Which thyroid nodules should biopsy be performed: a comparison of three recommendations from the literature

Sung Soo Ahn

Department of Medicine

The Graduate School, Yonsei University

Which thyroid nodules should biopsy be performed: a comparison of three recommendations from the literature

Sung Soo Ahn

Department of Medicine

The Graduate School, Yonsei University

Which thyroid nodules should biopsy be performed: a comparison of three recommendations from the literature

Directed by Professor Eun-Kyung Kim

The Master's Thesis
submitted to the Department of Medicine,
the Graduate School of Yonsei University
in partial fulfillment of the requirements for the degree
of Master of Medical Science

Sung Soo Ahn

June 2009

This certifies that the Master's Thesis of
Sung Soo Ahn is approved.

Thesis Supervisor : Eun-Kyung Kim

Dae Ryong Kang

Sung-Kil Lim

The Graduate School
Yonsei University

June 2009

ACKNOWLEDGEMENTS

I acknowledge my deep gratitude to Dr. Eun-Kyung Kim, who is training me in thyroid division and is also my thesis director, for supporting my efforts with total commitment and facilitating every step of the process. My appreciation for her guidance and encouragement is tremendous.

Also, I am indebted to Dr. Dae Ryong Kang, for his help for statistical consultation and analysis of the data as well as working enthusiastically to assure the superior quality of this paper.

Finally, I would like to thank Drs. Sung-Kil Lim, Jin Young Kwak, Min Jung Kim for their support of me personally.

<TABLE OF CONTENTS>

LIST OF FIGURES

LIST OF TABLES

<ABSTRACT>

Which thyroid nodules should biopsy be performed: a comparison of three recommendations from the literature

Sung Soo Ahn

Department of Medicine
The Graduate School, Yonsei University

(Directed by Professor Eun-Kyung Kim)

Purpose: To compare the results of three guidelines for fine needle aspiration biopsy (FNAB) of thyroid nodules.

Materials and Methods: A total of 1398 nodules, confirmed by FNAB or surgery, were included. We compared the diagnostic performance of several guidelines: 1) Kim's Criteria: a nodule with at least one of the following findings: marked hypoechogenicity, irregular or microlobulated margin, microcalcifications, or taller than wide shape; 2) the Society of Radiologists in Ultrasound (SRU): a nodule ≥ 1 cm with microcalcifications, a nodule ≥ 1.5 cm which is solid or has coarse calcifications, or a nodule ≥ 2 cm with a solid component; 3) American Association of Clinical Endocrinologists (AACE): a hypoechoic nodule with at least one additional feature such as irregular margins, taller than wide shape, or microcalcifications. The diagnostic value of these criteria was assessed.

Results: For all nodules, the diagnostic accuracy of Kim's criteria (0.868) and AACE (0.85) were more superior to that of SRU (0.551). Number of performing FNAB was lowest in AACE criteria (25.6%), whereas the number of missed malignancy was lowest in Kim's criteria (7.3%). When considering nodules larger than 1 cm, the results did not change.

Conclusions: Kim's criteria and AACE criteria are more accurate than SRU criteria. AACE's guideline is recommended to achieve high specificity, whereas Kim's criteria may be chosen for its higher sensitivity.

Key words : thyroid, ultrasound, fine needle aspiration biopsy

Which thyroid nodules should biopsy be performed: a comparison of three recommendations from the literature

Sung Soo Ahn

Department of Medicine
The Graduate School, Yonsei University

(Directed by Professor Eun-Kyung Kim)

I. INTRODUCTION

Fine needle aspiration biopsy (FNAB) is considered to be the most reliable diagnostic test for thyroid nodule evaluation and has a very low rate of complication, especially with ultrasound (US) guidance. FNAB increases the detection of thyroid cancer and reduces the number of unnecessary surgical procedures and the overall cost of medical care [1-3]. However, performing FNAB on every thyroid nodule detected on US may not be cost-effective due to their high prevalence. Therefore, previous studies have sought to identify sonographic features suggestive of malignancy. These findings include hypoechogenicity, irregular or microlobulated margins, calcifications, taller than wide shape, intranodular vascularity, solitary, and solid components [4-11]. However, no single US feature has a high positive predictive value for cancer that can be used to identify which nodules should undergo FNAB. Because of this, several studies defined guidelines with various combinations of US features that are both sensitive and specific for predicting cancer, though there is still extensive debate as to which guidelines are the most accurate and effective [4, 5, 7-9, 12, 13]. It is difficult to compare the accuracy of the different guidelines, however, because the populations included in each study were different.

In this study, we calculated risks for the malignancy of thyroid nodules with

known suspicious sonographic features, three sonographic criteria from previous reports were applied to the same population to assess the diagnostic performance of those criteria for predicting histologically-confirmed malignancy.

