

Experience With Standardized Thyroid Fine-Needle Aspiration Reporting Categories

Follow-Up Data From 529 Cases With “Indeterminate” or “Atypical” Reports

Min-En Nga, MD¹; Marian P. Kumarasinghe, MD²; Bibiana Tie, MD²; Gregory F. Sterrett, MD²; Benjamin Wood, MD²; John Walsh, MD³; Hue Nguyen, MD⁴; Andrew Whyte, MD⁵; and Felicity A. Frost, MD²

BACKGROUND. A significant number of thyroid fine-needle aspiration cytology (FNAC) cases yield inconclusive results. The recent National Cancer Institute guidelines and those published by other societies are important contributions to standardizing the diagnostic approach. Nevertheless, there are significant issues in the application of guidelines and the evaluation of their clinical efficacy. Data from individual departments can be useful in demonstrating the role of standardized reporting. **METHODS.** The authors followed 529 consecutive cases with inconclusive thyroid FNA results that were analyzed in a single laboratory in Western Australia. In that laboratory, standardized reporting in categories has been in place for a decade, and inconclusive cases have been subdivided into indeterminate and atypical groups. Follow-up data was obtained for 341 indeterminate cases (17.2% of total thyroid FNA accessions) and for 188 atypical cases (9.5% of accessions). **RESULTS.** In total, 127 nodules with atypical results (67.6%) underwent surgical excision compared with 131 nodules with indeterminate results (38.4%; $P < .0001$). In 96 excised nodules with atypical results (75.6%), the excised specimens were identified as neoplastic compared with 61 excised nodules with indeterminate results (46.6%; $P < .0001$). In addition, 31 excised nodules with atypical results (24.4%) had a malignancy proven compared with 17 excised nodules with indeterminate results (13%; $P < .05$). In addition, 51 of 82 repeat FNAs (62.2%) among patients who had indeterminate results yielded a more specific diagnosis compared with 2 of 9 repeat FNAs (22.2%) among patients who had atypical results ($P < .05$). **CONCLUSIONS.** The routine subcategorization of patients who had inconclusive thyroid FNA reports into indeterminate and atypical groups resulted in statistically significant differences in the likelihood of neoplasia and malignancy. Patients who had indeterminate results were more likely to benefit from repeat FNAC than patients who had atypical results. The current results indicated that patients who fall into these 2 categories are likely to benefit

Corresponding author: Min-En Nga, MD, Department of Pathology, National University of Singapore, 5 Lower Kent Ridge Road, Singapore 119074; Fax: (011) 6567780671; min_en_nga@nuhs.edu.sg

¹Department of Pathology, National University Health System, Singapore; ²Department of Anatomical Pathology, Sir Charles Gairdner Hospital, PathWest Queen Elizabeth II Medical Center, Perth, Western Australia; ³Department of Endocrinology, Sir Charles Gairdner Hospital, PathWest Queen Elizabeth II Medical Center, Perth, Western Australia; ⁴Department of Surgery, Sir Charles Gairdner Hospital, Queen Elizabeth II Medical Center, Perth, Western Australia; ⁵Perth Radiological Clinic, Subiaco, Perth, Western Australia

Presented in part at the 38th Annual Scientific Meeting of the Australian Society of Cytology; Sydney, Australia; October 24-27, 2008 and at the 17th International Congress of Cytology of the International Academy of Cytology and Annual Scientific Meeting of the British Society for Clinical Cytology; Edinburgh, Scotland; May 16-20, 2010.

This study was carried out by Dr. Nga during a Fellowship in the Cytopathology Unit of the Department of Anatomical Pathology, PathWest Queen Elizabeth II Medical Center in 2007 and 2008.

Received: May 10, 2010; **Revised:** July 24, 2010; **Accepted:** August 12, 2010

Published online October 13, 2010 in Wiley Online Library (wileyonlinelibrary.com)

DOI: 10.1002/cncy.20111, wileyonlinelibrary.com

from different clinical management protocols. *Cancer (Cancer Cytopathol)* 2010;118:423-33. © 2010 American Cancer Society.

KEY WORDS: thyroid fine-needle aspiration cytology, inconclusive thyroid nodules, thyroid nodules, follicular lesions, thyroid cytology diagnostic categories, thyroid fine-needle aspiration reporting.

The investigation of thyroid nodules using fine-needle aspiration (FNA) cytology (FNAC) has been used since the 1950s, when its use was established by Lopez Cardozo from Holland¹ and by Soderstrom in Sweden and the Karolinska Hospital group during the 1950s and 1960s.^{2,3} This diagnostic method has endured through the decades and yields higher sensitivity and predictive value for diagnosis than any other single diagnostic modality.⁴ The versatility of this simple diagnostic tool is evidenced by the finding that it is practiced widely in a variety of clinical settings, ranging from the primary healthcare clinic to specialist centers and tertiary referral centers.

