



# Malignancy risk of initially benign thyroid nodules: validation with various Thyroid Imaging Reporting and Data System guidelines

Su Min Ha<sup>1</sup> · Jung Hwan Baek<sup>2</sup> · Young Jun Choi<sup>2</sup> · Sae Rom Chung<sup>2</sup> · Tae Yon Sung<sup>3</sup> · Tae Yong Kim<sup>4</sup> · Jeong Hyun Lee<sup>2</sup>

Received: 16 March 2018 / Revised: 13 May 2018 / Accepted: 25 May 2018 / Published online: 19 June 2018  
© European Society of Radiology 2018

## Abstract

**Objective** Some authors have found little or no diagnostic benefit from repeated biopsy of benign thyroid nodules. However, to our knowledge, integration of Thyroid Imaging Reporting and Data System (TIRADS) guidelines with one biopsy for sufficient benign thyroid nodule diagnosis has not been previously described. We investigated malignancy rate and probability by using various malignancy stratification systems in initially biopsy-proven benign nodules and sought to determine their clinical relevance in management of benign thyroid nodules.

**Methods** This retrospective study collected 6762 thyroid nodules from 6493 consecutive patients who underwent biopsy between January 2013 and December 2013. The initial biopsy with  $\geq 1$  year of follow-up was used as the gold standard for benign diagnosis. For our study purpose, we analyzed 2747 (57.0%, 2747 of 4822, 532 women, 2111 men; 229 malignant and 2518 benign) thyroid nodules diagnosed by initial biopsy with  $28.2 \pm 9.1$  (range, 12–41) months of follow-up. We calculated the malignancy probability of thyroid nodules by using various malignancy risk stratification systems.

**Results** The overall calculated thyroid malignancy rate was 8.3% (229 of 2747). Initially biopsy-proven benign nodules exhibited a  $\leq 3.0\%$  malignancy probability when assessed as “low suspicion” by Korean-TIRADS (K-TIRADS), “low suspicion” by the ATA guideline, and “very probably benign” by the French TIRADS guideline and gave a score of  $\leq 3$  by the web-based TIRADS.

**Conclusion** When initially biopsy-proven benign nodules exhibit a “low suspicion” US pattern and low malignancy probability, as stratified by various TIRADS guidelines, imaging surveillance instead of second biopsy is warranted.

## Key Points

- One biopsy is sufficient for initially biopsy-proven benign nodules.
- Repetitive biopsy is necessary for imaging–pathology mismatched nodules.
- Scoring risk stratification permits personalized management.

**Keywords** Thyroid nodule · Core needle biopsy · Fine needle aspiration · Diagnostic imaging · Guideline

✉ Jung Hwan Baek  
radbaek@naver.com

- <sup>1</sup> Department of Radiology and Thyroid Center, Chung-Ang University Hospital, Chung-Ang University College of Medicine, Seoul 06973, Korea
- <sup>2</sup> Department of Radiology and the Research Institute of Radiology, University of Ulsan College of Medicine, Asan Medical Center, 88, Olympic-ro 43-gil, Songpa-gu, Seoul 05505, Korea
- <sup>3</sup> Department of Surgery, University of Ulsan College of Medicine, Asan Medical Center, 88, Olympic-ro 43-gil, Songpa-gu, Seoul 05505, Korea
- <sup>4</sup> Department of Internal Medicine, University of Ulsan College of Medicine, Asan Medical Center, 88, Olympic-ro 43-gil, Songpa-gu, Seoul 05505, Korea

## Abbreviations

ATA	American Thyroid Association
CNB	Core needle biopsy
FNA	Fine needle aspiration
TIRADS	Thyroid Imaging Reporting and Data System
US	Ultrasound

## Introduction

Fine needle aspiration (FNA) has reported false-negative rates between 0.7% and 11.0% [1], and this variability requires repeated biopsy of benign thyroid nodules determined at the first biopsy. However, some authors have reported low false

negativity [2, 3] and suggested that there is limited benefit from repeated biopsy. Indeed, in a study of 2000 benign thyroid nodules that were followed for 8 years, there were no patient deaths due to thyroid cancer, thyroidectomy was performed in 24.0%, and the false-negative rate was estimated to be 1.3% [4]. In the absence of clinical or sonographic features of concern, repeated biopsy rarely yields a different result; thus, in many cases, the first biopsy without further intervention is sufficient.

The revised American Thyroid Association (ATA) guidelines emphasize the concept of “less is more” [5], such that less extensive surgeries and less surveillance testing will become more common because it is reasonable to do less to provide responsibly individualized therapy for patients [6]. Indeed, several researchers [7, 8] have reported that suspicious sonographic features are better predictors of malignancy in thyroid nodules with previous benign cytology [9, 10]. This integrated management, encompassing pathology results and sonographic features, is now more frequently advocated because routine repeat biopsy of initially biopsy-proven benign nodules may not be necessary in all cases.

