AAU CHS Department of Pathology

Concordance between BETHESDA cytology classification and ACR-TIRADS ultrasound classification in evaluation of thyroid nodule, at Tikur Anbesa Specialized Hospital, Addis Ababa, Ethiopia: Prospective study from April,2023-october,2023.

A research thesis submitted to Addis Ababa University, college of Health science, School of Medicine, Department of Pathology; in partial Fulfillment for the Requirements of Post graduate certificate in Pathology.

By: Birhanu Kassie, MD (Pathology Resident)

Advisor: Dagnachew Tamrat, MD, Assistant Professor of Pathology

February, 2024

Addis Ababa, Ethiopia

Contents

1. Abbreviations 4
2. Abstract 5
2.1 Background and objective 5
2.2 Methodology 5
3.1 Background 6
3.2 Statement of the problem 8
4. Literature review and significance of the study 9
4.1 Literature review 9
4.2 Significance of the study 10
5. Objectives 11
5.1 General objective 11
5.2 Specific objectives 11
6.1 Study setting 12
6.2 Study design 12
6.3 Study duration 12
6.4 Population 12
Source population 12
Study population 12
6.5 Sample size and sampling technique 12
6.6 Data collection procedure 13
6.7 Study variables 13
Dependent variables 13
Independent variables 13
6.8 Eligibility criteria, inclusion criteria, exclusion criteria 13
6.8.1 Eligibility criteria 13
6.8.2 Inclusion criteria 13
6.8.3 Exclusion criteria 14
6.9 Data processing and analysis 14
7. **Results** 15
8. Discussion 20

9. Strengths and Limitations 23
10. Conclusion 24
11. Recommendations 25
12. Dissemination of Results 25
13. Compliance with Ethical Standards 25
14. Acknowledgments 25
15. Disclosure of Conflict of Interest 25
References 27

1. Abbreviations

AAU: Addis Ababa University

ACR: American College of Radiology

CHS: College of Health Sciences

FNA: Fine Needle Aspiration

MNG: Multinodular Goiter

MRN: Medical Record Number

NPV: Negative Predictive Value

PPV: Positive Predictive Value

TASH: Tikur Anbessa Specialized Hospital

TBSRTC: The Bethesda System for Reporting Thyroid Cytology

TIRADS: Thyroid Imaging, Reporting and Data System

US: Ultrasound

2. Abstract

2.1 Background and objective

The Bethesda System for Reporting Thyroid Cytopathology TBSRTC is widely used method currently helping in management of thyroid nodules which are common and mostly benign. The thyroid ultrasound ACR-TIRADS system is also widely used classification system which has helped in screening for nodules requiring subsequent FNAC, but the correlation between the two systems has not been established in our setup. This study aims to determine the level of concordance between TBSRTC cytology category and ACR-TIRADS ultrasound category.

2.2 Methodology

It was a prospective cross-sectional study over 6 months which included 92 patients with thyroid nodule. In all cases the thyroid ultrasound and cytology findings were reported based on ACR-TIRADS and the TBSRTC 2017 system respectively. The level of agreement between ACR-TIRADS and BSRTC systems was expressed as kappa value.

2.3: Result

Study subjects were mostly women (84%), with average age of 47.60 years. The highest frequency in Bethesda system was seen in category II (54/92) and category V (17/92). In contrast, the highest frequency in TIRADS was, in category 4 (45/92) and category 3(21/92). Maximum concordance was found among the category 2-II and 5-V of the classification systems. The overall observed agreement was 65.2% with a linear weighted kappa of 0.343 (95% CI: 0.213-0.472). A higher weighted kappa value was seen in males, nodules ≥4 cm, individuals aged <50 years and free handed FNAC cases.

2.4 Conclusion

The study revealed a fair concordance between cytologic diagnosis based on the TBSRTC and ultrasound diagnosis based on the ACR-TIRADS, with the highest concordance seen in the two ends of the classification categories (TIRADS5-Bethesda V and TIRADS2-Bethesda II). Careful interpretation of the ultrasound findings and cytologic specimens allow the treating physicians to decide nodules which needs surgery and follow up, so that there will be no unnecessary surgical procedures and complications. There should be an institutional monitoring system for nodules with discordant Bethesda and TIRADS category, especially for those with extreme categories are produced.

Key words: Thyroid nodule, Bethesda, Ultrasound, ACR-TIRADS, FNAC, cytologic diagnosis

3. Introduction

3.1 Background

In recent decades, a steady increase in the incidence of thyroid cancer has been observed worldwide, and the reasons for this increase remain controversial. The increase in thyroid cancer is almost exclusively due to the increase in papillary cancers, and there are no significant changes in other histological subtypes. [1, 2,3]. The typical presentation is small tumors, although the incidence of large tumors is increasing; it has been suggested that the increase in the incidence of thyroid cancer is largely due to better detection rather than an actual increase in incidence. [3]. Thyroid cancer is the most common endocrine malignancy.

A thyroid nodule can be defined as a discrete lesion within the thyroid gland that is radiologically distinct from the surrounding thyroid parenchyma. It may be solitary, multiple, solid or cystic and may or may not be functional. Thyroid nodules are common in the general population and ultrasonography (US) has greatly increased the number of cases detected. About 4–8% of the thyroid nodules are detected by manual palpation. Ultrasound is an accurate method for the detection of thyroid nodules; but its accuracy in differentiating between benign from malignant thyroid nodules is low [4,5]. Hypo echogenicity, increased intra-nodular vascularity, irregular margins, microcalcifications, absent halo, and a taller-than-wide shape measured in the transverse dimension are commonest sonographic features associated with a higher likelihood of malignancy include. Thus, several US Thyroid Imaging and Data Systems (TIRADS) have been proposed for risk stratification of thyroid nodules. [5].

The nodules are usually divided into different categories based on TIRADS and are then referred for FNAC or follow-up, according to the variable risk of malignancy. TIRADS was originally designed to improve patient management and cost-effectiveness by avoiding unnecessary FNA biopsies in patients with thyroid nodules (Table 1), with a sensitivity, specificity, positive predictive value, negative predictive value, and accuracy of 88, 49, 49, 88, and 94%, respectively.

