Ann Surg Oncol (2015) 22:152–157
DOI 10.1245/s10434-014-3952-7

Annals of
SURGICAL ONCOLOGY
OFFICIAL JOURNAL OF THE SOCIETY OF SURGICAL ONCOLOGY

ORIGINAL ARTICLE – ENDOCRINE TUMORS

False Negative Cytology in Large Thyroid Nodules

Wesley H. Giles, MD[1], Reid A. Maclellan, MD, MMSc[2], Atul A. Gawande, MD, MPH[1], Daniel T. Ruan, MD[1], Erik K. Alexander, MD[3], Francis D. Moore Jr., MD[1], and Nancy L. Cho, MD[1]

[1]Department of Surgery, Brigham and Women's Hospital, Boston, MA ; [2]Department of Surgery, Children's Hospital, Boston, MA ; [3]Division of Endocrinology, Diabetes, and Hypertension, Thyroid Unit, Department of Medicine, Brigham and Women's Hospital, Boston, MA

ABSTRACT

Background. Controversy exists regarding the accuracy of fine-needle aspiration (FNA) in large thyroid nodules. Recent surgical series have documented false-negative rates ranging from 0.7 to 13 %. We examined the accuracy of benign FNA cytology in patients with thyroid nodules ≥3 cm who underwent surgical resection and identified features characteristic of false-negative results.

Methods. We retrospectively studied all thyroidectomy specimens between January 2009 and October 2011 and identified nodules ≥3 cm with corresponding benign preoperative FNA cytology. We collected clinical information regarding patient demographics, nodule size, symptoms, sonographic features, FNA results, and final surgical pathology. For comparison, we analyzed nodules <3 cm from this cohort also with benign FNA cytology.

Results. A total of 323 nodules with benign preoperative cytology were identified. Eighty-three nodules were <3 cm, 94 nodules were 3–3.9 cm, and 146 nodules were ≥4 cm in size. The false-negative rate was 11.7 % for all nodules ≥3 cm and 4.8 % for nodules <3 cm ($p = 0.03$). Subgroup analysis of nodules ≥3 cm revealed a false-negative rate of 12.8 % for nodules 3–3.9 cm and 11 % for nodules ≥4 cm. Age ≥55 years and asymptomatic clinical status were the only patient characteristics that reached statistical significance as risk factors. Final pathology of the false-negative specimens consisted mainly of follicular variant of papillary thyroid cancer and follicular thyroid cancer.

Conclusions. When referred for thyroidectomy, patients with large thyroid nodules demonstrate a modest, yet significant, false-negative rate despite initial benign aspiration cytology. Therefore, thyroid nodules ≥3 cm may be considered for removal even when referred with benign preoperative cytology.

Thyroid nodules are common in the general population. While the majority of nodules are benign, the incidence of thyroid cancer has risen dramatically in recent years. Approximately 56,000 new cases of thyroid cancer were diagnosed in the United States in 2012.[1] Interestingly, this trend appears consistent across thyroid cancers of all sizes, suggesting that increased detection of smaller malignant nodules through ultrasonography or cross-sectional imaging is not the only reason for the rising cancer incidence.[2] Careful evaluation and surveillance are necessary to ensure accurate assessment of cancer risk in patients with thyroid nodules.

Fine-needle aspiration (FNA) is the "gold standard" for the preoperative diagnosis of thyroid nodules. FNA is commonly performed by multiple specialties and carries an overall false-negative rate of less than 5 %.[3] While FNA is widely accepted to be safe, rapid, and cost-effective, it is limited by the skill of the aspirator to obtain an appropriate tissue sample and the expertise of the cytologist to differentiate benign from malignant disease. Ultrasound guidance decreases nondiagnostic cytology and false-negative rates; however, controversy exists regarding the accuracy of FNA in large thyroid nodules.[4]

Prior reports have observed a high false-negative rate for FNA cytology in large thyroid nodules.[5–7] More recent surgical series have documented contradictory false-negative rates ranging from 0.7 % ($N = 145$, nodules ≥3 cm) to 13 % ($N = 71$, nodules ≥4 cm) in large nodules.[8,9] However, many of these studies only considered FNA cytology in patients who underwent surgery, thus reporting

