CANCER
CYTOPATHOLOGY

Long-term Assessment of a Multidisciplinary Approach to Thyroid Nodule Diagnostic Evaluation

Leila Yassa, MD[1]
Edmund S. Cibas, MD[2]
Carol B. Benson, MD[3]
Mary C. Frates, MD[3]
Peter M. Doubilet, MD, PhD[3]
Atul A. Gawande, MD, MPH[4]
Francis D. Moore Jr., MD[4]
Brian W. Kim, MD[1]
Vânia Nosé, MD[2]
Ellen Marqusee, MD[1]
P. Reed Larsen, MD[1]
Erik K. Alexander, MD[1]

[1] Department of Medicine, Brigham and Women's Hospital and Harvard Medical School, Boston, Massachusetts.

[2] Department of Pathology, Brigham and Women's Hospital and Harvard Medical School, Boston, Massachusetts.

[3] Department of Radiology, Brigham and Women's Hospital and Harvard Medical School, Boston, Massachusetts.

[4] Department of Surgery, Brigham and Women's Hospital and Harvard Medical School, Boston, Massachusetts.

Supported in part by a Thyroid, Head, and Neck Cancer Foundation Young Investigator Award to E.K.A.

We thank Dr. Robert Utiger for reviewing this article and Dr. Shilpa Rao for assistance with data collection.

Address for reprints: Erik K. Alexander, MD, Division of Endocrinology, Diabetes, and Hypertension; Brigham and Women's Hospital, 75 Francis Street, Boston, MA 02115; Fax: (617) 264-6346; E-mail: ekalexander@partners.org

Received March 28, 2007; revision received May 22, 2007; accepted June 18, 2007.

BACKGROUND. The diagnostic evaluation of patients with thyroid nodules is imprecise. Despite the benefits of fine-needle aspiration (FNA), most patients who are referred for surgery because of abnormal cytology prove to have benign disease. Recent technologic and procedural advances suggest that this shortcoming can be mitigated, although few data confirm this benefit in unselected patients.

METHODS. A total of 2587 sequential patients were evaluated by thyroid ultrasound and were offered ultrasound-guided FNA (UG-FNA) of all thyroid nodules that measured $\geq$1 cm during a 10-year period. Results of aspiration cytology were correlated with histologic findings. The prevalence of thyroid cancer in all patients and in those who underwent surgery was determined. Surgical risk was calculated.

RESULTS. Tumors that measured $\geq$1 cm were present in 14% of patients: Forty-three percent of patients had tumors that measured <2 cm in greatest dimension, and 93% had American Joint Committee on Cancer stage I or II disease. The cytologic diagnoses 'positive for malignancy' and 'no malignant cells' were 97% predictive and 99.7% predictive, respectively. Repeat FNA of initial insufficient aspirates, as well as more detailed classification of inconclusive aspirates, improved preoperative assessment of cancer risk and reduced surgical intervention. Fifty-six percent of patients who were referred for surgery because of abnormal cytology had cancer compared with from 10% to 45% of patients historically. An analysis of operative complications from a subset of 296 patients demonstrated a 1% risk of permanent surgical complications.

CONCLUSIONS. The current findings demonstrated the benefits of UG-FNA and of a more detailed classification of inconclusive aspirates in the preoperative risk assessment of thyroid nodules, supporting adherence to recently published guidelines. ***Cancer (Cancer Cytopathol)* 2007;111:508–16.** © *2007 American Cancer Society.*

KEYWORDS: thyroid, nodules, fine-needle aspiration, ultrasound, cytology, malignancy, neoplasm.

The diagnosis of thyroid nodular disease is increasingly common—a trend largely attributed to the advancing age of the population and the increased frequency of head, neck, and chest imaging procedures.[1] Although the majority of thyroid nodules are benign, an estimated 30,000 patients in the U.S. will be diagnosed with thyroid cancer this year, and an estimated 1500 patients will die of the disease.[2]

The primary objective of initial evaluation is to distinguish the majority of benign nodules from those that are cancer and require removal to limit morbidity and mortality. First introduced 40 years ago, fine-needle aspiration (FNA) has improved preoperative assess-

DOI 10.1002/cncr.23116
Published online 12 November 2007 in Wiley InterScience (www.interscience.wiley.com).

ment substantially, mostly through its high positive and negative predictive value.[3–5] However, FNA is limited by sampling difficulties, which result in "nondiagnostic" aspirates, and by the significant overlap in morphologic features between benign and malignant nodules, which can result in abnormal but not conclusive results.[6,7] During the last decade, several clinical advances have mitigated some of these limitations.

Thyroid ultrasonography identifies nonpalpable nodules for which aspiration may be indicated and defines the extent of nodules in multinodular glands.[8] Ultrasound-guided FNA has lowered the rates of "nondiagnostic" aspirations by allowing sampling of the cellular portions of predominantly cystic nodules. In addition, liquid-based processing of aspiration specimens improves the accuracy of cytology interpretation by eliminating obscuring blood and controlling the thickness of the preparation.[9–13] Finally, terminology for cytologic classification has been described that may differentiate more precisely between abnormal but not conclusive aspirates.[4] In combination, these advances would be expected to improve the diagnostic evaluation of thyroid nodules and reduce the need for surgery. However, to our knowledge, there are few data available demonstrating the impact of these techniques in an unselected population of patients with thyroid nodules.