It is widely accepted that FNAC for thyroid nodules functions mainly as a triaging test and is most useful in selecting cases for either clinical follow-up or surgical excision. The limitations of FNAC in specific diagnoses are well known, particularly for follicular neoplasms. However, precise performance criteria for this test and acceptable yields of neoplastic or malignant outcomes in cases referred for surgery have yet to be defined.⁵

The Cytopathology Unit of PathWest Queen Elizabeth II Medical Center (QE II MC) is a public-sector tertiary referral center in Western Australia with a major interest in thyroid cytology. The vast majority of material relates to the investigation of thyroid nodules. A standardized approach to all nongynecologic cytology reporting has been in place for over a decade. The conclusion of each report includes the diagnostic category and specific diagnosis sections followed by a recommendation for management, such as referral for a specialist opinion if appropriate. The 6 categories that are used for thyroid cytology reporting include nondiagnostic, benign, indeterminate, atypical, suspicious of malignancy, and malignant and are based on the general cytologic criteria discussed below. The recommendations given in reports generally are limited to a request for repeat FNA sampling (for indeterminate or nondiagnostic cases) or a referral for specialist surgical management (for all atypical, suspicious, or malignant cases).

In October 2007, the National Cancer Institute (NCI) hosted the “State-of-the-Science Conference” in Bethesda, Maryland, which was organized by Dr. Andrea Abati and comoderated by Drs. Susan Mandel and Edmund Cibas, to determine a consensus classification system for the reporting terminology and criteria in thyroid cytology.⁶ The newly proposed Bethesda System that was the result of this meeting is a landmark publication that proposes a system of 6 reporting categories—nondiagnostic/unsatisfactory, benign, atypia of undetermined significance, suspicious for follicular neoplasm, suspicious for malignancy, and malignant.⁶⁻⁸ Although we are unable to confirm an exact correlation between the Bethesda subcategories and our own categories, our view is that our “indeterminate” and “atypical” categories correspond most closely to the Bethesda categories of “atypia of undetermined significance” and “follicular neoplasm/suspicious for follicular neoplasm,” respectively. The other categories (nondiagnostic, benign, suspicious for malignancy, and malignant) essentially are concordant.

In our institution, approximately 25% of all thyroid FNAs yield an indefinite (indeterminate or atypical) result (unpublished data; QE II MC). We obtained follow-up data from a large series of cases that were subcategorized as indeterminate or atypical. Our objectives were 1) to examine the consistency of reporting by the multiple pathologists in the diagnostic service by reviewing their cytologic reports and cytologic descriptions, and 2) to examine the predictive value of the 2 subcategories in terms of the proportions of neoplasms and malignancies discovered at further follow-up.

MATERIALS AND METHODS

Case Setting

Currently, most thyroid FNA samples in Perth are obtained by radiologists under ultrasound guidance. The standard radiologic approach is to take 3 or 4 samples from nodules using a 25-gauge needle usually without

aspiration, at least initially unless a cystic lesion is suggested by imaging. All smears are air-dried and stained by the May-Grunwald-Giesma stain. Cell blocks are prepared if the radiologist submits excess blood-stained aspirate or if specifically requested by the pathologist, eg for repeat FNA samples. All smears are screened by cytotechnologists who prepare a draft report. The final report is constructed by the pathologist after all smears are examined. All reports are validated by pathologists.

A standard approach to smear preparation, in the form of an illustrated poster that can be pinned to notice boards, is recommended to our radiologists. Few FNAs of palpable nodules are performed by pathologists except for some inpatients and outpatients from the Sir Charles Gairdner Hospital (SCGH), which is the on-site hospital for the QE II MC. An on-the-spot reporting service is available for cases sampled in the Radiology Department of SCGH; however, currently, the majority of thyroid FNAs received in the department are from a large private radiology practice (Perth Radiological Clinic) and are reported by 4 cytopathologists (M.P.K., B.W., G.F.S., and F.A.F.) although up to 13 pathologists may be involved in reporting some thyroid FNAs from multiple sources.

Case Selection

Cytology reports of thyroid FNAs over a 3-year period from January 2004 to December 2006 were retrieved from the archives of QE II MC and reviewed. Cases that were categorized as indeterminate or atypical were selected for further study. The following diagnostic criteria were applied:

Cases are categorized as *indeterminate* when the interpreter is unable to resolve the cytologic features into those arising in a non-neoplastic setting (eg, nodular goiter, thyroiditis, hyperplastic nodule) versus those arising in a neoplasm (ie, neoplasm not excluded). Further investigation is necessary before definitive management; for example, a repeat FNA may be useful in confirming a diagnosis. These include:

1. a suboptimal sample because of hemodilution or low cellularity—few groups of follicular cells, and little colloid—not “nondiagnostic” but unable to definitively diagnose as benign;
2. a mixed pattern—some flat sheets, some microfollicular structures, and some colloid—but cannot exclude a follicular neoplasm (hyperplastic/adenomatous nodule vs follicular neoplasm);
3. findings that raise the possibility of thyroiditis (eg, the presence of moderate numbers of lymphocytes or lymphoid tangles but with an accompanying predominant microfollicular pattern), especially when there is clinical/imaging concern about neoplasm;
4. a history of Hashimoto thyroiditis but with a prominent Hurthle cell population (benign Hurthle cell nodule vs neoplasm), especially when there is clinical concern about neoplasm; and
5. a predominant pattern of colloid nodule but with a subpopulation of cells that exhibit nuclear atypia (eg, nuclear enlargement or pale chromatin or nuclear grooves; ie, a very low likelihood of a papillary lesion) but in which a repeat FNA may be useful in excluding malignancy.

In the *atypical* category, cytologic features strongly suggest neoplasm, and surgical excision is likely to be necessary for a definitive diagnosis. These include:

1. classic findings of a follicular neoplasm (predominantly microfollicular pattern with scant colloid);
2. predominantly Hurthle cells with little or no colloid and no background or history of Hashimoto thyroiditis; and
3. thyroid epithelial cells with some nuclear atypia (enlargement, pleomorphism, irregularity) but insufficient to categorize as suspicious or malignant.