II. MATERIALS AND METHODS

The Institutional Review Board of our university approved this study and informed consent was waived for the study. However, informed consent for FNAB was obtained from all patients prior to biopsy.

1. Study population

Between September 2002 and July 2004, US-guided FNABs were performed in 1492 consecutive patients with 1583 nodules composed of 1371 women (91.9%) and 120 men (8%), with a mean age of 49.0 years (range, 9-82). Among them, 626 (42%) patients had multiple nodules. At our institution, US-FNABs were performed on thyroid nodules with suspicious US features regardless of size. FNABs were performed on the largest of thyroid nodules in patients with multiple thyroid nodules that have similar US features, whereas, FNAB was performed in each nodule when multiple nodules showed several different US features. Among 1583 nodules, thyroidectomies were performed on 455 patients with thyroid nodules, including 33 nondiagnostic (20.5%), 111 benign (10.9%), 27 follicular neoplasm (51.9%), 78 nodules suspicious for papillary carcinoma (70.9%), and 206 malignant (84.4%) nodules on cytological evaluation.

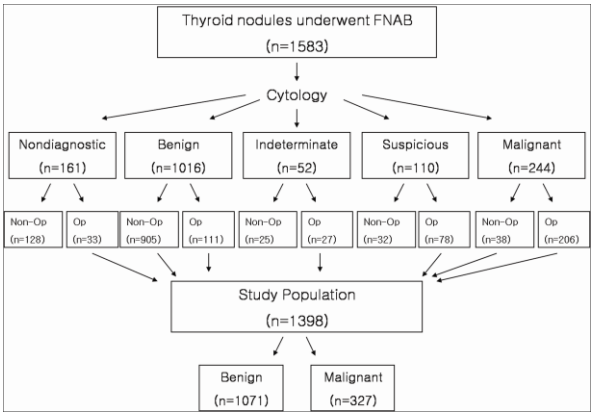

Figure 1. Study population. FNAB: fine needle aspiration biopsy, Op: thyroidectomy.

Because surgery following FNAB was not performed, 128 out of 161 nodules with nondiagnostic cytology, 25 of 52 nodules with follicular neoplasm cytology, and 32 of 110 with suspicious for papillary carcinomas were excluded. Nodules with benign (n=1016) or malignant (n=244) cytology were included, whether they underwent surgery or not.

Finally, 1398 nodules were included in this study. There were 1071 (76.6%) benign nodules and 327 (23.4%) malignant nodules.

Table 1. Size and pathology of thyroid nodules in our study population

	Benign (%)	Malignant (%)	Total
< 1 cm	233 (58)	169 (42)	402
≥ 1 cm	838 (84.1)	158 (15.9)	996
Total	1071 (76.6)	327 (23.4)	1398

The diagnosis of malignancy included papillary carcinoma (n=319), follicular carcinoma (n=3), lymphoma (n=1), medullary carcinoma (n=1), anaplastic carcinoma (n=1), metastasis from breast cancer (n=1), and poorly differentiated carcinoma (n=1).

2. Sonographic evaluation and US-guided FNAB

Thyroid US was performed with an HDI 3000 or 5000 scanner (Phillips, Bothell, Washington, USA) using a bandwidth of 7 to 12-MHz transducer. All images were sent to the local picture archiving and communication system for review. One experienced radiologist who performed FNABs prospectively described the sonographic characteristics of thyroid nodules with respect to the following parameters: 1) size, 2) multiplicity, 3) composition, 4) echogenicity, 5) margin, 6) calcification, 7) shape, and 8) abnormal cervical LN. The size was measured at the maximum dimension. Substantial growth was retrospectively assessed in 287 nodules that underwent US examinations at least 6 months prior to FNAB. The composition of the nodule was classified as purely cystic, cystic, mixed, or solid based on the ratio of the cystic portion to the solid portion in the nodule. A purely cystic nodule meant a completely anechoic nodule with or without a comet tail artifact. A nodule having both a cystic and solid portion was defined as a mixed echoic nodule. Echogenicity of the mixed echoic nodule was assessed based on the solid portion. The echogenicity of the nodule was compared to that of the surrounding parenchyma, and it was classified as hypoechogenicity, isoechogenicity, and hyperechogenicity. Marked hypoechogenicity was defined as decreased echogenicity compared to the cervical strap muscle. The margin of a nodule was described as well-circumscribed or not well-circumscribed, which included irregular or microlobulated. Microcalcifications were defined as tiny (less than 2 mm), punctuate and hyperechoic foci, with or without acoustic shadows. Macrocalcifications or coarse calcifications were defined as larger than 2 mm. Mixed calcifications were defined as a combination of microcalcifications and macrocalcifications, which were reclassified as the presence of microcalcifications. Shape was assessed by the ratio of anteroposterior (A) and transverse (T) diameter ($A/T \geq 1$ or $A/T < 1$). Abnormal lymph nodes (LN) were

defined as LN with heterogeneous echotexture, calcifications, cystic areas, and round shape.