Beginning in 1995, the Brigham and Women's Hospital began a multidiscipline, collaborative effort involving the departments of endocrinology, radiology, pathology, and surgery for the evaluation of patients with thyroid nodular disease. Primary care physicians were asked to refer patients with thyroid nodules or thyroid asymmetry and patients who had incidental nodules discovered at imaging. After documentation of normal thyroid function, ultrasonography was performed, and ultrasound-guided FNA (UG-FNA) of thyroid nodules that measured $\geq$1 cm in greatest dimension was offered during the same visit. ThinPrep slides were prepared and analyzed, and surgical consultation was arranged thereafter when indicated. In this report, we summarize our 10-year experience evaluating 2587 sequential patients.

MATERIALS AND METHODS

Since 1995, all primary care physicians in our hospital system have been asked to refer patients to the multispecialty Thyroid Nodule Clinic at Brigham and Women's Hospital upon clinical suspicion of thyroid nodular disease or incidental detection of a nodule by imaging. After documentation of normal thyroid function, all patients underwent ultrasound examination of the thyroid by a radiologist. Because sonographic findings are unable conclusively to distinguish benign from malignant disease, UG-FNA performed by an endocrinologist was offered when a nodule that measured $\geq$1 cm in greatest dimension was identified. Rarely, patients were offered FNA of nodules that measured <1 cm because of clinical or sonographic features suggestive of thyroid cancer, but these patients are not the subject of the current analysis. We reviewed the records of all patients who underwent evaluation and UG-FNA for 1 or more nodule(s) that measured $\geq$1 cm in greatest dimension between 1995 and 2004.

Thyroid ultrasonography was performed by 1 of 3 radiologists using a 5- to 15-megahertz transducer. The length, width, and depth of each nodule were reported, and each was classified subjectively as solid or as $\leq$25%, from >25% to 50%, from >50% to 75%, or >75% cystic. The location of each nodule was documented. Nodules were classified further as solitary (only 1 nodule $\geq$1 cm on ultrasound examination) or as part of a multinodular thyroid ($\geq$2 nodules each $\geq$1 cm on ultrasound examination). UG-FNA was performed by an endocrinologist and a radiologist after subcutaneous lidocaine administration. A 25-gauge needle was directed into the nodule by using ultrasound guidance, and the tip was moved in and out over a 5-mm to 10-mm path. This was repeated 3 or 4 times per nodule. For cystic nodules, grayscale and color-Doppler analysis were used to direct sampling of the solid portion of the nodule with evidence of tissue perfusion. If fluid was aspirated, then it was sent for separate cytologic evaluation.

All specimens were processed using the ThinPrep technique. Each needle was rinsed with CytoLyt solution (Cytyc Corporation, Marlborough, Mass) into a 50-mL conical tube and centrifuged at 1880 revolutions per minute for 10 minutes. If the cell pellet was bloody, then it was resuspended in CytoLyt and recentrifuged. Then, the pellet was transferred to a vial that contained PreservCyt solution (Cytyc Corporation), and 2 thin-layer slides were prepared with the ThinPrep 2000 (Cytyc Corporation). If sufficient cellular material was available, then a portion was retained and fixed in 10% formalin, embedded in paraffin, and processed for cell block sections. Both ThinPrep slides were stained with a modified Papanicolaou stain, and the cell block was stained with hematoxylin and eosin. All slides were interpreted by a cytopathologist at the Brigham and Women's Hospital.

FNA results were reported according to modification of the Mayo clinic terminology[4]: 1) insufficient for diagnosis (<6 groups of follicular cells, each containing >10 cells, without evidence of cellular

