Online ISSN: 2682-2628
Print ISSN: 2682-261X

IJCBR

International Journal of Cancer and Biomedical Research

https://jcbr.journals.ekb.eg

Editor-in-chief
Prof. Mohamed Labib Salem, PhD

Accuracy of Bethesda system in the diagnosis of thyroid nodules: Utility of combined histopathological and radiological reporting systems

Maha Fawzy, Dina Harb and Heba Sheta

PUBLISHED BY
EACR EGYPTIAN ASSOCIAN FOR CANCER RESEARCH
Since 2014

INTERNATIONAL JOURNAL OF CANCER AND BIOMEDICAL RESEARCH (IJCBR)

RESEARCH ARTICLE

Accuracy of Bethesda system in the diagnosis of thyroid nodules: Utility of combined histopathological and radiological reporting systems

Maha Fawzy[1], Dina Harb[2] and Heba Sheta[1]

[1]Pathology Department, Faculty of Medicine, Mansoura University, Egypt
[2]Diagnostic and Interventional Radiology Department, Faculty of Medicine, Mansoura University, Egypt

ABSTRACT

Background: Nodular thyroid lesions are common among the population, especially women. Four-step processes were needed to evaluate thyroid nodules; starting with a history and clinical examination, followed by thyroid function test, then thyroid ultrasound (US) examination, and lastly US-guided aspiration cytology. The availability of thyroid imaging reporting and data systems (TIRADs) allows for an accurate clinic-pathological correlation. Moreover, the Bethesda system for reporting thyroid cytopathology was introduced to standardize the communication of fine-needle aspiration cytology (FNAC) interpretation between pathologists and clinicians. **Aim**: We aimed to assess the diagnostic utility of TIRAD and the Bethesda systems, separately and in combination, to explore the accuracy of the combination of these two standardized grading methods in the differential diagnosis of thyroid nodules. **Materials & Methods**: This is a retrospective study including patients with thyroid nodules. Clinical data of the patients were collected from their reports. TIRADs system was used to classify solitary thyroid lesions. All FNAC and cell block slides were revised. The Bethesda system was used to classify the thyroid lesions. All radiological and pathological results were correlated statistically. **Results**: By combining sensitivity and specificity of both TIRADs and Bethesda systems, sensitivity for detecting the nature of thyroid nodule was raised to 90.8% and the specificity was increased to 98%. **Conclusion**: Combination of both Bethesda systems and TIRADs increases the accuracy of evaluation of thyroid nodules to take the appropriate surgical decision.

Keywords: Bethesda; Sensitivity; Specificity; TIRADs; Thyroid nodules

Editor-in-Chief: Prof. M.L. Salem, PhD - Article DOI: 10. 10.21608/jcbr.2021.71741.1199

ARTICLE INFO

Article history

Received: April 9, 2021
Revised: June 25, 2021
Accepted: October 14, 2021

Correspondence to

Dr. Heba Sheta, MD,
Pathology Department,
Faculty of Medicine,
Mansoura University
Tel.: 00201102000054
Email: heba_sheta@yahoo.com

Copyright

©2021 Maha Fawzy, Dina Harb, Heba Sheta.. This is an Open Access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

INTRODUCTION

Nodular thyroid lesions are common among the population especially women (Mora-Guzmán et al., 2018). Ultrasonography has great importance in thyroid imaging in the clinical approach to nodular thyroid lesions. However, there are no sure signs by US that can give a cut-off point to differentiate benign from malignant lesions. Therefore, FNAC is a well-established technique for the differentiation of benign from the malignant lesion with high sensitivity and specificity (Arpana et al., 2018).

A radiological system like that of the breast was established in diagnosing thyroid nodules. Thyroid imaging reporting and data systems (TIRADS) is a classification system based on ultrasound features that were introduced to allow for a better selection of thyroid nodules undergoing FNAC for avoiding unnecessary procedures. This system also unifies language between radiologists and endocrinologists all over the world (Horvath et al., 2009). TIRADs have six categories for diagnosis and each of which has special characters and different risks of malignancy (Tessler et al., 2017).

Fine needle aspiration cytology is an important method in diagnosing thyroid nodules as it is safe, cost-effective, and can differentiate between benign and malignant lesions (Tayde et al. 2017). Bethesda system was developed to have a uniform reporting system for thyroid

FNAC that help effective communication among pathologists, clinicians, and radiologist (Cibas et al., 2009). Six categories were established in the Bethesda system and each of which is linked to a different risk of malignancy and needs certain clinical management (Alshaikh et al., 2018). Clinical management concerning the Bethesda system ranges from clinical and sonographic follow-up in category II to near-total thyroidectomy or lobectomy in categories V and VI (Cibas et al., 2009).

So that the present study aimed to assess the diagnostic utility of TIRAD and the Bethesda systems separately and in combination to explore the accuracy of the combination of these two standardized grading methods in the differential diagnosis of thyroid nodules.

MATERIALS AND METHODS

A retrospective type of study during the period, from 1 January 2017 to 1 January 2018 was conducted in the Oncology center at the Faculty of Medicine, Mansoura. This study was approved by our Institutional Review Board (IRB code no: R.19.03448).

Inclusion and exclusion criteria

All patients having clinically palpable thyroid swellings, irrespective of their age and sex, and who underwent ultrasound imaging with subsequent FNAC and surgery for nodule removal were included in this study. However, cases who did not do further surgery after FNAC were excluded.

Data collection

Clinical data of the patient were collected from their reports as regard age, sex, and clinical presentation. All cases were assessed radiologically and FNAC was collected under US guidance. Assessment of thyroid nodules was done using the TIRAD system. This system provides standardized recommendations despite the expertise of the reader. Many studies supported that the TIRAD system decreased interobserver variability (Hoang et al., 2018; Itani et al., 2019).

TRIADS system was used to classify solitary Thyroid lesions into six categories. These include TI-RADS 1: normal thyroid gland. No focal lesion. TI-RADS 2: benign nodules. Noticeably benign pattern. TI-RADS 3: probably benign nodules. TI-RADS 4: 4a–undetermined nodules, 4b–suspicious nodules, 4c–highly suspicious nodules TI-RADS 5: Probably malignant nodules and TI-RADS 6: biopsy-proven malignancy (Sánchez, 2014). Radiological data according to the number of nodules, site, and size were revised. Regarding cases presented with more than one nodule, we commented on the more suspicious one.

All cases in the current study undergo US-guided FNAC. TRIADS 1cases in this study undergo further FNA and surgery due to Nodule size consideration as the size of TIRAD1 cases ranged from 1.3 to 11 cm. Most guidelines recommend FNA for nodules that are larger than 10 mm (Frates et al., 2005; American Thyroid Association (ATA) Guidelines Taskforce on Thyroid Nodules and Differentiated Thyroid Cancer et al., 2009; Luster et al., 2019).