Recently, various Thyroid Imaging Reporting and Data System (TIRADS) guidelines have been reported from several institutions and societies [11–15]. To our knowledge, the integration of TIRADS guidelines in initially biopsy-proven benign nodules has not been previously described. Therefore, our aim was to evaluate whether one biopsy is sufficient for benign thyroid nodule diagnosis via comprehensive analysis of initially biopsy-proven benign thyroid nodules by various TIRADS guidelines.

## Materials and Methods

### Patient selection

This retrospective study was approved by the institutional review board of the Asan Medical Center, and the institutional review board gave a waiver of informed consent for the use of the data. The study population was obtained from a historical cohort of 6762 thyroid nodules from 6493 consecutive patients, all of whom underwent core needle biopsy (CNB) or FNA between January 2013 and December 2013. These patients were enrolled in a previous study regarding the efficacy and safety of CNB in initially detected thyroid nodules [16]. Excluding nodules that previously underwent CNB or FNA ( $n = 1940$ ), the final diagnosis was obtained for 2114 nodules from 1928 patients in the CNB group and 2708 nodules from 2625 patients in the FNA group. As a reference standard, the final diagnoses of benign nodules were determined by (i) pathology results of

surgical resections, (ii) benign cytology results of FNA or CNB procedures that were repeated at least twice, (iii) an initial benign result from FNA or CNB and decreased or stable nodule size at an ultrasound (US) follow-up of  $\geq 1$  year. Malignancy was diagnosed after surgery or after repeated or initial biopsy. For the purpose of this study, we analyzed 2747 (57.0%, 2747 of 4822) thyroid nodules from 2643 patients that were diagnosed by initial biopsy with  $28.2 \pm 9.1$  months follow-up (range, 12–41 months).

### US-guided FNA and CNB procedures

US images were obtained for the evaluation of thyroid nodules before the US-guided procedure by using either an HDI 5000 (ATL Ultrasound, Philips) or Sequoia (Acuson, Siemens Healthineers) instrument equipped with a 5–12 MHz or an 8–15 MHz linear-array transducer. All US-guided procedures were performed by radiologists under the supervision of two faculty radiologists (J.H.B. and J.H.L., with 19 and 14 years of clinical experience, respectively, in performing and evaluating thyroid US). US-guided CNB and FNA procedures for thyroid nodules were performed according to current practice guidelines [17–20]. US-guided CNBs were performed under local anesthesia with 1% lidocaine by using a disposable 1.1- or 1.6-cm excursion, 18-gauge, double-action, spring-activated needle (TSK Ace-cut; Create Medic) [21, 22]. Before insertion of the core needle, the vessels along the approach route were evaluated by power Doppler US to prevent hemorrhage. Using a freehand technique, we advanced the core needle from the isthmus of the thyroid toward the nodule [22]. After the needle tip had been advanced toward the edge of the nodule, the distance of firing (1.1 or 1.6 cm) was measured before sequential firing of the stylet and cutting cannula of the needle [22]. A total of 1056 nodules in 983 patients were examined with CNB: 101 nodules had suspicious features seen on US, 270 heavily calcified nodules, 51 predominantly cystic nodules, and the remaining 634 nodules underwent CNB at the request of their referring physicians. Either US-guided FNA or CNB was used at the discretion of the US operator or referring physician without definite indications for selection.

US-guided FNAs were routinely performed by using a 23-gauge needle. In all cases, direct smears were made and immediately fixed with alcohol after the FNA procedure, then stained by using the Papanicolaou method [21]. The number of needle passes was determined by the operator during the procedure, and a maximum of four needle passes were permitted for each nodule. Additional FNAs were recommended in the case of incomplete visual assessment results. In the case of

US-guided CNB, the adequacy of the procedure was monitored via real-time US, and the adequacy of the tissue core was assessed by visual inspection [22]. An additional CNB was performed when the targeting of the lesion was considered to be inaccurate. The maximum number of CNBs performed during a single session was two. Each patient was observed after firm local compression of the biopsy site for 10–20 min after the biopsy. If patients complained of pain or neck swelling, a repeat US examination was performed to evaluate possible complications.

All CNB specimens and FNA cytological analysis were reviewed by a thyroid cytopathologist (D.E.S., with 11 years of clinical experience in thyroid cytopathology). FNA cytology diagnoses were categorized into six categories according to the Bethesda System for Reporting Thyroid Cytopathology [21, 23, 24]. Although standardization of CNB diagnostic criteria for thyroid nodules had not yet been established during our study period, the histological results of CNB were categorized into the six categories of the Bethesda System [20, 21, 23, 25, 26].

