


# Proposal for a novel management of indeterminate thyroid nodules on the basis of cytopathological subclasses

Martina Rossi<sup>1,2</sup> · Sabrina Lupo<sup>1</sup> · Roberta Rossi<sup>2</sup> · Paola Franceschetti<sup>2</sup> ·  
Giorgio Trasforini<sup>2</sup> · Stefania Bruni<sup>2</sup> · Federico Tagliati<sup>1,2</sup> · Mattia Buratto<sup>1</sup> ·  
Giovanni Lanza<sup>3</sup> · Luca Damiani<sup>1</sup> · Ettore degli Uberti<sup>1,2</sup> · Maria Chiara Zatelli <sup>1,2</sup>

Received: 26 May 2016 / Accepted: 25 August 2016  
© Springer Science+Business Media New York 2016

**Abstract** Indeterminate thyroid nodules include heterogeneous lesions that could benefit from a differential management. Our aim is to better define the management of the Bethesda System for Reporting Thyroid Cytopathology class III and IV nodules, by identifying cytological subcategories among Bethesda System for Reporting Thyroid Cytopathology class III associated with different clinical risk, by means of ultrasound, repeated FNAB, and *BRAFV600E* molecular analysis. We also evaluated the outcome of nodules not operated, over a 5-year follow-up. Out of 460 nodules (269 Bethesda System for Reporting Thyroid Cytopathology class III and 191 Bethesda System for Reporting Thyroid Cytopathology class IV), 344 were operated on surgical group and 116 followed-up conservatively (follow-up group). Bethesda System for Reporting Thyroid Cytopathology class III was divided into four subcategories on the basis of cytomorphological features (III-1, III-2, III-3, III-4). Clinical risk was defined on the basis of histological, cytological, and ultrasound data. Malignancy was higher in Bethesda System for Reporting Thyroid Cytopathology class III vs. Bethesda System for

Reporting Thyroid Cytopathology class IV (34.4 vs. 26.2 %;  $p < 0.01$ ). Papillary thyroid carcinoma was the most frequent cancer in each Bethesda System for Reporting Thyroid Cytopathology class (35 %). *BRAFV600E* diagnostic accuracy was 87 %. Repeated FNAB reclassified as benign nearly 40 % of nodules, selecting patients where surgery could be spared. Significant nodule growth occurred in 13.7 % of nodules, belonging mostly to Bethesda System for Reporting Thyroid Cytopathology class III-2 and Bethesda System for Reporting Thyroid Cytopathology class IV. Overall clinical risk was higher in Bethesda System for Reporting Thyroid Cytopathology III-1, III-4, and IV classes. We propose a differential management of Bethesda System for Reporting Thyroid Cytopathology III and IV classes and related subcategories: surgery may be indicated in Bethesda System for Reporting Thyroid Cytopathology class III-1, III-4, and IV; a conservative follow-up avoiding repeated FNAB may be appropriated in class III-3, while repeated FNAB may be useful in class III-2.

**Keywords** Thyroid nodule management · Indeterminate thyroid nodule · Cytological sub-categories · Thyroid nodule clinical risk

The authors Martina Rossi and Sabrina Lupo equally contributed to this work.

✉ Maria Chiara Zatelli  
ztlmch@unife.it

<sup>1</sup> Section of Endocrinology & Internal Medicine, Department of Medical Sciences, University of Ferrara, Via Ariosto 35, Ferrara 44121, Italy

<sup>2</sup> Endocrinology Unit, Azienda Ospedaliero-Universitaria di Ferrara, Via Aldo Moro 8, Ferrara 44124, Italy

<sup>3</sup> Section of Pathology and Biomolecular Diagnostics, Department of Morphology, Surgery and Experimental Medicine, University of Ferrara, Via Ariosto 35, Ferrara 44121, Italy

## Introduction

Ultrasound-guided fine-needle aspiration biopsy (US-FNAB) is a safe, accurate, and cost-effective method for pre-surgical evaluation of malignancy risk (MR) in thyroid nodules [1]. According to the Bethesda System for Reporting Thyroid Cytopathology (BSRTC) [2], FNAB identifies most nodules as benign (60–70 %) or malignant (3–7 %), directing choice of appropriate management as

indicated by current guidelines [1, 3]. However, a significant portion of FNAB samples are labeled as “non-diagnostic” (<10 %) or “indeterminate”, including BSRTC classes III, IV, and V. BSRTC class III is expected not to exceed 7 % of all cytological diagnoses [2], but real-world studies indicate that this category represents 2–20 % of the reported series [4]. These data support the evidence that BSRTC III includes a heterogeneous group of lesions with a wide range of cytomorphological features. As a consequence, malignancy rates vary from 6 to 69 % [5, 6], indicating the need for a better clinical characterization. Recent studies have evaluated the utility of subdividing BSRTC III on the basis of cytomorphological features in order to better stratify MR [7–10]. Repeated FNAB (RFNAB) may provide better information for the management of BSRTC III [2], but the utility of this approach is highly debated. The recently published American Thyroid Association (ATA) guidelines refer to RFNAB as a weak recommendation in both BSRTC III and IV, indicating this approach as an alternative to either US surveillance or surgery, depending on patient preference, clinical, and US risk factors [1]. The role of US, however, has limitations, since the few available studies report conflicting results [11–13]. The introduction of molecular markers, especially BRAF, in the work-up of thyroid nodules has increased the diagnostic power of FNAB [14–16], but this approach is not currently employed in all centers. In addition, no data are available concerning behavior and follow-up of indeterminate nodules that did not undergo surgery. Therefore, the management of indeterminate lesions, especially BSRTC III and IV, represents a significant clinical problem.

Our study aimed at a better definition of management of BSRTC III and IV thyroid nodules by identifying cytological subcategories among BSRTC class III nodules associated with different clinical risk (CR), by means of US, RFNAB, and histology, taking advantage of *BRAFV600E* molecular analysis. We also evaluated the outcome of the four BSRTC III subcategories and BSRTC IV nodules, verifying the consequences of a conservative approach over a 5-year follow-up.

## Materials and methods

### Patients

At the Section of Endocrinology of the University of Ferrara, from January 2008 to December 2009, a total of 3596 thyroid nodules, with a diameter of at least 4 mm, were assessed by US-FNAB; among these, we prospectively selected 460 BSRTC III and IV nodules (each from a single patient; 77 males and 383 females; mean age: 59.6 and 56.1 years, respectively), representing 12.8 % of all assessed nodules. After receiving patient consent, we

collected clinical (gender, age), biochemical (Thyroid Stimulating Hormone (TSH) and antibody levels, L-thyroxine replacement where applicable), and US data (nodularity, diameter, echogenicity, US structure, halo, calcifications, presence of suspicious neck lymph nodes), as well as information concerning type of surgery, where applicable, pathological (BSRTC classification and histology where applicable) and molecular data (*BRAFV600E* mutation). All patients not submitted to surgery were followed up for 5 years through monitoring of US parameters and TSH levels.

