Original Article

Boosting thyroid nodule diagnosis through ultrasound and molecular imaging integration with unsupervised learning

Federica Facchinetti[1]^, Vincenzo L'Imperio[2]^, Isabella Piga[3]^, Daniele M. Papetti[1,4]^, Giulia Capitoli[1,5]^

[1]Bicocca Bioinformatics Biostatistics and Bioimaging B4 Center, Department of Medicine and Surgery, University of Milano-Bicocca, Milano, Italy; [2]Department of Medicine and Surgery, Pathology Unit, IRCCS Fondazione San Gerardo dei Tintori, University of Milano-Bicocca, Monza, Italy; [3]Proteomics and Metabolomics Unit, Department of Medicine and Surgery, University of Milano-Bicocca, Monza, Italy; [4]Department of Informatics, Systems and Communication, University of Milano-Bicocca, Milano, Italy; [5]Biostatistics and Clinical Epidemiology, Fondazione IRCCS San Gerardo dei Tintori, Monza, Italy

Contributions: (I) Conception and design: F Facchinetti, DM Papetti, G Capitoli; (II) Administrative support: V L'Imperio, DM Papetti, G Capitoli; (III) Provision of study materials or patients: V L'Imperio, I Piga; (IV) Collection and assembly of data: V L'Imperio, I Piga; (V) Data analysis and interpretation: F Facchinetti, DM Papetti, G Capitoli; (VI) Manuscript writing: All authors; (VII) Final approval of manuscript: All authors.

Correspondence to: Daniele M. Papetti, PhD. Department of Informatics, Systems and Communication, University of Milano-Bicocca, Viale Sarca 336, Milano 20126, Italy; Bicocca Bioinformatics Biostatistics and Bioimaging B4 Center, Department of Medicine and Surgery, University of Milano-Bicocca, Milano, Italy. Email: daniele.papetti@unimib.it; Giulia Capitoli, PhD. Bicocca Bioinformatics Biostatistics and Bioimaging B4 Center, Department of Medicine and Surgery, University of Milano-Bicocca, Via Follereau 3, Vedano al Lambro (MB), Milano 20814, Italy; Biostatistics and Clinical Epidemiology, Fondazione IRCCS San Gerardo dei Tintori, Monza, Italy. Email: giulia.capitoli@unimib.it.

Background: In recent years, matrix-assisted laser desorption-ionization (MALDI) mass spectrometry imaging (MSI) has been applied to cytological thyroid specimens as a complementary tool for the diagnosis of thyroid nodules. Specifically, MALDI-MSI has been proven to be effective for classifying "indeterminate for malignancy" reports. In this work, we analyse the effectiveness of unsupervised methods to further explore whether unsupervised learning (UL) can reveal latent structures within MALDI-MSI, clinical, and ultrasound (US) data that may help better understand the heterogeneity of thyroid nodules, integrating different clinical and molecular information, such as demographic and echographic data.

Methods: This retrospective study involved 222 patients who underwent US-guided fine-needle aspiration (FNA), with MALDI-MSI analysis. This first dataset (named "Dataset MSI") comprises demographic, molecular, and clinical information. A second dataset of 82 patients ("Dataset MSI + ECHO") was extracted from the first one, containing additional information regarding the US characterization of the thyroid nodules. Unsupervised clustering was performed for each dataset through three distinct methods: k-means, partitioning around medoids (PAM), and hierarchical clustering (HC), and compared to MALDI-MSI classifications.

Results: Our results highlight the potential value of unsupervised approaches in exploring the underlying structure of the data. In Dataset MSI, the clustering analysis revealed patterns partially consistent with clinical outcomes and helped group cases that were inconclusive in MALDI-MSI reports, with 91% (20 out of 22) of patients without a clear MALDI-MSI result falling into clinically coherent clusters. Similarly, in Dataset MSI + ECHO, we observed a drastic increase in the overall sensitivity from 0.4 to 0.95 compared to the MALDI-MSI prediction. We also correctly clustered 94% of TIR4/Thy4 (5 out of 5) and TIR5/Thy5 (11 out of 12) patients that MALDI-MSI misclassified.

Conclusions: This research highlights the need to incorporate clinical, US, and molecular information

^ ORCID: Federica Facchinetti, 0009-0006-0970-3409; Vincenzo L'Imperio, 0000-0002-9284-2998; Isabella Piga, 0000-0001-8221-2493; Daniele M. Papetti, 0000-0002-3574-6027; Giulia Capitoli, 0000-0002-8178-2440.

into a single learning method in routine diagnosis, especially when US variables are available. This approach may represent a useful exploratory framework to investigate the biological and clinical heterogeneity of thyroid nodules and to guide future studies, potentially leading to more accurate and timely diagnoses. These findings pave the way for further research and applications in diagnosing thyroid nodules.

Keywords: Unsupervised learning (UL); thyroid nodules; non-invasive follicular thyroid neoplasm with papillary-like nuclear features; multimodal machine learning (multimodal ML)

Submitted May 13, 2025. Accepted for publication Apr 21, 2026. Published online Apr 30, 2026.
doi: 10.21037/qims-2025-1125
View this article at: https://dx.doi.org/10.21037/qims-2025-1125

Introduction

Fine-needle aspiration (FNA) is currently the gold standard, non-invasive cytology approach to evaluate the malignant nature of thyroid nodules (1). However, 20–30% of cases have an "indeterminate for malignancy" report, whose consequence is a diagnostic thyroidectomy, potentially leading to lifelong thyroid hormone replacement therapy (2). Of these, 70–80% have benign histology, stressing the need for markers able to prevent the "thyroid carnage" (2,3).