### Analysis of US findings and malignancy rate

When analyzing the US images, the two radiologists (S.M.H. and J.H.B. with 9 and 20 years of clinical experience, respectively, in performing and evaluating thyroid US) assessed the thyroid nodules by using criteria that were obtained from published reports [11–13, 27], including size, internal content (solid, predominantly solid, predominantly cystic, cystic), shape (ovoid to round, taller than wide, irregular), margin (well-defined smooth, microlobulated or spiculated, ill-defined), echogenicity of the solid portion (hyper- or iso-echogenicity, hypoechoogenicity or marked hypoechoogenicity), and the presence of microcalcification, macrocalcification, and/or rim calcification. The relationship between the final diagnosis and US features was assessed [13, 28]. The suspicious US features, as suggested by the Korean Society of Thyroid Radiology (KSThR) and Moon et al, are as follows: taller than wide shape, spiculated/microlobulated margin, marked hypoechoogenicity, and presence of microcalcifications [13, 27]. A thyroid nodule with at least one of these suspicious US findings was classified as a suspicious nodule during US examination. We calculated the probability of malignancy by using a web-based TIRADS ([www.gap.kr/thyroidnodule.php](http://www.gap.kr/thyroidnodule.php)) [14], the Korean-TIRADS (K-TIRADS) guideline [12], the revised ATA guidelines [11], and the French TIRADS guideline (proposed by Russ et al) [15]. For our analysis, suspicious nodules were defined as those with a score of  $\geq 8$  on the web-based TIRADS, high suspicion according to the K-TIRADS guideline, high suspicion according to the ATA

guidelines, and a score of  $\geq 4B$  according to the French TIRADS guideline.

### Statistical analysis

Statistical analysis was performed by using the SPSS software package (Version 19.0 for Windows; SPSS, IBM). Each of the US characteristics was analyzed to determine its association with a benign or malignant diagnosis. Categorical data were summarized by using frequencies and percentages. Either the chi-squared test or Student's *t*-test was used to evaluate the relationship between the variable factors and benignity or malignancy of the nodules. Two-tailed *p*-values of  $< 0.05$  were considered as indicative of statistical significance.

## Results

### Demographic data and US characteristics of all nodules

The size of nodules that were diagnosed by initial biopsy ranged from 2 mm to 96 mm (mean size, 14.4 mm). The proportion of nodules with a diameter  $< 1$  cm was 37.0% (1017/2747) and those with a diameter  $\geq 1$  cm was 63.0% (1730/2747). Two nodules that measured 2 mm underwent FNA because of contralateral cancer and to decide on performing either lobectomy or total thyroidectomy depending on the diagnosis of the nodule. The malignancy rate was significantly higher in the thyroid nodules with a diameter  $< 1$  cm than in the nodules with a diameter  $\geq 1$  cm (14.2% vs. 4.9%, respectively;  $p < 0.001$ ; Table 1). Significant differences were found in all US features representing benign and malignant nodules, regardless of size (Table 2). None of the included patients experienced any major complications associated with intervention or hospitalization. Four patients (0.4%; four of 1056) developed a hematoma after the CNB, but resolution of the hematoma occurred following compression and rest for 1 h.

Overall, the calculated thyroid malignancy rate was 8.3% (229/2747) when an initial biopsy with  $\geq 1$  year of follow-up was used as the gold standard for benign diagnosis. For malignant nodules ( $n = 229$ , 8.3%), the diagnoses upon histological examination included papillary thyroid carcinomas (PTC) ( $n = 222$ ), medullary carcinomas ( $n = 2$ ), anaplastic carcinoma ( $n = 3$ ), and metastasis ( $n = 2$ ). The diagnoses of benign lesions ( $n = 2518$ , 91.7%) upon histological examination comprised 2518 benign non-neoplastic nodules, including nodular hyperplasia, adenomatous goiter, follicular proliferating lesions, and thyroiditis.

