

# Clinicopathologic Agreement of ACR TI-RADS, Bethesda, and Histopathology in Thyroid Nodules at Clinica de Occidente, Cali, Colombia

E Manuel Barajas<sup>1</sup>, Eduardo Gutierrez Martin<sup>2</sup>, Ivo Siljic<sup>2</sup>, Alejandro Ramirez Florin<sup>3\*</sup>, Leslie Melissa Salazar<sup>4</sup>, Laura Camila Martinez<sup>4</sup>, Jose David Tamayo<sup>4</sup> and Jobany Castro Espinosa<sup>5</sup>

<sup>1</sup>Department of Endocrine and Advanced Laparoscopic Surgery, Clinica de Occidente, Cali, Colombia

<sup>2</sup>Department of Surgical Oncology, Clinica de Occidente, Cali, Colombia

<sup>3</sup>International Healthcare Services, Clinica de Occidente, Cali, Colombia

<sup>4</sup>Medical Internship Program, Fundacion Universitaria San Martin and Clinica de Occidente, Cali, Colombia

<sup>5</sup>Research Coordination, Fundacion Universitaria San Martin, Cali, Colombia

## \*Corresponding author:

Alejandro Ramirez Florin, MD.  
International Healthcare Services,  
Clinica de Occidente, Cali, Colombia

Received: 13 Aug 2026

Accepted: 25 Aug 2026

Published: 28 Aug 2026

J Short Name: COO

## Copyright:

©2026 Alejandro Ramirez Florin. This is an open access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and build upon your work non-commercially

## Keywords:

Thyroid Nodules; ACR TI-RADS; Bethesda System; Histopathology; Diagnostic Accuracy

## Citation:

Alejandro Ramirez Florin, Clinicopathologic Agreement of ACR TI-RADS, Bethesda, and Histopathology in Thyroid Nodules at Clinica de Occidente, Cali, Colombia. Clinics of Oncology® 2026; V11(1): 1-8

## 1. Abstract

### 1.1. Objective

To evaluate agreement and diagnostic accuracy of recalculated ACR TI-RADS, ultrasound-reported TI-RADS, and Bethesda cytology against final surgical histopathology in thyroid nodule patients managed at Clinica de Occidente, Cali, Colombia.

### 1.2. Methods

We conducted a retrospective diagnostic accuracy and agreement study using a cleaned institutional database. Each row represented one patient. ACR TI-RADS was recalculated from the recorded ultrasound score and compared with the TI-RADS category reported in the original ultrasound report. Bethesda cytology and ultrasound categories were evaluated against final surgical histopathology as the reference standard. Sensitivity, specificity, predictive values, accuracy, receiver operating characteristic curves, and kappa agreement were estimated. Results: The final cohort included 177 patients; 168 had final histopathology, of whom 77 (45.8%) had malignant disease. Bethesda showed the strongest discrimination for malignant histopathology (AUC 0.712), followed by recalculated ACR TI-RADS (AUC 0.621) and ultrasound-reported TI-RADS (AUC 0.579). ACR TI-RADS  $\geq 4$  was sensitive but poorly specific, whereas Bethesda  $\geq V$  was highly specific with limited sensitivity. Agreement between recalculated ACR TI-RADS and ultrasound-reported TI-RADS was low to moderate.

### 1.3. Conclusion

In this histopathology-anchored institutional cohort, Bethesda cytology outperformed ultrasound-based classifications. The findings invite local strategies to improve harmonization of ACR TI-RADS reporting.

## 2. Introduction

Thyroid nodules are common in clinical practice and on neck imaging. The central problem is not detection, but reliable pre-operative risk stratification: most nodules are benign, whereas a selected minority require cytology, surgery, and definitive histopathologic classification. A useful diagnostic pathway should identify clinically significant malignancy without expanding procedures in low-risk disease [1].

ACR TI-RADS and the Bethesda System address different points in this pathway. ACR TI-RADS standardizes ultrasound assessment through scoring of composition, echogenicity, shape, margins, and echogenic foci; however, size and clinical context remain essential for biopsy and follow-up decisions [2]. Bethesda cytology provides standardized cytopathologic categories and was updated in 2023; its indeterminate categories, particularly III and IV, remain clinically challenging when imaging and final pathology are discordant [3].

