[18F]FDG-PET/CT Texture Analysis in "Hot" Thyroid Nodules: Results of the validation Cohort

Martina Sollini*[1], Margarita Kirienko[1], Paola Magnoni[2], Luca, Cozzi[1,3], Luca di Tommaso[1,4], Andrea Gerardo Lania[1,5], Arturo Chiti[1,6]

[1]Department of Biomedical Sciences - Humanitas University, Pieve Emanuele (Milan), Italy

[2]Ultrasound Service - Humanitas Clinical and Research Center, Rozzano (Milan), Italy

[2]Radiation Oncology - Humanitas Clinical and Research Center, Rozzano (Milan), Italy

[3]Pathology - Humanitas Clinical and Research Center, Rozzano (Milan), Italy

[4]Endocrinology and Diabetology - Humanitas Clinical and Research Center, Rozzano (Milan), Italy

[5]Nuclear Medicine - Humanitas Clinical and Research Center, Rozzano (Milan), Italy

***Corresponding author:** Martina Sollini, Department of Biomedical Sciences, Italy

ARTICLE INFO

ABSTRACT

Background: significance of thyroid nodule 2-deoxy-2-[18F]fluoro-D-glucose ([18F]FDG) uptake on positron emission tomography/computed tomography (PET/CT) scans remains controversial. We aimed to validate our previous results on the ability of texture analysis to predict final diagnosis in [18F]FDG-avid thyroid nodules.

Methods: We retrospectively evaluated medical records of all patients who performed an [18F]FDG-PET/CT from January 2017 to December 2018. Those patients who presented a thyroid nodule described in the medical records and performed a fine needle aspiration in our institution were considered for the analysis. Cytological and/or histological results were used as reference standard to define the final diagnosis. In case of negative cytology, the nodule was considered benign. In case of non-diagnostic or inconclusive results ultrasound, follow-up and/or further cytology/histology was used as final diagnosis. LifeX software was used for image segmentation (threshold of 40% SUV_{max}) and for textural features extraction. The 7 features previously identified as possible diagnostic biomarkers were tested by logistic regression. Receiver operating characteristic (ROC) curves were plotted and for each curve, the area under the curve (AUC) was calculated. Statistical analysis was performed with STATA package.

Results: We identified 498 medical reports describing [18F]FDG uptake in a thyroid nodule. Fifty-two patients performed further diagnostic steps in our institution. Twelve patients/52, referred to ultrasonographers, did not perform cytology. Eight/52 patients were excluded from the analysis because a final diagnosis was not available. Accordingly, 32 patients were analyzed. Nineteen out of 32 patients had benign nodules while in 13/32 cases a malignancy (primary thyroid cancer = 12, metastases = 1) was diagnosed. Conventional PET parameters and histogram-based features were calculated for all patients, while second and higher-order features were available for only 15/32 patients. Skewness and kurtosis resulted significantly different in benign and malignant nodules (p = 0.03 and = 0.02, respectively) and showed the best AUC (=0.75).

Conclusions: texture analysis may be useful to characterize [18F]FDG-avid thyroid nodules and to refine the selection of patients to be referred for cytology.

Introduction

The incidence of thyroid incidentalomas is increasing according to the fact that PET/CT is increasingly used for oncological and non-oncological applications. The chance of malignancy among the incidentalomas detected by [18F]FDG is higher than incidental

nodules discovered with other imaging modalities and the 2015 American Thyroid Association Management Guidelines [1] recommend fine-needle aspiration (FNA) in case of focal [18F]FDG uptake within a thyroid nodule. However, if it is true that a focal [18F]FDG uptake within a thyroid nodule has an increased risk of malignancy, it is also true that not all thyroid nodules characterized by a focal [18F]FDG uptake are malignant. Therefore, it would be really useful an approach able to select patients proposed for FNA. In our previous publication, we explored the role of textural analysis of [18F]FDG-PET to refine the selection of the patients with "hot" nodules to be referred for cytology. Seven indices including SUVstd, SUVmax, metabolic tumor volume (MTV), total lesion glycolysis (TLG), Skewness, Kurtosis, and Correlation calculated from the grey level co-occurrence matrix (Correlation$_{GLCM}$) were selected as potential diagnostic biomarkers able to stratify the patients with thyroid incidentaloma identified on [18F]FDG-PET/CT scans, with respect to the risk of a malignant thyroid nodule [2]. However, the role of the identified candidate biomarkers needs to be tested on a separate patient cohort to be validated. Therefore, the aim of this study was to validate the ability of texture analysis on [18F]FDG-PET/CT to predict final diagnosis in "Hot" Thyroid Nodules.