**Table 1** Demographic data of patients in this study (n=2747)

Characteristics	Benign (n =2518, 91.7%)	Malignant (n =229, 8.3%)	p-value
Age (mean), years <sup>a</sup>	54.5 ± 11.8	51.4 ± 14.4	0.0027
Sex <sup>a</sup>			0.057
Male	1950 (92.4)	161 (7.6)	
Female	478 (89.9)	54 (10.1)	
Size of nodule			< 0.001
< 1 cm	873 (85.8)	144 (14.2)	
≥ 1 cm	1645 (95.1)	85 (4.9)	

Data indicate the number of lesions

Numbers in parentheses indicate percentages

<sup>a</sup> per patient analysis (n=2643)

### Comparison of the malignancy probability assessed by various malignant risk systems

Table 3 shows the malignancy probability of initially biopsy-proven benign thyroid nodules according to various TIRADS guidelines. The malignancy probability of the “low suspicion”

category by K-TIRADS (category 3) is 1.3%, the “low suspicion” category of ATA is 1.2%, and the “very probably benign” category of the French TIRADS guideline (category 3) is 1.2%, all lower than the ≤3.0% malignancy probability. Applying the web-based TIRADS [14], we could subdivide each malignancy probability by using a score range of 0–13. A cutoff

**Table 2** Ultrasonography characteristics of benign and malignant thyroid nodules in this study

Characteristics	Benign (n =2518)	Malignant (n =229)	p-value
Internal content			< 0.001
Solid	1687 (88.8)	213 (11.2)	
Predominantly solid	614 (97.8)	14 (2.2)	
Predominantly cystic	192 (99.0)	2 (1.0)	
Cystic	25 (100.0)	0 (0.0)	
Shape			< 0.001
Ovoid to round	2234 (94.2)	138 (5.8)	
Irregular	206 (85.1)	36 (14.9)	
Taller than wide	78 (58.7)	55 (41.3)	
Margin			< 0.001
Well-defined smooth	1581 (96.0)	66 (4.0)	
Ill-defined	870 (92.0)	76 (8.0)	
Spiculated/microlobulated	67 (43.5)	87 (56.5)	
Echogenicity			< 0.001
Isoechogenicity	1344 (98.3)	23 (1.7)	
Hypoechoogenicity	1002 (88.0)	137 (12.0)	
Marked hypoechoogenicity	127 (64.8)	69 (35.2)	
Hyperechogenicity	45 (100.0)	0 (0.0)	
Calcification			< 0.001
No calcification	2058 (94.9)	111 (5.1)	
Microcalcification	110 (64.3)	61 (35.7)	
Macrocalcification	262 (83.4)	52 (16.6)	
Rim	88 (94.6)	5 (5.4)	

Data indicate the number of lesions.

Numbers in parentheses indicate percentages.

US ultrasonography

**Table 3** Malignancy rate according to various Thyroid Image Reporting and Data System (TIRADS) guidelines and a web-based risk stratification system (n=2747)

Total	Total	Benign	Malignant	Calculated malignancy rate <sup>a</sup>	Suggested malignancy rate
<b>K-TIRADS</b>					
3 (Low)	1513	1493	20	1.3% <sup>b</sup>	3–15%
4 (Intermediate)	880	790	90	10.2%	15–50%
Isohyperechoic with any suspicious US feature	37	31	6	16.2%	
Solid hypoechoic nodule without any suspicious US feature	820	742	78	9.5%	
Partially cystic nodule with any suspicious US feature	23	17	6	26.1%	
5 (High)	278	160	118	42.4%	> 60%
<b>ATA guideline</b>					
Low	1358	1342	16	1.2% <sup>b</sup>	< 3%
Intermediate	644	619	45	7.0%	10–20%
High	292	169	123	42.1%	70–90%
<b>French TIRADS</b>					
3 (very probably benign)	1365	1349	16	1.2% <sup>b</sup>	0.25%
4a (low risk)	957	892	65	6.8%	6%
4b (high risk)	367	253	114	31.1%	69%
<b>Web-based</b>					
<8	2546	2409	137	5.4%	
0	429	426	3	0.7% <sup>b</sup>	0–4.8%
1	184	182	2	1.1% <sup>b</sup>	0–6.7%
2	633	624	9	1.4% <sup>b</sup>	6.8–9.6%
3	325	319	6	1.8% <sup>b</sup>	9.5–13.8%
4	466	432	34	7.3%	12.1–17.5%
5	278	250	28	10.1%	33.3–38.9%
6	130	98	32	24.6%	47.0–66.7%
7	101	78	23	22.8%	54.8–58.5%
≥8	201	109	92	45.8%	
8	75	40	35	46.7%	69.6–82.1%
9	49	30	19	38.8%	71.4–86.7%
10	41	21	20	48.8%	83.3–87.0%
11	22	11	11	50.0%	93.0–93.2%
12	5	4	1	20.0%	100.0%
13	9	3	6	66.7%	100.0%