First Received: 18 May 2014;
Published Online: 30 July 2014

N. L. Cho, MD
e-mail: nlcho@partners.org

a biased false-negative rate, because those patients who did not proceed to surgery were excluded from analysis. At our own institution, the false-negative rate for benign FNA cytology was recently reported as 1.1 % (79/7,348 nodules), although the authors are careful to note that this calculation is representative of FNA cytology as opposed to final histology from surgically resected nodules.[10]

In our practice, we routinely offer thyroid resection to patients with nodules ≥3 cm given our observation that large thyroid nodules carry a high false-negative cytology rate. In order to circumvent the selection bias observed in other studies, we examined the accuracy of benign FNA cytology in a large cohort of patients with thyroid nodules ≥3 cm who underwent surgery, regardless of FNA results. We hypothesized that the false-negative rate for FNA cytology is increased in nodules ≥3 cm and determined histopathologic, sonographic, and patient features that were characteristic of nodules with false-negative results.

METHODS

Under a protocol approved by the institutional review board at the Brigham and Women's Hospital (BWH), we collected clinical data on patients with thyroid nodules ≥3 cm with benign FNA cytology and subsequent thyroid surgery at our institution between January 2009 and October 2011. Within this cohort, additional nodules <3 cm with benign FNA cytology that were identified in the thyroidectomy specimen also were included for comparison. Clinical data included patient demographics, presence of symptoms, ultrasound characteristics, FNA cytology, nodule size, presence of chronic, lymphocytic thyroiditis, and final pathology results. Patients were placed in the symptomatic group if they presented with cervical dysphagia, neck tenderness, dyspnea, odynophagia, or laboratory evidence of hyperthyroidism.

Preoperative ultrasound was performed by a radiologist using a 5- to 17-MHz transducer. Nodules were classified as solid or cystic. Cystic lesions were subclassified as <25 % cystic, 25–50 % cystic, 50–75 % cystic, or >75 % cystic. Nodule size was reported in three dimensions: length, width, and depth. FNA was performed by an endocrinologist with guidance by a radiologist in our Thyroid Nodule Clinic. Ultrasound was used to guide the needle tip into the nodule and ensure an accurate specimen. All aspirations were processed using the Thin-Prep technique (Cytyc Corp., Boxborough, MA, USA). Three (or rarely four) needle sticks were obtained per nodule, constituting a single aspiration analysis. Each specimen was read by an experienced cytopathologist trained in thyroid cytopathology. Only nodules with benign cytology (6 or more groups of 10 or more benign appearing follicular cells) were included in the study. Most patients with thyroid nodules ≥4 cm in size were referred to Endocrine Surgery for consultation. In addition, patients with nodules ≥3 cm in size were referred to Surgery if continued growth or ongoing symptoms were documented.

Thyroidectomy was performed by one of four dedicated endocrine surgeons (F.D.M., A.A.G., D.T.R., and N.L.C.). All specimens were submitted to Pathology for permanent section with stitches to mark the side and upper pole. The final pathology report was then compared to the preoperative ultrasound and cytology results to determine the false-negative rate. Only nodules in which FNA cytology, ultrasound, and surgical histopathology could be accurately verified were included. False negatives were defined as nodules with benign FNA cytology but malignant final histopathology. Microcarcinomas (mPTC) incidentally found outside the nodule were considered true negatives.

Statistical data are presented as proportion or median with range. Nodules were arranged into two groups: ≥3 cm and <3 cm and characterized by ultrasound as mostly solid or cystic. Thyroid nodule size, sex, ultrasound characteristics, symptoms, and age were compared using Chi square or Fisher's exact test, as indicated. Patient age was analyzed using the Wilcoxon rank-sum test. Two-tailed values of $p \leq 0.05$ were considered statistically significant. Analysis was conducted using the SAS software package (version 9.3; SAS Institute, Inc., Cary, NC, USA).