US-FNABs were performed by one experienced radiologist using a 23-gauge needle attached to a 20-mL disposable plastic syringe and aspirator. Each lesion was aspirated at least twice. Materials obtained from the aspiration biopsy were expelled onto glass slides and smeared. All smears were placed immediately in 95% alcohol for Papanicolaou staining. The remainder of the material was rinsed in saline for processing as a cell block. The cytopathologist was not on site during the biopsy. Additional special staining was made on a case-by-case basis according to the requirements of the cytopathologist.

3. Comparisons of three guidelines

Three sonographic criteria were applied to our data to compare the accuracy of the various guidelines for predicting cancer. In addition, we regrouped the study population with nodules ≥1 cm in diameter and analyzed them in the same way.

A. Kim's criteria [4]

According to this guideline, FNAB for nodules with any single suspicious US feature, regardless of its size, is recommended. Suspicious sonographic features are defined as marked hypoechogenicity, irregular or microlobulated margins, microcalcifications, and A/T ≥ 1.

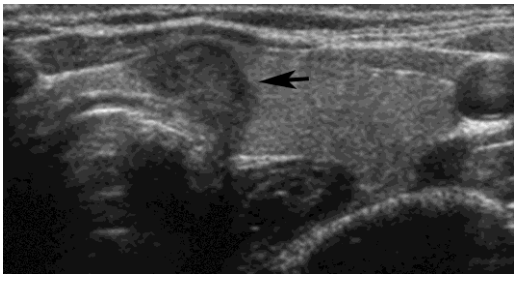

Figure 2. Transverse US image of thyroid gland in a 42-year-old woman demonstrates a 10 mm sized, wider than tall shaped, hypoechoic nodule with an irregular margin in the left thyroid gland. These US findings may suggest performing FNAB based on Kim's and AACE criteria. The nodule was

confirmed as papillary carcinoma with extrathyroidal invasion.

B. The Society of Radiologists in Ultrasound (SRU) criteria [7]

This guideline makes recommendations based on the size and US characteristics of thyroid nodules. The recommendations for FNAB are as follows: (i) a nodule ≥ 1.0 cm if microcalcifications are present; (ii) a nodule ≥ 1.5 cm if completely or almost entirely solid, or if there are coarse calcifications (classified as macrocalcifications); (iii) a nodule ≥ 2.0 cm if there are mixed solid and cystic components, or almost entirely cystic with a solid mural component; (iv) a nodule that has shown substantial growth since prior US examination.

Figure 3. Transverse US image of thyroid gland in a 53-year-old woman demonstrates a 8 mm sized, microlobulated, taller than wide shaped, isoechoic nodule. FNAB was recommended in this nodule based only on Kim's criteria and the nodule was confirmed as papillary carcinoma with extrathyroidal invasion.

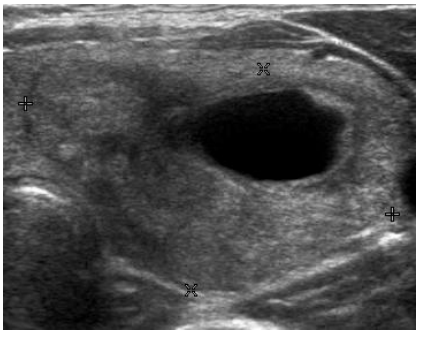

Figure 4. Transverse US image of thyroid gland in a 32-year-old woman shows a 4 cm, well-defined, mixed solid, cystic nodule, which could be recommend

for FNAB only by SRU. It proved to be adenomatous hyperplasia after surgery.

Since there was no consensus on the definition of substantial growth that would require biopsy, but we included nodules showing increase in maximal diameter of 3 mm or more [14]. This guideline also recommends FNAB when abnormal LNs are present regardless of US features of thyroid nodules. US features associated with a high risk of cancer are defined as LNs with heterogeneous echotexture, calcifications, cystic areas, round shape, and mass effect. LN size is not mentioned, and we therefore included morphologically abnormal LNs regardless of size.

C. American Association of Clinical Endocrinologists (AACE) criteria [8]

According to the AACE guidelines, FNAB should be performed on all hypoechoic nodules with at least one of the following additional US features: irregular margins, intranodular vascular spots, taller than wide shape, or microcalcifications (Fig. 2). However, the parameter of intranodular vascularity was excluded because color Doppler sonography was not assessed in our study.