In addition, FNA biopsy is the most accurate method for determining malignancies and is an integral part of the current evaluation of thyroid nodules.
The Bethesda System for Thyroid Cytopathology Reporting is a standardized reporting system for classifying thyroid FNA biopsy results and includes six diagnostic categories with unique risk of malignancy and clinical treatment recommendations. The Bethesda system has been widely adopted since its commencement, each category with malignancy risk and recommended next steps (Table 2). A recent meta-analysis showed that the Bethesda reporting system has a sensitivity (97%), specificity(50.7%), negative predictive value (96.3%) and positive predictive values of 55.9% [7, 8].

Despite the fact that both US and FNAC are widely recommended procedures to study patients with thyroid nodules, the value of the existing concordance between the two methods has not been established in our setup. Consequently, the purpose of this study was to assess the existing concordance between the two diagnostic methods used in the initial evaluation of individuals with thyroid nodule (TIRADS and Bethesda systems).

Table 1: TIRADS Categories with Risk of malignancy and subsequent recommended management (ref. 6, 9,21)

TIRADS category	ROM	Subsequent management
TIRAD1 (Normal thyroid)	No	No FNA
TIRADS 2 (Benign)	No	No FNA
TIRADS 3 (Mildly suspicious)	5%	FNAC if ≥ 2.5cm Follow up if ≥ 1.5cm
TIRADS 4 (Moderately suspicious)	5-80%	FNAC if ≥ 1.5cm Follow up if ≥ 1cm
TIRADS 5 (Highly suspicious)	>80%	FNAC if ≥ 1 cm Follow up if ≥ 0.5cm
TIRADS 6 (Malignant, including biopsy proven cases		

Table 2: TBSRTC classification (ref 7,8,10,20)

Bethesda Category	ROM mean% (Range)	Usual management
Bethesda I (Nondiagnostic)	13 (5–20)	Repeat FNA with ultrasound guidance
Bethesda II (Benign)	4 (2–7)	Clinical and ultrasound follow-up
Bethesda III (AUS)	22 (13–30)	Repeat FNA, molecular testing, diagnostic lobectomy, or surveillance
Bethesda IV (Follicular neoplasm)	30 (23–34)	Molecular testing, diagnostic lobectomy
Bethesda V (Suspicious for malignancy)	74 (67–83)	Molecular testing, lobectomy or near-total thyroidectomy
Bethesda VI (Malignant)	97 (97–100)	Lobectomy or near-total thyroidectomy

3.2 Statement of the problem

Thyroid nodules are common findings, and also thyroid enlargement is a common presenting symptom in our set up. Ultrasound based risk stratification according to the TIRADS lexicon and FNAC via TBSRTC are playing significant role in guiding management of patients with thyroid nodule. But in our setup, it is not uncommon to see cases diagnosed as malignant via ultrasound and benign with FNAc resulting in complicated presentation and delayed management, similarly there are also scenarios where ultrasound report as benign and FNAC as malignant. The level of agreement of these two systems in our setup has not been tested. In this study we determined the level of concordance between the two systems ensuring the ultrasound and cytology results are reported based on the ACR-TIRADS and The Bethesda System for Reporting Thyroid Cytology.

4. Literature review and significance of the study

4.1 Literature review

In study done by Vergas H. et al (11), evaluated the concordance between the TIRADS and the Bethesda reporting systems on the non-toxic thyroid nodule. The result showed a good concordance and maximum concordance was found between category 2-II. This study showed an observed agreement of 87.2%. The linear weighted kappa value generated in this study was of 0.69. This study also found that higher level of agreement in nodules with size ≥4 cm, males, and individuals aged ≥50 years,

In another study done in India by Biswas A, et al (12), which was a prospective cross-sectional study over 1 year on 69 patients, showed that TIRADS and BSRTC systems have substantial agreement in categories with low risk of malignancy (category 2-II). The calculated weighted kappa value in this study was 0.411, which implies that Bethesda system and TIRADS lexicon have a moderate degree of agreement

In another study done by Redmi s. et al. (13), which is a prospective cross-sectional study which evaluated total of 54 patients over a period of nine months, showed an overall agreement of 77.77%, between the TBSRTC and Ultrasound. This study revealed a kappa value of 0.633, which implies that the two systems have a good agreement in evaluation of thyroid nodules. 11.1% of the lesions in this study were nondiagnostic or unsatisfactory for evaluation.

A study done by Chumber S. et al. (14), which is prospective observational study conducted at a tertiary care hospital in India, included 80 subjects. This study generated kappa value of 0.38, which implies the Bethesda system and TIRADS lexicon have a fair agreement in evaluation of thyroid nodules. Observed agreement of 64% is seen in this study.

A study by Livani s. et al. (15), descriptive-analytical study, performed on 165 patients at Educational Hospital in northern Iran. This study revealed that categories 3 and 4 of TIRADS system and categories II and IV of the Bethesda system were the most frequent categories. The kappa value calculated in this study was 0.061, which implies that the two systems have a slight agreement in evaluation of thyroid nodules. The highest consistency was seen for categories of TIRADS 2 and Bethesda II.

In another study done by Singapore W. et al (16), conducted on 100 cases over a period of 1 year, showed an overall observed agreement 83%. between TBSRTC and TIRADS.

A study by Periakaruppan et al (17), has showed that TIRADS and Bethesda system have significant association in initial assessment of thyroid nodules. This study has revealed that ultrasound has a specificity of 94.15%, sensitivity of 92.3%, NPV of 99.38% and PPV of 54.54%.

Another study by Thattarakkal. et al (18), showed that TIRADS classification has a strong negative predictive value, good specificity and fairly good sensitivity. This study calculated a sensitivity, specificity, positive predictive value and negative predictive value of 77.8%, 89.6%, 66.6%, and 93.8% respectively.

A study done by Pudasaini et al (19), which evaluated the level of agreement the TBSRTC and TIRADS system. Has revealed that the benign categories (Bethesda II and TIRADS 2) showed the highest agreement.