Cytological and histopathological evaluation

In our cytopathology lab, cytology slides were stained by hematoxylin and eosin stains, then the residual hemorrhagic aspiration in the syringe and needle was fixed in 10% formalin and paraffin-embedded (cell block). The histological sections were examined as a complementary diagnostic tool. This because cytologic-histologic correlation is a reliable method for determining the outcome of FNAC diagnosis and is proved to reduce the rate of unsatisfactory samples and increase the accuracy of diagnosis (Cristo et al., 2016).

All FNAC and cell block slides were revised. Six categories of the Bethesda system were used to classify the thyroid lesions. These categories include (Bethesda I: non-diagnostic, Bethesda II: benign, Bethesda III: atypia of undetermined significance (AUS)/follicular lesion of undetermined significance (FLUS), Bethesda IV: follicular neoplasm/suspicious for follicular neoplasm, Bethesda V: suspicious for malignancy, and Bethesda VI: malignant). For a thyroid FNA specimen to be satisfactory for evaluation, at least six groups of benign follicular cells are required, and each group is composed of at least 10 cells (Alshaikh et al., 2018). Smears showing atypical cells were never considered inadequate regardless of cellularity.

H&E slides of the thyroid nodule after surgery were also revised.

Statistics

Sensitivity, specificity, positive predictive value, and negative predictive value were calculated for both TIRADS and Bethesda systems separately and in combination. The pathological results after thyroid nodule surgery were taken as the criterion standard for evaluating thyroid nodules. All radiological and pathological results including (FNAC and surgical specimens) were correlated statistically. Data were analyzed using the Statistical Package of Social Science (SPSS) program for Windows (Standard version 21). The normality of data was first tested with a one-sample Kolmogorov-Smirnov test. Qualitative data were described using the number and the percent. Continuous variables were presented as mean±SD (standard deviation).

RESULTS

Clinicopathologic data of the studied cases are summarized in table 1. A total number of 190 cases presented with thyroid nodules were studied. There was female predominance including 158 cases (83.2%). The age of presentation ranged from 19 years to 80 years old with a mean age of 45.67 years. Cases presented with one thyroid nodule were 126 cases (66.3%), while 35 cases were presented with multiple nodules (more than two) and only 29 (15.3%) were presented with two nodules. As regard nodule site, 71 cases (37.4%) were in the right lobe, 62 cases (32.6%) were in the left lobe, and 35 cases (30.0%) were in both lobes. The mean length of the nodule was 3.27±1.4 and the mean width was 3.73±1.3.

TIRAD system's results were shown in table 2. Benign results were found in 35 patients (18.4%) (TIRADS 1) & probably benign results were found in (79 & 31 cases) (41.6 & 16.3%) (TIRADS 2 and TIRADS 3 respectively), indeterminate results were found in 35 cases (TIRADS 4) (18.4%), and 10 were suspicious for malignancy (5.3%) (TIRADS 5). The percentage of a malignant FNAC (Bethesda Class V and Class VI) in TIRADS 2, 3, 4, and 5 classes were 16.5, 32.3, 48.6, and 100%, respectively. The percentage of a benign FNAC (Bethesda Class II) in TIRADS category 1 & 2 was 94.3, 83.5 % respectively, while for TIRADS 3, 4, and 5 classes were 67.7, 51.4, and 0%, respectively.

As regards Bethesda results (table 3), 166 cases (86.3%) were adequate. Ninety-seven cases (53.2%) were Bethesda II (Figure 1a), of which 94 cases (96.9%) were benign and 3 cases (3.1%) were malignant after surgery. Most of these cases were diagnosed as colloid nodular goiter (78 cases: 72.3%) (Figure 1b).

Bethesda III was diagnosed in 16 cases (8.4%) (Figure 1c), of which 11 cases were diagnosed benign; mostly follicular adenoma and hyperplastic nodule (Fig. 1d) and 5 cases were malignant by biopsy. Twenty-nine cases (13.2%) were Bethesda IV (Figure 2a), of which 18 cases (62.1%) were malignant; papillary thyroid carcinoma and 11 cases (37.9%) were benign mostly colloid nodular goiter and Hurthle cell adenoma (Figure 2b).

Bethesda V was detected in 12 cases (6.3%) of these cases (Figure 2c), 11 cases (91.7%) were malignant predominantly papillary thyroid carcinoma (Figure 2d) and medullary carcinoma, only one case was benign, and 10 cases (5.3%) were Bethesda VI. All these cases were malignant with 8 cases of papillary thyroid carcinoma. As regards surgical biopsies of the thyroid nodules (Table 1), 138 (72.6%) cases were benign; most of them were colloid nodular goiter. Fifty-two cases (26.8%) were malignant; most of them were papillary thyroid carcinoma.

Table 4 assessed the diagnostic accuracy of TIRADs and Bethesda in the diagnosis of thyroid nodules. The sensitivity of TIRADs was 51.9% and the specificity was 86.9% and its accuracy was 77.4%., however, Bethesda sensitivity was 82.9% and 89.7% specificity with 87.8% accuracy.

By combining the sensitivity and the specificity of both TIRADs and Bethesda systems (table 4), sensitivity for detecting the nature of thyroid nodule raised to be 90.8%, and specificity increased to 98%.

DISCUSSION

A four-step process was needed to evaluate and comment on thyroid nodules starting with history with clinical examination and followed