4. Statistical analyses

Statistical analyses were performed using the SAS software package (version 9.1.3) and MedCalc (version 9.3.6.0). The odds ratio (OR) of malignancy for each US feature was calculated using logistic regression analysis. We calculated sensitivity, specificity, positive predictive value (PPV), negative predictive value (NPV), and accuracy for individual sonographic characteristics as well as various combinations of sonographic features as follows: sensitivity, TP/(TP+FN); specificity, TN/(TN+FP); positive predictive value, TP/(TP+FP); negative predictive value, TN/(TN+FN); and accuracy, (TP+TN)/(TP+FP+FN+TN), where TP is the true-positive, TN the true-negative, FP the false positive and FN the false-negative.

The diagnostic accuracy of malignancy predictions was calculated by the receiver operating characteristics (ROC) curve and area under the curve (AUC)

based on each guideline. The number of FNABs and the rate of missed carcinomas were assessed according to different criteria and their statistical significances were evaluated based on Yate's correction for continuity.

III. RESULTS

1. US findings

The mean size of the nodules was 16.6 mm (range, 3-100 mm) and 402 nodules (28.8%) were less than 1 cm in diameter. The mean size of benign and malignant nodules was 17.9 mm (range, 3-100 mm) and 12.1 mm (range, 3-80mm), respectively. Among 287 nodules, which had US examinations at least 6 months prior to FNAB, 36 nodules (12.5%) showed increase in maximal diameter of 3 mm or more.

Independent predictors of malignant thyroid nodules on US are irregular or microlobulated margins, microcalcifications, marked hypoechogenicity, hypoechogenicity, taller than wide shape, solidity, and calcifications.

Table 2. Diagnostic accuracy of US findings and their combinations for malignant nodules

Characteristics	Odds Ratio	Sensitivity (%)	Specificity (%)	PPV (%)	NPV (%)	AUC
Irregular or microlobulated margin	30.811	276/327 (84.4)	911/1071 (85.1)	276/436 (63.3)	911/962 (94.7)	0.847
Microcalcifications	28.985	120/327 (36.7)	1050/1071 (98)	120/141 (85.1)	1050/1257 (83.5)	0.674
Marked hypoechogenicity	19.448	120/327 (36.7)	1040/1071 (97.1)	120/151 (79.5)	1040/1247 (83.4)	0.669
Hypoechogenicity	18.387	265/327 (81.0)	869/1071 (81.1)	265/467 (56.7)	869/931 (93.9)	0.811
Taller than wide	17.633	183/327 (56.0)	999/1071 (93.3)	183/255 (71.8)	999/1143 (87.4)	0.746
Solid	11.338	292/327 (89.3)	617/1071 (57.6)	292/746 (39.1)	617/652 (94.6)	0.735
Calcifications	9.836	149/327 (45.6)	987/1071 (92.2)	149/233 (63.7)	987/1165 (83.9)	0.689
Solitary	1.053	187/327 (57.2)	472/1071 (44.1)	187/786 (23.8)	472/612 (77.1)	0.506
≥1 cm	0.260	158/327 (48.3)	233/1071 (21.8)	158/996 (15.9)	233/402 (58)	0.65
Abnormal lymph nodes	NC	17/327 (5.2)	1071/1071 (100)	17/17 (100)	1071/1381 (77.6)	0.526
Substantial growth	NC	NC	NC	2/36 (5.6)	193/251 (76.9)	NC

PPV: positive predictive value; NPV: negative predictive value; AUC: area under the receiver operating characteristic curve; NC: non-calculable. These seven US findings showed significant associations with malignancy ($p<0.001$). Morphologically abnormal LNs were observed only in cases with malignancy

(17/327). Solitary or multiplicity was not associated with the risk of malignancy (OR=1.1). The frequency of each suspicious US feature based on nodule size is demonstrated in Table 3. All suspicious findings, except abnormal LN, were found significantly more frequently in nodules smaller than 1 cm in size.

Table 3. Frequency of suspicious US features according to the size of the nodule

Characteristics	<1 cm (n=402)	≥1 cm (n=996)	p-value
Irregular or microlobulated margin	215 (53.5)	221 (22.2)	<0.001
Microcalcifications	52 (12.9)	89 (8.9)	0.025
Marked hypoechogenicity	97 (24.1)	54 (5.4)	<0.001
Hypoechogenicity	226 (56.2)	241 (24.2)	<0.001
Taller than wide	159 (39.6)	96 (9.6)	<0.001
Solid	293 (72.9)	453 (45.5)	<0.001
Calcifications	83 (20.6)	150 (15.1)	0.014
Solitary	227 (56.5)	559 (56.1)	0.954
Abnormal lymph nodes	3 (0.7)	14 (1.4)	0.309

2. Diagnostic performance according to several guidelines

Sensitivity, specificity, PPV, and NPV of Kim's criteria were estimated to be 92.7%, 80.9%, 59.6%, and 97.3%, respectively. AUC based on Kim's criteria was 0.868. AUC based on SRU criteria was 0.551 ($p<0.001$), whereas it was 0.850 for the AACE criteria ($p=0.179$) (Fig. 5). Sensitivity, specificity, PPV, and NPV of each criterion are given in Table 4.

