Original Investigation

The Natural History of Benign Thyroid Nodules

Cosimo Durante, MD, PhD; Giuseppe Costante, MD; Giuseppe Lucisano, MScStat; Rocco Bruno, MD; Domenico Meringolo, MD; Alessandra Paciaroni, MD; Efisio Puxeddu, MD, PhD; Massimo Torlontano, MD; Salvatore Tumino, MD; Marco Attard, MD; Livia Lamartina, MD; Antonio Nicolucci, MD; Sebastiano Filetti, MD

IMPORTANCE Detection of asymptomatic thyroid nodules has increased. Consensus is lacking regarding the optimal follow-up of cytologically proven benign lesions and sonographically nonsuspicious nodules. Current guidelines recommend serial ultrasound examinations and reassessment of cytology if significant growth is observed.

OBJECTIVE To determine the frequency, magnitude, and factors associated with changes in thyroid nodule size.

DESIGN, SETTING, AND PARTICIPANTS Prospective, multicenter, observational study involving 992 consecutive patients with 1 to 4 asymptomatic, sonographically or cytologically benign thyroid nodules. Patients were recruited from 8 hospital-based thyroid-disease referral centers in Italy between 2006 and 2008. Data collected during the first 5 years of follow-up, through January 2013, were analyzed.

MAIN OUTCOMES AND MEASURES Baseline nodule growth (primary end point) was assessed with yearly thyroid ultrasound examinations. Size changes were considered significant for growth if an increase of 20% or more was recorded in at least 2 nodule diameters, with a minimum increase of 2 mm. Baseline factors associated with growth were identified. Secondary end points were the sonographic detection of new nodules and the diagnosis of thyroid cancer during follow-up.

RESULTS Nodule growth occurred in 153 patients (15.4% [95% CI, 14.3%-16.5%]). One hundred seventy-four of the 1567 original nodules (11.1% [95% CI, 10.3%-11.9%]) increased in size, with a mean 5-year largest diameter increase of 4.9 mm (95% CI, 4.2-5.5 mm), from 13.2 mm (95% CI, 12.1-14.2 mm) to 18.1 mm (95% CI, 16.7-19.4 mm). Nodule growth was associated with presence of multiple nodules (OR, 2.2 [95% CI 1.4-3.4] for 2 nodules; OR, 3.2 [95% CI, 1.8-5.6 for 3 nodules; and OR, 8.9 [95% CI, 4.4-18.0] for 4 nodules), main nodule volumes larger than 0.2 mL (OR, 2.9 [95% CI, 1.7-4.9] for volumes >0.2 to <1 mL and OR, 3.0 [95% CI, 1.8-5.1] for volumes ≥1 mL), and male sex (OR, 1.7 [95% CI, 1.1-2.6]), whereas an age of 60 years or older was associated with a lower risk of growth than age younger than 45 years (OR, 0.5 [95% CI 0.3-0.9]). In 184 individuals (18.5% [95% CI, 16.4%-20.9%]), nodules shrank spontaneously. Thyroid cancer was diagnosed in 5 original nodules (0.3% [95% CI, 0.0%-0.6%]). Only 2 had grown. An incidental cancer was found at thyroidectomy in a nonvisualized nodule. New nodules developed in 93 patients (9.3% [95% CI, 7.5%-11.1%]), with detection of one cancer.

CONCLUSIONS AND RELEVANCE Among patients with asymptomatic, sonographically or cytologically benign thyroid nodules, the majority of nodules exhibited no significant size increase during 5 years of follow-up and thyroid cancer was rare. These findings support consideration of revision of current guideline recommendations for follow-up of asymptomatic thyroid nodules.

Editorial page 903

Author Audio Interview at jama.com

Supplemental content at jama.com

Author Affiliations: Author affiliations are listed at the end of this article.

Corresponding Author: Sebastiano Filetti, MD, Dipartimento di Medicina Interna e Specialità Mediche Università di Roma Sapienza, Viale del Policlinico, 155, 00161 Rome, Italy (sebastiano.filetti@uniroma1.it).

JAMA. 2015;313(9):926-935. doi:10.1001/jama.2015.0956

Thyroid nodule diagnosis has become an increasingly frequent event in clinical practice. This trend stems largely from improved detection of small incidentally discovered nodules due to the increased use of diagnostic imaging for purposes unrelated to the thyroid.[1-3]

The discovery of a thyroid nodule may be stressful for the patient, but more than 90% of the detected nodules are clinically insignificant benign lesions.[4,5] Fine-needle aspiration cytology is the procedure of choice to identify suspicious lesions that require thyroid surgery.[6-8] Established criteria for initial biopsy include nodule size and sonographic characteristics.[6-8] Nodules measuring less than 1 cm, the majority of which are discovered incidentally, do not require initial aspiration unless they exhibit suspicious features on ultrasonography.[6,8]

When the initial assessment indicates that a nodule is cytologically benign or sonographically nonsuspicious, clinicians are faced with the challenge of prescribing evidence-based follow-up protocols for these patients.[9] This task is complicated because knowledge of the natural history of thyroid nodules is incomplete. Current guidelines recommend serial ultrasound examinations for benign thyroid nodules and reassessment of cytology if significant growth is observed.[7] Significant growth has been defined as an increase of 20% or more in at least 2 nodule diameters, with a minimum increase of 2 mm.[7] However, little is known about the actual frequency and magnitude of nodule growth, and there is no reliable method for identifying patients likely to experience growth. The assumption that growing nodules increase a patient's risk of malignancy is untested.

To resolve these questions, we conducted a prospective multicenter study of the natural history of cytologically benign and sonographically nonsuspicious thyroid nodules. We report data from the first 5 years of follow-up, including changes in nodule size, baseline factors associated with nodule growth, the appearance of new thyroid nodules, and the incidence of thyroid cancer diagnosis.

Box. Enrollment Criteria

Nodule Characteristics

Number: from 1 to 4[a]

Largest diameter: 4 to 40 mm

Structure: solid or mixed (solid-cystic) with a fluid component representing no more than 75% of the total volume

No evidence of malignancy[b]

Patient Characteristics

No nodule-related signs or symptoms

Euthyroidism (ie, thyrotropin levels between 0.4 and 4 mU/L and free thyroxine levels within normal ranges)

No thyroid hormone therapy during the study or in the 6 months preceding enrollment

No history of surgical or nonsurgical thyroid interventions[c]

Serum thyroid antibody (autoantibodies to thyroid peroxidase and thyroglobulin) levels within normal limits[d]

No sonographic evidence of chronic autoimmune thyroiditis[d,e]

No clinical or sonographic evidence of acute thyroiditis[d,f]

[a] The decision to select patients who had up to 4 nodules has been made to facilitate the sonographic follow-up overtime.

