The Journal of Clinical Endocrinology & Metabolism, 2022, **107**, 3309–3319
https://doi.org/10.1210/clinem/dgac537
Advance access publication 14 September 2022
Clinical Research Article

Clinical Use of Raman Spectroscopy Improves Diagnostic Accuracy for Indeterminate Thyroid Nodules

Andrea Palermo,[1,2] Armida Sodo,[3] Anda Mihaela Naciu,[1] Michael Di Gioacchino,[3] Alessio Paolucci,[3] Alessandra di Masi,[3] Daria Maggi,[4] Pierfilippo Crucitti,[5] Filippo Longo,[5] Eleonora Perrella,[6] Chiara Taffon,[6] Martina Verri,[6] Maria Antonietta Ricci,[3] and Anna Crescenzi[6]

[1]Unit of Metabolic Bone and Thyroid Disorders, Fondazione Policlinico Universitario Campus Bio-Medico, Via Alvaro del Portillo, 200 - 00128 Roma, Italy
[2]Unit of Endocrinology and Diabetes, Department of Medicine and Surgery, Università Campus Bio-Medico di Roma, Via Alvaro del Portillo, 21 - 00128 Roma, Italy
[3]Dipartimento di Scienze, Università Roma Tre, Rome, Italy
[4]Unit of Endocrinology and Diabetes, Fondazione Policlinico Universitario Campus Bio-Medico, Rome, Italy
[5]Unit of Thoracic Surgery, Fondazione Policlinico Universitario Campus Bio-Medico, Rome, Italy
[6]Unit of Pathology, Fondazione Policlinico Universitario Campus Bio-Medico, Rome, Italy

Correspondence: Andrea Palermo, MD, PhD, Unit of Metabolic Bone and Thyroid Disorders, Fondazione Policlinico Universitario Campus Bio-Medico, Via Alvaro del Portillo 200, 00128, Rome, Italy, Email: a.palermo@unicampus.it.

Abstract

Background and Objective: Molecular analysis of thyroid fine-needle aspiration (FNA) specimens is believed to improve the management of indeterminate nodules. Raman spectroscopy (RS) can differentiate benign and malignant thyroid lesions in surgically removed tissues, generating distinctive structural profiles. Herein, the diagnostic performance of RS was tested on FNA biopsies of thyroid gland.

Design: Prospective, blinded, and single-center study.

Methods: We enrolled 123 patients with indeterminate or more ominous cytologic diagnoses (TIR3A-low-risk indeterminate lesion, TIR3B-high-risk indeterminate lesion, TIR4-suspicious of malignancy, TIR5-malignant). All subjects were surgical candidates (defined by international guidelines) and submitted to FNA procedures for RS analysis. We compared RS data, cytologic findings, and final histologic assessments (as reference standard) using various statistical techniques.

Results: The distribution of our study population was as follows: TIR3A:37, TIR3B:32, TIR4:16, and TIR5:38. In 30.9% of patients, histologic diagnoses were benign. For predicting thyroid malignancy in FNA samples, the overall specificity of RS was 86.8%, with 86.5% specificity in indeterminate cytologic categories. In patients with high-risk ultrasound categories, the specificity of RS increased to 87.5% for TIR3A, reaching 100% for TIR3B. Benign histologic diagnoses accounted for 72.9% of patients classified as TIR3A and 31.3% of those classified as TIR3B. Based on positive RS testing, unnecessary surgery was reduced to 7.4% overall (TIR3A-33.3%, TIR3B-6.7%).

Conclusions: This premier use of RS for thyroid cytology confirms its role as a valuable diagnostic tool and a valid alternative to molecular studies, capable of improving the management of indeterminate nodules and reducing unnecessary surgery.

Key Words: Raman spectroscopy, thyroid, thyroid nodule, indeterminate cytology

Thyroid nodules have been increasing in incidence worldwide. To avoid their overdiagnosis and overtreatment, international guidelines are periodically revised. Neck ultrasonography (US) is generally the first diagnostic step, coupled with fine-needle aspiration (FNA) biopsy. Both techniques are sufficiently informative for a large part of such lesions to determine whether surgery is or is not indicated. However, cytologic assessments of the thyroid do have certain limitations. A lesser percentage (10-26%) of all nodules will display indeterminate attributes (equivocal or atypical nuclear or architectural features) in cytologic preparations (1–3). This is often problematic, because clinical management is then left unsettled, usually prompting diagnostic surgical referrals to date.

Full genomic assessment of thyroid carcinoma has subsequently identified driver gene mutations with specific roles in cancer subtypes (4). In an attempt to reduce the numbers of cytologically indeterminate thyroid nodules, much of the molecular testing performed during the past 10 years has involved FNA samples, developing several panels of genomic alterations for diagnostic purposes (3, 5). The risk of malignancy (based on mutational status) may thus be gauged by the frequency of a given genetic alteration in a specific neoplastic condition (6). However, the sensitivity, specificity, and expense of molecular diagnostics must be addressed to justify this approach in medical practice. It is not yet worldwide regularly used in clinical workups. Moreover, from a scientific point of view, molecular analysis (even if extensive) offers

Received: 23 May 2022. **Editorial Decision:** 12 September 2022. **Corrected and Typeset:** 30 September 2022

no insights into tumor metabolic functions. Although one may capture existing genetic alterations in cancer thyrocytes, the biochemical ramifications remain elusive.

Several experimental techniques currently used in chemical-physical research have promised substantial contributions to both biochemical investigations and clinical cancer diagnosis (7, 8). Among these, Raman spectroscopy (RS) has been successfully applied within the past 2 decades as diagnostic tool for various cancers (9–12). This technique detects molecular compositions of human samples (fluids, cells, tissues, etc) by analyzing inelastic light scatter after laser irradiation (13). RS is a nondestructive, high-sensitivity analytic technique requiring no sample preparation and quite suitable for in vivo analyses.

In the context of thyroid pathology, RS has served to distinguish between normal tissues and follicular thyroid lesions (14, 15) and between benign and malignant thyroid tumors in surgically removed tissues (16–18). To our knowledge, there have been no investigations to date of its use in the preoperative screening of thyroid FNA samples. This particular study was undertaken to evaluate the diagnostic performance of RS in FNA cytology of thyroid nodules, relying on surgical histopathology as reference standard. The potential to reduce unnecessary resection of cytologically indeterminate lesions was also assessed.

Materials and Methods

Study Design

This prospective, blinded, and single-center study took place at the thyroid outpatient clinic of the Metabolic Bone and Thyroid Disorders Unit of Fondazione Policlinico Universitario Campus Bio-Medico, Rome, between September 2019 and July 2021.

Ethics Approval and Consent to Participate

Our protocol adhered to the Declaration of Helsinki and to the International Conference on Harmonization Good Clinical Practice, receiving approval of Campus Bio-Medico University ethics committees (31/19 PAR ComEt CBM from July 26, 2019). All participants granted informed consent, allowing use of their anonymized information for data analysis.