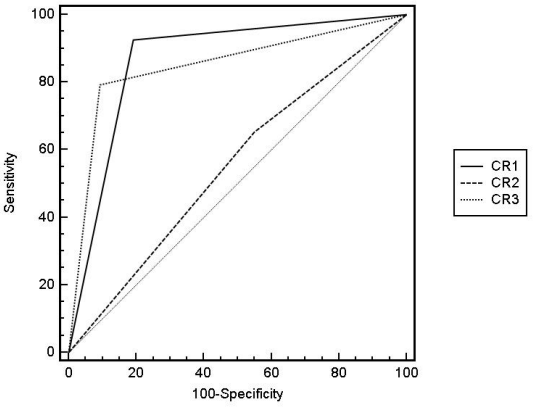

Fig. 5. Receiver operator characteristics for each criterion. CR1: Kim's Criteria; CR2: SRU Criteria; CR3: AACE Criteria.

Table 4. Diagnostic accuracy of each criterion

Criteria	Sensitivity (%)	Specificity (%)	PPV (%)	NPV (%)	AUC	*p*-value*
All nodules						
Kim's	92.7	80.9	59.6	97.3	0.868	
SRU	35.5	54.3	19.2	80.8	0.551	<0.001
AACE	79.2	90.8	72.3	93.5	0.85	0.179
Nodules ≥ 1 cm						
Kim's	89.2	85.2	53.2	97.7	0.872	
SRU	71.5	41.9	18.8	88.6	0.567	<0.001
AACE	74.1	94.4	71.3	95.1	0.842	0.117

PPV: positive predictive value; NPV: negative predictive value; AUC: area under the receiver operating characteristic curve (ROC); * *p*-value was from

ROC comparison with Kim's criteria.

When we compared the AUC of Kim's criteria with other criteria in nodules larger than or equal in size to 1 cm (Table 4), the AUC of Kim's criteria was 0.872, which was significantly higher than that of the SRU criteria (0.567, $p<0.001$). Kim's criteria appeared to be more accurate than the AACE criteria for nodules ≥ 1 cm (AUC 0.872 vs. 0.842, $p=0.117$), though this difference was not significant.

Three guidelines were compared to determine the number of FNAB that would have been performed and the number of cancers that would have been missed using these criteria (Table 5).

Table 5. Number of FNABs that would have been performed and number of thyroid carcinomas that would have been missed carcinomas according to each criteria

Criteria	All nodules (n=1398)		Nodules ≥1 cm (n=996)	
	Number of FNAB*	Missed carcinoma	Number of FNAB*	Missed carcinoma
Kim's	508 (36.3%)	24 (7.3%)	265 (26.6%)	17 (10.8%)
SRU	605 (43.8%)	211 (64.5%)	600 (60.2%)	45 (28.5%)
AACE	358 (25.6%)	68 (20.8%)	164 (16.5%)	41 (25.9%)

* The number of biopsies that would have been performed if the criteria were applied to the overall population of nodules.

For all nodules, the number of FNABs based on AACE criteria could be reduced to 25.6% (358/1398), which was significantly lower than those based on other criteria ($p < 0.001$). On the other hand, the number of missed cancer based on Kim's criteria (7.3%) was the lowest ($p < 0.001$). When we applied these criteria to thyroid nodules ≥1 cm, AACE criteria may lower the number of FNAB to 16.5% and Kim's criteria showed low risk of missed carcinoma of 10.8%.

IV. DISCUSSION

Detection of thyroid nodules has increased with the wide use of US, and it is important to exclude the presence of malignancy, which accounts for about 5% of all thyroid nodules, regardless of size [8]. Because of the high prevalence of thyroid nodules, it is essential to determine a reliable and cost-effective guideline for the management of thyroid nodules on the basis of their US features. In a multicenter retrospective study, spiculated margins, microcalcifications, marked hypoechogenicity, taller than wide shape, and macrocalcifications showed a significant association with malignancy [15]. These suspicious features had relatively high specificity (80.8-96.1%) and positive predictive value (64.8-81.3%) even though the sensitivity was low (9.7-48.3%). However, no single sonographic feature is accurate enough to recommend FNAB, therefore several guidelines with various combinations of suspicious US findings have been suggested [4, 7, 8].

In this study, AUC of Kim's and the AACE criteria were comparable (0.868 and 0.85, respectively), which was superior to that of the SRU criteria (0.551). Kim's criteria was originally suggested for recommending FNAB in nonpalpable solid nodules of the thyroid and revealed high sensitivity (93.8%) and NPV (95.9%). Tae et al. [16] applied the same FNAB criteria on palpable or nonpalpable thyroid nodules and also obtained a high NPV (97.8%) and diagnostic accuracy (86.5%). These results suggest that Kim's criteria may be applied not only to nonpalpable nodules but also palpable nodules [16].