In recent years, matrix-assisted laser desorption-ionization (MALDI) mass spectrometry imaging (MSI) has been successfully applied on cytological thyroid specimens as a complementary tool for the characterization of indeterminate thyroid nodules (4). MALDI-MSI has been shown to be effective in separating areas with benign/malignant cells in cellularly adequate cytological samples, still showing "gray zone" areas hampering its more widespread employment in the routine screening of indeterminate cases (5). To overcome this intrinsic limitation, a possible approach is represented by machine learning (ML) methods, especially for unsupervised learning (UL) models, allowing the self-clusterization of data without prior human intervention based on the characteristics of the cohort (6).

Moreover, previous work investigated supervised diagnostic models based on ultrasound (US) classification systems such as American College of Radiology-Thyroid Imaging Reporting and Data System (ACR-TIRADS) and European Thyroid Imaging Reporting and Data System (EU-TIRADS), showing moderate diagnostic performances and highlighting the complexity and heterogeneity of thyroid nodules (7,8). For this reason, in the present study, we adopted a different perspective, focusing on an unsupervised exploratory analysis aimed at identifying intrinsic patterns within the data without imposing predefined diagnostic labels.

In this work, we propose and evaluate different unsupervised methods to further improve MALDI-MSI analysis in distinguishing malignant and non-malignant thyroid nodules, supporting the classification process in diagnosing indeterminate cases, using additional information collected during the clinical workflow, such as echographic data. In particular, we applied three clustering methods on patients within malignant/non-malignant groups. We present this article in accordance with the TRIPOD+AI reporting checklist (available at https://qims.amegroups.com/article/view/10.21037/qims-2025-1125/rc).

Methods

Patients and datasets

This retrospective study included 222 patients who underwent US-guided FNA at the interventional radiology clinic, Fondazione IRCCS San Gerardo dei Tintori, University of Milano-Bicocca, Monza, Italy, during an Italian Association for Research on Cancer (AIRC)-granted project for the diagnosis of thyroid carcinoma. Clinical (presence of thyroiditis, multinodularity, and nodule dimensions) and demographic (age and gender) information were retrieved. Results of the FNA were collected and reported following the Italian Society of Pathology (SIAPEC) classification (9) and the Bethesda System for reporting thyroid cytopathology (10). In detail, cases were considered benign after a TIR2/Thy2 FNA result and by performing a US examination of patients 12 months after the first US-guided FNA confirming:

- Absence of new echographic malignant features;
- Absence of a significant increasing of nodule size;
- Absence of nodal metastasis;
- No incidence of new suspicious nodules.

For malignant cases, histological diagnoses were progressively collected after thyroidectomy to certify the nature of the nodules. All cytology cases underwent MALDI-MSI analysis, as previously described (Dataset MSI) (5).

From the original cohort, a second dataset, which combines MSI and echographic data (Dataset MSI + ECHO), was extracted (n=82) with available US characterization of the thyroid nodules following both EU-TIRADS and ACR-TIRADS (11). Specifically, for each patient, specific US features (e.g., nodule composition, echogenicity, shape, margins, and the presence of calcifications) were available (11). A binary approach (lower *vs.* higher malignancy probability) was adopted as output of the ML model on US features. In detail, characteristics related to lower *vs.* higher malignancy probability were identified for margins (smooth and ill-defined *vs.* irregular), shape (wider-than-tall *vs.* taller-than-wide), echogenicity (hyper/iso *vs.* hypoechoic), and nodular foci (e.g., punctate foci suggestive of malignancy), as previously described (11).

This study was conducted in accordance with the Declaration of Helsinki and its subsequent amendments. The study was approved by the regional ethics committee (Comitato Etico Brianza, via Pergolesi, 33, 20900 Monza, Italy) (No. FINAL-TIR PU 3581/21, approval date: January 14, 2021), and informed consent was taken from all the patients.

MALDI-MSI classification

The MALDI-MSI analysis was based on a pixel-by-pixel approach, thresholding selected percentages of malignant pixels (5). A multinomial penalized regression with the Lasso regularization model to pick the proteomic signals able to distinguish malignant cells [papillary thyroid carcinoma (PTC)] from normal thyrocytes [hyperplastic (HP)] or inflammatory background [Hashimoto thyroiditis (HT)] was used, as previously described (5,12). The resulting classification model was used to predict the probability to belong to one of the three classes. For each patient, all the single mass spectra of the imzML MALDI-MSI analysis (pixel-by-pixel) extracted from the FNA were classified. For the classification, a three-tiered system was applied using different thresholds for benign (<7.0% of malignant pixels) and malignant (>16.7%) cases, and those falling in between (7.0–16.7%) were considered as intermediate/"gray zone" per MALDI-MSI (5).

Statistical analysis

Continuous variables were described with mean ± standard deviation (SD), and categorical variables were reported as count and frequency (%).

Unsupervised clustering was performed for each dataset by means of three distinct methods: k-means (13,14), partitioning around medoids (PAM) (6), and hierarchical clustering (HC) (15). All three methods leveraged the Euclidean distance as a distance metric to compute the difference between the data representing the patients. In order to avoid the application of any pre-analysis labelling step, the number of clusters k—with k ranging from 2 to 10—was selected by maximizing the silhouette score, which indicates how well patients are grouped into homogeneous and separated clusters (16). All three methods were tested using the same k, selecting the clustering configuration that showed the most coherent separation between patient groups and clinical outcomes, as benign or malignant, based on sensitivity (Sn) and specificity (Sp). In addition, the performance metrics, positive and negative predictive values (PPV and NPV, respectively), were also reported. For every method used, each cluster is identified with a prevalence of either malignant or benign nodules, and consequently, it is associated with the respective benign or malignant label. These labels are subsequently compared with the pathologist's labels, allowing an assessment of the UL output. Variables from the two datasets were used as input for the three UL models (*Figure 1*). The results obtained from the MALDI-MSI classification were compared with the outcomes of the UL model proposed.