[b] Benign cytology (Bethesda class II)[28] for any lesion with (1) sonographic features that were suggestive of malignancy (ie, at least 1 of the following: hypoechogenicity, irregular margins, a more tall (anteroposterior) than wide (transverse) shape of the nodule, intranodular vascular spots, microcalcifications)[6-8]; (2) largest diameter of 10 mm or more (in patients with multiple nonsuspicious nodules of this size, only the largest lesion was aspirated); or (3) both. Nonsuspicious lesions measuring less than 10 mm were classified as presumably benign, without cytological confirmation.

[c] Nonsurgical interventions include radioactive iodine ablation and ultrasound-guided minimally invasive procedures, such as ethanol injection and radiofrequency, laser, or ultrasound ablation.[29]

[d] These criteria were designed to exclude patients with chronic or acute thyroiditis because (1) acute or chronic inflammation of the thyroid parenchyma makes it difficult to distinguish true nodules from pseudonodules on ultrasound, and (2) positive thyroid antibody titers represent a potential confounding factor in a study of the natural history of thyroid nodules.

[e] Hypoechoic and heterogeneous thyroid parenchyma.

[f] Clinical (painful thyroid swelling) and sonographic evidence (hypoechoic and heterogeneous thyroid parenchyma) of acute thyroiditis.

Methods

Study Design, Population, and Protocol

This study was conducted in 8 hospital-based thyroid-disease referral centers in Italy, including 4 academic sites. The protocol was approved by the institutional review boards of all participating centers, and written informed consent was obtained from all participants.

Since January 1, 2006, on the basis of a written protocol, all centers consecutively enrolled patients with 1 to 4 asymptomatic, presumably benign thyroid nodules, using the criteria detailed in the **Box**. Baseline assessments included complete medical histories and physical examinations; measurement of serum levels of thyroid-stimulating hormone, free thyroxine, and autoantibodies to thyroid peroxidase and thyroglobulin; and thyroid ultrasound documentation of nodule number, size, volume, echogenicity, and sonographic features. Malignancy was excluded by fine-needle aspiration cytology for all nodules with suspicious sonographic features (at least 1 of the following: hypoechogenicity, irregular margins, taller-than-wide shape, intranodular vascular spots, microcalcifications).[6-8] In the absence of such features, aspiration was limited to nodules measuring 10 mm or more in the largest diameter. In patients with multiple nonsuspicious nodules of this size, only the largest nodule was aspirated.

Yearly follow-up visits at the center of enrollment included a color Doppler ultrasound examination of the thyroid and neck performed by an endocrinologist with specific training and at least 10 years' experience in thyroid and neck ultrasonography. All examinations at a given center were performed by the same operator throughout the whole study. Multifrequency linear transducers were used to ensure optimal axial and lateral resolution of superficial (12 MHz) and deeper

(7.5-10 MHz) structures. Examiners measured the transverse, anteroposterior, and longitudinal diameters of each nodule without consulting data from the previous follow-up scan. Volumes were calculated with the ellipsoid formula (the 3 largest perpendicular diameters multiplied by 0.525).[10] Examiners also noted the intraglandular location of each nodule, its structure (solid, spongiform, cystic, mixed), the echogenicity of its solid component, margin characteristics, and the presence of intranodular echogenic foci and flow signals on color Doppler imaging. Thyroid function tests were measured at each visit. In the absence of clinical indications, thyroid antibody titers were remeasured only at the 5-year assessment to confirm the absence of autoimmune thyroiditis. Fine-needle aspiration was performed or repeated when suspicious findings (hypoechogenicity, irregular margins, taller-than-wide shape, intranodular vascular spots, microcalcifications)[6-8] developed subsequent to the baseline ultrasound examination. In addition, at the 5-year visit, patients whose nodules had been cytologically diagnosed as benign at baseline and not reaspirated during follow-up were asked, but not required, to undergo repeat aspiration to confirm the accuracy of the baseline diagnosis.

Data, End Points, and Definitions

We analyzed sonographic findings to determine the incidence, magnitude, and temporal characteristics of spontaneous changes in the size and number of thyroid nodules during the first 5 years of follow-up (through January 31, 2013). If a participant had shorter follow-up times, the outcomes in these cases were those recorded when the patient was last seen. The prespecified primary end point was the proportion of patients showing a growth in the original nodule (ie, the nodule present at baseline) or, in case of multinodular disease, in at least 1 of the original nodules. It was adjudicated by the study coordinator after review of measurements recorded by local examiners at baseline and the final follow-up visit. To this purpose, size changes were considered significant if the increase or decrease involved at least 2 nodule dimensions, each amounting to at least 2 mm and representing at least 20% of the baseline diameter.[7] This approximates a nodule volume change of 50%, which is the minimal reproducible change that could be accurately measured. It has been reported that the interobserver variability for thyroid nodule volume measurement is 48.9% (95% CI, 39.8%-58.6%) using the ellipsoid method.[11] This definition was applied exclusively to the solid component of mixed nodules. As secondary end points, we also evaluated the sonographic detection of new nodules during follow-up and the diagnosis of thyroid cancer arising in an original nodule.

Statistical Analysis

The characteristics of participants with and without nodule growth were compared with the χ^2 statistic for categorical variables and the Mann-Whitney *U* test for continuous variables. An analysis of variance for repeated measures using a longitudinal linear model with an unstructured correlation-type matrix[12] was applied to estimate changes over 5 years of original nodule volume and diameter. Baseline factors associated with nodule growth were identified by multivariable backward logistic regression analysis and by recursive partitioning and amalgamation (RECPAM).[13,14] The latter method, which combines the advantages of standard logistic regression and tree-growing techniques, allowed definition of homogeneous patient subgroups with distinct risks for nodule growth. The variables entered into the RECPAM model were those used in the multivariable logistic regression analysis (listed in the footnotes of **Table 1**), with the exception of the nodule volume. The maximum diameter of the largest nodule was preferred over nodule volume because it is easier to obtain in clinical practice. Categorization of values for continuous variables was omitted in RECPAM to allow algorithm-based selection of the natural cut-off points. Tests were 2-sided, and a *P* value <.05 was considered statistically significant. Statistical analyses were performed with SAS software, version 9.2 (SAS Institute Inc).

Results

Patient Characteristics

Between January 1, 2006, and January 31, 2008, we enrolled 1009 patients. Seventeen were excluded from analysis because they developed thyroid dysfunction or autoantibody positivity during follow-up. The analytic population comprised 992 patients (**Table 2**; eFigure in the Supplement) with a total of 1567 nodules at baseline. The mean age of the patients was 52.4 years, 82% were women, and almost half had family histories of nodular thyroid disease. None were receiving levothyroxine therapy, although 30% had used it in the past. Approximately 60% had solitary nodules at baseline. More than 80% of the nodules were solid. Six hundred thirty nodules (40.2%), corresponding to 579 patients, were classified as benign on the basis of cytological findings. For the remaining 937 nodules (59.8%), the exclusion of malignancy was based solely on the absence of suspicious ultrasound features. This group included 852 nodules less than 1 cm in diameter and 85 nodules measuring 1 cm or more that were nontarget lesions in a multinodular thyroid gland.