The study was conceived as an institutional clinicopathologic agreement project. The expected sequence is straightforward: a clinically detected or incidental nodule is characterized by ultrasound; suspicious sonographic findings should increase the

likelihood of FNAC; Bethesda cytology should refine risk; and final surgical histopathology should adjudicate whether the pre-operative pathway anticipated benign or malignant disease.

Although this relationship has been studied internationally, Colombian evidence remains limited and has addressed different parts of the pathway. Prior national reports have evaluated TI-RADS-Bethesda concordance, Bethesda-histopathology correlation, ACR TI-RADS agreement, and thyroid nodule characterization in high-complexity settings [7,11-13]. Few local reports have integrated ultrasound-reported TI-RADS, recalculated ACR TI-RADS, Bethesda cytology, and final surgical histopathology in the same institutional cohort.

Clinica de Occidente is a high-complexity oncology centre in Cali, Colombia, corresponding to a level III-IV institution in the national health system. Its thyroid pathway includes radiology, cytology, pathology, endocrine surgery, and surgical oncology. Not all steps were performed in-house for every patient, but the institutional capacity to complete the full pathway makes this cohort suitable for auditing routine multidisciplinary risk stratification.

### 3. Materials and Methods

#### 3.1. Study Design, Setting, and Data Source

We performed a retrospective diagnostic accuracy and agreement study, reported according to STARD principles where applicable [8]. Data came from an institutional medical record database of thyroid nodule patients evaluated at Clinica de Occidente, Cali, Colombia.

The project began with an institutional universe of approximately 3,000 potentially relevant thyroid nodule records. Because many lacked linked clinical, ultrasound, cytologic, surgical, or histopathologic information, analysis was limited to records with sufficient recoverable data to evaluate the ultrasound-cytology-histopathology pathway.

Patients entered the pathway through heterogeneous referral routes. Some had external ultrasound, cytology, or pathology reports, whereas others completed more steps at Clinica de Occidente. Original reports were retained to preserve the real-world pathway; ACR TI-RADS was recalculated when recorded data permitted it.

The study used secondary data only and did not involve patient contact or intervention.

#### 3.2. Participants and Unit of Analysis

Each database row represented one patient. After cleaning, 177 records had recoverable diagnostic information. Diagnostic accuracy analyses were restricted to patients with final surgical histopathology, the reference standard for benign versus malignant disease.

Nodule size was not recoverable from the available database. This prevented assessment of size-based ACR TI-RADS biopsy and follow-up recommendations, and it limited interpretation to category-level risk stratification rather than guideline-concordant management decisions.

#### 3.3. Index Tests and Reference Standard

The index tests were recalculated ACR TI-RADS, TI-RADS as reported in the original ultrasound report, and Bethesda cytology from FNAC. Reported and recalculated TI-RADS were analyzed separately to evaluate both diagnostic performance and the consistency of routine ultrasound reporting with the standardized ACR framework.

Final surgical histopathology was the reference standard. Papillary, follicular, and anaplastic carcinoma were classified as malignant; benign nodule, follicular adenoma, and Hashimoto thyroiditis were classified as benign unless malignant behaviour was explicitly reported.

#### 3.4. Data Quality Control

Data were reviewed for missing values, duplicate entries, non-standardized categories, displaced summary tables, and discrepancies between ACR score and recorded category. Terms indicating absent information were treated as missing. Summary tables embedded below patient records were excluded.

A probable duplicate with discordant age and sex was recorded as a data-quality limitation rather than corrected by assumption.