The first considered method is the k-means, which seeks an optimal data partition into k clusters such that the squared distances between each data point and its assigned cluster centroid (mean) are minimized, while the distance from the other cluster centroids is maximized. The second considered method is PAM. In such a method, each cluster is represented by one of the data points belonging to the cluster, namely the medoid of the data points belonging to the considered cluster. The third method is the agglomerative HC, which seeks to build a hierarchy of clusters in a bottom-up approach. Each patient starts in its personal cluster, represented as a leaf in the clustering dendrogram. The clusters are iteratively built by merging together the cluster of patients more similar to each other. The choice of which clusters to merge together is performed by computing the distance between the clusters

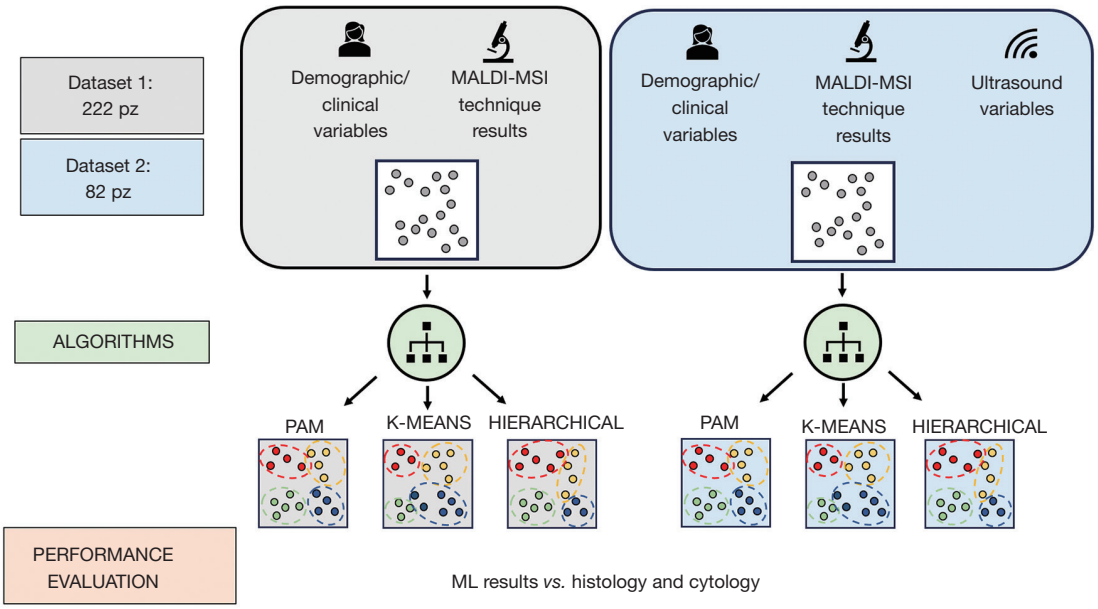

Figure 1 General outline of the study workflow: the two datasets with different input variables (left and right top squares) are used as input in three different UL methods. Dataset 1: Dataset MSI, underwent MALDI-MSI analysis. Dataset 2: Dataset MSI + ECHO, combines MSI and echographic data. MALDI, matrix-assisted laser desorption-ionization; ML, machine learning; MSI, mass spectrometry imaging; PAM, partitioning around medoids; pz, patients; UL, unsupervised learning.

and selecting the two closest clusters.

Importantly, diagnostic labels (benign/malignant) were not used during the clustering process. These labels were only used a posteriori to interpret the clusters and to describe the clinical composition of each group.

All the statistical analyses were performed using the open-source R software v.4.4.2 (R Foundation for Statistical Computing, Vienna, Austria).

Results

Cohort and datasets details

The entire cohort (Dataset MSI) had an average age of 59.10±15.2 years; 75.2% of the patients were females and 24.8% males; and in the final thyroid disease, 149 (67.1%) and 73 (32.9%) were benign and malignant, respectively (*Table 1*). The Dataset MSI + ECHO was characterized by an average age of 58.92±15.0 years; 75.6% of the patients were females and 24.4% males; and in the final thyroid disease, 60 (73.2%) and 22 (26.8%) cases were benign and malignant, respectively. The different frequency of US features predicting the malignant likelihood in this last cohort is reported in *Figure 2*.

Cluster interpretation results

The optimal silhouette score was set at k=2. PAM clustering showed the best results on the "Dataset MSI" (*Figure 3*, Figure S1), outperforming the MALDI-alone classification overall (Sn =0.57 *vs.* 0.53 and Sp =0.95 *vs.* 0.93, respectively) (*Table 2*). Focusing on the cytological diagnostic categories (*Table 3*), the UL-based model was superior to MALDI-alone classification in the benign TIR2/Thy2 category (Sp =0.92 *vs.* 0.85), correctly assigning to this class 11% of cases classified as "gray zone" by MALDI-alone. On the same page, improvement was noted within the malignant TIR5/Thy5 group (Sn =0.64 *vs.* 0.53), rescuing 7 (13%) cases otherwise classified as "gray zone" by MALDI-alone. A slight increase in Sn was achieved within the TIR3/Thy3 category (0.54 *vs.* 0.45), and no significant improvement in the suspicious for malignancy category was recorded.