Participants

Patients eligibility criteria were as follows: (1) ≥18 years old; (2) one or more thyroid nodules with medium-high ultrasound risk of malignancy [Thyroid Imaging Reporting & Data System (TI-RADS)] score ≥3) (19); (3) referral to our outpatient clinic for thyroid FNA, according to clinical guidelines (20); and (4) willingness to provide material for RS. Once thyroid FNA cytology was reported, only those subjects with at least 1 nodule categorized as indeterminate, suspicious, or malignant and warranting thyroid surgery (21) were included in the final analysis (Fig. 1). Although the malignancy rate in noninvasive follicular thyroid neoplasm with papillary-like nuclear features is still debated (22), we considered it as a malignant thyroid lesion. Patients with high-risk thyroid nodules received indication to surgery independently of lesion volume, as suggested by national guidelines. In case of cytological diagnosis of low-risk thyroid nodules the patient received indication to surgery only in the presence of other criteria such as nodule volume and/or compressive symptoms and/or cosmetic discomfort, as suggested by national guidelines. Participants were recruited from and managed at Fondazione Policlinico Universitario Campus Bio-Medico, Rome; all of them granted written informed consent.

Procedures

We screened consecutive patients with thyroid nodular pathology who were referred to our outpatient clinic for FNA. Before FNA procedures, all subjects submitted to thyroid US evaluations in a frequency range of 10 to 12 MHz, using a MyLab 50 system (Esaote, Genova, Italy). US scans of thyroid glands and neck areas were performed by 2 experienced physician specialists the endocrinology unit of our hospital. Nodules were classified according to American College of Radiology TI-RADS (19) risk stratification criteria, without prior knowledge of cytologic results. For any disagreement, both met in separate session to reach a consensus. Patient medical records supplied demographic data, including age, gender, family history of thyroid cancer, prior thyroid peroxidase antibody or thyroglobulin antibody positivity, and most recent serum thyroid stimulating hormone levels.

Patients harboring 1 or more thyroid nodules of medium-high malignancy risk by US (TI-RADS score ≥3) and who submitted to thyroid FNA were eligible for study. FNA was done freehand under US guidance, using a 25-gauge needle. On average, 2 to 3 passes were made in each nodule. A single drop from 1 of these passes was utilized for RS analysis. Passage across thyroid vessels was avoided to prevent localized bleeding. In mixed nodules, solid areas were selected. Specimens were withdrawn into syringes then routinely smeared onto glass slides, sprayed with ethanol-based fixative, and stained by Papanicolaou method. All slides were evaluated by expert cytopathologists of the pathology unit, applying the Italian Reporting System for Thyroid Cytology (23) as follows: TIR1 (nondiagnostic), TIR1C (nondiagnostic-cystic), TIR2 (nonmalignant/benign), TIR3A (low-risk indeterminate lesion), TIR3B (high-risk indeterminate lesion), TIR4 (suspicious of malignancy), or TIR5 (malignant). As previously published (24), the 2 diagnostic subcategories of indeterminate nodules with low (TIR3A) or high risk (TIR3B) of malignancy corresponded with classes III and IV, respectively, of the Bethesda System for Reporting Thyroid Cytopathology (25). After producing smears for diagnostic purposes, single drops (maximum diameters ~4 mm) of cell suspensions were placed on positively charged microscopy slides (without smearing) and stored unstained at −80°C for later RS analysis. All data and specimens were anonymized per approved protocol (Fig. 2A). Slides were regularly carried on cold containers.

RS analyses were conducted at the Science Department of Roma Tre University (Rome, Italy), using an inVia Micro-Raman Spectrometer (Renshaw, Wotton-under-Edge, UK) equipped with a solid-state diode laser source at 532 nm (100 mW nominal output power) and a confocal microscope [50× long working distance objective (Fig. 2B and 2C); Leica Microsystems, Wetzlar, Germany] to focalize the incident laser beam and collect back-scattered light. Elastic light scatter was eliminated by using a holographic edge filter, and inelastic scatter intensity was dispersed by a diffraction grating (1800 grooves/mm) on a Peltier cooled 1024 × 256 pixel CCD detector. The final instrumental resolution was on the order of 1 cm^{-1}. To prevent photodamage, the laser power at the sample was controlled by neutral

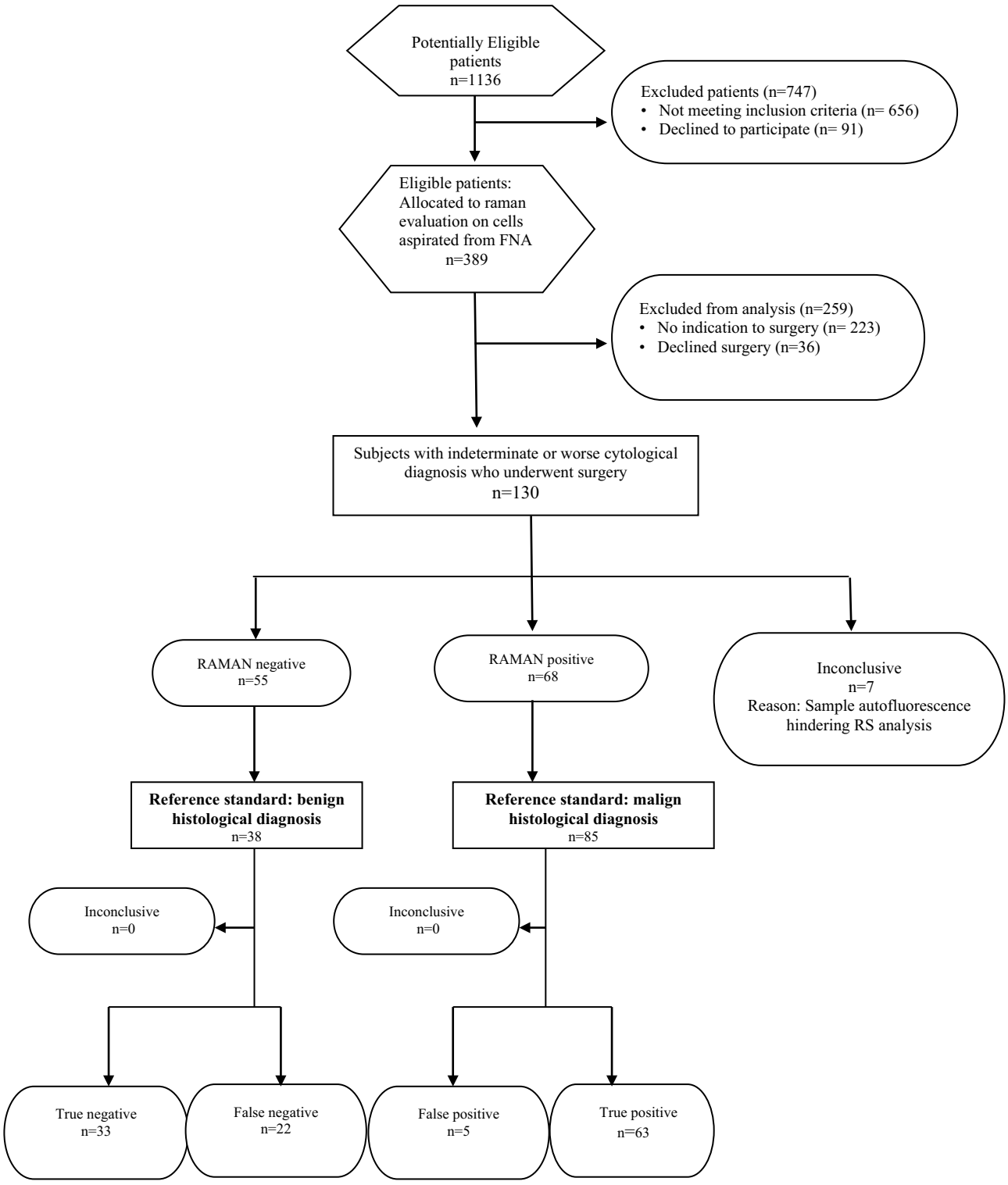

Figure 1. STARD flow diagram. STARD flow diagram for a study of 130 patients with indeterminate or worse cytological diagnosis who underwent surgery.

density filters. Spectra were collected in the 100 to 3600 cm^{-1} range. For each sample, 5 measurement points were selected, and 5 scans per point were recorded (total integration time, 50 seconds). Experimental conditions and data acquisition were driven by Wire software (Renishaw), which also enabled preliminary data reduction (eg, background and fluorescence subtraction). MATLAB (MathWorks, Natick, MA, USA) and Origin 9.0 (OriginLab Corp, Northampton, MA, USA) software applications served for data analysis.

