ill-defined ones. Halo sign and extra-thyroid extension were rarely found (respectively 0.6 and 0.1 %). Echogenic foci were absent in 388 nodules

Table 2
Distribution of gender and nodule sonographic characteristics according to Bethesda classes.

	Bethesda Class							
	0 N(%)	1 N(%)	2 N(%)	3 N(%)	4 N(%)	5 N(%)	Test	P-value
Gender								
Female	56(78.9 %)	432(81.1 %)	86(83.5 %)	30(78.9 %)	18(64.3 %)	13(72.2 %)	MW	.518
Male	15(21.1 %)	101(18.9 %)	17(16.5 %)	8(21.1 %)	10(35.7 %)	5(27.8 %)		
Composition								
Solid	50(70.4 %)	294(55.2 %)	81(78.6 %)	27(71.1 %)	24(85.7 %)	18(100 %)	KW	< .001
Predominantly solid	10(14.1 %)	117(22 %)	16(15.5 %)	10(26.3 %)	4(14.3 %)	0.0 %)	JT	< .001
Predominantly cystic	6(8.5 %)	60(11.3 %)	3(2.9 %)	0.0 %)	0.0 %)	0.0 %)		
Cystic	2(2.8 %)	10(1.9 %)	0.0 %)	0.0 %)	0.0 %)	0.0 %)		
Spongiform	3(4.2 %)	52(9.8 %)	3(2.9 %)	1(2.6 %)	0.0 %)	0.0 %)		
Echogenicity								
Hyperechoic	3(4.2 %)	44(8.3 %)	10(9.7 %)	5(13.2 %)	0.0 %)	1(5.6 %)	KW	.431
Isoechoic	27(38 %)	297(55.7 %)	50(48.5 %)	20(52.6 %)	4(14.3 %)	4(22.2 %)	JT	.356
Hypoechoic	37(52.1 %)	141(26.5 %)	36(35 %)	12(31.6 %)	19(67.9 %)	9(50 %)		
Severely hypoechoic	4(5.6 %)	51(9.6 %)	7(6.8 %)	1(2.6 %)	5(17.9 %)	4(22.2 %)		
Shape								
Height > Width	3(4.2 %)	40(7.5 %)	18(17.5 %)	6(15.8 %)	13(46.4 %)	11(61.1 %)	MW	< .001
Width > Height	68(95.8 %)	493(92.5 %)	85(82.5 %)	32(84.2 %)	15(53.6 %)	7(38.9 %)		
Margins								
Regular	58(81.7 %)	412(77.3 %)	72(69.9 %)	32(84.2 %)	14(50 %)	12(66.7 %)	KW	< .001
Irregular	5.7 %)	17(3.2 %)	10(9.7 %)	1(2.6 %)	8(28.6 %)	2(11.1 %)	JT	.053
Lobulated	1(1.4 %)	12(2.3 %)	9(8.7 %)	1(2.6 %)	3(10.7 %)	3(16.7 %)		
Ill-defined	6(8.5 %)	89(16.7 %)	11(10.7 %)	3(7.9 %)	3(10.7 %)	1(5.6 %)		
Halo	1(1.4 %)	2(0.4 %)	1.1 %)	1(2.6 %)	0.0 %)	0.0 %)		
Extra-thyroid extension	0.0 %)	1(0.2 %)	0.0 %)	0.0 %)	0.0 %)	0.0 %)		
Echogenic foci								
Absence	38(53.5 %)	252(47.3 %)	59(57.3 %)	23(60.5 %)	10(35.7 %)	6(33.3 %)	KW	.953
Punctiform	22(31 %)	180(33.8 %)	28(27.2 %)	10(26.3 %)	12(42.9 %)	5(27.8 %)		
Macrocalcifications	9(12.7 %)	77(14.4 %)	13(12.6 %)	4(10.5 %)	4(14.3 %)	5(27.8 %)		
Peripheral	1(1.4 %)	14(2.6 %)	2(1.9 %)	1(2.6 %)	1(3.6 %)	0.0 %)		
Comet tail	1(1.4 %)	10(1.9 %)	1.1 %)	0.0 %)	1(3.6 %)	2(11.1 %)		

MW: Mann-Whitney U test; KW: Kruskall-Wallis test; JT: Jonckheere-Terpstra.

Table 3
Distribution of age and nodule size according to Bethesda classes.