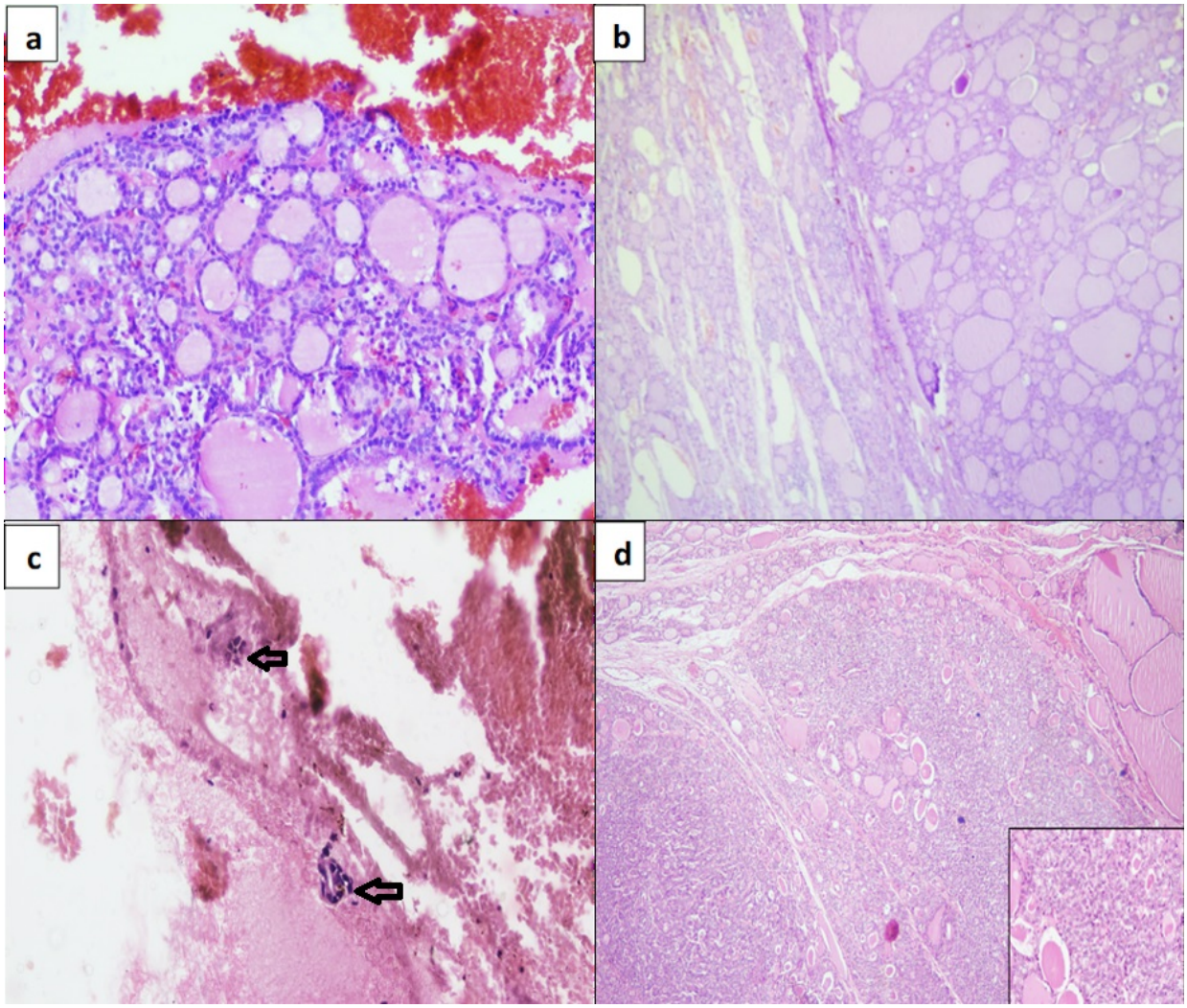

Figure 1. Cytomorphological findings seen in Bethesda categories II & III and histopathological sections of excision specimens a) Smear of thyroid nodule that shows variable sized thyroid follicles lined by bland-looking follicular cells and filled with colloid (Bethesda category II) (H&E x200). b) Excision of the previous case that shows dominant colloid nodule in a background of colloid nodular goiter (H&E x40). c) Hypocellular smear of thyroid nodule that shows two groups of follicular cells arranged in microfollicular pattern (arrows) with absent colloid (Bethesda category III) (H&E x200). d) Excision of the previous case that shows hyperplastic nodule of multinodular goiter (H&E x40), with inset that focus on hyperplastic thyroid follicles (H&E x400).

by thyroid function test, thyroid US examination, and lastly US-guided aspiration cytology (Singaporewalla et al., 2017). This work aimed to assess the diagnostic utility of TIRAD and the Bethesda systems separately and in combination to explore the accuracy of the combination of these two standardized grading methods in the differential diagnosis of thyroid nodules.

In the current study, there was female predominance (83.2%). This agreed with almost all the reviewed studies (Nandedkar et al., 2018; Jabar et al., 2019). The mean age of presentation was 45.67±12.73 in our study. This comes in agreement with the results of some reports (Grace, 2017; Acar et al., 2017), however, other studies found younger age; 37.6 years (Nandedkar et al., 2018), and others found older age (51.8 years) (Mora-Guzmán et al., 2018). This difference may be due to different sample sizes and different ages at which the patients presented to the health care units.

This study found that the risk of malignancy for different TIRADS categories was 16.5% (TIRADS 2), 32.3% (TIRADS 3), 48.46% (TIRADS 4), and 100% (TIRADS 5). Horvath et al., 2009 found that the rate of malignancy in cases classified from TIRADS 2 to 5 were 0%, 14.1%, 45%, and 89.6% respectively.