atypia), 2) benign (sufficient for diagnosis without atypical cells), 3) atypical cells of undetermined significance (rare and/or mildly abnormal cells, not definitively benign or suspicious for malignancy), 4) suspicious for a follicular or Hurthle cell neoplasm (a predominance of microfollicles, classically a ringlet of thyroid follicular cells; a predominance of crowded trabecular arrangements of follicular cells; or a specimen composed exclusively of Hurthle cells), 5) suspicious for papillary carcinoma (some, but not all, of the features of papillary carcinoma, including crowded cells; enlarged nuclei with pale, powdery chromatin; nuclear grooves, nuclear pseudoinclusions; distinct nucleoli; papillary structures; and psammoma bodies) or, 6) positive for malignancy (most of the above features of papillary carcinoma or other malignancy). Subsequent patient management was based on the cytology results. Patients who had insufficient samples or cytology with 'atypical cells of undetermined significance' were advised to return for repeat UG-FNA of the nodules in 2 to 6 months. Occasionally, repeat aspiration also was recommended in patients who had previously benign nodules that demonstrated substantial growth on follow-up assessment. For patients who underwent repeated biopsies, cytology results were recorded and labeled as the initial and repeat, or final (most recent), aspiration. Patients whose aspirates had cytology that was 'positive for malignancy' or 'suspicious for papillary carcinoma' usually were advised to undergo near-total thyroidectomy. Patients with cytology 'suspicious for a follicular or Hurthle cell neoplasm' were advised to undergo either lobectomy or thyroidectomy. Surgical pathology specimens were reviewed and interpreted by a staff pathologist according to widely accepted standards.[14] When available, the final classification of each nodule was based on histopathologic analysis of the surgical specimen. In nodules with a discrepancy between the FNA and histopathologic interpretations, the histopathologic interpretation was accepted as correct for the purposes of the current analysis without review of the original materials. Because the majority of patients with benign aspirates were not referred for surgery, the classification of their nodules was based only on cytology. A small number of cancers were diagnosed without histologic confirmation, for example, if the cytology demonstrated anaplastic thyroid cancer or nonthyroid cancer and the patient was not a candidate for surgery. The size of a malignant nodule was reported as the longest dimension of the gross pathology specimen. Rarely, a specimen measured <1 cm despite ultrasound measurement ≥ 1 cm, as described below. The extent of disease at diagnosis was classified by using the American Joint Committee on Cancer (AJCC) (sixth edition) and the Metastases, Age, Completeness of resection, Invasion, and Size (MACIS) staging systems.[15,16]

To obtain a representative sample of our current surgical practice, data detailing surgical complications were collected for all patients who underwent surgery between 2002 and 2004. Surgical and clinical records were reviewed, including follow-up records of all patients at least 1 month after surgery. Rates of permanent hypoparathyroidism and recurrent laryngeal nerve paralysis were recorded.

Findings were entered into a computerized database. Interim analyses of various aspects of these patients have been reported previously.[8,17–20] Permission from the Investigational Review Board of Brigham and Women's Hospital was obtained to perform these investigations.

RESULTS

Patient Population and Nodule Characteristics

From 1995 to 2004, 2587 patients who were referred to the Thyroid Nodule Clinic had 1 or more thyroid nodule(s) that measured ≥ 1 cm in greatest dimension identified and underwent UG-FNA (Table 1). Of the 2587 patients, 2268 (88%) were women, and the mean age was 50 years (range, 18–96 years).

In total, 4595 thyroid nodules that measured ≥ 1 cm in greatest dimension were detected in the 2587 patients. One thousand four hundred seventy-six patients (57%) had a solitary thyroid nodule, and 1111 patients (43%) had 2 or more nodules each that measured ≥ 1 cm in greatest dimension. The mean greatest nodule dimension was 2.2 cm (range, 1–9.4 cm), and 2438 nodules (53%) measured <2 cm in greatest dimension (Table 1). Of the 4595 nodules, 3589 nodules (78%) were aspirated. An additional 325 nodules did not undergo FNA but were evaluated histologically after the patients underwent thyroidectomy because of abnormal cytology in a separate nodule. In total, 3914 of 4595 nodules (85%) were evaluated. Approximately 200 of the remaining nodules (4%) were functional on thyroid scintigraphy, precluding the need for FNA.

FNA Cytology

Of 3589 *initial* aspirations, 3113 aspirates (87%) satisfied the criteria for adequacy and were classified into 1 of 5 cytologic categories (Table 2), whereas 476 aspirates (13%) were insufficient for diagnosis. Four hundred fifty-three aspirates (13%) were 'positive for malignancy' or 'suspicious for papillary carcinoma', and 2141 aspirates (60%) were benign. The remain-

TABLE 1
Characteristics of Patients and Nodules

Characteristic	No. of patients (%)
Patient demographics (N = 2587)	
Sex	
Women	2268 (88)
Men	319 (12)
Age, y	
Mean, y	50
Range, y	18–96
18–30	258 (10)
31–40	470 (18)
41–50	684 (26)
51–60	590 (23)
61–70	362 (14)
>70	223 (9)
No. of nodules per patient (≥1 cm in size)	
1	1476 (57)
2	592 (23)
3	277 (11)
≥4	242 (9)
Nodule characteristics (N = 4595)	
Nodule size, mm	
Mean	22.3
Range	10–94
10–14	1439 (31)
15–19	999 (22)
20–24	667 (15)
25–29	471 (10)
30–39	596 (13)
40–49	247 (5)
≥50	176 (4)
Cystic content, %	
0–25	3282 (71)
>25–75	727 (16)
>75	528 (12)
Unknown	58 (1)

TABLE 2
Cytologic Diagnoses in Initial and Final Aspirates From 3589 Nodules Measuring ≥1 cm in Greatest Dimension

	No. of aspirates (%)	
Cytology	Initial diagnosis	Final diagnosis
Positive for malignancy	168 (5)	173 (5)
Suspicious for papillary carcinoma	285 (8)	314 (9)
Suspicious for a follicular neoplasm	292 (8)	328 (9)
Atypical cells of undetermined significance	227 (6)	144 (4)
No malignant cells	2141 (60)	2361 (66)
Insufficient for diagnosis	476 (13)	269 (7)
Total	3589	3589