RESULTS

During the study period, 420 nodules ≥3 cm were excised. Within this cohort, an additional 83 nodules

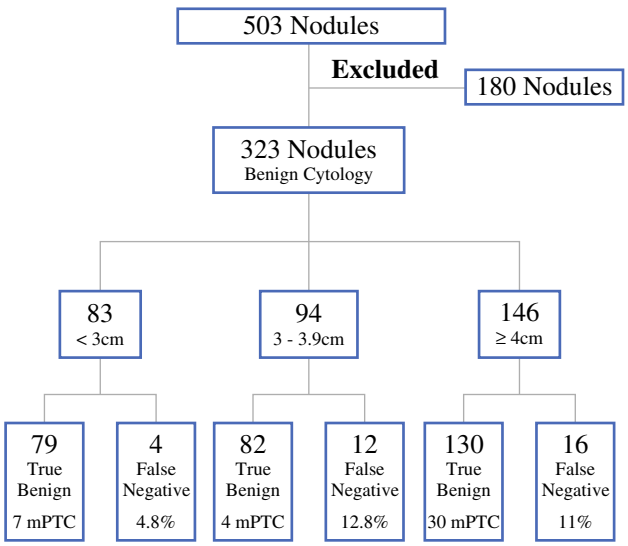

FIG. 1 Study algorithm. Three hundred twenty-three nodules with benign preoperative FNA cytology were included for analysis. A false-negative rate of 4.8 % was found for thyroid nodules <3 cm compared with 11.7 % for thyroid nodules ≥3 cm ($p = 0.03$)

TABLE 1 False-negative FNA cytology rates by nodule size

Nodule size	*N*	True benign	False negative	False-negative rate (%)	*p*
≥2 cm	278	248	30	10.8	0.08
≥2.5 cm	258	230	28	10.9	0.26
≥3 cm	241	213	28	11.6	**0.03**
≥3.5 cm	191	167	24	12.6	**0.05**

Statistically significant values $p \leq 0.05$ are given in bold

p values were calculated by comparing the nodule at each designated size to the rest of the cohort

<3 cm were identified in the thyroidectomy specimen allowing comparison of 503 nodules. Only nodules with benign preoperative cytology were eligible for the study. Therefore, 180 nodules were excluded, leaving 323 benign nodules which formed our study cohort (Fig. 1). Of these 323 nodules, 146 measured ≥4 cm (45 %), 94 measured 3–3.9 cm (29 %), and 83 measured <3 cm (26 %). The mean size of nodules ≥3 cm was 4.7 cm (range 3.0–9.8). The mean size of nodules <3 cm was 2 cm (range 1.0–2.9).

Subgroup analysis of nodules ≥3 cm revealed a false-negative rate of 12.8 % for nodules 3–3.9 cm and 11 % for nodules ≥4 cm (Fig. 1). The false-negative rate of nodules <3 cm was 4.8 %. Statistical significance was not obtained until nodules reached 3 cm in size ($p = 0.03$; Table 1). Taken together, the false-negative rate was 11.7 % for all nodules ≥3 cm and 4.8 % for nodules <3 cm ($p = 0.03$). Of note, an additional 7 incidental micropapillary thyroid cancers (mPTC) were identified in patients with nodules <3 cm, 4 incidental mPTC in nodules 3–3.9 cm, and 30 incidental mPTC in nodules ≥4 cm; however, these were considered true negatives due to their location outside the nodule of interest.

We next used patient, pathology, and ultrasound characteristics to identify risk factors for false-negative cytology in nodules ≥3 cm (Table 2). Patients ≥55 years old had a false-negative rate of 16.7 % compared with 8 % in patients <55 years old, reaching marginal significance ($p = 0.05$). Asymptomatic patients also had a similar false-negative rate of 16.5 % compared with 7.6 % in symptomatic patients ($p = 0.05$). The false-negative rate in males was 17 versus 10 % in females ($p = 0.08$). We observed no difference between races when controlling for size ($p = 0.22$); however, the lowest false-negative rate was seen in African American patients (2.5 %). We found no difference in false-negative rates in patients with Hashimoto's thyroiditis confirmed on final pathology (12.5 %) compared with those without the disease (11.4 %; $p = 0.82$). No difference in false-negative rates was observed in solid nodules (11.6 %) compared with cystic nodules (10.1 %; $p = 0.79$).