Cases originally classified as *indeterminate* but in which there were findings of cyst contents were excluded. We place these cases in the nondiagnostic category if no further information was available or in the benign category if complete emptying could be achieved and if there were no atypical cytologic or imaging findings.

Review and Reclassification of Reports

The diagnostic cytology reports from all cases were reviewed by 2 observers (M.-E.N. and B.T.), and the reclassification was determined by consensus. Reclassification of cases was based mainly on discordance between the

use of the diagnostic category and the specific diagnosis in the original cytology reports. For example, cases that were categorized initially as “indeterminate” for which the specific diagnosis was “suggestive of follicular neoplasm,” were reclassified as “atypical.” Likewise, cases that were categorized originally as “atypical” with a specific diagnosis of “follicular neoplasm not fully excluded” or “features are between those of a hyperplastic nodule and neoplasm” were reclassified as “indeterminate.” Therefore, reclassified cases were those for which an incorrect category had been assigned by the pathologist but for which a specific diagnosis was given, indicating the likelihood of neoplasia. The reclassification exercise was performed without knowledge of the histologic outcome for an individual case. Glass slides were not reviewed, because our objective was to test consistency of the use of categories and specific diagnoses along guidelines already agreed on within the Cytopathology Unit rather than the accuracy of cytologic interpretation of cases by individual pathologists.

Patient Follow-Up

A single thyroid FNA accession may include sampling of more than 1 nodule, but data about this were not extracted for all thyroid FNA cases, and the most significant cytologic abnormality detected was used for the purposes of follow-up and statistical evaluation. When repeat FNA had been performed on a particular nodule, the results from both samples were included in the total number of index cases with follow-up. Therefore, the total represents the number of FNA samples rather than the total number of patients or nodules.

Follow-up data were obtained as follows: Subsequent histology or cytology reports for all patients were obtained from the common electronic archives of public sector PathWest pathology laboratories in Western Australia. If no cytology or histology follow-up was available in the PathWest database, then letters were sent to referring clinicians to obtain follow-up information. Only histologic or cytologic follow-up data from the database of the PathWest laboratory or healthcare providers were included in the follow-up analysis. Patients who had only clinical and/or radiologic follow-up were excluded.

Acquisition of follow-up data was during the period from July to December 2008. Thus, the period of follow-up for individual cases ranged from a minimum of 18 months to a maximum of 47 months.

Selected Case Review

A review of the cytologic material from 16 cases that were classified as indeterminate and 30 cases that were reported as atypical but had subsequent malignant histologic diagnoses was performed by 4 pathologists (M.-E.N, M.P.K., G.F.S., and F.A.F.).

Statistical Analyses

Kappa statistics were used to evaluate the degree of inter-observer agreement between original report categories and the review classification. Chi-square tests (2-tailed) were used to assess statistical significance for differences in outcomes for the 2 subcategories in the follow-up study.

RESULTS

Review and Reclassification of Reports

During the study period, all thyroid FNA reports were in the standard format described above, ie, reporting category, followed by specific diagnosis and comments, including recommendations for management, where appropriate. In total, 529 inconclusive reports on individual lesions were studied from a total of 517 patients, representing 26.6% of the 1986 thyroid FNA accessions that were received during the study period. This is an overestimate of the proportion of such cases, because an individual accession may include sampling from multiple nodules. However, data for the total numbers of lesions or nodules sampled during the study period were not extracted from the files. Samples from 338 lesions were originally classified as indeterminate, 183 were originally classified as atypical, and 8 were originally were coded as indeterminate/atypical. Upon review, 19 (5.6%) of the original 338 indeterminate reports were reclassified as atypical, and 15 (8.2%) of the original 183 atypical reports were reclassified as indeterminate. Seven of the 8 cases that were categorized originally as indeterminate/atypical were reclassified as indeterminate, and 1 was reclassified as atypical.

The kappa coefficient of agreement with the original report category after review was 0.85 (0.935 agreement; standard error, 0.024; range, 0.811-0.904). The final numbers of indeterminate and atypical cases for the follow-up study were 341 (17.2% of the total thyroid FNAs) and 188 (9.5% of the total thyroid FNAs), respectively.

Table 1. Histologic Follow-Up of Cases

Cytologic Category	Histologic Diagnosis										Total
	Malignant				Benign						
	PTC	FC	HCC	Other	Neoplastic		Non-Neoplastic				
	FA	HCA	PTA	CN	HT	DQT					
Indeterminate	9	7	0	1 ^a	36	6	2	64	5	1	131
Atypical	16	10	4	1 ^b	54	10	1	25	6	0	127
Total	25	17	4	2	90	16	3	89	11	1	258

PTC indicates papillary thyroid carcinoma; FC, follicular carcinoma; HCC, Hurthle cell carcinoma; FA, follicular adenoma; HCA, Hurthle cell adenoma; PTA, parathyroid adenoma; CN, colloid nodule; HT, Hashimoto thyroiditis; DQT, De Quervain thyroiditis.

^aCarcinoma not otherwise specified

^bCarcinoma (?)metastatic.

The indeterminate group included 267 females and 68 males who ranged in age between 14 years and 94 years (mean age, 51.9 years); and the atypical group included 137 females and 45 males who ranged in age from 15 years to 90 years (mean age, 49.7 years).