Variable Bethesda class	Age (years) Mean ± standard deviation	Box-Cox nodule size Mean ± standard deviation
Bethesda 0	52.6 ± 12.9	0.812 ± 0.655
Bethesda 1	48.9 ± 13.4	1.024 ± 0.654
Bethesda 2	46.8 ± 14.8	0.839 ± 0.597
Bethesda 3	46.9 ± 14.6	0.953 ± 0.599
Bethesda 4	40.6 ± 11.8	0.472 ± 0.823
Bethesda 5	45.8 ± 16.4	0.447 ± 0.92
ANOVA (p-value)	0.003	< .001

Box-Cox transformation of nodule size in this case is $5(\sqrt[5]{size} - 1)$, size denoting the nodule size.

(49.1 % of cases) and, when present, largely dominated by punctate foci (257 nodules or 32.5 %) and macrocalcifications (112 nodules or 14.2 %). On pathological analysis, there was a majority of Bethesda 1 (533 nodules or 67.4 %), followed by Bethesda 2 (103 nodules or 13 %). Bethesda 0 (71 nodules or 9 %) represents the cases where pathological analysis was inconclusive, either because of insufficient cells or the presence of too much blood. Bethesda 3, 4 and 5 classes that require surgery are made of 84 nodules or 10.6 %. The detailed characteristics of the nodules are depicted in Table 1.

3.2. Analytical part

In univariate analysis, Bethesda classes distribution was not influenced by gender (Mann-Whitney p value = 0.518, Table 2). However, Bethesda classes distribution was statistically different when considering nodule composition (Kruskall-Wallis p value < 0.001), with transition for upper Bethesda classes being associated with the solid component (Jonckheere-Terpstra p value < 0.001). Bethesda classes distribution was not statistically associated with echogenicity (Kruskall-Wallis p value = 0.431, Jonckheere-Terpstra p value = 0.356). There is however a significant difference of Bethesda classes distribution in relation to nodules shape (Mann-Whitney p value < 0.001) and it appears clearly that Bethesda class is higher in the taller-than-wide group. When nodules margins are considered, Bethesda classes distribution is not the same (Kruskall-Wallis test p value < 0.001), without being necessarily an ordinal relationship (Jonckheere-Terpstra p value = 0.053). The decomposition of Kruskall-Wallis test into homogeneous subtypes shows in fact two different profiles of margins: halo, irregular and lobulated margins represent a separate group with relatively high Bethesda score, whereas other types of margins have low Bethesda score. There was no statistically significant relation between Bethesda classes and echogenic foci types (Kruskall-Wallis p value = 0.953). Age distribution was significantly different among Bethesda classes as shown by ANOVA (Table 3) with a pre-specified linear contrast (p = 0.003), suggesting a negative linear decrease of mean age with the increase of Bethesda class. Likewise, the Box-Cox transform of nodule size was different among the different Bethesda classes, as shown by ANOVA with a pre-specified linear contrast (Table 3, p < 0.001). Using CART (Classification and Regression Trees) methodology as defined in the statistical analysis section, a decision tree with 13 nodes and 7 terminal nodes was derived as depicted in the detailed decision tree shown in Fig. 3. Node 0 represents the initial distribution of the nodules according to Bethesda classification before any sonographic criteria had been applied. It replicates exactly the elements in Table 1 for Bethesda classes. Nodes 1 and 2 reflect the effect of introducing 'shape' as a first decision criterion. The category 'Height > Width' relocates most of the Bethesda 3, 4, 5 to node 2 with increased frequencies compared to node 0 (6.6 %, 14.3 %, and 12.1 % respectively). Node 1 contains less Bethesda 4 and 5 classes than nodes 0 and 2. The residual Bethesda classes