The UL method with the best performance on the Dataset MSI + ECHO was k-means (*Figure 3*, Figure S1), providing a significantly better Sn (0.95 *vs.* 0.40) despite a slight drop in Sp (0.88 *vs.* 0.96) as compared to the MALDI-alone approach (*Table 2*). As per the previous dataset, k-means provided an improvement in the detection of benign cases, allowing the correct TIR2/Thy2 assignment

Table 1 Clinical information of the nodules and the patients for both datasets included in the study

Clinical characteristics	Dataset MSI (n=222)	Dataset MSI + ECHO (n=82)
Age (years)	59.10±15.2	58.92±15.0
Sex		
Female	167 (75.2)	62 (75.6)
Male	55 (24.8)	20 (24.4)
Thyroiditis	n=221	
Yes	27 (12.2)	10 (12.2)
No	194 (87.8)	72 (87.8)
Nodularity		
Single	115 (51.8)	51 (62.2)
Pseudonodules	8 (3.6)	2 (2.4)
Multiple	99 (44.6)	29 (35.4)
Nodule dimensions (mm)	21±11.5	19.5±9
Cytology	n=223	
TIR2/Thy2	82 (36.8)	28 (34.1)
TIR3/Thy3	62 (27.8)	37 (45.1)
TIR4/Thy4	16 (7.2)	5 (6.1)
TIR5/Thy5	53 (23.8)	12 (14.6)
Outcome		
Benign	149 (67.1)	60 (73.2)
Malignant	73 (32.9)	22 (26.8)

Continuous variables are presented as mean ± SD, and qualitative variables are presented as n (%). All the patients that presented one of the following outcomes: "no follow-up", "Hürthle cell adenoma", "follicular adenoma", "non-invasive follicular thyroid neoplasm with papillary-like nuclear features", and "multinodular goiter" that have been categorized as having a benign outcome, or presented one of the following: "minimally invasive Hürthle cell carcinoma", "metastasis of spindle cell melanoma", "anaplastic carcinoma", and "papillary thyroid carcinoma" and thus have been considered as having a malignant outcome. Dataset MSI: underwent MALDI-MSI analysis. Dataset MSI + ECHO: combines MSI and echographic data. MALDI, matrix-assisted laser desorption-ionization; MSI, mass spectrometry imaging; SD, standard deviation.

to one of the two "gray zone" cases by MALDI-alone (Sp =0.93 *vs.* 0.89). Similarly, k-means adequately clustered 11 out of 12 malignant cases in the TIR5/Thy5 category, as compared to the MALDI-alone approach (5 out of 12, improving Sp from 0.41 to 0.92). Improved performances were obtained in the TIR4/Thy4 as well, with all cases correctly assigned by the UL-approach as compared to the MALDI-alone one (2 out of 5). The most impressive result was obtained within the TIR3/Thy3 category, with a significant improvement of Sn (1 *vs.* 0.25) despite a minor drop in Sp (0.94 *vs.* 1). Indeed, two out of the three patients of the TIR3/Thy3 "gray zone" were included in the wrong cluster. The clustering of TIR4/Thy4 patients resulted in a perfect score, assigning all five patients to the malignant cluster since all the patients had a malignant outcome, while MALDI-alone classified 2 out of 5 nodules as benign. Finally, we reported in Figure S2 four cases of different cytological diagnoses that MALDI-alone misclassified, while the harmonization of MALDI with US (Dataset MSI + ECHO) examination correctly classified.

Discussion

An important aspect of this study is that the objective was

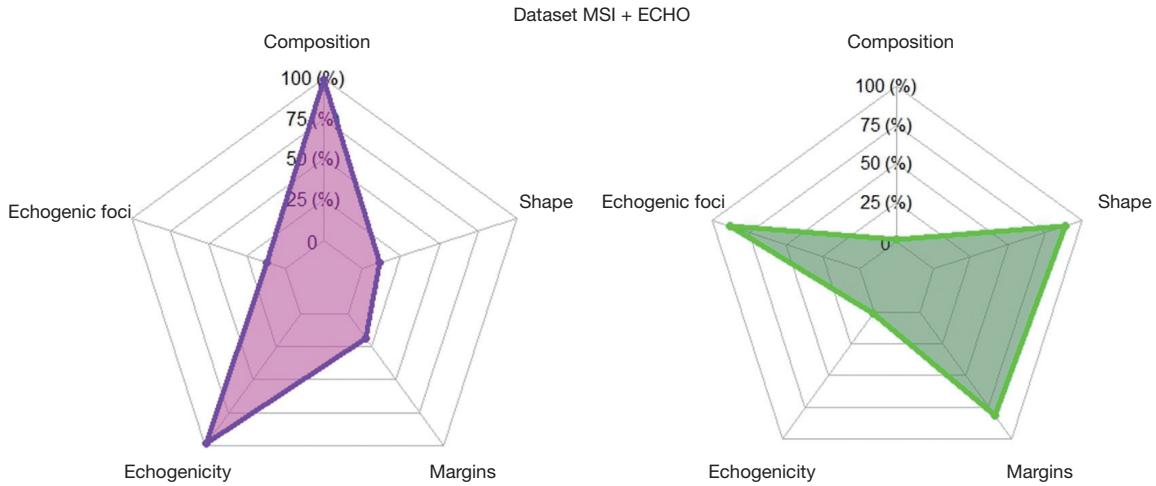

Figure 2 Spider plot depicting the absolute frequency of each US variable outcome typically associated with a higher probability of malignancy (purple, on the left) and those associated with a higher likelihood of benignity (green, on the right) for Dataset MSI + ECHO. Dataset MSI + ECHO: combines MSI and echographic data. MSI, mass spectrometry imaging; US, ultrasound.