Materials and Methods

Patients

We retrospectively evaluated the medical records of all the patients who performed an [18F]FDG-PET/CT from January 2017 to December 2018. Methods for patient selection and data collection were previously detailed [2].

Among all subjects who underwent a [18F]FDG-PET/CT (8465 scans) and had a thyroid uptake described in the medical record (398 scans), we selected 52 patients who performed further diagnostic steps in our institution. Twelve patients, referred to ultrasonographers, did not perform cytology. Eight patients were excluded from the analysis because a final diagnosis was not available. Accordingly, 32 patients (male:female=5:27; age 64±12 years) were included in the analysis. Indication for [18F]FDG-PET/CT was oncological in 30 cases and non-oncological in the remaining 2 cases (1 vasculitis, 1 Langerhans cell histiocytosis). In the oncologic group, [18F]FDG-PET/CT was performed for breast tumor (8 cases), lymphoma (6 cases), gastro-intestinal tumors (5 cases), head and neck (2 cases), thymic masses (2 cases), bladder cancer (1 case), mesenchymal tumor (case 1) and lung cancer (1 case). In the remaining 4 cases [18F]FDG-PET/CT was performed in metastatic patients to search for the primary tumor (4 cases). This retrospective observational study was reviewed and approved by the Local Ethics Committee.

Cytology

Methods for fine-needle aspiration and cytology were previously detailed [2]. All cytological results were reported according the SIAPEC-IAP 2014 classification [3].

In case of negative cytology (TIR2), no further studies were immediately performed, and the nodule was considered benign. In case of non-diagnostic (TIR1) or doubtful (TIR3) results, further cytological samples or histology were used as final diagnosis when available (n=6), otherwise patients were excluded (n=8).

[18F]FDG PET/CT acquisition protocol and analysis

[18F]FDG-PET/CT image acquisition and visual interpretation were previously detailed [2]. PET images were segmented using LifeX software (http://www.lifexsoft.org) [4] settled with a threshold of 40% SUVmax. The region of interest (ROI), set in the target tissue (i.e., incidentaloma), was delineated in 3D and then textural features were extracted using the same package (http://www.lifexsoft.org) [5]. LifeX was settled up using the following parameters for calculation of the features: 64 gray levels to resample the ROI content which was performed in absolute terms between a minimum of 0 and a maximum of 20 [6]. A total of 43 features were extracted from the analysis of the volumes inspected. The correction for partial volume effect was not applied. As previously detailed, LifeX calculated shape, second order and higher-order features only for a ROI including at least 64 voxels.

Statistical Analysis

Statistical analysis was performed using the STATA package (StataCorp. 2017. Stata Statistical Software: Release 15. College Station, TX: StataCorp LLC). Continuous normally distributed data were expressed as mean ± standard deviation. Conventional PET parameters and textural features previously identified as possible diagnostic biomarkers (SUVstd, SUVmax, MTV, TLG, Skewness, Kurtosis, CorrelationGLCM) [2] were compared to final diagnosis using logistic regression. For each predictor the truth table (true and false positive and negative values) was determined. Receiver operating characteristic (ROC) curves were plotted and for each curve, the area under the curve (AUC) was calculated. Additionally, Compacity, previously identified as predictor able to identify TIR classes [2], was assessed by multinomial logistic regression. A p-value < 0.05 was considered statistically significant.

Results

Overall 32 patients were included in the study Table 1 summarizes the characteristics of the cohort of patients included in the analysis. FNA was non-diagnostic in 6 cases, negative in 17 cases, doubtful in 2 cases (TIR3), suspected in 2 cases, and positive for malignant tumor cells in 6 cases. Patients classified as TIR1, repeated the FNA which resulted non-diagnostic (TIR1) in 5 cases and suspicious (TIR4) in one case. The cases that resulted non-diagnostic twice, presented suspicious characteristics on ultrasonography and underwent surgery; pathology resulted positive for DTC in all cases. Both TIR3 cases underwent surgery resulting negative for malignancy (oxyphil cell adenoma = 1, thyroid hyperplasia =1). All TIR4 or TIR5 were proven to be malignant. Nodules resulted malignant (n=13) included 12 thyroid cancers (papillary thyroid cancer = 11, follicular thyroid cancer =1), and 1 metastasis from neuroendocrine tumor.