TABLE 2 Characteristics of true benign versus false-negative nodules

	True benign	False negative	False-negative rate (%)	*p*
Nodule size				**0.03**
<3 cm	83	4	4.8	
≥3 cm	212	28	11.7	
Age				**0.05**
<55 years	58	7	8.0	
≥55 years	85	17	16.7	
Gender				0.08
Female	174	20	10.0	
Male	38	8	17.0	
Race				0.22
White	155	23	12.9	
Black	39	1	2.5	
Asian	5	1	16.7	
Hispanic	13	3	18.8	
Hashimoto's				0.82
Yes	49	7	12.5	
No	163	21	11.4	
Symptomatic				**0.05**
Yes	121	10	7.6	
No	91	18	16.5	
US characteristics				0.79
Mostly solid	38	5	11.6	
Mostly cystic	89	10	10.1	

Statistically significant values $p \leq 0.05$ are given in bold

The final pathology of nodules with false-negative cytology is depicted in Fig. 2. Follicular variant of papillary thyroid cancer (FvPTC) was diagnosed in 57 % (16/28) of false-negative nodules ≥3 cm and 100 % (4/4) of nodules <3 cm. Follicular thyroid carcinoma (FTC) was the next most common finding in nodules ≥3 cm at 25 % (7/28) followed by Hürthle cell cancer, oxyphilic variant of papillary thyroid cancer (PTC), tall cell variant of PTC, macrofollicular variant of PTC, and poorly differentiated thyroid cancer, all at 3.6 % (1/28).

DISCUSSION

FNA is the accepted standard for preoperative assessment of malignancy in clinical relevant (>1 cm) thyroid nodules. A low false-negative rate enables physicians to use FNA as a reliable tool for guiding operative decision making. However, several studies have demonstrated conflicting data regarding the accuracy of benign cytology in large thyroid nodules. Meko et al. reviewed records from 90 patients with preoperative FNA and thyroid surgery and found a false-negative rate of 17 % (5/30) in nodules