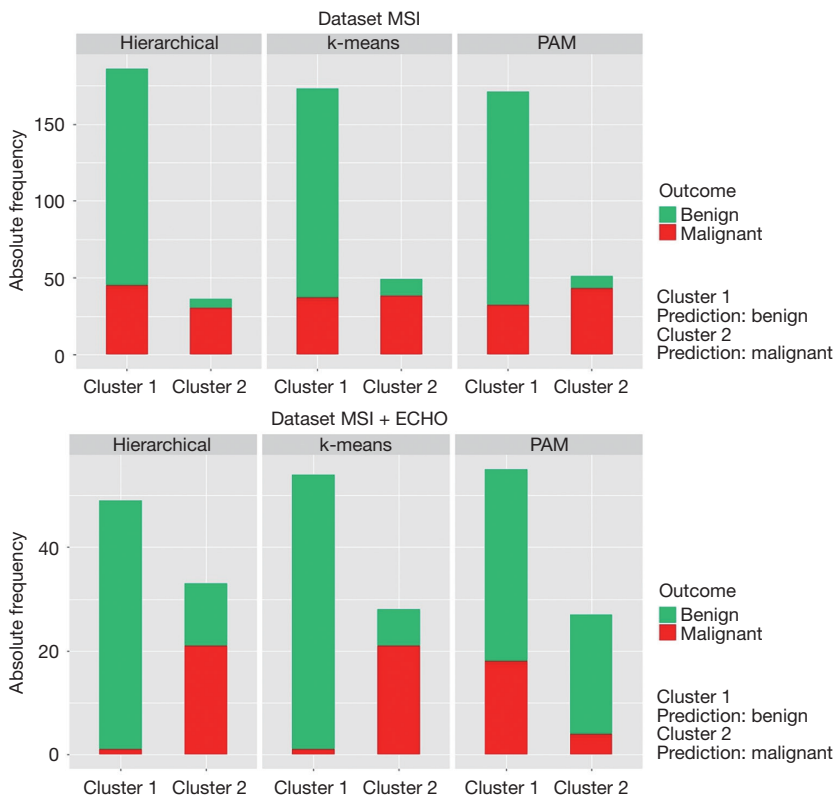

Figure 3 Bar plot of the distribution of malignant and benign nodules across various UL methods (HC, PAM, and k-means) within both Dataset MSI and Dataset MSI + ECHO. The bar plot of the distribution of thyroid classification nodules across various UL methods is reported in Figure S1. Dataset MSI: underwent MALDI-MSI analysis. Dataset MSI + ECHO: combines MSI and echographic data. HC, hierarchical clustering; MALDI, matrix-assisted laser desorption-ionization; MSI, mass spectrometry imaging; PAM, partitioning around medoids; UL, unsupervised learning.

Table 2 Performances of both MALDI-alone and UL-assisted methods in the Dataset MSI and Dataset MSI + ECHO

Dataset	Leveraged features	TP, n	TN, n	Sn	Sp	PPV, %	NPV, %
Dataset MSI	MALDI-alone performances	35	125	0.53	0.93	80	80
	UL-assisted performances	43	139	0.57	0.95	84	81
Dataset MSI + ECHO	MALDI-alone performances	8	54	0.4	0.96	80	82
	UL-assisted performances	21	53	0.95	0.88	70	98

Dataset MSI: underwent MALDI-MSI analysis. Dataset MSI + ECHO: combines MSI and echographic data. MALDI, matrix-assisted laser desorption-ionization; MSI, mass spectrometry imaging; NPV, negative predictive value; PPV, positive predictive value; Sn, sensitivity; Sp, specificity; TN, true negative; TP, true positive; UL, unsupervised learning.

Table 3 Performance of both MALDI-alone and UL-assisted methods in the Dataset MSI and Dataset MSI + ECHO with a focus on the cytological diagnostic categories

Dataset	Cytology	MALDI-alone performances			UL-assisted performances	
		Benign	"Gray zone"	Malignant	Cluster 1 (prediction: benign)	Cluster 2 (prediction: malignant)
Dataset MSI	TIR2/Thy2	70 [85]	8 [10]	4 [5]	78 [95]	4 [5]
	TIR3/Thy3	60 [85]	6 [8]	5 [7]	65 [92]	6 [8]
	TIR4/Thy4	8 [50]	1 [6]	7 [44]	9 [56]	7 [44]
	TIR5/Thy5	18 [34]	7 [13]	28 [53]	19 [36]	34 [64]
Dataset MSI + ECHO	TIR2/Thy2	25 [89]	2 [7]	1 [4]	26 [93]	2 [7]
	TIR3/Thy3	33 [89]	3 [8]	1 [3]	30 [81]	7 [9]
	TIR4/Thy4	2 [40]	0	3 [60]	0	5 [100]
	TIR5/Thy5	6 [55]	1 [9]	5 [46]	1 [9]	11 [91]

Data are presented as n [%]. Dataset MSI: underwent MALDI-MSI analysis. Dataset MSI + ECHO: combines MSI and echographic data. MALDI, matrix-assisted laser desorption-ionization; MSI, mass spectrometry imaging; UL, unsupervised learning.

not to develop a diagnostic prediction model, but rather to explore the intrinsic structure of the data through an unsupervised approach. This perspective was motivated by previous supervised studies conducted on MALDI-MSI and US-based classification systems, where the diagnostic performances were moderate and suggested that current labels may not fully capture the biological variability of thyroid nodules.

Specifically, MALDI-MSI has proven successful in its application to cytological thyroid specimens, serving as a complementary tool for diagnosing and characterizing indeterminate thyroid nodules; however, the results are limited by the need of a cytological sample that is cellularly adequate, and by the fact that MALDI-MSI output still contemplates a "gray zone", representing a limit on the number of FNAs with a clear report (5).