SRU reported recommendations based on the size and US characteristics of thyroid nodules of 1 cm or more at the largest diameter. They did not recommend FNAB in a nodule less than 1 cm based on the assumption that undergoing thyroid surgery for microcarcinoma may not improve life expectancy, considering its indolent clinical course. However, the prevalence of thyroid cancer does not differ between nodules greater or smaller than 1 cm [5, 6].

Furthermore, microcarcinomas can also have aggressive courses, including extracapsular growth and nodal metastasis [5, 6, 17-21].

The overall malignancy rate in our study was 23.4% (327/1398), and this high rate of malignancy is due to the high prevalence of cancer in nodules smaller than 1 cm (169/402, 42%). This high rate of malignancy in nodules less than 1 cm is probably due to the fact that many small sized suspicious thyroid nodules were referred to our tertiary hospital for FNAB. In other words, patients with small nodules that had obviously benign US findings might have not been referred for FNAB. Others have also reported higher prevalence rates of malignancy in subcentimeter nodules compared to larger nodules [6, 22]. In this series, suspicious findings were more prevalent in nodules less than 1 cm, compared with nodules larger than 1 cm (Table 3). This observation suggests that the size of nodules may be a poor indicator for predicting malignancy.

SRU also suggested that FNAB should be considered for nodules demonstrating substantial growth on serial US. However, only two among 36 nodules that grew more than 3 mm proved to be malignant and consequently had a low positive predictive value (5.6%). Our findings support previous study results that suggest the presence or absence of growth is not a reliable marker of a nodule's malignant or benign nature [8, 9, 23-26].

According to the AACE criteria, additional suspicious US features includes intranodular vascular spots. However, we could not apply intranodal vascular spots, because our data did not contain any color Doppler findings. Color Doppler US is useful to evaluate intranodular and perinodular vascularity on thyroid US, but its ability to differentiate benign from malignant nodules is still the subject of debate [5, 27-30]. Several previous studies suggested that intranodular vascularity is a significant indicator of malignancy, and that the sensitivity of intranodular vascularity was relatively high, ranging from 66.7 to 91.7% [5, 28, 29]. However, in their studies, specificity values varied (range, 34.7-80.8%) and the positive predictive values were low (range, 23.2-34.5%), which might have

been due to the identification of intranodular vascularity in many benign nodules. Therefore, if we had included color Doppler sonography data, it is likely that it would have increased the false positive rate and consequently lowered the specificity and PPV of the AACE criteria, although sensitivity would have increased. Further study with data including the types of vascularity is needed to directly compare the diagnostic accuracy of these various criteria in the near future.

The American Thyroid Association (ATA) recommends routine FNAB in nodules ≥ 1 cm in diameter unless serum TSH is suppressed [9]. They mentioned that nodules smaller than 1 cm may require evaluation if there are suspicious US findings. However, since they did not define what the suspicious US findings are specifically, we could not include the ATA guideline in this analysis. If we had performed FNAB in nodules equal or larger than 1 cm based on ATA guidelines, 996 nodules would have underwent FNAB, for identifying 158 malignancies.

In a recent report, several guidelines including SRU, AACE, ATA, and routine observation were evaluated in terms of relative desirability [31]. They reported that all guidelines, with the exception of ATA's, revealed similar outcomes and FNAB based on ATA guidelines appearing to be the least desirable. However, their baseline model assumptions were based on likelihoods suggested in previous studies and they calculated the probability of every potential outcome statistically instead of estimating it in real patients. Therefore, their findings may not necessarily be applied to the general population. In contrast, we applied US criteria to a large set of real nodules in patients.

Our study has several limitations. First, although many patients had multiple nodules, we did not perform FNAB on all nodules, which might have caused selection bias. However, since we do not want to perform FNAB on all thyroid nodules detected on US in clinical practice, it may be inevitable limitation. Second, all sonographic evaluations were done by one experienced radiologist

and interobserver variation data was not included in this study. Moon et al. reported that reviews had fair to higher agreement for US characteristics of thyroid nodules (κ value > 0.5) [15]. A large-scale prospective study would be warranted to confirm these results in the near future. Lastly, we did not have any clinical information such as prior head and neck irradiation, family history of thyroid cancer, or serum TSH or thyroglobulin that might have affected the decision to perform FNAB. However, this may not necessarily affect the results, because all criteria were applied to the same nodules and under the same conditions in a manner blinded to clinical history.

V. CONCLUSION

In summary, Kim's criteria and AACE criteria are more accurate than SRU criteria. AACE's guideline is recommended to achieve high specificity, whereas Kim's criteria can be chosen for its high sensitivity.