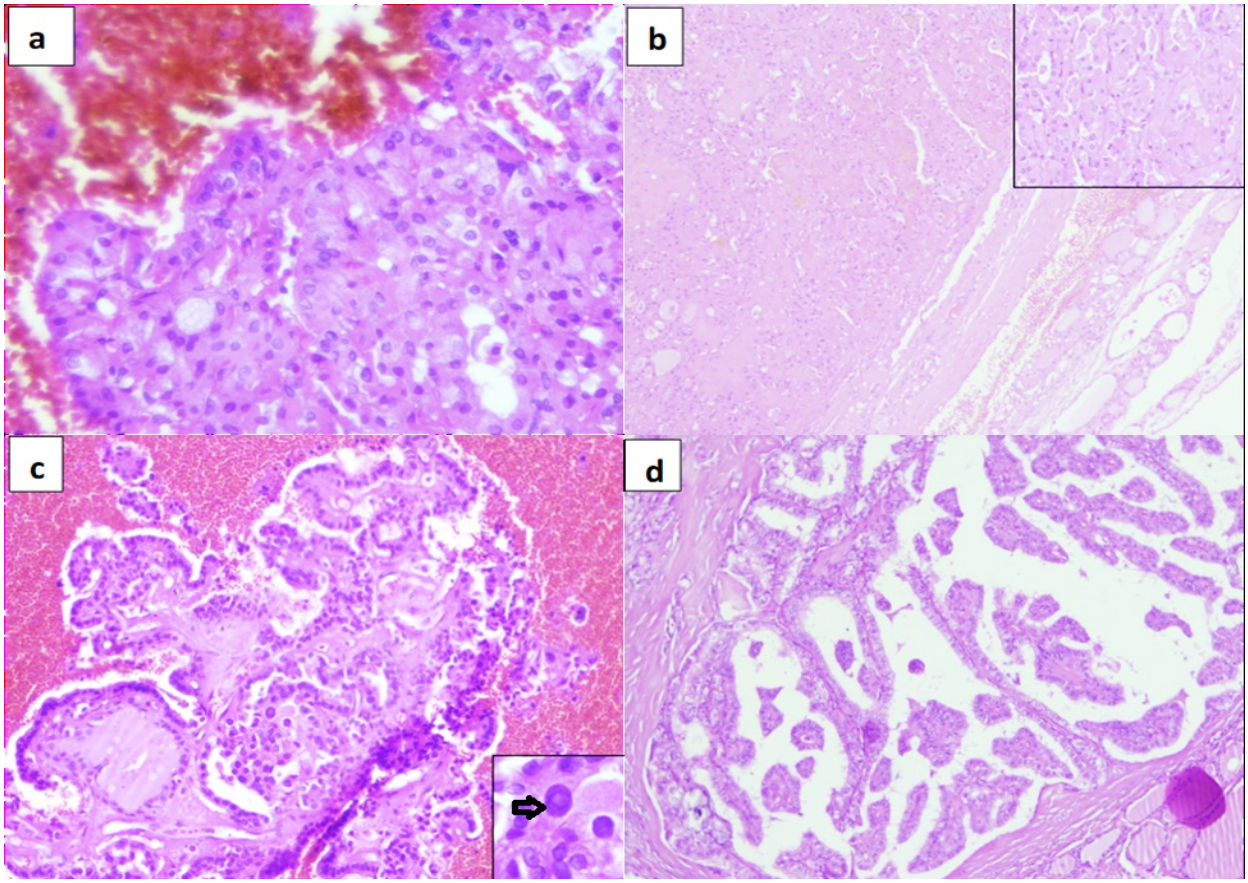

Figure 2. Cytomorphological findings seen in Bethesda categories IV & V and histopathological sections of excision specimens a) Highly cellular smear of thyroid nodule formed of Hurthle cells with abundant eosinophilic cytoplasm arranged mainly in microfollicular pattern with absent colloid (Bethesda category IV) (H&E x200). b) Excision of the previous case that shows Hurthle cell adenoma (H&E x40), with inset to illustrate cytologic features of Hurthle cells (H&E x400). C) Hypercellular smear of thyroid nodule arranged mainly in papillary structures lined by atypical follicular cells that shows focal nuclear features of papillary thyroid carcinoma as nuclear inclusion (Bethesda category V) (H&E x200), with inset to illustrate nuclear inclusion (arrow) (H&E x400). D) Excision of the previous case that shows papillary thyroid carcinoma (H&E x200).

In addition, the risk of malignancy reported by Kwak et al., 2011 for TIRADS 2, 3, 4a, 4b, 4c, and 5 were 0%, 1.1%, 1.7%, 3.3%, 9.2%, and 87.5% respectively.

Jabar et al., 2019 worked on 127 cases. None of them were TIRADs 1 nor 5. Also, they found that the percentage of TIRADs 2, 3, and 4 was 21.2%, 33.8%, and 32.2% respectively. These results were different from our results in the current study. As we reported 22.6% and 4.2% to be TIRADs 1 and TIRADs 5 respectively. As regards TIRADS 2, 3, and 4 results in our study, there were 41.6%, 15.8%, and 15.8% respectively. This difference may be due to differences in sample size and inter-observer variability in nodule assessment by radiology.

Periakaruppan et al., 2018 worked on 184 cases and most of them (63.6%) were TIRADs 2 with a comparison with Bethesda system including 90% were Bethesda II and none of them were malignant after surgery. Our results were different from Periakaruppan et al. Results found that 41.6% to be TIRADs 2, of the 60.8% were Bethesda II, and 16.5% were malignant. This discrepancy can be explained by inter-observer variability in nodule assessment by radiology. Moreover, Periakaruppan et al. reported 5% of cases to be TIRADs 5; of which 78% were malignant compared to 100% in our result.

When comparing Horvath et al. 2009, results with our result, it is found that lower insensitivity in malignancy prediction by US examination of different thyroid nodules (51.9% compared to 88%), but higher in specificity in the current study (86.9% compared to 49%).

Table 1. Clinico-pathologic characteristics of the studied cases.

Clinical Characteristics	Study Groups (n=190)
Gender	
Male	32 (16.8%)
Female	158 (83.2%)
Age (years)	
Mean ± SD	45.67±12.73
Range	19-80
Number on nodules	126 (66.3%)
One	29 (15.3%)
Two	35 (18.4%)
Multiple	
Site	71 (37.4%)
Right lobe	62 (32.6%)
Left lobe	57 (30.0%)
Both	
Size (length)	3.27±1.4
Mean ± SD	
Size (width)	2.73±1.3
Mean ± SD	
TIRADs	
1	35 (18.4%)
2	79 (41.6%)
3	31 (16.3%)
4	35 (18.4%)
5	10 (5.3%)
Bethesda	
I	26 (13.7%)
II	97 (51.1%)
III	16 (8.4%)
IV	29 (15.3%)
V	12 (6.3%)
VI	10 (5.3%)
Surgery	
Colloid nodular goiter	99 (52.1%)
Follicular adenoma	17 (8.9%)
Hyperplastic nodule	12 (6.3%)
Hurthle cell adenoma	4 (2.1%)
Controlled toxic goiter	1 (0.5%)
Papillary thyroid	34 (17.9%)
carcinoma	3 (1.6%)
Papillary microcarcinoma	6 (3.2%)
Medullary carcinoma	1 (0.5%)
Hashimoto thyroiditis	1 (0.5%)
Granulomatous	3 (1.6%)
thyroiditis	2 (1.1%)
Lymphocytic thyroiditis	3 (1.6%)
Anaplastic carcinoma	2 (1.1%)
Poorly differentiated	1 (0.5%)
carcinoma	1 (0.5%)
Follicular carcinoma	
Non-Hodgkin lymphoma	
Minimally invasive	
follicular carcinoma	
Histopathology	
Malignant	52 (27.4%)
Benign	138 (72.6%)

SD: standard deviation; TIRADs: Thyroid imaging reporting and data systems.

Our results had been shown higher positive predictive value (60% compared to 49%) and lower negative predictive value (82.7% compared to 88%, respectively) of TIRADS score malignancy prediction as well. Overall, this study found an accuracy of 77.4% that is higher than Mohandas et al. 2019 who reported 69.2%.

According to Bethesda assessment in our study and inadequate Bethesda I category, it was detected in 13.7% of our studied cases. There was a wide range of incidence of Bethesda I category among previous studies that ranged from 1.2 % (Mondal et al., 2013) to as high as 35.3% (Park et al., 2014). It may be attributed to technical and interpretative factors (Raab et al., 2006).