Patient Follow-Up

When subsequent histology or cytology results were available, the mean and median follow-up was 4.2 months and 3 months, respectively (range, from 2 days to 18 months). For patients who did not undergo a subsequent biopsy, the maximal clinical follow-up was up to 47 months.

Table 1 summarizes the histologic follow-up data. In total, 127 of 188 atypical cases (67.6%) underwent excision. Neoplasm (including follicular or Hurthle cell adenoma/carcinoma, papillary carcinoma or other malignancy) was identified in 96 cases, representing 75.6% of excised lesions and 51.1% of the total number of atypical cases. Malignancy was identified in 31 cases, representing 24.4% of the excised cases and 16.5% of the total.

One hundred thirty-one of 341 indeterminate cases (38.4%) underwent surgical excision. Neoplasm was identified in 61 cases, representing 46.6% of the excised lesions and 17.9% of the total number of indeterminate cases. Thus, the positive predictive value (PPV) for neoplasia in excised atypical cases was 75.6% (95% confidence interval [CI], 67.4%-82.3%), whereas the PPV for neoplasia in excised indeterminate cases was 46.6% (95% CI, 38.2%-55.1%). Malignancy was identified in 17 indeterminate cases, representing 13% of the excised lesions and 5% of the total. Chi-square tests revealed a statistically significant difference in the prevalence of neoplasm (chi-square test, 21.6; $P < .0001$) and malignancy

(chi-square test, 4.84; $P < .05$) between the excised atypical and indeterminate cases.

Repeat FNA sampling was carried out in 82 of 341 indeterminate cases (24%) (Table 2). This yielded a more specific diagnosis in 51 cases (62.2%), including a benign diagnosis in 38 cases (46.3%) and an atypical diagnosis in 13 cases (15.9%). Repeat FNA sampling was performed in 9 of the 188 atypical cases (4.8%). This resulted in a more specific (benign) diagnosis in 2 cases (22.2%). In the indeterminate group, 2 of 34 cases that were classified as colloid nodule on repeat FNA were excised, and both revealed nodular goiter. One case that was reported as non-diagnostic on repeat FNA was excised and revealed nodular goiter. Six cases in which the repeat FNA was indeterminate underwent excision, which revealed 3 nodular goiters, 2 follicular adenomas, and 1 papillary carcinoma. Four cases that were reported as atypical on repeat FNA were excised, and all were identified as follicular adenomas. In the atypical group, 1 case that was reported as indeterminate on repeat FNA was excised and revealed a Hurthle cell follicular adenoma, and 2 repeat atypical cases were excised and revealed 1 nodular goiter and 1 Hurthle cell carcinoma.

Cytologic Review of Cases With Malignancy on Excision

A review of 16 available cases that were classified as indeterminate but had a subsequent malignant histologic diagnosis was performed by 4 pathologists (M.-E.N., M.P.K., G.F.S., and F.A.F.). The histologic diagnoses were papillary thyroid carcinoma (PTC) in 9 cases (2 classic, 3 follicular variant, 1 cystic, and 3 PTCs, variant not specified) and follicular carcinoma in 7 cases. Most of the smears

Table 2. Results of Repeat Fine-Needle Aspiration

Initial FNA Category	Result of Repeat FNA					Total
	Benign		Nondiagnostic	Inconclusive		
	CN	HT		IND	ATYP	
Indeterminate	34 ^a	4	8	23	13	82
Atypical	2 ^a	0	0	1	6	9
Total	36	4	8	24	19	91

FNA indicates fine-needle aspiration; CN, colloid nodule; HT, Hashimoto thyroiditis; IND, indeterminate; ATYP, atypical.

^aOne patient had a benign cytologic diagnosis on the second repeat FNA. All other patients underwent only 1 repeat FNA.

were of low-to-moderate cellularity with marked hemodilution. However, 10 of the 16 cases were upgraded to “atypical” based on the presence of repetitive, microfollicular structures; rounded, crowded clusters; and minimal colloid. In 2 cases that were confirmed as PTC on excision, there were features suspicious for PTC, such as significant nuclear enlargement in 1 case and possible psammoma bodies and the presence of monolayered sheets with rare nuclear-intranuclear inclusions in another. The histologic diagnoses in the atypical cases that had a subsequent malignant histologic diagnosis included 16 papillary carcinomas, 10 follicular carcinomas, 4 Hurthle cell carcinomas, and 1 poorly differentiated carcinoma of uncertain origin. The review diagnoses of the 30 available atypical cases with subsequent malignant histologic diagnoses were in complete agreement with the original reports in 26 cases, whereas 4 were upgraded (1 to “suspicious for papillary carcinoma” and 3 to “malignant” [papillary carcinoma]). The cases were upgraded based on the presence of cellular areas composed of sheets of cells (some exhibiting a papillary configuration), fibrovascular cores, and nuclear grooves and inclusions.

