



Core-Needle Biopsy Does Not Show Superior Diagnostic Performance to Fine-Needle Aspiration for Diagnosing Thyroid Nodules

Ilah Shin¹, Eun-Kyung Kim¹, Hee Jung Moon¹, Jung Hyun Yoon¹,
Vivian Youngjean Park¹, Si Eun Lee¹, Hye Sun Lee², and Jin Young Kwak¹

¹Department of Radiology, Severance Hospital, Research Institute of Radiological Science, Yonsei University College of Medicine, Seoul;

²BioStatistics Collaboration Unit, Yonsei University College of Medicine, Seoul, Korea.

Purpose: To compare the diagnostic performances of fine-needle aspiration (FNA) and core-needle biopsy (CNB) for thyroid nodules according to nodule size.

Materials and Methods: This retrospective study included 320 thyroid nodules from 320 patients who underwent both FNA and CNB at outside clinics and proceeded with surgery in our institution between July 2012 and May 2019. According to nodule size, the diagnostic performances of FNA and CNB were calculated using various combinations of test-negatives and test-positives defined by the Bethesda categories and were compared using the generalized estimated equation and the Delong method.

Results: There were 279 malignant nodules in 279 patients and 41 benign nodules in 41 patients. The diagnostic performance of FNA was mostly not different from CNB regardless of nodule size, except for negative predictive value, which was better for FNA than CNB when applying Criteria 1 and 2. When applying Criteria 3, the specificity and positive predictive value of FNA were superior to CNB regardless of size. When applying Criteria 4, diagnostic performance did not differ between FNA and CNB regardless of size. After applying Criteria 5, diagnostic performance did not differ between FNA and CNB in nodules ≥ 2 cm. However, in nodules ≥ 1 cm and all nodules, the sensitivity, accuracy, and negative predictive value of CNB were better than those of FNA.

Conclusion: CNB did not show superior diagnostic performance to FNA for diagnosing thyroid nodules.

Key Words: Ultrasonography; biopsy, fine-needle; thyroid neoplasms; biopsy, large-core needle

INTRODUCTION

Ultrasonography (US) guided-fine-needle aspiration (FNA) has been used as the primary diagnostic tool for thyroid nodules, owing to its safety, simplicity, cost-effectiveness, and reliability.¹ Since early 2010, several articles have demonstrated

the safety and tolerability of core-needle biopsy (CNB) for the evaluation of thyroid nodules, as well as its diagnostic efficacy.²⁻⁸ Unlike FNA, which has been validated as a standard diagnostic method for over half a century, the value of CNB has not yet been fully evaluated for diagnosing thyroid nodules.⁹⁻¹¹

The main limitation of FNA is nondiagnostic and indeterminate results, corresponding to categories I, III, and IV of the Bethesda System for Reporting Thyroid Cytopathology (BSRTC).¹² To overcome this weakness, some investigators have suggested CNB as alternative or complementary diagnostic tool for thyroid nodules,^{5,8,13-16} even though CNB does not show superior diagnostic performance to FNA for diagnosing thyroid cancer and poses a higher rate of post-biopsy hematoma than FNA.^{1,17} Many favorable results for CNB have been found in studies conducted by experienced radiologists and CNB has even been used to diagnose small nodules,^{5,8,15,16,18} even micronodules not indicated for FNA in prestigious guide-

Received: November 22, 2019 **Revised:** December 24, 2019

Accepted: December 24, 2019

Corresponding author: Jin Young Kwak, MD, PhD, Department of Radiology, Severance Hospital, Yonsei University College of Medicine, 50-1 Yonsei-ro, Seodaemun-gu, Seoul 03722, Korea.

Tel: 82-2-2228-7413, Fax: 82-2-2227-8337, E-mail: docjin@yuhs.ac

•The authors have no potential conflicts of interest to disclose.

© Copyright: Yonsei University College of Medicine 2020

This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<https://creativecommons.org/licenses/by-nc/4.0>) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

lines.^{19,20} Considering the thickness of the core needle, the potential risks of CNB might be greater than the expected benefits in patients with thyroid nodules, especially in those with small nodules, even though CNB shows better diagnostic accuracy than FNA.

The purpose of our study was to compare the diagnostic performances of FNA and CNB performed simultaneously at outside clinics for thyroid nodules according to nodule size.

MATERIALS AND METHODS

Our Yonsei University Institutional Review Board approved this retrospective study (IRB No. 4-2019-0745), and the requirement for informed consent was waived for review of images and medical records.

Study population

Between July 2012 and May 2019, patients who underwent both FNA and CNB for a thyroid nodule from one of 71 outside clinics were included. Among these patients, 320 patients who underwent surgery in our institution were finally included in our study. The median age of patients was 42 years (range, 19–74 years), and the median size of nodules was 10 mm (range, 3–62 mm). There were 250 female patients and 70 male patients (Table 1).