Benign category II was diagnosed in 53.2% of our studied cases. Among the reviewed previous studies, this category was predominant with different percentages that ranged from 32.9% (Acar et al., 2017), 49.6% (Mora-Guzmán et al., 2018), 59% (Jo et al., 2010) up to 87.5% (Mondal et al., 2013). This difference depends on whether the institute where the study was done by a tertiary care center, where patients come only on a referral basis and, hence, is not exactly representative of the general population or primary care center that was representative of the general population with a high percentage of benignancy. The risk of malignancy of category II was 3.1% in the current study. Malignant risk of this category was lower among most of the reviewed studies and was ranged from 0% (Garg et al., 2015), 0.3% (Yassa et al., 2007), 1.1% (Jo et al., 2010), 3.1% (Mufti and Molah, 2012), 4.5% (Mondal et al., 2013) and 5.6% (Park et al., 2014).

Bethesda III (AUS category) was diagnosed in 8.4% of the studied cases with a malignant risk of 31.2%. Mufti et al., 2012 and Salillaas et al., 2015 reported 50% malignant risk. Meta-analysis was performed including 51 articles and a total of 145,928 FNA specimens was done (Straccia et al., 2015). The malignant risk of the AUS group in this analysis was 27% that is slightly lower than our result. This difference may be explained by that not all AUS/FLUS cases were submitted to surgery (surgeon's selection bias).

Table 2. TIRADs results of the studied cases.

TIRADs	No of cases (%)	Bethesda	No of cases (%)	Type	No of cases (%)	Pathology	No (%)
1	35(18.4)	I II III IV V	5(14.3) 22(62.9) 4(11.4) 3(8.6) 1(2.9)	Malignant Benign	2(5.7) 33(94.3)	Colloid nodular goiter Follicular adenoma Hyperplastic nodule Hurthle cell adenoma Controlled toxic goiter Papillary thyroid carcinoma Papillary microcarcinoma Hashimoto thyroiditis Lymphocytic thyroiditis	17(48.6) 7(20) 3(8.6) 3(8.6) 1(2.9) 1(2.9) 1(2.9) 1(2.9) 1(2.9)
2	79(41.6)	I II III IV V VI	10(12.7 48(60.8) 3(3.8) 11(13.9) 2(2.5) 5(6.3)	Malignant Benign	13(16.5) 66(83.5)	Colloid nodular goiter Follicular adenoma Hyperplastic nodule Hurthle cell adenoma Papillary thyroid carcinoma Papillary microcarcinoma Medullary carcinoma Granulomatous thyroiditis Anaplastic carcinoma Follicular carcinoma Minimally invasive follicular carcinoma	53(67.1) 6(7.6) 5(6.3) 1(1.3) 8(10.1) 1(1.3) 1(1.3) 1(1.3) 1(1.3) 1(1.3) 1(1.3)
3	31(16.3)	I II III IV VI	5(16.1) 16(51.6) 1(3.2) 8(25.8) 1(3.2)	Malignant Benign	10(32.3) 21(67.7)	Colloid nodular goiter Follicular adenoma Hyperplastic nodule Papillary thyroid carcinoma Medullary carcinoma Lymphocytic thyroiditis Anaplastic carcinoma.	16(51.6) 1(3.2) 3(9.7) 8(25.8) 1(3.2) 1(3.2) 1(3.2)
4	35(18.4)	I II III IV V VI	4(11.4) 10(28.6) 8(22.9) 5(14.3) 7(20) 1(2.9)	Malignant Benign	17(48.6) 18(51.4)	Colloid nodular goiter Follicular adenoma Hyperplastic nodule Papillary thyroid carcinoma Medullary carcinoma Lymphocytic thyroiditis Poorly differentiated carcinoma Follicular carcinoma	13(37.1) 3(3.6) 1(2.9) 10(28.6) 3(8.6) 1(2.9) 3(8.6) 1(2.9)
5	10(5.3)	I II IV V VI	2(20) 1(10) 2(20) 2(20) 3(30)	Malignant Benign	10(100) 0(0)	Papillary thyroid carcinoma Papillary microcarcinoma Medullary carcinoma Non-Hodgkin lymphoma	7(70) 1(10) 1(10) 1(10)

TIRADs: Thyroid imaging reporting and data systems.

Regarding Bethesda IV, it was diagnosed in 15.3% of cases with a malignant risk of 62.1%. The range of distribution of Bethesda IV is wide and ranged from 0.16% (Bohacek et al., 2012), 16.5% (Sinna and Ezzat, 2012), and up to 88.6% (Bongiovanni et al., 2014). Among previous studies, the malignant risk among this category was ranged from 11.11% (Yoo et al., 2013), 42.86% (Onder et al., 2014), 61.29% (Kim et al., 2014), and 65.71% (Park et al., 2014). This is due to the variable distribution of thyroid diseases among different populations.

As regards Bethesda categories V and VI, their frequencies were 6.3% and 5.3% of our studied cases. This was close to Acar et al. who found 7.5% and 6.6% of both categories respectively and this is close to Arul et al. who also found 5.3% and 6.3% respectively. On the other hand, our result was higher than some studies (Mufti and Molah, 2012; Williams et al., 2013) and lower than others (Park et al., 2014). Malignant risk of Bethesda categories V and VI of our study was 91.7% and 100% respectively.

Table 3. Bethesda results of the studied cases.

Bethesda	No of cases (%)	TIRADs	No of cases (%)	Type	No of cases (%)	Pathology	No (%)
I	26(13.7)	1 2 3 4 5	5(19.2) 10(38.5) 5(19.2) 4(15.4) 2(7.7)	Malignant Benign	5(19.2) 21(80.8)	Colloid nodular goiter Follicular adenoma Papillary thyroid carcinoma Papillary microcarcinoma Lymphocytic thyroiditis Non-Hodgkin lymphoma	15(57.7) 4(15.4) 3(11.5) 1(3.8) 2(7.7) 1(3.8)
II	97(51.1)	1 2 3 4 5	22(22.7) 48(49.5) 16(16.5) 10(10.3) 1(1)	Malignant Benign	3(3.1) 94(96.9)	Colloid nodular goiter Follicular adenoma Hyperplastic nodule Hurthle cell adenoma Controlled toxic goiter Papillary thyroid carcinoma Papillary microcarcinoma Hashimoto thyroiditis Granulomatous thyroiditis Lymphocytic thyroiditis	78(80.4) 6(6.2) 5(5.2) 1(1) 1(1) 2(2.1) 1(1) 1(1) 1(1) 1(1)
III	16(8.4)	1 2 3 4 5	4(25) 3(18.8) 1(6.2) 8(50) 0(0)	Malignant Benign	5(31.2) 11(68.8)	Colloid nodular goiter Follicular adenoma Hyperplastic nodule Papillary thyroid carcinoma Medullary carcinoma Poorly differentiated carcinoma	2(12.5) 5(31.2) 4(25) 3(18.8) 1(6.2) 1(6.2)
IV	29(15.3)	1 2 3 4 5	3(10.3) 11(37.9) 8(27.6) 5(17.2) 2(6.9)	Malignant Benign	18(62.1) 11(37.9)	Colloid nodular goiter Follicular adenoma Hyperplastic nodule Hurthle cell adenoma Papillary thyroid carcinoma Medullary carcinoma Poorly differentiated carcinoma Follicular carcinoma Minimally invasive follicular carcinoma	4(13.8) 2(6.9) 2(6.9) 3(10.3) 12(41.4) 1(3.4) 1(3.4) 1(3.4) 2(6.9) 1(3.4)
V	12(6.3)	1 2 4 5	1(8.3) 2(16.7) 7(58.3) 2(16.7)	Malignant Benign	11(91.7) 1(8.3)	 Hyperplastic nodule Papillary thyroid carcinoma Medullary carcinoma Poorly differentiated carcinoma	 1(8.3) 6(50) 4(33.3) 1(8.3)
VI	10(5.3)	2 3 4 5	5(50) 1(10) 1(10) 3(30)	Malignant Benign	10(100) 0(0)	Papillary thyroid carcinoma Anaplastic carcinoma	8(80) 2(20)

