Research Article

Clinical Implication of the New AI-TIRADS Classification of Thyroid Nodules; Our Real Clinical Experience

Mohammad Hadi Gharib [iD] [1], Reza Afghani [iD] [2, *], Siamak Rajaei [iD] [1], Gholamreza Roshandel [iD] [3], Ahmad Alijani [iD] [1, **], Zolaykha Karamollahi [iD] [1], Mahin Tatari [iD] [4], Sina Mohajernoei [iD] [5], Sepideh Sadat Hosseini [iD] [6], Seyyedeh Atefeh Rezazadeh [iD] [7]

[1] Clinical Research Development Unit (CRDU), Department of Radiology, 5 Azar Hospital, Golestan University of Medical Sciences, Gorgan, Iran
[2] Clinical Research Development Unit (CRDU), Department of Surgery, 5 Azar Hospital, Golestan University of Medical Sciences, Gorgan, Iran
[3] Golestan Research Center of Gastroenterology and Hepatology, Golestan University of Medical Sciences, Gorgan, Iran
[4] Department of Statistical Sciences, Sapienza University of Rome, Rome, Italy
[5] Department of Medical Sciences, Faculty of Medicine, Golestan University of Medical Sciences, Gorgan, Iran
[6] Department of Education, 5 Azar Hospital, Golestan University of Medical Sciences, Gorgan, Iran
[7] Department of Surgery, 5 Azar hospital, Golestan University of Medical Sciences, Gorgan, Iran

[*] **Corresponding Author:** Clinical Research Development Unit (CRDU), Department of Surgery, 5 Azar Hospital, Golestan University of Medical Sciences, Gorgan, Iran. Email: af_med75@yahoo.com
[**] **Corresponding Author:** Clinical Research Development Unit (CRDU), 5 Azar Hospital, Golestan University of Medical Sciences, Gorgan, Iran. Email: ahmadalijani6483@gmail.com

Abstract

Background: Ultrasound and fine needle aspiration biopsy (FNAB) play a primary role in determining the nature of nodules.

Objectives: The aim of our study is to determine the diagnostic accuracy, sensitivity, and specificity of artificial intelligence-thyroid imaging reporting and data system (AI-TIRADS) in differentiating malignant from benign nodules to establish the treatment plan for patients and avoid unnecessary FNAB.

Methods: This cross-sectional study was conducted between September 2020 and September 2021 (one year). Patients (n = 133) who had an indication for surgery due to nodular thyroid disease and were referred to the therapeutic training center of 5 Azar and Falsafi private section hospitals, Gorgan, Iran, were included. The diagnostic accuracy of FNAB and AI-TIRADS was calculated.

Results: The mean age of patients with malignant pathology (39.65 ± 11.69) was lower than that of patients with benign pathology (44.14 ± 13.54), which was statistically significant (P = 0.042). Fifty-two out of 72 nodules less than 25 mm were malignant, demonstrating a higher prevalence of benign pathology in larger nodules. A pooled analysis of thyroid reporting (TR) = 2 and TR points ≥ 7 for their power in predicting the final pathology outcome showed figures of 0.97, 0.93, 0.97, and 0.97 for sensitivity, specificity, positive predictive value (PPV), and negative predictive value (NPV), respectively, with an overall pooled accuracy of 0.97 for these two groups of sonographic results.

Conclusions: We concluded that AI-TIRADS has a high sensitivity and diagnostic accuracy in differentiating benign from suspicious nodules.

Keywords: Diagnosis, Accuracy, AI-TIRADS, Thyroid Nodule

1. Background

A thyroid nodule is a solid or fluid-filled mass that may or may not produce thyroid hormone. These nodules may be small and asymptomatic, found incidentally or during ultrasound follow-up of a patient with thyroiditis or hypothyroidism, or may grow to an extent that creates cosmetic concerns or puts pressure on the airway or esophagus (1, 2).