DISCUSSION

The inconclusive thyroid FNA result is a subject of considerable recent interest.⁹⁻¹¹ There is broad consensus that such cases should be subdivided into groups with lower and higher PPVs/likelihoods of associated neoplasm and malignancy, although there is controversy regarding which terms are best to describe the groups. Published follow-up data for such cases are relatively limited.⁶⁻⁸ To our knowledge, this is 1 of the largest follow-up studies in the English literature to date for cytologically inconclusive thyroid FNAs in the setting of a thyroid nodule. We use

the term *indeterminate* when we are unable to resolve the cytologic features into those arising in a non-neoplastic setting (eg. nodular goiter or thyroiditis) versus those observed in a neoplasm. Therefore, this category is used when it is difficult to exclude neoplasm rather than when the findings strongly suggest neoplasm. We use the term *atypical* when the cytologic findings strongly suggest follicular neoplasm. We do not use follicular neoplasm as a stand-alone cytologic category, because hyperplastic nodules in nodular goiter and Hurthle cell nodules in nodular goiter and Hashimoto thyroiditis may have cytologic features similar to or even indistinguishable from those observed follicular neoplasms.¹² Although we may agree on standard terminology, it is another matter to use this terminology consistently. Nevertheless, in the current study, a separation of cases into indeterminate and atypical categories based on generally recognized cytologic criteria appeared to be applied relatively consistently by pathologists. Diagnostic categories were mostly in agreement with the specific diagnoses given in reports. A review and reclassification of the 529 descriptive reports revealed >90% agreement with the original report subcategory. A blinded review of glass slides would have provided more information about this aspect of reporting but was outside the scope of the current study.

Our follow-up data revealed a statistically significant difference between excised atypical and indeterminate cases in terms of both the likelihood of malignancy (24.4% of excised atypical cases vs 13% of excised indeterminate cases; $P < .05$) and neoplasia (75.6% of excised atypical cases vs 46.6% of excised indeterminate cases; $P < .0001$) (Table 1). The prevalence of malignancy was 16.5% for all atypical cases compared with 5% for all indeterminate cases ($P < .0001$), and the prevalence of neoplasia was 51.1% for all atypical cases and 17.9% for all indeterminate cases ($P < .0001$). Repeat FNA sampling

of 82 indeterminate cases yielded a more specific result (ie, a benign or atypical report) in 62.2% of those cases.

We believe that the current results validate the reporting approach that is used in our laboratory and provide valuable information for our clinicians and radiologists. The findings support a management approach of excision for lesions that are classified as atypical and an initial repeat FNA sampling in the indeterminate cases, especially for patients in whom clinical follow-up rather than excision is preferred for medical reasons.

The limitations of our study include a much lower excision rate for indeterminate cases compared with atypical cases and a short and incomplete clinical follow-up (<48 months) for patients with nonexcised lesions. There also was limited opportunity to review histologic material, because many cases were reported outside of public-sector laboratories. Nevertheless, interobserver variation in the assessment of vascular and capsular invasion in follicular neoplasms and in diagnosing the follicular variant of papillary carcinoma are recognized problems for all pathologists and would not necessarily have been improved by a review in our department. Review of cytologic material also was restricted to a small subset of cases in which malignancy was identified on excision; therefore, a detailed correlation of imaging findings, precise sites of sampling, and precise sites of tumors within the thyroid was not possible. Despite all of these caveats, we believe that the quantum of cases studied is large enough to confirm good results.

The recent literature includes several follow-up studies that had objectives comparable to ours. Recently, Nayar and Ivanovic presented the histologic follow-up of lesions that were “indeterminate for neoplasm” (defined as cases in which the differential cytologic diagnoses included hyperplasia, follicular neoplasm, and follicular variant of papillary carcinoma) and identified neoplasms in 48% of cases and malignancy in 6%. In lesions that were reported as “neoplasm,” their rates of neoplasia and malignancy were 77% and 53%, respectively.⁹ Layfield et al analyzed the histologic outcomes of 127 “follicular lesions of undetermined significance (FLUS)” and identified malignant lesions in 28.3% of cases.⁷ This malignancy yield is similar to that of our atypical category. However, the recently published *Bethesda System for Reporting of Thyroid Cytology*⁶⁻⁸ estimates a malignancy rate of 5% to 10% for the FLUS subcategory, which is closer to outcomes for our indeterminate subcategory. Layfield et al concluded that application of the FLUS category will require

criteria that are defined more rigorously before it can become useful as a guide in clinical management.

Yang et al conducted histologic follow-up of 52 cases that were reported as “atypical cellular lesion” with a malignancy rate of 19.2%¹³ compared with a malignancy rate of 32.3% from their 326 cases that were reported as “follicular neoplasm.” Bahar et al reported a malignancy rate of 26% in 58 cases that were diagnosed cytologically as “follicular lesion,” which they defined as lesions in which the cytologic findings could not distinguish reliably between follicular neoplasm and adenomatous goiter.¹¹

In the major classification systems, the term “follicular lesion” has been used rather variably. The 1996 Papanicolaou Society of Cytopathology guidelines and the NCI Bethesda group both use this term for cases that have features straddling those of non-neoplastic and neoplastic nodules.^{7,14} The Papanicolaou Society classification subdivides “cellular follicular lesion” into “cellular follicular lesion, favor hyperplastic/adenomatous nodule” and “favor follicular neoplasm,” which have an overall estimated malignancy rate of 15% to 20%. The former lesions should prompt conservative management, whereas surgery is suggested for the latter.¹⁴ A revised version is presented in Table 3.¹⁵ A variety of terms synonymous with “follicular lesion” include FLUS, atypical follicular lesion, and cellular follicular lesion. At the NCI “State-of-the-Science Conference” in 2007, Baloch et al indicated that the intended use of these terms was to denote that the findings were neither convincingly benign nor sufficiently atypical to render a diagnosis of “follicular neoplasm” or “suspicious for malignancy.”^{6,7} The recommended incidence of this category was not to exceed 7% of all FNA results. However, this recommendation is significantly less than our results for the indeterminate category (17.2%) and for the findings reported by Layfield et al in a recent, large, triple-center review of 6782 cases in which the incidence of FLUS varied widely among the 28 reporting pathologists (range, 2.5%-28.6%) and between participating institutions (range, 3.3%-14.9%).¹⁰ It appears likely that there still is significant interobserver variation associated with these terms.