Patients with BSRTC IV lesions, large goiters, or nodules bearing the *BRAFV600E* mutation were addressed to first-line surgery as were those with BSRTC III lesions that displayed US features highly suspicious of malignancy. Surgery was also indicated for patients whose second cytology after RFNAB was referred to as BSRTC I, III, IV, V, and VI (surgical group, SG), after considering nodule size, US pattern, and patient preference. We then defined SG1 as those patients undergoing surgery after the first FNAB, and SG2 as those patients undergoing surgery after RFNAB. SG1 and SG2 patients submitted to lobectomy underwent yearly US, to monitor the contra lateral lobe.

Yearly follow-up (follow-up group, FG) was proposed for elderly patients with multiple co-morbidities, patients with nodules that displayed US features highly suggestive of benign nature, patients with a BSRTC II after RFNAB (performed within 6–12 months after the first FNAB), and patients who refused surgery. All FG patients underwent yearly thyroid hormone function assessment and US. Nodules were defined as stationary, increased or reduced, in keeping with US evidence of growth, as indicated by the ATA 2015 guidelines [1], that describe as evidence of growth a 20 % increase in at least two nodule dimensions with a minimal increase of 2 mm, corresponding to  $\geq 50$  % change in volume at US. At least one RFNAB control was indicated for FG patients with BSRTC III and IV lesions.

### US examination, US-FNAB procedure, and *BRAFV600E* mutation analysis

Each patient underwent US examination at the Section of Endocrinology of the University of Ferrara by two endocrinologists (GT, RR), using the Hi Vision Preirus US machine (Hitachi, Wallingford, CT, USA). Solid structure, marked hypoechogenicity, microcalcifications, increased intranodular vascularity, and irregular margins were considered as suspected for thyroid malignancy. In contrast, cystic US appearance was considered as predicting a benign condition. Nodules with two or more suspicious US features were considered as being associated with a higher MR, regardless of diameter.

To calculate nodule volume, the ellipsoid formula was applied. Variations in nodule volume were measured as

growth rates, calculated as  $(V2-V1)/V1$ , where  $V1$  is the nodule volume assessed at the first US, and  $V2$  the nodule volume at the last US. Using the ellipsoid method, the inter-observer variability for thyroid nodule volume measurement was 48.9 % (95 % CI, 39.8–58.6 %), as previously reported [17].

FNABs were performed under US guidance with a 22-gauge needle attached to a 10 ml syringe. Each lesion was aspirated one to three times. The aspirate was smeared on glass slides and sent to an experienced cytopathologist for cytology analysis. The remaining material in the needle was washed out with 5 ml of normal saline into a collection tube, obtaining enough material for both cytology and molecular analysis. The criterion for an adequate smear was the presence of six groups of cells with  $>10$  cells/group [18]. *BRAFV600E* mutation analysis was performed as previously described [16].

Cytological (following BSRTC indications) and histological evaluations were performed independently by two pathologists at the Section of Pathology and Biomolecular Diagnostics, Dept. of Morphology, Surgery and Experimental Medicine, University of Ferrara, with an inter-operator and intra-operator variability of ~20 %.

In order to obtain a detailed analysis of the variability in clinical behavior of BSRTC III, we subcategorized this class on the basis of cytomorphological features, applying the definitions provided by Cibas [2], simplified by modification into four subcategories as follows:

- (1) Lesions with nuclear atypia suggestive of papillary thyroid carcinoma (PTC) including cells with intranuclear pseudoinclusions and/or nuclear grooves, rare atypical follicular cells with nuclear enlargement, crowding, and pallor; thyroid aspirates containing rare follicular cells and malignant-appearing large lymphocytes with mild nuclear atypia (III-1, including E, G, and H Cibas subcategories).
- (2) Lesions with a predominant microfollicular pattern having a predominant population of microfollicles with scant colloid that did not match the criteria for category IV (III-2, including Cibas subcategory A).
- (3) Lesions with an oncocyctic pattern having a predominance of Hurtle cells or atypical cyst lining cells without other supporting features of PTC (III-3, including B, D, and F Cibas subcategories).
- (4) Lesions not otherwise specified, including drying artefacts, obscuring blood samples, or scant atypical cells (III-4, including C and I Cibas subcategories).

#### Definition of malignancy rate, MR, and CR

We defined malignancy rate as the ratio between the number of thyroid cancers and the number of patients submitted

to surgery, and MR as the ratio between the number of thyroid cancers and the total number of patients. CR of each BSRTC class was calculated as follows:  $(A + B + C)/\text{total nodules number}$ , where A is the number of nodules with malignant histology, B is the number of nodules not undergoing surgery with a diagnosis which is different from BSRTC II after RFNAB, and C is the number of nodules that grew significantly and/or acquired US suspicious features during follow-up.

#### Statistical analysis

Data analysis was performed by calculating the odds ratio (OR) to assess the presence of an association of clinical-US features with pathology findings and with volumetric nodule changes. Statistical significance was assessed by the chi-square ( $\chi^2$ ) test or Fisher's exact test. A  $p$ -value  $<0.05$  was considered significant. The paired Student  $t$ -test was employed to compare mean  $V1$  and  $V2$ . Sensitivity, specificity, positive predictive value (PPV), negative predictive value (NPV), and accuracy of the molecular test were calculated. The 95 % CI for the diagnostic value was calculated.