4.2 Significance of the study

Due to the high prevalence of thyroid nodules, it is obvious doing FNACs or following nodules will create a significant burden in a country with limited resources like our country. The limitation on health care budget, human resource and well-equipped hospitals is evident. Additionally, researches have shown that vast majority of the thyroid nodules are benign; and only about 5-7% of thyroid nodules are malignant. Therefore, it is important to identify nodules that are malignant and manage accordingly; while avoiding unnecessary burden created by the benign nodules.

US examination through ACR-TIRADS system is the modality of choice in examining the thyroid gland; and it is also relatively cheap and widely available. It allows evaluation of thyroid nodules, and formation of an impression as to which nodules have suspicious features and need tissue diagnosis. Fine needle aspiration cytology is the gold standard test in identifying benign and malignant thyroid nodules through the organized reporting form of Bethesda system (TBSRTC). Both systems, ACR-TIRADS and TBSRTC, are proven to be effective in decreasing unnecessary thyroid surgeries and guiding in nodules which needs surgery. The level of concordance between the two systems is seen to be good in studies done at best centers of western community. But the level of concordance has not been tested in our hospital; which is significantly different from the set ups where these studies were done. Some of the differences include the experience of the practitioners doing the ultrasound characterization, and the practice of US guided FNAC as well as free hand FNAc, the experience of the pathologists examining the cytology specimens. Also, the demographic factors and the epidemiology of different diseases, including pathologies of the thyroid disease may have significant impact on the outcome of such studies. Therefore, this study was aimed at determining the level of agreement of the two classification systems at TASH and if good concordance is seen to recommend their concordance in other institutions of our country, so that cautions to be paid whenever there is discrepancies in the categories of the two systems in evaluating thyroid nodule.

5. Objectives

5.1 General objective

-To determine the level of concordance between the ultrasound criteria established under TIRADS (The Thyroid Imaging Reporting and Data System for US of the thyroid); and the cytology criteria according to The Bethesda System for Reporting Thyroid Cytopathology at TASH.

5.2 Specific objectives

- To characterize the study population and concordance of the classification systems from the socio-demographic point of view.
- To characterize the Bethesda category in terms of frequency and pathologic diagnosis.
- To assess prevalence of benign and malignant lesions based on the cytologic diagnosis.

6. Research Methodology

6.1 Study setting

The study was conducted at Tikur Anbessa comprehensive specialized and teaching hospital, Pathology department, A.A, Ethiopia. TASH is under college of health sciences campus of AAU, which is one of the pioneer universities in the country. The hospital is a tertiary level referral and teaching hospital providing service to people from the all corners of the country in its various departments. The study was conducted in the Pathology department with collaboration with the TASH Radiology department. Both departments are equipped with qualified human resource, including consultant pathologists and radiologists that involved in conducting and guiding the research;

6.2 Study design

Institution based, prospective cross-sectional study is conducted in patients referred for free handed FNAc at Pathology department and ultrasound guided FNAc at Radiology department of TASH during the time period of April 2023-October 2023.

6.3Study duration

The study is conducted during the time period of April 2023-October 2023.

6.4 Population

Source population

All patients referred to TASH Pathology department for free handed FNAc and radiology department for ultrasound guided FNAC from thyroid nodule during the study period of April 2023-October 2023.

Study population

All patients referred to TASH Pathology department for free handed FNAc with ultrasound report and radiology department for ultrasound guided FNAC from thyroid nodule during the study period of April 2023-October 2023.

6.5 Sample size and sampling technique

Since the number of both free handed and US guided FNACs for thyroid nodules done at both departments of TASH is small, we use non probability sampling technique; and all patients for whom ultrasound guided or free handed FNAC is done were included in our study. Upon consecutive non-probabilistic sampling initially 115 patients were included, however the final analysis was limited to 92 patients and 23 patients were excluded due to:

1. Bethesda category I cytology result in 13 patients (2 cases free handed and 11 cases ultrasound guided)
2. Cytology report not based on TBSRTC in 1 patient (result reported as see description)
3. TIRADS category 1 in 1 patient
4. Ultrasound report done outside of TASH in 1 case
5. Post thyroidectomy in 7 cases (4 cases follicular carcinoma, 3 cases for papillary thyroid carcinoma, 1 case for hurthl'e cell carcinoma)

6.6 Data collection procedure

Data was collected by using structured questioner (annex-1). The questioner included the MRN, a demographic data (the age and sex) of the patient, duration of the nodule, size of the nodule from measurement of ultrasound, presence or absence of toxic symptoms, the TIRADS category, Bethesda category, the specific pathologic diagnosis rendered.

6.7 Study variables

Dependent variables

-Bethesda (TBSRTC) category

-ACR-TIRADS category

Independent variables

-Demographic data

Age

Sex

-Duration of lesion

- Toxicity symptoms

- Nodule size

- Mode of aspiration

6.8 Eligibility criteria, inclusion criteria, exclusion criteria

6.8.1 Eligibility criteria

-All thyroid nodules for which free handed and ultrasound guided FNAC is done at TASH, pathology and radiology department were included in this study.

6.8.2 Inclusion criteria

-All thyroid nodules for which free handed or ultrasound guided FNAC is done, provided that ultrasound report with ACR-TIRADS category and pathologic results with Bethesda (TBSRTC) category were obtained.

-IN FNAc done in more than 1 thyroid nodules of the same patient; the nodules were recorded as separate data set; provided that the ultrasound report has separate

TIRADS and Bethesda category.

6.8.3 Exclusion criteria

- All thyroid nodules for which the ultrasound report and cytologic result could not be produced.
- Those with repeat aspiration within same study period
- Those with the ultrasound examination done outside TASH
- TIRADS 1 and Bethesda category I cases
- All nodules in a patient with known thyroid malignancy or thyroidectomy for unknown diagnosis.

6.9 Data processing and analysis

All the nodules for which ultrasound report consists ACR-TIRADS category and pathologic results in form of Bethesda category were included in the study.

The weighted Kappa statistical method with a 95% confidence interval and the statistical Z-test were used to estimate the level of concordance between the two systems. A p-value of < 0.05 was considered as statistically significant.