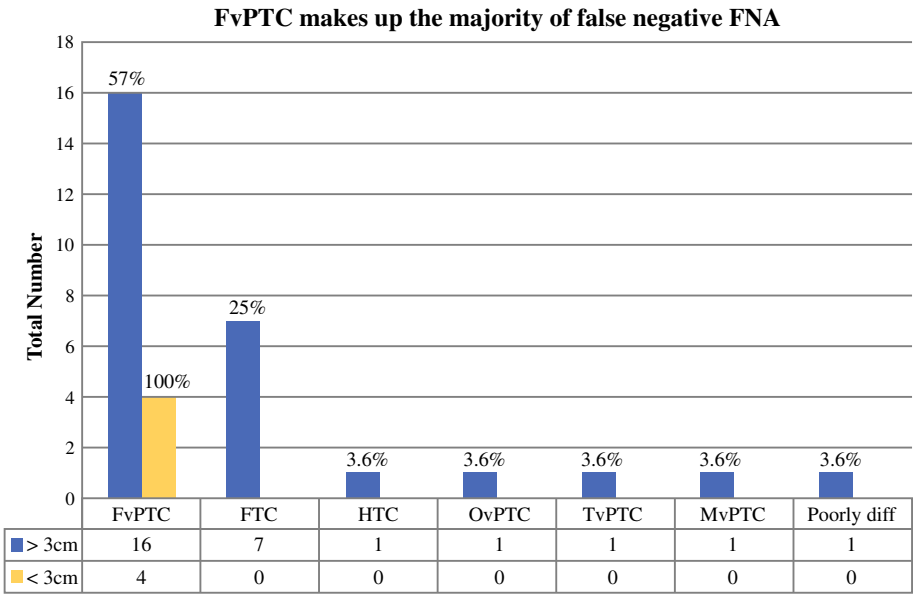

FIG. 2 Final pathology results of nodules with false-negative cytology. The majority of false-negative nodules demonstrated follicular variant of papillary thyroid cancer (FvPTC) or follicular cancer (FTC) on final pathology. Hürthle cell thyroid cancer (HTC), oxyphilic variant of PTC (OvPTC), tall cell variant of PTC (TvPTC), macrofollicular variant of PTC (MvPTC), and poorly differentiated thyroid cancer (poorly diff) comprised the remainder of the false-negative nodules. Follicular variant of PTC was the only pathology result identified in false-negative nodules <3 cm

≥3 cm.[6] McCoy et al.[8] examined 223 patients with nodules ≥4 cm with preoperative FNA and subsequent thyroidectomy. They identified a false-negative rate of 13 % (9/71) as well as higher incidence of cancer in larger nodules. Both studies recommended that large thyroid nodules should be considered for diagnostic lobectomy, regardless of FNA results. In contrast, Porterfield et al.[9] reviewed data from 743 ultrasound-guided FNA specimens with benign cytology. Twenty percent (145/743) underwent thyroidectomy and only one false-negative result (0.7 %) was identified on final pathology. They concluded that with appropriate aspiration and expert cytopathologic interpretation, the false-negative rate of FNA is extremely low and that diagnostic resection is unnecessary. However, it is important to note that a large proportion of patients with benign FNA cytology did not undergo surgery, thereby potentially decreasing the false-negative rate by a significant margin.

Recently, our institution began routinely offering thyroidectomy to patients referred for surgery with nodules ≥3 cm. This was based on suggestion of a higher than expected false-negative rate observed in such patients. Because all referred patients with large nodules were offered surgery regardless of FNA results, we attempted to avoid the selection bias discussed by Tee et al.[7] by including patients with benign FNA cytology who may otherwise have been managed conservatively. In our study, we identified a false-negative rate of 11.7 % (28/240) for nodules ≥3 cm and 4.8 % (4/83) for nodules <3 cm ($p = 0.03$). We consider this to be a significant finding, because it is the largest study to date to examine nodules ≥ 3 cm with both benign preoperative FNA and final surgical pathology for confirmation.

In 2013, Kamran et al.[10] analyzed more than 7,000 thyroid nodules and showed that a size of 2 cm was the threshold at which a significant increase in cancer was detected. Given that the majority of thyroid cancers are indolent processes, missing the diagnosis of a small, early cancer is unlikely to be a life-threatening situation. However, our results show that cancers are more frequently missed in larger nodules, which may delay diagnosis and treatment of a very curable disease and lead to cancer upstaging. In all major thyroid cancer scoring systems, tumor size is used to determine prognosis and survival; for example, the TNM system uses size to determine the pathologic stage.[11–13] Whereas 5-year survival for patients with Stage I or II well-differentiated carcinoma approaches 100 %, this number decreases to 71–93 % for such cancers when Stage III. In the absence of nodal or distant metastasis, the difference between Stage II and III disease rests solely in tumor size (>4 cm). In addition to decreased survival, tumor size correlates with higher incidence of cervical nodal metastasis.[14–16] Podnos et al.[17] reviewed 9,904 patients in the Surveillance, Epidemiology, and End Results database and found that large tumor size in association with lymph node metastases predicted poorer outcome. Taken together, we recommend that clinicians maintain a healthy index of suspicion for thyroid nodules ≥3 cm as false-negative cytology may result in cancer upstaging, with associated increased risk.

Tumors with follicular elements, namely FvPTC and follicular thyroid cancer, comprised the majority of false-negative specimens. Interestingly, FvPTC accounted for 57 % (16/28) of the false negatives in nodules ≥3 cm and 100 % (4/4) of nodules <3 cm. These tumors are inherently more difficult to diagnose with FNA. Whereas PTC

has an FNA sensitivity of 60 % to >90 %, FvPTC has a reported FNA sensitivity of 25–42 %.[18–20] FvPTC lacks many of the cytologic criteria, such as fine chromatin, nuclear grooves, overlapping nuclei, and intranuclear inclusions, necessary for diagnosis. FvPTC also demonstrates significant histopathologic overlap with benign lesions, such as nodular hyperplasia and follicular adenoma, further confounding diagnosis. More accurate tools are needed to improve the cytologic diagnosis of FvPTC and optimize patient selection for surgery.