Staging US and cytopathology reports

Preoperative staging US was performed by one of 33 radiologists with 1–22 years of experience in thyroid imaging. The ra-

diologists prospectively recorded the size and US features of thyroid nodules that had undergone FNA or CNB at outside clinics into our institutional database. The recorded US features were composition, echogenicity, margin, calcifications, and shape.²¹

In our institution, all cytopathologic slides from outside clinics were reviewed by one of the 15 cytopathologists in our institution according to the review schedule before surgery. Cytology reports were based on the six categories of the BSRTC.¹² The CNB pathologic reports were categorized into the same six categories of the Bethesda System according to a recently proposed guideline by a radiologist (K.J.Y.), because the diagnostic categories of CNB have yet to be standardized for thyroid nodules (Fig. 1).²²

Data and statistical analysis

Surgical pathologic diagnosis was defined as the standard reference. One nodule was interpreted as “malignancy” on FNA and “suspicious for malignancy (SUSP)” on CNB, although no residual cancer was noted in the surgical specimen. In this study, the nodule was considered as a cancer because central lymph node metastases were seen on the pathologic samples. We compared demographics between benign and malignant nodules, using the chi-square or Fisher’s exact test for categorical variables and the Mann-Whitney U test for continuous variables.

To calculate the diagnostic performances of both FNA and CNB, we defined various combinations of test-negatives and test-positives according to the Bethesda categories to assess malignancy and designated these combinations as Criteria 1

Table 1. Demographic Data of 320 Nodules in 320 Patients

| | Total (n=320) | Benign (n=41) | Malignant (n=279) | p value |
|-------------------------|---------------|---------------|-------------------|---------|
| Median age (yr) | 42 (19–74) | 39 (22–68) | 42 (19–74) | 0.783 |
| Median nodule size (mm) | 10 (3–62) | 23 (5–62) | 10 (3–47) | <0.001 |
| Male sex | 70 (21.9) | 11 (26.8) | 59 (21.2) | 0.411 |
| FNA results | | | | <0.001 |
| Nondiagnostic | 23 (7.2) | 8 (19.5) | 15 (5.4) | |
| Benign | 18 (5.6) | 12 (29.3) | 6 (2.2) | |
| AUS/FLUS | 75 (23.4) | 18 (43.9) | 57 (20.4) | |
| FN/SFN | 2 (0.6) | 2 (4.9) | 0 (0) | |
| SUSP | 88 (27.5) | 0 (0) | 88 (31.5) | |
| Malignancy | 114 (35.6) | 1 (2.4) | 113 (40.5) | |
| CNB results | | | | <0.001 |
| Nondiagnostic | 18 (5.6) | 2 (4.9) | 16 (5.7) | |
| Benign | 8 (2.5) | 1 (2.4) | 7 (2.5) | |
| AUS/FLUS | 59 (18.4) | 23 (56.1) | 36 (12.9) | |
| FN/SFN | 18 (5.6) | 15 (36.6) | 3 (1.1) | |
| SUSP | 32 (10.0) | 0 (0) | 32 (11.5) | |
| Malignancy | 185 (57.8) | 0 (0) | 185 (66.3) | |

FNA, fine-needle aspiration; CNB, core-needle biopsy; AUS/FLUS, atypia of undetermined significance or follicular lesion of undetermined significance; FN/SFN, follicular neoplasm or suspicious for follicular neoplasm; SUSP, suspicious for malignancy. Values are number (%) or number (range).

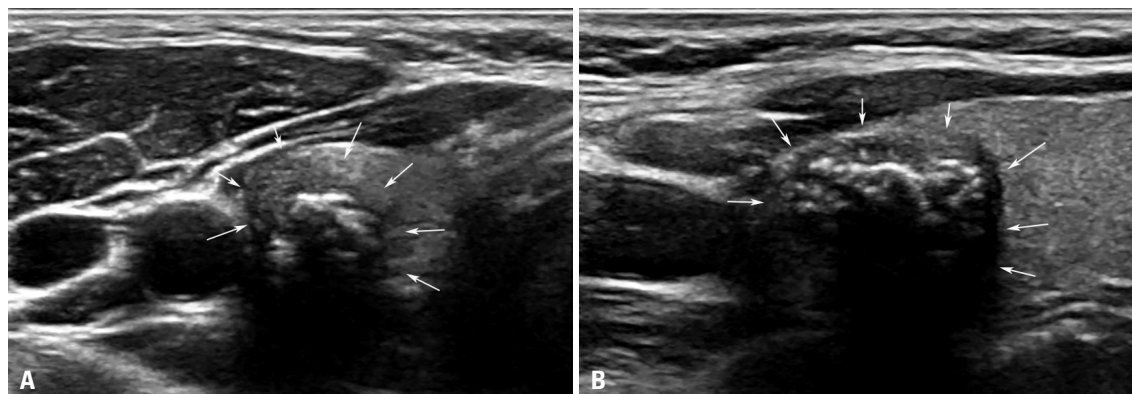


Fig. 1. US images of a 35-year-old female with a papillary thyroid carcinoma, conventional type, which was diagnosed as “malignancy” on fine-needle aspiration and “nondiagnostic” on core-needle biopsy detected at an outside clinic. A 1.3-cm solid hypoechoic nodule (arrows) with microlobulated margins and mixed calcifications was seen on the axial (A) and longitudinal (B) scans of US. US, ultrasonography.