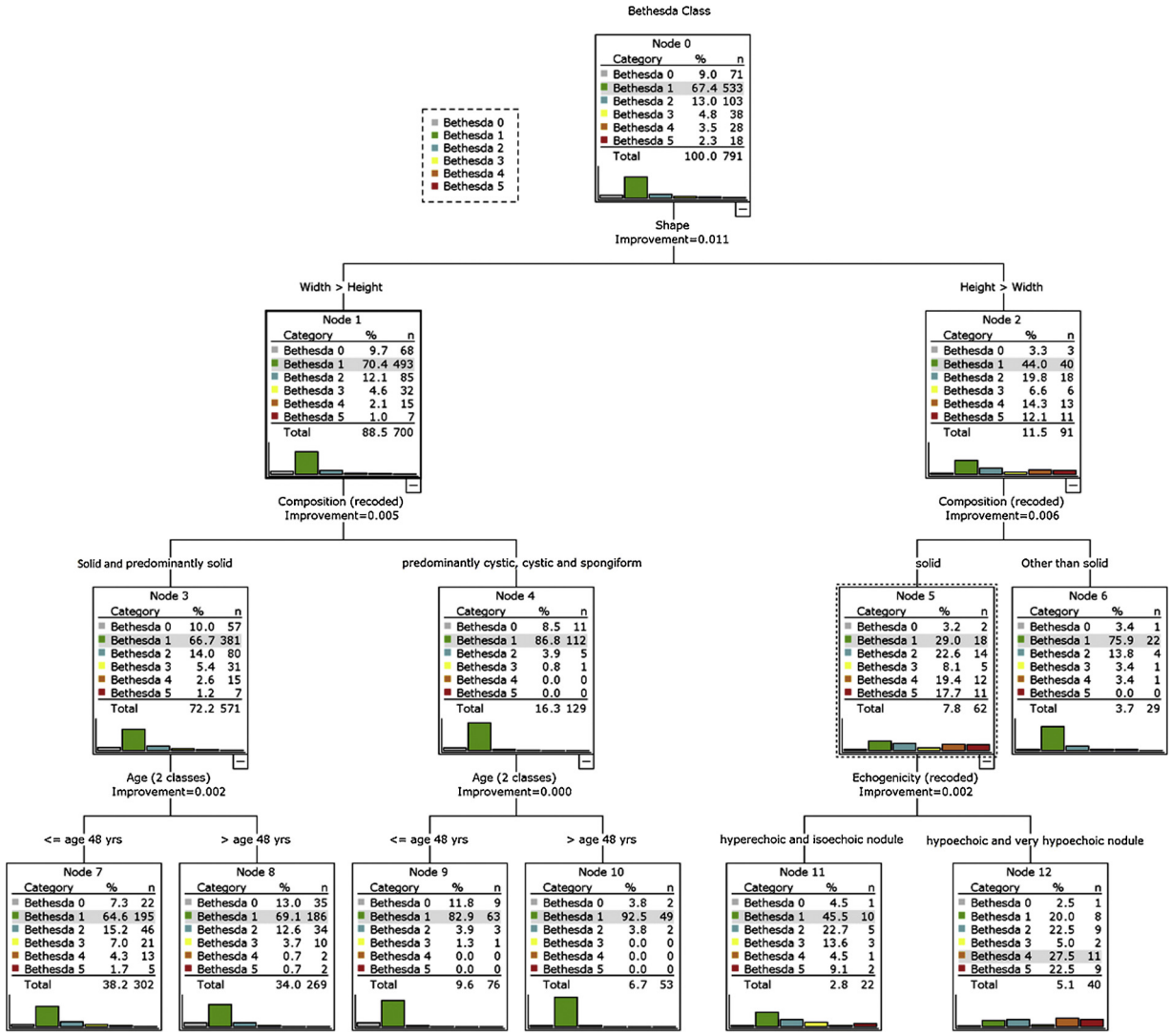

Fig. 3. Flow chart of a Classification and Regression Trees (CART) model showing detailed decision tree for the sonographic features leading to a concentration of nodules having Bethesda classes 3, 4 and 5.

Table 4
Gains of the CART analysis by node, and CART error rate.

Gains for nodes						
Node	Node		Gain		Response	Index
	N	Percent	N	Percent		
12	40	5.1 %	9	50.0 %	22.5 %	988.8 %
11	22	2.8 %	2	11.1 %	9.1 %	399.5 %
7	302	38.2 %	5	27.8 %	1.7 %	72.8 %
8	269	34.0 %	2	11.1 %	0.7 %	32.7 %
9	76	9.6 %	0	0.0 %	0.0 %	0.0 %
10	53	6.7 %	0	0.0 %	0.0 %	0.0 %
6	29	3.7 %	0	0.0 %	0.0 %	0.0 %

Risk method	**Estimate**	**Std. Error**
Resubstitution	.322	.017
Cross-Validation	.330	.017

Growing method: CART.
Dependent variable: Bethesda Class.