For 875 patients, data representing 5 full years of follow-up were analyzed but shorter follow-up times were considered in 117 cases (11.8%; eFigure in the Supplement). These included 71 patients (7.1%) who were lost to follow-up, 4 (0.4%) who died due to nonthyroid-related causes, and 42 (4.2%) who had thyroidectomies for nodule growth, suspected thyroid cancer, or both before the 5-year assessment. There were no significant differences in the demographic or clinical characteristics between these patients, and the study cohort that had 5 years of follow-up (eTable 1 in the Supplement).

Changes in the Size, Number, and Benign Status of Nodules

In 686 of 992 patients (69%), the size of the original nodules remained stable during follow-up. In 184 others (18.5% [95% CI, 16.4%-20.9%]) , 1 or more nodules (205 of 1567; 13.1% [95% CI, 12.2%-14.0%]) shrank spontaneously, with mean reductions in the largest diameter of 3.7 mm (95% CI, −4.4 to −3.0 mm) and in volumes of 0.5 mL (95% CI, −0.6 to −0.4). Signifi-

Table 1. Baseline Features Associated With Original Thyroid Nodule Growth: Results of Logistic Regression Analyses[a]

Covariates	No. (%) of Patients	Odds Ratio (95% CI)[b]
Model 1: Total Cohort (n = 992)[c]		
Sex		
Women	814 (82.1)	1 [Reference]
Men	178 (17.9)	1.7 (1.1-2.6)
Age, y		
<45	301 (30.3)	1 [Reference]
≥45 to <60	358 (36.1)	0.7 (0.4-1.1)
≥60	333 (33.6)	0.5 (0.3-0.9)
BMI		
<25	372 (37.5)	1 [Reference]
≥25 to <27	173 (17.4)	0.4 (0.2-0.7)
≥27	447 (45.1)	1.00 (0.7-1.5)
No. of nodules		
1	594 (59.9)	1 [Reference]
2	262 (26.4)	2.2 (1.4-3.4)
3	95 (9.6)	3.2 (1.8-5.6)
4	41 (4.1)	8.9 (4.4-18.0)
Volume of largest nodule, mL		
≤0.2	320 (32.3)	1 [Reference]
>0.2 to <1	348 (35.1)	2.9 (1.7-4.9)
≥1	324 (32.7)	3.0 (1.8-5.1)
Model 2: Female Subcohort (n = 814)[d]		
BMI		
<25	318 (39.1)	1 [Reference]
≥25 to <27	138 (17.0)	0.4 (0.2-0.8)
≥27	358 (44.0)	1.0 (0.6-1.6)
No. of pregnancies		
>3	134 (16.5)	1 [Reference]
3	152 (18.7)	3.0 (1.1-8.6)
2	258 (31.7)	3.8 (1.4-10.2)
1	116 (14.3)	4.1 (1.5-11.6)
0	154 (18.9)	4.9 (1.8-13.6)
No. of nodules		
1	477 (58.6)	1 [Reference]
2	222 (27.3)	2.3 (1.4-3.6)
3	78 (9.6)	3.0 (1.6-5.8)
4	37 (4.6)	7.1 (3.3-14.9)
Volume of largest nodule, mL		
<0.2	273 (33.5)	1 [Reference]
>0.2 to <1	291 (35.8)	2.4 (1.3-4.4)
≥1	250 (30.7)	2.7 (1.5-4.9)

Abbreviation: BMI, body mass index, calculated as weight in kilograms divided by height in meters squared.

[a] Nodule growth was defined according to American Thyroid Association recommendations (an increase ≥20% in ≥2 nodule diameters, with a minimum increase of 2 mm)[7] and measured at the 5-year follow-up visit.

[b] Odds ratios (ORs) were estimated with respect to the reference category (ie, the first category listed under each variable, OR, 1.0)

[c] Other baseline covariates included in model 1 were family history of nodular thyroid disease (yes/no), previous levothyroxine treatment (yes/no), smoking during the 6 months preceding enrollment (yes/no), serum thyrotropin level (<1, 1-2, or ≥2 mIU/L), largest diameter of largest nodule (≤10 or >10 mm), structure of the largest nodule (solid or mixed), thyroid gland volume (≤10, 10.1-15, or >15 mL), area of residency (northern or southern Italy, which are characterized by mild and moderate iodine deficiency, respectively).[26,27]

[d] Other baseline covariates included in model 2 were oligomenorrhea (yes/no), menopause (yes/no), use of estrogen-progestin drugs during the 6 months preceding enrollment (yes/no), age in years (<45, ≥45 to <60 or ≥60), family history of nodular thyroid disease (yes/no), previous levothyroxine treatment (yes/no), smoking during the 6 months preceding enrollment (yes/no), serum thyrotropin level (<1, 1-2, or ≥2 mIU/L), diameter of largest nodule (≤10 or >10 mm), structure of the largest nodule (solid or mixed), thyroid gland volume (≤10, 10.1-15, or >15 mL), area of residency (northern or southern Italy, which are characterized by mild and moderate iodine deficiency, respectively).[26,27]

cant growth (a 20% increase ≥2 nodule diameters, with a minimum increase of 2 mm) occurred in 153 patients (15.4% [95% CI, 14.3%-16.5%]), 94 (61.4% [95% CI, 59.9%-62.9%]) of whom had multiple nodules at baseline. In 73 patients (77.6% [95% CI, 75.5%-79.7%]) with multinodular glands, the growth involved only 1 nodule. This nodule was the largest baseline nodule in 50 out of 73 patients (68.5% [95% CI, 63.1%-73.9%]). Significant growth of 2 nodules occurred in 21 patients (22.3% [95% CI, 20.2%-24.4%]) with multinodular glands. In 31 patients (3.1% [95% CI, 2.2% -4.0%]) with multiple nodules, both growth and shrinkage were detected in the same gland. Overall, 174 of the 1567 original nodules (11.1% [95% CI, 10.3%-11.9%]) increased in size, with a mean change in the largest diameter of more than 4.9 mm (95% CI, 4.2-5.5 mm) from 13.2 mm (95% CI, 12.1-14.2 mm) at baseline to 18.1 mm (95% CI, 16.7-19.4 mm) at the end of follow-up ($P < .001$). As shown in **Figure 1**, nodule size changes occurred in a linear fashion, starting from the first year of observation. Mean yearly increases in the largest diameter were less than 1 mm, and nodule shrinkage was more gradual.