Table 2. Diagnostic Criteria

| Criteria | Test-negative | Test-positive |
|-------------------------------|---|--------------------------|
| Criteria 1 ^{1,23,24} | Benign | SUSP, malignancy |
| Criteria 2 ^{1,24} | Benign | FN/SFN, SUSP, malignancy |
| Criteria 3 ^{1,8,24} | Nondiagnostic, benign, AUS/FLUS | FN/SFN, SUSP, malignancy |
| Criteria 4 ^{1,4,14} | Nondiagnostic, benign, AUS/FLUS, FN/SFN | SUSP, malignancy |
| Criteria 5 ^{8,16} | Nondiagnostic, benign, AUS/FLUS, FN/SFN, SUSP | Malignancy |

AUS/FLUS, atypia of undetermined significance or follicular lesion of undetermined significance; FN/SFN, follicular neoplasm or suspicious for follicular neoplasm; SUSP, suspicious for malignancy.

to 5 (Table 2).^{4,8,14,16,23,24} In Criteria 1, “benign” was considered as test-negative, and “SUSP” and “malignancy” were considered as test-positive. In Criteria 2, “benign” was considered as test-negative, and “follicular neoplasm or suspicious for follicular neoplasm (FN/SFN),” “SUSP,” and “malignancy” were considered as test-positive. In Criteria 3, “nondiagnostic,” “benign,” and “atypia of undetermined significance or follicular lesion of undetermined significance (AUS/FLUS)” were considered as test-negative, and the remaining categories were considered as test-positive. In Criteria 4, “SUSP” and “malignancy” were considered as test-positive, and the remaining categories were considered as test-negative. In Criteria 5, “malignancy” was considered as test-positive, and the remaining categories were considered as test-negative. We calculated the sensitivities, specificities, accuracies, positive predictive values (PPVs), negative predictive values (NPVs), and areas under the receiver operating characteristic curve (AUCs) along with 95% confidence intervals (CIs) of FNA and CNB to diagnose thyroid malignancy. These variables were compared using the generalized estimated equation and the AUCs using the De-long method. All diagnostic performances of FNA and CNB were calculated and compared separately for all nodules, nod-

ules ≥ 10 mm, and nodules ≥ 20 mm, respectively. All analyses were performed with SPSS version 25 (IBM Corp., Armonk, NY, USA) and SAS (version 9.4, SAS Institute Inc., Cary, NC, USA). A *p* value of less than 0.05 was considered to indicate a significant difference.

RESULTS

Baseline clinicopathological characteristics

There were 279 malignant nodules in 279 patients and 41 benign nodules in 41 patients (Tables 1 and 3). The median age (42 years, range; 19–74 years) of the patients with malignant nodules was not significantly different from that (39 years, range; 22–68 years) of patients with benign nodules. The median size of benign nodules was 23 mm, which was significantly larger than that of malignant nodules (10 mm) ($p < 0.001$). There was no statistical association between malignancy risk and gender. Significant differences were evident between benign and malignant nodules for the FNA and CNB results ($p < 0.001$).

The proportions of the “benign” and “SUSP” categories were higher in FNA than in CNB ($p = 0.049$ and $p < 0.001$, respectively). In contrast, the proportions of the “FN/SFN” and “malignancy” categories were higher in CNB than in FNA ($p < 0.001$).

Diagnostic performances of FNA and CNB

Table 4 demonstrates the diagnostic performances of FNA and CNB according to nodule size using Criteria 1–5.

When applying Criteria 1 and 2, diagnostic performance mostly did not differ between FNA and CNB in all nodules, except for NPV, which was better for FNA FNA (66.7%; 95% CI: 44.9–88.4 for Criteria 1 and 48.8%; 95% CI: 33.5–64.1 for Criteria 2) than for CNB (12.5%; 95% CI: 0–35.4 for Criteria 1 and 11.5%; 95% CI: 0–23.8 for Criteria 2) ($p < 0.001$). When applying Criteria 3, specificity and PPV were higher for FNA (92.7%; 95% CI: 84.7–100 for specificity and 98.5%; 95% CI: 96.9–100 for

Table 3. Final Surgical Diagnosis According to FNA and CNB Results

| FNA | CNB | Final diagnosis | |
|--------------------|-------------------|--|------------------------------------|
| | | Malignant (n=279) | Benign (n=41) |
| Nondiagnostic (23) | Nondiagnostic (1) | | AH (1) |
| | Benign (1) | PTC FV (1) | |
| | AUS/FLUS (9) | PTC conv (3), FC MI (1) | FA (2), AH (3) |
| | FN/SFN (2) | | AH (1), HCA (1) |
| | SUSP (1) | PTC FV (1) | |
| Benign (18) | Malignancy (9) | PTC conv (7), PTC FV (1), anaplastic carcinoma (1) | |
| | AUS/FLUS (6) | | AH (3), FA (1), HCA (1), NIFTP (1) |
| | FN/SFN (7) | FC MI (1) | AH (1), FA (5) |
| | SUSP (1) | PTC conv (1) | |
| AUS/FLUS (75) | Malignancy (4) | PTC conv (3), PTC FV (1) | |
| | Nondiagnostic (8) | PTC conv (4), PTC FV (3) | FA (1) |
| | Benign (1) | | AH (1) |
| | AUS/FLUS (25) | PTC conv (5), PTC FV (5), FC MI (5) | AH (5), FA (4), SAT (1) |
| | FN/SFN (8) | FC MI (1), PTC FV (1) | AH (1), FA (4), HCA (1) |
| | SUSP (8) | PTC conv (8) | |
| FN/SFN (2) | Malignancy (25) | PTC conv (22), PTC FV (3) | |
| | AUS/FLUS (1) | | FA (1) |
| SUSP (88) | FN/SFN (1) | | AH (1) |
| | Nondiagnostic (7) | PTC conv (7) | |
| | Benign (3) | PTC conv (3) | |
| | AUS/FLUS (8) | PTC conv (6), PTC FV (2) | |
| | SUSP (11) | PTC conv (9), PTC FV (2) | |
| Malignancy (114) | Malignancy (59) | PTC conv (56), PTC FV (2), MTC (1) | |
| | Nondiagnostic (2) | PTC conv (2) | |
| | Benign (3) | PTC conv (3) | |
| | AUS/FLUS (10) | PTC conv (9) | AH (1) |
| | SUSP (11) | PTC conv (9), PTC FV (1), no residual carcinoma (1)* | |
| | Malignancy (88) | PTC conv (86), PTC FV (1), PTC solid variant (1) | |