in node 1 will be subjected to the second criterion for further segregation. The second criterion for the decision process is the composition of the nodule: When the nodule is wider-than-tall (Width > Height, node 1), its composition being predominantly cystic, cystic or spongiform ensures no Bethesda 4 and 5 classes are remaining, and a marginal 0.8 % frequency of Bethesda 3 class (node 4). The Bethesda 3, 4 and 5 classes inherited from node 1 are being relocated in node 3 by the solid and predominantly solid composition criterion. When the nodule is taller-than-wide (Height > Width, node 2), its composition being solid relocates all Bethesda 3, 4 and 5 classes but 2 cases in node 5, and drastically increases their likelihood: 8.1 %, 19.4 %, and 17.7 % respectively. Node 6 corresponds to a composition other than solid and contains 2 cases (Bethesda 3 and 4). The third level for the decision tree helps refining further the classification. For the nodules that are wider-than-tall (Width > Height) and depending on their composition (nodes 3 and 4), there remains a minimal risk of Bethesda 3, 4 and 5 in node 3. Applying the 'age ≤ 48 years' criterion will relocate Bethesda 3, 4 and 5 classes in node 7, thus leaving null or quasi-null frequencies for nodes 8, 9 and 10. For the nodules that are taller-than-wide (Height > Width) and depending on their composition (nodes 5 and 6), there is a dismal risk of Bethesda 3, 4 and 5 in node 6 with no further break down, node 6 being terminal. For node 5 that contains much of the Bethesda 3, 4 and 5 classes, applying 'echogenicity' as a 3rd criterion will increase further the likelihood of these classes in the presence of hypoechoic and severely hypoechoic nodules (5 %, 27.5 %, and 22.5 % respectively in node 12), thus leaving lower frequencies for Bethesda 3, 4 and 5 classes

in node 11 than in node 5. The gains for nodes 11 and 12 were maximal (9.9 % and 4.0 % respectively) as shown in Table 4. The overall error rate of this decision tree, assessed by two methods, is around 30 %. For a given nodule, before applying any criterion, the probability of being Bethesda class 5 is 2.3 %. If the nodule is taller-than-wide, with solid composition and hypoechoic to severely hypoechoic, the probability of being Bethesda class 5 is increased to 27.5 % according to this CART analysis.