TABLE 3
Final Cytologic Diagnoses in 282 Nodules With Aspirates Initially Diagnosed as 'Insufficient for Diagnosis' That Underwent Repeat Aspiration(s)

Final cytologic diagnosis	No. of nodules (%)
Positive for malignancy	3 (1)
Suspicious for papillary carcinoma	14 (5)
Suspicious for a follicular neoplasm	16 (6)
Atypical cells of undetermined significance	14 (5)
No malignant cells	161 (57)
Insufficient for diagnosis	74 (26)
Total	282

TABLE 4
Final Cytologic Diagnoses in 120 Nodules With 'Atypical Cells of Undetermined Significance' on Initial Aspiration That Underwent Repeat Aspiration(s)

Final cytologic diagnosis	No. of nodules (%)
Positive for malignancy	2 (2)
Suspicious for papillary carcinoma	15 (13)
Suspicious for a follicular neoplasm	20 (17)
Atypical cells of undetermined significance	23 (19)
No malignant cells	59 (49)
Insufficient for diagnosis	1 (<1)
Total	120

ing aspirates were 'suspicious for a follicular neoplasm' (292 nodules; 8%) or revealed 'atypical cells of undetermined significance' (227 nodules; 6%).

Among the 476 nodules in which the aspirate was inadequate, 282 nodules (59%) were aspirated again. Reaspiration resulted in adequate samples in 208 of 282 nodules (74%) (Table 3). One aspiration of a nodule that initially was diagnosed as 'atypical cells of undetermined significance' also was 'insufficient for diagnosis' on repeat FNA. Thus, aspirates from only 269 of the 3589 nodules (7%) persistently were inadequate despite repeat aspiration. These 269 patients were referred for consideration of surgery, which was performed on 77 nodules (29%). The remaining patients either refused further intervention or were lost to follow-up.

The initial aspirates from 227 nodules revealed 'atypical cells of undetermined significance'. Of these, 120 nodules (53%) were aspirated again 1 or more times (Table 4). The repeat aspiration(s) resulted in a definitive diagnosis in 96 nodules (80%). The diagnosis was benign in 59 nodules (49%) and malignant in 2 nodules (2%), whereas only 23 nodules (20%) were classified as 'atypical' cytology on repeat aspiration.

There were no major complications after FNA. Specifically, there were no major bleeding events. No complications related to vascular, tracheal, or esoph-

TABLE 5
Correlation of Final Cytology and Nodule Histology From 3589 Nodules That Measured ≥1 cm in Greatest Dimension

	No. of nodules (%)			
Final cytology	Total	Resected	Malignant	Nodule histology, No.*
Positive for malignancy	173 (5)	156	152 (97)	PC ≥1 cm, 138; PC < 1 cm, 8; FC, 2; poorly diff, 1; anaplastic, 1; MTC, 2
Suspicious for papillary carcinoma	314 (9)	288	173 (60)	PC ≥1 cm, 149; PC < 1 cm, 22; poorly diff, 1; MTC, 1
Suspicious for follicular neoplasm	328 (9)	268	74 (28)	PC ≥1 cm, 37; PC < 1 cm, 4; FC, 27; poorly diff, 5; anaplastic, 1
Atypical cells of undetermined significance	144 (4)	84	20 (24)	PC ≥1 cm, 15; PC < 1 cm, 3; FC, 2
Subtotal: Abnormal cytology	959 (27)	796	419 (53)	
No malignant cells	2361 (66)	369	6† (0.3‡)	PC ≥1 cm, 5; FC, 2
Insufficient for diagnosis	269 (7)	77	8 (10)	PC ≥1 cm, 5; PC < 1 cm, 2; FC, 1
Total	3589	1242	433	

PC indicates papillary thyroid carcinoma; FC, follicular thyroid carcinoma; MTC, medullary thyroid carcinoma; Poorly diff, poorly differentiated thyroid carcinoma.

* Seven additional cancers were documented by fine-needle aspiration but were not resected based on the clinical scenario. These tumors, which include 1 anaplastic tumor, 1 MTC, and 5 nonthyroid tumors, were not included in this Table.

† Does not include 7 separate benign aspirates from multinodular glands in which a malignant tumor was present but for which insufficient data were available to definitively assign the malignancy to the aspirated nodule.

‡ This rate was determined by using the entire population of nodules with benign cytology (N = 2361).

ageal puncture were documented. Similarly, no soft tissue infection of the neck was identified. A few patients had self-limited bleeding into the nodule or surrounding soft tissues at the time of aspiration. Two patients had persistent discomfort for >4 weeks after aspiration and ultimately chose to undergo thyroidectomy for symptom relief.

Predictive Value of FNA Cytology

Seven hundred three patients (with 796 nodules) were referred for surgery because of abnormal FNA cytology. Three hundred ninety-one of those patients had cancer (in 419 separate nodules). This resulted in a 56% surgical yield for thyroid cancer *per patient* (defined as the number of patients with cancer divided by the number of patients who underwent surgery). Because some patients had a solitary cancer within a multinodular gland, the surgical yield *per nodule* was slightly lower (419 of 796 nodules; 53%) (Table 5).