In order to undergo the Kappa analysis, the number of diagnostic categories in TIRADS system and TBSRTC should be equal. The first category of both systems are excluded because category 1 in TIRADS is normal thyroid and, in Bethesda is non diagnostic. And category 5-V and 6-VI of both systems are combined into one, because the highest risk of malignancy is described in those two categories. As a result, the analysis categories are:

- Bethesda category (II, III, IV and V)
- TIRADS category (2, 3, 4 and 5)

In this study we used cohens kappa with linear weight, to determine the level of concordance. A further analysis was done to assess factors resulting in variation in weighted kappa for the variables of: gender, age, nodule size, presence or absence of toxic symptoms and mode of aspiration. All the analyses were based on SPSS version 27.

7. Results

A total of 92 patients are included in this study, in 45.65%(n=42) of cases the FNAc was done under ultrasound guidance whereas 54.35%(n=50) of cases were aspirated free handed. Majority of patients (84.8%, n=78) were female with female to male ration of 5:1.

Each Bethesda categories show exclusive female predominance (>80% were female), except for category V where male to female ratio is 1.4:1 (41.2% and 58.8%). The average age was 47.60 years (SD=14.63) with minimum and maximum of 14 and 90 years respectively. 43.5% of patients were above the age of 50 year.

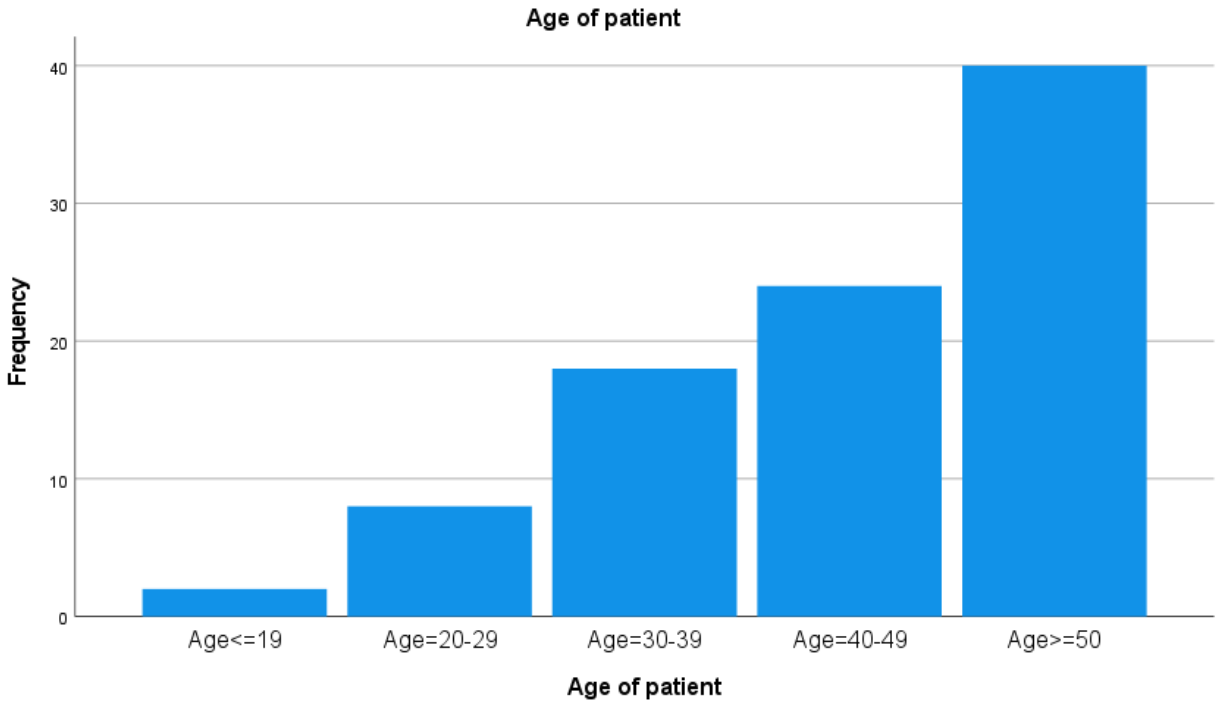

Fig 1: study subjects according to their age distribution

Mean size of nodules was 3.72cm (SD=1.26, minimum=1.8 and maximum=7.4cm). In 63.04% (n=58) of nodules the size was below 4cm. There was no statically significant size variation between Bethesda categories (p value=0.082).

Majority of patients (77.2%, n=71) were seen to have toxic symptoms. The prevalence of toxic symptoms in cases categorized as Bethesda category (IV, V and VI) was 3.125%. Out of 32 cases reported as Bethesda category IV, V and VI, toxicity is seen in only 1 case.

Table 3: Age, sex, size nodule, presence or absence of toxic symptoms and duration of lesion in each Bethesda categories

	Age	Sex	Size of nodule in cm	Toxicity prevalence	Duration in year
Bethesda II	Min=14 Mean=48.04 Maximum=90 SD=15.7	Female=96.6% Male=7.4%	Minimum=1.8 Mean=3.57 Maximum=7 SD=1.09	33.3%	Min=1 Mean=6.28 Max=30 SD=6.43
Bethesda III	Min=29 Mean=45 Maximum=68 SD=13.5	Female=100%	Minimum=2.20 Mean=3.12 Maximum=4.1 SD=0.82	33.3%	Minimum=1 Mean=5.17 Maximum=12 SD=3.92
Bethesda IV	Minimum=20 Mean=47.47 Maximum=70 SD=14.67	Female=80% Male=20%	Minimum=1.80 Mean=3.88 Maximum=6.5 SD=1.49	6.7%	Minimum=1 Mean=4.33 Maximum=20 SD=5.06
Bethesda IV	Minimum=25 Mean=47.24 Maximum=69 SD=12.31	Female=58.8% Male=41.2%	Minimum=2.00 Mean=4.25 Maximum=7.4 SD=1.54	0%	Minimum=1 Mean=3.06 Maximum=10 SD=2.95

The frequency of ultrasound results in TIRADS category was, TIRAD2 (20.7%, n=19), TIRAD3 (22.8%, n=21), TIRAD4 (38.0%, n=34), and TIRAD5 (18.5%, n=17).