4. Discussion

Having a majority of benign (Bethesda 1) and inconclusive/non-diagnostic (Bethesda 0) in our study is largely found in many published series, and it is widely admitted that thyroid nodules with a higher Bethesda cytological score of 3, 4 or 5 present a malignancy risk of 15–99% and require a surgical intervention, whether a lobectomy or a total thyroidectomy. Our aim is to perform FNAC on all of these nodules and spare the others. In the statistical analysis of our series, the use of variables shape, composition, age and echogenicity according to a logical decision tree is able to depict the large majority of these cytologically suspicious nodules. The taller-than-wide shape (Height > Width) is usually evident on visual inspection in the axial plane images and there is no need to measure the height and width of the nodule in order to demonstrate it. In a large study of 831 patients, Moon and al [4] have shown that taller-than-wide nodule was a sonographic sign in favor of malignancy. They have later confirmed it in a consensus revue published in 2011 on the management of nodules [5]. Other authors also consider the taller-than-wide shape as being a malignant sign [6]. In our study, 11.5 % of the nodules had a taller-than-wide shape, and the application of this criterion attracts most of the Bethesda 3, 4 and 5 nodules with the best statistical improvement. In 2011, Moon and al [7] published a study on 471 thyroid nodules. Three criteria regarding the shape were formulated as follows: criterion 1, a taller-than-wide shape in any plane as a suspicious feature; criterion 2, the same in the transverse plane only; criterion 3, the same in the longitudinal plane only. They concluded that a taller-than-wide shape was a useful sonographic feature for predicting thyroid malignancy. In particular, a taller-than-wide shape in either transverse or longitudinal plane was most accurate and sensitive for predicting thyroid malignancy among the three criteria. This feature's sensitivity varies between 40 and 68 %, specificity between 82 and 93 %, positive predictive value ranges between 58 and 73 %, and negative predictive value between 77 % and 88 % [3,8–12]. This is compatible with our results where Bethesda classes are statistically higher in the taller-than-wide group. The second most important variable found in our study was nodule composition. The largely prevailing type in our series is the solid nodule, which represents 62.5 % of the cases, followed by the predominantly solid type that represents 19.8 % of the cases. This is not unlikely, since the solid or predominantly solid composition of a nodule is a main indication for FNAC. Many studies have reported a large majority of solid nodules in their series as well, like Naoun [13] with 58.2 % of the cases, Solbiati [14] with 63 % and Leenhard [15] with 61.9 %. When this variable is applied in the decision process, after the variable 'shape', we find that almost all the Bethesda 3, 4 or 5 are solid or predominantly solid if they are wider-than-tall. Also, almost all of these are solid if they are taller-than-wide. Our analysis has also demonstrated that the Bethesda score tends to increase when the nodule's composition goes from spongiform to solid. In other words, the Bethesda categories 3, 4 or 5 are more represented within the solid and predominantly solid nodules types. The category Bethesda 5 is exclusively composed of solid nodules. We therefore agree with Bonavita et al. [16] about the correlation between solid nodules and malignancy. In addition to that, they have isolated four sonographic aspects correlated with 100 % benign lesions: spongiform configuration, colloid cyst, diffuse hyperechogenicity (as in goiter) and a 'giraffe' aspect of the thyroid parenchyma (as in chronic thyroiditis). When no real problem is found in identifying solid and predominantly solid, as well as cystic nodules on sonography, it is important to recognize the spongiform nodule as is and also the mixed nodule with a suspicious solid component. The solid component of a mixed nodule has to be evaluated. If it is peripherally located, has an acute angle with the wall of the nodule, is hypoechoic, lobulated, has an irregular border or punctuate echogenic foci line, the risk of malignancy is increased [3,17]. In the case of taller-than-wide and solid nodules, a significant majority of Bethesda 3, 4 or 5 lesions can be isolated using echogenicity, hypoechoic or severely hypoechoic nodules being the most suspicious. Further statistical evaluation showed us that applying age with a cutoff of 48 years old could isolate the remaining suspicious nodules in the wider-than-tall and solid or predominantly solid sub-group. Patients aged 48 or less are more prone to have Bethesda 3, 4 or 5 nodules. This is interesting, since young age is not known to represent a significant variable of malignant risk and is not included in workup recommendations. However, Bessey et al. [18] in 2013 have tried to determine the influence of age and gender on the rate of thyroid nodule malignancy diagnosed by FNAC. They concluded that the risk of thyroid nodule malignancy on FNAC varies depending on patient's age and gender. Specifically, the rates of thyroid cancer diagnosed by FNAC were higher in patients younger than 45 years. The highest incidences of malignancy were between ages 30 and 39 in women and ages 40 and 49 in men. They nonetheless pointed out that most of the studies that correlated risk of malignancy and age usually only examine patients that underwent surgical excision, which creates an inherent bias, especially in regards to the risk of malignancy. On the other hand, our study shows that size of the nodule is not significantly the same among the different Bethesda categories, and furthermore, that there is a negative linear relation between nodule size and Bethesda score. In other words, the smaller the nodule size, the lower the Bethesda score and the lower the risk of malignancy. In the literature, the data correlating nodule size and malignancy are variable. For some authors, a size superior to 4 cm increases the risk of malignancy [19,20]. Other series, such as Nguyen's et al. [21] have demonstrated that the size of the nodule affects rather the survival rate than malignancy. Indeed, for tumors measuring between less than 5 mm and 39 mm, the 10 years survival rate varied from 99.4–99.9 %. This rate was only 96.6 % for the tumors measuring 39–40 mm, and it dropped to 84 % for tumors measuring 40 mm or more. When considering margins of the nodules, we found that 75.9 % of the nodules in our series had regular ones, which is a well-known benign sign. Having a peripheral halo is also considered to be in favor of a benign nodule; this sign was found in only 5 nodules in our series (0.6 % of the cases). There was no ordinal relation between the type of margins and the Bethesda score. Nonetheless, nodules with halo, lobulated margins or irregular margins are associated with higher Bethesda score compared to other types of margins. We thus agree with authors [4,8,16,22] associating lobulated or irregular margins with malignancy. Nevertheless, Moon and al [4] have in 2008 hypothesized that a spiculated margin is suggestive of malignancy, while an ill-defined margin can be seen in both benign and malignant nodules, even if previous studies have suggested that blurred or ill-defined margins favor a diagnosis of malignancy. According to them, with the development of sonographic techniques, a previously described ill-defined margin could actually be a spiculated margin with sharp demarcation or a poorly defined margin in which the tumor cannot be discriminated from the normal parenchyma. We feel that the assessment of nodule margins on ultrasound is difficult and might be unreliable. In fact, Pang et al. showed in a recent study, a high inter-observer discrepancy when assessing nodule margins on ultrasound [23]. No correlation was found in our study between echogenic foci and Bethesda score. This appears strange because the presence of microcalcifications is presented in the literature as predictive of malignancy [4,10] and we found echogenic foci in 50.9 % of the nodules, and among those, 32.5 % were microcalcifications. The discrepancy between our series and the published data can pertain to the difficulty of identifying microcalcifications in thyroid

