Long-term outcomes of cytologically benign thyroid tumors: a retrospective analysis of 3,102 patients at a single institution

Minoru Kihara¹, Akira Miyauchi¹, Mitsuyoshi Hirokawa², Hiroo Masuoka¹, Takuya Higashiyama¹, Naoyoshi Onoda¹, Yasuhiro Ito¹ and Akihiro Miya¹

¹Departments of Surgery, Kuma Hospital, Hyogo 650-0011, Japan
²Departments of Diagnostic Pathology, Kuma Hospital, Hyogo 650-0011, Japan

Abstract. Some thyroid tumors that are cytologically diagnosed as benign may be pathologically diagnosed as malignant. Here, we investigated the long-term outcomes of patients with thyroid tumors with benign cytology, and the factors for malignancy. We retrospectively reviewed the cases of 3,102 patients with thyroid tumors >1 cm cytologically diagnosed as benign at our hospital during a 1-year period from January 2007. The median follow-up duration for all patients was 68.7 (range 0.0–168.7) months. Immediate surgery and delayed surgery were performed in 393 and 148 patients, respectively. Eventually, 541 (17.4%) of the 3,102 patients underwent a thyroidectomy, and 2,561 (82.6%) were observed without surgery. Among the surgically treated patients, the tumors of 525 (97.0%) and 16 (3.0%) were pathologically diagnosed as benign and malignant, respectively. There was no significant difference in age, gender, tumor size, serum thyroglobulin level at surgery, or the tumor volume-doubling rate (TV-DR) between the benign and malignant cases. Only the ultrasonographic findings based on our hospital’s classification system were directly and significantly linked to pathological diagnosis (p < 0.01).

Among the tumors of the 667 patients who were followed without surgery for >10 years, 89.9% remained unchanged and 7.2% were reduced in size. Ultrasonographic evaluation provides important information for therapeutic decision-making regarding surgery versus observation for cytologically benign tumors.

Key words: Cytologically benign thyroid tumor, Fine needle aspiration cytology, Ultrasonographic finding, Long-term outcome

FINE NEEDLE ASPIRATION CYTOLOGY (FNAC) is the most sensitive and specific test for the diagnosis of thyroid tumors. The decision regarding surgical intervention versus observation for thyroid tumors requires conclusive cytological evidence from a FNAC specimen. Indeed, in the absence of specific surgical indications, most cytologically benign thyroid tumors are simply observed, and thus the frequency of surgical removal for these tumors is low. Nonetheless, because neither the sensitivity nor the specificity of FNAC is 100% [1-5], some thyroid tumors that are cytologically diagnosed as benign will later be revealed to be malignant after surgical resection and pathological examination. According to the recent American Thyroid Association (ATA) guidelines [5], the risk of a false-negative interpretation, although very low, necessitates that benign thyroid tumors be followed up over the long term with periodic ultrasonography (US) monitoring every 6–18 months from the initial FNAC.

Among the several modalities available for the imaging of thyroid tumors, US has the advantage of facilitating both tumor detection and qualitative evaluation [2, 3, 6-11]. Moreover, we previously demonstrated that US is specifically useful for qualitative evaluation of thyroid tumors that were cytologically diagnosed as follicular or benign [2, 3]. We conducted the present study to retrospectively investigate the long-term (≤14 years) outcomes of 3,102 patients with thyroid tumors cytologically diagnosed as benign. We also investigated several factors, i.e., US evaluation, age, gender, tumor size, serum thyroglobulin level, changes in tumor size and changes in US diagnoses, for their potential association
with malignancy. This is the first study to use the tumor volume-doubling rate (TV-DR) to evaluate the kinetics of changes in cytologically benign tumors.

Patients and Methods

We routinely perform US for all patients with thyroid diseases who visit our hospital. When a thyroid tumor is revealed by US, we evaluate the ultrasonographic findings based on our classification system, the Kuma Hospital Ultrasound Classification, as described in previous reports [2, 3, 6, 9]. Briefly, the ultrasonographic classification system is as follows. Class 1: Round or oval anechoic lesion. Class 2: Regular-shaped tumor with cystic change; the echo level of a solid lesion is similar to that of a normal thyroid. Class 3: Solid and regular-shaped tumor; the internal echo is homogeneous or may have strong echoes internally or at the capsule. Class 4: Solid and regular-shaped tumor; the internal echo is usually low and may have fine strong echoes internally. Class 5: Solid and irregular-shaped tumor with extrathyroid extension. The ultrasonographic diagnoses are benign (Class 1 or 2), borderline (Class 3), suspicious for malignancy (Class 4), and malignant (Class 5).