### 4. Statistical Analysis

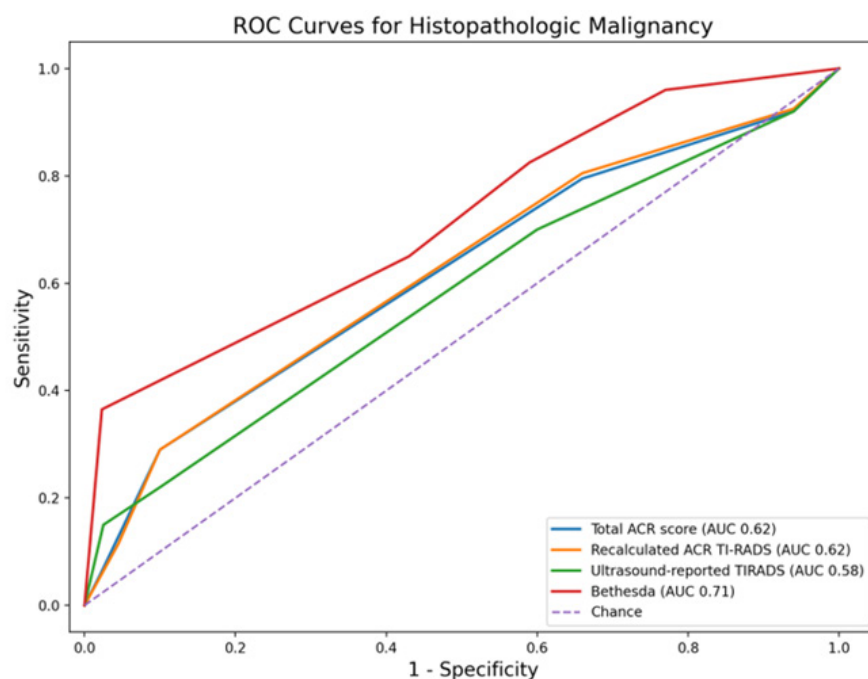
Categorical variables were summarized as n (%). Quantitative variables were summarized as mean and standard deviation or, when skewed, as median, interquartile range, and range. Missing data were handled with analysis-specific denominators.

Diagnostic accuracy was assessed using sensitivity, specificity, positive predictive value (PPV), negative predictive value (NPV), and accuracy, with 95% confidence intervals estimated using the Wilson method [9]. Thresholds included recalculated ACR TI-RADS  $\geq 4$  and  $\geq 5$ , ultrasound-reported TI-RADS  $\geq 4$  and  $\geq 5$ , and Bethesda  $\geq \text{III}$ ,  $\geq \text{IV}$ , and  $\geq \text{V}$ .

Discrimination was evaluated using AUC, treating ACR score, recalculated ACR TI-RADS, ultrasound-reported TI-RADS, and Bethesda as ordinal predictors. Agreement was assessed using Cohen kappa and weighted kappa as appropriate [10]. Exploratory logistic regression assessed associations with malignant histopathology after adjustment for age and sex; these models were considered hypothesis-generating.

#### 4.1. Ethical Considerations

The study used secondary data from institutional medical records, without patient contact or clinical intervention. The protocol was approved for retrospective medical record review, and the requirement for individual informed consent was waived because the analysis used anonymized secondary clinical data.



**Figure legend:** Bethesda had the highest AUC, followed by recalculated ACR TI-RADS and total ACR score. Ultrasound-reported TI-RADS showed lower discriminative performance in this dataset.

**Table 1:** Baseline characteristics and data availability.

| Variable                                 | n / statistic                           | % or median [IQR]                   |
|------------------------------------------|-----------------------------------------|-------------------------------------|
| Patients included                        | 177                                     | 100                                 |
| Female sex                               | 157                                     | 88.7                                |
| Male sex                                 | 20                                      | 11.3                                |
| Age, years                               | n=175; mean $\pm$ SD 53.1 $\pm$ 16.0    | 54.0 [40.5-65.0]; range 12.0-91.0   |
| Preoperative TSH                         | n=124; mean $\pm$ SD 26.16 $\pm$ 114.76 | 2.66 [1.43-4.14]; range 0.01-828.00 |
| Postoperative TSH                        | n=122; mean $\pm$ SD 16.36 $\pm$ 41.34  | 3.45 [1.51-8.32]; range 0.01-292.37 |
| Final histopathology available           | 168                                     | 94.9                                |
| Malignancy among patients with histology | 77/168                                  | 45.8                                |
| ACR TI-RADS recalculable                 | 169                                     | 95.5                                |

**Table 2:** Final histopathologic diagnosis.

| Diagnosis               | n  | % of total |
|-------------------------|----|------------|
| Benign nodule           | 88 | 49.7       |
| Follicular adenoma      | 2  | 1.1        |
| Hashimoto thyroiditis   | 1  | 0.6        |
| Papillary carcinoma     | 62 | 35.0       |
| Follicular carcinoma    | 14 | 7.9        |
| Anaplastic carcinoma    | 1  | 0.6        |
| No final histopathology | 9  | 5.1        |

**Table 3:** Recalculated ACR TI-RADS distribution.

| Category         | n  | % of total |
|------------------|----|------------|
| TR1              | 1  | 0.6        |
| TR2              | 9  | 5.1        |
| TR3              | 35 | 19.8       |
| TR4              | 93 | 52.5       |
| TR5              | 31 | 17.5       |
| Not recalculable | 8  | 4.5        |