The FNA interpretation 'positive for malignancy' was accurate in 152 of 156 nodules (97%) (Table 5). The 4 false-positive FNA diagnoses included 3 follicular adenomas and 1 hyalinizing trabecular adenoma. One of the follicular adenomas was reviewed by multiple histopathologists who disagreed on the final diagnosis. One observer interpreted this tumor as a follicular variant of papillary carcinoma, and the patient was treated on the basis of that interpretation. We classified this aspirate as false-positive, however, because the majority of observers (2 of 3) believed that the nodule was a follicular adenoma. A second follicular adenoma was associated with lymphocytic thyroiditis and nuclear features "suggestive but not sufficient for the diagnosis of papillary carcinoma."

The FNA interpretation 'no malignant cells' was accurate in 2355 of 2361 nodules (>99%). In total, 6 of 2361 nodules (0.3%) with initially benign aspirates were identified as malignant at some later time (Table 5). We recommend that all patients with benign aspirates undergo repeat clinical or sonographic evaluation in 1 year. The majority of these patients have been followed from 2 to 9 years without clinical or sonographic evidence of substantial nodule growth. However, 369 of 2361 nodules with a benign cytologic diagnosis ultimately were removed in 1 of several clinical scenarios: One hundred sixty-one nodules were resected as part of surgery for a separate nodule or parathyroid abnormality, 119 nodules were resected because their greatest dimension had exceeded 4 cm, 37 nodules were resected because of persistent local symptoms, 15 nodules were resected for substantial growth over time, and the remaining 37 nodules were resected for unique clinical indications or patient preferences. Of the 369 nodules in this highly selected subgroup, 6 were interpreted as cancer on histopathologic analysis; 2 were follicular carcinoma and 4 were papillary carcinomas (1 measured <1 cm in greatest dimension on histopathology). Of the 2 follicular carcinomas, 1 was a 2-cm Hurthle cell carcinoma with a single focus of minimal capsular penetration. The other was a bilateral and multifocal follicular-cell proliferation with pro-

minent signet ring cells that was interpreted as a "low-grade malignancy." The 4 papillary carcinomas measured 0.8 cm, 3 cm, 3.3 cm, and 3.9 cm in greatest dimension (3 follicular variants of papillary carcinoma and 1 unspecified). Among the nodules that had cytologic diagnoses of 'suspicious for papillary carcinoma,' 'suspicious for a follicular neoplasm,' and 'atypical cells of undetermined significance,' 173 of 288 nodules (60%), 74 of 268 nodules (28%), and 20 of 84 nodules (24%), respectively were diagnosed as cancers (Table 5).

Prevalence of Malignancy

In total, 373 of 2587 patients (14.4%) had thyroid cancer that measured ≥1 cm in greatest dimension, as determined by pathologic examination. Thirty-nine additional patients (1.5%) had a papillary carcinoma that measured <1 cm in greatest dimension on pathologic examination, although their nodules measured ≥1 cm on ultrasound studies. Among the 373 patients who had thyroid cancer that measured ≥1 cm, 322 patients (86%) had papillary carcinoma, and 32 patients (9%) had follicular carcinoma (Table 6). Five nonthyroid tumor metastases to the thyroid also were identified, including 2 renal cell carcinomas, 1 esophageal cancer, 1 Langerhan cell histiocytosis, and 1 chronic lymphocytic leukemia that involved the thyroid gland.

In 162 patients (43%) who had thyroid cancer, the cancer measured <2 cm in greatest dimension, and 329 patients (93%) were classified with AJCC stage I or II disease.[15] Similarly, 282 of 322 patients (88%) who had with papillary thyroid cancer had MACIS scores <6, and 95% of patients had MACIS scores <7 (Table 6).[16] All patients aged <45 years had AJCC stage I disease at diagnosis, and 99% had MACIS scores <6. Among patients aged ≥45 years, 74% had MACIS scores <6, 15% had MACIS scores from 6 to 6.99, 8% had MACIS scores from 7 to 7.99, and 4% had MACIS scores >8.

Surgical Risk

During the years 2002 through 2004, 313 patients underwent partial or near-total thyroidectomy. Complete adverse event data were available on 296 of 313 patients (95%). The remaining 17 procedures were performed outside the Brigham and Women's Hospital, and data were unavailable. Two of 296 patients had permanent hypoparathyroidism (0.7%), and 1 patient had permanent unilateral recurrent laryngeal nerve section (0.3%). The remaining patients (99%) had no permanent sequelae.

TABLE 6
Characteristics of 373 Patients With Cancer ≥1 cm in Greatest Dimension Detected Over a 10-year Period

Characteristic	No. of patients (%)
Primary histology (N = 373)	
Papillary carcinoma	322 (86)
Follicular carcinoma	32 (9)
Poorly differentiated	7 (2)
Anaplastic	3 (1)
Medullary carcinoma	4 (1)
Nonthyroid	5 (1)
Histopathologic size, cm	
Mean, cm	2.4
Range, cm	1–10
1–1.4	93 (25)
1.5–1.9	69 (18)
2–2.4	60 (16)
2.5–2.9	47 (13)
3–3.9	54 (14)
≥4	43 (12)
Unknown (not resected)	7
AJCC stage*	
I	253 (72)
II	76 (21)
III	21 (6)
IV	4 (1)
MACIS score†	
<6	282 (88)
6–6.99	25 (8)
7–7.99	11 (3)
≥8	4 (1)

AJCC indicates American Joint Committee on Cancer (see AJCC, 2002[15]); MACIS, Metastases, Age, Completeness of Resection, Invasion, and Size (see Hay et al., 1992[16]).