<sup>a</sup> Calculated malignancy rate of our results

<sup>b</sup> Malignancy rate  $\leq$  3.0%

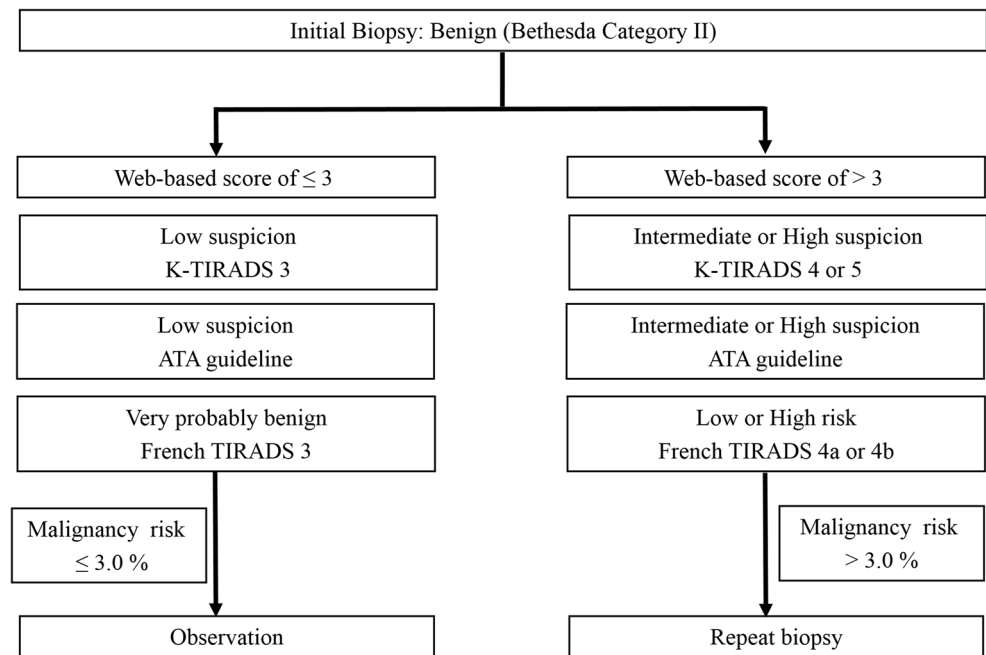
US ultrasonography, TIRADS Thyroid Image Reporting and Data System, K-TIRADS Korean-TIRADS, ATA American Thyroid Association

of  $\leq$ 3.0% malignancy probability was observed at a score of 3 (1.8%) and increased substantially at a score of 4 (7.3%) (Fig. 1). We additionally analyzed subgroups of nodules that matched the criteria for the “intermediate suspicion” category by K-TIRADS guidelines (category 4), divided into three categories: i) iso- or hyperechoic nodule with any suspicious US feature, ii) solid hypoechoic nodule without any suspicious US feature, and iii) partially cystic nodule with any suspicious US feature [12]. The lowest malignancy probability was observed in solid hypoechoic nodules without any suspicious US feature (9.5%).

## Discussion

Our study results demonstrated that initially biopsy-proven benign nodules exhibited a  $\leq$ 3.0% malignancy probability when assessed as “low suspicion” by K-TIRADS (category 3) guideline, “low suspicion” by the ATA guideline, the “very probably benign” category by the French TIRADS guideline (category 3), and a score of  $\leq$ 3 by the web-based TIRADS. Therefore, one biopsy was found to be sufficient for these thyroid nodules. However, even when the initial biopsy result

**Fig. 1** Management pathway algorithm of initially biopsy-proven benign thyroid nodule by various US-based risk stratification systems



was benign, the lowest malignancy probability was 6.8% in nodules with a “high suspicion” US pattern. Therefore, one biopsy was found to be insufficient for the initially biopsy-proven benign nodules with a higher TIRADS assessment category. These results stressed the value of a combination of biopsy results and TIRADS assessment categories. Moreover, current thyroid radiofrequency ablation (RFA) guidelines require pathological confirmation of benign status at two separate US-guided FNAs before RFA. The current study results showed that two biopsy procedures before RFA were unnecessary. This more conservative paradigm shift regarding initially proven benign nodules is concordant with the “less is more” theme suggested by the 2015 ATA guideline [5], which suggests that less surveillance testing will become more common because it is reasonable to do less.