Table 2. Baseline Characteristics of the Study Cohort and Subcohorts[a]

Characteristics	Total Cohort (N = 992)	Nodule Growth		P Value[b]
		Without (n = 839)	With (n = 153)	
Patient characteristics				
Age, mean (SD), y	52.4 (13.7)	52.7 (13.8)	50.8 (13.1)	.09
Residence, No. (%)[c]				
Northern Italy	425 (42.8)	350 (41.7)	75 (49.0)	.11
Southern Italy	567 (57.2)	489 (58.3)	78 (51.0)	
Women, No. (%)	814 (82.1)	695 (82.8)	119 (77.8)	.13
No. pregnancies, No. (%)				
0	154 (18.9)	124 (17.8)	30 (25.2)	<.001
1-2	374 (45.9)	311 (44.7)	63 (52.9)	
>2	286 (35.1)	260 (37.4)	26 (21.8)	
Oligomenorrhea, No. (%)	57 (7.0)	45 (6.5)	12 (10.1)	.15
Menopause, No. (%)	447 (54.9)	392 (56.4)	55 (46.2)	.03
Estrogen-progestin therapy, No. (%)[d]	71 (8.7)	61 (8.8)	10 (8.4)	.96
BMI, mean (SD)	26.9 (4.8)	27.0 (4.8)	26.8 (4.8)	.71
Smokers, No. (%)[e]	163 (16.4)	130 (15.5)	33 (21.6)	.06
Thyroid gland volume, mean (SD), mL	14.5 (7.2)	14.4 (7.4)	14.5 (6.1)	.59
Family history of nodular thyroid disease, No. (%)[f]	458 (46.2)	384 (45.8)	74 (48.4)	.55
Thyroid-stimulating hormone, mean (SD), mIU/L	1.46 (0.75)	1.45 (0.74)	1.56 (0.80)	.10
Previous levothyroxine treatment, No. (%)[g]	297 (29.9)	248 (29.6)	49 (32.0)	.54
Nodule characteristics[h]				
Multiple nodules, No. (%)	398 (40.1)	304 (36.2)	94 (61.4)	<.001
Maximum diameter of the largest nodule, mean (SD), mm	14.2 (8.1)	13.9 (8.3)	15.7 (7.3)	<.001
Volume of the largest nodule, mL				
Mean (SD)	1.47 (2.81)	1.46 (2.89)	1.53 (2.33)	<.001
Median (IQR)	0.38 (0.15-1.45)	0.35 (0.14-1.37)	0.61 (0.27-1.71)	
Structure of the largest nodule, No. (%)				
Solid	780 (78.6)	665 (79.3)	115 (75.2)	.26
Mixed[i]	212 (21.4)	174 (20.7)	38 (28.4)	

Abbreviations: BMI, body mass index, calculated as weight in kilograms divided by height in meters squared; IQR, interquartile range.

[a] Nodule growth at the end of follow-up was defined in accordance with the 2009 Guidelines of the American Thyroid Association (an increase of ≥20% in at least 2 nodule diameters, with a minimum increase of 2 mm).[7]

[b] Significant differences between the 2 subcohorts assessed with the Pearson χ^2 test (categorical variables) and Mann-Whitney *U* tests (continuous variables).

[c] Residents of northern Italy included patients from the regions of Emilia-Romagna, Umbria, and Lazio; residents of southern Italy were from Basilicata, Calabria, Apulia, or Sicily. Iodine deficiency has been classified as mild in northern Italy and moderate in the southern regions (median urine iodine levels: 90 µg/L and 72 µg/L, respectively. To convert iodine from µg/L to nmol/L, multiply by 7.88).[26,27]

[d] Includes use of oral contraceptives, therapy for oligomenorrhea, or hormone replacement therapy during the 6 months preceding enrollment.

[e] Includes current smokers and those who had stopped smoking less than 6 months before study enrollment.

[f] Self-reported.

[g] Discontinued at least 6 months before enrollment. In 90% of the cases, levothyroxine had been discontinued at least 2 years prior to enrollment.

[h] Patients with multiple nodules had 2-4 lesions. In these cases, size and structure are given only for the main (ie, largest) nodule.

[i] Solid with cystic components representing <75% of total nodule volume. Spongiform nodules are included in mixed nodules.

No changes in nodule number were noted during follow-up in 876 of the 992 patients (88.3% [95% CI, 87.3%-89.3%]). In 23 patients (2.3% [95% CI, 1.8%-2.8%]) with multinodular thyroids at baseline, 1 of the original thyroid nodules disappeared. New nodules were discovered in 93 patients (9.3% [95% CI, 8.4%-10.2%]): 1 nodule in 76 patients, 2 in 15 patients, and 3 in 2 patients.

Of 365 patients whose nodules were reaspirated, the benign baseline diagnosis was confirmed in 361 cases (98.9% [95% CI, 98.4%-99.4%]). In 245 patients (67.9% [95% CI, 65.4%-70.4%]) of 361 cases, fine-needle aspiration cytology was repeated at the 5-year follow-up visit to confirm the accuracy of the baseline diagnosis, 109 patients (30.2% [95% CI, 27.8%-32.6%]) because of nodule growth and 7 patients (1.9% [95% CI, 1.2%-2.6%]) based on the appearance of suspicious ultrasound features. Thyroid cancer was identified in 5 (0.3% [95% CI, 0.2%-0.4%]) of the 1567 original nodules (**Table 3**). Four nodules had been classified as benign on the basis of cytological findings, and the fifth was a nodule less than 1 cm, with the benign status at baseline assessed by ultrasound features. The latter displayed no growth or sonographic changes during follow-up until the fifth year, when it became hypoechoic with

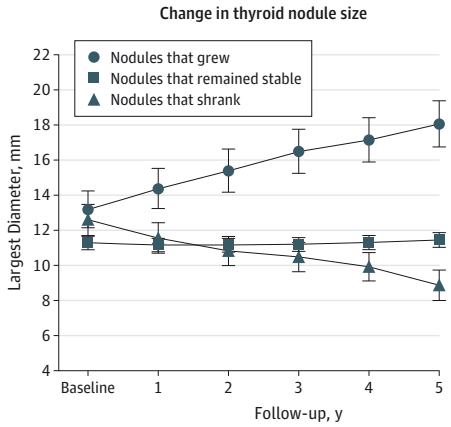
irregular margins, and the fine-needle aspirate was suspicious. In 2 other patients, malignancy was diagnosed in a nodule that had not been visualized at baseline. In one case, it was noted for the first time during the 3-year follow-up ultrasound study. In the other patient, the malignant nodule was never visualized sonographically. It was discovered incidentally after thyroidectomy prompted by changes (suspicious ultrasound features, indeterminate cytology) in the single nodule visualized at baseline. The latter lesion ultimately proved to be a follicular adenoma, but a previously undetected nodule was histologically diagnosed as a 3-mm papillary thyroid carcinoma.