Classical PTC and FvPTC demonstrate similar survival curves; however, advanced age, tumor stage, and male gender independently correlate with reduced overall survival.[21] We also found that these risk factors (age ≥55, tumor size ≥3 cm, and male gender) corresponded with increased rates of false-negative FNA results, so appropriate monitoring is critical in these patients. Of note, although we acknowledge that the survival rate for differentiated thyroid cancer is high, we identified aggressive cancers in nearly half of the false-negative nodules ≥3 cm (Fig. 2). Once nodules reach a certain size, it may be difficult to accurately sample the entire nodule; thus, a false-negative result may be purely a sampling error rather than a diagnostic dilemma. Furthermore, thyroids with large nodules may harbor separate, incidental microcarcinomas of unknown malignant potential (Fig. 1). Although we defined these incidental lesions as true negatives, the increased presence of microcarcinomas found in conjunction with large nodules suggests a microenvironment that supports proliferation, inflammation, and tumorigenesis. Overall, our results demonstrate that diagnostic lobectomy should be considered for thyroid nodules ≥3 cm.

Our study is limited by its retrospective design. We correlated each FNA result with the final pathology specimen by reviewing the cytology and pathology reports in conjunction with preoperative imaging. To minimize error, we excluded any specimen where an obvious connection between the two reports was absent. Another potential limitation is the referral bias to our surgical clinics. Many of our patients are referred by a group of in-house endocrinologists. A 2003 paper from this group observed that most solid nodules with benign FNA cytology grow over time and that increasing size does not necessarily indicate a false-negative FNA or malignant transformation.[22] Despite this finding, many patients referred to surgery by this group of endocrinologists were asymptomatic, suggesting that clinical suspicion played an important role in the physician's decision to recommend surgery. Subtle findings on ultrasound, clinical experience, or an overall *gestalt*—difficult to identify in the medical record—could impact this decision process. This referral pattern could influence a higher false-negative rate in this cohort as it represents a group of patients deemed worrisome and thus selected for referral by experienced endocrinologists. Our findings may not be reproducible at other centers where surgical referrals are made by physicians without significant experience in managing thyroid nodules.

CONCLUSIONS

Benign FNA cytology for thyroid nodules ≥3 cm has a modest, but significant, false-negative rate of 11.7 %. FvPTC is the most common missed cancer on final pathology; however, we also identified a number of more aggressive cancers and separate, incidental microcarcinomas in our cohort. Missed or delayed diagnosis can lead to cancer upstaging and poorer outcomes. As such, surgical resection should be considered for nodules ≥3 cm regardless of benign FNA results, especially when patients are referred by experienced endocrinologists. Future studies should focus on improving the cytologic diagnosis of FvPTC.

