Ultrasound-guided fine-needle aspiration biopsy of thyroid nodules <10 mm in the maximum diameter: does size matter?

Yi-jun Lyu1,2,*
Fang Shen3,*
Yun Yan4,5
Ming-zhu Situ4,5
Wei-zhu Wu1,2
Guo-qiang Jiang3
Ya-ya Chen4,5

1Department of Thyroid and Breast Surgery, Ningbo Medical Center Lihuili Eastern Hospital, Ningbo 315040, Zhejiang, China; 2Department of Thyroid and Breast Surgery, Taipei Medical University Ningbo Medical Center, Ningbo 315040, Zhejiang, China; 3Department of Orthopaedic Surgery’s Spine Division, The Affiliated Hospital of Medical School of Ningbo University, Ningbo 315020, Zhejiang, China; 4Department of Thyroid and Breast Surgery, Taipei Medical University Ningbo Medical Center, Ningbo 315040, Zhejiang, China; 5Department of Ultrasound, Ningbo Medical Center Lihuili Eastern Hospital, Ningbo 315040, Zhejiang, China

*These authors contributed equally to this work

Objective: Ultrasound-guided fine-needle aspiration biopsy (US-FNAB) is a safe and effective method of screening malignant thyroid nodules such as papillary thyroid carcinoma. However, not much data are available regarding the diagnostic efficacy of US-FNAB for papillary thyroid microcarcinoma (<10 mm in diameter). We aim to compare the diagnostic efficacy of US-FNAB on thyroid nodules between two groups divided by a diameter of 10 mm by correlating the cytological results of US-FNAB with the histopathologic diagnoses in selected patients.

Patients and methods: Eight hundred twenty-two thyroid nodules (Group A: diameter ≤10 mm, n=620; Group B: diameter >10 mm, n=202) from 797 patients treated between March 2014 and June 2017 were retrospectively evaluated. Only nodules with Thyroid Imaging Reporting and Data System (TIRADS) categories 4–6 were enrolled and sampled by US-FNAB, followed by surgical resection.

Results: According to The Bethesda System for Reporting Thyroid Cytopathology (TBSRTC) diagnostic categories, 94 thyroid nodules were classified as I, III and IV, and were excluded from the analysis. The resultant 728 thyroid nodules from 721 patients were analyzed. The malignant tendency (TBSRTC V and VI) rates on US-FNAB were 88.2% and 84.6% (P=0.020) in Group A and Group B, respectively, and the malignant rates were 89.5% and 86.9% (P=0.330), respectively, on histopathology. There was a high concordance between cytology and histopathology diagnoses (kappa value =0.797), and no statistical difference in terms of US-FNAB accuracy was found between the two groups (P=0.533).

Conclusion: For thyroid nodules of TIRADS category 4–6, the diagnostic efficacy of US-FNAB is similar for thyroid nodules either smaller or greater than 10 mm in their maximum diameter.

Keywords: ultrasound-guided fine-needle aspiration biopsy, thyroid nodules, microcarcinoma, histopathology, cytopathology

Introduction

Thyroid nodules are common entities that are detected in 19%–68% of the population by using high-resolution ultrasound (US).1 Nevertheless, only a relatively small percentage (~5%) of thyroid nodules are malignant, with papillary thyroid carcinoma (PTC) being the most common pathological type.2 According to the WHO classification system, papillary thyroid microcarcinomas (PTMCs) are PTCs <10 mm in diameter.3

Early cancer detection and intervention has been suggested to reduce patient’s mortality and morbidity. Given its minimal invasiveness and technical simplicity, ultrasound-guided fine-needle aspiration biopsy (US-FNAB) has been widely adopted for characterizing thyroid nodules cytopathologically.4 The cytology results will help to determine whether or not subsequent thyroidectomy is necessary. According to the
American Thyroid Association (ATA) guideline, US-FNAB is recommended for those thyroid nodules with a diameter >10 mm those with intermediate to high suspicion US pattern. On the other hand, for patients with nodules ≤10 mm of suspicious US pattern, active sonographic surveillance is recommended instead. Despite the fact that the prognostic advantage of PTMC has been an issue of debate in recent studies, early diagnosis and treatment of PTMC might be beneficiary to patient prognosis as small size alone does not guarantee low risk in incidentally found thyroid cancers.