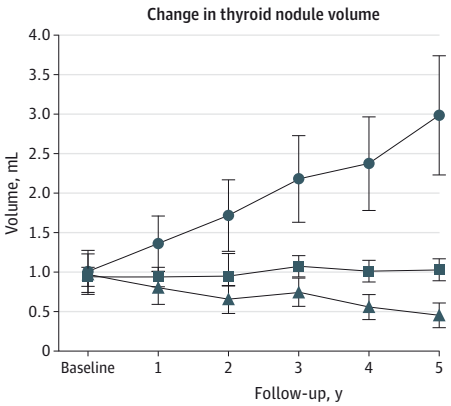
Baseline Variables Associated With Original Nodule Growth

In bivariable analysis (Table 2), nodule growth during follow-up was significantly associated with lesion number and size at baseline and nulliparity and premenopausal age in women. All baseline variables were entered into an adjusted multivariable logistic regression model (Table 1). The presence of multiple nodules, main nodule volumes of more than 0.2 mL, and male sex were associated with nodule growth, whereas an age of 60 years or more was associated with a lower risk of growth. In women, multiple nodules, larger nodule volumes, and nulliparity were associated with nodule growth. Neither sonographic features nor serum thyrotropin levels were associated with nodule growth.

RECPAM analysis identified 6 patient subgroups characterized by increasing risk of original nodule growth (**Figure 2** and eTable 2 in the Supplement). The largest diameter of the main nodule at baseline was the most relevant correlate of subsequent growth. Growth was least likely in patients whose largest nodule diameter was 7.5 mm or less (class 1), the referents for the other 5 classes. By class, the odds ratios (OR) ranged from 2.8 (95% CI, 1.2-6.8) for class 2, 9.2% of patients, to 20.7 (95% CI, 8.6-50.1) for class 6, 42.7% of patients. Patients 43 years or older with multinodular disease and with the largest nodule diameter exceeding 7.5 mm had the highest likelihood of nodule growth during follow-up (class 6). Older patients with comparable disease features exhibited a lower odds, which were dependent on body mass index (BMI), calculated as weight in kilograms divided by height in meters squared: among 32.5% of patients who had a BMI of 28.6 or higher had an OR of 13.4 (95% CI, 5.5-32.6) (class 5) and among 18.0% of patients with a BMI of 28.6 or lower had an OR of 6.1 (95% CI, 2.6-14.2) (class 4). Patients who had presented with a solitary nodule whose largest diameter exceeded 7.5 mm, growth was more likely among the 14.9% of patients aged 51 years or younger (OR, 4.9; 95% CI, 2.1-11.2) than 9.2% of those older than 51 years (OR, 2.8; 95% CI, 1.2-6.8) (classes 3 and 2). When multiple logistic regression was repeated with the RECPAM classes forced into the model (eTable 3 in the Supplement), no additional variables were associated with nodule growth.

Figure 1. Changes in Thyroid Nodule Size and Volume During the First 5 Years of Follow-up





No. of nodules that						
Grew	174	174	170	157	144	131
Remained stable	1188	1188	1170	1155	1115	1096
Shrank	205	205	197	188	173	161
No. of patients	992	992	963	946	902	875

Nodule growth occurred in 174 (11.1%) of the 1567 nodules present at baseline; 1188 (75.8%) remained stable and 205 (13.1%) shrank. Graphs represent the estimated mean with 95% CIs of the maximum diameters and volumes of thyroid nodules. An analysis of variance for repeated measures was carried out to evaluate the change in thyroid nodule size over 5 years of follow-up.

Discussion

This prospective study was undertaken to define the natural history of cytologically benign and small sonographically nonsuspicious asymptomatic thyroid nodules over 5 years of follow-up. We found that the vast majority of nodules present at baseline exhibited no significant size change during this period or they actually decreased in size. Growth considered significant by American Thyroid Association standards was observed in only 15% of the patients. Growth was slow, steady, and limited in magnitude, with a mean 5-year largest diameter increase of 4.9 mm, and which was generally restricted to the main nodule in patients with multinodular disease. Nodule size changes occurred early, starting from the 1-year follow-up visit. On the basis of the RECPAM analysis, the most relevant baseline characteristics associated with nodule growth were multinodular disease, nodule diameters of 7.5 mm or more, and age at diagnosis of 43 years or less. Growth of solitary nodules was less likely, and also inversely associated with age. Among older patients with multiple

Table 3. Patients With Diagnoses of Thyroid Cancer During Follow-up

					Characteristics of Cancerous Nodule					
					At Baseline			During Follow-up		
Patient No.	Sex	Age at Cancer Diagnosis, y	No. Nodules at Baseline	Time to Diagnosis, years[a]	Largest Diameter, mm	Sonographic Features	Cytology[28]	Behavior	Cytology[28]	Surgical Histology
Cancers Arising in Original Nodules										
1	Woman	40	Multiple	2	15.2	Solid, hypoechoic	Benign (Bethesda class II)	Growth	Suspicious (Bethesda class V)	PTC-FV
2	Woman	31	One	2	18.1	Solid, hypoechoic	Benign (Bethesda class II)	No growth[b]	Malignant (Bethesda class VI)	PTC
3	Woman	60	Multiple	4	8.0	Solid, isoechoic, microcalcifications	Benign (Bethesda class II)	Growth	Malignant (Bethesda class VI)	PTC
4	Woman	73	One	5	22.0	Solid, hypoechoic	Benign (Bethesda class II)	No growth[c]	Indeterminate (Bethesda class IV)	PTC-FV
5	Woman	47	One	5	9.2	Mixed, hypoechoic	Not done	No growth[b]	Suspicious (Bethesda class V)	PTC
Cancers Arising in New Nodules										
6	Woman	43	One	4	Not present			First detected (with US) at visit 3; growing nodule[d]	Suspicious (Bethesda class V)	PTC
7	Woman	62	One	4	Not present			Incidentally detected during surgery[e]	Not done	PTC

Abbreviations: PTC, papillary thyroid cancer, classic variant; PTC-FV, papillary thyroid cancer, follicular variant; US, ultrasonography.

[a] Time from study enrollment to cancer diagnosis. Year of diagnosis coincides with final follow-up visit and outcome classification (original nodule growth).

[b] Fine-needle aspiration cytology was performed because of the appearance of suspicious ultrasound features in the original nodule.

[c] Fine-needle aspiration cytology was repeated at the 5-year follow-up visit as per protocol (see the Methods section) even though no signs of growth or suspicious ultrasound findings had been observed.

[d] The new nodule was first detected at the 3-year follow-up visit: it had suspicious sonographic features (hypoechogenicity, irregular margins) from the outset and a largest diameter of 4 mm. At the 4-year visit, the diameter had increased to almost 10 mm, and a biopsy disclosed suspicious cytology (Bethesda class V). The histologic examination revealed a 9-mm papillary carcinoma. The solitary nodule present at baseline never showed any signs of growth or malignancy during follow-up and was in fact benign at final histology.

