Classic Variant Papillary Thyroid Carcinoma in Graves Disease and Multinodular Goiter Backgrounds: A Propensity Score-Matched Comparison

Kubilay Dalcı[1], *Burak Aydoğan*[1]

1 Department of General Surgery, Faculty of Medicine, Çukurova University, Adana, Türkiye

Abstract

Aim: There are conflicting data in the literature regarding the prognosis and aggressiveness of papillary thyroid carcinoma arising in the background of Graves disease. This study aimed to compare the histopathologic features of incidentally discovered classic variant papillary thyroid carcinomas in Graves disease and multinodular goiter backgrounds using propensity score matching.

Methods: Patients who underwent total thyroidectomy for benign indications and were found to have classic variant papillary thyroid carcinoma on histopathologic examination at Çukurova University Balcalı Hospital between 2015 and 2025 were retrospectively evaluated. Graves disease diagnosis was established based on the criteria of at least 2 years of antithyroid drug use, TSH receptor antibody positivity, and presence of clinical/radiologic ophthalmopathy. Groups were balanced by propensity score matching at a 1:3 ratio for age and sex. Histopathologic aggressiveness parameters were compared using non-parametric tests.

Results: A total of 44 patients with incidentally discovered classic variant papillary thyroid carcinoma following thyroidectomy for benign indications were categorized into Graves disease (n=11) and multinodular goiter (n=33) groups.After matching, standardized mean difference of <0.10 was achieved for all covariates. Median tumor size was significantly smaller in the Graves disease group compared to the multinodular goiter group (6.0 mm vs 10.0 mm, p=0.007). While lymphatic invasion was not detected in any patient in the Graves disease group, it was observed at a rate of 33.3% in the multinodular goiter group (p=0.041). Extrathyroidal extension, perineural invasion, and vascular invasion rates were found to be lower in the Graves disease group. No recurrence was observed in any patient during a median 44-month follow-up.

Conclusions: Classic variant papillary thyroid carcinomas arising in the background of Graves disease demonstrate smaller size and less invasive features compared to those arising in multinodular goiter background. These findings suggest that the autoimmune microenvironment in Graves disease may have a potential protective effect on tumor invasiveness.

Keywords: *Graves disease; multinodular goiter; papillary thyroid carcinoma; classic variant; propensity score matching; lymphatic invasion*

1. Introduction

Papillary thyroid carcinoma (PTC) accounts for 80-85% of all thyroid malignancies.[1] In some cases, it is incidentally detected in thyroidectomy specimens performed for benign indications. Concurrent PTC has been reported in 0.5-15% of patients operated for Graves disease (GD) and in 3-16% of those operated for multinodular goiter (MNG).[2,3] Despite similar rates of incidental PTC detection in these two pathologic backgrounds, they possess different pathogenesis and tumor microenvironments.

The relationship between GD and PTC has remained controversial for a long time. Early hypotheses proposed that chronic stimulation of follicular cells by TSH receptor antibodies (TRAb) could increase cellular proliferation, creating a potentially tumorigenic microenvironment.[4] This mechanism theoretically suggests that PTCs arising in the background of GD may exhibit more aggressive biological behavior; however, the evidence is insufficient, and current clinical data demonstrate significant heterogeneity.

Factors involved in MNG pathogenesis, such as chronic nodular hyperplasia, iodine deficiency, and variable TSH stimulation, constitute a different carcinogenesis process from the autoimmune activation present in GD.[5] Despite this biological heterogeneity, studies comparing tumor behavior between the two groups have frequently been conducted on methodologically unbalanced cohorts. Most of

Corresponding Author: Burak Aydoğan, baydogan1902@gmail.com, Received: 03.02.2026, Accepted: 14.03.2026, Available Online Date: 15.03.2026.
https://doi.org/10.36516/jocass.1881300

these studies have been inadequate in controlling for selection bias, demographic factors such as age and sex, and histopathologic variant differences.[6-8]

As a reflection of this heterogeneity, conflicting results have been reported regarding the clinical course and prognosis of PTCs arising in the background of GD. While some studies report that PTCs arising in this background exhibit aggressive clinical behavior such as multifocality/multicentricity and distant metastasis at diagnosis, other studies support that GD-associated PTCs are mostly small, indolent foci that do not alter survival outcomes.[6,9,10] For accurate analysis of this biological difference and tumor behavior, comparison groups must be balanced in terms of clinical and histopathologic data. Therefore, the independent effect of background pathology should be examined through a model homogenized in terms of variant distribution and other clinical variables.

The aim of this study was to compare the histopathologic features of classic variant papillary thyroid carcinomas (CV-PTC) incidentally detected in GD and MNG operated for benign indications. By achieving demographic homogeneity between groups through propensity score matching (PSM), we aimed to evaluate the effect of concomitant background pathology on tumor biology and aggressiveness parameters.

2. Materials and Methods

Study Design and Setting

This retrospective, matched cohort study was conducted on patients who underwent total thyroidectomy and were diagnosed with CV-PTC on histopathologic examination at the Department of General Surgery, Çukurova University Balcalı Hospital between 2015 and 2025. Clinical and imaging data were obtained from the hospital electronic medical record system, and histopathologic data were obtained from the pathology archive. Approval for the study protocol was granted by the Clinical Research Ethics Committee of Çukurova University Faculty of Medicine under decision number 25, dated January 9, 2026.