Several investigators [10, 11, 29–32] have emphasized the combination of US features and pathological information rather than merely requiring nodule growth as an indication for repeat biopsy. The ATA guideline [11] states that given the low false-negative rate of nodules with prior benign cytology and higher yield of missed malignancy based on US pattern rather than on nodule growth, management should be determined by risk stratification based on US pattern. In a prospective study [31], the cytologically benign nodules without suspicious features exhibited only a 0.7% incidence of cancer and 1.1% false-negative rate during 5 years of follow-up. Therefore, repeated biopsy is unnecessary for these nodules. In contrast, the malignancy risk (20.4–56.6%) has been found to be significantly higher in benign nodules with accompanying suspicious US features [9, 10, 29, 32]; thus, repetitive biopsy or even diagnostic surgery is

necessary for US–pathology mismatched nodules (28, 29, 38, 52). This approach is suggested by the ATA guidelines as well [11]. To justify our results, we applied various TIRADS guidelines and achieved a relatively low malignancy probability in initially biopsy-proven benign nodules with “low suspicion” US patterns. Similarly, Hong et al [30] reported a malignancy probability of <3.0% (0%, 0.4%, and 2.4%, respectively) in thyroid nodules with benign cytology results stratified according to K-TIRADS categories 2, 3, and 4 and recommended observation in lieu of repeated biopsy. Nevertheless, the initially biopsy-proven benign nodules with higher TIRADS assessment categories by various TIRADS guidelines exhibited a malignancy probability of ≥6.8%, which is higher than the malignancy probability of the benign category (0–3.0%) [24].

For malignancy risk stratification, several attempts have been made to convert the TIRADS “pattern-based” approach to a “score-based” approach [33, 34]. Such scoring risk stratification models can be implemented in clinical practices with varying volumes of patients and in sites with varying levels of professional expertise. Moreover, they permit more personalized management, with >10 different ranges of malignancy risk scores for thyroid nodules [35]. More recently, the American College of Radiology (ACR) developed the TIRADS [36] by allocating points to more suspicious features, summing the points, and determining the TIRADS category of the nodules. With a lower malignancy risk assigned to benign nodules without any suspicious US features or a score of 0, a lower biopsy rate can be expected. According to the web-based scoring risk stratification model ([www.gap.kr/thyroidnodule.php](http://www.gap.kr/thyroidnodule.php)), the overall risk of malignancy for score

0 nodules was <5% [14]. Our current risk results equated to 1.8% in score 3 nodules and increased to 7.3% in score 4 nodules based on the web-based scoring system. Thus, by integrating the scoring systems, prediction of the malignancy probability of initially biopsy-proven benign thyroid nodules will increase biopsy efficacy regarding repeated biopsy, thereby enabling more personalized and optimized management.

Given the potentially higher false-negative rate of biopsy in larger nodules, some groups have suggested performing lobectomy when thyroid nodules are above a certain size [37]. In addition to surgery, RFA is one of management options for reducing nodule volume and relieving nodule-related clinical problems. US-guided RFA is a nonsurgical technique that has been used for the treatment of benign nodules [38]. Pathological confirmation of benignity at two separate US-guided FNAs is usually a requirement before an RFA procedure [39]. When we analyzed initially biopsy-proven benign thyroid nodules according to size, we observed consistently low malignancy probabilities when assessed as “low suspicion” by K-TIRADS (category 3, 1.4% malignancy probability in nodules >3 cm), “low suspicion” by ATA guideline (1.5%), “very probably benign” in French TIRADS guideline (category 3, 1.5%), and a score of  $\leq 3$  (1.2%) by web-based TIRADS. With further accumulation of evidence indicating that one biopsy result is sufficient for benign thyroid nodule diagnosis, patients may proceed to ablation therapy and be relieved of nodule-related clinical problems more promptly.

There were several limitations in our study. First, its retrospective design may have induced selection bias; however, such a problem would be overcome to a certain extent by the large study population. Second, we assumed that benign cases were “truly benign” after  $\geq 1$  year of follow-up; however, the fact that no test has a 100.0% negative predictive value may explain the relatively high calculated thyroid malignancy rate of 8.3% observed in our study. In addition, 2.8% (77/2,747) of the nodules measured >3 cm and were palpable. Thus, larger prospective cohorts are required to determine the number of benign nodules that behave as “truly benign” and to investigate any significance regarding an asymptomatic patient population for screening. Because of the relatively short follow-up period, we could not assess the long-term outcome or growth of these benign nodules. In fact, a previous study [40] that compared the rates of new diagnosis of malignancy between long-term and short-term follow-up of patients with biopsy-proven benign nodules found no significant difference in outcomes. Last, we did not evaluate the interobserver variability.

In conclusion, when initially biopsy-proven benign nodules exhibiting a “low suspicion” US pattern and low malignancy probability are stratified by various TIRADS, imaging surveillance, rather than second biopsy, is warranted. However, repeated biopsy should be performed for thyroid nodules with suspicious US features or those that belong to

higher TIRADS assessment categories even when initial biopsy results indicate that the lesion is benign.

**Funding** The authors state that this work has not received any funding.

## Compliance with ethical standards

**Guarantor** The scientific guarantor of this publication is Jung Hwan, Baek

**Conflict of interest** The authors of this manuscript declare no relationships with any companies, whose products or services may be related to the subject matter of the article.