Data in the spectral range of 600 to 1800 cm^{-1} provided the so-called molecular fingerprints of investigated samples, reflecting vibrational modes of intramolecular bonds. Each Raman active mode peaks at a characteristic frequency, which distinguishes 1 molecule from another (eg, carotenoid

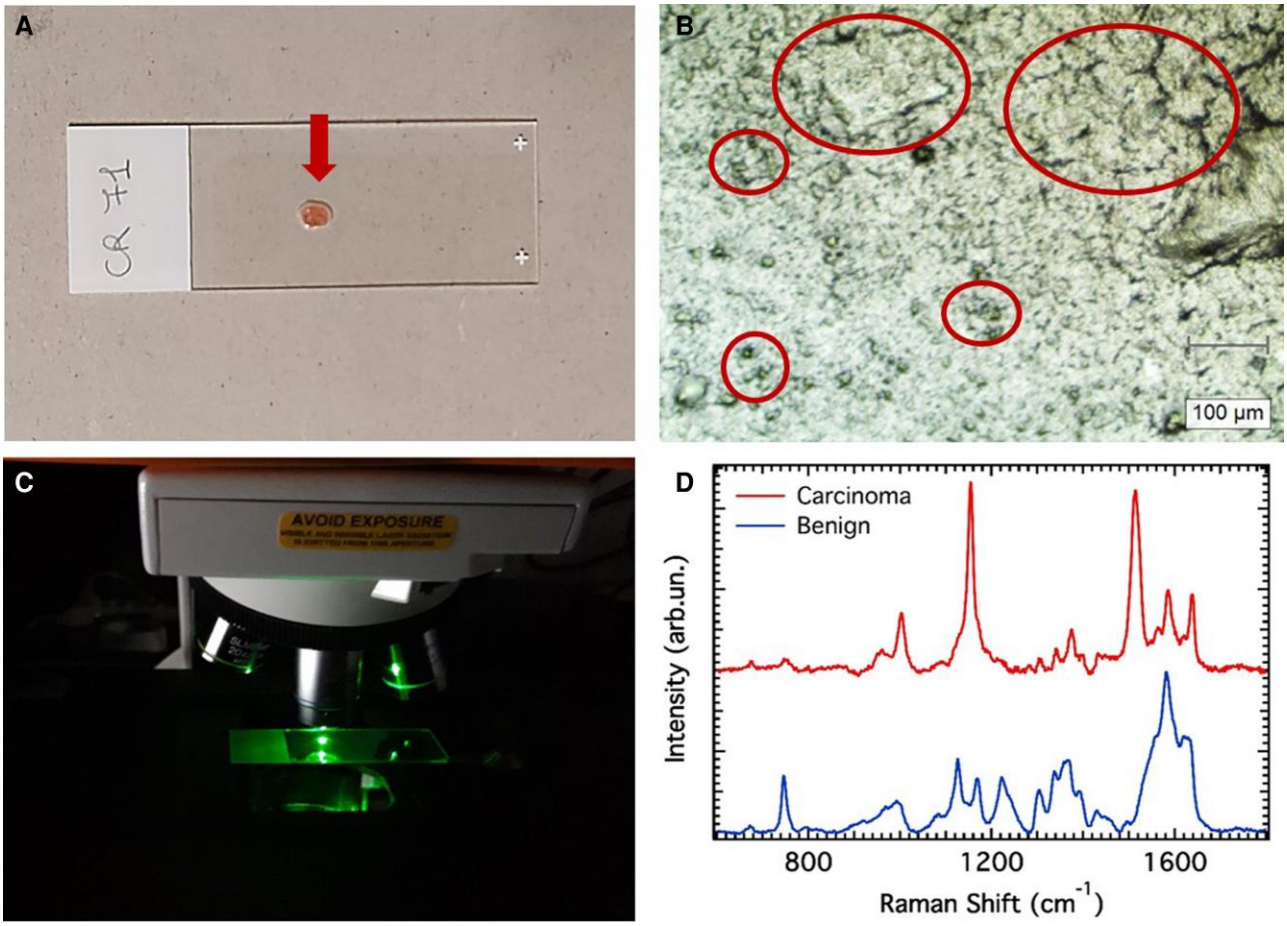

Figure 2. Procedure for Raman spectroscopy. (A) A single drop (maximum diameters, ~4 mm) of cell suspensions from fine-needle aspiration is placed on positively charged microscopy slides, without smearing (arrow). (B) Bright field evaluation at low power shows unstained cellular material with cell clusters of different size (circles). (C) Ramon spectroscopy analyses are conducted on the cell clusters using the inVia Micro-Raman Spectrometer equipped with a solid-state diode laser source at 532 nm. (D) Data in the spectral range of 600 to 1800 cm^{-1} provided the so-called molecular fingerprints of investigated samples. Differences in the spectral profile are evident between benign and carcinoma cells.

molecules from cytochrome *c*) or from other cell components of interest. This observation is key in spectroscopic diagnosis of thyroidal cancer (17), due to similarity/dissimilarity among Raman spectra (Fig. 2D). Our analysis was rooted in 2 statistical applications, namely agglomerative hierarchical clustering analysis (AHCA) and K-means (KM) methods. In these techniques, each spectrum is defined by its distance with respect to an origin, thus generating a dissimilarity matrix. AHCA analysis was done by choosing either the Euclidean or the correlation distance, whereas KM analysis was based on Euclidean distance alone.

The AHCA method seeks to sort Raman spectra into homogeneous groups (or clusters) by their measured distances. We used the complete-linkage algorithm to enhance differences among clusters. It does so by initially pairing the most similar spectra (those of lowest dissimilarity value in the matrix) and then searching for greatest distances between such pairs and other data. This algorithm is iterated until all spectra are incorporated into well-separated clusters. The output is presented as a dendrogram.

The KM method initially sorts spectra among an a priori number of random centroids (k) and iteratively minimizes the variance within each cluster by updating the centroids and their spectral distances. Each data entry (ie, Euclidean distance associated with an experimental Raman spectrum) then gravitates to a cluster with the closest centroid. This process continues until convergence. We have found that for the set of data analyzed here, $k = 3$ proved sufficient. Outputs of this method were 3 classes of spectra and the Raman spectrum of each centroid.

Final histological diagnosis was reported in agreement with the IV edition of WHO Classification of Tumours of Endocrine Organs (26).

Outcomes

Analyzed patient characteristics included age, sex, nodular volume (measured by US), Italian/Bethesda cytologic categories, thyroid function, peroxidase antibody or thyroglobulin antibody positivity, and pathology reports following surgical resection.