TIRADs: Thyroid imaging reporting and data systems.

As regards category V, our result was higher than many other studies' results (Arul and Masilamani, 2015; Nandedkar et al., 2018). Regarding the malignant risk of category VI, our result agreed with most of the previous studies (Grace, 2017).

Regarding surgical biopsy specimen results in our study, 73.2% of cases were benign and 26.8% were malignant. These results were quite like Jabar's results et al., 2019. They found that the number of benign cases after surgery (81.8%) is more than malignant cases (18.1%). Similar results are reported by Grace Dy et al., 2017 who found 66.4% of cases were benign. This difference in the percentage of malignant cases could be explained by sample size difference and racial issues that regulate neoplastic cell behavior.

As regards the assessment of Bethesda system sensitivity and specificity, Naz et al., 2014 reported 66.3% sensitivity, 85.1% specificity, and 80.3% accuracy.

Table 4. Diagnostic accuracy of TIRADs and Bethesda system compared to histopathological result in diagnosing of thyroid nodules

	True positive	True negative		False positive	False negative
TIRADs	27	120		18	25
Bethesda	39	105		12	8
	Sensitivity	**Specificity**	**PPV**	**NPV**	**Accuracy**
TIRADs	51.9%	86.9%	60%	82.7%	77.4%
Bethesda	82.9%	89.7%	76.5%	92.9%	87.8%
Combined	90.8%	98%	-	-	-

TIRADs: Thyroid imaging reporting and data systems; PPV: positive predictive value; NPV: Negative predictive value.

Table 5. Frequency and Risk of Malignancy for 6 Categories of The Bethesda System for Reporting Thyroid Cytopathology in Western and Asian Series in comparison to the current study [37]

FNAC category	Western series (n=22)	Asian series (n=16)	The present study
I			
Frequency	11.9 (9.1-14.7)	12.6 (6.7-18.5)	13.7
ROM	13.2 (9.6-16.7)	26.5 (16.4-36.6)	5
II			
Frequency	64.2 (60.0-68.4)	59.8 (51.6-67.9)	51.1
ROM	4.1 (2.8-5.4)	13.8 (9.0-18.6)	3.1
III			
Frequency	7.7 (5.1-10.2)	8.4 (5.5-11.4)	8.4
ROM	21.5 (17.0-26.0	45.0 (33.4-56.5	31.2
IV			
Frequency	7.9 (5.7-10.1)	3.5 (1.9-5.1)	15.3
ROM	27.3 (24.4-30.2)	32.8 (27.5-38.1)	62.1
V			
Frequency	3.3 (2.6-4.1)	4.3 (2.6-6.1)	6.3
ROM	75.1 (69.8-80.4)	88.1 (82.8-93.4	91.7
VI			
Frequency	4.9 (3.8-6.0)	10.9 (7.1-14.7)	5.3
ROM	99.2 (98.8-99.5)	98.6 (97.6-99.5	100

FNAC: fine-needle aspiration cytology; ROM: Risk of Malignancy

This was quite different from our study that found 77.8% sensitivity, 88.9% specificity, and 85.9% accuracy. Naz worked only on 61 cases that may underlie this mild difference in results based on the sample size difference. On the other hand, our result was quite like Aravinthan et al. 2007 (80.2% sensitivity and 87.2% specificity).

Meta-analysis was done to investigate the differences in diagnosis frequency, resection rate, and risk of malignancy (ROM) between Western (i.e., American, and European) and Asian cytopathology practices. This study included a total of 38 studies with 145,066 fine-needle aspirations. Compared with Asian practice, the Western series had a significantly lower ROM in most of Bethesda categories. Focusing on indeterminate nodules, the ROM was significantly lower in the Western series (25.4% vs 41.9%; P = .002) compared with those in the Asian series (Vuong et al., 2020). Table (5) demonstrates the results of the current study in comparison to the previously mentioned meta-analysis. These heterogeneities might stem from diverse clinical approaches among the institutions, different prevalence of thyroid cancer among geographic areas, or differences in diagnostic concepts among institutions and individual pathologists.

A high-quality service requires close co-operation between biomedical/healthcare scientists, pathologists, radiologists/ sonographers, and clinicians managing the patients so that appropriate procedures are set up, implemented, and monitored (Porterfield et al., 2008). According to the British Thyroid

Association Guidelines for the Management of Thyroid Cancer, a multidisciplinary discussion is recommended when the clinical, cytological, and ultrasound findings are inconsistent. Moreover, the follow-up of thyroid nodules should depend upon integration between the initial US appearances and associated cytology (Perros et al., 2014). Moreover, many previous studies demonstrated the role of TIRAD in reaching a diagnosis of cytologically indeterminate thyroid nodules (Bethesda category III) (Gao et al., 2017; Grani et al., 2018).

The previous discussion reveals how the TIRADS and Bethesda systems were integrated and used for clinical decision-making. So, for this reason, we assessed the sensitivity and specificity of TIRADs and Bethesda statistically together and we found that both sensitivity and specificity had been raised to 90.8%, 98% respectively. To the best of our knowledge, there is only one previous study that investigates this combination (Tan et al., 2019). Tan et al. concluded that the combination of high-resolution ultrasonography TI-RADS classification and US-FNAC (Bethesda classification) can improve the accuracy of malignant thyroid nodules diagnosis.

So, this study recommends assessing any thyroid nodule with both TIRADs and Bethesda to decrease unnecessary surgery and not to miss malignant nodules.

ABBREVIATIONS

AUS	atypia of undetermined significance
FLUS	follicular lesion of undetermined significance
FNAC	fine-needle aspiration cytology
H&E	Hematoxylin and eosin.
ROM	risk of malignancy
SD	standard deviation
TIRADs	Thyroid imaging reporting and data systems
US	Ultrasound.