REFERENCES

1. Hamberger B, Gharib H, Melton LJ, Goellner JR, Zinsmeister AR. Fine-needle aspiration biopsy of thyroid nodules. Impact on thyroid practice and cost of care. The American journal of medicine 1982;73(3):381-4.

2. Mittendorf EA, Tamarkin SW, McHenry CR. The results of ultrasound-guided fine-needle aspiration biopsy for evaluation of nodular thyroid disease. Surgery 2002;132(4):648-53.

3. Hegedüs L. Clinical practice. The thyroid nodule. The New England journal of medicine 2004;351(17):1764-71.

4. Kim EK, Park CS, Chung WY, Oh KK, Kim DI, Lee JT, et al. New sonographic criteria for recommending fine-needle aspiration biopsy of nonpalpable solid nodules of the thyroid. AJR, American journal of roentgenology 2002;178(3):687-91.

5. Papini E, Guglielmi R, Bianchini A, Crescenzi A, Taccogna S, Nardi F, et al. Risk of malignancy in nonpalpable thyroid nodules: predictive value of ultrasound and color-Doppler features. The Journal of clinical endocrinology & metabolism 2002;87(5):1941-6.
6. Nam-Goong IS, Kim HY, Gong G, Lee HK, Hong SJ, Kim WB, et al. Ultrasonography-guided fine-needle aspiration of thyroid incidentaloma: correlation with pathological findings. Clinical endocrinology 2004;60(1):21-8.
7. Frates MC, Benson CB, Charboneau JW, Cibas ES, Clark OH, Coleman BG, et al. Management of thyroid nodules detected at US: Society of Radiologists in Ultrasound consensus conference statement. Radiology 2005;237(3):794-800.
8. American Association of Clinical Endocrinologists and Associazione Medici Endocrinologi medical guidelines for clinical practice for the diagnosis and management of thyroid nodules. Endocr Pract 2006 Jan-Feb;12(1):63-102.
9. Cooper DS, Doherty GM, Haugen BR, Kloos RT, Lee SL, Mandel SJ, et al. Management guidelines for patients with thyroid nodules and differentiated thyroid cancer. Thyroid 2006;16(2):109-42.
10. Chan BK, Desser TS, McDougall IR, Weigel RJ, Jeffrey RB. Common and uncommon sonographic features of papillary thyroid carcinoma. Journal of ultrasound in medicine 2003;22(10):1083-90.
11. Peccin S, de Castsro JA, Furlanetto TW, Furtado AP, Brasil BA, Czepielewski MA. Ultrasonography: is it useful in the diagnosis of cancer in thyroid nodules? Journal of endocrinological investigation 2002;25(1):39-43.
12. Butros R, Boyvat F, Ozyer U, Bilezikci B, Arat Z, Aytekin C, et al. Management of infracentimetric thyroid nodules with respect to ultrasonographic features. European radiology 2007;17(5):1358-64.
13. Cappelli C, Castellano M, Pirola I, Cumetti D, Agosti B, Gandossi E, et al. The predictive value of ultrasound findings in the management of thyroid nodules. QJM 2007 Jan;100(1):29-35.
14. Alexander EK, Hurwitz S, Heering JP, Benson CB, Frates MC, Doubilet PM,

et al. Natural history of benign solid and cystic thyroid nodules. Annals of internal medicine 2003;138(4):315-8.

15. Moon WJ, Jung SL, Lee JH, Na DG, Baek JH, Lee YH, et al. Benign and malignant thyroid nodules: US differentiation--multicenter retrospective study. Radiology 2008;247(3):762-70.

16. Tae HJ, Lim DJ, Baek KH, Park WC, Lee YS, Choi JE, et al. Diagnostic value of ultrasonography to distinguish between benign and malignant lesions in the management of thyroid nodules. Thyroid 2007;17(5):461-6.

17. Mazzaferri EL, de los Santos ET, Rofagha-Keyhani S. Solitary thyroid nodule: diagnosis and management. The Medical clinics of North America 1988;72(5):1177-211.

18. Noguchi S, Yamashita H, Murakami N, Nakayama I, Toda M, Kawamoto H. Small carcinomas of the thyroid. A long-term follow-up of 867 patients. Archives of surgery 1996;131(2):187-91.

19. Berker D, Aydin Y, Ustun I, Gul K, Tutuncu Y, Işık S, et al. The Value of Fine-Needle Aspiration Biopsy in Subcentimeter Thyroid Nodules. Thyroid 2008;18(6):603-8.

20. Wada N, Duh QY, Sugino K, Iwasaki H, Kameyama K, Mimura T, et al. Lymph node metastasis from 259 papillary thyroid microcarcinomas: frequency, pattern of occurrence and recurrence, and optimal strategy for neck dissection. Annals of surgery 2003;237(3):399-407.