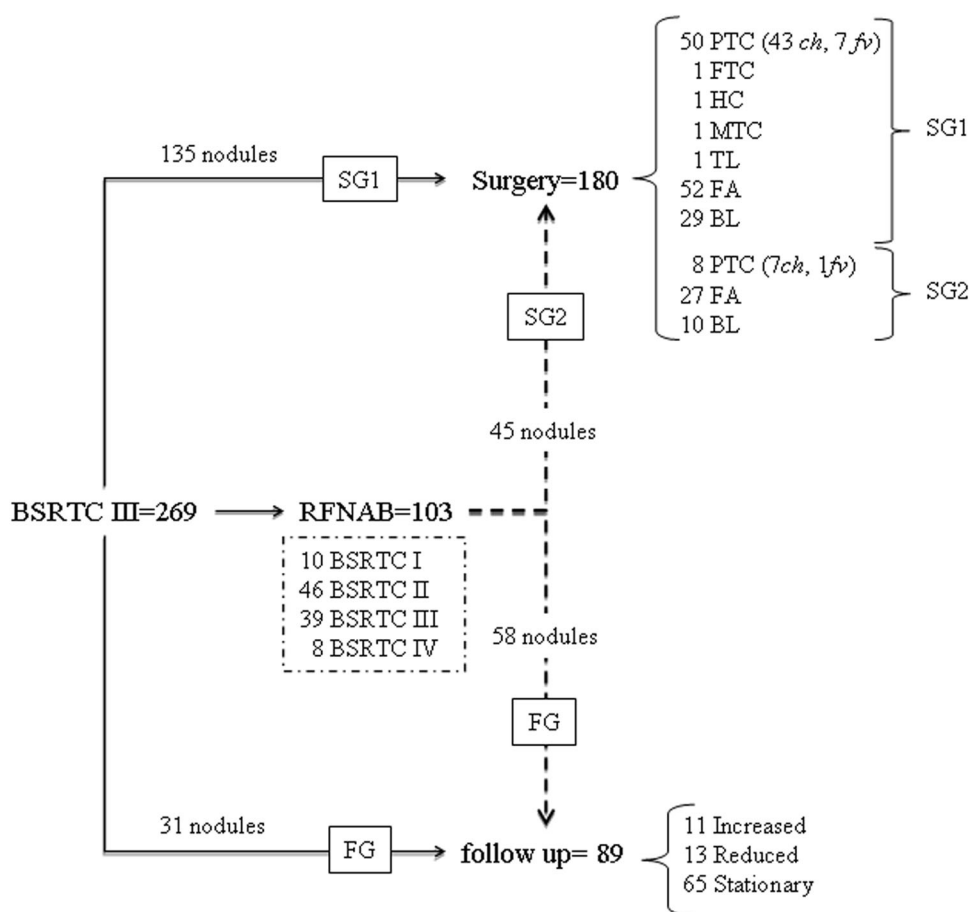
#### Results

A total of 460 patients were enrolled in this study: 269 had a nodule with cytological diagnosis of BSRTC III (III-1 = 46; III-2 = 140; III-3 = 76; III-4 = 7) and 191 of BSRTC IV; 103 patients with BSRTC III nodules and 25 patients with BSRTC IV nodules underwent RFNAB. Patients were then divided into SG1, SG2, and FG.

#### *BRAF* molecular analysis

*BRAFV600E* mutation was found in 30 samples (18 BSRTC III and 12 BSRTC IV). Among BSRTC III, mutated cases almost exclusively belonged to subcategory III-1 (17/18 cases, 94.4 %) and only one case to III-2. Overall, the incidence of *BRAFV600E* mutation was similar in BSRTC III and IV (6.7 and 6.3 %, respectively), also among PTC (33.3 vs. 38.7 %, respectively). The presence of the *BRAFV600E* mutation allowed us to reclassify as malignant 36.9 % of BSRTC III-1 and 0.7 % of BSRTC III-2 nodules. *BRAFV600E* was significantly associated with PTC with classical histology (*ch*) (27/30 cases;  $p < 0.01$ ). Diagnostic accuracy of the molecular analysis was high in both BSRTC III and IV (86.6 and 87.9 %, respectively), with NPV of 85.7 % [95 % CI, 77–94 %] and 87.1 % [95 % CI, 77–97 %], respectively. PPV and specificity were 100 % in both classes. Sensitivity was low in both BSRTC III

**Fig. 1** Management and outcome of the 269 nodules diagnosed as BSRTC III. BSRTC Bethesda System for Reporting Thyroid Cytopathology, RFNAB repeated FNAB, SG1 patients operated in after the first FNAB, SG2 patients operated in after RFNAB, FG follow-up group, PTC papillary thyroid carcinoma, *ch* classical histology, *fv* follicular variant, FTC follicular thyroid carcinoma, HC Hurtle carcinoma, MTC medullary thyroid carcinoma, TL thyroid lymphoma, FA follicular adenoma, BL benign lesion



(33.3 % [95 % CI, 21–45 %]) and BSRTC IV (34.1 % [95 % CI, 19–49 %]).

### Repeated FNAB

As shown in Figs. 1 and 2, RFNAB led to a benign cytology in 44.6 % of BSRTC III nodules, mostly represented by BSRTC III-2 and III-3 nodules, and in 32 % of BSRTC IV nodules. RFNAB confirmed an indeterminate diagnosis in nearly 50 % of these nodules. In addition, >10 % were not diagnostic after RFNAB, while only 1 BSRTC IV nodule was re-classified as cytologically malignant. Malignancy rate varied widely among BSRTC III subcategories, ranging from 15.1 % in III-3 to 80 % in III-4. In addition, malignancy rate decreased in SG2 group (i.e., patients addressed to surgery after RFNAB) as compared to SG1 group in all BSRTC III subcategories, except for III-4 (Table 1), indicating that RFNAB may be useful to select patients in which surgery may be spared, especially in BSRTC III-2 subcategory. The impact of RFNAB in the other subcategories seems to be negligible. Nodules with non-diagnostic or indeterminate RFNAB displayed nearly the

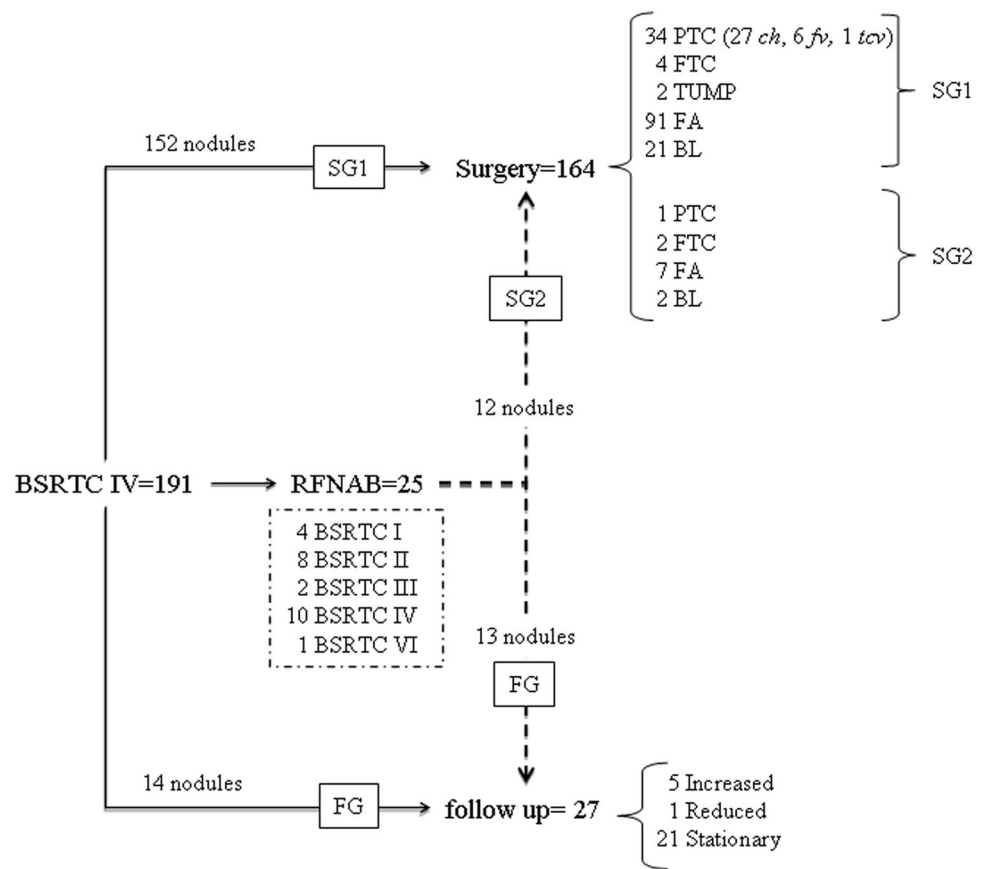
same malignancy rate as compared to nodules with a single indeterminate diagnosis.