In this context, clustering can be considered a hypothesis-generating strategy that may help identify patterns that deserve further investigation in future larger and prospective studies.

Consequently, in this study, we investigated the role of unsupervised clustering on MALDI-MSI data to provide the pathologist with an additional support tool to distinguish malignant and non-malignant thyroid nodules. Specifically, we tested k-means, PAM, and HC performance on two datasets, comparing the obtained results against the MALDI-MSI outcome. These UL methods (k-means, PAM, and HC) have been chosen as they represent widely adopted and well-established clustering methods (17,18).

In Dataset MSI, which consists of 222 patients characterized by demographic and MALDI-MSI variables, 91% (20 out of 22) of the patients that could not receive a clear MALDI-MSI report were properly grouped in the appropriate cluster. It is important to note that 100% of

TIR2/Thy2 and TIR5/Thy5 were all correctly clustered: specifically, the ability in detecting "gray zone" TIR5/Thy5 nodules as malignant is an important asset in supporting the decision of the pathologist, filling the diagnostic gaps of MALDI-alone. Similarly, in Dataset MSI + ECHO, which was composed of 82 patients with available US variables, we drastically improved Sn from 0.4 to 0.95. Moreover, we correctly clustered 94% (16 out of 17) TIR4/Thy4 and TIR5/Thy5 patients that MALDI-alone failed to properly label. This improvement highlights how the additional information provided by the US variables improved the general capability of the model in detecting nodules, particularly malignant ones.

Conclusions

These research findings demonstrate that unsupervised clustering may be able to enhance the diagnostic power of MALDI-MSI, particularly in cases within the "gray zone". By integrating US features, the model's performance significantly improves, offering a promising supplementary tool to assist pathologists in the classification of thyroid nodules. While the data obtained from this cohort of patients has provided valuable insights, it is crucial to recognize that such results are one of the first steps to show the feasibility of exploiting ML methods to analyse data to support the pathologist in the diagnostic process. Future studies with larger and more diverse patient populations, preferably adopting a multicentric and prospective approach, will be essential to further assess the robustness and applicability of these findings. Moreover, since previous studies showed different shortcomings of the ACR-TIRADS and the EU-TIRADS scores (7,8), the performance of the proposed UL-based methods could benefit from the harmonization of US-features report and classification. We also envision that an integration of multiple ML models by means of an ensemble based on a majority voting approach will help in properly classifying new patients.

In conclusion, the incorporation of ML methods into the diagnostic routine, particularly when US variables are available, emerges as a promising exploratory approach to better understand the multidimensional characteristics of thyroid nodules.

Acknowledgments

None.

Footnote

Reporting Checklist: The authors have completed the TRIPOD+AI reporting checklist. Available at https://qims.amegroups.com/article/view/10.21037/qims-2025-1125/rc

Data Sharing Statement: Available at https://qims.amegroups.com/article/view/10.21037/qims-2025-1125/dss

Funding: This work was partially funded by the National Plan for NRRP Complementary Investments (PNC, established with the decree-law 6 May 2021, No. 59, converted by law No. 101 of 2021) in the Call for the Funding of Research Initiatives for Technologies and Innovative Trajectories in the Health and Care Sectors (Directorial Decree No. 931 of 06-06-2022)—AdvaNced Technologies for Human-centrEd Medicine (ANTHEM) (Project No. PNC0000003). This work was partially supported by the MUR under the grant "Dipartimenti di Eccellenza 2023-2027" of the Department of Informatics, Systems and Communication of the University of Milano-Bicocca, Milan, Italy. This work reflects only the authors' views and opinions; neither the Ministry for University and Research nor the European Commission can be considered responsible for them. The research leading to these results has received funding from AIRC under MFAG 2025—P.I. Capitoli Giulia (ID 32620 Project).

Conflicts of Interest: All authors have completed the ICMJE uniform disclosure form (available at https://qims.amegroups.com/article/view/10.21037/qims-2025-1125/coif). The authors have no conflicts of interest to declare.

Ethical Statement: The authors are accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved. This study was conducted in accordance with the Declaration of Helsinki and its subsequent amendments. The study was approved by the regional ethics committee (Comitato Etico Brianza, via Pergolesi, 33, 20900 Monza, Italy) (No. FINAL-TIR PU 3581/21, approval date: January 14, 2021), and informed consent was taken from all the patients.