From a surgical/clinical perspective, the term “follicular lesion” may harbor some uncertainty. Bahar et al, as mentioned above, reported a high rate of malignancy of 26% in their “follicular lesions,” sufficient to recommend surgical excision for all nodules that fall into this

Table 3. Proposed Classification Schemes for Thyroid Fine-Needle Aspiration Reporting

British Thyroid Association and Royal College of Physicians, 2007¹⁷	Papanicolaou Society, 2007¹⁵ (Yang 2007¹³)	NCI-Bethesda System, 2009 (Baloch 2008⁶ and Layfield 2009¹⁰)
Thy 1: Inadequate	Unsatisfactory	Unsatisfactory
Thy 2: Benign	Benign	Benign
Thy 3: Follicular lesion/suspected follicular neoplasm—discuss at multidisciplinary team meeting; probable surgical removal	Atypical cellular lesion, follicular neoplasm not excluded (19.2% malignant; n=52)→follow-up or repeat FNA	Atypia of undetermined significance/follicular lesion of undetermined significance/cellular follicular lesion/atypical follicular lesion (estimated 5%-15% malignant)→repeat FNA
Thy 4: Suspicious of malignancy→excise	Follicular neoplasm/indeterminate (32.2% malignant; n=326)→excise Suspicious for PTC→excise	Suspicious for neoplasm/follicular neoplasm (15%-30% malignant)→excise Suspicious for malignancy (50%-70% malignant)→excise
Thy 5: Malignant	Malignant	Malignant

NCI indicates National Cancer Institute; Thy, British Thyroid Association thyroid classification; FNA, fine-needle aspiration; PTC, papillary thyroid carcinoma.

category.¹¹ This is in contradistinction to the malignancy rate of 5% to 10% put forward by Baloch et al⁶ and to our own findings in the indeterminate group. For these reasons, we believe that use of the term “follicular lesion” may have limited value. It is regarded differently among pathologists and clinicians, causes considerable confusion for clinicians in our community, has no precise diagnostic meaning without further qualification, and can be applied to many pathologic processes in the thyroid. Therefore, it is clear that a more precise, simple, and universally agreed terminology is required for a better understanding among all physicians involved in the management of these patients, including pathologists, general practitioners, endocrinologists, and surgical specialists.

The new Bethesda System proposes a system of 6 reporting categories.⁶⁻⁸ The 2 inconclusive categories are “atypia of undetermined significance” and “suspicious for follicular neoplasm,” and the estimated risks of malignancy correspond well to our actual follow-up data of indeterminate and atypical categories, respectively (Table 3). Although we are in agreement with the categories proposed, we have observed that use of the indeterminate and atypical categories allows a more flexible approach for our pathologists and clinicians. In our particular clinical setting and practice, the term “atypia” in thyroid cytology reports, as in cytology reports in general, implies the need for histologic diagnosis to prove or exclude neoplasm or malignancy. If applied to our “indeterminate” group, this may lead to over-reaction both by clinicians and patients and may force the surgeon to advise surgery when, in fact, the risk of malignancy in that category is low, ranging from

5% to 13% (QE II MC) and from 5% to 10% (suggested as satisfactory by the NCI review), and when repeat FNA sampling may resolve the need for immediate operation.

In September 2009, a working group of 28 members from 14 European countries (the European Federation of Cytology Societies) met at the European Congress of Cytology in Lisbon to discuss the reporting nomenclature in thyroid cytology.¹⁶ The discussion concluded that, although thyroid FNA reporting should be standardized, and the NCI Bethesda terminology offers a “risk-stratification” approach with clinical recommendations, more than half of the group favored a translation of local terminology as the first step toward a unified European nomenclature. The individual national societies for cytology would be tasked to match their local terminologies to the Bethesda classification, observe its acceptance by clinicians, and audit its correlation with outcome.

In the United Kingdom, a national reporting system has been in use for some years. The British Thyroid Association, in collaboration with the Royal College of Physicians, recommend a 5-category system of thyroid FNA reporting using the categories Thy 1 through Thy 5¹⁷ (Table 3). In this system, which was revisited last in 2007, the category Thy 3 encompasses our indeterminate and atypical categories.

Combining the 2 subcategories from our current series would result in a PPV for malignancy of 18.6% for excised lesions and 9.1% for all lesions. The PPV for neoplasm was 60.9% for excised lesions and 29.7% for all lesions. The 9.1% figure is similar to the often quoted figure of approximately 10% for the prevalence of

Table 4. Thyroid Fine-Needle Aspiration Reporting at PathWest Queen Elizabeth II Medical Center Information for Clinicians and Radiologists