In clinical practice, US examiners might categorize marginal tumors between Classes 2 and 3 or between Classes 3 and 4 as Class “2.5” or Class “3.5” tumors, respectively. In the present study, Class 2.5 or less was considered benign, Class 3 as borderline, and Class 3.5 or higher as suspicious for malignancy or malignant. In the study period described below, we routinely performed US-guided FNAC for solid tumors ≥1 cm and suspicious tumors ≥0.5 cm for cytological diagnosis in order to determine how to manage these tumors.

Between January 2007 and December 2007, 3,102 adult patients had a thyroid tumor with a maximum diameter of ≥1 cm with cytological findings classified as benign (equivalent to Category II in the Bethesda System (Be-II) [12]) at our hospital. All 3,102 patients were enrolled in this study. During that year, our cytologic classification was based on five suggested categories: benign (Be-II equivalent), indeterminate (Be-III or IV equivalent), suspicious for malignancy (Be-V equivalent), malignant (Be-VI equivalent), and inadequate (Be-I equivalent). The tumors diagnosed as indeterminate, suspicious for malignancy, or malignant on cytological examination were excluded from the present study. Incidental microscopic malignancies such as papillary microcarcinoma were not included in this study.

At our institution, a cytologically benign thyroid tumor usually warrants observation without immediate surgery. However, we have recommended surgery to patients with a cytologically benign tumor when they had at least one of the following features: (1) a tumor that was ultrasonographically suspicious for malignancy, (2) a solid tumor ≥40 mm in diameter, (3) a tumor with increased size during follow-up, (4) serum thyroglobulin >1,000 ng/mL, (5) a tumor compressing the trachea or esophagus, (6) a tumor extending into the mediastinum, (7) an autonomously functioning thyroid tumor, (8) a tumor that presents cosmetic problems, (9) a preference for surgery on the part of the patient, or (10) the presence of surgical indications for other associated thyroid disease. The patients were followed at our outpatient clinic by a US examination essentially 1×/year. A repeat FNAC was performed in 429 patients because the physician in charge deemed it necessary. A repeat FNAC revealed the following findings: benign (n = 381 patients), a suspicious follicular tumor (n = 46), and undeniable malignancy (n = 2).

Of the 3,102 patients, 393 patients underwent a thyroidectomy within 1 year of the cytological diagnosis (the immediate-surgery group). The remaining 2,709 patients were followed without surgery; however, 148 of them underwent a thyroidectomy ≥12 months later (the delayed-surgery group) because of changes in US features, an increase in tumor size, or the patient’s preference for surgery. Surgical specimens of the 541 thyroidectomy patients (393 + 148) were evaluated histopathologically according to the WHO classification [13].

We also calculated the TV-DR to evaluate the kinetics of changes in tumor volume during the follow-up. In the patients with multiple cytologically benign nodules, the nodule largest at the final follow-up point was used to determine the TV-DR. Details of the calculation of the TV-DR have been described previously [14].

Statistical analyses

The χ² (Chi-squared) test or Mann-Whitney U-test was used for statistical analysis. A p-value <0.05 was regarded as significant. All analyses were performed using StatFlex 7.0 software (Artech, Osaka, Japan).

Results

The median follow-up duration for all patients was 68.7 (range 0.0–168.7) months. The median period until surgery for the surgical cases was 3.9 months (range 0.3–161.4 months). The reasons for the immediate and delayed surgeries are summarized in Table 1. In total, 541 (17.4%) of the 3,102 patients underwent a thyroidectomy, and the other 2,561 (82.6%) patients were observed without surgery (Fig. 1). Table 2 summarizes the clinical characteristics of all patients at their initial cytological diagnoses. The median period between the initial diagnosis and the delayed surgery was 57.8
months (range 12.1–161.4 months). One of the patients with malignancy was diagnosed 111.8 months after the initial diagnosis by FNAC. The median time interval to delayed surgery for the patients who were eventually diagnosed with thyroid cancer was >6 years. During this period of observation, none of these patients developed advanced thyroid cancer. They all had T1, T2, or T3 tumors without lymph node or distant metastases. To date there have been no apparent recurrences. Table 3 provides the clinical characteristics of the patients who underwent immediate surgery or delayed surgery (Supplementary Table 1).