### SG patients outcome

As shown in Figs. 1 and 2, overall, 344 patients were operated on (SG). Patient and nodule characteristics are described in Table 2. Among these patients, 132 underwent total thyroidectomy (T), 212 lobectomy (L), and 34 lobectomy followed by completion thyroidectomy. Overall malignancy rate was 30.5 % (105/344) and was significantly higher in BSRTC III as compared to BSRTC IV (34.4 vs. 26.2 %, respectively;  $p < 0.01$ ). Compared with BSRTC III, BSRTC IV contained a significantly greater number of lesions with a follicular pattern, including follicular thyroid carcinoma (FTC), tumors of undetermined malignant potential (TUMP) (4.9 % in BSRTC IV vs. 1.1 % in BSRTC III;  $p < 0.05$ ), and follicular adenoma (FA) (59.8 vs. 43.8 %;  $p < 0.01$ ).

Yearly US follow-up of the 212 patients submitted to L (93 BSRTC III and 119 BSRTC IV) showed no change, except for a single patient, previously diagnosed with FA. This patient displayed a significant growth of the contra

**Fig. 2** Management and outcome of the 191 nodules diagnosed as BSRTC IV. *BSRTC* Bethesda System for Reporting Thyroid Cytopathology, *RFNAB* repeated FNAB, *SG1* patients operated in after the first FNAB, *SG2* patients operated in after RFNAB, *FG* follow-up group, *PTC* papillary thyroid carcinoma, *ch* classical histology, *fv* follicular variant, *tcv* tall cell variant, *ov* oncocytic variant, *FTC* follicular thyroid carcinoma, *HC* Hurtle carcinoma, *TUMP* tumors of undetermined malignant potential, *FA* follicular adenoma, *BL* benign lesion



**Table 1** MR and malignancy rate in SG1 and in SG2 according to BSRTC

	MR (%)	Malignancy rate (%)		
	Total	Total	SG1	SG2
BSRTC III	62/269 (23)	62/180 (34.4)	54/135 (40)	8/45 (17.8)
III-1	26/46 (56.5)	26/40 (65)	25/34 (73.5)	1/6 (16.6)
III-2	27/140 (19.3)	27/102 (26.5)	21/73 (28.8)	6/29 (20.7)
III-3	5/76 (6.6)	5/33 (15.1)	5/24 (20.8)	0/9 (0)
III-4	4/7 (57.1)	4/5 (80)	3/4 (75)	1/1 (100)
BSRTC IV	43/191 (22.5)	43/164 (26.2)	40/152 (26.3)	3/12 (25)

*SG1* patients undergoing surgery after the first FNAB, *SG2* patients undergoing surgery after RFNAB, *BSRTC* Bethesda System for Reporting Thyroid Cytopathology

lateral lobe and pathological neck lymph nodes, consistent with a FTC diagnosis.

**FG patients outcome**

As shown in Figs. 1 and 2, overall, 116 patients underwent a 5-year conservative follow-up (FG) and nearly half were submitted at least once to RFNAB. During follow-up, 87 nodules (75%) remained stable and 13 shrank (11.2%). Significant nodule growth occurred in 16 nodules (13.7% [95% CI, 7.5–20%]): 11/89 BSRTC III nodules (12.3%

[95% CI, 5.5–19.2%]) and 5/27 BSRTC IV (18.5% [95% CI, 3.8–33.1%]), respectively. Growing nodules almost exclusively belonged to subcategory III-2 (81.8%), while those that had shrunk mostly belonged to subcategory III-3 (84.6%). At baseline, the mean volume was  $1.3 \pm 0.4$  ml in BSRTC III and  $2.5 \pm 1.5$  ml in BSRTC IV (V1); at the end of follow-up, volume was  $3.7 \pm 2.4$  ml in BSRTC III and  $13.4 \pm 12.7$  ml in BSRTC IV (V2), with a mean change in the largest diameter of 6.6 and 7.7 mm, respectively. Therefore, the mean increase in volume was 2-fold and 4-fold, respectively.



**Table 2** SG patients clinical features and nodules US characteristics, according to BSRTC class and final histology

	BSRTC III (n 180)		BSRTC IV (n 164)	
	Malignant (n 62)	Benign (n 118)	Malignant (n 43)	Benign (n 121)
Age				
<45 years	11.3 %	14.4 %	25.6 %	23.2 %
>45 years	88.7 %	85.6 %	74.4 %	76.8 %
Gender				
Male	1.6 %	4.2 %	30.2 %	22.3 %
Female	98.4 %	95.8 %	69.8 %	77.7 %
Nodularity				
Single	46.7 %	59.3 %	55.8 %	57.8 %
Multiple	53.3 %	40.7 %	44.2 %	42.2 %
Diameter				
≤5 mm	32.3 %	11.9 %	27.7 %	12.4 %
5–10 mm	32.3 %	22.1 %	34.7 %	19.9 %
10–20 mm	25.8 %	31.3 %	18.8 %	28.9 %
>20 mm	9.6 %	34.7 %	18.8 %	38.8 %
Echogenicity				
Hypoechoic	56.5 %	52.6 %	72 %	76.9 %
Isoechoic	38.7 %	46.6 %	25.7 %	20.7 %
Hyperechoic	1.6 %	0.8 %	2.3 %	0.8 %
Anaechoic	3.2 %	0 %	0 %	1.6 %
US structure				
Solid	95.1 %	90.7 %	95.4 %	87.6 %
Mixed	4.9 %	7.6 %	4.6 %	10.7 %
Cystic	0 %	1.7 %	0 %	1.7 %
US features				
Halo	34. %	76.2 %	5.9 %	66.4 %
Irregular margins	12 %	4.2 %	21 %	2.1 %
Micro-calcifications	30.6 %	10.1 %	4.6 %	2.5 %
Intranodular vascularity	22.2 %	11.3 %	24 %	7 %
Thyroiditis	27.4 %	29.6 %	11.6 %	19 %
L-thyroxine				
Suppressive	0 %	5.1 %	2.3 %	5.8 %
Replacement	4.8 %	5.9 %	2.3 %	6.6 %
No therapy	95.2 %	89 %	95.4 %	87.6 %

SG surgical group, US ultrasound, BSRTC Bethesda System for Reporting Thyroid Cytopathology

Over time, US characteristics did not change; in particular, growing nodules did not acquire US characteristics suspicious of malignancy, but BSRTC IV nodules were larger and grew significantly more than BSRTC III nodules. Three patients with growing nodules underwent surgery and turned out to be BL.