sonography, as well as the other echogenic foci. Microcalcifications are defined as hyperechoic punctuations, round or linear, measuring less than 1 mm in diameter, without acoustic shadowing, whereas macrocalcifications measure more than 3 mm and often have a strong acoustic shadow. The “comet tail” artifacts are echogenic foci with V-shaped posterior echoes. They are associated with colloid granulations and are strongly indicative of benignity. Malhi and al [12] have in their study found no statistical difference in the prevalence of malignant lesions between nodules with and those without echogenic foci. They also highlight that very few authors have precisely defined echogenic foci before studying the association between these foci and malignant nodules. Most of them on the contrary, have considered echogenic foci as a group of features, largely considered as microcalcifications, probably as we did in our study. Regarding the use of a CART model in the management of thyroid nodules, there is only one available study in which clinical as well as sonographic features were used but the latter are not as detailed as in our study. It lacks for example the most significant feature we found in our study, which is the taller-than-wide shape [24].

5. Conclusions

High-resolution thyroid ultrasound is nowadays the preferred tool for detection, analysis and management of thyroid nodules. The objective of this study was to find a statistical value of sonographic features in predicting malignancy. We also have modelized a Classification And Regression Tree (CART) based on those features in order to predict the Bethesda scores 3, 4 and 5. We noted first that there is a significant correlation between the shape taller-than-wide and Bethesda scores 3, 4 and 5. We noted second that there is a significant correlation between solid or predominantly solid nodules and Bethesda scores 3, 4 and 5. Third, there is no significant correlation between echostructure and Bethesda scores but we did find more nodules classified Bethesda 4 and 5 in the categories hypoechoic and severely hypoechoic. And last, there is a negative linear relation between the nodule’s size and Bethesda score. In the CART model we developed, the sequence leading to most nodules classified Bethesda 4 and 5 is: taller-than-wide shape, solid composition and hypoechoic or severely hypoechoic nodule. In case of wider-than-tall shape and solid or predominantly solid nodule, age of 48 years old and less holds more risk of having a high Bethesda score. No other study had shown in a CART model the importance of a major feature, which is the taller-than-wide sign in the assessment of thyroid nodules. Our study is however biased since our patients were not randomized but selected by their referring physician to undergo FNAC and we are unable to know what do they represent among general population. On the other hand, we believe that more studies involving the taller-than-wide feature could enhance ultrasound specificity in thyroid cancer.

IRB approval

Our study was submitted to our institution’s Ethics committee, called “Comité d’Ethique de l’HDF” who gave its approval under the reference CEHDF 1403.

Declaration of Competing Interest

None.