Primary outcome measures were sensitivity, specificity, positive predictive value (PPV), negative predictive value (NPV), benign call rate (BCR, % of RS screens showing benign or negative results) and positive call rate (PCR, the ratio of positive RS screens to total subjects in a cytologic category) for RS analysis of cytologic samples (surgical histopathology being the reference standard), and the ability of RS to reduce unnecessary surgery in patients with indeterminate cytologic reports (TIR3A/Bethesda III or TIR3B/Bethesda IV).

As a secondary outcome measure, we assessed the diagnostic performance of RS in subjects with high TI-RADS scores (4 and 5) and indeterminate thyroid nodule cytology.

Statistical Analysis

We based our calculated sample size on a study by Bongiovanni et al (27). These authors had cited a specificity of 61.3% at a cutpoint of TIR3B for standard FNA detection of thyroid cancer before surgery. To reach a specificity ≥80% via the new technique (vs 61.3% by standard procedure), 123 subjects were needed (power, 0.90; alpha, 0.05). Therefore, we closed enrollment once 123 recruits with good-quality Raman spectra were enlisted.

Having checked all variables initially for duplicate, missing, or illogic values and searching for outliers or for nonnormal distributions of continuous variables, the samples were then grouped by malignant status and FNA results. Two-sample *t*-test or analysis of variance served to compare normally distributed variables (log-normal variables were log-transformed beforehand). To compare categorical variables, Pearson's chi-squared test was applied. There were no variables showing other types of distribution.

Next, we assessed the 3 RS analyses (ie, AHCA based on either the Euclidean or the correlation distance and KM method based on Euclidean distance alone; pairwise, total of 3 comparisons), using Cohen's kappa coefficient and Cronbach's alpha with bootstrapped 95% confidence intervals (95% CI). After 1000 replications, AHCAcorr and KM showed exceptional agreement, so KM sufficed as the single best measure going forward.

We also evaluated the performance of RS (sensitivity, specificity, PPV, and NPV) in identifying malignant nodules preoperatively for differing subsets of cytologically indeterminate thyroid nodules (TIR3A and TIR3B) and in patients overall, and we repeated this analysis for those nodules with TI-RADS scores of 4 or 5.

Finally, we examined rates of unnecessary surgery in patients overall and in those with positive RS screens, looking also at differing subsets of cytologically indeterminate thyroid nodules (TIR3A and TIR3B).

Results

Seven of the 130 eligible nodules proved inadequate for RS analysis due to sample autofluorescence. Ultimately, 123 subjects (age range, 21-79 years; men, 66.7%) qualified for analysis, each undergoing thyroid cancer screening and subsequent thyroidectomy (with histologic diagnosis).

Tables 1 and 2 list patient samples according to thyroid malignancy (histologic diagnosis) and FNA results, respectively. Overall, 85 of 123 patients (69.1%) had confirmed thyroid cancers (30.9% follicular adenoma, 58.53% papillary carcinoma, 7.32% follicular variant of papillary carcinoma, 0.81% follicular carcinoma, 0.81% medullary thyroid carcinoma, 1.63% Hurthle cell carcinoma). Tg-Ab positivity was associated with both thyroid malignancy and FNA positivity, as were lesion echogenicity, echostructure, vascularization, microcalcifications, irregular margins, dimensions, TI-RADS score, and RS analysis. Patients with histologically confirmed thyroid cancers had higher preoperative TI-RADS scores ($P = 0.003$) and suspicious or malignant FNA status ($P < 0.001$) and were more likely to test positive by RS ($P < 0.001$). Likewise, those with suspicious or malignant FNA categories had higher TI-RADS scores ($P = 0.036$) and were more likely to show RS positivity ($P < 0.001$). In 27.0% (10 of 37) of our patients grouped as TIR3A, excisions were confirmatory of thyroid cancer, compared with 68.8% in the TIR3B group. Cancer was excluded in just 1 of 16 patients with TIR4 designations.

Three separate multivariate statistical analyses of Raman spectra are shown for comparison. Both AHCAcorr and KM methods revealed exceptional agreement, with values of 0.97 (95% CI: 0.92-1.00) for Cohen's kappa coefficient and 0.98 (95% CI: 0.96-1.01) for Cronbach's alpha. KM only was then applied in subsequent evaluations (Table 3).

Table 4 shows the performance of RS in cytologically indeterminate thyroid nodules and in patients overall. Altogether, RS analysis showed a specificity of 86.8% in predicting thyroid malignancy before surgery. In patients classified as TIR3B, specificity reached 90.0%, whereas in those classified TIR3A, specificity was still 85.2%. In pooled TIR3A and TIR3B subsets, the specificity of RS analysis was 86.5%. BCR was 67.5% for TIR3A, 53.1% for TIR3B, and 60.8% for both (TIR3A + TIR3B). PCR values were 32.4%, 46.8%, and 39.1% for TIR3A, TIR3B, and TIR3A + TIR3B, respectively. All supplementary materials are located in a digital research materials repository (28).

When restricting analysis to those patients scored as 4 or 5 by TI-RADS, the specificity of RS increased to 87.5% for TIR3A lesions, reaching 100% at TIR3B level (Table 5).

To address the issue of unnecessary surgery, Table 6 indicates proportions of patients whose thyroid nodules did not prove cancerous after surgery. Overall, 30.9% of surgically treated patients had benign tissue diagnoses. That proportion was as high as 72.9% for TIR3A lesions, falling to 31.3% in the TIR3B subset. However, in patients testing positive by RS analysis, unnecessary surgeries (ie, false-positive rates) were 7.4% overall, 33.3% for TIR3A, and 6.7% for TIR3B. *P*-values for differences displayed ranged from 0.064 to <0.001.

Discussion

Herein, we have demonstrated for the first time the diagnostic utility of RS in thyroid cytology, primarily in nodules with indeterminate features. The latter represent a growing clinical problem. Through accurate thyroidal US-based triage, we managed to reject prospective nodules of benign inclination, so the number of FNAs performed on US-determined suspicious nodules rose proportionally, with a parallel increase in indeterminate cytologic diagnoses. By design, this approach to thyroid cytology was intended to improve the diagnostic accuracy of FNA and to reduce the need for diagnostic surgery.

In recent years, molecular testing has been further improved by a quest for multiple gene alterations, thus significantly enhancing sensitivity and NPV. However, molecular testing suffers from relatively low specificity and PPV. Particularly in Hürthle cell tumors, there is limited clinical validation and little specific cancer risk assessment for the various molecular alterations entailed (5). Consequently, the use of molecular profiling to define malignancy risk in indeterminate lesions has yet to enter mainstream clinical practice, due to inherent drawbacks mentioned here, high cost, and limited availability. These shortcomings and the need to improve the diagnostic