There are several studies that have evaluated the efficacy of US-FNAB for small thyroid nodules. We focused here on the diagnostic efficacy of US-FNAB for ultrasonographically “suspicious”, that is, Thyroid Imaging Reporting and Data System (TIRADS) categories 4–6, and small, that is, measuring ≤10 mm as the maximum diameter on US, thyroid nodules. To the best of our knowledge, our research is the first study of this kind with the largest sample size ever reported. This knowledge is important to determine whether US-FNAB should be routinely performed in this subgroup of patients in future clinical practices.

**Patients and methods**

**Patient selection**

We retrospectively reviewed 721 patients who had received thyroid US, thyroid nodule US-FNAB and subsequent thyroid surgery at Ningbo Medical Center Lihuili Eastern Hospital and Taipei Medical University Ningbo Medical Center (a tertiary medical center) between March 2014 and June 2017. Thyroid nodules were given cytopathological and histopathologic diagnoses with the criteria listed in the following sections and were grouped according to their maximum diameter as measured on US: Group A (≤10 mm) and Group B (>10 mm). Written informed consents were obtained from all patients before performing US-FNABs and surgeries. Written consents were also obtained from the patients regarding the report of their medical data. Ethical approval for conducting this retrospective study was obtained from the Ethics Committee of Ningbo Medical Center Lihuili Eastern Hospital and Taipei Medical University Ningbo Medical Center (ethical approval no. DYLL2016001) in compliance with the Declaration of Helsinki.

**Thyroid US evaluation**

US examinations were performed at the Department of Ultrasound with Philips Q5 US equipment and a 5–12 MHz linear probe. The reasons for thyroid US scan are as follows: 1) palpable cervical mass and/or enlarged lymph nodes; 2) patients with a known history of thyroid nodes and referred from other hospitals and 3) on routine health checks that included thyroid US. US patterns of the thyroid lobes and nodules (eg, calcification, echogenicity, volume, shape, dimensions, long axis/short axis ratio, vascularity) were recorded and all nodules were classified according to TIRADS by using five sonographic parameters (ie, composition, echogenicity, shape, margin and echogenic foci). Patients having TIRADS 4–6 thyroid nodules were included for US-FNAB, while the exclusion criteria were: 1) patients who rejected US-FNAB or were not cooperative for the procedure; 2) accompanied with severe cardiovascular or pulmonary conditions and 3) patients with bleeding history of unknown causes or coagulation disorders (prothrombin time >18 seconds, platelet count <50×10^9/L, prothrombin activity <40%).

**US-FNAB procedure for thyroid nodules**

US-FNAB was performed by the same pair of experienced diagnostic sonographers who were licensed to perform the procedure, in order to ensure similarity in techniques employed. Specifically, one was responsible for US guidance and the other for fine-needle aspiration (FNA). Under US guidance, the target nodule was punctured with a 24-G needle connected to a 5 mL syringe without local anesthesia. After confirmation of reaching the target nodule by the needle tip on US, the needle was moved forward and backward ten times under negative pressure to aspirate the sample. Then the needle was withdrawn, negative pressure was released and specimen within the needle was immediately transferred to a liquid-based cytology medium. Generally, only one puncture was needed to sample each individual nodule and a second attempt could be delivered in case of unsatisfactory gross tissue yield by the first attempt as judged by the sonographer performing aspiration as a precaution.

**Pathological review of thyroid nodules**

Cytological and histopathologic diagnoses were performed at the Ningbo Diagnostic Pathology Center by pathologists specialized in thyroid cancer. The cytological diagnoses were made by two independent specialized cytologists (Dr Jue Zhou and Dr Xian-fa Xu; kappa value =0.83), and release of the final report was subjected to the approval of a third senior cytologist (Dr Deng Pan), who was also in charge of making the final decision when there was discrepancy between the two. Cytological results were reported according to The Bethesda System for Reporting Thyroid Cytopathology (TBSRTC) as follows: 1) nondiagnostic or unsatisfac-
tory; II, benign; III, atypia of undetermined significance or follicular lesion of undetermined significance; IV, follicular neoplasm or suspicious for a follicular neoplasm; V, suspicious for malignancy and VI, malignant. Surgery was carried out in patients with US TIRADS categories 4–6 at their own choice with signed consents. The histological diagnoses were reported according to the WHO histological classification of thyroid tumors.17