* AJCC stage was determined only for the 354 patients who had differentiated thyroid carcinoma.

† MACIS scores were calculated only for the 322 patients who had papillary carcinoma.

DISCUSSION

The current data describe our 10-year experience evaluating >2500 patients with 3589 thyroid nodules that measured ≥1 cm in greatest dimension who were evaluated by UG-FNA. This approach resulted in a low rate of inadequate aspirates (7%) and had high positive and negative predictive values (>97% and >99.7%, respectively). Differentiating abnormal but not conclusive aspirates into 3 categories ('atypical cells of undetermined significance,' 'suspicious for a follicular or Hurthle cell neoplasm,' or 'suspicious for papillary carcinoma') resulted in improved preoperative risk assessment, and the majority of patients (56%) who were referred for surgery because of abnormal cytology had thyroid cancer. The risks associated with diagnostic aspiration and surgical intervention were very small. Together, these findings provide extensive observational data demonstrating the benefits of current strategies compared with historic controls and support adherence to recently

published guidelines by the American and European Thyroid Associations.[21,22]

Fourteen percent of our cohort was diagnosed with thyroid cancer ≥1 cm. This rate was similar to what has been reported in some recent studies[23–25] but was more than double that reported in numerous earlier series and reviews.[5,26–28] We believe this is a more accurate assessment of the true extent of this disease attributable to several factors—broader use of head and neck imaging, which identifies a substantial population of nonpalpable (yet still potentially malignant) thyroid nodules; and increased use of ultrasound-guidance for FNA, which significantly improves the accurate preoperative determination of benign versus malignant nodules.[8,29,30] In support of this, nearly 50% of the thyroid cancers diagnosed in our study measured from 1 cm to 2 cm in greatest dimension, whereas an additional 39 cancers measured <10mm (despite measuring ≥1 cm on ultrasound). Early-stage disease, as determined with the AJCC and MACIS staging systems,[15,16] was identified in a greater proportion of our patients compared with historic series.[16,31,32]

We diagnosed cancer in 56% of the patients who underwent surgery, a rate that compared favorably to large historic series. In previous reports of patients who were evaluated by FNA, from 10% to 45% of patients who underwent surgery had cancer.[5,28,31–33] We attribute this improvement to several factors-an increased accuracy afforded by ultrasound-guidance during FNA,[29,30] a liquid-based cytology preparation method, and a more detailed classification of inconclusive or "indeterminate" aspirates, which historically comprised up to 25% of aspirates.[5] We believe that the unqualified "indeterminate" diagnosis may do patients a disservice. This term is ambiguous: Sometimes it means insufficient cellularity, and other times it suggests mild atypia. Others use this diagnosis to mean suspicion for carcinoma or concern for a follicular neoplasm. The diagnosis fails to stratify patients meaningfully into distinct categories associated with corresponding risk. We sought to differentiate these specimens better, thus allowing for more specific treatment recommendations. We have labeled abnormal but low-risk aspirates 'atypical cells of undetermined significance,' and we recommend repeat FNA of such nodules. Upon repeat aspiration, nearly half (49%) of all aspirates were benign, obviating the need for surgery. It appears clear that a significant reduction in unnecessary surgery can be achieved by making this cytologic distinction.

The accuracy of negative aspirates is difficult to validate, because most patients with benign aspirates do not undergo surgery. However, others have reported similar data describing false-negative cytologic diagnosis rates <1% using UG-FNA.[29,30,34,35] Patients with benign aspirates in our study have been followed for 2 to 9 years with serial clinical or sonographic examinations. Those whose nodules demonstrate substantial growth, persistent or worsening symptoms, or growth beyond 4 cm were considered for repeat aspiration or resection. Even among this highly selected subset of 192 patients (with 369 nodules), only 6 had thyroid cancer identified despite a benign cytologic diagnosis. No additional thyroid cancers have been detected among patients with FNA-proven benign disease during the last 10 years.

Although, to our knowledge, the long-term benefit of detecting small papillary cancers has not been studied prospectively, retrospective analysis suggests that effective detection and treatment of thyroid cancer before it reaches 1.5 cm in greatest dimension is associated with decreased mortality.[37] Furthermore, several studies have demonstrated a low but persistent rate of distant metastasis associated with papillary microcarcinomas, confirming that it cannot always be presumed that such disease is indolent.[38–40] Thus, detection and removal of thyroid cancers that measure between 1 cm and 2 cm may result in a reduction in disease-related morbidity and mortality in the future.