Thyroid nodules can be benign or malignant. In general, most nodules are small and do not produce symptoms. It is important to determine the exact nature of the nodule because it affects the treatment plan and the outcome of the patient (2, 3). Ultrasound and fine needle aspiration biopsy (FNAB) play primary roles in

determining the nature of thyroid nodules. The main imaging modality for evaluating thyroid tissue, particularly thyroid nodules, is ultrasound because it is safe, cost-effective, widely available, and does not involve radiation exposure (4, 5).

To date, various guidelines have been introduced to determine whether to perform FNAB or to plan for follow-up of thyroid nodules. Some of these guidelines include the 2015 American Thyroid Association guideline, the American College of Radiology-TIRADS (ACR-TIRADS), and European-TIRADS. The sensitivity, specificity, and diagnostic accuracy of these guidelines have been reviewed and presented in the literature (6, 7). Despite revisions to these guidelines, improvements in radiologists' experience, and advancements in technology, unnecessary FNABs and surgeries are sometimes performed when follow-up would have been sufficient. Artificial intelligence-thyroid imaging reporting and data system (AI-TIRADS) is one of the revised guidelines, and it is associated with improved specificity while maintaining sensitivity compared to ACR-TIRADS. Based on AI-TIRADS features and nodule size, the indication for FNAB is determined (8).

2. Objectives

The purpose of this study is to investigate and determine the diagnostic accuracy, sensitivity, and specificity of AI-TIRADS (2022) in differentiating malignant from benign nodules, to establish a treatment plan and avoid unnecessary FNAB.

3. Methods

This cross-sectional descriptive analytical study was conducted to investigate the diagnostic value of AI-TIRADS in patients with thyroid nodules. The study included patients referred to the therapeutic training center of 5 Azar and Falsafi private section hospitals in Gorgan, Iran, between September 2020 and September 2021 (one-year period). A total of 133 patients who had indications for surgery due to nodular thyroid disease were included. Indications for surgery included a diagnosis of thyroid cancer, symptoms due to pressure on the airway or digestive tract from a large thyroid nodule, cosmetic reasons, or FNAB results showing malignancy or suspected malignancy.

All patients underwent ultrasound, and a second thyroid and neck ultrasound was performed by an expert in head and neck radiology to determine the AI-TIRADS score. Information from 35 patients with bilateral nodules was recorded, and the nodule with the higher TIRADS score was included in the study. If two nodules had equal TIRADS scores, the larger nodule was selected.

The majority of patients had undergone FNAB before referral. However, 21 cases did not undergo FNAB or had non-diagnostic results (7 cases were non-diagnostic, and 14 cases did not have FNAB). Reasons for not performing FNAB included multinodular goiter with external symptoms or pressure, a small nodule size (< 10 mm), or a low TIRADS score, with patients opting for a lobectomy or isthmectomy.

Pathology results were reported by two pathologists, and the data were analyzed using SPSS version 16 with *t*-test, Pearson chi-square, and Fisher's exact test. A significance level of 0.05 was used. Sensitivity, specificity, positive predictive value (PPV), negative predictive value (NPV), and diagnostic accuracy of AI-TIRADS were calculated. The study was approved by Golestan University of Medical Sciences with ethical code number IR.GOUMS.REC.1400.193.

4. Results

A total of 133 patients were recruited, including 121 women (91%) and 12 men (9%). The mean age was 38.83 ± 12.9 for men and 42.07 ± 12.75 for women, with no statistically significant difference. The mean age among those with benign pathology was 44.14 ± 13.54, while those with malignant nodules had a mean age of 39.65 ± 11.69, showing a lower mean age for malignant pathologies (P = 0.042). Body Mass Index analysis between the benign and malignant groups did not show significant differences (P = 0.524) (Table 1).

Family history of cancer did not reveal significant differences between benign and malignant groups. Further analysis of variables such as chief complaints, size and site of nodules, AI-TIRADS scores, FNAB results, and final pathology results are provided in Table 2.