**Table 4:** Ultrasound-reported TI-RADS distribution.

| Reported category              | n  | % of total |
|--------------------------------|----|------------|
| TI-RADS I                      | 1  | 0.6        |
| TI-RADS II                     | 9  | 5.1        |
| TI-RADS III                    | 44 | 24.9       |
| TI-RADS IV/4A/4B               | 88 | 49.7       |
| TI-RADS V                      | 12 | 6.8        |
| Not reported/non-interpretable | 23 | 13.0       |

**Table 5:** Bethesda cytology distribution.

| Bethesda category              | n  | % of total |
|--------------------------------|----|------------|
| I                              | 26 | 14.7       |
| II                             | 27 | 15.3       |
| III                            | 28 | 15.8       |
| IV                             | 56 | 31.6       |
| V                              | 24 | 13.6       |
| VI                             | 6  | 3.4        |
| Not reported/non-interpretable | 10 | 5.6        |

**Table 6:** Observed malignant histopathology among surgically verified cases by recalculated ACR TI-RADS category.

| Category | n with histology | Malignant | Observed malignant histopathology %, 95% CI |
|----------|------------------|-----------|---------------------------------------------|
| TR1      | 1                | 1         | 100.0 (20.7-100.0)                          |
| TR2      | 9                | 4         | 44.4 (18.9-73.3)                            |
| TR3      | 34               | 9         | 26.5 (14.6-43.1)                            |
| TR4      | 86               | 37        | 43.0 (33.1-53.6)                            |
| TR5      | 30               | 21        | 70.0 (52.1-83.3)                            |

**Table 7:** Observed malignant histopathology among surgically verified cases by Bethesda cytology category.

| Category | n with histology | Malignant | Observed malignant histopathology %, 95% CI |
|----------|------------------|-----------|---------------------------------------------|
| I        | 23               | 3         | 13.0 (4.5-32.1)                             |
| II       | 25               | 10        | 40.0 (23.4-59.3)                            |
| III      | 27               | 13        | 48.1 (30.7-66.0)                            |
| IV       | 56               | 21        | 37.5 (26.0-50.6)                            |
| V        | 23               | 22        | 95.7 (79.0-99.2)                            |
| VI       | 6                | 5         | 83.3 (43.6-97.0)                            |

**Table 8:** Diagnostic performance of ultrasound and cytologic thresholds against histopathology.

| Test / threshold                     | n   | TP/FP/TN/FN | Sensitivity % (95% CI) | Specificity % (95% CI) | PPV % (95% CI)   | NPV % (95% CI)   | Accuracy % (95% CI) |
|--------------------------------------|-----|-------------|------------------------|------------------------|------------------|------------------|---------------------|
| Recalculated ACR TI-RADS $\geq 4$    | 160 | 58/58/30/14 | 80.6 (70.0-88.0)       | 34.1 (25.0-44.5)       | 50.0 (41.0-59.0) | 68.2 (53.4-80.0) | 55.0 (47.3-62.5)    |
| Recalculated ACR TI-RADS $\geq 5$    | 160 | 21/9/79/51  | 29.2 (19.9-40.5)       | 89.8 (81.7-94.5)       | 70.0 (52.1-83.3) | 60.8 (52.2-68.7) | 62.5 (54.8-69.6)    |
| Ultrasound-reported TI-RADS $\geq 4$ | 149 | 46/50/33/20 | 69.7 (57.8-79.4)       | 39.8 (29.9-50.5)       | 47.9 (38.2-57.8) | 62.3 (48.8-74.1) | 53.0 (45.0-60.9)    |
| Ultrasound-reported TI-RADS $\geq 5$ | 149 | 10/2/81/56  | 15.2 (8.4-25.7)        | 97.6 (91.6-99.3)       | 83.3 (55.2-95.3) | 59.1 (50.8-67.0) | 61.1 (53.1-68.5)    |
| Bethesda $\geq$ III                  | 160 | 61/51/35/13 | 82.4 (72.2-89.4)       | 40.7 (30.9-51.3)       | 54.5 (45.2-63.4) | 72.9 (59.0-83.4) | 60.0 (52.3-67.3)    |
| Bethesda $\geq$ IV                   | 160 | 48/37/49/26 | 64.9 (53.5-74.8)       | 57.0 (46.4-66.9)       | 56.5 (45.9-66.5) | 65.3 (54.1-75.1) | 60.6 (52.9-67.9)    |
| Bethesda $\geq$ V                    | 160 | 27/2/84/47  | 36.5 (26.4-47.9)       | 97.7 (91.9-99.4)       | 93.1 (78.0-98.1) | 64.1 (55.6-71.8) | 69.4 (61.8-76.0)    |