The remaining 2,561 patients underwent only a follow-up during the study period. The median follow-up period in these patients without surgery was 66.4 months (range 0.0–168.7 months), and the follow-up rates at 5, 7, and 10 years in this group were 44.6% (1,141 patients), 36.5% (936 patients) and 26.0% (667 patients), respectively. The median tumor size at cytological diagnosis in the patients with surgery was significantly larger

### Table 1  Reasons for immediate and delayed surgery

| Reason(s) for surgery                                      | Immediate surgery (n = 393) | Delayed surgery (n = 148) |
|------------------------------------------------------------|----------------------------|--------------------------|
| Ultrasonographical suspicious malignancy or borderline     | 73                         | 53                       |
| Tumors size ≥40 mm                                         | 171                        | 102                      |
| Increasing tumor size during follow-up                     | NA                         | 106                      |
| Serum thyroglobulin >1,000 ng/mL                           | 35                         | 12                       |
| Tumors compressing the trachea or esophagus                | 24                         | 8                        |
| Tumors extending into the mediastinum                      | 28                         | 13                       |
| Autonomously functioning thyroid tumors                    | 25                         | 1                        |
| Cosmetic problems                                          | 16                         | 8                        |
| Be-III or IV on repeat FNAC                                | 16                         | 12                       |
| Patient’s preference                                       | 56                         | 10                       |
| Surgical indication for other associated thyroid diseases   | 72                         | 4                        |
| Unclear                                                    | 39                         | 2                        |

The values (n) are numbers of patients. Some cases had more than one reason for surgery. Be, Bethesda System category; FNAC, fine needle aspiration cytology; NA, not applicable.

![Patient flowchart](image-url)
| Table 2 | Clinical characteristics of the patients at the initial cytological examination |
|---------|--------------------------------------------------------------------------------|
|         | Surgery | Observation | \( p \)-value |
| No. of patients | 541 (17.4%) | 2,561 (82.6%) |  |
| Age, yrs * | 55 (20–83) | 59 (20–92) | <0.01 |
| Gender | | |  |
| Male | 75 (13.9%) | 351 (13.7%) | 0.945 |
| Female | 466 (86.1%) | 2,210 (86.3%) |  |
| Max. tumor size, mm a | 46 (13–113) | 20 (10–90) | <0.01 |
| Serum thyroglobulin, ng/mL b | 90.6 (0.7–8,000) | 41.3 (0.5–8,000) | <0.01 |
| Kuma Hospital US classification: | | |  |
| Class 2 (benign) | 433 (80.0%) | 2,439 (95.2%) | <0.01 |
| Class 3 (borderline) | 104 (19.2%) | 120 (4.7%) |  |
| Class 4 (susp. malignant) | 4 (0.7%) | 2 (0.1%) |  |
| Results of repeated FNAC: | | |  |
| Be-II | 138 | 291 |  |
| Be-III or IV | 110 (79.7%) | 271 (93.1%) |  |
| Be-V | 26 (18.8%) | 20 (6.9%) |  |
| | 2 (1.4%) | 0 (0.0%) |  |

a Median (range). b Patients who were positive for thyroglobulin antibody were excluded. US, Ultrasonographic.

| Table 3 | Clinical characteristics of patients who underwent immediate surgery and delayed surgery |
|---------|--------------------------------------------------------------------------------|
|         | Immediate surgery | Delayed surgery | \( p \)-value |
| No. of patients | 393 (72.6%) | 148 (27.4%) |  |
| Gender | | |  |
| Male | 59 (15.0%) | 16 (10.8%) | 0.26 |
| Female | 334 (85.0%) | 132 (89.2%) |  |
| At initial cytological diagnosis: | | |  |
| Age, yrs * | 56 (22–83) | 50 (20–74) | <0.01 |
| Max. tumor size, mm a | 38 (13–113) | 35 (10–77) | 0.60 |
| Serum thyroglobulin, ng/mL b | 91.5 (0.5–8,000) | 88.3 (14–3,985) | 0.08 |
| Kuma Hospital US classification: | | |  |
| Class 2 (benign) | 320 (81.4%) | 113 (76.4%) | 0.21 |
| Class 3 (borderline) | 69 (17.6%) | 35 (23.6%) |  |
| Class 4 (susp. for malignancy) | 4 (1.0%) | 0 (0.0%) |  |
| At surgery: | | |  |
| Age, yrs * | 57 (23–83) | 56 (24–79) | 0.43 |
| Max. tumor size, mm a | 36 (10–119) | 46 (20–119) | <0.01 |
| Serum thyroglobulin, ng/mL b | 103.9 (1.2–6,921) | 127.0 (0.7–6,638) | 0.23 |
| Kuma Hospital US classification: | | |  |
| Class 2 (benign) | 320 (81.4%) | 95 (64.2%) | <0.01 |
| Class 3 (borderline) | 71 (18.1%) | 51 (34.5%) |  |
| Class 4 (susp. for malignancy) | 2 (0.5%) | 2 (1.4%) |  |
| Histopathological diagnosis: | | |  |
| Benign | 384 (97.7%) | 141 (95.3%) | 0.13 |
| PTC | 2 (0.5%) | 2 (1.4%) |  |
| FTC, minimally invasive | 6 (1.5%) | 2 (1.4%) |  |
| FTC, widely invasive | 1 (0.3%) | 3 (2.0%) |  |