REFERENCES

1. Siegel R, Naishadham D, Jemal A. Cancer statistics, 2012. *CA Cancer J Clin.* 2012;61(1):10–29.
2. Chen AY, Jemal A, Ward EM. Increasing incidence of differentiated thyroid cancer in the United States, 1988–2005. *Cancer.* 2009;115(16):3801–7.
3. Ogilvie JB, Piatigorsky EJ, Clark OH. Current status of fine needle aspiration for thyroid nodules. *Adv Surg.* 2006;40:223–38.
4. Danese D, Sciacchitano S, Farsetti A, Andreoli M, Pontecorvi A. Diagnostic accuracy of conventional versus sonography-guided fine-needle aspiration biopsy of thyroid nodules. *Thyroid.* 1998;8:15–21.
5. Gharib H, Goellner JR. Fine needle aspiration biopsy of the thyroid: an appraisal. *Ann Intern Med.* 1993;118:282–9.
6. Meko KB, Norton JA. Large cystic/solid thyroid nodules: a potential false-negative fine-needle aspiration. *Surgery.* 1995;118:996–1004.
7. Tee YY, Lowe AJ, Brand CA, Judson RT. Fine-needle aspiration may miss a third of all malignancy in palpable thyroid nodules: a comprehensive literature review. *Ann Surg.* 2007;246(5):714–20.
8. McCoy, K. The incidence of cancer and rate of false negative cytology in thyroid nodules greater than or equal to 4 cm in size. Surgery. 2007.
9. Porterfield JR. Reliability of benign fine needle aspiration cytology of large thyroid nodules. *Surgery.* 2008;144(6):963–8.
10. Kamran SC, Marqusee E, Kim MI, Frates, MC, Ritner J, Peters H, Benson CB, Doubilet PM, Cibas ES, Barletta J, Cho N, Gawande A, Ruan D, Moore FD Jr, Pou K, Larsen PR, Alexander EK. Thyroid nodule size and prediction of cancer. *J Clin Endocrinol Metab.* 2013;98(2):564–70.
11. Hay ID, Grant CS, Taylor WF, et al. Ipsilateral lobectomy versus bilateral lobar resection in papillary thyroid carcinoma: a retrospective analysis of surgical outcome using a novel prognostic scoring system. *Surgery.* 1987;102:1088–95.
12. Cady B, Rossi R. An expanded view of risk: group definition in differentiated thyroid carcinoma. *Surgery.* 1988;104:947–53.
13. Hay ID, Bergstralh EJ, Goellner JR, et al. Predicting outcome in papillary thyroid carcinoma: development of a reliable prognostic scoring system in a cohort of 1779 patients surgically treated at

one institution during 1940 through 1989. *Surgery.* 1993;114:1050–8.
14. Lim YS, Lee JC, Lee YS, Lee BJ, Wang SG, Son SM, Kim IJ. Lateral cervical lymph node metastases from papillary thyroid carcinoma: predictive factors of nodal metastasis. *Surgery.* 2011;150(1):116–21.
15. Konturek A, Barczynski M, Nowak W, Richter P. Prognostic factors in differentiated thyroid cancer: a 20-year surgical outcome study. *Langenbecks Arch Surg.* 2012;397(5):809–15.
16. Gomez NR, Kouniavsky G, Tsai HL, Somervell H, Pai SI, Tufano RP, Umbricht C, Kowalksi J, Dackiw APB, Zeiger MA. Tumor size and presence of calcifications on ultrasonography are preoperative predictors of lymph node metastases in patients with papillary thyroid cancer. *J Surg Oncol.* 2011;104:613–6.
17. Podnos YD, Smith D, Wagman LD, et al. The implication of lymph node metastasis on survival in patients with well-differentiated thyroid cancer. *Am Surg.* 2005;71:731–4.
18. Yoon JH, Kim EK, Hong SW, Kwak JY, Kim MJ. Sonographic features of the follicular variant of papillary thyroid carcinoma. *J Ultrasound Med.* 2008;27:1431–7.
19. Shih SR, Shun CT, Su DH, Hsiao YL, Chang TC. Follicular variant of papillary thyroid carcinoma: Diagnostic limitations of fine needle aspiration cytology. *Acta Cytol.* 2005;49:383–6.
20. Kurian EM, Dawlett M, Wang J, Gong Y, Guo M. The triage efficacy of fine needle aspiration biopsy for follicular Varian of papillary thyroid carcinoma using the Bethesda reporting guidelines. *Diag Cytopathol.* 2011;10:E69–73.
21. Lin HW, Bhattacharyya N. Clinical behavior of follicular variant of papillary thyroid carcinoma: presentation and survival. *Laryngoscope.* 2010;120:Suppl 4:S163.
22. Alexander EK, Hurwitz S, Heering BA, Benson CB, Frates MC, Doubilet PM, Cibas ES, Larsen PR, Marqusee E. Natural history of benign solid and cystic thyroid nodules. *Ann Intern Med.* 2003;138:315–8.