**Table 9:** Area under the receiver operating characteristic curve.

| Ordinal predictor           | n   | AUC (95% bootstrap CI) |
|-----------------------------|-----|------------------------|
| Total ACR score             | 162 | 0.619 (0.530-0.704)    |
| Recalculated ACR TI-RADS    | 160 | 0.621 (0.541-0.701)    |
| Ultrasound-reported TI-RADS | 149 | 0.579 (0.492-0.661)    |
| Bethesda                    | 160 | 0.712 (0.633-0.786)    |

**Table 10:** Agreement analysis.

| Comparison                                      | n   | Statistic                | Value |
|-------------------------------------------------|-----|--------------------------|-------|
| Recalculated ACR vs ultrasound-reported TI-RADS | 148 | Unweighted kappa         | 0.245 |
| Recalculated ACR vs ultrasound-reported TI-RADS | 148 | Linear weighted kappa    | 0.287 |
| Recalculated ACR vs ultrasound-reported TI-RADS | 148 | Quadratic weighted kappa | 0.328 |
| Bethesda $\geq$ V vs histopathology             | 160 | Binary kappa             | 0.357 |
| ACR $\geq$ 5 vs histopathology                  | 160 | Binary kappa             | 0.200 |

**Table 11:** Exploratory multivariable logistic regression for malignant histopathology.

| Variable                               | Adjusted OR | 95% CI    | p      |
|----------------------------------------|-------------|-----------|--------|
| Recalculated ACR TI-RADS, per category | 1.56        | 1.00-2.45 | 0.052  |
| Bethesda, per category                 | 1.91        | 1.42-2.56 | <0.001 |
| Age, per year                          | 0.99        | 0.97-1.01 | 0.298  |
| Male sex                               | 0.77        | 0.26-2.27 | 0.638  |

**Table 12:** Combined diagnostic strategies using Bethesda and recalculated ACR TI-RADS.

| Rule                                           | n   | TP/FP/<br>TN/FN | Sensitivity<br>% (95% CI) | Specificity<br>% (95% CI) | PPV %<br>(95% CI)     | NPV %<br>(95% CI)   | Accuracy %<br>(95% CI) |
|------------------------------------------------|-----|-----------------|---------------------------|---------------------------|-----------------------|---------------------|------------------------|
| Bethesda $\geq$ V OR ACR<br>TI-RADS $\geq$ 5   | 154 | 37/9/74/34      | 52.1 (40.7-63.3)          | 89.2 (80.7-94.2)          | 80.4 (66.8-89.3)      | 68.5<br>(59.3-76.5) | 72.1<br>(64.5-78.6)    |
| Bethesda $\geq$ V AND ACR<br>TI-RADS $\geq$ 5  | 154 | 11/0/83/60      | 15.5 (8.9-25.7)           | 100.0 (95.6-100.0)        | 100.0<br>(74.1-100.0) | 58.0<br>(49.8-65.8) | 61.0<br>(53.2-68.4)    |
| Bethesda $\geq$ IV OR ACR<br>TI-RADS $\geq$ 4  | 154 | 63/65/18/8      | 88.7 (79.3-94.2)          | 21.7 (14.2-31.7)          | 49.2 (40.7-57.8)      | 69.2<br>(50.0-83.5) | 52.6<br>(44.7-60.3)    |
| Bethesda $\geq$ IV AND ACR<br>TI-RADS $\geq$ 4 | 154 | 41/25/58/30     | 57.7 (46.2-68.5)          | 69.9 (59.3-78.7)          | 62.1 (50.1-72.9)      | 65.9<br>(55.5-75.0) | 64.3<br>(56.5-71.4)    |

## 5. Results

### 5.1. Patient Flow and Baseline Characteristics

From the broader institutional universe initially considered for the project, data recovery and cleaning yielded 177 patient records suitable for analysis of the ultrasound-cytology-histopathology pathway. Final surgical histopathology was available for 168 patients, defining the primary analytic cohort for diagnostic accuracy estimates. This surgically verified subset should not be interpreted as the denominator for institutional thyroid nodule prevalence.