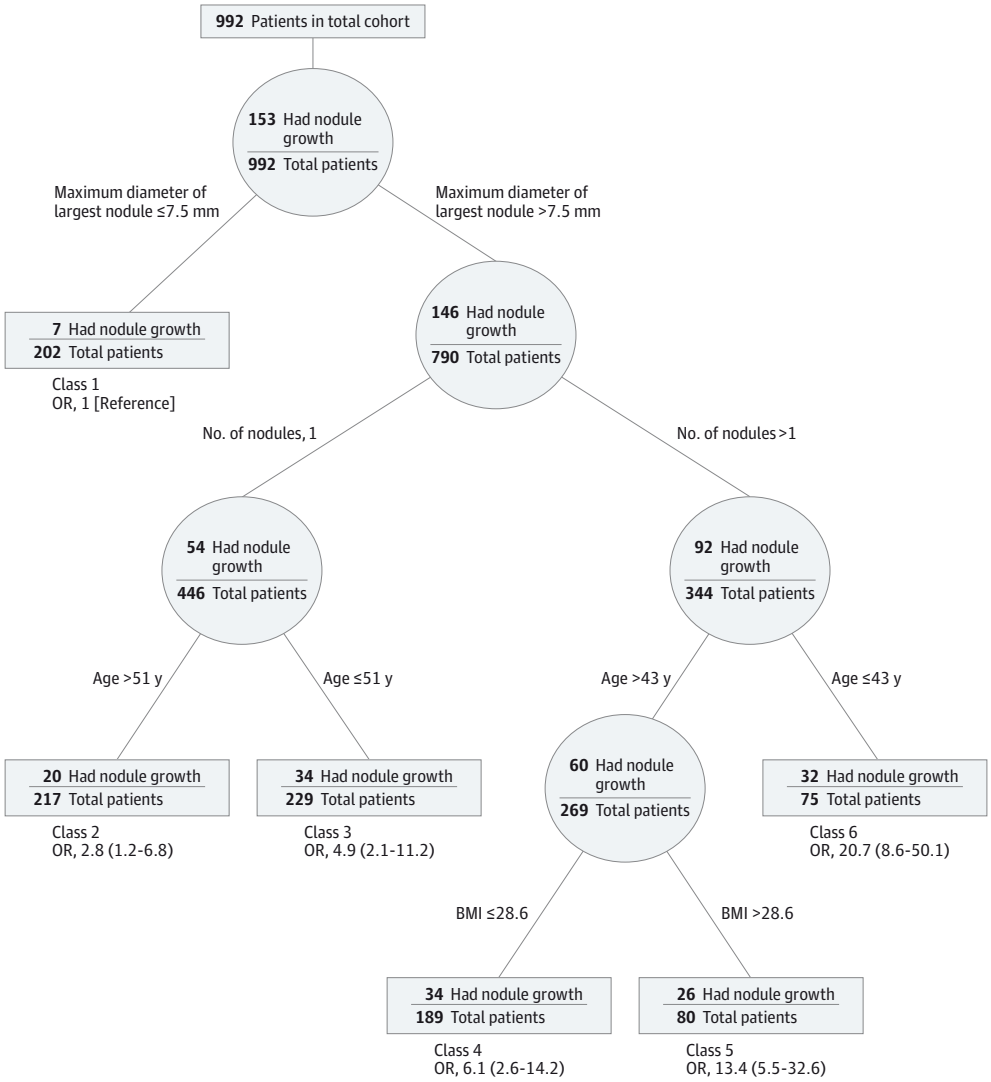
[e] The cancerous nodule (maximum diameter 3 mm) was never visualized sonographically. It was discovered incidentally during thyroidectomy prompted by suspected malignancy in the single nodule present at baseline (based on the appearance of suspicious ultrasound features during follow-up and indeterminate cytology findings [Bethesda class IV]). The suspicious nodule ultimately proved to be a follicular adenoma.

nodules and larger dominant nodules, a higher BMI was associated with nodule growth, an observation that is consistent with recent reports linking thyroid nodular disease to obesity and insulin resistance.[15-17] Smoking has been reported to be associated with an increased risk of developing nodular thyroid disease.[18] In this series it was not significantly associated with nodule growth in either the multivariable logistic regression or RECPAM analysis.

Previous studies pointed to the existence of a subgroup of benign nodules that can be expected to grow over time. However, they include relatively small retrospective series of patients, with short follow-up times and widely differing criteria for defining nodule growth. Conflicting results have emerged, with some authors insisting that most nodules (39% to 61.2%) grow[19,20] and others maintaining that the majority (68% to 96%) remain stable in size and may even shrink or disappear entirely.[21] To define nodule growth they used cut offs of a 15% and 30% volume increase, which are not easily applicable in the clinical setting. A 50% cut off for nodule volume growth is the minimal significant and reproducible change in nodule size to be applied in clinical investigations and in clinical practice,[11] and it was endorsed as such by the American Thyroid Association guidelines.[7]

One of the goals of surveillance is the prompt detection and treatment of thyroid cancers that arise during follow-up or have been missed on the initial assessment. In the population we studied, these events were rare. Like others,[22,23] we found that fine-needle aspiration cytology of thyroid nodules is associated with a very low false-negative rate (1.1%). Moreover, only 7 patients (0.7% of the study population) were diagnosed with thyroid cancer during the 5-year of follow-up, and in 2 of these cases, the malignancy was located in a nodule that had not been present at baseline. In one of the latter patients the malignant nodule was never visualized on ultrasound, but it was discovered incidentally after thyroidectomy. In addition to cytologically benign thyroid nodules, we followed sonographically nonsuspicious lesions that did not meet the criteria for fine-needle aspiration cytology. Among the latter, we cannot exclude the presence of latent, asymptomatic, undiagnosed thyroid cancers. The frequency with which sonographically

Figure 2. Identification of Patient Subgroups With Different Risks for Original Thyroid Nodule Growth: Results of Recursive Partitioning and Amalgamation (RECPAM) Analysis



The RECPAM tree-growing algorithm models the odds of presenting a growth in the original nodule (ie, the nodule present at baseline) or, in case of multinodular disease, in at least 1 of the original nodules during the first 5 years of follow-up based on a multivariable logistic analysis. At each partitioning step, the method selects the covariate with the binary split that maximizes the difference in the risk. The algorithm stops when user-defined conditions (stopping rules) are met. The minimum set considered in our analysis comprised 5 nodule-growth cases and 50 patients per node. Splitting baseline variables were age in years; body mass index (BMI), calculated as weight in kilograms divided by height in meters squared; maximum diameter of the largest nodule in millimeters; and the number of nodules. Values associated with patient assignment to each of the 2 subgroups (circles or rectangles, the latter representing final RECPAM classes) are shown on the branches leading to the subgroup. The data in the circles and rectangles represent the ratios of the number of patients with growth to the total number of patients in subgroup. Odds ratios (ORs) for growth with 95% CIs are shown for each of the 6 classes. Class 1, for which nodule growth is least likely, is the reference category (OR, 1).

