Pathology – Research and Practice 211 (2015) 345–348

Contents lists available at ScienceDirect

Pathology – Research and Practice

journal homepage: www.elsevier.com/locate/prp

Original article

Examining the Bethesda criteria risk stratification of thyroid nodules

Ahmed Deniwar[a], Catherine Hambleton[a], Tina Thethi[b], Krzysztof Moroz[c], Emad Kandil[a,*]

[a] Division of Endocrine and Oncologic Surgery, Department of Surgery, Tulane University School of Medicine, New Orleans, LA, United States
[b] Division of Endocrinology and Metabolism, Department of Medicine, Tulane University School of Medicine, New Orleans, LA, United States
[c] Department of Pathology, Tulane University School of Medicine, New Orleans, LA, United States

ARTICLE INFO

Article history:
Received 23 July 2014
Received in revised form 5 January 2015
Accepted 4 February 2015

Keywords:
Thyroid cancer
Thyroid cytology/FNA
Thyroid nodule evaluation
Thyroid pathology
Thyroid surgery

ABSTRACT

Background: The Bethesda criteria are proposed for appropriate stratification of malignancy risk in thyroid nodules, but controversy exists regarding their accuracy and reliability in decision making. Additionally, previous studies have suggested higher rates of both malignancy and false negative fine needle aspiration biopsy (FNA) associated with increasing nodule size. This study aims to determine the accuracy of ultrasound (US)-guided FNA using the current Bethesda criteria in surgical practice. We also aimed to investigate the relationship between nodule size and malignancy.
Methods: A retrospective analysis of US-guided FNAs by a single surgeon during a 4.5 year period. FNA results using Bethesda criteria were compared to final surgical pathology.
Results: 611 patients with thyroid nodules underwent US-guided FNA. FNA results in 375 subsequently excised thyroid nodules were recorded according to the Bethesda criteria: 192 (51%) benign, 65 (17%) atypia of unknown significance/follicular lesion of undetermined significance (AUS/FLUS), 42 (11%), suspicious for follicular neoplasm (SFN), 17 (5%) suspicious for malignancy (SM), 28 (8%) malignancy, and 31 (8%) non-diagnostic. Malignancy was confirmed by surgical pathology in 15%, 34%, 50%, 88%, 100%, and 39% of the above groups respectively. Sensitivity, specificity, and false-negative rate were 61%, 99%, and 15% respectively. No correlation existed between the size of nodules with indeterminate FNA results and malignancy rate ($p = 0.89$), or size of nodules with non-diagnostic FNA and malignancy rate ($p = 0.50$).
Conclusion: The current Bethesda risk stratification system underestimated malignancy rates in benign, indeterminate and non-diagnostic cytopathologic categories in our experience. There was no positive linear correlation between nodule size and malignancy rate in these cytopathologic categories.

Introduction

Thyroid nodules are clinically palpable in 4–7% of adults in the United States [1–6]. With ultrasound (US) examination, thyroid nodules are identified in up to 70% of adults [7]. Less than 5% of thyroid nodules are malignant; however incidence of both thyroid nodules and thyroid cancer are rapidly increasing [8,9].

Thyroid surgery can be associated with multiple complications, such as post-operative thyroid hormone imbalance, hypoparathyroidism, recurrent laryngeal nerve injury, bleeding, or infection; thus, there has been an effort to limit unnecessary surgery in asymptomatic patients with benign lesions [5,6,9–13].

Fine needle aspiration (FNA) biopsy is now the gold standard diagnostic test as per the American Thyroid Association (ATA) guidelines for the initial evaluation of a thyroid nodule [8,14]. Reported sensitivity and specificity for US-guided FNA range from 65 to 98% and 72 to 100%, respectively [3,15–18]. FNA diagnostic accuracy is limited by skill of the aspirator and expertise of the cytologist [6].

US-guided FNA accuracy is also limited by nodule size; nodules larger than 4 cm in diameter have been reported to be associated with increased false-negative FNA cytology [2,3,19–21].

It is imperative that clinicians minimize false-negative results. The Bethesda criteria for reporting thyroid cytopathology were developed by a committee at the National Cancer Institute meeting in 2007 to ameliorate this problem. Each cytopathological category is risk stratified for malignancy and corresponds to specific recommendations for patient management [11,12,22]. Though some

* Corresponding author at: Department of Surgery, Tulane University School of Medicine, 1430 Tulane Avenue, Room 8510 (Box SL-22), New Orleans, LA 70112, United States. Tel.: +1 504 988 7520; fax: +1 504 988 4762.
E-mail addresses: adeniwar@tulane.edu (A. Deniwar), chamblet@tulane.edu (C. Hambleton), tthethi@tulane.edu (T. Thethi), kmoroz@tulane.edu (K. Moroz), ekandil@tulane.edu (E. Kandil).

http://dx.doi.org/10.1016/j.prp.2015.02.005
0344-0338/

studies have concluded that the Bethesda criteria appropriately stratifies malignancy risk in thyroid nodules [16,23–25], controversy continues to exist regarding their accuracy and reliability in decision-making [26–29]. This study aims to determine the accuracy and outcome of US-guided FNA in the diagnosis of thyroid nodules using the current Bethesda criteria in surgical practice.

Methods

We performed a retrospective analysis of US-guided FNAs by a single surgeon during a four and half year period from January 2010 to June 2014 at Tulane University Medical Center (New Orleans, LA). Patients with thyroid nodules were included in the study if they had undergone FNA and surgical excision or follow up with serial US imaging. Patients who received surgery for large compressive goiters or Graves' disease with no concomitant nodules did not undergo preoperative FNA and were excluded from the current study. Patients with benign nodules on FNA biopsy did proceed to surgery if they were hyperthyroid or had compressive symptoms. In surgically excised lesions, FNA cytology was compared to surgical histopathology in order to analyze the accuracy of FNA using the Bethesda criteria reporting system.