Table 1: Descriptive statistics of the cohort of patients in the study.

Parameter	Group		Number of cases
FNA results	TIR1		6
	TIR2		17
	TIR3		2
	TIR4		2
	TIR5		6
Final diagnosis	Negative		19
	Positive		13
Histological type*	DTC		12
		Papillary	11
		Follicular	1
	Metastases		1
FNA=Fine needle aspiration, DTC=Differentiated thyroid cancer *Only for patients with a final diagnosis positive for malignancy (n=13)			

Thyroid [^{18}F]FDG uptake was unilateral in the majority of patients (n=31) occurring in the isthmus in 1 case, in the right and left lobe in 15 and 13 cases, respectively. In the remaining case [^{18}F] FDG uptake involved both thyroid lobes. Fifteen patients (8 with a final diagnosis of DTC) had small ROI (i.e., < 64 voxels), therefore all features were available only for 17 patients (5 with a final diagnosis of DTC). Table 2 summarizes results of the 7 features tested as predictors. Skewness and kurtosis were significantly different in benign and malignant nodules Table 2. Figure 1 shows the box-plots of the 7 tested features stratified for the two subgroups within the

cohort. Figure 2 shows the ROC curves (together with the reference 50% AUC line) for the metabolic volume and the texture related features. Table 3 summarizes the predictors and their power in the identification of the negative or positive subgroups in the cohort. All features showed higher specificity than sensitivity. SUVstd classified as correct the highest number of nodules. Compacity failed in differentiating TIR classes (p=0.16). Figure 3 shows examples of four patients with thyroid nodules characterized by different patterns of [^{18}F]FDG uptake (low-moderate uptake in a and b, high uptake in c and d).

Table 2: Summary of tested predictors previously identified as potential diagnostic biomarkers.

Predictor	Mean ± St. dev	Range	AUC	p
Conventional				
SUVstd	1.2±1.12	0.34 – 4.61	0.57	0.06
SUVmax	7.6±5.5	2.55 – 22.34	0.56	0.06
MTV	131.03±148.85	13 - 583	0.7	0.12
TLG	14.18±10.35	2.65 – 44.84	0.64	0.13
Histogram-based				
Skewness	0.85±0.46	0.02 – 1.68	0.75	0.03
Kurtosis	3.13±1.05	1.55 – 5.41	0.75	0.02
GLCM-based				
Correlation^	0.31±0.22	0.02 – 0.63	0.61	0.64
AUC = area under curve; MTV=metabolic tumor volume; TLG=total lesion glycolysis;				
GLCM= gray-level co-occurrence matrix; SUV=standardized uptake value.				
^Data calculated only for 15/32 patients				

Table 3: Summary of predictive power of the significant predictors previously identified as potential diagnostic biomarkers.

Predictor	Specificity[%]	Sensitivity[%]	Positive predictive value [%]	Negative Predictive value [%]	Correctly Classified [%]
Conventional					
SUV$_{std}$	95	46	86	72	75
SUV$_{max}$	95	38	83	69	72
MTV	79	53	64	71	69
TLG	74	38	50	63	59