Among nodules less than 25 mm in size, 52 out of 72 were malignant, showing a higher prevalence of benign pathology in larger nodules. Bilateral thyroid nodules were present in 35 patients, with 24 cases of multinodular goiter, 11 of which showed malignancy. Of these, 7 cases involved bilateral malignancy.

Artificial intelligence-thyroid imaging reporting and data system scores showed that 46 patients (34.6%) had TR ≤ 3, while 87 (65.5%) had TR ≥ 4. Among 30 patients with TR = 2, 29 had benign pathology, while one small (8 mm) nodule showed malignancy. In 59 patients with TR ≥ 7, 57 had malignant pathology (93.4%), reflecting the statistical significance of this score (P < 0.001). In a pooled analysis, only 4 out of 46 patients with TR ≤ 3 had malignant pathology (8.7%) (Figures 1 and 2).

Table 1. Number of the Patients in Body Mass Index Subgroups and Mean Age of the Patients According to Final Pathology

Variables and Category	N	Mean ± SD
BMI		
< 25	104	-
≥ 25	29	-
Final pathology		
Benign	63	22.16 ± 3.47
Malignant	70	22.58 ± 4.03

Among 87 patients with TR ≥ 4, 21 had benign results (24.4%), further demonstrating a significant correlation between TR score and final pathology (P < 0.001). Sensitivity, specificity, PPV, and NPV for AI-TIRADS were 0.94, 0.67, 0.76, and 0.91, respectively, with an overall accuracy of 0.81. A pooled analysis of TR = 2 and TR ≥ 7 showed sensitivity, specificity, PPV, and NPV of 0.97, 0.93, 0.97, and 0.97, respectively, with an overall accuracy of 0.97.

5. Discussion

In our study, the majority of patients (91%) were women, and out of 70 malignant nodules, 7 cases (10%) were men, and 63 cases (90%) were women. This gender predominance in suffering from nodular thyroid diseases and malignant thyroid nodules is likely due to differences in estrogen hormone levels between the sexes and its effect on hormone receptors in thyroid cells (6, 7). In our study, the mean age of patients with malignant nodules was significantly lower compared to those with benign nodules, consistent with findings from other studies (9).

Fine needle aspiration has long been considered the gold standard for evaluating thyroid nodules. However, ultrasound evaluation has recently gained increasing attention in diagnosing malignant nodules and determining the treatment plan in patients with nodular thyroid diseases. In older guidelines, FNA was routinely performed on nodules larger than 10 mm to distinguish benign from malignant nodules. This approach has led to unnecessary FNAs of benign nodules, which are not only unpleasant for the patient but also costly, and sometimes result in indeterminate cytopathologic reports. There are also challenges in differentiating among various follicular pathologies (e.g., follicular neoplasms including follicular adenoma, follicular carcinoma, or follicular variant of papillary thyroid cancer), Hurthle cell neoplasm, and new Bethesda classification pathologies such as atypia of undetermined significance (AUS) and follicular lesion of undetermined significance (FLUS) in FNA, which may lead to lobectomy, isthmectomy, and decision-making based on frozen sections during surgery.

With the introduction of ultrasonography and new classification guidelines, especially TIRADS, it has become possible to predict the potential malignancy of nodules based on composition, echogenicity, margins, microcalcifications, and echogenic foci. Consequently, unnecessary FNAs have been reduced, and the predictive accuracy of the nature of nodules has significantly improved before definitive surgery (2, 10). With the revision of these guidelines, the sensitivity and diagnostic accuracy of TIRADS have gradually increased (11). The latest guideline in this field, AI-TIRADS, has higher sensitivity, predictive value, and diagnostic accuracy compared to ACR-TIRADS, as demonstrated in multiple studies (6, 7, 12). Sensitivity and diagnostic accuracy are even relatively higher in cases with TIRADS 2 and 5 sonographic evaluations.