21. Cappelli C, Castellano M, Braga M, Gandossi E, Pirola I, De Martino E, et al. Aggressiveness and outcome of papillary thyroid carcinoma (PTC) versus microcarcinoma (PMC): a mono-institutional experience. Journal of surgical oncology 2007;95(7):555-60.

22. Sahin M, Sengul A, Berki Z, Tutuncu NB, Guvener ND. Ultrasound-guided fine-needle aspiration biopsy and ultrasonographic features of infracentimetric nodules in patients with nodular goiter: correlation with pathological findings. Endocr Pathol 2006 Spring;17(1):67-74.

23. Filetti S, Durante C, Torlontano M. Nonsurgical approaches to the management of thyroid nodules. Nature clinical practice endocrinology & metabolism 2006;2(7):384-94.
24. Quadbeck B, Pruellage J, Roggenbuck U, Hirche H, Janssen OE, Mann K, et al. Long-term follow-up of thyroid nodule growth. Experimental and clinical endocrinology & diabetes 2002;110(7):348-54.
25. Alexander EK, Marqusee E, Orcutt J, Benson CB, Frates MC, Doubilet PM, et al. Thyroid nodule shape and prediction of malignancy. Thyroid 2004;14(11):953-8.
26. Roti E, degli Uberti EC, Bondanelli M, Braverman LE. Thyroid papillary microcarcinoma: a descriptive and meta-analysis study. Eur J Endocrinol 2008 Dec;159(6):659-73.
27. Shimamoto K, Endo T, Ishigaki T, Sakuma S, Makino N. Thyroid nodules: evaluation with color Doppler ultrasonography. Journal of ultrasound in medicine 1993;12(11):673-8.
28. Rago T, Vitti P, Chiovato L, Mazzeo S, De Liperi A, Miccoli P, et al. Role of conventional ultrasonography and color flow-doppler sonography in predicting malignancy in 'cold' thyroid nodules. European journal of endocrinology 1998;138(1):41-6.
29. Appetecchia M, Solivetti FM. The association of colour flow Doppler sonography and conventional ultrasonography improves the diagnosis of thyroid carcinoma. Hormone research 2006;66(5):249-56.
30. Frates MC, Benson CB, Doubilet PM, Cibas ES, Marqusee E. Can color Doppler sonography aid in the prediction of malignancy of thyroid nodules? Journal of ultrasound in medicine 2003;22(2):127-31.
31. McCartney CR, Stukenborg GJ. Decision analysis of discordant thyroid nodule biopsy guideline criteria. J Clin Endocrinol Metab 2008;93(8):3037-44.

<ABSTRACT (IN KOREAN)>

초음파에서 보이는 갑상선 결절 중 세침흡인 세포검사를 시행해야 하는 적응증: 기존의 세 가지 지침 비교

<지도교수 김은경>

연세대학교 대학원 의학과

안성수

목적: 갑상선 결절에서 세침흡인 세포검사를 시행하기 위한 세 가지 지침을 비교하고자 함.

재료 및 방법: 세침흡인 세포검사 또는 수술로 확진 된 1398 개의 갑상선 결절을 대상으로 다음 세 가지 지침을 비교하였다. 1) Kim's: 현저한 저에코, 불규칙한 경계, 미세 석회화, 전후 직경/가로 직경 ≥ 1 중 적어도 한 가지를 가지는 결절; 2) the Society of Radiologists in Ultrasound (SRU): 1 cm 이상의 미세 석회화를 포함한 결절, 1.5 cm 이상의 고형 결절 또는 거친 석회화를 가진 결절, 2 cm 이상의 고형과 낭성의 혼합성 결절, 점진적으로 자라는 결절, 비정상적인 경부 림프절; 3) American Association of Clinical Endocrinologists (AACE): 불규칙한 경계, 전후 직경/가로 직경 ≥ 1 , 미세 석회화 중 한 가지 이상을 가지는 저에코성 결절에서 세침흡인 세포검사를 시행하도록 하였다. 이 세 가지 지침의 진단적 가치를 ROC분석을 통해 AUC값을 비교하였다.

결과: 모든 결절에서 Kim's criteria (0.868)와 AACE (0.85)의 AUC값은 SRU (0.551)에 비해 높았다. 세침흡인 세포검사의 빈도는 AACE 지침이 가장 낮았으며 (25.6%), 악성 결절을 놓치는 빈도는 Kim's 지침이 가장 낮았다 (7.3%). 1 cm 이상의

결절에서도 동일한 결과를 얻을 수 있었다.

결론: Kim’s와 AACE 지침이 SRU 지침보다 정확하며 AACE 지침은 높은 특이도를 보이고 Kim’s 지침은 높은 민감도를 보인다.

핵심되는 말 : 갑상선, 초음파, 세침흡인 세포검사