A significant volume reduction was observed in 14 nodules (12.6 % [95 % CI, 6.1–18 %]): 13/89 BSRTC III nodules (14.6 % [95 % CI, 8.2–21 %]) and 1/27 BSRTC IV nodules (3.7 % [95 % CI, 0.3–7.1 %]), respectively. At baseline, the mean volume was  $2.9 \pm 1.17$  ml (V1) and at the end of follow-up was  $0.14 \pm 0.072$  ml (V2).

L-thyroxine was prescribed for 34 FG patients, for replacement purposes in 25 cases (21 BSRTC III and 4 BSRTC IV) and to suppress TSH in 9 cases (4 BSRTC III and 5 BSRTC IV). L-thyroxine treatment did not affect nodule volume, in keeping with the finding that TSH levels overlapped in patients displaying an increased or a reduced nodule volume (1.5  $\mu$ UI/ml vs. 1.4  $\mu$ UI/ml, respectively). Concerning clinical and US characteristics, no significant difference was found either between BSRTC III and IV class or between FG and SG patients.

### Clinical risk

To better define the management of indeterminate nodules and select those cases needing a more stringent follow-up, we calculated CR in each BSRTC and related subcategories (Table 3). The highest CR was found in BSRTC III-4 and BSRTC III-1 nodules. The results obtained by calculating CR were then taken into account to draw up an algorithm suggesting a novel management on the basis of BSRTC III subcategories (Fig. 3).

### Clinical correlations

We then investigated whether clinical and US features may associate with malignancy and/or nodule growth. Malignancy was significantly associated with nodule size <1 cm in both BSRTC III (OR 3.15;  $p < 0.01$ ) and BSRTC IV (OR 3.6;  $p < 0.01$ ), while it significantly correlated with micro-calcifications only in BSRTC III (OR 4.5;  $p < 0.01$ ). Nodule growth was associated with the presence of multinodular goiter in BSRTC III (OR 33.4;  $p < 0.01$ ). Conversely, nodule size <1 cm (OR 0.2;  $p < 0.05$ ) and hypoechoogenicity (OR 0.2;  $p < 0.05$ ) were protective toward nodule growth in BSRTC III.

### Management proposal for BSRTC III and IV nodules

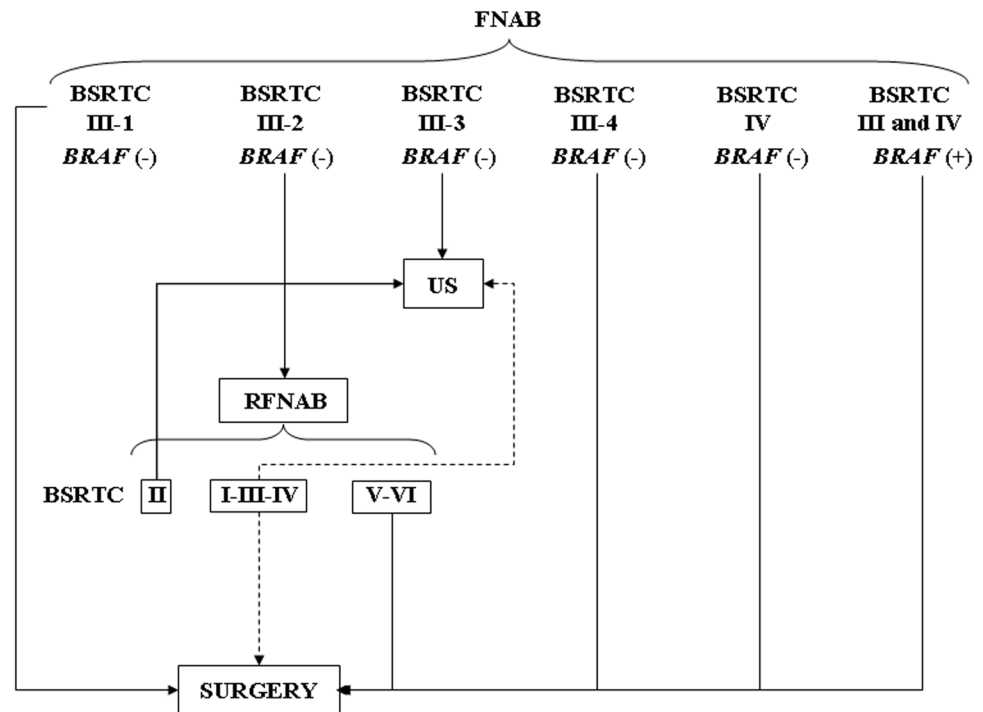
On the basis of our results, we propose a differential management of indeterminate thyroid nodules, as depicted in Fig. 3. Nodules displaying BSRTC III-1, III-4, and BSRTC IV with a negative *BRAFV600E* analysis may be addressed to surgery after the first FNAB, since their CR is high. On the contrary, BSRTC III-3 nodules with US features not suspicious of cancer and a negative *BRAFV600E* analysis should be addressed to conservative follow-up with yearly US. In BSRTC III-2, since CR is not negligible (~30 %),

**Table 3** CR and MR according to BSRTC class

Total patients (n)	Criteria (n)	CR (%)	MR (%)
BSRTC III (n 269)	Histology (62) RFNAB (13) US (11)	86/269 (32)	62/269 (23)
III-1 (n 46)	Histology (26) RFNAB (0) US (0)	26/46 (56.5)	26/46 (56.5)
III-2 (n 140)	Histology (27) RFNAB (7) US (9)	43/140 (30.7)	27/140 (19.3)
III-3 (n 76)	Histology (5) RFNAB (5) US (2)	12/76 (15.8)	5/76 (6.6)
III-4 (n 7)	Histology (4) RFNAB (1) US (0)	5/7 (71.4)	4/7 (57.1)
BSRTC IV (n 191)	Histology (43) RFNAB (6) US (5)	54/191 (28.3)	43/191 (22.5)

BSRTC Bethesda System for Reporting Thyroid Cytopathology, CR clinical risk, MR malignancy risk, RFNAB repeated fine-needle aspiration biopsy, US ultrasound

**Fig. 3** Proposed algorithm for differential management of BSRTC III and IV nodules and related subcategories. BSRTC Bethesda System for Reporting Thyroid Cytopathology, RFNAB repeated FNAB, US ultrasonography, BRAF negative (-) or positive (+) BRAFV600E analysis



RFNAB is indicated to identify nodules with a lower risk of being a cancer in the absence of the BRAFV600E mutation.