nonsuspicious or sonographically undetectable thyroid carcinomas progress to clinical disease is currently unknown. Indeed, in a Japanese series including biopsy-proven papillary thyroid microcarcinomas that were not surgically removed, 6.4% showed an increase in tumor size of 3 mm or more by ultrasound on 5-year follow-up, and 1.4% showed evidence for lymph node metastases.[24] Collectively, these data point to the existence of a minority of thyroid carcinomas that can be expected to grow over time, in general very slowly, and to progress to clinical disease. In our series, nodule growth alone was a nonspecific marker of malignancy. Only 2 of the 5 diagnoses of cancer in an established nodule were preceded by significant growth of the cancerous nodule. These data suggest that the American Thyroid Association's recommendation for indication for repeat cytology should be revised.[7] Clinical and sonographic findings should probably play larger roles in the decision-making process.[25]

Iodine intake is a major environmental determinant of thyroid nodular disease. The patients in our study all live in regions characterized by mild (northern Italy) or moderate (south-

ern Italy) iodine deficiency, despite an iodine supplementation program in place for the past 20 years.[26] We cannot exclude the possibility that in areas with more marked iodine deficiency, or none at all, nodule behavior might differ from that observed in our population. However, the absence of significant differences between the subgroups from northern and southern Italy, which are characterized by median urine iodine levels of 90 µg/L and 72 µg/L, respectively,[26,27] suggests that the iodine supply plays a minor role in the evolution of thyroid nodular disease. Different trends might therefore be observed in populations with different levels of iodine nutrition and genetic profiles, but even in these settings, the majority of nodules can probably be expected to exhibit no growth at all.

The predictive modeling that we applied to identify the characteristics associated with nodule growth has not been externally validated. This is an important factor in all prediction models, but may be particularly the case with the methods used in this study. Because recursive partitioning findings are derived from the study population, they are particularly prone to overfitting. This is a major limitation of this component of the study. Therefore, all descriptions of these findings should be considered as preliminary until and unless they are replicated in an independent study population.

Aside from this consideration, our study has several strengths, most notably, its size and prospective nature. Our data are representative of hospital-based outpatient clinics. The participating centers included academic centers and community hospitals, serving populations from rural as well as urban areas. The nodules considered included cytologically benign nodules as well as small, subcentimeter lesions that are not routinely biopsied, mirroring current clinical practice. Finally, the size changes were presented using criteria recommended by the American Thyroid Association.[7]

Current guidelines suggest, based on expert opinion, repeating thyroid ultrasonography after 6 to 18 months and, if nodule size is stable, every 3 to 5 years.[7] The indolent behavior and limited growth observed in our study confirm that nodules that were benign based on initial fine-needle aspiration or subcentimeter and sonographically nonsuspicious can be safely managed with a second ultrasound examination 1 year after the first (early follow-up) and in the absence of changes, reassessment after 5 years (long-term follow-up). This approach should be suitable for 85% of patients, whose risk of disease progression is low. Closer surveillance may be appropriate for nodules occurring in younger patients or older overweight individuals with multiple nodules, large nodules (>7.5 mm), or both.

Conclusions

Among patients with asymptomatic thyroid nodules that were sonographically or cytologically benign, the majority of nodules exhibited no significant size change during 5 years of follow-up or they actually decreased in size. Significant nodule growth occurred in 15% of cases, new nodules developed in 9.3% of patients, and thyroid cancer was diagnosed in 0.3% of nodules. These findings justify reconsideration of the current guideline recommendations for follow-up of asymptomatic thyroid nodules.

ARTICLE INFORMATION

Author Affiliations: Dipartimento di Medicina Interna e Specialità Mediche, Università di Roma Sapienza, Roma, Italy (Durante, Paciaroni, Lamartina, Filetti); Dipartimento di Scienze della Salute, Università di Catanzaro Magna Graecia, Catanzaro, Italy (Costante); Center for Outcomes Research and Clinical Epidemiology, Pescara, Italy (Lucisano, Nicolucci); Unità di Endocrinologia, Ospedale di Tinchi-Pisticci, Matera, Italy (Bruno); Unità Operativa Semplice Dipartimentale di Endocrinologia, Ospedale di Bentivoglio, Bologna, Italy (Meringolo); Dipartimento di Medicina, Università di Perugia, Perugia, Italy (Puxeddu); Unità Operativa di Endocrinologia, Istituto di Ricovero e Cura a Carattere Scientifico, Casa Sollievo della Sofferenza, San Giovanni Rotondo, Italy (Torlontano); Dipartimento di Medicina Clinica e Sperimentale, Università di Catania, Catania, Italy (Tumino); Unità Operativa di Endocrinologia, Azienda Ospedaliera Ospedali Riuniti Villa Sofia-Cervello, Palermo, Italy (Attard); Dipartimento di Scienze Mediche di Base, Neuroscienze ed Organi di Senso, Università di Bari Aldo Moro, Bari, Italy (Lucisano).

Author Contributions: Drs Durante and Filetti had full access to all of the data in the study and take responsibility for the integrity of the data and the accuracy of the data analysis.
Study concept and design: Durante, Costante, Nicolucci, Filetti.
Acquisition, analysis, or interpretation of data: Durante, Costante, Lucisano, Bruno, Meringolo, Paciaroni, Puxeddu, Torlontano, Tumino, Attard, Lamartina, Filetti.
Drafting of the manuscript: Durante, Nicolucci, Filetti.
Critical revision of the manuscript for important intellectual content: All the authors.
Statistical analysis: Lucisano, Nicolucci.
Obtained funding: Filetti.
Administrative, technical, or material support: Lucisano, Paciaroni, Puxeddu, Lamartina.
Study supervision: Durante, Costante, Bruno, Meringolo, Torlontano, Tumino, Attard, Nicolucci.

Conflict of Interest Disclosures: All authors have completed and submitted the ICMJE Form for Disclosure of Potential Conflicts of Interest and none were reported.

Funding/Support: The study was funded by research grants from the Umberto Di Mario Foundation, Banca d'Italia, and the Italian Thyroid Cancer Observatory (ITCO) Foundation.

Role of the Funder/Sponsor: The funders had no role in the design and conduct of the study; in the collection, analysis, and interpretation of the data; in the preparation, review, or approval of the manuscript; or in the decision to submit the manuscript for publication.

Additional Contributions: The authors thank David Cooper, MD, Division of Endocrinology, Diabetes, and Metabolism, Johns Hopkins University School of Medicine, for invaluable comments and suggestions, and to Marian Everett Kent, BSN (Rome, Italy), for medical editing services. No compensation for the contribution was received by Dr Cooper. Ms Kent's contribution was funded by research grants from the Umberto Di Mario Foundation. We also thank all the medical doctors, residents, students, and nurses who participated in this study for their important contributions.