Category, % of Total Cases ^a	Specific Diagnosis	Recommendation	PPV/Risk of Malignancy ^a	Explanatory Notes/ Follow-Up Data
Nondiagnostic, 10%-15%	1) Insufficient material for diagnosis; 2) Cyst/hemorrhage with no epithelium, no colloid, and no information about whether a mass remained—cystic change/hemorrhage into neoplasm not excluded	1) Repeat sample, eg, in 6-8 wk; 2) Refer to specialist	Low, <5%	1) If a satisfactory sample is not obtained on repeat FNA, then refer for specialist surgical opinion; 2) Papillary carcinoma can be predominantly cystic
Benign, 40%-60%	1) Colloid nodule; 2) Cyst aspirated to dryness—no residual mass by palpation or ultrasound and no atypical imaging or cytologic features; 3) Thyroiditis (Hashimoto or de Quervain)	1) A repeat sample in 6 mo is used by some centers to reinforce a benign diagnosis; 2) Refer to specialist if cyst recurs; 3) Refer for medical treatment	Very low, <1%	1) Precise local follow-up data not available, adjacent incidental tumor may not be sampled, papillary carcinomas rarely have abundant colloid in FNA samples, and purely Hurthle cell lesions generally need excision; 2) Papillary carcinoma can be predominantly cystic; 3) Dual pathology (with neoplasm/malignancy) may occur
Indeterminate, 15%-20%	Unable to distinguish between colloid nodule and neoplasm or between thyroiditis and neoplasm	Refer to specialist or repeat FNA in 6-8 wk; QE II MC 2004-2006: 62.2% of repeat FNAs produce a more definitive answer	Low, 5%-13%	QE II MC 2004-2006: 17 malignancies (5% of all 341 cases; 13% of 131 cases with histologic follow-up), and 61 neoplasms (17.9% of total patients; 46.6% of excisions); the difference between the PPV for malignancy and neoplasm on histologic follow-up in the indeterminate and atypical categories was highly significant ($P=0.01$).
Atypical, ~10%	Suggests a follicular neoplasm or atypical cells seen, malignancy not excluded; the terms "follicular lesion" or "follicular nodule" are not used in reports because of low specificity	Refer to specialist surgeon	Moderate, 16.5%-24.4%	QE II MC 2004-2006: 31 malignancies (16.5% of all 188 cases; 24.4% of 127 cases with histologic follow-up) and 96 neoplasms (51.1% of total cases; 75.6% of excisions); excision is needed to distinguish between follicular adenoma, carcinoma or hyperplastic nodule; follicular variant of papillary carcinoma often is diagnosed as "suggests follicular neoplasm" on cytology
Suspicious of malignancy, 2%-3%	Findings suggest malignancy (specify type); histologic confirmation needed before definitive management	Refer to specialist surgeon; repeat FNA for ancillary tests is required to confirm lymphoma, medullary carcinoma, anaplastic carcinoma, or metastases	High, ~90%	High predictive value, but further sampling or histopathology is required for definitive management
Malignant, 3%-4%	Unequivocal diagnosis of malignancy, including tumor type	Refer to specialist surgeon; ancillary testing (repeat FNA) usually is required to confirm lymphoma, medullary carcinoma, anaplastic carcinoma, or metastases	Very high, ~100%	Can be used for definitive management; a diagnosis of medullary carcinoma needs cell-block immunocytochemistry and serum calcitonin levels for proof; repeat FNA usually is requested if no material for cell-block is submitted

PPV indicates positive predictive value; FNA, fine-needle aspiration; QE II MC, PathWest Queen Elizabeth II Medical Center.

^aThe estimated frequency (unpublished data; PathWest QE II MC).

malignancy in cold nodules before the introduction of FNA. The main objective of FNA is to provide more effective triaging of cases; and, in our practice, combining both categories probably would result in less clinically effective results.

Table 4 provides an information sheet for our pathologists, clinicians, and radiologists that summarizes our approach to reporting thyroid FNA, current management recommendations, and expected outcomes. We use a 2 tiered/Linnaean system of category and specific diagnosis that can cover almost all eventualities and allows detailed comment and discussion in reports if required. The table includes the results from the current study. For our specialist-referred cases, recommendations generally are not provided in reports. Thus, the management guidelines will be most useful for general practitioners, but the table also may be useful as a summary for surgeons, endocrinologists, and radiologists, because it can be pinned to notice boards and read easily, it provides a basic framework for teaching students or trainees, and it includes local outcome data that can be expanded or updated.

Our general diagnostic approach has been reviewed by a group of endocrine surgeons from multiple institutions in Perth and is supported by them. In particular, the clear-cut recommendation of referral for specialist opinion for all indeterminate or atypical cases has been strongly endorsed by surgeons. To our knowledge, this is adhered to by radiologists and general practitioners. Thus, decisions about further management or the possible role of repeat FNA are made by the surgeon in consultation with the radiologist and pathologist, as required.

Other refinements in thyroid FNA reporting, including more critical cytologic evaluation and using clinical factors to aid in triage, have been considered in the past and were proposed again recently. Kelman et al described the surgical outcomes of 98 inconclusive thyroid nodules using nuclear atypia as a marker for neoplasia.¹⁸ Those authors defined atypia as nuclear enlargement, pleomorphism, abnormal chromatin, and prominent nucleoli. Cases were classified into “microfollicular pattern without atypia” and “atypia with or without microfollicular pattern.” Malignant outcomes were present in 3 of 46 patients (6.5%) in the first group and in 31 of 52 patients (60%) in the second group. Thus, Kelman et al concluded that significant nuclear atypia was helpful in triaging cases for surgical management. Other

authors have not reported that the presence of nuclear atypia as defined here was a reliable and consistent feature of follicular carcinoma.¹⁹ Baloch et al stratified 167 patients with cytologically diagnosed “follicular neoplasms” by age, sex, and nodule size. Those authors observed that age >40 years, male sex, and nodule size >3 cm conferred a statistically significant increased risk of a malignant outcome.²⁰ These parameters also may be taken into account by clinicians in the management of individual patients. Finally, our selected, retrospective case review of those with malignant outcomes revealed some under calling in the presence of overall low cellularity and significant hemodilution, which hampered optimal cytologic evaluation. There also was an over-representation of nonclassic PTC in tumors that were confirmed histologically (3 follicular variant PTCs and 1 cystic PTC vs 2 classic PTCs), highlighting the difficulty in diagnosing PTC variants, especially the follicular variant, by cytology.