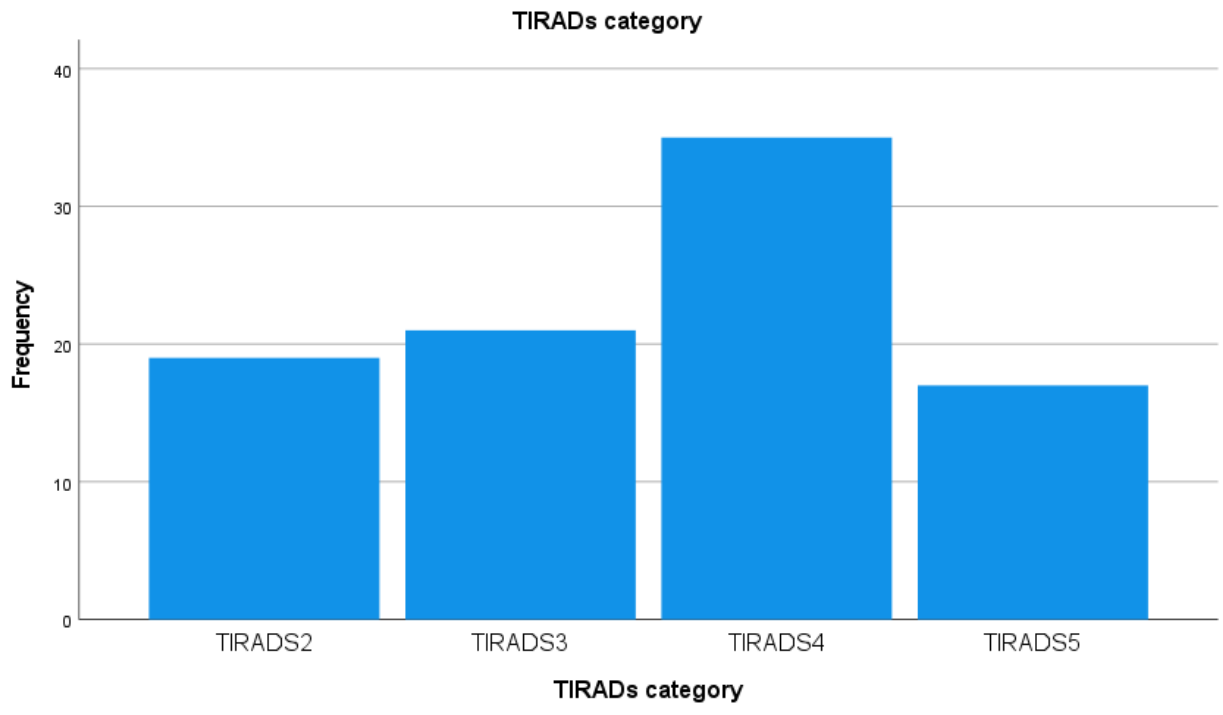

Fig 2: Frequency distribution of study subjects based on TIRADS Category

The frequency of Bethesda category II (58.7%, n=54), III (6.5%, n=6), IV (16.3%, n=15), and V (18.5%, n=17).

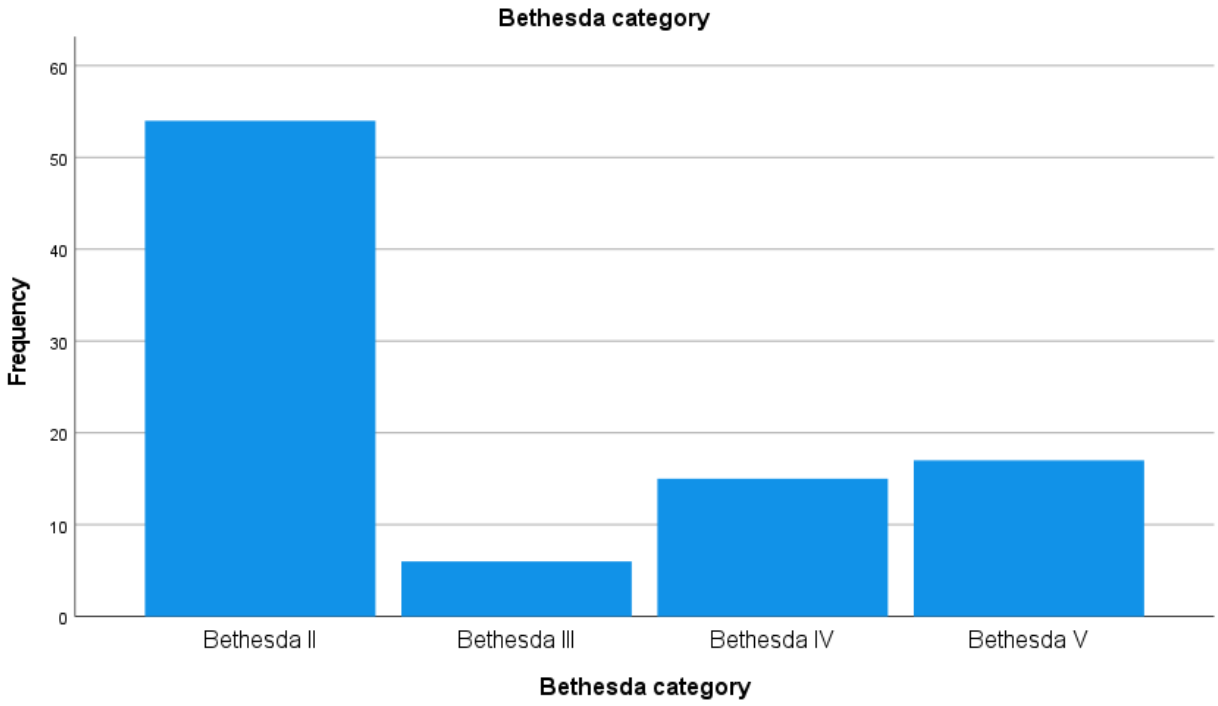

Fig 3: Frequency distribution of study subjects according to Bethesda category

Table 4: Joint distribution of TIRADS and Bethesda categories

		Bethesda category				
		II	III	IV	V	Total
TIRADs category	TIRADS2	16	1	2	0	19
	TIRADS3	14	3	3	1	21
	TIRADS4	22	2	7	4	35
	TIRADS5	2	0	3	12	17
Total		54	6	15	17	92

Non diagnostic rate (Bethesda category I) in our study was 11.3% (13/115), and majority of thyroid nodules turned out to be nondiagnostic are those aspirated under ultrasound guidance (84.6%).