Statistical analyses
Data were analyzed using the SPSS 22.0 software package (IBM Corporation, Armonk, NY, USA). Continuous variables were presented as mean ± SD. Categorical variables were presented as percentages. The sensitivity, specificity, positive predictive value (PPV) and negative predictive value (NPV) were given as percentages. The diagnostic performance of US-FNAB was defined by the consistency between cytological and histopathologic results: US-FNAB was considered as accurate when the two results matched and vice versa. Data comparisons between groups were carried out with the Mann–Whitney U test. Chi-squared test or Fisher’s exact test was used to compare categorical variables. The kappa statistic was used to measure the agreement between the US-FNAB and histopathology diagnoses. The kappa can range from -1 to +1: values ≤ 0 indicate no agreement and values 0.01–0.20 mean slight agreement, 0.21–0.40 mean moderate agreement, 0.61–0.80 mean substantial agreement, 0.41–0.60 mean fair agreement, 0.21–0.40 indicate no agreement and values 0.01–0.20 mean none to slight agreement, 0.21–0.40 mean fair agreement, 0.41–0.60 mean moderate agreement, 0.61–0.80 mean substantial agreement and 0.81–1.00 mean almost perfect agreement. A P-value <0.05 was considered statistically significant.

Ethics approval and consent to participate
Ethical approval was obtained from the Ethics Committee of Ningbo Medical Center Lihuili Eastern Hospital and Taipei Medical University Ningbo Medical Center (ethical approval no. DYLL2016001).

Results
Demographics, US and US-FNAB results
A total of 822 thyroid nodules of TIRADS categories 4–6 from 797 patients (772 had 1 nodule and 25 had 2) were initially reviewed during the study period. There were 620 nodules in Group A and 202 nodules in Group B. Cytologically, 94 thyroid nodules were classified as TBSRTC I, III and IV, and were excluded from subsequent statistical analysis. The resultant 728 thyroid nodules from 721 patients (714 had 1 nodule and 7 had 2) were classified as TBSRTC II (n=92, 12.6%), V (n=168, 23.1%) and VI (n=468, 64.3%). Among those 721 patients, there were 557 (77.3%) females and 164 (22.7%) males. There were 553 nodules in Group A and 175 nodules in Group B. The median nodule size was 6 mm (range: 2–10 mm) for Group A and 15 mm (range: 11–52 mm) for Group B. There was no significant difference in the mean age between the two groups (Group A: 42.3±15.4 vs Group B: 48.9±17.2, P=0.018).

Diagnostic performance of US-FNAB
According to US-FNAB cytology, the percentages of nodes that fell in the diagnostic categories of V (suspicious for malignancy) and VI (malignant), which indicate malignant tendency, were 88.2% in Group A and 84.6% in Group B (P=0.202), whereas the histological malignant rates were 89.5% in Group A and 86.9% in Group B (P=0.330). There was a high agreement between the US-FNAB cytology and final histopathology (kappa value =0.797). There was no significant difference in sensitivity, specificity, accuracy, PPV and NPV of US-FNAB between the two groups (Table 1). The diagnostic performance of US-FNAB did not exhibit significant difference between the two groups either (P=0.533).

Discussion
Given its minimal invasiveness, technical simplicity and high concordance with histology results, US-FNAB has become a routine method for biopsy of many benign and malignant pathologies, including thyroid nodules.18 The current recommendations on the size selection criteria for US-FNAB of thyroid nodules have been set at a cutoff value of 10 mm with suspicious sonographic features.19,20 ATA guidelines recommend that nodules ≥10 mm with high to intermediate suspicious US pattern (or ≥10 mm with very low suspicious pattern or ≥20 mm with very low suspicious pattern) be evaluated by US-FNAB.3 Similarly, the Society of Radiologists in Ultrasound recommend performing US-FNAB on thyroid nodules that are >10 mm and when there

| Table 1 | Sensitivity, specificity, accuracy, PPV and NPV of US-FNAB for thyroid nodules of TIRADS categories 4–6 |
|---|---|---|---|---|
| Diagnostic parameters | Group A (≤10 mm), % | Group B (>10 mm), % | P-value |
| Sensitivity | 96.8 | 96.7 | 1.000 |
| Specificity | 84.5 | 95.7 | 0.268 |
| Accuracy | 95.5 | 96.6 | 0.533 |
| PPV | 98.2 | 99.3 | 0.467 |
| NPV | 75.4 | 81.5 | 0.596 |