**Statistics and biometry** No complex statistical methods were necessary for this paper.

**Informed consent** Written informed consent was waived by the Institutional Review Board.

**Ethical approval** Institutional Review Board approval was obtained.

**Study subjects or cohorts overlap** Some study subjects or cohorts have been previously reported in Scientific Report 7:8242 by Suh CH et al.

## Methodology

- retrospective
- case-control study
- performed at one institution

## References

1. Wang CC, Friedman L, Kennedy GC et al (2011) A large multicenter correlation study of thyroid nodule cytopathology and histopathology. *Thyroid* 21:243–251
2. Lucas A, Llatjos M, Salinas I, Reverter J, Pizarro E, Sanmarti A (1995) Fine-needle aspiration cytology of benign nodular thyroid disease. Value of re-aspiration. *Eur J Endocrinol* 132:677–680
3. Kessler A, Gavriel H, Zahav S et al (2005) Accuracy and consistency of fine-needle aspiration biopsy in the diagnosis and management of solitary thyroid nodules. *Isr Med Assoc J* 7:371–373
4. Nou E, Kwong N, Alexander LK, Cibas ES, Marqusee E, Alexander EK (2014) Determination of the optimal time interval for repeat evaluation after a benign thyroid nodule aspiration. *J Clin Endocrinol Metab* 99:510–516
5. Kim BW, Yousman W, Wong WX, Cheng C, McAninch EA (2016) Less is More: Comparing the 2015 and 2009 American Thyroid Association Guidelines for Thyroid Nodules and Cancer. *Thyroid* 26:759–764
6. Kim TY, Shong YK (2017) Active Surveillance of Papillary Thyroid Microcarcinoma: A Mini-Review from Korea. *Endocrinol Metab (Seoul)* 32:399–406
7. Oertel YC, Miyahara-Felipe L, Mendoza MG, Yu K (2007) Value of repeated fine needle aspirations of the thyroid: an analysis of over ten thousand FNAs. *Thyroid* 17:1061–1066
8. Mazzawi SJ, Rosen G, Luboshitzky R, Dharan M (2000) Management of benign thyroid nodules based on the findings of fine-needle aspiration. *J Otolaryngol* 29:95–97