FNA, fine-needle aspiration; CNB, core-needle biopsy; AUS/FLUS, atypia of undetermined significance or follicular lesion of undetermined significance, FN/SFN, follicular neoplasm or suspicious for follicular neoplasm; AH, adenomatous hyperplasia; FA, follicular adenoma; HCA, Hürthle cell adenoma; MTC, medullary thyroid carcinoma; NIFTP, noninvasive follicular thyroid neoplasm with papillary-like nuclear features; PTC FV, papillary thyroid carcinoma, follicular variant; PTC conv, papillary thyroid carcinoma, conventional; SAT, subacute granulomatous thyroiditis; SUSP, suspicious for malignancy; FC MI, follicular carcinoma, minimally invasive.

Data in parentheses are number of nodules.

*One nodule was "malignancy" on FNA and "SUSP" on CNB. On the pathologic sample, there was no residual cancer. However, central lymph node metastases were seen on the pathologic samples. Therefore, the nodule was considered as a cancer in this study.

PPV) than for CNB (63.4%; 95% CI: 48.7–78.2 for specificity and 93.6%; 95% CI: 90.5–96.7 for PPV) ($p < 0.001$ and $p = 0.003$, respectively). Accuracy and NPV were not statistically different between FNA and CNB. The sensitivity (78.9%; 95% CI: 74.1–83.6) of CNB was higher than that (72%; 95% CI: 66.8–77.3) of FNA ($p = 0.036$). When applying Criteria 4, the diagnostic performance of FNA was similar to that of CNB. When applying Criteria 5, sensitivity, accuracy, and NPV were significantly higher for CNB than FNA.

In nodules ≥ 1 cm, FNA showed superior diagnostic performance over CNB in terms of NPV with Criteria 1 and 2 and specificity and PPV with Criteria 3. The other measures of diagnostic performance were not different between FNA and CNB when Criteria 1–3 were used. When Criteria 4 was ap-

plied, none of the diagnostic performance measures of FNA significantly differed from CNB. When applying Criteria 5, sensitivity, accuracy, and NPV were significantly higher for CNB than for FNA.

In nodules ≥ 2 cm, FNA showed superior diagnostic performance over CNB in terms of NPV with Criteria 1 and 2 and specificity and PPV with Criteria 3. The other measures of diagnostic performance were not different between FNA and CNB when Criteria 1–3 were used. When Criteria 4 and 5 were applied, the diagnostic performance of FNA did not statistically differ from that of CNB.