## Discussion

To date, clinical management of indeterminate thyroid nodules is a very controversial issue. Indeed, cytological interpretation of indeterminate nodules is subjective, as witnessed by the significant inter-observer and intra-observer disagreement [19–22]. In order to simplify this issue, we clustered the nine cytological patterns of BSRTC III described by Cibas [2] into four subcategories, on the basis of cytological pattern similarity and MR, as suggested by Mathur et al. [21]. In this prospective clinical study, we evaluated the outcome of the four BSRTC III subcategories and BSRTC IV nodules (undergoing surgery or not) and

explored the contribution of BRAFV600E analysis, RFNAB, and US in the therapeutic decision. Moreover, we verified the consequences of a conservative approach during a mean follow-up of 5 years.

Therefore, we proposed an algorithm for the management of BSRTC III and IV nodules, based on the results obtained after calculating CR for each BSRTC class. CR was considered as a parameter that, taking into account also the nodules not submitted to surgery, may help in the management of each BSRTC class and related subcategories depending on MR, second cytology, and nodule changes. Our data support the indication for surgery as first-line treatment for subcategory III-1 and III-4, due to high CR. Conversely, subcategory III-2, displaying a moderate but not negligible CR (~30%), requires a RFNAB to plan an appropriate management. The latter may include conservative follow-up when second cytology displays a

benign result. On the other hand, in the case of BSRTC I, III, and IV after RFNAB, the decision between surgery and follow-up should be made by taking into account clinical and US risk factors, as well as patient preference. Finally, since BSRTC III-3 displays a low CR, our data suggest that a conservative follow-up may be appropriate, avoiding RFNAB. As for BSRTC IV, although CR is <30 %, our results indicate surgery as an appropriate management choice since the incidence of FTC and FA in this category is high. For patients who refuse surgery or with US features suggestive of a benign lesion, RFNAB and molecular analysis may be considered, in keeping with current guidelines [1].

Little investigation has been carried out regarding the outcome of nodules not submitted to surgery. US is the first-line method for thyroid nodule follow-up. Several US features have been found to be good predictors of malignancy. In our experience, none of these US features alone is 100 % specific, but a combination of at least two of them is associated with a higher likelihood of malignancy [15]. Many attempts have been made recently to classify the nodules on the basis of US features, as also indicated by the current ATA guidelines. In this respect, US features were indirectly taken into account in the calculation of CR in the present study, since the only variation that was prospectively detected was related to nodule diameter.

The growth of benign thyroid nodules is rare and not usually associated with malignancy [23]. On the other hand, growing BSRTC IV nodules display an increased MR [24]. In our settings, nodule growth was an infrequent occurrence that more often involved BSRTC III-2 and BSRTC IV nodules. Although these nodules did not acquire US characteristics suspicious of malignancy, BSRTC IV nodules grew more than BSRTC III nodules, often causing compressive effects. Therefore, we suggest lobectomy in patients with BSRTC IV nodules. Furthermore, in BSRTC III nodule size <1 cm and hypoechoogenicity were protective features toward nodule growth. The latter, on the contrary, was significantly associated with goiter.

Finally, l-thyroxine therapy did not seem to be useful in controlling nodule growth, since the latter did not correlate with TSH levels. Our findings are not in accordance with recent literature [25], which reports that L-thyroxine treatment associates with a significant reduction in nodule volume. This discrepancy may be due to the small number of patients treated with suppressive L-thyroxine doses in our study.

Overall, in our analysis BSRTC III and IV lesions were often benign, suggesting the need to adopt a conservative approach, where possible. In our settings, malignancy rate was higher in BSRTC III as compared to BSRTC IV, in keeping with previous reports [26, 27]. However, malignancy rate varied widely in BSRTC III, ranging from

15.1 % in BSRTC III-1 to 80 % in BSRTC III-4. PTC was the most frequent cancer diagnosed in both BSRTC and related subcategories, with a higher prevalence of PTC *ch*. This finding partly explains the high incidence of *BRAFV600E* mutation also in BSRTC IV. Moreover, as expected, we found a higher incidence of lesions with a follicular pattern in BSRTC IV as compared to BSRTC III. Our data confirm, once again, the diagnostic value of *BRAFV600E* analysis, demonstrating that it may be useful to identify PTC, especially among BSRTC III-1 nodules that were reclassified as malignant in 36.9 % of cases on the basis of molecular results. However, *BRAFV600E* analysis has a low sensitivity, which highlights the need for further development of new molecular tests with a better NPV [28, 29]. In our experience, RFNAB allowed us to reclassify as benign ~40 % of BSRTC III and IV cases, in line with previous results [30, 31] showing that RFNAB is able to assign a definitive diagnosis to the majority of indeterminate lesions. None of the lesions that were benign after RFNAB was a cancer at histology and most of them remained stationary during follow-up. Cases confirmed as indeterminate after RFNAB did not display a significant increase in MR, but had a higher tendency to grow compared to benign nodules. In addition, nodules with a nondiagnostic cytology after RFNAB displayed a higher MR as compared to nodules with a single indeterminate diagnosis. Overall, our data suggest that RFNAB may be used in BSRTC III-2 subcategory to select patients with lower MR, where diagnostic surgery could be avoided in favor of a conservative management. Our data further indicate that RFNAB is not useful in the management of BSRTC III subclasses which are different from III-2, while US allows us to monitor all indeterminate nodules and detect early changes.

US follow-up of nodules with an indeterminate cytology is advisable, since it has been reported that also nodules with a benign cytology after RFNAB have a higher MR as compared to nodules with a single benign cytology [32]. US alone cannot be considered a diagnostic tool for cancer, since none of the US characteristics accurately predicts malignancy, as also reported by Frasoldati et al. [33]. On the other hand, the combination of several suspicious characteristics increases nodule MR [34].

In conclusion, management of the BSRTC III and BSRTC IV nodules could be diversified on the basis of cytological subcategories, *BRAFV600E* analysis, US features, and RFNAB. Calculating CR for each BSRTC class and related subcategories helped to obtain the results that, in turn, allowed us to draw the proposed management algorithm.

If we had applied this approach to our series, 74 patients might have been followed up exclusively with US, avoiding 34 diagnostic surgeries and 31 RFNAB, impacting the clinical management in 16 % of the patients, with a consequent reduction in patient and healthcare system costs.



**Acknowledgments** This work was supported by grants from the Italian Ministry of Education, Research and University (FIRB RBAP11884 M, RBAP1153LS), Associazione Italiana per la Ricerca sul Cancro (AIRC) in collaboration with “Laboratorio in rete del Tecnopolo Tecnologie delle terapie avanzate” (LTTA) of the University of Ferrara.

#### Compliance with ethical standards

**Conflict of interest** MCZ has received consultant fees from Novartis and Genzyme. EdU has received consultant fees from Novartis and grant support from Sanofi. The remaining authors declare that they have no conflict of interest.