Acknowledgements

None

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

References

[1] R.A. Levine, Something old and something new: a brief history of thyroid ultrasound technology, Endocr. Pract. 10 (2004) 227–233.
[2] J. Tramalloni, J.L. Wémeau, Consensus français sur la prise en charge du nodule thyroïdien: ce que le radiologue doit connaître, EMC 0 (2012) 1–18.
[3] E.G. Grant, F.N. Tessler, J.K. Hoang, et al., Thyroid ultrasound reporting lexicon: white paper of the ACR Thyroid imaging, Reporting and Data System (TI-RADS) Committee, J. Am. Coll. Radiol. 12 (2015) 1272–1279.
[4] W.J. Moon, S.L. Jung, J.H. Lee, et al., Benign and malignant thyroid nodules: US differentiation-multicenter retrospective study, Radiology 247 (2008) 762–770.
[5] W.J. Moon, J.H. Baek, S.L. Jung, et al., Ultrasonography and the ultrasound-based management of thyroid nodules: consensus statement and recommendations, Korean J. Radiol. 12 (2011) 1–14.
[6] M.T. Heller, C. Gilbert, N.P. Ohori, M.E. Tublin, Correlation of ultrasound findings with the Bethesda cytopathology classification for thyroid nodule fine-needle aspiration: a primer for radiologists, AJR Am. J. Roentgenol. 201 (2013) 487–494.
[7] H.J. Moon, J.Y. Kwak, E.K. Kim, M.J. Kim, A taller-than-wide shape in thyroid nodules in transverse and longitudinal ultrasonographic planes and prediction of malignancy, Thyroid 21 (2001) 1249–1253.
[8] J.Y. Kwak, K.H. Han, J.H. Yoon, et al., Thyroid imaging reporting and data system for US features of nodules: a step in establishing better stratification of cancer risk, Radiology 260 (2011) 892–899.
[9] S.S. Ahn, E.K. Kim, D.R. Kang, S.K. Lim, J.Y. Kwak, M.J. Kim, Biopsy of thyroid nodules: comparison of three sets of guidelines, AJR Am. J. Roentgenol. 194 (2010) 31–37.
[10] H.J. Moon, J.M. Sung, E.K. Kim, J.H. Yoon, J.H. Youk, J.Y. Kwak, Diagnostic performance of gray-scale US and elastography in solid thyroid nodules, Radiology 262 (2012) 1002–1013.
[11] H.G. Kim, H.J. Moon, J.Y. Kwak, E.K. Kim, Diagnostic accuracy of the ultrasonographic features of subcentimeter thyroid nodules suggested by the revised American Thyroid Association guidelines, Thyroid 23 (2013) 1583–1590.
[12] H. Malhi, M.D. Beland, S.Y. Cen, et al., Echogenic foci in thyroid nodules: significance of posterior acoustic artifacts, AJR Am. J. Roentgenol. 203 (2014) 1310–1316.
[13] A. Naoun, Evaluation of ultrasonography in the diagnosis of thyroid nodules, Ann. Endocrinol. (Paris) 54 (1993) 232–234.
[14] L. Solbiati, L. Volterrani, G. Rizzatto, et al., The thyroid gland with low uptake lesions: evaluation by ultrasound, Radiology 155 (1985) 187–191.
[15] L. Leenhardt, Comments on the ultrasonographic evaluation of thyroid nodules, Ann. Endocrinol. (Paris) 54 (1993) 237–240.
[16] J.A. Bonavita, J. Mayo, J. Babb, et al., Pattern recognition of benign nodules at ultrasound of the thyroid: which nodules can be left alone? AJR Am. J. Roentgenol. 193 (2009) 207–213.
[17] M.C. Frates, C.B. Benson, J.W. Charboneau, et al., Management of thyroid nodules detected at US: Society of Radiologists in Ultrasound consensus conference statement, Radiology 237 (2005) 794–800.
[18] L.J. Bessey, N.B. Lai, N.E. Coorough, H. Chen, R.S. Sippel, The incidence of thyroid cancer by fine needle aspiration varies by age and gender, J Surj Res. 184 (2013) 761–765.
[19] J. Tourniaire, Management of the solitary thyroid nodule: clinical evaluation, Ann. Endocrinol. (Paris) 54 (1993) 226–229.
[20] C. Capelli, M. Castellano, I. Pirola, et al., Thyroid nodules shape suggests malignancy, Eur. J. Endocrinol. 155 (2006) 27–31.
[21] X.V. Nguyen, K.R. Choudhury, J.D. Eastwood, et al., Incidental thyroid nodules on CT: evaluation of 2 risk-categorization methods for work-up of nodules, AJNR Am. J. Neuroradiol. 34 (2013) 1812–1817.
[22] G. Besbes, N. Beltaief, M. Oukhaï, et al., Les facteurs prédictifs de malignité des nodules thyroïdiens à propos de 412 cas, J. Tunis. D Orl Chir. Cerv.-fac. 19 (2007) 14–18.
[23] Z. Pang, M. Margolis, R.J. Menezes, H. Maan, S. Ghai, Diagnostic performance of 2015 American Thyroid Association guidelines and inter-observer variability in assigning risk category, Eur. J. Radiol. Open 6 (2019) 122–127.
[24] S. Taghipour Zahir, F. Binesh, M. Mirouliaei, E. Khajeh, S. Noshad, Malignancy risk assessment in patients with thyroid nodules using classification and regression trees, J. Thyroid Res. 2013 (2013), https://doi.org/10.1155/2013/983953 Epub 2013 Sep 11.