a Median (range). b Patients who were positive for thyroglobulin antibody were excluded. FTC, follicular thyroid carcinoma; PTC, papillary thyroid carcinoma.
than that of the patients without surgery (46 mm vs. 20 mm, p < 0.01). The median serum thyroglobulin level at the cytological diagnosis in the patients with surgery was significantly higher than that in the patients without surgery (90.6 ng/mL vs. 41.3 ng/mL, p < 0.01), excluding the thyroglobulin antibody-positive cases. At the presentation, the nodules of all 3,102 patients were classified with US as benign, borderline, or suspicious for malignancy in 2,872 patients (92.6%), 224 patients (7.2%), and 6 patients (0.2%), respectively. The surgery group had significantly higher US classes than the observation group (p < 0.01).

Table 4 shows the clinicopathological parameters at surgery according to pathological diagnosis in the 541 patients who underwent surgery. The tumors of 525 of these patients (97.0%) were diagnosed as benign (49 follicular adenomas or 476 adenomatous nodules) and the remaining 16 (3.0%) tumors were diagnosed as malignant. The malignancies included 12 (75%) follicular thyroid carcinomas (FTCs) and 4 (25%) papillary thyroid carcinomas (PTCs). The PTCs consisted of well-differentiated PTC (n = 2) and follicular variants of PTC (n = 2). Of the clinical features, only the US classification at surgery was significantly different between the benign and malignant cases, with the latter having higher US classifications. The TNM classifications in these patients were T1 (n = 1), T2 (n = 8) or T3 (n = 7), and N0 and M0 in all. All 4 patients with widely FTC underwent completion thyroidectomy and adjuvant radioactive iodine therapy at a later date.

The relationships between the US evaluation at surgery and pathological diagnoses are summarized in Fig. 2. The tumors of these patients were evaluated by US at surgery as benign in 415 patients (76.7%), borderline in 122 patients (22.6%), and suspicious for malignancy in 4 patients (0.7%). Our analysis revealed that 50% of the tumors that were evaluated as suspicious for malignancy on US were pathologically diagnosed as malignant. In contrast, only 0.7% of the benign tumors and 9.0% of the borderline tumors on US were pathologically diagnosed as malignant. Evaluation of suspicious for malignancy by US was significantly linked to pathological diagnosis of thyroid carcinoma (p < 0.01) (Table 4, Fig. 2).

There were no significant differences between the benign and malignant cases in patient age, gender, tumor size, or serum thyroglobulin level at surgery (Table 4). In the patients with delayed surgery, there was also no significant difference in the TV-DR between the histopathologically benign and malignant cases (median 0.245/year vs. 0.220/year, p = 0.528) (Fig. 3).

Of the 2,709 patients without immediate surgery, 667 were followed for >10 years without surgical treatment. The median follow-up period of these 667 patients was