Similarly, in our study, the sensitivity and positive predictive value of AI-TIRADS were found to be 0.94 and 0.76 for all TR scores, while these values were 0.97 and 0.97 in patients with TR = 2 and TR points ≥ 7 (classic TIRADS 2 and 5 sonographic reports). Meanwhile, the negative predictive value of AI-TIRADS was 0.91 across all TR groups, and 0.97 in patients with TR = 2 and TR points ≥ 7. The overall diagnostic accuracy of this index was 0.81, and it was 0.97 for the TIRADS 2 and 5 groups. All of these indices were comparatively stronger than other well-known classification guidelines, such as ACR-TIRADS or British Thyroid Association (BTA), and notably more pronounced in the extreme ends of classification, where benign lesions and high cancer probabilities were found (TR = 2 and TR ≥ 7) (11).

Many clinicians recommend surgical resection when the thyroid nodule is larger than 4 cm, as larger nodules have an increased probability of being follicular thyroid cancer or the follicular variant of papillary thyroid cancer, and the false-negative rate of FNAB increases (13). However, the size of the nodule is not an absolute criterion for thyroidectomy; surgery is determined based on pressure or aesthetic symptoms, or the

Table 2. Frequency of Patients in Each Subgroup of Variables

Variables and Sub Groups	Number (%)
Chief complaint	
Compression symptoms	24 (18)
Cosmetic couse	62 (47)
Accidental finding	18 (13)
Follow up of thyroid disease	29 (22)
Site of the nodule	
Right lobe	60 (45)
Left lobe	36 (27)
Isthmus	2 (1.5)
Bilateral nodule (both lobs)	35 (26.5)
Size of the nodule (mm)	
5 - 9	15 (11)
10 - 14	25 (19)
15 - 24	32 (24)
25 - 34	25 (19)
40 ≤	36 (27)
AI-TIRADS score	
1 point	2 (1.5)
2 points	30 (23)
3 points	14 (10.5)
4 - 6 points	28 (21)
≥ 7 points	59 (44)
FNA result	
Non-diagnostic	7 (5.5)
Benign (colloid or adenomatoid)	32 (24)
Aus-flus	4 (3)
Suspicious for follicular neoplasm	17 (13)
Suspicios for malignancy	20 (15)
Malignant	39 (29.5)
Final pathology result	
Multinodular goiter	31 (23)
Adenomatoid or colloid nodule	29 (22)
Thyroiditis (atrophic-hashimoto-riedel)	3 (2.5)
Papillary thyroid cancer	61 (46)
Follicular thyroid cancer	7 (5)
Medullary thyroid cancer	2 (1.5)

suspicion of malignancy based on clinical presentation or malignant manifestations in ultrasound.