Table 1. Sample description by histological diagnosis

Factor and category		Histological diagnosis		*P* value
	Total N = 123	Benign N = 38	Malignant N = 85	
Male	66.7% (82/123)	57.9% (22/38)	70.6% (60/85)	0.17
Age	53.3 ±SD 12.8	53.6 ±SD 12.2	53.2 ±SD 13.1	0.85
AbTpo positive	16.3% (20/123)	18.4% (7/38)	15.3% (13/85)	0.66
AbTg positive	11.4% (14/123)	2.6% (1/38)	15.3% (13/85)	0.041
TSH	1.3 (×/1.9)	1.1 (×/2.0)	1.3 (×/1.8)	0.15
Calcitonine positive	1.6% (2/123)	0.0% (0/38)	2.4% (2/85)	0.34
On levothyroxine	11.4% (14/123)	7.9% (3/38)	12.9% (11/85)	0.42
On methimazole	0.0% (0/123)	0.0% (0/38)	0.0% (0/85)	
Familiarity for thyroid cancer	9.8% (12/123)	7.9% (3/38)	10.6% (9/85)	0.64
Localisation				0.38
Right lobe	55.3% (68/123)	55.3% (21/38)	55.3% (47/85)	
Left lode	41.5% (51/123)	44.7% (17/38)	40.0% (34/85)	
Isthmus	3.3% (4/123)	0.0% (0/38)	4.7% (4/85)	
Echogenicity				<0.001
Hypoechogenic	39.0% (48/123)	15.8% (6/38)	49.4% (42/85)	
Isoechogenic	26.8% (33/123)	50.0% (19/38)	16.5% (14/85)	
Mixed hypoechogenic and isoechogenic	34.1% (42/123)	34.2% (13/38)	34.1% (29/85)	
Echostructure				0.002
Solid	65.0% (80/123)	44.7% (17/38)	74.1% (63/85)	
Mixed solid and liquid	35.0% (43/123)	55.3% (21/38)	25.9% (22/85)	
Vascularisation				0.009
Avascular	8.9% (11/123)	0.0% (0/38)	12.9% (11/85)	
Perilesion	27.6% (34/123)	18.4% (7/38)	31.8% (27/85)	
Perilesion and intralesion	63.4% (78/123)	81.6% (31/38)	55.3% (47/85)	
Microcalcifications	12.2% (15/123)	2.6% (1/38)	16.5% (14/85)	0.030
Macrocalcifications	22.8% (28/123)	13.2% (5/38)	27.1% (23/85)	0.089
Irregular margins	24.4% (30/123)	5.3% (2/38)	32.9% (28/85)	<0.001
Dimensions				
Ap (mm)	11.1 (×/1.8)	16.2 (×/1.9)	9.4 (×/1.7)	<0.001
LL (mm)	12.6 (×/1.9)	19.2 (×/1.9)	10.4 (×/1.7)	<0.001
Long (mm)	15.9 (×/1.9)	25.3 (×/1.8)	12.9 (×/1.7)	<0.001
Volume (mL)	2.3 (×/6.1)	8.2 (×/6.0)	1.3 (×/4.6)	<0.001
TI-RADS				0.003
2	1.6% (2/122)	5.4% (2/37)	0.0% (0/85)	
3	22.1% (27/122)	32.4% (12/37)	17.6% (15/85)	
4	59.0% (72/122)	59.5% (22/37)	58.8% (50/85)	
5	17.2% (21/122)	2.7% (1/37)	23.5% (20/85)	
FNA				<0.001
3A	30.1% (37/123)	71.1% (27/38)	11.8% (10/85)	
3B	26.0% (32/123)	26.3% (10/38)	25.9% (22/85)	
4	13.0% (16/123)	2.6% (1/38)	17.6% (15/85)	
5	30.9% (38/123)	0.0% (0/38)	44.7% (38/85)	
FNA 3B or higher	69.9% (86/123)	28.9% (11/38)	88.2% (75/85)	<0.001
RAMAN—positive				
AHCA (correlation distance)	56.9% (70/123)	13.2% (5/38)	76.5% (65/85)	<0.001
AHCA (euclidean distance)	79.7% (98/123)	42.1% (16/38)	96.5% (82/85)	<0.001
KM	55.3% (68/123)	13.2% (5/38)	74.1% (63/85)	<0.001

For normally distributed variables, data are presented as mean (SD) with *P* value from 2-sample *t* tests. For log-normal variables, data are presented as geometric mean (×/geometric SD) with *P* value from 2-sample *t* tests on logged values. For categorical variables, data are presented as % (n/total) with *P* value from Pearson's chi-squared test.

Table 2. Sample description by FNA results

Factor and category	FNA				*P* value
	TIR 3A N = 37	TIR 3B N = 32	TIR 4 N = 16	TIR 5 N = 38	
Male	59.5% (22/37)	68.8% (22/32)	50.0% (8/16)	78.9% (30/38)	0.14
Age	54.9 ±SD 12.7	49.7 ±SD 13.1	54.6 ±SD 12.0	54.3 ±SD 12.8	0.33
AbTpo positive	18.9% (7/37)	15.6% (5/32)	6.2% (1/16)	18.4% (7/38)	0.68
AbTg positive	2.7% (1/37)	9.4% (3/32)	6.2% (1/16)	23.7% (9/38)	0.029
TSH	1.1 (×/2.1)	1.5 (×/1.6)	1.0 (×/1.7)	1.4 (×/1.9)	0.039
Calcitonine positive	0.0% (0/37)	6.2% (2/32)	0.0% (0/16)	0.0% (0/38)	0.12
On levothyroxine	13.5% (5/37)	12.5% (4/32)	6.2% (1/16)	10.5% (4/38)	0.88
On methimazole	0.0% (0/37)	0.0% (0/32)	0.0% (0/16)	0.0% (0/38)	
Familiarity for thyroid cancer	2.7% (1/37)	12.5% (4/32)	0.0% (0/16)	18.4% (7/38)	0.062
Localisation					0.28
Right lobe	64.9% (24/37)	59.4% (19/32)	50.0% (8/16)	44.7% (17/38)	
Left lode	35.1% (13/37)	40.6% (13/32)	43.8% (7/16)	47.4% (18/38)	
Isthmus	0.0% (0/37)	0.0% (0/32)	6.2% (1/16)	7.9% (3/38)	
Echogenicity					<0.001
Hypoechogenic	13.5% (5/37)	40.6% (13/32)	56.2% (9/16)	55.3% (21/38)	
Isoechogenic	48.6% (18/37)	12.5% (4/32)	18.8% (3/16)	21.1% (8/38)	
Mixed hypoechogenic and isoechogenic	37.8% (14/37)	46.9% (15/32)	25.0% (4/16)	23.7% (9/38)	
Echostructure					0.008
Solid	45.9% (17/37)	62.5% (20/32)	87.5% (14/16)	76.3% (29/38)	
Mixed solid and liquid	54.1% (20/37)	37.5% (12/32)	12.5% (2/16)	23.7% (9/38)	
Vascularisation					0.13
Avascular	0.0% (0/37)	9.4% (3/32)	6.2% (1/16)	18.4% (7/38)	
Perilesion	27.0% (10/37)	34.4% (11/32)	18.8% (3/16)	26.3% (10/38)	
Perilesion and intralesion	73.0% (27/37)	56.2% (18/32)	75.0% (12/16)	55.3% (21/38)	
Microcalcifications	2.7% (1/37)	12.5% (4/32)	25.0% (4/16)	15.8% (6/38)	0.11
Macrocalcifications	16.2% (6/37)	15.6% (5/32)	43.8% (7/16)	26.3% (10/38)	0.11
Irregular margins	10.8% (4/37)	28.1% (9/32)	37.5% (6/16)	28.9% (11/38)	0.12
Dimensions					
Ap (mm)	20.3 (×/1.5)	8.4 (×/1.7)	9.0 (×/1.6)	8.6 (×/1.6)	<0.001
LL (mm)	24.6 (×/1.6)	9.2 (×/1.5)	9.6 (×/1.5)	9.5 (×/1.6)	<0.001
Long (mm)	31.7 (×/1.5)	12.3 (×/1.6)	11.1 (×/1.5)	11.8 (×/1.6)	<0.001
Volume (mL)	16.4 (×/3.5)	1.0 (×/3.8)	1.0 (×/3.2)	1.0 (×/3.9)	<0.001
TI-RADS					0.036
2	5.4% (2/37)	0.0% (0/31)	0.0% (0/16)	0.0% (0/38)	
3	32.4% (12/37)	19.4% (6/31)	18.8% (3/16)	15.8% (6/38)	
4	59.5% (22/37)	67.7% (21/31)	56.2% (9/16)	52.6% (20/38)	
5	2.7% (1/37)	12.9% (4/31)	25.0% (4/16)	31.6% (12/38)	
Histologically-confirmed malignancy	27.0% (10/37)	68.8% (22/32)	93.8% (15/16)	100.0% (38/38)	<0.001
RAMAN positivity					
AHCA (correlation distance)	32.4% (12/37)	46.9% (15/32)	81.2% (13/16)	78.9% (30/38)	<0.001
AHCA (euclidean distance)	51.4% (19/37)	81.2% (26/32)	100.0% (16/16)	97.4% (37/38)	<0.001
KM	32.4% (12/37)	46.9% (15/32)	81.2% (13/16)	73.7% (28/38)	<0.001