The procedure for FNA biopsy is standardized in our practice as we described previously [30]. FNA cytology was reported according to the Bethesda criteria: non-diagnostic or unsatisfactory, benign, atypia of unknown significance or follicular lesion of unknown significance (AUS/FLUS), suspicious for follicular neoplasm (SFN), suspicious for malignancy, or malignant. Other cytological variants of Bethesda category IV as lesions suspicious for Hürthle cell neoplasm were enrolled under SFN category. All FNA samples suspicious for indeterminate FNA cytology (AUS/FLUS and SFN) were reviewed by a single cytopathologist (KM) who is a co-author on this paper.

FNA cytology was then compared to surgical histopathology in order to analyze accuracy of FNA using the Bethesda criteria reporting system. Non-surgically managed patients were followed up with serial US imaging as per the ATA guidelines.

The numbers of patients in each category of the Bethesda criteria were identified. Malignancy rates which were confirmed by final surgical pathology. Sensitivity, specificity, false-positive rate, and false-negative rate were calculated using Bethesda criteria as the gold standard. In addition, the Cochran–Armitage trend test was performed to assess the association between nodule size and rates of malignancy. A p value of <0.05 was considered statistically significant.

False-negative results were defined as nodules with benign FNA cytology and malignant surgical histology. False-positive results were defined as nodules with malignant cytology on FNA and benign surgical histology. For follicular neoplasms, false-negative results were defined as nodules with suspected follicular neoplasm on FNA and malignancy on surgical histology. False-positive results were defined as nodules with cytology suspected for follicular neoplasm on FNA and benign lesion on surgical histology. Microcarcinomas, incidentally found malignant lesions with size <1 cm, were not classified as true malignant lesions [31].

Results

611 patients were included in the study with 723 thyroid nodules. The study population had an average age of 53 years ± 12.8 years, and the majority were female patients (81%). Thyroid nodules were classified according to the Bethesda criteria: 506 benign (70%), 94 atypia of unknown significance or follicular lesion of unknown significance (AUS/FLUS) (13%), 43 suspicious for follicular neoplasm (SFN) (6%), 17 suspicious for malignancy (2%), 28 malignancy (4%), and 35 non-diagnostic (5%). 284 patients with 314 benign nodules on cytology were followed with serial US examinations. None of these patients have developed subsequent malignancy on repeat FNA if needed. 375 thyroid nodules in 297 patients were excised. After excluding microcarcinomas, malignancy was confirmed by surgical pathology in 15% (28 of 192 excised benign nodules), 34% (22 of 65 AUS/FLUS), 50% (21 of 42 SFN), 88% (15 of 17 suspicious for malignancy), 100% (28 of 28 malignant) and 39% (12 of 31 nodules with non-diagnostic FNA) (Table 1). In patients with benign FNA results, 28 of 192 (15%) surgically excised nodules were malignant. False-negative result rate decreased to 5.5% (28/506) when all the 506 benign nodules are included in the analysis. There were no false-positive results in our study. Patients with suspicious or malignant cytology on FNA corresponded to malignant thyroid diagnosis on final surgical pathology (43/45, 96%). Surgeon performed US-guided FNA had a sensitivity of 50%, specificity of 100%, positive predictive value of 100%, and negative predictive value of 85% (Table 2).

Papillary thyroid microcarcinoma was identified in 17 nodules, and follicular variant micropapillary thyroid carcinoma was identified in 11 nodules. On FNA cytopathology, 11 lesions were benign, 8 were AUS/FLUS, 5 were non-diagnostic, 3 were SFN, 1 was suspicious for malignancy, and 0 were malignant. When microcarcinomas were included, 22.0% of nodules were malignant and 78.0% were benign.

In size sub-analysis of surgical specimens, there was no linear trend for malignancy rate with larger nodule size, with 28% in <1 cm, 40% in 1–1.9 cm, 35% in 2–2.9 cm, and 40% in 3–3.9 cm, and 19% in ≥4 cm groups, respectively ($p = 0.89$).

The validity of FNA for diagnosis of follicular neoplasms was analyzed. When compared to nodules with malignant FNA and considering SFN nodules as a benign, it showed a sensitivity of 57%, specificity of 100%, positive predictive value (PPV) of 100%, and negative predictive value (NPV) of 89.4% (Table 3). While comparing it to benign FNA and considering SFN nodules as malignant yielded a sensitivity of 43%, specificity of 89%, PPV of 50% and NPV of 85% (Table 4).

Table 1
FNA pathology of thyroid nodules and malignancy rate in excised nodules.

FNA pathology	*N*	%	Excised nodules	Malignant nodules
Benign	506	69.99	192	28 (15%)
AUS/FLUS	94	13	65	22 (34%)
SFN	43	5.95	42	21 (50%)
Suspicious	17	2.35	17	15 (88%)
Malignant	28	3.87	28	28 (100%)
Non-diagnostic	35	4.84	31	12 (39%)
Total	723		375	

FNA = fine needle aspiration, AUS = atypia of undeterminate significance, FLUS = follicular lesion of undeterminate significance, SFN = suspicious for follicular neoplasm.

Table 2
Validity analysis: FNA pathology compared to surgical pathology.

	Surgical pathology	
	Benign	Malignant
Benign FNA	164	28
Malignant FNA	0	28
Sensitivity	50%	
Specificity	100%	
PPV	100%	
NPV	85%	

FNA = fine needle aspiration, PPV = positive predictive value, NPV = negative predictive value.