Most patients were female (157/177; 88.7%). The median age was 54 years (interquartile range, 40.5-65 years; range, 12-91 years). Preoperative TSH was markedly skewed, with a median of 2.66 and a range of 0.01-828.00, supporting median-based reporting rather than reliance on the mean alone.

### 5.2. Histopathologic Outcomes

The most frequent final histopathologic diagnosis was benign thyroid nodule, followed by papillary thyroid carcinoma. Malignant tumours were mainly papillary carcinoma, with fewer follicular carcinoma cases and one anaplastic carcinoma. Because histopathology was available primarily among surgically managed patients, the observed malignancy proportion should be interpreted as conditional on surgical verification rather than as the prevalence of malignancy among all institutional thyroid nodules.

### 5.3. Distribution Of Ultrasound and Cytologic Classifications

After recalculation from the ACR total score, TR4 was the most frequent category, followed by TR3 and TR5. Ultrasound-reported TI-RADS was unavailable or non-interpretable in 32

patients. This distinction between reported and recalculated categories was clinically relevant because it reflects a real-world documentation gap: the institution may have the clinical capacity and specialist expertise to complete the diagnostic pathway, but standardized reporting elements are not always captured uniformly across reports.

#### 5.4. Observed Histopathology by Preoperative Category

Observed malignant histopathology varied across preoperative ultrasound and cytology categories. These proportions describe surgically verified cases only; they should not be interpreted as intrinsic malignancy risks for each ACR TI-RADS or Bethesda category in an unselected thyroid nodule population.

#### 5.5. Diagnostic Accuracy

Recalculated ACR TI-RADS  $\geq 4$  had high sensitivity but low specificity. Raising the threshold to ACR TI-RADS  $\geq 5$  increased specificity but markedly reduced sensitivity. Bethesda  $\geq V$  showed high specificity and PPV, making it useful as a confirmatory threshold, but its low sensitivity limits its use as a stand-alone rule-out strategy. Bethesda  $\geq III$  was sensitive but nonspecific, as expected when indeterminate categories are considered positive.

#### 5.6. Discrimination and Agreement

Bethesda showed the highest discriminative performance, with an AUC of 0.712. Recalculated ACR TI-RADS and total ACR score showed modest discrimination, whereas ultrasound-reported TI-RADS performed less strongly. Agreement between recalculated ACR TI-RADS and ultrasound-reported TI-RADS was low to moderate, indicating imperfect translation of structured ultrasound components into the category documented in routine clinical reports.

#### 5.7. Exploratory Multivariable Model and Combined Strategies

In exploratory multivariable analysis, Bethesda category remained independently associated with malignancy, whereas recalculated ACR TI-RADS showed borderline association after adjustment for Bethesda, age, and sex. Combined diagnostic rules showed predictable trade-offs between sensitivity and specificity and should be interpreted as exploratory rather than as validated clinical algorithms.

### 6. Discussion

These findings show partial, not complete, agreement across the thyroid nodule diagnostic pathway. Bethesda cytology had the strongest discrimination for malignant histopathology; recalculated ACR TI-RADS performed modestly; and ultrasound-reported TI-RADS was less consistent with recalculated ACR categories. This pattern supports integrating ultrasound, cytology, and histopathology rather than treating any single preoperative category as definitive.

This interpretation is consistent with the intended roles of these tools. ACR TI-RADS supports standardized sonographic risk stratification and biopsy/follow-up thresholds, but it does not

diagnose cancer without size, clinical context, cytology, and surgical selection [2]. Bethesda provides morphologic cytologic evidence, especially in categories V and VI, but indeterminate and lower-risk categories remain affected by sampling, multilocularity, and selection bias [3].

Because this was a surgically enriched, histology-verified cohort, the 45.8% malignancy proportion should not be interpreted as the institutional prevalence of malignant thyroid nodules. The study is better understood as a histopathology-anchored audit of preoperative risk stratification within recoverable records from a larger institutional universe.

The most actionable signal was the gap between recalculated ACR TI-RADS and ultrasound-reported TI-RADS. In an experienced multidisciplinary centre, clinical expertise should complement structured reporting, not replace it. These findings do not mandate a single template, but they support discussion of practical strategies to harmonize ACR TI-RADS reporting, including consistent capture of the five domains, total score, final category, size, laterality, and nodule location [19].