AUTHORS' CONTRIBUTION

Maha Fawzy: Designed the experiment, collected the results. Dina Harb: Collected the results, did the statistical analysis.Heba Sheta: suggested the study idea, collected the results. All authors wrote and approved the manuscript.

CONFLICT OF INTEREST

All authors declare no conflicts of interest.

FUNDING

No fund was received for this work.

REFERENCES

Acar Y, Doğan L, Güven HE, Aksel B, Karaman N, Özaslan C & Gülçelik MA (2017). Bethesda Made It Clearer: A Review of 542 Patients in a Single Institution. Oncol Res Treat.;40(5):277-280.

Alshaikh s.,Harb Z., Aljufairi E. & Almahari a. (2018). Classification of thyroid fine needle aspiration cytology into Bethesda categories: An institutional experience and review of the literature. Cytojournal., 15:4.

American Thyroid Association (ATA) Guidelines Taskforce on Thyroid Nodules and Differentiated Thyroid Cancer, Cooper DS, Doherty GM, Haugen BR, Kloos RT, Lee SL, Mandel SJ, Mazzaferri EL, McIver B, Pacini F, Schlumberger M, Sherman SI, Steward DL & Tuttle RM. (2009). Revised American Thyroid Association management guidelines for patients with thyroid nodules and differentiated thyroid cancer. Thyroid.;19(11):1167-1214.

Aravinthan T, Banagala AS & Gamage KJ (2007). Use of fine needle aspiration cytology on thyroid lumps. Galle Med J ;12:25-29.

Arpana, Om Biju Panta, Ghanashyam Gurung & Pradhan S (2018). Ultrasound Findings in Thyroid Nodules: A Radio-Cytopathologic Correlation. J Med Ultrasound; 26(2):90-93.

Arul P & Masilamani (2015). A correlative study of solitary thyroid nodules using the bethesda system for reporting thyroid cytopathology. S. J Cancer Res Ther.;11(3):617-622.

Bohacek L, Milas M, Mitchell J, Siperstein A & Berber E. Diagnostic accuracy of surgeon-performed ultrasound-guided fine-needle aspiration of thyroid nodules. Ann Surg Oncol. 2012 Jan;19(1):45-51.

Bongiovanni M, Piana S, Spitale A, Valli R, Carlinfante G & Gardini G. (2014). Comparison of the diagnostic accuracy of thyroid fine-needle aspiration in follicular-patterned lesions using a 5-tiered and a 6-tiered diagnostic system: a double-blind study of 140 cases with histological confirmation. Diagn Cytopathol. ;42(9):744-750.

Cibas E. & Ali S. (2009). The Bethesda system for reporting thyriod cytopathology. Am J pathological,132:658-665.

Cristo AP, Goldstein HF, Faccin CS, Maia AL & Graudenz MS. (2016). Increasing diagnostic effectiveness of thyroid nodule evaluation by implementation of cell block preparation in routine US-FNA analysis. Arch Endocrinol Metab.;60(4):367-373.

Frates MC, Benson CB, Charboneau JW, Cibas ES, Clark OH, Coleman BG, Cronan JJ, Doubilet PM, Evans DB, Goellner JR, Hay ID, Hertzberg BS, Intenzo CM, Jeffrey RB, Langer JE, Larsen PR, Mandel SJ, Middleton WD, Reading CC, Sherman SI, Tessler FN & Society of Radiologists in Ultrasound. (2005). Management of thyroid nodules detected at US: Society of Radiologists in Ultrasound consensus conference statement. Society of Radiologists in Ultrasound.Radiology;237(3):794-800.

Gao LY, Wang Y, Jiang YX, Yang X, Liu RY, Xi XH, Zhu SL, Zhao RN, Lai XJ, Zhang XY & Zhang B. (2017). Ultrasound is helpful to differentiate Bethesda class III thyroid nodules: A PRISMA-compliant systematic review and meta-analysis. Medicine (Baltimore);96(16):6564

Garg S, Desai NJ, Mehta D & Vaishnav M. (2015). To Establish Bethesda System for Diagnosis of Thyroid Nodules on the Basis of Fnac with Histopathological Correlation. J Clin Diagn Res; 9(12):17-21.

Grace Dy J (2017). Thyroid Imaging Reporting and Data System (TIRADS) in Stratifying Risk of Thyroid Malignancy at The Medical City. VOL. 32 NO. 2 (2017): NOVEMBER ISSUE.

Grani G, Lamartina L, Cantisani V, Maranghi M, Lucia P & Durante C. (2018). Interobserver agreement of various thyroid imaging reporting and data systems. Endocr Connect.;7(1):1-7.

Hoang JK, Middleton WD, Farjat AE, Teefey SA, Abinanti N, Boschini FJ, Bronner AJ, Dahiya N, Hertzberg BS, Newman JR, Scanga D, Vogler RC & Tessler FN. (2018). Interobserver Variability of Sonographic Features Used in the American College of Radiology Thyroid Imaging Reporting and Data System. AJR Am J Roentgenol.;211(1):162-167.

Horvath E, Majlis S, Rossi R, Franco C, Niedmann JP & Castro A (2009). An ultrasonogram reporting system for thyroid nodules stratifying cancer risk for clinical management. J Clin Endocrinol Metab ; 94:1748–51.

Itani M, Assaker R, Moshiri M, Dubinsky TJ & Dighe MK. (2019). Inter-observer Variability in the American College of Radiology Thyroid Imaging Reporting and Data System: In-Depth Analysis and Areas for Improvement. Ultrasound Med Biol.;45(2):461-470.

Jabar ASS, Koteshwara P & Andrade J. (2019). Diagnostic reliability of the Thyroid Imaging Reporting and Data System (TI-RADS) in routine practice. Pol J Radiol. ;84: 274-280.

Jo VY, Stelow EB, Dustin SM & Hanley KZ. (2010). Malignancy risk for fine-needle aspiration of thyroid lesions according to the Bethesda system for reporting thyroid cytopathology. Am J Clin Pathol ;134:450-456.

Kim GR, Yoon JH, Kim EK, Moon HJ & Kwak JY (2014). Benign Aspirates on Follow-Up FNA May Be Enough in Patients with Initial Atypia of Undetermined Significance/Follicular Lesion of Undetermined Significance. Int J Endocrinol.;2014:354612.

Kwak JY, Han KH, Yoon JH, Moon HJ, Son EJ, Park SH, Jung HK, Choi JS, Kim BM & Kim EK. (2011). Thyroid Imaging Reporting and Data System for Ultrasound Features of Nodules: A Step in Establishing Better Stratification of Cancer Risk. Radiology; 260:892-899.