Table 4  Clinicopathological features at surgery in patients with benign or malignant pathological diagnosis

|                          | Benign   | Malignant | p-value |
|--------------------------|----------|-----------|---------|
| No. of patients          | 525 (97.0%) | 16 (3.0%) |         |
| Age, yrs a               | 57 (23–84) | 65 (35–78) | 0.06    |
| Gender                   |          |           |         |
| Male                     | 73 (14.0%) | 2 (12.5%) | 0.99    |
| Female                   | 452 (86.1%) | 14 (87.5%) |         |
| Max. tumor size, mm b    | 47 (13–113) | 38 (16–78) | 0.53    |
| Max. tumor size <40 mm   | 251 (47.8%) | 9 (56.3%) | 0.61    |
| ≥40 mm                   | 274 (52.2%) | 7 (43.8%) |         |
| Serum thyroglobulin, ng/mL b | 110.1 (0.7–6,921) | 213.5 (14–3,985) | 0.10 |
| Kuma Hospital US classification: |          |           |         |
| Class 2 (benign)         | 412 (78.5%) | 3 (18.8%) | <0.01   |
| Class 3 (borderline)     | 111 (21.1%) | 11 (68.8%) |         |
| Class 4 (sus, for malignancy) | 2 (0.4%) | 2 (12.5%) |         |

a Median (range). b Patients who were positive for thyroglobulin antibody were excluded.
148.9 months (range 120.2–168.7 months). There was no appearance of suspicious features on follow-up US in any of the patients. Repeat FNACs were performed in 25 of these patients, but there were no findings suggestive of follicular tumor/possible malignancy. The median tumor size at the latest point in these patients was significantly larger than that at the cytological diagnosis (24 mm vs. 20 mm, \( p < 0.01 \)). The TV-DRs (/year) in these patients were >0.5 (indicating moderate growth), 0.1 to 0.5 (slow growth), –0.1 to 0.1 (stable disease), and less than –0.1 (decrease in tumor size) over time in 0 patients (0%), 20 patients (3.0%), 599 patients (89.8%), and 48 patients (7.2%) (Table 5). In addition, US showed no evidence of cervical lymph node metastasis. The median serum thyroglobulin level at the latest point in the 667 patients was significantly higher than that at cytological diagnosis (41.3 ng/mL vs. 37.1 ng/mL, \( p < 0.01 \)), excluding thyroglobulin antibody-positive cases. However, the method of thyroglobulin measurement changed during the study period at our hospital.

**Discussion**

In particular, PTCs misdiagnosed as benign tumors by FNAC are indolent and very slow-growing [15]. Because of the good prognosis in such cases, it is not as imperative that the malignancies be detected as soon as possible. The above-described data confirm that an initially benign finding based on FNAC confers a negligible mortality risk during long-term follow-up despite a low but real risk of false negatives in this cytologic category. This in turn supports the idea that an initial finding of benign cytology conveys an overall excellent prognosis that can be managed with a conservative follow-up strategy. Based on all the above, we consider that (1) it is not necessary to perform surgery for all cases of benign thyroid tumor, and (2) it is important to select patients who need surgical resection. Our basic policy for thyroid tumors that are cytologically defined as benign is observation without surgery, and our criteria for surgical indications are as described above.

The issue of whether the maximum tumor size or a tumor showing enlargement should be used as a
Outcomes of benign thyroid tumors

separate criterion for recommending surgical therapy or determining the extent of thyroidectomy is controversial [2, 4, 16-34]. Several surgical series have reported higher malignancy rates in nodules >3–4 cm, but these studies suffered from both selection bias and potential sampling error [19, 33]. In a 2014 single-center study, the practice was to offer thyroidectomy to all patients with nodules ≥4 cm [34]; the study investigators reported that thyroid nodules underwent a preoperative FNAC, and of the 125 cytologically benign nodules, 10.4% were malignant on final histopathology. Thus, they suggested that patients with a >4 cm cytologically benign nodule of the thyroid might be at increased risk of malignancy. However, in the present study, the time of the patients’ initial cytological diagnoses and surgery, the benign and malignant nodules did not significantly differ in their median maximum measured diameters, and the malignancy risk was low in the cytologically benign thyroid nodules even when they were >4 cm. Larger nodules may require monitoring for growth that could result in symptoms and thus prompt surgical intervention despite benign cytology. Even if larger nodules are not malignant, it makes sense that larger nodules would cause more compressive symptoms and therefore lead to a thyroidectomy.

Both malignant and benign tumors of the thyroid gland are known to be slow-growing tumors [5, 14, 35, 36], but we observed that the malignancy risk was high in FNAC-diagnosed follicular tumors when they had grown previously [4]. Among the present study’s patients with delayed surgery, there was no significant difference in the growth that occurred in the malignant versus benign nodules over time. These findings may indicate that tumor growth or larger tumor size does not constitute a suitable criterion for surgical indication in order to prevent overlooking a malignancy. Moreover, the present study is the first to demonstrate that there is no association between malignancy and tumor growth when using the tumor volume-doubling rate measured during the follow-up of benign tumors. In terms of the reasons that tumor growth was not necessarily associated with malignancy in our cytologically benign cases, we suggest two possible explanations. 1. Tumors diagnosed as benign by cytology, even if malignant, are low grade and have a slow growth rate. 2. Even in benign tumors, as shown in Fig. 3, DR is distributed from slow reduction to slow increase, and the DR of the cancer group is buried within the width of this distribution. It is thus possible that longer-term observation may result in more malignant tumors in the enlarged group.