Category-level discrepancies require caution. Because the dataset was organized by patient rather than by nodule, the imaged nodule, sampled nodule, and malignant surgical focus could not always be matched lesion by lesion. This is particularly relevant in multinodular glands and in patients with external imaging or cytology.

Combined diagnostic rules showed expected sensitivity-specificity trade-offs and should not be converted into management rules without prospective, nodule-level validation including nodule size and clinical risk factors.

The literature is best interpreted by context rather than direct numerical benchmarking. Colombian studies have evaluated TIRADS-Bethesda concordance, Bethesda-histopathology correlation, ACR TI-RADS agreement, and thyroid nodule characterization from different perspectives [7,11-13]. Regional and international studies similarly show that ultrasound risk stratification and Bethesda cytology are complementary, with histopathology remaining the decisive endpoint in surgically verified cohorts [4-6,15-18]. The contribution of this study is its institutional audit of the full preoperative pathway while separating reported TI-RADS from recalculated ACR TI-RADS.

### 7. Limitations

This study is limited by its retrospective design, secondary-data source, incomplete reconstruction of a larger initial institutional universe, and heterogeneous referral pathway in which not all reports were generated in-house. The analysis was organized by patient rather than nodule, limiting lesion-specific concordance. Nodule size was not recoverable, preventing assessment of size-based ACR TI-RADS biopsy/follow-up recommendations. The cohort was surgically enriched; malignancy proportions and predictive values are therefore not generalizable to all thyroid nodules evaluated at the institution.

Despite these limitations, distinguishing ultrasound-reported TI-RADS from recalculated ACR TI-RADS provides a practical quality signal for local review. The study did not measure avoidable FNAC, avoidable surgery, or delayed indicated surgery; these outcomes require prospective evaluation before management recommendations are adopted.

### Conclusion

In this retrospective cohort from Clinica de Occidente, Cali, Colombia, agreement across ACR TI-RADS, Bethesda cytology, and final surgical histopathology was partial. Bethesda cytology showed the strongest discrimination for malignant histopathology, recalculated ACR TI-RADS showed modest discrimination, and ultrasound-reported TI-RADS showed lower agreement with recalculated ACR categories. The findings support histopathology-anchored institutional audit and invite local strategies to harmonize ACR TI-RADS reporting, while recognizing the greater uniformity of Bethesda and final pathology in this dataset.

### 8. Acknowledgments

The authors acknowledge the institutional support of Clinica de Occidente and Fundacion Universitaria San Martin in the development of this retrospective medical record review.

### 9. Conflicts of Interest

The authors declare no conflicts of interest.

### 10. Data Availability Statement

The data that support the findings of this study are derived from institutional medical records and are not publicly available because of privacy and ethical restrictions. De-identified aggregate data may be available from the corresponding author upon reasonable request and with institutional approval.

### 11. Author Contributions

Conceptualization: E. Manuel Barajas, Eduardo Gutierrez Martin, Ivo Siljic, and Alejandro Ramirez Florin. Data curation and database review: Alejandro Ramirez Florin, Leslie Melissa Salazar, Laura Camila Martinez, and Jose David Tamayo. Statistical review and methodological interpretation: Alejandro Ramirez Florin and Jobany Castro Espinosa. Clinical interpretation: E. Manuel Barajas, Eduardo Gutierrez Martin, Ivo Siljic, and Alejandro Ramirez Florin. Writing - original draft and English editorial preparation: Alejandro Ramirez Florin. Writing - review and editing: all authors. Supervision: E. Manuel Barajas, Eduardo Gutierrez Martin, Ivo Siljic, and Jobany Castro Espinosa. All authors reviewed and approved the final version of the manuscript.

### 12. Editorial Support Statement

AI-assisted tools were used for editorial support. The authors reviewed and approved all analyses, interpretations, and manuscript content, and take full responsibility for the final article.