In conclusion, the results from this large follow-up study indicated that a 6-tier classification system with subdivision of inconclusive thyroid FNA results into indeterminate and atypical classes (see Tables 1 and 4) resulted in a high PPV for malignancy and neoplasia in the atypical group and confirmed that this group warrants surgical intervention. The lower PPV for malignancy and neoplasia in the indeterminate group may justify repeat FNA and specialist follow-up rather than immediate excision. The application of this reporting system was reasonably consistent among pathologists in our practice and yielded satisfactory clinical outcomes.

CONFLICT OF INTEREST DISCLOSURES

The authors made no disclosures.

REFERENCES

1. Cardozo LP. *Clinical Cytology*. Leyden, Netherlands: Stafleu; 1954.
2. Soderstrom N. Puncture of goitres for aspiration biopsy: a preliminary report. *Acta Med Scand*. 1952;144:235-244.
3. Einhorn J, Franzen S. Thin-needle biopsy in the diagnosis of thyroid disease. *Acta Radiol*. 1962;58:321-336.
4. Ashcroft MW, van Herle AJ. Management of thyroid nodules II. *Head Neck Surg*. 1981;3:297-322.
5. Layfield LJ, Cibas ES, Gharib H, Mandel SJ. Thyroid aspiration cytology: current status. *CA Cancer J Clin*. 2009;59:99-110.

6. Baloch ZW, LiVolsi VA, Asa SL, et al. Diagnostic terminology and morphologic criteria for cytologic diagnosis of thyroid lesions: a synopsis of the National Cancer Institute Thyroid Fine-Needle Aspiration State of the Science Conference. *Diagn Cytopathol*. 2008;36:425-437.
7. Cibas ES, Ali SZ. The Bethesda System for Reporting Thyroid Cytopathology. *Am J Clin Pathol*. 2009;132:655-657.
8. Ali SZ, Cibas ES, eds. The Bethesda System for Reporting Thyroid Cytopathology. Definitions, Criteria and Explanatory Notes. New York: Springer Science + Business Media; 2010.
9. Nayar R, Ivanovic M. The indeterminate thyroid fine-needle aspiration. Experience from an academic centre using terminology similar to that proposed in the 2007 National Cancer Institute Thyroid Fine-Needle Aspiration State of the Science Conference. *Cancer (CancerCytopathol)*. 2009;117:195-202.
10. Layfield LJ, Morton MJ, Cramer HM, Hirschowitz S. Implications of the proposed thyroid fine-needle aspiration category of "follicular lesion of undetermined significance": a 5-year multi-institutional analysis. *Diagn Cytopathol*. 2009;37:710-714.
11. Bahar G, Braslavsky D, Shpitzer T, et al. The cytological and clinical value of the thyroid "follicular lesion." *Am J Otolaryngol*. 2003;24:217-220.
12. DeMay R, ed. Thyroid. In: DeMay R, ed. The Art and Science of Cytopathology. Chapter 17. Chicago, IL: American Society of Clinical Pathologists; 1996:724-740.
13. Yang J, Schnadig V, Logrona R, Wasserman PG. Fine-needle aspiration of thyroid nodules: a study of 4703 patients with histologic and clinical correlations. *Cancer (Cancer Cytopathol)*. 2007;111:306-315.
14. The Papanicolaou Society of Cytopathology Task Force on Standards of Practice. Guidelines of the Papanicolaou Society of Cytopathology for the examination of fine-needle aspiration specimens from thyroid nodules. *Mod Pathol*. 1996;9:710-715.
15. Papanicolou Society of Cytopathology Recommendations for Thyroid Fine-Needle Aspiration. Available at: <http://www.pathology.org/guidelines.html>. Accessed March 2, 2010.
16. Kocjan G, Cochand-Priollet B, de Agustin PP, et al. Diagnostic terminology for reporting thyroid fine needle aspiration cytology: European Federation of Cytology Societies Thyroid Working Party Symposium, Lisbon 2009. *Cytopathology*. 2010;21:86-92.
17. British Thyroid Association, Royal College of Physicians; Perros P, ed. Guidelines for the Management of Thyroid Cancer. 2nd ed. Report of the Thyroid Cancer Guidelines Update Group. London, United Kingdom: Royal College of Physicians; 2007.
18. Kelman AS, Rathan A, Leibowitz J, Burstein DE, Haber RS. Thyroid cytology and the risk of malignancy in thyroid nodules: importance of nuclear atypia in indeterminate specimens. *Thyroid*. 2001;11:271-277.
19. Orell SR, Sterrett GF, Whitaker D. Fine Needle Aspiration Cytology. 4th ed. Philadelphia, Pa: Churchill Livingstone; 2005.
20. Baloch ZW, Fleisher S, LiVolsi VA, Gupta PK. Diagnosis of "follicular neoplasm": a gray zone in thyroid fine-needle aspiration cytology. *Diagn Cytopathol*. 2001;26:41-44.