Among the factors potentially related to malignancy, only the ultrasonographic classification was found to exhibit a significant difference between benign and malignant tumors. Our hospital’s US classification may thus be very helpful for predicting malignant tumors in patients with a diagnosis of benign tumor on findings from FNAC. Similarly, recent investigations of repeat US-guided FNAC in nodules with initial benign cytology showed higher detection rates for missed malignancy for the nodules with a high-suspicion sonographic pattern rather than tumor size increase [37, 38].

In the present study, during the follow-up for patients with observation, the delayed surgery was prompted mainly by either nodule growth or the development of a new suspicious US feature. However, as above, there was no significant difference in the growth that occurred in malignant versus benign nodules over time in the patients with delayed surgery. The present results indicate that the use of suspicious US characteristics rather than nodule growth should be the indication for malignancy despite an initial benign cytology diagnosis. We thus consider that suspicious US findings cannot be ignored, even when the result of a biopsy is benign.

It seems that most misdiagnoses or inconclusive diagnoses are the result of examining material that is unsatisfactory for diagnosis [39]. A sample from a firm area such as a calcification, which is less likely to contain diagnostic cells, may still be difficult to classify. In the present study, PTCs were accompanied by gross calcification in the 4 patients with PTC. Cells of the thyroid tissue around the tumor were collected in these cases, and the samples contained few cells for diagnosis. There were 12 patients with FTC in this study, and unfortunately, an FNAC cannot distinguish between a benign tumor and a follicular carcinoma because the diagnosis of adenoma/carcinoma is based on histopathologic criteria (such as capsular or vascular invasion) to which cytology does not contribute.

Generally, the measurement of serum thyroglobulin levels is currently only performed postoperatively as a marker of recurrent disease or distant metastases in the follow-up of a patient with differentiated thyroid cancer. In the present study, we found no difference in the level of serum thyroglobulin at surgery between the benign and malignant tumors. We feel that a high thyroglobulin level may not be an appropriate indication for the surgical treatment of cytologically benign tumors.

In 1992, we reported that in 140 patients with untreated palpable thyroid nodules, after a mean of 15 years 13.5% of the nodules had increased in size, 41.5% had decreased, and 11.4% disappeared [40]. In the present study, 89.8% and 7.2% of the tumors showed stable disease and tumor regression, respectively, among the 667 patients who were followed without surgery for >10 years.

This study has some limitations. First, its design was retrospective, and it was thus not exempt from the risk of
selection bias inherent to retrospective studies. Second, not all of the patients could undergo surgery, and the decisions regarding whether and when surgical resection was recommended to a patient with a growing nodule were dependent on the judgement of the treating physician. The true incidence of malignancy for nodules that are cytologically benign cannot be accurately evaluated, because surgery is not performed for all patients, introducing selection bias. The follow-up examination intervals, the length of the observation periods, and the ratios of nodule growth at surgery also varied among the patients. In addition, as we could not trace patients who were transferred to other hospitals, a considerable number of patients were lost to follow-up.

In conclusion, 541 (17.4%) of 3,102 patients with a cytological diagnosis of benign tumor underwent a thyroidectomy within a maximum of 14 years from their FNAC, and 16 (3.0%) of these patients underwent surgery and were histologically diagnosed as having a malignant tumor. The tumor size, TV-DR, and serum thyroglobulin level do not appear to be independent predictors of thyroid malignancy. An ultrasonographic evaluation provides important information for therapeutic decision-making regarding surgery versus observation for cytologically benign thyroid tumors.

**Conflict of Interest**

The authors have no conflict of interest to disclose.