Table 5: Nondiagnostic rates based on mode of aspiration

Mode of aspiration	Number of nondiagnostic cases (%)
Ultrasound guided	11 (84.6%)
Free handed	2 (15.4%)

Colloid goiter is the most prevalent cytologic diagnosis, accounting for 56.5% (n=52) of total cases and 96.3% of cases reported as Bethesda category II. Majority of cases categorized as Bethesda Category VI, turned out to be papillary thyroid carcinoma (67%, n=8), the remaining cases in this category are Anaplastic thyroid carcinoma (16.7%, n=2), metastatic carcinoma suggestive of renal cell carcinoma (8.3%, n=1) and squamous cell carcinoma (8.3%, n=1). The prevalence of malignancy as per the cytologic diagnosis was 13%.

Table 6: Bethesda categories with cytology diagnosis

Bethesda category (Number of cases and percentage out of total)	Cytologic diagnosis, number of cases with percentage within category
Bethesda II, 54(58.7%)	-Colloid Goiter, 52(96.3%) -Adenomatoid Goiter, 1(1.85%) -Reidel's Thyroiditis, 1(1.85%)
Bethesda III, 6(6.5%)	AUS
Bethesda IV, 15(16.3%)	-Follicular neoplasm, 10(66.7%) -Suspicious for follicular neoplasm, 4(26.7%) -Hurthle cell neoplasm, 1(6.6%)
Bethesda V, 5(5.43%)	-Suspicious for papillary thyroid carcinoma, 3(60%) -Suspicious for malignancy, 2(40%)
Bethesda VI, 12 (13.04%)	-Papillary thyroid carcinoma, 8(66.7%) -Anaplastic Thyroid carcinoma, 2(16.7% -Squamous cell Carcinoma, 1(8.3%) -Suggestive of metastatic Renal cell carcinoma 1(8.3%)

The overall observed agreement between the TIRADS and Bethesda categories was 65.2% with estimated agreement (agreement by chance) was 48%. The highest concordance was found for categories TIRADS5-Bethesda V (70.6%). Out of 17 cases classified as TIRADS 5 category, 12 were rated as Bethesda V, 3 as Bethesda IV, and 2 as Bethesda II. The frequency of Bethesda V

was 18.5% (17/92), 12 of the patients were rated as TIRADS 5, 4 as TIRADS 4, and 1 rated as TIRAD 3.

The next highest agreement was between categories of Bethesda II-TIRADS 2. Out of 19 cases categorized under TIRADS2, 16 turned out to be Bethesda II, 1 Bethesda III and 2 Bethesda V.

The weighted kappa value calculated was 0.343 (95% CI, 0.213-0.472, P<0.01, Z=5.124). And a higher weighted kappa value was seen in male patients, nodules ≥ 4cm, individuals <50 years of age, and patients whose FNAC was done free handed. ACR-TIRADS categorization of thyroid nodules is found to have sensitivity (81.25%), specificity (56.67%), positive predictive value (50%) and negative predictive value (85%)

Table 7: The kappa value variation for age, sex, nodule size, toxicity, and mode of aspiration

Variable	Kappa	Standard error	IC 95%
Overall kapa	0.343	0.066	0.213-0.472
Sex Male	0.556	0.138	0.286-0.825
Female	0.280	0.070	0.143-0.416
Age <50	0.456	0.094	0.272-0.641
≥50	0.218	0.074	0.072-0.364
Size <4cm	0.284	0.087	0.113-0.455
≥4cm	0.419	0.104	0.216-0.623
Toxicity Yes	0.114	0.072	0.028-0.256
No	0.380	0.073	0.237-0.524
Mode of aspiration			
Ultrasound guided	0.196	0.068	0.063-0.329
Free handed	0.436	0.100	0.24-0.632

8. Discussion

This study evaluated the concordance between the TIRADS and the Bethesda reporting systems on thyroid nodule. The result showed a "fair" concordance and the most frequent agreement was found for categories 5-V and 2-II. The kappa index measures the level of inter-observer agreement, or as in this study the concordance between the category produced by ultrasound TIRADS and cytology result produced by TBSRTC. Kappa is a measure of reliability, not a measure of validity.

Majority of patients in our study were female (84.8%). Other similar studies also reported predominance of the female population presenting with thyroid lesion. (11) (12) (13) (14) (17) (18) (19), Probably this trend is due to the fact that autoimmune thyroid disease is significantly more frequent in females than in males, so these patients with autoimmune thyroid disease visit the physician more often increasing the probability of detecting the nodules either through palpation or ultrasound; clinically this situation may be defined as a "medical surveillance bias" (11). Increased prevalence of thyroid lesions in females, according to studies, might be related to hormonal influences of estrogen and progesterone. (13)

The mean age of our patients was 47.60 (SD=14.67), this finding is similar with those reported previously by Pudasaini et al. (19) (48.9, SD 14.2) and Redmi S. et al (13) (50.74, SD=17.8), and older and younger than those reported by Biswas A. et al (12) (41.54, SD=11.86) and Vargas H. et al (11) (57, SD=14) respectively.