Abbreviations: NPV, negative predictive value; PPV, positive predictive value; TIRADS, Thyroid Imaging Reporting and Data System; US-FNAB, ultrasound-guided fine-needle aspiration biopsy.
are suspicious US features suggesting cancer.\(^\text{21}\) On the other hand, for nodules <10 mm, ATA recommends further evaluation only for patients with clinical symptoms or associated lymphadenopathy and no routine sonographic follow-up for those with very low suspicious US pattern (weak recommendation, low-quality evidence).\(^\text{22}\) Therefore, for nodules \(\leq 10\) mm that exhibit intermediate to highly suspicious US patterns, routine US follow-ups are justifiable. However, nodule size alone is not predictive of malignancy in patients with Bethesda category III, IV and V thyroid nodules, and due to the existence of thyroid microcarcinomas such as PTMC, early identification by US-FNAB and subsequent surgical intervention might provide clinical benefits for those selected patients.\(^\text{23,24}\) Certain US patterns have been associated with potential malignancy and thus indicate the necessity of US-FNAB. For example, a recent meta-analysis suggested that US features such as microcalcification, a taller than wide shape, irregular margins and absence of elasticity are associated with higher risk of malignancy and have the most satisfactory diagnostic performances.\(^\text{25}\) Similarly, the Revised Korean Society of Thyroid Radiology Consensus Statement and Recommendations suggest that for nodules >5 mm, FNAB should be performed if the target nodule is solid and hypoechoic, together with any of the following suspicious US features: microcalcification, nonparallel orientation (taller than wide), spiculated or microlobulated margin.\(^\text{26}\)

According to TBSRTC, the malignancy risk is 60%–75% for of category V (suspicious for malignancy) and 97%–99% for of category VI (malignant).\(^\text{16}\) On the other hand, there are only a limited number of reports that have evaluated the diagnostic efficacy of US-FNAB for thyroid nodules <10 mm.\(^\text{13,27,28}\) Unal and Sezer found a concordance rate of 52.4% between US-FNAB cytology and histopathology in 21 nodules <10 mm in diameter. Also, among those inconsistent cases (false negative: cytological category of benign, but histologically malignant), PTMC accounted for the majority of cases.\(^\text{14}\) Rosario et al recently specifically addressed the FNA results for a subgroup of “highly suspicious” thyroid nodules (\(\leq 10\) mm, highly suspicious US features, restricted to the thyroid) in patients who are candidates for active surveillance according to ATA.\(^\text{29}\) They reviewed a total of 198 nodules and found a very high rate of malignancy on histology for nodules with suspicious/malignant (100%) and indeterminate (81.4%) cytology results. Zhong et al compared the efficacy of US-FNAB for 344 thyroid nodules with different sizes (\(\leq 5.0, 5.1–10.0\) and >10.0 mm) and found similar diagnostic efficacy regardless of size.\(^\text{30}\)

The main novelty of the current study is the focus on TIRADS 4–6 thyroid nodules and enrolment of the largest sample size to date for comparing the diagnostic efficacy of US-FNAB for nodules that were either \(\leq 10\) or >10 mm. Besides, unlike previously studies where the benign nature of the nodule was verified by repetition of cytology or simply by clinical follow-up, all nodes in our study were surgically removed with definite histopathologic diagnoses. Our results showed that the inconsistency rate between US-FNAB and histology was much lower than what has been reported previously: only 21 cytologically “benign” nodules identified by the US-FNAB were later confirmed as PTCs after surgery (false negative) and 10 cytologically “malignant” nodules were later verified as benign nature (false positive). Among the 21 nodes with false-negative US-FNAB results, 5 were >10 mm and 16 were \(\leq 10\) mm. This might be caused by sampling of normal thyroid tissue during US-FNAB procedure. Besides, there were false-negative US-FNAB cases in both groups, implicating the need of follow-up of these patients with cytologically “benign” results.