References

1. Kitahara CM, Sosa JA. The changing incidence of thyroid cancer. Nat Rev Endocrinol 2016;12:646-53.
2. Balentine CJ, Domingo RP, Patel R, Laucirica R, Suliburk JW. Thyroid lobectomy for indeterminate FNA: not without consequences. J Surg Res 2013;184:189-92.
3. Nardi F, Basolo F, Crescenzi A, Fadda G, Frasoldati A, Orlandi F, Palombini L, Papini E, Zini M, Pontecorvi A, Vitti P. Italian consensus for the classification and reporting of thyroid cytology. J Endocrinol Invest 2014;37:593-9.
4. Capitoli G, Piga I, Galimberti S, Leni D, Pincelli AI, Garancini M, Clerici F, Mahajneh A, Brambilla V, Smith A, Magni F, Pagni F. MALDI-MSI as a Complementary Diagnostic Tool in Cytopathology: A Pilot Study for the Characterization of Thyroid Nodules. Cancers (Basel) 2019;11:1377.
5. Capitoli G, Piga I, L'Imperio V, Clerici F, Leni D, Garancini M, Casati G, Galimberti S, Magni F, Pagni F. Cytomolecular Classification of Thyroid Nodules Using Fine-Needle Washes Aspiration Biopsies. Int J Mol Sci 2022;23:4156.
6. Kaufman L, Rousseeuw PJ. Finding groups in data: an introduction to cluster analysis. Hoboken: Wiley-Interscience; 2005.
7. Seminati D, Capitoli G, Leni D, Fior D, Vacirca F, Di Bella C, Galimberti S, L'Imperio V, Pagni F. Use of Diagnostic Criteria from ACR and EU-TIRADS Systems to Improve the Performance of Cytology in Thyroid Nodule Triage. Cancers (Basel) 2021;13:5439.
8. Leni D, Seminati D, Fior D, Vacirca F, Capitoli G, Cazzaniga L, Di Bella C, L'Imperio V, Galimberti S, Pagni F. Diagnostic Performances of the ACR-TIRADS System in Thyroid Nodules Triage: A Prospective Single Center Study. Cancers (Basel) 2021;13:2230.
9. Fadda G, Basolo F, Bondi A, Bussolati G, Crescenzi A, Nappi O, Nardi F, Papotti M, Taddei G, Palombini L; SIAPEC-IAP Italian Consensus Working Group. Cytological classification of thyroid nodules. Proposal of the SIAPEC-IAP Italian Consensus Working Group. Pathologica 2010;102:405-8.
10. Ali SZ, Baloch ZW, Cochand-Priollet B, Schmitt FC, Vielh P, VanderLaan PA. The 2023 Bethesda System for Reporting Thyroid Cytopathology. Thyroid 2023;33:1039-44.
11. Tessler FN, Middleton WD, Grant EG, Hoang JK. Re: ACR Thyroid Imaging, Reporting and Data System (TI-RADS): White Paper of the ACR TI-RADS Committee. J Am Coll Radiol 2018;15:381-2.
12. Capitoli G, Piga I, Clerici F, Brambilla V, Mahajneh A, Leni D, Garancini M, Pincelli AI, L'Imperio V, Galimberti S, Magni F, Pagni F. Analysis of Hashimoto's thyroiditis on fine needle aspiration samples by MALDI-Imaging. Biochim Biophys Acta Proteins Proteom 2020;1868:140481.
13. Hartigan JA, Wong MA. Algorithm AS 136: A k-means clustering algorithm. J R Stat Soc Ser C Appl Stat 1979;28:100-8.
14. Xu R, Wunsch DC 2nd. Clustering algorithms in biomedical research: a review. IEEE Rev Biomed Eng 2010;3:120-54.
15. Nielsen F. Introduction to HPC with MPI for Data Science. Switzerland: Springer; 2016.
16. Rousseeuw PJ. Silhouettes: a graphical aid to the interpretation and validation of cluster analysis. Journal of Computational and Applied Mathematics 1987;20:53-65.
17. Tian X, Zhang G, Shao Y, Yang Z. Towards enhanced metabolomic data analysis of mass spectrometry image: Multivariate Curve Resolution and Machine Learning. Anal Chim Acta 2018;1037:211-9.
18. Sengupta A, Naresh G, Mishra A, Parashar D, Narad P. Proteome analysis using machine learning approaches and its applications to diseases. Adv Protein Chem Struct Biol 2021;127:161-216.

Cite this article as: Facchinetti F, L'Imperio V, Piga I, Papetti DM, Capitoli G. Boosting thyroid nodule diagnosis through ultrasound and molecular imaging integration with unsupervised learning. Quant Imaging Med Surg 2026;16(6):479. doi: 10.21037/qims-2025-1125