Table 4. Diagnostic Performance of FNA versus CNB According to Nodule Size

| | Criteria 1 | | | Criteria 2 | | | Criteria 3 | | | Criteria 4 | | | Criteria 5 | | |
|-----------------------|---------------------|---------------------|---------|---------------------|---------------------|---------|---------------------|---------------------|---------|---------------------|---------------------|---------|---------------------|---------------------|---------|
| | FNA | CNB | p value | FNA | CNB | p value | FNA | CNB | p value | FNA | CNB | p value | FNA | CNB | p value |
| All data (n=320) | | | | | | | | | | | | | | | |
| TP | 201 | 217 | | 201 | 217 | | 201 | 220 | | 201 | 217 | | 201 | 113 | 185 |
| TN | 12 | 1 | | 20 | 3 | | 38 | 26 | | 40 | 41 | | 40 | 41 | |
| FP | 1 | 0 | | 1 | 0 | | 3 | 15 | | 1 | 0 | | 1 | 0 | |
| FN | 6 | 7 | | 21 | 23 | | 78 | 59 | | 78 | 62 | | 166 | 94 | |
| Sensitivity | 97.1 (94.8–99.4) | 96.9 (94.6–99.2) | 0.892 | 90.5 (86.7–94.4) | 90.4 (86.7–94.1) | 0.964 | 72.0 (66.8–77.3) | 78.9 (74.1–83.6) | 0.036 | 72.0 (66.8–77.3) | 77.8 (72.9–82.7) | 0.072 | 40.5 (34.7–46.3) | 66.3 (60.8–71.9) | <0.001 |
| Specificity | 92.3 (77.8–100) | 100 (100–100) | 0.298 | 95.2 (86.1–100) | 100 (100–100) | 0.306 | 92.7 (84.7–100) | 63.4 (48.7–78.2) | <0.001 | 97.6 (92.8–100) | 100 (100–100) | 0.311 | 97.6 (92.8–100) | 100 (100–100) | 0.311 |
| Accuracy | 96.8 (94.5–99.1) | 96.9 (94.6–99.2) | 0.966 | 90.9 (87.3–94.6) | 90.5 (86.9–94.2) | 0.876 | 74.7 (69.9–79.5) | 76.9 (72.3–81.5) | 0.481 | 75.3 (70.6–80.0) | 80.6 (76.3–84.9) | 0.058 | 47.8 (42.3–53.3) | 70.6 (65.6–75.6) | <0.001 |
| PPV | 99.5 (98.5–100) | 100 (100–100) | 0.316 | 99.5 (98.5–100) | 100 (100–100) | 0.316 | 98.5 (96.9–100) | 93.6 (90.5–96.7) | 0.003 | 99.5 (98.5–100) | 100 (100–100) | 0.316 | 99.1 (97.4–100) | 100 (100–100) | 0.315 |
| NPV | 66.7 (44.9–88.4) | 12.5 (0–35.4) | <0.001 | 48.8 (33.5–64.1) | 11.5 (0–23.8) | <0.001 | 32.8 (24.2–41.3) | 30.6 (20.8–40.4) | 0.585 | 33.9 (25.4–42.4) | 39.8 (30.4–49.3) | 0.054 | 19.4 (14.0–24.8) | 30.4 (22.6–38.1) | <0.001 |
| Nodules ≥1 cm (n=177) | | | | | | | | | | | | | | | |
| TP | 102 | 105 | | 102 | 105 | | 102 | 108 | | 102 | 105 | | 102 | 65 | 91 |
| TN | 11 | 0 | | 19 | 2 | | 34 | 23 | | 36 | 37 | | 36 | 37 | |
| FP | 1 | 0 | | 1 | 0 | | 3 | 14 | | 1 | 0 | | 1 | 0 | |
| FN | 3 | 3 | | 10 | 11 | | 38 | 32 | | 38 | 35 | | 75 | 49 | |
| Sensitivity | 97.1 (93.9–100) | 97.2 (94.1–100) | 0.972 | 91.1 (85.8–96.4) | 90.5 (85.2–95.8) | 0.888 | 72.9 (65.5–80.2) | 77.1 (70.2–84.1) | 0.302 | 72.9 (65.5–80.2) | 75 (67.8–82.2) | 0.589 | 46.4 (38.2–54.7) | 65 (57.1–72.9) | <0.001 |
| Specificity | 91.7 (76.0–100) | NA | - | 95 (85.4–100) | 100 (100–100) | 0.305 | 91.9 (83.1–100) | 62.2 (46.5–77.8) | 0.001 | 97.3 (92.1–100) | 100 (100–100) | 0.311 | 97.3 (92.1–100) | 100 (100–100) | 0.311 |
| Accuracy | 96.6 (93.3–99.9) | 97.2 (94.1–100) | 0.783 | 91.7 (87–96.4) | 90.7 (85.4–95.9) | 0.789 | 76.8 (70.6–83.1) | 74.0 (67.6–80.5) | 0.474 | 78 (71.9–84.0) | 80.2 (74.4–86.1) | 0.479 | 57.1 (49.8–64.4) | 72.3 (65.7–78.9) | <0.001 |
| PPV | 99.0 (97.1–100) | 100 (100–100) | 0.315 | 99 (97.1–100) | 100 (100–100) | 0.315 | 97.1 (93.9–100) | 88.5 (82.9–94.2) | 0.004 | 99.0 (97.1–100) | 100 (100–100) | 0.315 | 98.5 (95.5–100) | 100 (100–100) | 0.314 |
| NPV | 78.6 (57.1–100) | 0 (0–0) | <0.001 | 65.5 (48.2–82.8) | 15.4 (–4.2–35) | <0.001 | 47.2 (35.7–58.8) | 41.8 (28.8–54.9) | 0.310 | 48.6 (37.3–60.0) | 51.4 (39.8–62.9) | 0.479 | 32.4 (23.7–41.1) | 43.0 (32.6–53.5) | <0.001 |
| Nodules ≥2 cm (n=58) | | | | | | | | | | | | | | | |
| TP | 17 | 19 | | 17 | 19 | | 17 | 22 | | 17 | 19 | | 15 | 17 | |
| TN | 9 | 0 | | 16 | 2 | | 24 | 14 | | 25 | 25 | | 25 | 25 | |
| FP | 0 | 0 | | 0 | 0 | | 1 | 11 | | 0 | 0 | | 0 | 0 | |
| FN | 1 | 1 | | 3 | 4 | | 16 | 11 | | 16 | 14 | | 18 | 16 | |

Table 4. Diagnostic Performance of FNA versus CNB According to Nodule Size (Continued)