Because this was an observational, retrospective study, it had limitations. However, we have tried to address these as follows. Analysis of our study population confirms that it represents 85% of all patients who were evaluated for thyroid nodular disease (by FNA) within our institution during the 10-year study period. In-depth examination of referral indications in a random subset of patients further demonstrated no evidence of biased selection. Thus, we believe our study cohort is unselected and representative of the population in the northeastern U.S. We acknowledge that only 59% and 53% of nodules that initially were deemed inadequate or 'atypical of an undetermined significance' underwent repeat aspiration, respectively. In part, this may be because most nodules with inadequate cytology were primarily cystic. Such nodules often are drained of fluid or may decrease in size over time.[18] This often alleviates the need for repeat FNA given the high likelihood of benign disease. Furthermore, some patients with initial cytology 'atypical of an undetermined significance' underwent surgery because of abnormal FNA cytology in a separate nodule that measured >1 cm in greatest dimension. However, whereas the above may in part explain the 50% to 60% rate of repeat aspiration, some impact of selection cannot be excluded.

Finally, it is possible our data may differ from those of others because of the geography of our cohort population, although no data firmly support this notion.

In summary, the current data demonstrate the increased accuracy that can be achieved by the use of UG-FNA in the evaluation of thyroid nodules. The results also suggest that a more precise characterization of "indeterminate" cytologic findings, together with repeated aspiration of nodules that demonstrate 'atypical cells of undetermined significance,' can predict postoperative cancer risk better, reducing the need for unnecessary surgery. Together, an interdisciplinary approach that encourages the use of ultrasound in conjunction with improved cytologic analysis increases the surgical yield for the detection of thyroid cancer beyond 50%, and supports recently published guidelines.[21,22]

REFERENCES

1. Brander A, Viikinkoski P, Tuuhea J, Voutilainen L, Kivisaari L. Clinical versus ultrasound examination of the thyroid gland in common clinical practice. *J Clin Ultrasound*. 1992;20:37–42.
2. American Cancer Society. Cancer Facts and Figures, 2006. Atlanta, Ga: American Cancer Society; 2006. Available at URL: http://www.cancer.org/docroot/STT/stt_0.asp Accessed October 1, 2007.
3. Gharib H, Goellner JR. Fine-needle aspiration biopsy of thyroid nodules. *Endocr Pract*. 1995;1:410–417.
4. Gharib H, Goellner JR, Johnson DA. Fine-needle aspiration cytology of the thyroid. A 12-year experience with 11,000 biopsies. *Clin Lab Med*. 1993;13:699–709.
5. Gharib H, Goellner JR. Fine-needle aspiration biopsy of the thyroid: an appraisal. *Ann Intern Med*. 1993;118:282–289.
6. Castro MR, Gharib H. Thyroid fine-needle aspiration biopsy: progress, practice, and pitfalls. *Endocr Pract*. 2003;9: 128–136.
7. Castro MR, Gharib H. Continuing controversies in the management of thyroid nodules. *Ann Intern Med*. 2005;142: 926–931.
8. Marqusee E, Benson CB, Frates MC, et al. Usefulness of ultrasonography in the management of nodular thyroid disease. *Ann Intern Med*. 2000;133:696–700.
9. Biscotti CV, Hollow JA, Toddy SM, Easley KA. ThinPrep versus conventional smear cytologic preparations in the analysis of thyroid fine-needle aspiration specimens. *Am J Clin Pathol*. 1995;104:150–153.
10. Afify AM, Liu J, Al-Khafaji BM. Cytologic artifacts and pitfalls of thyroid fine-needle aspiration using ThinPrep: a comparative retrospective review. *Cancer (Cancer Cytopathol)*. 2001;93:179–186.
11. Frost AR, Sidawy MK, Ferfelli M, et al. Utility of thin-layer preparations in thyroid fine-needle aspiration: diagnostic accuracy, cytomorphology, and optimal sample preparation. *Cancer (Cancer Cytopathol)*. 1998;84:17–25.
12. Malle D, Valeri RM, Pazaitou-Panajiotou K, Kiziridou A, Vainas I, Destouni C. Use of a thin-layer technique in thyroid fine needle aspiration. *Acta Cytol*. 2006;50:23–27.
13. Scurry JP, Duggan MA. Thin layer compared to direct smear in thyroid fine needle aspiration. *Cytopathology*. 2000;11:104–115.
14. DeLellis RA, Lloyd RV, Heitz PU, Eng C, editors. World Health Organization Classification of Tumours. Pathology and Genetics of Tumours of Endocrine Organs. Lyon, France: IARC Press; 2004.
15. American Joint Committee on Cancer. Thyroid. In: Greene FL, Page DL, Fleming ID, editors. AJCC Cancer Staging Manual. 6th ed. New York, NY: Springer-Verlag; 2002:77–87.
16. Hay ID, Bergstralh EJ, Goellner JR, Ebersold JR, Grant CS. Predicting outcome in papillary thyroid carcinoma: development of a reliable prognostic scoring system in a cohort of 1779 patients surgically treated at 1 institution during 1940 through 1989. *Surgery*. 1993;114:1050–1057; discussion 1057–1058.
17. Alexander EK, Heering JP, Benson CB, et al. Assessment of nondiagnostic ultrasound-guided fine needle aspirations of thyroid nodules. *J Clin Endocrinol Metab*. 2002;87:4924–4927.
18. Alexander EK, Hurwitz S, Heering JP, et al. Natural history of benign solid and cystic thyroid nodules. *Ann Intern Med*. 2003;138:315–318.
19. Alexander EK, Marqusee E, Orcutt J, et al. Thyroid nodule shape and prediction of malignancy. *Thyroid*. 2004;14:953–958.
20. Frates MC, Benson CB, Doubilet PM, et al. Prevalence and distribution of carcinoma in patients with solitary and multiple thyroid nodules on sonography. *J Clin Endocrinol Metab*. 2006;91:3411–3417.
21. Cooper DS, Doherty GM, Haugen BR, et al. Management guidelines for patients with thyroid nodules and differentiated thyroid cancer. *Thyroid*. 2006;16:109–142.
22. Pacini F, Schlumberger M, Dralle H, Elisei R, Smit JW, Wiersinga W. European consensus for the management of patients with differentiated thyroid carcinoma of the follicular epithelium. *Eur J Endocrinol*. 2006;154:787–803.
23. Barroeta JE, Wang H, Shiina N, Gupta PK, Livolsi VA, Baloch ZW. Is fine-needle aspiration (FNA) of multiple thyroid nodules justified? *Endocr Pathol*. 2006;17:61–65.
24. Sachmechi I, Miller E, Varatharajah R, et al. Thyroid carcinoma in single cold nodules and in cold nodules of multinodular goiters. *Endocr Pract*. 2000;6:5–7.
25. Cochand-Priollet B, Guillausseau PJ, Chagnon S, et al. The diagnostic value of fine-needle aspiration biopsy under ultrasonography in nonfunctional thyroid nodules: a prospective study comparing cytologic and histologic findings. *Am J Med*. 1994;97:152–157.
26. Belfiore A, La Rosa GL, La Porta GA, et al. Cancer risk in patients with cold thyroid nodules: relevance of iodine intake, sex, age, and multinodularity. *Am J Med*. 1992;93: 363–369.
27. Mazzaferri EL. Management of a solitary thyroid nodule. *N Engl J Med*. 1993;328:553–559.
28. Werga P, Wallin G, Skoog L, Hamberger B. Expanding role of fine-needle aspiration cytology in thyroid diagnosis and management. *World J Surg*. 2000;24:907–912.
29. Carmeci C, Jeffrey RB, McDougall IR, Nowels KW, Weigel RJ. Ultrasound-guided fine-needle aspiration biopsy of thyroid masses. *Thyroid*. 1998;8:283–289.
30. 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.