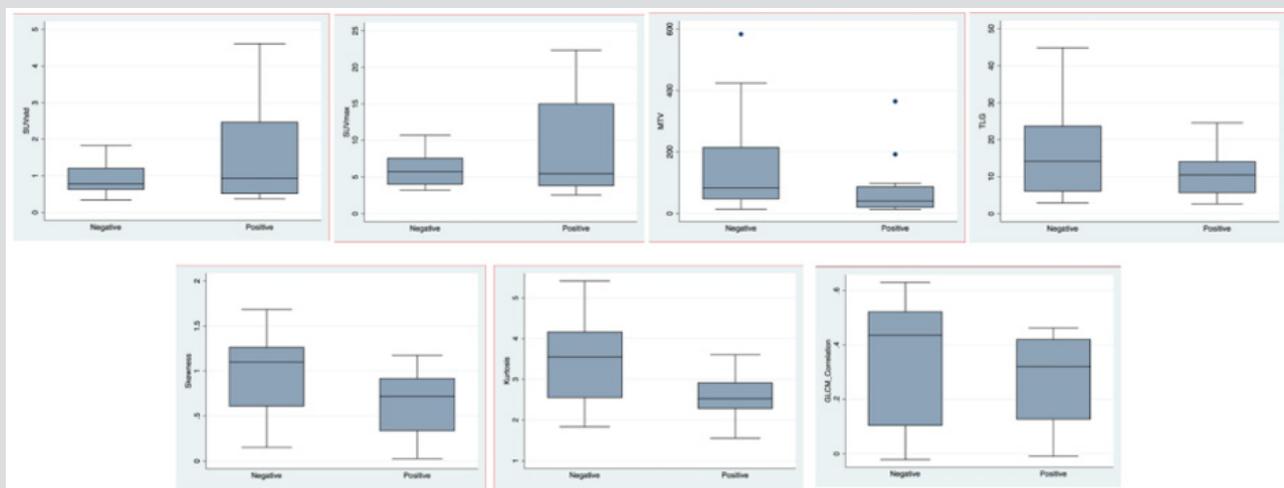| Histogram-based					
Skewness	74	46	54	66	62
Kurtosis	68	53	54	68	62
GLCM-based					
Correlation	100	-	-	73	73

Figure 1: Box plots of the 7 tested features stratified for the two subgroups (benign versus malignant) within the cohort.
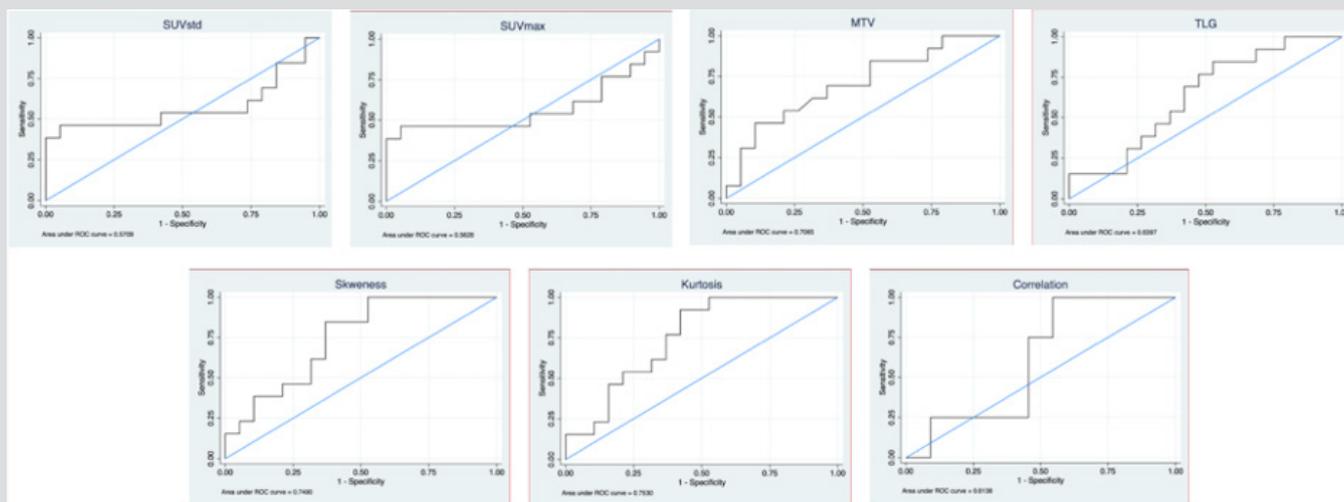

Figure 2: ROC curves (together with the reference 50% AUC line) for the 7 tested features.
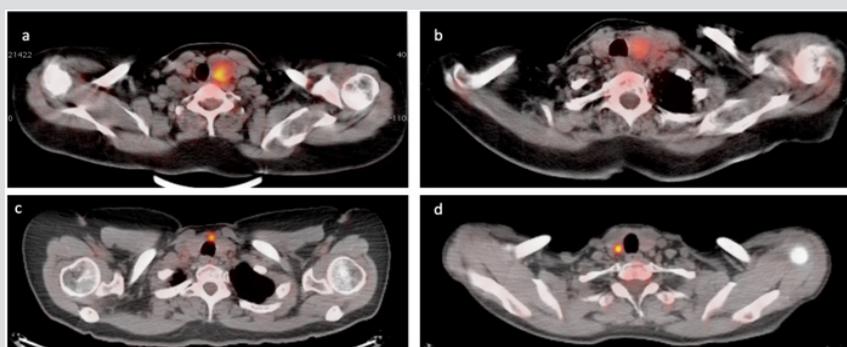