| | Criteria 1 | | | Criteria 2 | | | Criteria 3 | | | Criteria 4 | | | Criteria 5 | | |
|-------------|--------------------|------------------|---------|--------------------|---------------------|---------|---------------------|---------------------|---------|---------------------|---------------------|---------|---------------------|---------------------|---------|
| | FNA | CNB | p value | FNA | CNB | p value | FNA | CNB | p value | FNA | CNB | p value | FNA | CNB | p value |
| Sensitivity | 94.4 (83.9–100) | 95 (85.4–100) | 0.939 | 85 (69.4–100) | 82.6 (67.1–98.1) | 0.832 | 51.5 (34.5–68.6) | 66.7 (50.6–82.8) | 0.118 | 51.5 (34.5–68.6) | 57.6 (40.7–74.4) | 0.476 | 45.5 (28.5–62.4) | 51.5 (34.5–68.6) | 0.525 |
| Specificity | 100 (100–100) | NA | - | 100 (100–100) | 100 (100–100) | >0.999 | 96 (88.3–100) | 56 (36.5–75.5) | <0.001 | 100 (100–100) | 100 (100–100) | >0.999 | 100 (100–100) | 100 (100–100) | >0.999 |
| Accuracy | 96.3 (89.2–100) | 95 (85.4–100) | 0.832 | 91.7 (82.6–100) | 84 (69.6–98.4) | 0.382 | 70.7 (58.9–82.4) | 62.1 (49.6–74.6) | 0.270 | 72.4 (60.9–83.9) | 75.9 (64.8–86.9) | 0.478 | 69 (57.1–80.9) | 72.4 (60.9–83.9) | 0.526 |
| PPV | 100 (100–100) | 100 (100–100) | >0.999 | 100 (100–100) | 100 (100–100) | >0.999 | 94.4 (83.9–100) | 66.7 (50.6–82.8) | <0.001 | 100 (100–100) | 100 (100–100) | >0.999 | 100 (100–100) | 100 (100–100) | >0.999 |
| NPV | 90 (71.4–100) | 0 (0–0) | <0.001 | 84.2 (67.8–100) | 33.3 (-4.4–71.1) | 0.013 | 60 (44.8–75.2) | 56 (36.5–75.5) | 0.591 | 60.9 (46.0–75.9) | 64.1 (49.0–79.2) | 0.478 | 58.1 (43.4–72.9) | 61 (46.0–75.9) | 0.527 |

FNA, fine-needle aspiration; CNB, core-needle biopsy; TP, true-positive; TN, true-negative; FP, false-positive; FN, false-negative; PPV, positive predictive value; NPV, negative predictive value; NA, not applicable; SUSP, suspicious for malignancy; FN/SFN, follicular neoplasm or suspicious for follicular neoplasm; AUS/FLUS, atypia of undetermined significance or follicular lesion of undetermined significance. Values are numbers or percentages (95% confidence intervals). In Criteria 1, “benign” was considered as test-negative, and “SUSP” and “malignancy” were considered as test-positive. In Criteria 2, “benign” was considered as test-negative, and “FN/SFN,” “SUSP” and “malignancy” were considered as test-positive. In Criteria 3, “nondiagnostic,” “benign,” and “AUS/FLUS” were considered as test-negative, and the remaining categories were considered as test-positive. In Criteria 4, “SUSP” and “malignancy” were considered as test-positive, and the remaining categories were considered as test-negative. In Criteria 5, “malignancy” was considered as test-positive, and the remaining categories were considered as test-negative.

DISCUSSION

In this study, the diagnostic performance of FNA was mostly the same as that of CNB, regardless of nodule size, except for NPV, which was better for FNA than for CNB when applying Criteria 1 and 2. When applying Criteria 3, the specificity and PPV of FNA were superior to CNB regardless of size. When applying Criteria 4, diagnostic performance did not differ between FNA and CNB, regardless of size, unlike the results of a previous study that found CNB to have higher sensitivity and accuracy than FNA.⁴ After applying Criteria 5, diagnostic performance did not differ between FNA and CNB in nodules ≥2 cm (not small nodules). However, in nodules ≥1 cm and all nodules, the sensitivity, accuracy, and NPV of CNB were better than those of FNA, findings which were similar to those of a published study.⁸

Several investigators have suggested that CNB is superior to FNA and that it may be a potential first-line diagnostic tool.^{4,8,14,16} In these studies, the “nondiagnostic,” “benign,” “AUS/FLUS,” and “FN/SFN” categories were regarded as test-negative, and the “nondiagnostic,” “benign,” “AUS/FLUS,” and “FN/SFN” categories with and without the “SUSP” category were regarded as test-positive, which were the same designations used for Criteria 4 and 5, respectively, in our study.^{4,8,14,16} In other studies, the diagnostic performance of FNA was calculated by regarding the “benign” category with and without the “nondiagnostic” category as test-negative (Criteria 1 and 2 in our study, respectively) or by regarding the combination of the “nondiagnostic,” “benign,” and “AUS/FLUS” categories as test-negative (Criteria 3 in our study).^{23,24} According to BSRTC, the risk of malignancy was 15–30% and 60–75% for thyroid nodules categorized as “FN/SFN” and “SUSP,” respectively.¹² Considering that “FN/SFN” and “SUSP” nodules have relatively high risks of malignancy and are recommended for surgery according to the BSRTC, a nodule with the above cytologic categories should not be considered as test-negative when calculating the diagnostic performances of FNA or CNB (Criteria 4 and 5). Based on the results of our study, we need to reconsider previous research that found CNB to be more reliable or accurate than FNA, and CNB cannot be considered as a first-line diagnostic tool for diagnosing thyroid nodules without more supportive evidence.

To obtain tissue samples without damage to the adjacent tissues, it is important to visualize the entire needle. Therefore, the core needle should be positioned parallel to the transducer. Unlike the core needle in CNB, the fine needle in FNA can be inserted in various directions with to-and-fro motions to obtain cells from different sites of the thyroid nodule for representative sampling.²⁵ Differences in the methods of FNA and CNB may explain their different diagnostic performances.

Percutaneous CNB is a useful diagnostic tool for diagnosing soft tissue lesions, especially breast lesions, and surgical excision biopsies can be avoided with its use.^{26,27} CNB has also been

reported as a safe tool for diagnosing thyroid nodules with low complication rates.² However, as the core needle is thicker than the fine needle, it might cause more severe and permanent damage to the adjacent major organs than the thinner fine needle (Supplementary Figs. 1 and 2, only online).