Luster M, Aktolun C, Amendoeira I, Barczyński M, Bible KC, Duntas LH, Elisei R, Handkiewicz-Junak D, Hoffmann M, Jarząb B, Leenhardt L, Musholt TJ, Newbold K, Nixon IJ, Smit J, Sobrinho-Simões M, Sosa JA, Tuttle RM, Verburg FA, Wartofsky L & Führer D. (2019). European Perspective on 2015 American Thyroid Association Management Guidelines for Adult Patients with Thyroid Nodules and Differentiated Thyroid Cancer: Proceedings of an Interactive International Symposium.Thyroid.;29(1):7-26.

Mohandas M., Nadarajan S., Hubert N., Abd El Shahid M, Ali MMA & Lotfy AE (2019). Diagnostic accuracy of TIRADS in evaluation of thyroid nodules. Poster in ECR 2019 / C-0774.

Mondal SK, Sinha S, Basak B, Roy DN & Sinha SK. (2013).
The Bethesda system for reporting thyroid fine n eedle aspirates:A cytologic study with histologic follow-up. J Cytol.;30(2):94-9.

Mora-Guzmán I, José Luis Muñoz de Nova , Marín-Campos C, Jiménez-Heffernan JA, Cuesta Pérez JJ, Lahera Vargas M, Torres Mínguez E & Martín-Pérez E. (2018). Efficiency of the Bethesda System for Thyroid Cytopathology. Cir ESP 96(6):363-368.

Mufti ST & Molah R. (2012). The bethesda system for reporting thyroid cytopathology: a five-year retrospective review of one center experience. Int J Health Sci (Qassim) ;6(2):159-173.

Nandedkar SS, Dixit M & Malukani K (2018). Evaluation of Thyroid Lesions by Fine-needle Aspiration Cytology According to Bethesda System and Its Histopathological Correlation. Int J Appl Basic Med Res ;8(2):76-82.

Naz S, Hashmi AA, Khurshid A, Faridi N, Edhi MM, Kamal A & Khan M. (2014). Diagnostic accuracy of Bethesda system for reporting thyroid cytopathology: an institutional perspective. Int Arch Med.;7:46.

Onder S, Firat P & Ates D. (2014). The Bethesda system for reporting thyroid cytopathology: an institutional experience of the outcome of

indeterminate categories. Cytopathology. ;25(3):177-184.

Park JH, Yoon SO, Son EJ, Kim HM, Nahm JH & Hong S. (2014). Incidence and malignancy rates of diagnoses in the bethesda system for reporting thyroid aspiration cytology: an institutional experience. Korean J Pathol.;48(2):133-139.

Periakaruppan G, Seshadri KG, Krishna GMV, Mandava R, Sai VPM & Rajendiran S (2018). Correlation Between Ultrasound-based TIRADS and Bethesda System for Reporting Thyroid-cytopathology: 2-year Experience at a Tertiary Care Center in India. Indian J Endocrinol Metab; 22(5):651-655.

Perros P, Boelaert K, Colley S, Evans C, Evans RM, Gerrard Ba G & British Thyroid Association. Guidelines for the management of thyroid cancer. (2014). Clin Endocrinol (Oxf).;81:1-122.

Porterfield JR Jr, Grant CS, Dean DS, Thompson GB, Farley DR, Richards ML, Reading CC, Charboneau JW, Vollrath BK & Sebo TJ. (2008). Reliability of benign fine needle aspiration cytology of large thyroid nodules. Surgery;144(6):963-968.

Raab SS, Vrbin CM, Grzybicki DM & Trzebińska A. (2006). Errors in thyroid gland fine-needle aspiration. Am J Clin Pathol; 125: 873–882.

Salillas AL, Sun FC & Almocera EG. (2015). Review of the Bethesda System for Reporting Thyroid Cytopathology: a local study in Bohol Island, Philippines. Acta Cytol. ;59(1):77-82.

Sánchez J.F. TI-(2014). TI-RADS classification of thyroid nodules based on a score modified according to ultrasound criteria for malignancy. Rev.Argent. Radiol., 78(3): 138-148.

Singaporewalla R. M., Hwee J., Lang T. U. & Desai V. (2017). Clinico-pathological Correlation of Thyroid Nodule Ultrasound and Cytology Using the TIRADS and Bethesda Classifications. World J Surg 41:1807–1811.

Sinna EA & Ezzat N. (2012). Diagnostic accuracy of fine needle aspiration cytology in thyroid lesions. J Egypt Natl Canc Inst.;24(2):63-70.

Straccia P, Rossi ED, Bizzarro T, Brunelli C, Cianfrini F, Damiani D & Fadda G. (2015). A meta-analytic review of the Bethesda System for Reporting Thyroid Cytopathology: Has the rate of malignancy in indeterminate lesions been underestimated?. Cancer Cytopathol;123(12):713-722

Tan H, Li Z, Li N, Qian J, Fan F, Zhong H, Feng J, Xu H & Li Z. (2019). Thyroid imaging reporting and data system combined with Bethesda classification in qualitative thyroid nodule diagnosis. Medicine (Baltimore) ;98(50): 18320.

Tayde P., Dalwadi P., Sharma B., Chavhan J., Pawal P., Joshi A.,Bhagwat N.,Chadha M. & Varthakavi p.(2014) Solitary Thyroid Nodule: A Clinical Approach. An International Journal, 6(1):23-29.

Tessler FN, Middleton WD, Grant EG, Hoang JK & Berland LL (2017). ACR Thyroid Imaging, Reporting and Data System (TI-RADS): White Paper of the ACR TI-RADS Committee. Journal of the American College of Radiology: JACR ;14 (5): 587-595.

Vuong HG, Ngo HTT, Bychkov A, Jung CK, Vu TH, Lu KB, Kakudo K & Kondo T. (2020) Differences in surgical resection rate and risk of malignancy in thyroid cytopathology practice between Western and Asian countries: A systematic review and meta-analysis. Cancer Cytopathol;128(4):238-249

Williams BA, Bullock MJ, Trites JR, Taylor SM & Hart RD (2013). Rates of thyroid malignancy by FNA diagnostic category. J Otolaryngol Head Neck Surg. ;42(1):61.

Yassa L, Cibas ES, Benson CB, Frates MC, Doubilet PM, Gawande AA, Moore FD Jr, Kim BW, Nosé V, Marqusee E, Larsen PR & Alexander EK. (2007). Long-term assessment of a multidisciplinary approach to thyroid nodule diagnostic evaluation. Cancer;111(6):508-516.

Yoo C, Choi HJ, Im S, Jung JH, Min K, Kang CS & Suh YJ. (2013). Fine needle aspiration cytology of thyroid follicular neoplasm: cytohistologic correlation and accuracy. Korean J Pathol.;47(1):61-6.