For normally distributed variables, data are presented as mean (SD) with *P* value from analysis of variance. For log-normal variables, data are presented as geometric mean (×/geometric SD) with *P* value from analysis of variance on logged values. For categorical variables, data are presented as % (n/total) with *P* value from Pearson's chi-squared test.

performance of FNA cytology in the context of indeterminate results have fueled the pursuit of new biomarkers.

RS seems particularly appealing in this regard, given its ability to simultaneously identify differing molecular species in complex substrates, such as human tissues or cells. In recent years, there has been growing scientific publication support for the merit of RS in histologic diagnosis of thyroid tumors, defining the main spectra of various lesions (16, 17).

Table 3. A comparison of the 3 RAMAN samples

Comparison	Kappa (95% CI)			Alpha (95% CI)		
AHCAcorr vs AHCAeucl	0.50	(0.36	0.64)	0.72	(0.63	0.82)
AHCAcorr vs KM	0.97	(0.92	1.00)	0.98	(0.96	1.01)
AHCAeucl vs KM	0.48	(0.34	0.62)	0.71	(0.61	0.80)

Nevertheless, data are still lacking for RS with regard to cytologic samples and preoperative diagnoses. The first study of RS analysis to differentiate benign from malignant thyroid cytology was published in 2018 and was based cell lines (29). One recent study used cells obtained from surgical samples to evaluate RS performance in medullary thyroid carcinomas (30). The present investigation, to the best of our knowledge, is the first to involve human thyroid FNA samples.

As with other studies of molecular test performance in thyroid cytology, we compared our RS analytic results with final histologic diagnoses (the gold standard) to gauge appropriateness of operative intervention. There was good performance of RS overall in the aggregate of indeterminate cytologic samples (PPV, 81.5%; NPV, 76.2%; sensitivity, 68.8%; specificity, 86.5%). Hence, RS may be used to rule out or rule in malignancy, regardless of cancer prevalence within specific categories. However, subsets of indeterminate cytology did show distinct diagnostic benefit. In low-risk indeterminate lesions (TIR3A/Bethesda category III) specifically, RS clearly enhanced the capacity to rule out malignancy (NPV, 92% that become 93.3% when RS is combined with TI-RADS score), thereby reducing the need for unnecessary surgery. RS residual malignancy risk (7-8%) was significantly lower than the rate of cytological TIR3A diagnosis; moreover, it was slightly higher than the risk (3-7%) attached to benign cytologic status reported in the literature (25, 31, 32). In high-risk indeterminate nodules (TIR3B/Bethesda category IV), the ability to rule in malignancy was positively impacted by RS. For this category, the PPV of RS (93.3%) and the rate of proven cancer after malignant cytologic diagnosis overlapped. Such patients may thus be properly managed, based on expected histologic cancerous outcomes.

Of note, the steps in preparing cytologic material for spectroscopic evaluation are brief and unique. Cells are analyzed in a fresh state, avoiding the pitfalls of various cytologic processing methods. Cytologic material is actually a useful and valuable source for various molecular diagnostic tests that help optimize patient care, although the samples obtained and their processing are highly variable (33). Our RS protocol standardized preanalytical workflow, improving the reproducibility of results.

Table 4. Performance of RAMAN in cytologically indeterminate thyroid nodules

In those with TIR 3A						(95% CI)	
	Histology						
RAMAN	Negative	Positive	Total	Sensitivity	80.0%	(67.1%	92.9%)
Negative	23	2	25	Specificity	85.2%	(73.7%	96.6%)
Positive	4	8	12	Positive predictive value	66.7%	(51.5%	81.9%)
				Negative predictive value	92.0%	(83.3%	100.0%)
Total	27	10	37	Prevalence	27.0%	(12.7%	41.3%)
In those with TIR 3B							
	Histology						
RAMAN	Negative	Positive	Total	Sensitivity	63.6%	(47.0%	80.3%)
Negative	9	8	17	Specificity	90.0%	(79.6%	100.0%)
Positive	1	14	15	Positive predictive value	93.3%	(84.7%	100.0%)
				Negative predictive value	52.9%	(35.7%	70.2%)
Total	10	22	32	Prevalence	68.8%	(52.7%	84.8%)
In those with TIR 3A or 3B							
	Histology						
RAMAN	Negative	Positive	Total	Sensitivity	68.8%	(57.8%	79.7%)
Negative	32	10	42	Specificity	86.5%	(78.4%	94.6%)
Positive	5	22	27	Positive predictive value	81.5%	(72.3%	90.7%)
				Negative predictive value	76.2%	(66.1%	86.2%)
Total	37	32	69	Prevalence	46.4%	(34.6%	58.1%)
Overall sample							
	Histology						
RAMAN	Negative	Positive	Total	Sensitivity	74.1%	(66.4%	81.9%)
Negative	33	22	55	Specificity	86.8%	(80.9%	92.8%)
Positive	5	63	68	Positive predictive value	92.7%	(88.0%	97.3%)
				Negative predictive value	60.0%	(51.3%	68.7%)
Total	38	85	123	Prevalence	69.1%	(60.9%	77.3%)

Table 5. Performance of RAMAN in cytologically indeterminate thyroid nodules having a US-TIRADS score of 4 or 5