This study has several limitations. First, patient selection bias exists because we included patients who underwent both FNA and CNB before surgery. Second, we could not evaluate the experience levels of the physicians who performed both FNA and CNB as these procedures took place at outside clinics. However, the results of this study can reflect the real clinical value of FNA and CNB. Third, the results of the first procedure may have affected the next procedure. Commonly, FNA is performed before CNB even when they are both performed during the same session/visit. Therefore, the results of CNB can be influenced by the procedures undertaken for FNA. Fourth, although our findings indicate that CNB cannot be a main diagnostic tool for thyroid nodules, it can still play a role in diagnosing thyroid lymphoma, anaplastic carcinoma, and metastasis from other organs.²⁹⁻³⁰ In this study, we were unable to evaluate the selective indications of CNB. Finally, we did not assess diagnostic performance for predicting neoplasm because there were only a small number of FNs in this study.

In conclusion, CNB did not show superior diagnostic performance to FNA for diagnosing thyroid nodules.

ACKNOWLEDGEMENTS

There was no funding for this research.

AUTHOR CONTRIBUTIONS

Conceptualization: Jin Young Kwak and Ilah Shin. **Data curation:** Jin Young Kwak and Ilah Shin. **Formal analysis:** Hye Sun Lee. **Investigation:** Ilah Shin, Si Eun Lee, Eun-Kyung Kim, Hee Jung Moon, Jung Hyun Yoon, and Vivian Youngjean Park. **Methodology:** Ilah Shin and Hye Sun Lee. **Project administration:** Jin Young Kwak. **Resources:** Eun-Kyung Kim, Hee Jung Moon, Jung Hyun Yoon, Vivian Youngjean Park, and Si Eun Lee. **Software:** Ilah Shin and Hye Sun Lee. **Supervision:** Jin Young Kwak and Eun-Kyung Kim. **Validation:** Eun-Kyung Kim, Hee Jung Moon, Jung Hyun Yoon, and Vivian Youngjean Park. **Visualization:** Ilah Shin and Si Eun Lee. **Writing—original draft:** Jin Young Kwak and Ilah Shin. **Writing—review and editing:** Eun-Kyung Kim, Hee Jung Moon, Jung Hyun Yoon, Vivian Youngjean Park, and Si Eun Lee. **Approval of final manuscript:** All authors.

ORCID iDs

| | |
|-----------------------|---|
| Ilah Shin | https://orcid.org/0000-0002-2046-9426 |
| Eun-Kyung Kim | https://orcid.org/0000-0002-3368-5013 |
| Hee Jung Moon | https://orcid.org/0000-0002-5643-5885 |
| Jung Hyun Yoon | https://orcid.org/0000-0002-2100-3513 |
| Vivian Youngjean Park | https://orcid.org/0000-0002-5135-4058 |
| Si Eun Lee | https://orcid.org/0000-0002-3225-5484 |
| Hye Sun Lee | https://orcid.org/0000-0001-6328-6948 |
| Jin Young Kwak | https://orcid.org/0000-0002-6212-1495 |

REFERENCES

- Kim SY, Lee HS, Moon J, Kim EK, Moon HJ, Yoon JH, et al. Fine-needle aspiration versus core needle biopsy for diagnosis of thyroid malignancy and neoplasm: a matched cohort study. *Eur Radiol* 2017;27:801-11.
- Ha EJ, Suh CH, Baek JH. Complications following ultrasound-guided core needle biopsy of thyroid nodules: a systematic review and meta-analysis. *Eur Radiol* 2018;28:3848-60.
- Park KT, Ahn SH, Mo JH, Park YJ, Park DJ, Choi SI, et al. Role of core needle biopsy and ultrasonographic finding in management of indeterminate thyroid nodules. *Head Neck* 2011;33:160-5.
- Sung JY, Na DG, Kim KS, Yoo H, Lee H, Kim JH, et al. Diagnostic accuracy of fine-needle aspiration versus core-needle biopsy for the diagnosis of thyroid malignancy in a clinical cohort. *Eur Radiol* 2012;22:1564-72.
- Yeon JS, Baek JH, Lim HK, Ha EJ, Kim JK, Song DE, et al. Thyroid nodules with initially nondiagnostic cytologic results: the role of core-needle biopsy. *Radiology* 2013;268:274-80.
- Paja M, del Cura JL, Zabala R, Corta I, Lizarraaga A, Oleaga A, et al. Ultrasound-guided core-needle biopsy in thyroid nodules. A study of 676 consecutive cases with surgical correlation. *Eur Radiol* 2016; 26:1-8.
- Suh CH, Baek JH, Park C, Choi YJ, Lee JH. The role of core needle biopsy for thyroid nodules with initially indeterminate results on previous fine-needle aspiration: a systematic review and meta-analysis. *AJNR Am J Neuroradiol* 2017;38:1421-6.
- Hong MJ, Na DG, Kim SJ, Kim DS. Role of core needle biopsy as a first-line diagnostic tool for thyroid nodules: a retrospective cohort study. *Ultrasonography* 2018;37:244-53.
- Xiong Y, Yan L, Nong L, Zheng Y, Li T. Pathological diagnosis of thyroid nodules based on core needle biopsies: comparative study between core needle biopsies and resected specimens in 578 cases. *Diagn Pathol* 2019;14:10.
- Yoon JH, Kwak JY, Moon HJ, Kim EK. Ultrasonography-guided core needle biopsy did not reduce diagnostic lobectomy for thyroid nodules diagnosed as atypia of undetermined significance/follicular lesion of undetermined significance. *Ultrasound Q* 2019;35: 253-8.
- Yoon JH, Kim EK, Kwak JY, Moon HJ. Effectiveness and limitations of core needle biopsy in the diagnosis of thyroid nodules: review of current literature. *J Pathol Transl Med* 2015;49:230-5.
- Cibas ES, Ali SZ. The Bethesda System for Reporting Thyroid Cytopathology. *Thyroid* 2009;19:1159-65.
- Lee HJ, Kim YJ, Han HY, Seo JY, Hwang CM, Kim K. Ultrasound-guided needle biopsy of large thyroid nodules: core needle biopsy yields more reliable results than fine needle aspiration. *J Clin Ultrasound* 2019;47:255-60.
- Kim HC, Kim YJ, Han HY, Yi JM, Baek JH, Park SY, et al. First-line use of core needle biopsy for high-yield preliminary diagnosis of thyroid nodules. *AJNR Am J Neuroradiol* 2017;38:357-63.
- Choi YJ, Baek JH, Suh CH, Shim WH, Jeong B, Kim JK, et al. Core-needle biopsy versus repeat fine-needle aspiration for thyroid nodules initially read as atypia/follicular lesion of undetermined significance. *Head Neck* 2017;39:361-9.
- Suh CH, Baek JH, Lee JH, Choi YJ, Kim JK, Sung TY, et al. The role of core-needle biopsy as a first-line diagnostic tool for initially detected thyroid nodules. *Thyroid* 2016;26:395-403.
- Chae IH, Kim EK, Moon HJ, Yoon JH, Park VY, Kwak JY. Ultrasound-guided fine needle aspiration versus core needle biopsy: comparison of post-biopsy hematoma rates and risk factors. *Endocrine* 2017;57:108-14.