**Trial Registration**

No. 20200709-1 by the Institutional Review Board of Kuma Hospital

**References**

1. Kihara M, Hirokawa M, Ito Y, Kobayashi K, Miya A, et al. (2011) Final pathology findings after immediate or delayed surgery in patients with cytologically benign or follicular thyroid nodules. *World J Surg* 35: 558–562.
2. Kihara M, Ito Y, Hirokawa M, Masuoka H, Yabuta T, et al. (2011) Role of ultrasonography in patients with cytologically follicular thyroid tumor. * Auris Nasus Larynx* 38: 508–511.
3. Kihara M, Hirokawa M, Masuoka H, Yabuta T, Shindo H, et al. (2013) Evaluation of cytologically benign solitary thyroid nodules by ultrasonography: a retrospective analysis of 1,877 cases. *Auris Nasus Larynx* 40: 308–311.
4. Nakamura H, Hirokawa M, Ota H, Kihara M, Miya A, et al. (2015) Is an increase in thyroid nodule volume a risk factor for malignancy? *Thyroid* 25: 804–811.
5. Haugen BR, Alexander EK, Bible KC, Doherty GM, Mandel SJ, 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.
6. Yokozawa T, Fukata S, Kuma K, Matsuzuka F, Kobayashi A, et al. (1996) Thyroid cancer detected by ultrasound-guided fine-needle aspiration biopsy. *World J Surg* 20: 848–853.
7. Kobayashi K, Fukata S, Miyauuchi A (2005) Diagnosis of follicular carcinoma of the thyroid: role of sonography in preoperative diagnosis of follicular nodules. *J Med Ultrason 32*: 153–158.
8. Ito Y, Miyauchi A, 2007) A therapeutic strategy for incidentally detected papillary microcarcinoma of the thyroid. *Nat Clin Pract Endocrinol Metab 3*: 240–248.
9. Ito Y, Amino N, Yokozawa T, Ota H, Ohshita M, et al. (2007) Ultrasonographic evaluation of thyroid nodules in 900 patients: comparison among ultrasonographic, cytological, and histological findings. *Thyroid* 17: 1269–1276.
10. Raber W, Kaseret K, Niederle B, Vierhapper H (2000) Risk factors for malignancy of thyroid nodules initially identified as follicular neoplasia by fine-needle aspiration: results of a prospective study of one hundred twenty patients. *Thyroid* 10: 709–712.
11. Goldstein RE, Netterville JL, Burkey B, Johnson JE (2002) Implications of follicular neoplasms, atypia, and lesions suspicious for malignancy diagnosed by fine-needle aspiration of thyroid nodules. *Ann Surg* 235: 656–664.
12. Cíbas ES, Ali SZ (2017) The 2017 Bethesda system for reporting thyroid cytopathology. *Thyroid* 27: 1341–1346.
13. Lloyd RV, Osamura RY, Klöppel G, Rosai J (eds.) (2017) WHO classification of Tumours of Endocrine Organs (4th). World Health Organization, Lyon, France.
14. Miyauuchi A, Kudo T, Ito Y, Oda H, Yamamoto M, et al. (2019) Natural history of papillary thyroid microcarcinoma: kinetic analyses on tumor volume during active surveillance and before presentation. *Surgery* 165: 25–30.
15. Ito Y, Higashiyama T, Takamura Y, Miya A, Kobayashi K, et al. (2007) Long-term follow-up for patients with papillary thyroid carcinoma treated as benign nodules. *Anticancer Res* 27: 1039–1043.
16. McHenry CR, Huh ES, Machekano RN (2008) Is nodule size an independent predictor of thyroid malignancy? *Surgery* 144: 1062–1069.
17. Ylagan LR, Farkas T, Delhner LP (2004) Fine needle aspiration of the thyroid: a cytohistologic correlation and study of discrepant cases. *Thyroid* 14: 35–41.
18. Kim DW, Park AW, Lee EJ, Choo HJ, Kim SH, et al. (2009) Ultrasound-guided fine-needle aspiration biopsy of
thyroid nodules smaller than 5 mm in the maximum diameter: assessment of efficacy and pathological findings. *Korean J Radiol* 10: 435–440.

19. Pinchot SN, Al-Wagih H, Schaefer S, Sippel R, Chen H (2009) Accuracy of fine-needle aspiration biopsy for predicting neoplasm or carcinoma in thyroid nodules 4 cm or larger. *Arch Surg* 144: 649–655.

20. Chehade JM, Silverberg AB, Kim J, Case C, Mooradian AD (2001) Role of repeated fine-needle aspiration of thyroid nodules with benign cytologic features. *Endocr Pract* 7: 237–243.