REFERENCES

1. Davies L, Ouellette M, Hunter M, Welch HG. The increasing incidence of small thyroid cancers: where are the cases coming from? *Laryngoscope*. 2010;120(12):2446-2451.

2. Sosa JA, Hanna JW, Robinson KA, Lanman RB. Increases in thyroid nodule fine-needle aspirations, operations, and diagnoses of thyroid cancer in the United States. *Surgery*. 2013;154(6):1420-1426.

3. Smith-Bindman R, Lebda P, Feldstein VA, et al. Risk of thyroid cancer based on thyroid ultrasound imaging characteristics: results of a population-based study. *JAMA Intern Med*. 2013;173(19):1788-1796.

4. Papini E, Guglielmi R, Bianchini A, et al. Risk of malignancy in nonpalpable thyroid nodules: predictive value of ultrasound and color-Doppler features. *J Clin Endocrinol Metab*. 2002;87(5):1941-1946.

5. Mandel SJ. Diagnostic use of ultrasonography in patients with nodular thyroid disease. *Endocr Pract*. 2004;10(3):246-252.

6. Gharib H, Papini E, Paschke R, et al; AACE/AME/ETA Task Force on Thyroid Nodules. American Association of Clinical Endocrinologists, Associazione Medici Endocrinologi, and EuropeanThyroid Association medical guidelines for clinical practice for the diagnosis and management of thyroid nodules. *Endocr Pract*. 2010;16(suppl 1):1-43.

7. Cooper DS, Doherty GM, Haugen BR, et al; American Thyroid Association (ATA) Guidelines Taskforce on Thyroid Nodules and Differentiated Thyroid Cancer. Revised American Thyroid Association management guidelines for patients with thyroid nodules and differentiated thyroid cancer. *Thyroid*. 2009;19(11):1167-1214.

8. Pacini F, Schlumberger M, Dralle H, Elisei R, Smit JW, Wiersinga W; European Thyroid Cancer Taskforce. European consensus for the management of patients with differentiated thyroid carcinoma of the follicular epithelium. *Eur J Endocrinol*. 2006;154(6):787-803.

9. Filetti S, Durante C, Torlontano M. Nonsurgical approaches to the management of thyroid nodules. *Nat Clin Pract Endocrinol Metab*. 2006;2(7):384-394.

10. Knudsen N, Bols B, Bülow I, et al. Validation of ultrasonography of the thyroid gland for epidemiological purposes. *Thyroid*. 1999;9(11):1069-1074.

11. Brauer VF, Eder P, Miehle K, Wiesner TD, Hasenclever H, Paschke R. Interobserver variation for ultrasound determination of thyroid nodule volumes. *Thyroid*. 2005;15(10):1169-1175.

12. Cnaan A, Laird NM, Slasor P. Using the general linear mixed model to analyse unbalanced repeated measures and longitudinal data. *Stat Med*. 1997;16(20):2349-2380.

13. Ciampi A. *Constructing Prediction Trees From Data: The RECPAM Approach Proceedings From the Prague 1991 Summer School on Computational Aspects of Model Choice*. Heidelberg, Germany: Physica-Verlag; 1992:165-178.

14. Franciosi M, Pellegrini F, De Berardis G, et al; QuED Study Group—Quality of Care and Outcomes in Type 2 Diabetes. Self-monitoring of blood glucose in non–insulin-treated diabetic patients: a longitudinal evaluation of its impact on metabolic control. *Diabet Med*. 2005;22(7):900-906.

15. Anil C, Akkurt A, Ayturk S, Kut A, Gursoy A. Impaired glucose metabolism is a risk factor for increased thyroid volume and nodule prevalence in a mild-to-moderate iodine deficient area. *Metabolism*. 2013;62(7):970-975.

16. Sousa PA, Vaisman M, Carneiro JR, et al. Prevalence of goiter and thyroid nodular disease in patients with class III obesity. *Arq Bras Endocrinol Metabol*. 2013;57(2):120-125.

17. Rendina D, De Filippo G, Mossetti G, et al. Relationship between metabolic syndrome and multinodular non-toxic goiter in an inpatient population from a geographic area with moderate iodine deficiency. *J Endocrinol Invest*. 2012;35(4):407-412.

18. Brix TH, Hansen PS, Kyvik KO, Hegedüs L. Cigarette smoking and risk of clinically overt thyroid disease: a population-based twin case-control study. *Arch Intern Med*. 2000;160(5):661-666.

19. Quadbeck B, Pruellage J, Roggenbuck U, et al. Long-term follow-up of thyroid nodule growth. *Exp Clin Endocrinol Diabetes*. 2002;110(7):348-354.

20. Alexander EK, Hurwitz S, Heering JP, et al. Natural history of benign solid and cystic thyroid nodules. *Ann Intern Med*. 2003;138(4):315-318.

21. Erdogan MF, Gursoy A, Erdogan G. Natural course of benign thyroid nodules in a moderately iodine-deficient area. *Clin Endocrinol (Oxf)*. 2006;65(6):767-771.

22. Layfield LJ, Cibas ES, Gharib H, Mandel SJ. Thyroid aspiration cytology: current status. *CA Cancer J Clin*. 2009;59(2):99-110.

23. Nou E, Kwong N, Alexander LK, Cibas ES, Marqusee E, Alexander EK. Determination of the optimal time interval for repeat evaluation after a benign thyroid nodule aspiration. *J Clin Endocrinol Metab*. 2014;99(2):510-516.

24. Ito Y, Miyauchi A, Inoue H, et al. An observational trial for papillary thyroid microcarcinoma in Japanese patients. *World J Surg*. 2010;34(1):28-35.

25. Horvath E, Majlis S, Rossi R, et al. An ultrasonogram reporting system for thyroid nodules stratifying cancer risk for clinical management. *J Clin Endocrinol Metab*. 2009;94(5):1748-1751.

26. Aghini-Lombardi F, Vitti P, Antonangeli L, et al; Southern Italy Study Group for Iodine Deficiency Disorders. The size of the community rather than its geographical location better defines the risk of iodine deficiency: results of an extensive survey in Southern Italy. *J Endocrinol Invest*. 2013;36(5):282-286.

27. Watutantrige Fernando S, Barollo S, Nacamulli D, et al. Iodine status in schoolchildren living in northeast Italy: the importance of iodized-salt use and milk consumption. *Eur J Clin Nutr*. 2013;67(4):366-370.

28. Cibas ES, Ali SZ; NCI Thyroid FNA State of the Science Conference. The Bethesda system for reporting thyroid cytopathology. *Am J Clin Pathol*. 2009;132(5):658-665.

29. Gharib H, Hegedüs L, Pacella CM, Baek JH, Papini E. Clinical review: nonsurgical, image-guided, minimally invasive therapy for thyroid nodules. *J Clin Endocrinol Metab*. 2013;98(10):3949-3957.