#### References

- B.R. Haugen, E.K. Alexander, K.C. Bible, G.M. Doherty, S.J. Mandel, Y.E. Nikiforov, F. Pacini, G.W. Randolph, A.M. Sawka, M. Schlumberger, K.G. Schuff, S.I. Sherman, J.A. Sosa, D.L. Steward, R.M. Tuttle, L. Wartofsky, 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*. **26**, 1–133 (2016). doi:[10.1089/thy.2015.0020](https://doi.org/10.1089/thy.2015.0020)
- E.S. Cibas, S.Z. Ali, NCI Thyroid FNA state of the science conference. The Bethesda System for Reporting Thyroid Cytopathology. *Am. J. Clin. Pathol.* **132**, 658–665 (2009). doi:[10.1309/AJCPHLM3JV4LA](https://doi.org/10.1309/AJCPHLM3JV4LA)
- H. Gharib, E. Papini, R. Paschke, D.S. Duick, R. Valcavi, L. Hegedüs, P. Vitti, AACE/AME/ETA task force on thyroid nodules. American Association of Clinical Endocrinologists, Associazione Medici Endocrinologi, and European Thyroid Association Medical guidelines for clinical practice for the diagnosis and management of thyroid nodules: executive summary of recommendations. *Endocr. Pract.* **16**, 468–475 (2010)
- N.P. Ohori, K.E. Schoedel, Variability in the atypia of undetermined significance/follicular lesion of undetermined significance diagnosis in the Bethesda System for Reporting Thyroid Cytopathology: sources and recommendations. *Acta Cytol.* **55**, 492–498 (2011). doi:[10.1159/000334218](https://doi.org/10.1159/000334218)
- H.H. Wu, C. Rose, T.M. Elsheikh, The Bethesda System for Reporting Thyroid Cytopathology: an experience of 1,382 cases in a community practice setting with the implication for risk of neoplasm and risk of malignancy. *Diagn. Cytopathol.* **40**, 399–403 (2012). doi:[10.1002/dc.21754](https://doi.org/10.1002/dc.21754)
- J.H. Park, S.O. Yoon, E.J. Son, H.M. Kim, J.H. Nahm, S. Hong, Incidence and malignancy rates of diagnoses in the Bethesda System for Reporting Thyroid Aspiration Cytology: an institutional experience. *Korean J. Pathol.* **48**, 133–139 (2014). doi:[10.4132](https://doi.org/10.4132)
- M. Bongiovanni, J.F. Krane, E.S. Cibas, W.C. Faquin, The atypical thyroid fine-needle aspiration: past, present, and future. *Cancer Cytopathol.* **120**, 73–86 (2012). doi:[10.1002/cncy.20178](https://doi.org/10.1002/cncy.20178)
- H.J. Park, J.H. Moon, C.K. Yom, K.H. Kim, J.Y. Choi, S.I. Choi, S.H. Ahn, W.J. Jeong, W.W. Lee, S.Y. Park, Thyroid “atypia of undetermined significance” with nuclear atypia has high rates of malignancy and BRAF mutation. *Cancer Cytopathol.* **122**, 512–520 (2014). doi:[10.1002/cncy.21411](https://doi.org/10.1002/cncy.21411)
- A.A. Renshaw, Should “atypical follicular cells” in thyroid fine-needle aspirates be subclassified?. *Cancer Cytopathol.* **118**(4), 186–189 (2010). doi:[10.1002/cncy.20091](https://doi.org/10.1002/cncy.20091)
- A.A. Renshaw, Subclassification of atypical cells of undetermined significance in direct smears of fine-needle aspirations of the thyroid: distinct patterns and associated risk of malignancy. *Cancer Cytopathol.* **119**, 322–327 (2011). doi:[10.1002/cncy.20154](https://doi.org/10.1002/cncy.20154)
- R.S. Mehta, S.E. Carty, N.P. Ohori, S.P. Hodak, C. Coyne, S.O. LeBeau, M.E. Tublin, M.T. Stang, J.T. Johnson, K.L. McCoy, M. N. Nikiforova, Y.E. Nikiforov, L. Yip, Nodule size is an independent predictor of malignancy in mutation-negative nodules with follicular lesion of undetermined significance cytology. *Surgery*. **154**, 730–738 (2013). doi:[10.1016/j.surg.2013.05.015](https://doi.org/10.1016/j.surg.2013.05.015)
- S.M. Khoncarly, S.W. Tamarkin, C.R. McHenry, Can ultrasound be used to predict malignancy in patients with a thyroid nodule and an indeterminate fine-needle aspiration biopsy? *Surgery*. **56**, 967–970 (2014). doi:[10.1016/j.surg.2014.06.043](https://doi.org/10.1016/j.surg.2014.06.043)
- W.S. Yoo, H.S. Choi, S.W. Cho, J.H. Moon, K.W. Kim, H.J. Park, S.Y. Park, S.I. Choi, S.H. Choi, S. Lim, K.H. Yi, J. Park do, H.C. Jang, Y.J. Park, The role of ultrasound findings in the management of thyroid nodules with atypia or follicular lesions of undetermined significance. *Clin. Endocrinol. (Oxf.)*. **80**, 735–742 (2014). doi:[10.1111/cen.12348](https://doi.org/10.1111/cen.12348)
- Y.E. Nikiforov, S.E. Carty, S.I. Chiose, C. Coyne, U. Duvvuri, R.L. Ferris, W.E. Gooding, S.P. Hodak, S.O. LeBeau, N.P. Ohori, R.R. Seethala, M.E. Tublin, L. Yip, M.N. Nikiforova, Highly accurate diagnosis of cancer in thyroid nodules with follicular neoplasm/suspicious for a follicular neoplasm cytology by ThyroSeq v2 next-generation sequencing assay. *Cancer*. **120**, 3627–3634 (2014). doi:[10.1002/cncr.29038](https://doi.org/10.1002/cncr.29038)
- M. Rossi, M. Buratto, S. Bruni, C. Filieri, F. Tagliati, G. Trasforini, R. Rossi, M.D. Beccati, E.C. Degli Uberti, M.C. Zatelli, Role of ultrasonographic/clinical profile, cytology, and BRAF V600E mutation evaluation in thyroid nodule screening for malignancy: a prospective study. *J. Clin. Endocrinol. Metab.* **97**, 2354–2361 (2012). doi:[10.1210/jc.2011-3494](https://doi.org/10.1210/jc.2011-3494)
- M. Rossi, M. Buratto, F. Tagliati, R. Rossi, S. Lupo, G. Trasforini, G. Lanza, P. Franceschetti, S. Bruni, E. Degli Uberti, M.C. Zatelli, Relevance of BRAF(V600E) mutation testing versus RAS point mutations and RET/PTC rearrangements evaluation in the diagnosis of thyroid cancer. *Thyroid*. **25**, 221–228 (2015). doi:[10.1089/thy.2014.0338](https://doi.org/10.1089/thy.2014.0338)
- V.F. Brauer, P. Eder, K. Miehle, T.D. Wiesner, H. Hasenclever, R. Paschke, Interobserver variation for ultrasound determination of thyroid nodule volumes. *Thyroid* **15**, 1169–1175 (2005)
- M.C. Zatelli, G. Trasforini, S. Leoni, G. Frigato, M. Buratto, F. Tagliati, R. Rossi, L. Cavazzini, E. Roti, E.C. degli Uberti, BRAF V600E mutation analysis increases diagnostic accuracy for papillary thyroid carcinoma in fine-needle aspiration biopsies. *Eur. J. Endocrinol.* **161**, 467–473 (2009). doi:[10.1530/EJE-09-0353](https://doi.org/10.1530/EJE-09-0353)
- E.S. Cibas, Z.W. Baloch, G. Fellegara, V.A. LiVolsi, S.S. Raab, J. Rosai, J. Diggans, L. Friedman, G.C. Kennedy, R.T. Kloos, R.B. Lanman, S.J. Mandel, N. Sindy, D.L. Steward, M.A. Zeiger, B.R. Haugen, E.K. Alexander, A prospective assessment defining the limitations of thyroid nodule pathologic evaluation. *Ann. Intern. Med.* **159**, 325–332 (2013). doi:[10.7326/0003-4819-159-5-201309030-00006](https://doi.org/10.7326/0003-4819-159-5-201309030-00006)
- R. Gerhard, G. da Cunha Santos, Inter- and intraobserver reproducibility of thyroid fine needle aspiration cytology: an analysis of discrepant cases. *Cytopathology* **18**, 105–111 (2007)
- A. Mathur, A. Najafian, E.B. Schneider, M.A. Zeiger, M.T. Olson, Malignancy risk and reproducibility associated with atypia of undetermined significance on thyroid cytology. *Surgery*. **156**, 1471–1476 (2014). doi:[10.1016/j.surg.2014.08.026](https://doi.org/10.1016/j.surg.2014.08.026)
- M.T. Olson, T. Boonyarunnate, P. Aragon Han, C.B. Umbricht, S.Z. Ali, M.A. Zeiger, A tertiary center’s experience with second review of 3885 thyroid cytopathology specimens. *J. Clin. Endocrinol. Metab.* **98**, 1450–1457 (2013). doi:[10.1210/jc.2012-3898](https://doi.org/10.1210/jc.2012-3898)
- C. Durante, G. Costante, G. Lucisano, R. Bruno, D. Meringolo, A. Paciaroni, E. Puxeddu, M. Torlontano, S. Tumino, M. Attard, L.