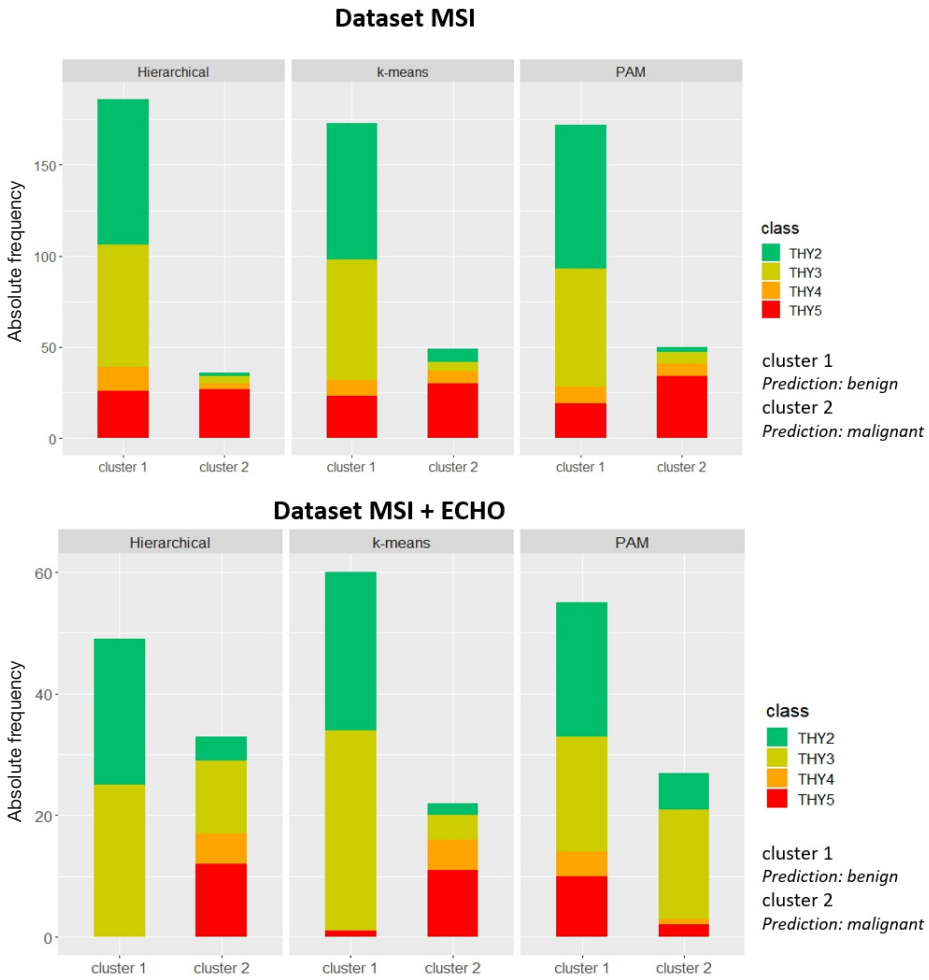

Figure S1 Bar plot showing the distribution of nodules categorized by class across various ML methods (HC, PAM, and k-means) within both Dataset MSI and Dataset MSI + ECHO. Dataset MSI: underwent MALDI-MSI analysis. Dataset MSI + ECHO: combines MSI and echographic data. HC, hierarchical clustering; MALDI, matrix-assisted laser desorption-ionization; ML, machine learning; MSI, mass spectrometry imaging; PAM, partitioning around medoids.

https://dx.doi.org/10.21037/qims-2025-1125

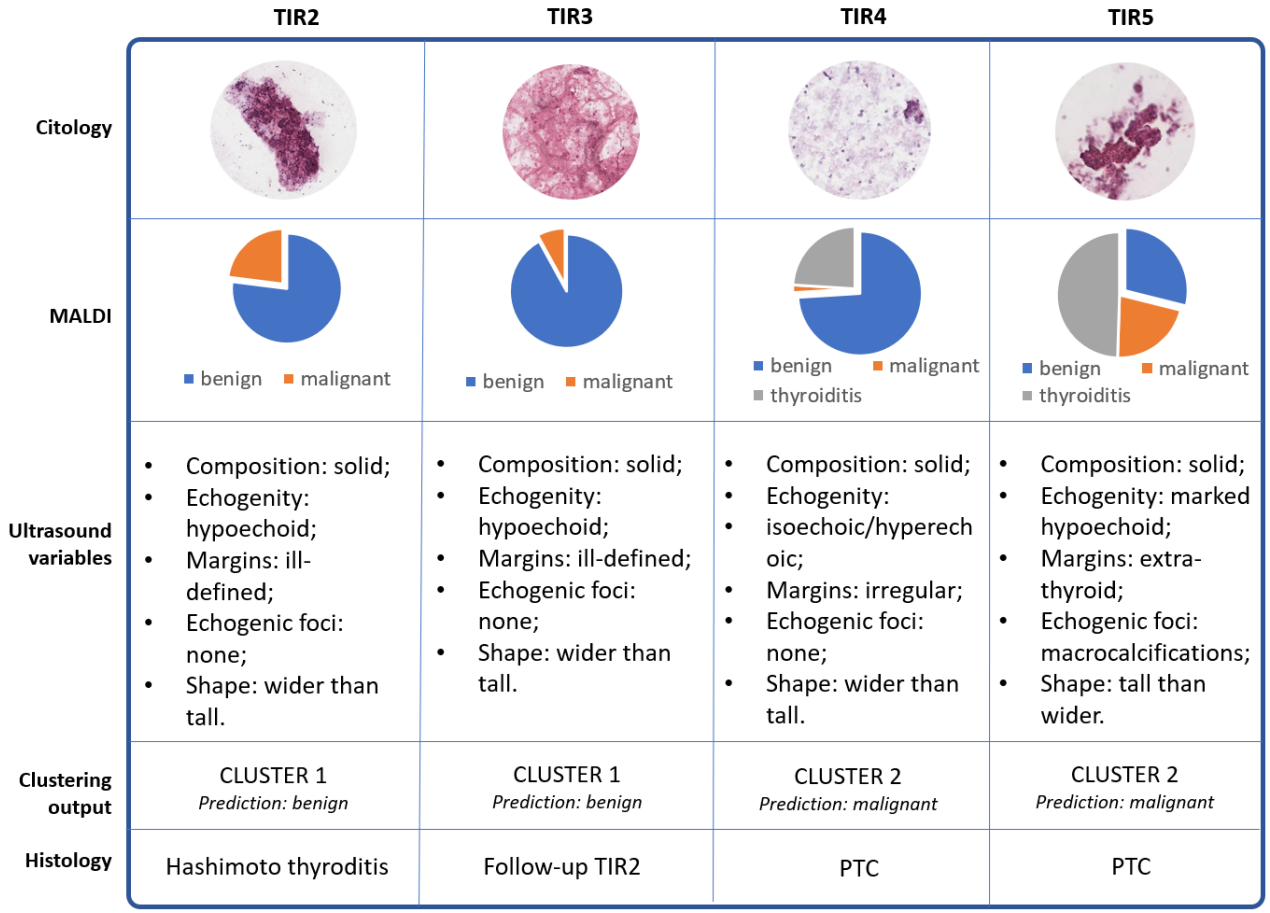

Figure S2 Four cases of different cytological diagnoses that MALDI-alone misclassified while the harmonization of MALDI with US (Dataset MSI + ECHO) examination correctly classified. Dataset MSI + ECHO: combines MSI and echographic data. MALDI, matrix-assisted laser desorption-ionization; MSI, mass spectrometry imaging; PTC, papillary thyroid carcinoma; US, ultrasound.

https://dx.doi.org/10.21037/qims-2025-1125