9. Kwak JY, Koo H, Youk JH et al (2010) Value of US correlation of a thyroid nodule with initially benign cytologic results. *Radiology* 254:292–300
10. Kwak JY, Kim EK, Kim HJ, Kim MJ, Son EJ, Moon HJ (2009) How to combine ultrasound and cytological information in decision making about thyroid nodules. *Eur Radiol* 19:1923–1931
11. Haugen BR, Alexander EK, Bible KC et al (2016) 2015 American Thyroid Association Management Guidelines for Adult Patients with Thyroid Nodules and Differentiated Thyroid Cancer: The American Thyroid Association Guidelines Task Force on Thyroid Nodules and Differentiated Thyroid Cancer. *Thyroid* 26:1–133
12. Shin JH, Baek JH, Chung J et al (2016) Ultrasonography Diagnosis and Imaging-Based Management of Thyroid Nodules: Revised Korean Society of Thyroid Radiology Consensus Statement and Recommendations. *Korean J Radiol* 17:370–395
13. Kwak JY, Han KH, Yoon JH et al (2011) Thyroid imaging reporting and data system for US features of nodules: a step in establishing better stratification of cancer risk. *Radiology* 260:892–899
14. Choi YJ, Baek JH, Baek SH et al (2015) Web-Based Malignancy Risk Estimation for Thyroid Nodules Using Ultrasonography Characteristics: Development and Validation of a Predictive Model. *Thyroid* 25:1306–1312
15. Russ G (2016) Risk stratification of thyroid nodules on ultrasonography with the French TI-RADS: description and reflections. *Ultrasonography* 35:25–38
16. Suh CH, Baek JH, Choi YJ et al (2017) Efficacy and safety of core-needle biopsy in initially detected thyroid nodules via propensity score analysis. *Sci Rep* 7:8242
17. Lee YH, Baek JH, Jung SL, Kwak JY, Kim JH, Shin JH (2015) Ultrasound-guided fine needle aspiration of thyroid nodules: a consensus statement by the Korean Society of Thyroid Radiology. *Korean J Radiol* 16:391–401
18. Na DG, Baek JH, Jung SL et al (2017) Core Needle Biopsy of the Thyroid: 2016 Consensus Statement and Recommendations from Korean Society of Thyroid Radiology. *Korean J Radiol* 18:217–237
19. Baek JH (2017) Current status of core needle biopsy of the thyroid. *Ultrasonography* 36:83–85
20. Jung CK, Baek JH (2017) Recent Advances in Core Needle Biopsy for Thyroid Nodules. *Endocrinol Metab (Seoul)* 32:407–412
21. Na DG, Kim JH, Sung JY et al (2012) Core-Needle Biopsy Is More Useful Than Repeat Fine-Needle Aspiration in Thyroid Nodules Read as Nondiagnostic or Atypia of Undetermined Significance by the Bethesda System for Reporting Thyroid Cytopathology. *Thyroid* 22:468–475
22. Suh CH, Baek JH, Lee JH et al (2016) The Role of Core-Needle Biopsy as a First-Line Diagnostic Tool for Initially Detected Thyroid Nodules. *Thyroid* 26:395–403
23. Cibas ES, Ali SZ, Conference NCITFSS (2009) The Bethesda System For Reporting Thyroid Cytopathology. *Am J Clin Pathol* 132:658–665
24. Cibas ES, Ali SZ (2017) The 2017 Bethesda System for Reporting Thyroid Cytopathology. *Thyroid* 27:1341–1346
25. Choi YJ, Baek JH, Ha EJ et al (2014) Differences in Risk of Malignancy and Management Recommendations in Subcategories of Thyroid Nodules with Atypia of Undetermined Significance or Follicular Lesion of Undetermined Significance: The Role of Ultrasound-Guided Core-Needle Biopsy. *Thyroid* 24:494–501
26. Yeon JS, Baek JH, Lim HK et al (2013) Thyroid nodules with initially nondiagnostic cytologic results: the role of core-needle biopsy. *Radiology* 268:274–280
27. Moon WJ, Jung SL, Lee JH et al (2008) Benign and malignant thyroid nodules: US differentiation—multicenter retrospective study. *Radiology* 247:762–770
28. Na DG, Kim DS, Kim SJ, Ryoo JW, Jung SL (2016) Thyroid nodules with isolated macrocalcification: malignancy risk and diagnostic efficacy of fine-needle aspiration and core needle biopsy. *Ultrasonography* 35:212–219
29. Shin JH, Han BK, Ko K, Choe YH, Oh YL (2006) Value of repeat ultrasound-guided fine-needle aspiration in nodules with benign cytological diagnosis. *Acta Radiol* 47:469–473
30. Hong MJ, Na DG, Baek JH, Sung JY, Kim JH (2017) Cytology-Ultrasonography Risk-Stratification Scoring System Based on Fine-Needle Aspiration Cytology and the Korean-Thyroid Imaging Reporting and Data System. *Thyroid* 27:953–959
31. Durante C, Costante G, Lucisano G et al (2015) The natural history of benign thyroid nodules. *JAMA* 313:926–935
32. Chemyavsky VS, Shanker BA, Davidov T et al (2012) Is one benign fine needle aspiration enough? *Ann Surg Oncol* 19:1472–1476
33. Kwak JY, Jung I, Baek JH et al (2013) Image reporting and characterization system for ultrasound features of thyroid nodules: multicentric Korean retrospective study. *Korean J Radiol* 14:110–117
34. Park JY, Lee HJ, Jang HW et al (2009) A proposal for a thyroid imaging reporting and data system for ultrasound features of thyroid carcinoma. *Thyroid* 19:1257–1264
35. Ha SM, Ahn HS, Baek JH et al (2017) Validation of Three Scoring Risk-Stratification Models for Thyroid Nodules. *Thyroid* 27:1550–1557
36. Tessler FN, Middleton WD, Grant EG et al (2017) ACR Thyroid Imaging, Reporting and Data System (TI-RADS): White Paper of the ACR TI-RADS Committee. *J Am Coll Radiol* 14:587–595
37. McCoy KL, Jabbour N, Ogilvie JB, Ohori NP, Carty SE, Yim JH (2007) The incidence of cancer and rate of false-negative cytology in thyroid nodules greater than or equal to 4 cm in size. *Surgery* 142:837–844 discussion 844.e831-833
38. Bernardi S, Dobrinja C, Fabris B et al (2014) Radiofrequency ablation compared to surgery for the treatment of benign thyroid nodules. *Int J Endocrinol* 2014:934595
39. Baek JH, Kim YS, Lee D, Huh JY, Lee JH (2010) Benign predominantly solid thyroid nodules: prospective study of efficacy of sonographically guided radiofrequency ablation versus control condition. *AJR Am J Roentgenol* 194:1137–1142
40. Lee S, Skelton TS, Zheng F et al (2013) The biopsy-proven benign thyroid nodule: is long-term follow-up necessary? *J Am Coll Surg* 217:81–88 discussion 88-89