Bethesda II was the most frequent category 54/92 (58.7%), and colloid goiter is the most prevalent cytologic diagnosis accounting for 96.3% of Bethesda II cases and 56.5% of all cytologic diagnoses in this study. Similar finding of predominance of Bethesda II category is seen in other studies by Vergas H. et al (36.1%), Biswas A. et al (56.52%), Redmi S, et al (68.5%) and Singapore W. et al (74.65%), (11) (12) (13) (16)

The overall agreement between TBSRTC and ACR-TIRADS categories found in our study was 65.2% with expected agreement of 48%. Similar observed agreement is seen in a study done by Chumber S. et al (64%) with lower expected agreement of 41.6%. A higher overall agreement is seen in studies done by Redmi S. et al (77.77%) and Singapore W. et al (83%)

Maximum agreement between Bethesda and TIRADS categories is seen between categories TIRADS 5- Bethesda V and TIRADS 2-Bethesda II. A study done in India by Chumber S. et al (14), has shown similar higher concordance rate between Categories of TIRADS 2-Bethesda II and TIRADS 5-Bethesda V. Other studies have also revealed higher agreement level between categories of TIRADS 2-Bethesda II. (11) (12) (15)

The weighted kappa value generated in our study is 0.343 which implies the two diagnostic systems has fair concordance rate. Similar kappa level of agreement is seen in a study done by Chumber S. et al (14), which is 0.38. A higher kappa value is revealed by other similar

studies done by Vergas H. et al (kappa=0.69 substantial or good agreement) and Redmi S. et al (kappa=0.633 substantial or good agreement).

Our study has found that higher weighted kappa value in those with nodule size greater than 4cm, age less than 50-year, male sex and those cases with free handed aspiration. Similar trend of higher kappa value for nodule size greater than 4cm is seen in a study done by Vergas H. et al (11)

In our study ultrasound is found to have 81.25% sensitivity, 56.67% specificity, 50% positive predictive value and 85% negative predictive value. These findings are lower compared to studies done by Periakaruppan et al (17), where the sensitivity, specificity, NNP and PPV were 92.3%, 94.15%, 99.38% and 54.54% respectively. A study done by Singapore W. et al (16) showed lower sensitivity (70.6%), and higher specificity (90.4%) and NPV (93.8%). Another study by Thattarakkal et al (18) also showed higher specificity (89.6%), NPV (93.8%) and PPV (66.6%), but lower sensitivity (77.8%).

Table 8: A comparation of previously published studies to the current study

Study and year	Country	Study Design	Sample size	M: F	Mean age	Concordance	Conclusion
Peri Karuppan. et al (17) 2018	India	Prospective	184 over 2 years	28: 156	3rd -6th Decade	Not mentioned	TIRADS and Bethesda systems showed remarkable agreement especially for the category II cases.
Vargas. et al (11) 2017	Columbia	Prospective	180	1: 2	57	TIRADS2 and Bethesda II Kappa =0.690	TIRADS criteria have substantial agreement with Bethesda system
Biswas. et al (12) 2020	India	Prospective	69 over 1 year	1: 5.3	41.54	TIRADS 2 and Bethesda II	Careful interpretation of both systems is essential in classifying thyroid nodules, to insure appropriate management.
Chumber et al (14) 2023	India	Prospective	80	12: 66	65% (20-40 years)	Categories 2-II and 5-V	The study shows a fair concordance between the TIRADS

						Kappa=0.380	and TBSRTC. With higher agreement in categories 2/II and 5/V of both classification systems.
Regmi et al (13) 2018	Nepal	Prospective	54 over 9 months	4: 50	50.74	Kappa= 0.633	The study showed good concordance between ultrasound TIRADS categories and cytology TBSRTC categories. FNAC should only be recommended for suspicious nodules.
Present study	Ethiopia	prospective	92 over 6 months	1: 5	47.60	Categories TIRADS 5 - Bethesda V and TIRADS 2-Bethesda II	The TIRADS and Bethesda system has a fair concordance rate with maximum agreement seen between the two ends of the classification systems (TIRADS5-Bethesda V and TIRADS2-Bethesda II).

9. Strengths and Limitations

This study is the first of its type in our country, to the best of our knowledge.

The prospective nature of the study has helped us to get the complete data about the thyroid nodules.

The fact that the ultrasound reports and the cytology results are signed out by different specialists, limited our ability to test true validity of the study.

10. Conclusion

Our study revealed a fair concordance between cytologic diagnosis based on the TBSRTC and ultrasound diagnosis based on the ACR-TIRADS, with the highest concordance seen in the two ends of the classification categories (TIRADS5-Bethesda V and TIRADS2-Bethesda II). Careful interpretation of the ultrasound findings and cytologic specimens allow the treating physicians to decide nodules which needs surgery and follow up, so that there will be no unnecessary surgical procedures and complications. There should be an institutional monitoring system for nodules with discordant Bethesda and TIRADS category, especially for those with extremes categories are produced.

11. Recommendations

We recommend consistent use of the TBSRTC for reporting thyroid FNAc results in our institution.

Maximum attention should be paid in signing thyroid FNAc cases with discordant categories.

Our study showed higher non diagnostic rate and relatively lower level of agreement between the two systems in the thyroid nodules aspirated with ultrasound guidance. So, ultrasound guided FNAC should be done with experienced hands and it would be beneficial to develop interdepartmental interactions between pathology and radiology team to optimize patient management

We recommend to do further studies with a larger sample size over multiple institutions so as to substantiate the results of our study

We also recommend follow up of histopathology results of the study subjects included in this study to further analyze the correlation of the ultrasound ACR-TIRADS and cytology Bethesda categories with final Histopathology diagnosis.

12. Dissemination of Results

Results of the study will be submitted to the department of pathology of TASH as part of dissertation requirement for the postgraduate certificate program and the thesis work will be presented for all pathology department staff members and residents. It will also be forwarded to journals for possible publication.

13. Compliance with Ethical Standards

All the procedures performed in this study were in accordance with the ethical standards of the department's research committee.

14. Acknowledgments

I would like to acknowledge and give my warmest thanks to my advisor and mentor Dr. Dagnachew Tamrat (Assistant professor of pathology) who made this work possible. His guidance and advice carried me through all the stages of writing this paper.

I would like to thank all TASH pathology and radiology department staff members and my fellow residents, for their valuable support in the data collection process.

15. Disclosure of Conflict of Interest

No conflicts of interest are declared by the outers.