In those with TIR 3A							
	Histology						(95% CI)
RAMAN	Negative	Positive	Total	Sensitivity	85.7%	(42.1%	99.6%)
Negative	14	1	15	Specificity	87.5%	(61.7%	98.4%)
Positive	2	6	8	Positive predictive value	75.0%	(34.9%	96.8%)
				Negative predictive value	93.3%	(68.1%	99.8%)
Total	16	7	23	Prevalence	30.0%	(13.0%	52.9%)
In those with TIR 3B							
	Histology						
RAMAN	Negative	Positive	Total	Sensitivity	60.0%	(36.1%	80.9%)
Negative	5	8	13	Specificity	100.0%	(47.8%	100.0%)
Positive	0	12	12	Positive predictive value	100.0%	(73.5%	100.0%)
				Negative predictive value	38.5%	(13.9%	68.4%)
Total	5	20	25	Prevalence	80.0%	(59.0%	93.2%)
In those with TIR 3A or 3B							
	Histology						
RAMAN	Negative	Positive	Total	Sensitivity	66.7%	(46.0%	83.5%)
Negative	18	9	27	Specificity	90.5%	(69.6%	98.8%)
Positive	2	19	21	Positive predictive value	90.0%	(68.3%	98.8%)
				Negative predictive value	67.9%	(47.6%	84.1%)
Total	20	28	48	Prevalence	56.0%	(41.0%	70.5%)
Overall sample							
	Histology						
RAMAN	Negative	Positive	Total	Sensitivity	76.4%	(64.9%	85.6%)
Negative	19	17	36	Specificity	90.5%	(69.6%	98.8%)
Positive	2	55	57	Positive predictive value	96.5%	(87.9%	99.6%)
				Negative predictive value	52.8%	(35.5%	69.6%)
Total	21	72	93	Prevalence	77.0%	(68.0%	85.4%)

Determining the clinical validity of a new test is a relevant point in this discussion. Under these circumstances, the BCR is a parameter proposed by Marti et al (34), having investigated performance of the Afirma gene expression classifier (Veracyte, South San Francisco, CA, USA). BCR is the percentage of molecular analyses resulting in benign or negative test results, given a high NPV. It corresponds with patient percentages where profiles are akin to benign cytology, thus avoiding diagnostic surgery. In our study, RS showed a BCR of 67.5% for low-risk indeterminate lesions and a BCR of 60.8% for indeterminate lesions overall, underscoring the relevance of RS in patient management. A PCR of 46.8% in high-risk indeterminate lesions demonstrates the utility of RS in identifying patients with neoplastic lesions. Interestingly, a recent meta-analysis (35) has shown the rate of malignancy for high-risk indeterminate lesions to be 47% in actual practice.

Table 6. False positive rates (unnecessary surgery) by FNA result

Subgroup	In all patients	In those with positive RAMAN	*P* value for the difference
TIR 3A	72.9% (27/37)	33.3% (4/12)	0.013
TIR 3B	31.3% (10/32)	6.7% (1/15)	0.064
TIR 3A or 3B	53.6% (37/69)	18.5% (5/27)	0.002
Overall sample	30.9% (38/123)	7.4% (5/68)	<0.001

From the standpoint of adequacy, we failed to obtain interpretable RS profiles in 7 of 130 cases (0.6%). This is lower than the nondiagnostic results of conventional cytology, where 10% sample inadequacy is deemed acceptable (25). The prevailing cause of inadequate RS interpretation was sampling autofluorescence. It is known that lipofuscin, the chief source of fluorescence in thyroid cells, accumulates in benign thyrocytes (36) and may imply nonmalignant status. Our histologic diagnoses confirm this hypothesis, although it cannot be construed as a diagnostic result.

Of course, for regular use in clinical practice, a software-based system of RS profiling must be developed to recognize relevant spectral sets and to output results as medical terms (ie, benign vs malignant). Finally, RS information may be integrated with other data capable of guiding clinical decisions in specific patients. When we linked RS data to US-based TI-RADS high-risk scores, overall performance was still improved in indeterminate nodules with low (TIR3A) or high (TIR3B) risk of malignancy. This means that a computational approach may generate a diagnostic/prognostic algorithm to be applied onsite in outpatients with thyroid nodules.

The issue of false-negative results (17.8% overall) remains an open question. Its likely cause is low numbers of thyrocytes in samples used for RS. We are now embarking on 2 separate investigational paths: (1) examining such samples through in-

depth sequencing and mRNA expression to detect the presence of thyrocytes, if any, and (2) characterizing typical RS profiles of thyroid epithelia and possibly of other cells present within thyroid glands (C-cells, parathyroid cells) to ascertain if inadequate specimens are lacking in these spectra. False-positive RS results (4.06% overall) are attributable to adenomas with altered levels of protein expression (positive immunostaining of galectin-3 or HBME-1 or loss of CD56 expression) (18). We have just published an analysis of this condition (18). Similar to what is known on a molecular level for adenomas with RAS or TERT gene alterations (6), one may speculate that these are lesions with atypical features and possibly evolving malignant potential. A larger body of data will likely resolve this dilemma.

One limitation of our study could be represented by the higher prevalence of malignancy in the cytological categories TIR3B and TIR4 compared to the literature that could affect the PPV of our test.

In conclusion, our data may support the clinical use of RS to improve the diagnostic accuracy of thyroid cytology, thereby effecting a significant reduction in unnecessary surgery. However, larger and prospective studies are needed to confirm the ability of RS to improve the diagnostic accuracy of thyroid cytology, in particular of indeterminate cytologic diagnoses. This perspective opens up new study spaces, given that a Raman apparatus tailored for use in clinical environments must be equipped with user-friendly software for data acquisition, analysis, and subsequent diagnostic output. The latter may be achieved by means of modern artificial intelligence techniques and through close cooperation among medical doctors, engineers, and physicists. The goal here is transforming RS into a bona fide clinical device.

Acknowledgments

The authors thank Antonio Ivan Lazzarino from epistata.org for the statistical analysis.

Funding

This study was supported by Ministero della Salute, through the TIRAMA project (RF-2018-12366568). The funding agency had no role in study design; data collection, analysis, or interpretation; draft review; manuscript approval; or publication submission. The chief investigator (A.C.) had full access to all data, once the database was locked, and had final responsibility for the decision to submit for publication.

Author Contributions

A.P. and A.C. had the original idea, wrote the study protocol, coordinated the study procedures, and critically revised the report. A.P., A.M.N., D.M., P.C., F.L., E.P., C.T., and A.C. selected, monitored, and cared for patients and collected data. M.V. prepared cellular samples for both microscopic analysis and Raman spectroscopy. E.P., C.T., and A.C. performed cytological and histological diagnosis. A.S., M.D.G., Al.P., A.dM., and M.A.R. performed Raman spectroscopy analysis and interpreted data. A.S., M.V., M.D.G., Al.P., and M.A.R. performed statistical analysis. A.P., A.S., A.M.N., M.D.G., M.V., M.A.R., and A.C. interpreted data and wrote the first draft of the report. All authors revised the first draft and wrote the final version of the manuscript.

Conflict of Interest

All authors declare no competing interests.

Data Availability

We will consider sharing deidentified, individual participant-level data that underlie the results reported in this article on receipt of a request detailing the study hypothesis and statistical analysis plan. All requests should be sent to the corresponding author. The corresponding author and lead investigators of this study will discuss all requests and make decisions about whether data sharing is appropriate based on the scientific rigor of the proposal. All applicants will be asked to sign a data access agreement.