Figure 3: Cases of four patients with thyroid nodules characterized by different patterns of [18F]FDG uptake. Upper panels show benign (TIR2 in a) and malignant (TIR5, final diagnosis of papillary DTC - pT4apN1apM1 - in b) thyroid nodules with low-moderate [18F]FDG uptake. Lower images show benign (TIR3b, final diagnosis of oxyphil cell adenoma in c) and malignant (TIR4, final diagnosis of papillary DTC - pT1a - in d) thyroid nodules with high [18F]FDG uptake.

Discussion

Our results confirmed that textural analysis on [18F]FDG-PET/CT might be used to refine the selection of the patients with "hot" nodules to be referred for cytology. The prevalence of thyroid incidentaloma on [18F]FDG-PET/CT was higher than previously reported and literature data (about 4.7% versus 1.2-4.3%) [2,7-9] while the percentage of patients who performed additional workup was lower than our previous findings (ultrasound available in 13% versus 19%). However, the number of patients who underwent an FNA was comparable to our previous findings (10% versus 12% of subjects). In fact, the prognosis associated with underlying malignancy should be considered before investigation of [18F]FDG–avid thyroid incidetalomas and certainly before any treatment [10].

Skeweness and kurtosis confirmed to be significantly different in negative and positive nodules with a better discriminative power than previously reported (AUC=0.75 versus AUC=0.66 and AUC=0.55, respectively). Interestingly, the AUCs of most of the tested parameters resulted higher than in the "discovery" cohort (MTV, skeweness, kurtosis, correlation). In fact, generally, validation analysis present lower performance than obtained in the training phase. Distribution of image-derived indices within benign and malignant nodules were confirmed except that for MTV and TLG. In the present cohort, volumes of malignant nodules were smaller than benign ones, possibly negatively affecting the results. In fact, small volumes prevent the possibility to calculate shape and higher-order features by LifeX. In our population, all features were available for 17/32 patients (only 5 with a final diagnosis of DTC). However, small size is a typical feature of thyroid incidentaloma.

These results are promising since they confirmed most of our previous findings. Different approaches have been tested to identify [18F]FDG-avid nodules at high risk of malignancy with some disappointing results, mostly lacking of any type of validation [10-18].

Essential requirement for biomarkers research is the validation phase. In fact, novel parameters need to be confirmed reliable using an internal approach (e.g. cross-validation or bootstrapping) or an independent one (e.g. split sample, temporal or geographical), being the latter more robust [19]. We adopted a temporal inclusion strategy. The parameters that resulted significant in the discovery study have been tested in the present and resulted significantly associated with malignancy. The present study suffered from some limitations. Firstly, it was retrospective and evaluated a limited number of patients. Secondly, textural features were not available for the entire cohort. Thirdly, partial volume correction was not applied, however it should be considered according the small size of nodules.

Conclusion

Our results confirmed that textural features differed in benign and malignant [18F]FDG-avid nodules. Accordingly, textural analysis on [18F]FDG-PET/CT might be used to refine the selection of the patients with "hot" nodules to be referred for cytology.

Ethics Approval

Acknowledgement

Conflict-of-Interest Disclosure

Authors' Contributions

The manuscript has been seen and approved by all authors, whose individual contributions were as follows: MK and MS conceptualized the paper; PM performed ultrasound images and FNA; LDT provided cytological results; MS contributed to data selection; MS and CL performed the images and statistical analysis; MS and MK wrote the paper and provided figures and tables; AGL managed and followed patients and provided clinical data; AC supervised the project and paper writing.

References

1. Haugen BR, Alexander EK, Bible KC, Doherty GM, Mandel SJ, et al. (2016) 2015 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 26(1): 1-133.

2. Sollini M, Cozzi L, Pepe G, Antunovic L, Lania A, et al. (2017) [18F] FDG-PET/CT texture analysis in thyroid incidentalomas: preliminary results. Eur J Hybrid Imaging 1(1): 3.

3. Nioche C, Orlhac F, Boughdad S, Reuze S, Goya Outi J, et al. (2018) Lifex: A freeware for radiomic feature calculation in multimodality imaging

to accelerate advances in the characterization of tumor heterogeneity. Cancer Res 78(16): 4786-4789.