Annex-1

Ser No.	MRN	FNAc no.	Age	Sex	duration	Toxic symptom	Size	TIRAD	Bethesda	Pathologic DX

References

1. Ferlay J, Seromata I, Dikshit R, Eser S, Mathers C, Rebelo M, Parkin DM, Forman D, Bray F. Cancer incidence and mortality worldwide: sources, methods and major patterns in GLOBOCAN 2012. Int J Cancer. 2015;136(5):359–86.

2. Torre LA, Bray F, Siegel RL, Ferlay J, Lortet-Tieulent J, Jemal A. Global cancer statistics, 2012. CA Cancer J Clin. 2015;65(2):87–108.

3. Wiltshire JJ, Drake TM, Uttley L, Balasubramanian SP. Systematic review of trends in the incidence rates of thyroid cancer. Thyroid. 2016;26(11):1541–52.

4. Gharib H, Papini E, Garber JR, Duick DS, Harrell RM, Hegedüs L, Paschke R, Valcavi R, Vitti P, AACE/ACE/AME Task Force on Thyroid Nodules. American Association of Clinical Endocrinologists, American College of Endocrinology, And Association Medici Endocrinologic Medical Guidelines for Clinical Practice for The Diagnosis and Management of Thyroid Nodules--2016 Update. Endocrine Practice. 2016;22(5):622–39.

5. Yoon JH, Lee HS, Kim EK, Moon HJ, Kwak JY. Malignancy risk stratification of thyroid nodules: comparison between the thyroid imaging reporting and data system and the 2014 American thyroid association management guidelines. Radiology. 2016;278(3):917–24.

6. Horvath E, Majlis S, Rossi R, Franco C, Niedmann JP, Castro A, Dominguez M. An ultrasonogram reporting system for thyroid nodules stratifying cancer risk for clinical management. J Clin Endocrinol Metab. 2009;94(5):1748–51.

7. Pusztaszeri M, Rossi ED, Auger M, Baloch Z, Bishop J, Bongiovanni M, Chandra A, Cochand-Priollet B, Fadda G, Hirokawa M, Hong S, Kakudo K, Krane JF, Nayar R, Parangi S, Schmitt F, Faquin WC. The Bethesda system for reporting thyroid cytopathology: proposed modifications and updates for the second edition from an international panel. Acta Cytol. 2016;60(5):399–405.

8. Garg S, Desai NJ, Mehta D, Vaishnav M. To establish Bethesda system for diagnosis of thyroid nodules on the basis of fnac with histopathological correlation. J Clin Diagn Res. 2015;9(12):EC17–21.

9. Russ G, Bigorgne C, Royer B, Rouxel A, Bienvenu-Perrard M. The Thyroid Imaging Reporting and Data System (TIRADS) for ultrasound of the thyroid. *J Radiol.* 2011;92(7–8):701–713. Doi: 10.1016/j.jradio.2011.03.022.
10. Cibas ES, Ali SZ. NCI Thyroid FNA State of the Science Conference. The Bethesda System for reporting thyroid cytopathology. *Am J Clin Pathol.* 2009;132:658–665. Doi: 10.1309/AJCPPHLWMI3JV4LA.

11. Vargas-Uricoechea H, Meza-Cabrera I, Herrera-Chaparro J. Concordance between the TIRADS ultrasound criteria and the BETHESDA cytology criteria on the nontoxic thyroid nodule. Thyroid Res. 2017 Feb 2;10(1):1–9.
12. Biswas A, Basu K, De S, Karmakar S, De D, Sengupta M, et al. Correlation between thyroid imaging reporting and data system and Bethesda system of reporting of thyroid cytopathology of thyroid nodule: A single center experience. J Cytol. 2020 Oct 1;37(4):193–9.
13. Regmi S, Tiwari A, Sharma R. Comparison of Fine Needle Aspiration Cytology in Thyroid Lesions using The Bethesda System for Reporting Thyroid Cytopathology with Ultrasonography using Thyroid Imaging Reporting and Data System. Journal of Lumbini Medical College. 2018 Dec 30;6(2).
14. Chumber S, Vyas S, Kataria K, Agarwal S, Rathore YS, Puri G, et al. Evaluation of Concordance of Ultrasound, Cytology, and Histopathology in Solitary Thyroid Nodules. Indian Journal of Endocrine Surgery and Research. 2023 Jun 30;18(1):17–23.
15. Livani S, Naeimi E, Taghavi N. Agreement between thyroid nodules ultrasound and cytology of fine needle aspiration (FNA) based on TIRADS and Bethesda system. Vol. 112, Journal of Gorgan University of Medical Sciences.
16. Singapore Walla RM, Hwee J, Lang TU, Desai V. Clinico-pathological Correlation of Thyroid Nodule
Ultrasound and Cytology Using the TIRADS and Bethesda Classifications. World J Surg. 2017 Jul 1;41(7):1807–11.
17. Periakaruppan G, Seshadri KG, Vignesh Krishna GM, Mandava R, Venkata Sai PM, Rajendiran S.
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. 2018 Sep 1;22(5):651–5.
18. Thattarakkal VR, Ahmed TSF, Saravanam PK, Murali S. Evaluation of Thyroid Nodule: Thyroid Imaging Reporting and Data System (TIRADS) and Clinicopathological Correlation. Indian Journal of Otolaryngology and Head and Neck Surgery. 2022 Dec 1; 74:5850–5.
19. Tuladhar AS, Pudasaini S, Simkhada S, Shrestha A, Pradhan S. Correlation of American College of
Radiology (ACR)-Thyroid Imaging Reporting and Data System (TIRADS) findings in Ultrasonogram (USG) of thyroid nodules with FNAC or Biopsy findings. Nepal Medical College Journal. 2022 Apr 4;24(1):23–9.
20. Najeeb DF, Ali HH. Bethesda system of reporting cytopathology of thyroid nodules in correlation with radiological and clinicopathological findings. Int J Health Sci (Qassim). 2022 Sep 1;956–66.
21. Tessler FN, Middleton WD, Grant EG, Hoang JK, Berland LL, Teefey SA, et al. ACR Thyroid Imaging, Reporting and Data System (TI-RADS): White Paper of the ACR TI-RADS Committee. Journal of the American College of Radiology. 2017 May 1;14(5):587–95.