18. Chung SR, Baek JH, Lee JH, Lee YM, Sung TY, Chung KW, et al. Risk of malignancy according to the sub-classification of atypia of undetermined significance and suspicious follicular neoplasm categories in thyroid core needle biopsies. *Endocr Pathol* 2019;30:146-54.
19. Tessler FN, Middleton WD, Grant EG, Hoang JK, Berland LL, Teefey SA, et al. ACR Thyroid Imaging, Reporting and Data System (TI-RADS): white paper of the ACR TI-RADS Committee. *J Am Coll Radiol* 2017;14:587-95.
20. Haugen BR, Alexander EK, Bible KC, Doherty GM, Mandel SJ, Nikiforov YE, et al. 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* 2016;26:1-133.
21. Yoon JH, Lee HS, Kim EK, Moon HJ, Kwak JY. Malignancy risk stratification of thyroid nodules: comparison between the Thyroid Imaging Reporting and Data System and the 2014 American Thyroid Association management guidelines. *Radiology* 2016;278:917-24.
22. Jung CK, Min HS, Park HJ, Song DE, Kim JH, Park SY, et al. Pathology reporting of thyroid core needle biopsy: a proposal of the Korean Endocrine Pathology Thyroid Core Needle Biopsy Study Group. *J Pathol Transl Med* 2015;49:288-99.
23. Eszlinger M, Ullmann M, Ruschenburg I, Böhme K, Görke F, Franzius C, et al. Low malignancy rates in fine-needle aspiration cytologies in a primary care setting in Germany. *Thyroid* 2017;27:1385-92.
24. Ke J, Jianyong L, Ying L, Genpeng L, Linlin S, Zhihui L, et al. The use of The Bethesda System for Reporting Thyroid Cytopathology in a Chinese population: an analysis of 13351 specimens. *Diagn Cytopathol* 2019;47:876-80.
25. Gharib H, Papini E, Paschke R, Duick DS, Valcavi R, Hegedüs L, et al. American Association of Clinical Endocrinologists, Associazione Medici Endocrinologi, and European Thyroid Association Medical guidelines for clinical practice for the diagnosis and management of thyroid nodules: executive summary of recommendations. *Endocr Pract* 2010;16:468-75.
26. Schizas N, Lazopoulos A, Krimiotis D, Rallis T, Paliouras D, Gogakos A, et al. Beware of hemopneumothorax following core needle breast biopsy. *Respir Med Case Rep* 2018;25:49-51.
27. Kim SY, Chung HW, Oh TS, Lee JS. Practical guidelines for ultrasound-guided core needle biopsy of soft-tissue lesions: transformation from beginner to specialist. *Korean J Radiol* 2017;18:361-9.
28. Kwak JY, Kim EK, Ko KH, Yang WI, Kim MJ, Son EJ, et al. Primary thyroid lymphoma: role of ultrasound-guided needle biopsy. *J Ultrasound Med* 2007;26:1761-5.
29. Ha EJ, Baek JH, Lee JH, Kim JK, Song DE, Kim WB, et al. Core needle biopsy could reduce diagnostic surgery in patients with anaplastic thyroid cancer or thyroid lymphoma. *Eur Radiol* 2016;26:1031-6.
30. Song OK, Koo JS, Kwak JY, Moon HJ, Yoon JH, Kim EK. Metastatic renal cell carcinoma in the thyroid gland: ultrasonographic features and the diagnostic role of core needle biopsy. *Ultrasonography* 2017;36:252-9.