In this study, 86 patients were referred with pressure or aesthetic symptoms (24 patients with pressure symptoms and 62 with external symptoms). In the subgroup with pressure symptoms, 8 patients (33%) had malignant pathology, while 47 patients who were either incidentally found or followed up due to thyroid disease had malignant nodules. Among patients with nodules larger than 25 mm, 18 (29.5%) had malignancy, with
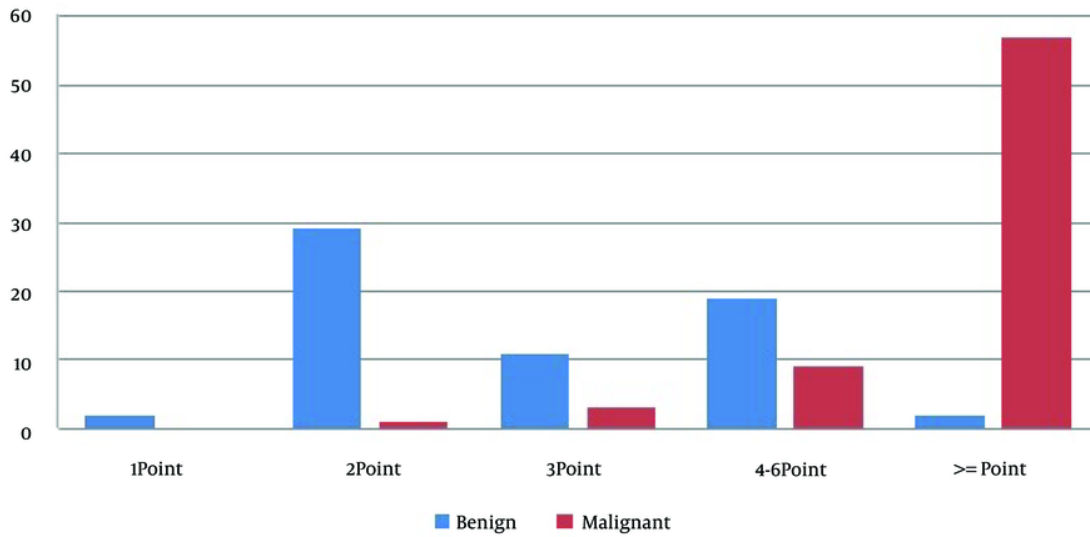
Figure 1. This figure shows the frequency of the number of patients based on artificial intelligence-thyroid imaging reporting and data system (AI-TIRADS) score.
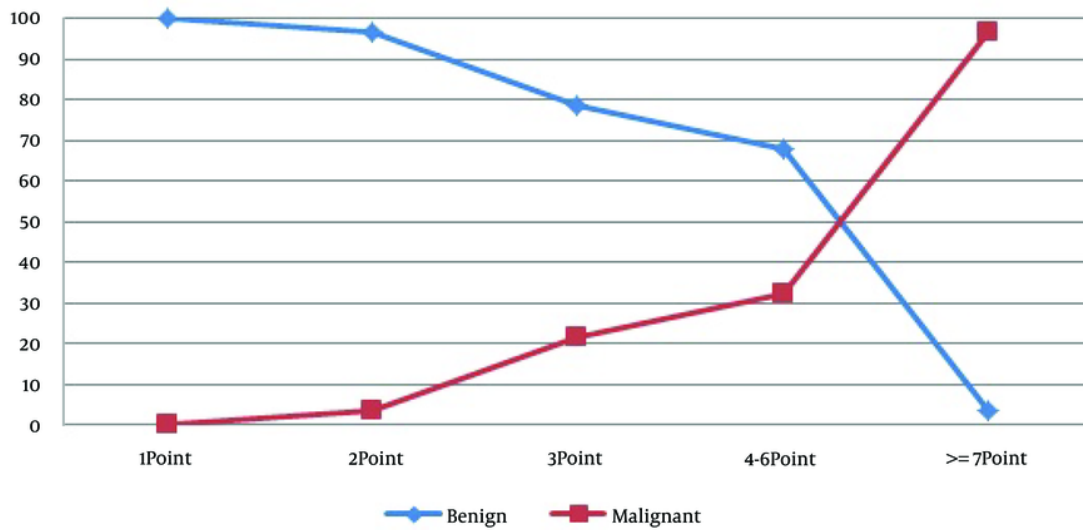

Figure 2. This figure shows linear graph of the number of patients based on artificial intelligence-thyroid imaging reporting and data system (AI-TIRADS) score.

papillary thyroid cancer (PTC) being the most common type in all size subgroups. Nodules that present with compression or external symptoms tend to be larger,

while those discovered incidentally or during follow-up of thyroid disease are smaller and asymptomatic.

In this study, malignant pathology was inversely related to nodule size, meaning that patients with

pressure symptoms often had benign pathology, whereas nodules found incidentally or during follow-up were more frequently malignant. In the majority of articles, smaller nodules are more likely to be malignant (14, 15). However, there are conflicting results in the literature, and relying solely on size for decision-making is not a reliable approach (13). It is now more accepted to make decisions based on ultrasound features, even for large nodules with benign FNA results. A higher probability of underlying malignancy should be considered if suspicious sonographic features are present. The new AI-TIRADS guideline, as demonstrated in our study, performs well in these problematic cases.