4. Nardi F, Basolo F, Crescenzi A, Fadda G, Frasoldati A, Orlandi F, et al. Italian consensus for the classification and reporting of thyroid cytology. J Endocrinol Invest. Springer International Publishing; 2014;37:593–9.

5. Orlhac F, Nioche C, Buvat I (2016) Technical Appendix - Local Image Features Extraction - LIFEx.

6. Orlhac F, Soussan M, Chouahnia K, Martinod E, Buvat I (2015) 18F-FDG PET-derived textural indices reflect tissue-specific uptake pattern in non-small cell lung cancer. PLoS One 10: 1-16.

7. Rigo P, Paulus P, Kaschten BJ, Hustinx R, Bury T, et al. (1996) Oncological applications of positron emission tomography with fluorine-18 fluorodeoxyglucose. Eur J Nucl Med 23(12): 1641-1674.

8. Mitchell J, Parangi S (2005) The thyroid incidentaloma: an increasingly frequent consequence of radiologic imaging. Semin Ultrasound CT MR 26(1): 37-46.

9. Liu Y, Ghesani NV, Zuckier LS (2010) Physiology and pathophysiology of incidental findings detected on FDG-PET scintigraphy. Semin Nucl Med 40(4): 294-315.

10. Pattison DA, Bozin M, Gorelik A, Hofman MS, Hicks RJ, et al. (2018) 18F-FDG-Avid Thyroid Incidentalomas: The Importance of Contextual Interpretation. J Nucl Med Society of Nuclear Medicine 59(5): 749-755.

11. Thuillier P, Bourhis D, Roudaut N, Crouzeix G, Alavi Z, et al. (2019) Diagnostic Value of FDG PET-CT Quantitative Parameters and Deauville-Like 5 Point-Scale in Predicting Malignancy of Focal Thyroid Incidentaloma. Front Med 6: 24.

12. Sager S, Vatankulu B, Sahin O, Cınaral F, Uslu L, et al. (2018) Clinical significance of standardized uptake values in thyroid incidentaloma discovered by F-18 fluorodeoxyglucose positron emission tomography/computed tomography. J Cancer Res Ther 14(5): 989-993.

13. Chung SR, Choi YJ, Suh CH, Kim HJ, Lee JJ, et al. (2018) Thyroid Incidentalomas Detected on 18F-Fluorodeoxyglucose Positron Emission Tomography with Computed Tomography: Malignant Risk Stratification and Management Plan. Thyroid 28(6): 762-768.

14. Shi H, Yuan Z, Yuan Z, Yang C, Zhang J, et al. (2018) Diagnostic Value of Volume-Based Fluorine-18-Fluorodeoxyglucose PET/CT Parameters for Characterizing Thyroid Incidentaloma. Korean J Radiol 19(2): 342-351.

15. Kim BH, Kim SJ, Kim K, Kim H, Kim SJ, et al. (2015) High metabolic tumor volume and total lesion glycolysis are associated with lateral lymph node metastasis in patients with incidentally detected thyroid carcinoma. Ann Nucl Med 29(8): 721-729.

16. Brindle R, Mullan D, Yap BK, Gandhi A (2014) Thyroid incidentalomas discovered on positron emission tomography CT scanning – Malignancy rate and significance of standardised uptake values. Eur J Surg Oncol 40(11): 1528-1532.

17. Hagenimana N, Dallaire J, Vallée É, Belzile M (2017) Thyroid incidentalomas on [18]FDG-PET/CT: a metabolico-pathological correlation. J Otolaryngol Head Neck Surg 46: 22.

18. Soelberg KK, Bonnema SJ, Brix TH, Hegedüs L (2012) Risk of malignancy in thyroid incidentalomas detected by 18F-fluorodeoxyglucose positron emission tomography: a systematic review. Thyroid 22(9): 918-925.

19. Park SH, Han K (2018) Methodologic Guide for Evaluating Clinical Performance and Effect of Artificial Intelligence Technology for Medical Diagnosis and Prediction. Radiology 286(3): 800-809.

Assets of Publishing with us

- Global archiving of articles
- Immediate, unrestricted online access
- Rigorous Peer Review Process
- Authors Retain Copyrights
- Unique DOI for all articles

https://biomedres.us/

BIOMEDICAL RESEARCHES

ISSN: 2574-1241