21. Kawai T, Nishihara E, Kudo T, Ota H, Morita S, et al. (2012) Histopathological diagnoses of “accessory” thyroid nodules diagnosed as benign by fine-needle aspiration cytology and ultrasonography. *Thyroid* 22: 299–303.

22. Alexopoulou O, Beguin C, Buysschaert M, Squifflet JP, de Burbure C, et al. (2004) Predictive factors of thyroid carcinoma in non-toxic multinodular goitre. *Acta Clin Belg* 59: 84–89.

23. Choi KU, Kim JY, Park DY, Lee CH, Sol MY, et al. (2005) Recommendations for the management of cystic thyroid nodules. *ANZ J Surg* 75: 537–541.

24. McCoy KL, Jabbour N, Ogilvie JB, Ohori NP, Carty SE, et al. (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.e1–e3.

25. Cappelli C, Castellano M, Pirola I, Gandossi E, De Martino E, et al. (2006) Thyroid nodule shape suggests malignancy. *Eur J Endocrinol* 155: 27–31.

26. Sippel RS, Elaraj DM, Khanafshar E, Zarnegar R, Kebebew E, et al. (2008) Tumor size predicts malignant potential in Hürthle cell neoplasms of the thyroid. *World J Surg* 32: 702–707.

27. Frates MC, Benson CB, Doubilet PM, Kunreuther E, Contreras M, et al. (2006) Prevalence and distribution of carcinoma in patients with solitary and multiple thyroid nodules on sonography. *J Clin Endocrinol Metab* 91: 3411–3417.

28. Sahin M, Gursoy A, Tutuncu NB, Guvener DN (2006) Prevalence and prediction of malignancy in cytologically indeterminate thyroid nodules. *Clin Endocrinol (Oxf)* 65: 514–518.

29. Alaeddin DI, Khiyami A, McHenry CR (2005) Fine-needle aspiration biopsy specimen with a predominance of Hürthle cells: a dilemma in the management of nodular thyroid disease. *Surgery* 138: 650–657.

30. Kim DW, Lee EJ, Kim SH, Kim TH, Lee SH, et al. (2009) Ultrasound-guided fine-needle aspiration biopsy of thyroid nodules: comparison in efficacy according to nodule size. *Thyroid* 19: 27–31.

31. Raparia K, Min SK, Mody DR, Anton R, Amrikachi M (2009) Clinical outcomes for “suspicious” category in thyroid fine-needle aspiration biopsy: patient’s sex and nodule size are possible predictors of malignancy. *Arch Pathol Lab Med* 133: 787–790.

32. Besic N, Sesek M, Peric B, Zgajnar J, Hocevar M (2008) Predictive factors of carcinoma in 327 patients with follicular neoplasm of the thyroid. *Med Sci Monit* 14: CR459–467.

33. Kuru B, Gulcelik NE, Gulcelik MA, Dincer H (2010) The false-negative rate of fine-needle aspiration cytology for diagnosing thyroid carcinoma in thyroid nodules. *Langenbecks Arch Surg* 395: 127–132.

34. Wharry LI, McCoy KL, Stang MT, Armstrong MJ, LeBeau SO, et al. (2014) Thyroid nodules (≥4 cm): can ultrasound and cytology reliably exclude cancer? *World J Surg* 38: 614–621.

35. Ito Y, Miyauchi A, Kudo T, Higashiyama T, Masuoka H, et al. (2019) Kinetic analysis of growth activity in enlarging papillary thyroid microcarcinomas. *Thyroid* 29: 1765–1773.

36. Miyauchi A, Ito Y (2019) Conservative surveillance management of low-risk papillary thyroid microcarcinoma. *Endocrinol Metab Clin North Am* 48: 215–226.

37. Kwak JY, Koo H, Youk JH, Kim MJ, Moon HJ, et al. (2010) Value of US correlation of a thyroid nodule with initially benign cytologic results. *Radiology* 254: 292–300.

38. Rosario PW, Purisch S (2010) Ultrasonographic characteristics as a criterion for repeat cytology in benign thyroid nodules. *Arq Bras Endocrinol Metabol* 54: 52–55.

39. Werga P, Wallin G, Skoog L, Hamberger B (2000) Expanding role of fine-needle aspiration cytology in thyroid diagnosis and management. *World J Surg* 24: 907–912.

40. Kuma K, Matsuoka F, Kobayashi A, Hirai K, Morita S, et al. (1992) Outcome of long standing solitary thyroid nodules. *World J Surg* 16: 583–588.