Given the slow-growing nature of PTC, there is still controversy over the clinical value of its early diagnosis by US-FNAB. It was reported that there was a high frequency of occult papillary microcarcinomas in an autopsy study (35.6%).\(^\text{31}\) In our study, the number of nodules \(\leq 10\) mm was three times higher than those >10 mm, suggesting the importance of appropriate clinical management of these subcentimeter nodules. The 2015 ATA guidelines, however, do not recommend US-FNAB for subcentimeter nodules unless there is extrathyroidal extension or suspicious lymphadenopathy. Besides, ATA also favors active surveillance in selected low-risk PTMC, instead of immediate surgery. These recommendations are largely based on observational studies done by Japanese researchers that suggested the relatively indolent nature of certain subtypes of PTMC.\(^\text{32–34}\) Nevertheless, perithyroidal lymph node metastasis is a feature seen in certain subtypes of PTMCs which are shown to have poor differentiation.\(^\text{12,35,36}\) However, without US-FNAB, there is no better way to stratify patients with subcentimeter thyroid nodules of intermediate to highly suspicious features into either active surveillance or surgery. Neither there is any reliable method to predict which subsets of PTMC will be more aggressive and thus need to have more timely surgery. A recent study by Gweon et al showed that, among patients with subcentimeter thyroid nodules of highly suspicious US features, those of younger age (<45 years), male gender and with certain US features (microcalcification and a taller-
than-wide shape) might not be good candidates for active surveillance as they are associated with malignancy rate and aggressive biological behavior.\textsuperscript{37} Besides, prolonged active surveillance itself is not without associated costs in terms of clinical, psychological and economic burdens on the health care system and patients.\textsuperscript{38} Therefore, ongoing research is needed to provide evidence for the optimal management of patients with subcentimeter pathologies. However, our study also has several limitations. First, similar to previous work,\textsuperscript{37} patients with nodules of TIRADS categories 4–6 were offered with US-FNAB and subsequent surgery. Although this provides valuable information on the concordance between cytology and histology results, some physicians would doubt the necessity of surgery for patients with benign US-FNAB cytological findings and would suggest follow-up instead.\textsuperscript{32} This is at least partially because of the associated surgical complications.\textsuperscript{43,44} However, in our practice, we did provide explanations to every patient regarding current ATA guidelines and related researches by Ito et al\textsuperscript{33} and offered conservative options including active surveillance.\textsuperscript{34} After thorough considerations, some patients opted to have US follow-ups, while others still had severe psychological stress and proceeded with surgeries despite our advice. Second, thyroid nodules with TBSRTC I, III and IV cytology (11\% of the total cases) were excluded from the analysis because, according to TBSRTC guidelines, a repeated FNA is needed for category I and III nodules, while category IV (folicular or suspicious) is not the focus of the current study. Therefore, our study only focused on thyroid nodules with definite cytology results (either benign or malignant, but excluding follicular neoplasms) yielded by single biopsy attempt. However, this might also lead to selection bias to some extent. Third, we did not exclude patients with coexistence of diffuse thyroid disease, such as thyroiditis or diffuse goiter, and these thyroid conditions might influence the performance of the US-FNABs, especially for small nodules.\textsuperscript{43} Last but not least, our study is retrospective and future prospective study will yield clinical evidence at a higher level to answer the controversy over the benefit of surgery on the prognostic advantage of PTMC.

To conclude, US-FNAB is an effective diagnostic method for thyroid nodules with TIRADS categories >4 regardless of their size. Prospective studies are needed to determine the natural history of thyroid malignancy <10 mm and identify those at particular risk of exhibiting aggressive behavior, in order to further support the value of its early diagnosis and surgical intervention as well as for providing more optimal and personalized care.

**Consent for publication**

This manuscript does not contain any individual person’s data in any form.

**Data sharing statement**

The data sets used and/or analyzed during the current study are available from the corresponding author on reasonable request.

**Acknowledgments**

We would like to thank Dr Jue Zhou, Dr Xian-fa Xu and Dr Deng Pan from Ningbo Diagnostic Pathology Center for their help in reviewing pathological diagnosis of the cases.

**Disclosure**

The authors report no conflicts of interest in this work.

**References**