31. Sherman SI, Brierley JD, Sperling M, et al. Prospective multicenter study of thyroid carcinoma treatment: initial analysis of staging and outcome. National Thyroid Cancer Treatment Cooperative Study Registry Group. *Cancer*. 1998;83:1012–1021.
32. Loh KC, Greenspan FS, Gee L, Miller TR, Yeo PP. Pathological tumor-node-metastasis (pTNM) staging for papillary and follicular thyroid carcinomas: a retrospective analysis of 700 patients. *J Clin Endocrinol Metab*. 1997;82:3553–3562.
33. Sclabas GM, Staerkel GA, Shapiro SE, et al. Fine-needle aspiration of the thyroid and correlation with histopathology in a contemporary series of 240 patients. *Am J Surg*. 2003;186:702–709; discussion 709–7010.
34. Court-Payen M, Nygaard B, Horn T, et al. US-guided fine-needle aspiration biopsy of thyroid nodules. *Acta Radiol*. 2002;43:131–140.
35. Galloway JW, Sardi A, DeConti RW, Mitchell WT Jr, Bolton JS. Changing trends in thyroid surgery. 38 years' experience. *Am Surg*. 1991;57:18–20.
36. Grant CS, Hay ID, Gough IR, McCarthy PM, Goellner JR. Long-term follow-up of patients with benign thyroid fine-needle aspiration cytologic diagnoses. *Surgery*. 1989;106: 980–985; discussion 985–986.
37. Mazzaferri EL, Jhiang SM. Long-term impact of initial surgical and medical therapy on papillary and follicular thyroid cancer. *Am J Med*. 1994;97:418–428.
38. Chow SM, Law SC, Chan JK, Au SK, Yau S, Lau WH. Papillary microcarcinoma of the thyroid—prognostic significance of lymph node metastasis and multifocality. *Cancer*. 2003;98:31–40.
39. Pellegriti G, Scollo C, Lumera G, Regalbuto C, Vigneri R, Belfiore A. Clinical behavior and outcome of papillary thyroid cancers smaller than 1.5 cm in diameter: study of 299 cases. *J Clin Endocrinol Metab*. 2004;89:3713–3720.
40. Roti E, Rossi R, Trasforini G, et al. Clinical and histological characteristics of papillary thyroid microcarcinoma: results of a retrospective study in 243 patients. *J Clin Endocrinol Metab*. 2006;91:2171–2178.