### References

1. Haugen BR, Alexander EK, Bible KC, Doherty GM, Mandel SJ. 2015 American Thyroid Association management guidelines for adult patients with thyroid nodules and differentiated thyroid cancer. *Thyroid*. 2016; 26(1): 1-133.
2. Tessler FN, Middleton WD, Grant EG, Hoang JK. ACR Thyroid Imaging, Reporting and Data System (TI-RADS): White paper of the ACR TI-RADS Committee. *J Am Coll Radiol*. 2017; 14(5): 587-595.
3. Ali SZ, Baloch ZW, Cochand-Priollet B, Schmitt FC, Vielh P. The 2023 Bethesda System for Reporting Thyroid Cytopathology. *J Am Soc Cytopathol*. 2023; 12(5): 319-325.
4. Osseis M, Jammal G, Kazan D, Noun R. Comparison between fine needle aspiration cytology with histopathology in the diagnosis of thyroid nodules. *J Pers Med*. 2023; 13(8): 1197.
5. Dhar L, Singh M, Jain SL, Rana D, Kumar J. Correlation of ultrasound-based TIRADS and the Bethesda system for reporting of thyroid cytopathology: A study in a tertiary care centre. *Cytopathology*. 2023; 34(6): 590-596.
6. George NA, Suresh S, Jiji V, Renu S, Thomas S. Correlation of TI-RADS and Bethesda scoring systems with final histopathology of thyroid nodules: An institutional experience. *Indian J Otolaryngol Head Neck Surg*. 2022; 74(Suppl 3): 5753-5758.
7. 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; 10: 1.
8. Bossuyt PM, Reitsma JB, Bruns DE, Gatsonis CA, Glasziou PP. STARD 2015: An updated list of essential items for reporting diagnostic accuracy studies. *BMJ*. 2015; 351: h5527.
9. Wilson EB. Probable inference, the law of succession, and statistical inference. *J Am Stat Assoc*. 1927; 22(158): 209-212.
10. Cohen J. A coefficient of agreement for nominal scales. *Educ Psychol Meas*. 1960; 20(1): 37-46.
11. Melo-Urbe MA, Sanabria A, Romero-Rojas A, Perez G, Vargas EJ. The Bethesda system for reporting thyroid cytopathology in Colombia: Correlation with histopathological diagnoses in oncology and non-oncology institutions. *J Cytol*. 2015; 32(1): 12-16.
12. Gonzalez Vasquez CM, Munoz Duran JA, Isaza Zapata S, Gonzalez Londono JF, Garcia Gomez V. Concordance of the ACR TI-RADS. *Radiologia (Engl Ed)*. 2021; 63(6): 469-475.
13. Buitrago-Gomez N, Garcia-Ramos A, Salom G, Cuesta-Castro DP. Sociodemographic, clinical and ultrasound characterization of thyroid nodule pathology and its association with malignancy in a Colombian high-complexity center. *Medicina de Familia SEMERGEN*. 2023; 49(6): 102015.
14. Santivanez JJ, Garcia-Lozano CA, Betancourt C, Sanchez JG, Mendieta LV. A real-world experience of active surveillance in Bethesda IV thyroid nodules. *Endocrine*. 2026; 91: 233.
15. Huang EYF, Kao NH, Lin SY. Concordance of the ACR TI-RADS Classification with Bethesda Scoring and Histopathology Risk Stratification of Thyroid Nodules. *JAMA Netw Open*. 2023; 6(9): e2331612.

16. Eissa MS, Sabry RM, Abdellateif MS. Evaluating the diagnostic role of ACR-TIRADS and Bethesda classifications in thyroid nodules highlighted by cyto-histopathological studies. *Exp Clin Endocrinol Diabetes*. 2024; 132(11): 596-606.
17. Rahal Junior A, Falsarella PM, Rocha RD, Lima JPBC, Iani MJ. Correlation of Thyroid Imaging Reporting and Data System [TIRADS] and fine needle aspiration: experience in 1,000 nodules. *Einstein (Sao Paulo)*. 2016; 14(2): 119-123.
18. Kraus-Fischer G, Alvarado-Bachmann R, de Rienzo-Madero B, Nunez-Garcia E. Correlation between the Bethesda system for thyroid nodules and post-thyroidectomy histopathological diagnosis. *Rev Med Inst Mex Seguro Soc*. 2020; 58(2): 114-121.
19. Griffin AS, Mitsky J, Rawal U, Bronner AJ, Tessler FN, Hoang JK. Improved quality of thyroid ultrasound reports after implementation of the ACR Thyroid Imaging Reporting and Data System nodule lexicon and risk stratification system. *J Am Coll Radiol*. 2018; 15(5): 743-748.