- Lamartina, A. Nicolucci, S. Filetti, The natural history of benign thyroid nodules. *J. Am. Med. Assoc.* **313**, 926–935 (2015). doi:[10.1001/jama.2015.0956](https://doi.org/10.1001/jama.2015.0956)
24. H. Nakamura, M. Hirokawa, H. Ota, M. Kihara, A. Miya, A. Miyachi Is an increase in thyroid nodule volume a risk factor for malignancy? *Thyroid.* **25**, 804–811 (2015). doi:[10.1089/thy.2014.0567](https://doi.org/10.1089/thy.2014.0567)
25. E. Bandeira-Echtler, K. Bergerhoff, B. Richter, Levothyroxine or minimally invasive therapies for benign thyroid nodules. *Cochrane Database Syst. Rev.* Jun 18;(6):CD004098 (2014). doi:[10.1002/14651858](https://doi.org/10.1002/14651858)
26. J.T. Broome, C.C. Solorzano, The impact of atypia/follicular lesion of undetermined significance on the rate of malignancy in thyroid fine-needle aspiration: evaluation of the Bethesda System for Reporting Thyroid Cytopathology. *Surgery.* **150**, 1234–1241 (2011). doi:[10.1016/j.surg.2011.09.006](https://doi.org/10.1016/j.surg.2011.09.006)
27. W.C. Faquin, Z.W. Baloch, Fine-needle aspiration of follicular patterned lesions of the thyroid: diagnosis, management, and follow-up according to National Cancer Institute (NCI) recommendations. *Diagn. Cytopathol.* **38**, 731–739 (2010). doi:[10.1002/dc.21292](https://doi.org/10.1002/dc.21292)
28. Y.E. Nikiforov, S.E. Carty, S.I. Chiosea, C. Coyne, U. Duvvuri, R.L. Ferris, W.E. Gooding, S.O. LeBeau, N.P. Otori, R.R. Seethala, M.E. Tublin, L. Yip, M.N. Nikiforova, Impact of the multi-gene ThyroSeq next-generation sequencing assay on cancer diagnosis in thyroid nodules with atypia of undetermined significance/follicular lesion of undetermined significance cytology. *Thyroid.* **25**, 1217–1223 (2015). doi:[10.1089/thy.2015.0305](https://doi.org/10.1089/thy.2015.0305)
29. S. Beaudenon-Huibregtse, E.K. Alexander, R.B. Guttler, J.M. Hershman, V. Babu, T.C. Blevins, P. Moore, B. Andruss, E. Labourier, Centralized molecular testing for oncogenic gene mutations complements the local cytopathologic diagnosis of thyroid nodules. *Thyroid.* **24**, 1479–1487 (2014). doi:[10.1089/thy.2013.0640](https://doi.org/10.1089/thy.2013.0640)
30. R. Gerhard, S.L. Boerner, Evaluation of indeterminate thyroid cytology by second-opinion diagnosis or repeat fine-needle aspiration: which is the best approach? *Acta Cytol.* **59**, 43–50 (2015). doi:[10.1159/000369332](https://doi.org/10.1159/000369332)
31. R. Gerhard, S.L. Boerner, The value of second opinion in thyroid cytology: a review. *Cancer Cytopathol.* **122**, 611–619 (2014). doi:[10.1002/cncy.21436](https://doi.org/10.1002/cncy.21436)
32. A.A. Renshaw, Does a repeated benign aspirate change the risk of malignancy after an initial atypical thyroid fine-needle aspiration? *Am. J. Clin. Pathol.* **134**, 788–792 (2010). doi:[10.1309/AJCPRA9Y2XQVFOFV](https://doi.org/10.1309/AJCPRA9Y2XQVFOFV)
33. A. Frasoldati, R. Valcavi, Challenges in neck ultrasonography: lymphadenopathy and parathyroid glands. *Endocr. Pract.* **10**(3), 261–268 (2004)
34. M. Bongiovanni, A. Spitale, W.C. Faquin, L. Mazzucchelli, Z.W. Baloch, The Bethesda System for Reporting Thyroid Cytopathology: a meta-analysis. *Acta Cytol.* **56**, 333–339 (2012). doi:[10.1159/000339959](https://doi.org/10.1159/000339959)