References

1. Valderrabano P, McIver B. Evaluation and management of indeterminate thyroid nodules: the revolution of risk stratification beyond cytological diagnosis. *Cancer Control*. 2017;24(5):1073274817729231.
2. Baloch ZW, LiVolsi VA, Asa SL, *et al*. Diagnostic terminology and morphologic criteria for cytologic diagnosis of thyroid lesions: a synopsis of the National Cancer Institute Thyroid Fine-Needle Aspiration State of the Science Conference. *Diagn Cytopathol*. 2008;36(6):425-437.
3. Livhits MJ, Zhu CY, Kuo EJ, *et al*. Effectiveness of molecular testing techniques for diagnosis of indeterminate thyroid nodules: a randomized clinical trial. *JAMA Oncol*. 2021;7(1):70-77.
4. Agrawal N, Akbani R, Aksoy BA, *et al*. Integrated genomic characterization of papillary thyroid carcinoma. *Cell*. 2014;159(3):676-690.
5. Nikiforov YE, Steward DL, Carty SE, *et al*. Performance of a multigene genomic classifier in thyroid nodules with indeterminate cytology: a prospective blinded multicenter study. *JAMA Oncol*. 2019;5(2):204-212.
6. Acquaviva G, Visani M, Repaci A, *et al*. Molecular pathology of thyroid tumours of follicular cells: a review of genetic alterations and their clinicopathological relevance. *Histopathology*. 2018;72(1):6-31.
7. Kumar S, Srinivasan A, Nikolajeff F. Role of infrared spectroscopy and imaging in cancer diagnosis. *Curr Med Chem*. 2018;25(9):1055-1072.
8. D'Acunto M, Cioni P, Gabellieri E, Presciuttini G. Exploiting gold nanoparticles for diagnosis and cancer treatments. *Nanotechnology*. 2021;32(19):192001.
9. Santos IP, Barroso EM, Bakker Schut TC, *et al*. Raman spectroscopy for cancer detection and cancer surgery guidance: translation to the clinics. *Analyst*. 2017;142(17):3025-3047.
10. Abramczyk H, Brozek-Pluska B. New look inside human breast ducts with Raman imaging. Raman candidates as diagnostic markers for breast cancer prognosis: mammaglobin, palmitic acid and sphingomyelin. *Anal Chim Acta*. 2016;909:91-100.
11. Ibrahim O, Toner M, Flint S, Byrne HJ, Lyng FM. The potential of Raman spectroscopy in the diagnosis of dysplastic and malignant oral lesions. *Cancers (Basel)*. 2021;13(4):1-14.
12. Depciuch J, Barnaś E, Skręt-Magierło J, *et al*. Spectroscopic evaluation of carcinogenesis in endometrial cancer. *Sci Rep*. 2021;11(1):9079.
13. Berne BJ, Pecora R. *Dynamic Light Scattering : with Applications to Chemistry, Biology, and Physics*. Wiley; 1976.
14. Teixeira CSB, Bitar RA, Martinho HS, *et al*. Thyroid tissue analysis through Raman spectroscopy. *Analyst*. 2009;134(11):2361-2370.
15. Rau J V, Graziani V, Fosca M, *et al*. RAMAN spectroscopy imaging improves the diagnosis of papillary thyroid carcinoma. *Sci Rep*. 2016;6:35117.

16. Rau J V, Fosca M, Graziani V, *et al*. Proof-of-concept Raman spectroscopy study aimed to differentiate thyroid follicular patterned lesions. *Sci Rep*. 2017;7(1):14970.
17. Sbroscia M, Di Gioacchino M, Ascenzi P, *et al*. Thyroid cancer diagnosis by Raman spectroscopy. *Sci Rep*. 2020;10(1):13342.
18. Sodo A, Verri M, Palermo A, *et al*. Raman spectroscopy discloses altered molecular profile in thyroid adenomas. *Diagnostics (Basel)*. 2021;11(1):43.
19. Tessler FN, Middleton WD, Grant EG, *et al*. TI-RADS. *J Am Coll Radiol*. 2017;14(5):587-595.
20. Haugen BR, Alexander EK, Bible KC, *et al*. American Thyroid Association management guidelines for adult patients with thyroid nodules and differentiated thyroid cancer: the American Thyroid Association guidelines task force on thyroid nodules and differentiated thyroid cancer. *Thyroid*. 2016; 26(1):1-133.
21. Gharib H, Papini E, Garber JR, *et al*. Diagnosis and management of thyroid nodules: AACE/ACE/AME 2016. *Endocr Pract*. 2016;22(Suppl. 1):1-60.
22. Kholová I, Haaga E, Ludvik J, Kalfert D, Ludvikova M. *Noninvasive Follicular Thyroid Neoplasm with Papillary-Like Nuclear Features (NIFTP): Tumour Entity with a Short History. A Review on Challenges in our Microscopes, Molecular and Ultrasonographic Profile. Vol. 12, Diagnostics*. MDPI; 2022.
23. Nardi F, Basolo F, Crescenzi A, *et al*. Italian Consensus for the Classification and Reporting of Thyroid Cytology. *J Endocrinol Invest*. 2014;37(6):593-599.
24. Poller DN, Cochand-Priollet B, Trimboli P. Thyroid FNA terminology: the case for a single unified international system for thyroid FNA reporting. *Cytopathology*. 2021;32(6):714-717.
25. Cibas ES, Ali SZ. The 2017 Bethesda System for Reporting Thyroid Cytopathology. *Thyroid*. 2017;27(11):1341-1346.
26. Lloyd RV, Osamura RY, Klöppel G, Rosai J. World Health Organization classification of tumours of endocrine organs. In: Lloyd RV, Osamura RY, Klöppel G, Rosai J, eds. *International Agency for Research on Cancer*. 4th Edition; 2017.
27. Bongiovanni M, Spitale A, Faquin WC, Mazzucchelli L, Baloch ZW. The Bethesda System for Reporting Thyroid Cytopathology: a meta-analysis. *Acta Cytol*. 2012;56(4):333-339.
28. Palermo A. Supplementary table with false positive and negative results. Uploaded July 27, 2022. *Figshare*. https://doi.org/10.6084/m9.figshare.20387538
29. O'Dea D, Bongiovanni M, Sykiotis GP, *et al*. Raman spectroscopy for the preoperative diagnosis of thyroid cancer and its subtypes: an in vitro proof-of-concept study. *Cytopathology*. 2019;30(1):51-60.
30. Soares de Oliveira MA, Campbell M, Afify AM, Huang EC, Chan JW. Raman-based cytopathology: an approach to improve diagnostic accuracy in medullary thyroid carcinoma. *Biomed Opt Express*. 2020;11(12):6962.
31. Torregrossa L, Poma AM, Macerola E, *et al*. The Italian Consensus for the Classification and Reporting of Thyroid Cytology: cytohistologic and molecular correlations on 37,371 nodules from a single institution. *Cancer Cytopathol*. Published online July 5, 2022. https://doi.org/10.1002/cncy.22618
32. Sutton W, Canner JK, Rooper LM, Prescott JD, Zeiger MA, Mathur A. Is patient age associated with risk of malignancy in a ≥4 cm cytologically benign thyroid nodule? *Am J Surg*. 2021;221(1):111-116.
33. Aisner DL, Sams SB. The role of cytology specimens in molecular testing of solid tumors: techniques, limitations, and opportunities. *Diagn Cytopathol*. 2012;40(6):511-524.
34. Marti JL, Avadhani V, Donatelli LA, *et al*. Wide inter-institutional variation in performance of a molecular classifier for indeterminate thyroid nodules. *Ann Surg Oncol*. 2015;22(12):3996-4001.
35. Trimboli P, Crescenzi A, Giovanella L. Performance of Italian Consensus for the Classification and Reporting of Thyroid Cytology (ICCRTC) in discriminating indeterminate lesions at low and high risk of malignancy. A systematic review and meta-analysis. *Endocrine*. 2018;60(1):31-35.
36. Landas SK, Schelper RL, Tio FO, Turner JW, Moore KC, Bennett-Gray J. Black thyroid syndrome: exaggeration of a normal process? *Am J Clin Pathol*. 1986;85(4):411-418.