In Bethesda III cases, molecular and genetic tests are used to determine whether to follow up, repeat FNA, or consider surgery. These tests are not widely available, are expensive, and require expert interpretation. Their diagnostic accuracy is still under investigation, with varying results reported in the literature (16). Therefore, in cases where FNA results fall into this category, ultrasound findings can guide the treatment plan, and surgery may be strongly recommended if the nodule has TR points ≥ 4.

In Bethesda IV cases, differentiating benign from malignant pathology in FNA results of follicular neoplasms can be challenging. Thyroidectomy remains a common recommendation for these nodules. Nodule size and sonographic features can also help guide treatment decisions, with short-term follow-up and repeat FNA recommended for nodules smaller than 25 mm with no suspicious sonographic features. Surgery is reserved for larger nodules (> 25 mm) or those with suspicious imaging findings (17).

The future looks promising. With the use of artificial intelligence and the validation of ultrasound features for thyroid nodules, particularly through AI-TIRADS compared to ACR-TIRADS, it will be possible to differentiate malignant from benign nodules more accurately, especially in difficult thyroid nodule cases, as reviewed briefly above.

Acknowledgements

Footnotes

References

1. Grani G, Sponziello M, Pecce V, Ramundo V, Durante C. Contemporary Thyroid Nodule Evaluation and Management. *J Clin Endocrinol Metab*. 2020;**105**(9):2869-83. [PubMed ID: 32491169]. [PubMed Central ID: PMC7365695]. https://doi.org/10.1210/clinem/dgaa322.

2. Shi M, Nong D, Xin M, Lin L. Accuracy of Ultrasound Diagnosis of Benign and Malignant Thyroid Nodules: A Systematic Review and Meta-Analysis. *Int J Clin Pract*. 2022;**2022**:5056082. [PubMed ID: 36160289]. [PubMed Central ID: PMC9489364]. https://doi.org/10.1155/2022/5056082.

3. Singaporewalla RM, Hwee J, Lang TU, Desai V. Clinico-pathological Correlation of Thyroid Nodule Ultrasound and Cytology Using the TIRADS and Bethesda Classifications. *World J Surg*. 2017;**41**(7):1807-11. [PubMed ID: 28251273]. https://doi.org/10.1007/s00268-017-3919-5.

4. Haugen BR, Alexander EK, Bible KC, Doherty GM, Mandel SJ, Nikiforov YE, et al. 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*. 2016;**26**(1):1-133. [PubMed ID: 26462967]. [PubMed Central ID: PMC4739132]. https://doi.org/10.1089/thy.2015.0020.

5. Jiang S, Xie Q, Li N, Chen H, Chen X. Modified Models for Predicting Malignancy Using Ultrasound Characters Have High Accuracy in Thyroid Nodules With Small Size. *Front Mol Biosci*. 2021;**8**:752417. [PubMed ID: 34901151]. [PubMed Central ID: PMC8662815]. https://doi.org/10.3389/fmolb.2021.752417.

6. Koc AM, Adıbelli ZH, Erkul Z, Sahin Y, Dilek I. Comparison of diagnostic accuracy of ACR-TIRADS, American Thyroid Association (ATA), and EU-TIRADS guidelines in detecting thyroid malignancy. *Eur*

J Radiol. 2020;**133**:109390. [PubMed ID: 33181485]. https://doi.org/10.1016/j.ejrad.2020.109390.

7. Kovatcheva RD, Shinkov AD, Dimitrova ID, Ivanova RB, Vidinov KN, Ivanova RS. Evaluation of the Diagnostic Performance of EU-TIRADS in Discriminating Benign from Malignant Thyroid Nodules: A Prospective Study in One Referral Center. *Eur Thyroid J.* 2021;**9**(6):304-12. [PubMed ID: 33718254]. [PubMed Central ID: PMC7923902]. https://doi.org/10.1159/000507575.

8. Wildman-Tobriner B, Buda M, Hoang JK, Middleton WD, Thayer D, Short RG, et al. Using Artificial Intelligence to Revise ACR TI-RADS Risk Stratification of Thyroid Nodules: Diagnostic Accuracy and Utility. *Radiology.* 2019;**292**(1):112-9. [PubMed ID: 31112088]. https://doi.org/10.1148/radiol.2019182128.

9. Salamat F, Aryannia A, Rajaei S, Naeimi-Tabiei M, Afghani R, Sedaghat SM, et al. Temporal and Geographical Trends of Incidence of Thyroid Cancer in Golestan, Iran, 2004-2013. *Arch Iran Med.* 2021;**24**(1):1-6. [PubMed ID: 33588561]. https://doi.org/10.34172/aim.2021.01.

10. Yoon JH, Lee HS, Kim EK, Moon HJ, Park VY, Kwak JY. Pattern-based vs. score-based guidelines using ultrasound features have different strengths in risk stratification of thyroid nodules. *Eur Radiol.* 2020;**30**(7):3793-802. [PubMed ID: 32088739]. https://doi.org/10.1007/s00330-020-06722-y.

11. Watkins L, O'Neill G, Young D, McArthur C. Comparison of British Thyroid Association, American College of Radiology TIRADS and Artificial Intelligence TIRADS with histological correlation: diagnostic performance for predicting thyroid malignancy and unnecessary fine needle aspiration rate. *Br J Radiol.* 2021;**94**(1123):20201444. [PubMed ID: 33989038]. [PubMed Central ID: PMC8248201]. https://doi.org/10.1259/bjr.20201444.

12. Tan L, Tan YS, Tan S. Diagnostic accuracy and ability to reduce unnecessary FNAC: A comparison between four Thyroid Imaging Reporting Data System (TI-RADS) versions. *Clin Imaging.* 2020;**65**:133-7. [PubMed ID: 32470834]. https://doi.org/10.1016/j.clinimag.2020.04.029.

13. Kim HK, Kim SY, Lee YS, Soh EY, Chang HS, Park CS. Suspicious thyroid nodules 4 cm require a diagnostic lobectomy regardless of their benign fine needle aspiration results. *Asian J Surg.* 2022;**45**(5):1113-6. [PubMed ID: 34511361]. https://doi.org/10.1016/j.asjsur.2021.08.005.

14. Al-Hakami HA, Alqahtani R, Alahmadi A, Almutairi D, Algarni M, Alandejani T. Thyroid Nodule Size and Prediction of Cancer: A Study at Tertiary Care Hospital in Saudi Arabia. *Cureus.* 2020;**12**(3). e7478. [PubMed ID: 32351856]. [PubMed Central ID: PMC7188016]. https://doi.org/10.7759/cureus.7478.

15. Chung SR, Baek JH, Choi YJ, Sung TY, Song DE, Kim TY, et al. The relationship of thyroid nodule size on malignancy risk according to histological type of thyroid cancer. *Acta Radiol.* 2020;**61**(5):620-8. [PubMed ID: 31554409]. https://doi.org/10.1177/0284185119875642.

16. Khan TM, Zeiger MA. Thyroid Nodule Molecular Testing: Is It Ready for Prime Time? *Front Endocrinol (Lausanne).* 2020;**11**:590128. [PubMed ID: 33162941]. [PubMed Central ID: PMC7581778]. https://doi.org/10.3389/fendo.2020.590128.

17. Boonrod A, Akkus Z, Castro MR, Zeinodini A, Philbrick K, Stan M, et al. Thyroid Nodule Size as a Predictor of Malignancy in Follicular and Hurthle Neoplasms. *Asian Pac J Cancer Prev.* 2021;**22**(8):2597-602. [PubMed ID: 34452575]. [PubMed Central ID: PMC8629458]. https://doi.org/10.31557/apjcp.2021.22.8.2597.