Patient Selection and Inclusion Criteria

-GD Group

At least 2 years of antithyroid drug use (methimazole/propylthiouracil), hyperthyroidism uncontrolled by medical therapy, acceptance of surgery in patients for whom radioactive iodine ablation or surgery was recommended, TRAb positivity, presence of clinical or radiologic ophthalmopathy, regular follow-up, undergoing total thyroidectomy, detection of CV-PTC on postoperative histopathology.

These criteria were aimed at selecting GD cases in a more homogeneous manner clinically and immunologically.

-MNG Group

Clinical-radiologic diagnosis of MNG, detection of CV-PTC arising in MNG background, undergoing total thyroidectomy.

Exclusion Criteria

PTC subtypes other than CV-PTC (follicular variant, tall-cell, diffuse sclerosing, etc.), incomplete/unavailable pathologic data, patients who underwent lobectomy or subtotal thyroidectomy.

Ultrasonography and Pathologic Evaluation

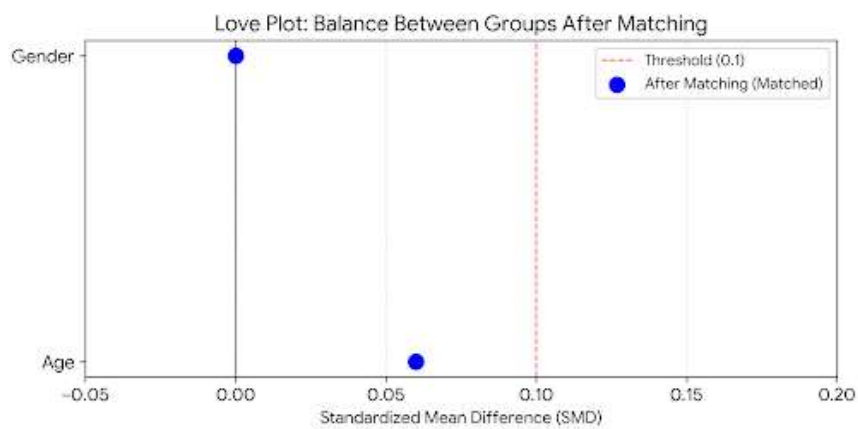
All ultrasonographic examinations were evaluated by a single experienced radiologist blinded to the pathology results, thereby minimizing potential observer bias. Histopathologic examinations were performed and reported by two independent pathologists.

Statistical Analysis

SPSS 25.0 software package was used for the statistical analysis of the data. Propensity scores were calculated using a logistic regression model based on age and sex. To achieve demographic balance, patients were matched at a 1:3 ratio using the nearest neighbor matching algorithm without replacement. The success of the matching process was confirmed by ensuring that the standardized mean difference (SMD) for all covariates remained below the 0.10 threshold Categorical measurements were summarized as numbers and percentages, while continuous measurements were summarized as Median (IQR: 25th-75th percentile). The conformity of variables to normal distribution was examined using the Kolmogorov-Smirnov test. Fisher's Exact test was applied to compare categorical expressions.

Figure 1

Love Plot: Covariate Balance Between Groups After Propensity Score Matching



This plot illustrates the covariate balance achieved following a 1:3 propensity score matching based on age and sex. The blue dots represent the standardized mean difference (SMD) for each variable. The SMD values for all covariates remain below the 0.10 threshold (indicated by the red dashed line), demonstrating a successful statistical balance between the GD and MNG groups.

For parameters that did not show a normal distribution, the Mann-Whitney U test was used to examine the differences between the two groups. The statistical significance level was set at 0.05 for all tests.

3. Results

A total of 44 patients included in the study were divided into two groups using PSM at a 1:3 ratio. The groups consisted of 11 patients with CV-PTC detected in GD background and 33 patients with CV-PTC detected in MNG background.

After matching, Standardized Mean Difference (SMD) values for all variables were found to be below the targeted threshold of 0.10 (Figure 1). This confirmed that complete covariate balance was achieved between groups prior to the main analyses.

In the analysis performed after matching, no statistically significant difference was found between the two groups in terms of age and sex distribution ($p > 0.05$). The median age of the GD group was 55.0 (IQR: 40.0 - 58.5) years, while the median age of the MNG group was 51.0 (IQR: 42.0 - 57.0) years ($p = 0.724$). Homogeneity was achieved in both groups in terms of sex distribution, with the female/male ratio in the GD group matching that in the MNG group one-to-one.

In the preoperative evaluation, the demographic data of the groups and the presence of nodules on ultrasonography are summarized in Table 1.

Table 1

Demographic characteristics and presence of preoperative nodules between the groups

Variable	GD Group (n=11)	MNG Group (n=33)	p-value
Age (Years), Median (IQR)	55.0 (40.0 - 58.5)	51.0 (42.0 - 57.0)	0.724
Gender, n (%)			
Female	5 (45.5%)	15 (45.5%)	1.000
Male	6 (54.5%)	18 (54.5%)	
Presence of Nodule, n (%)			
Present	9 (81.8%)	33 (100.0%)	0.059
Absent	2 (18.2%)	-	

IQR: Interquartile range (25th–75th percentiles). n: Number of patients. For statistical analyses, the Mann-Whitney U test was used for continuous variables, and Fisher's exact test was used for categorical variables. GD: Graves Disease. MNG: Multinodular Goiter

When preoperative ultrasonographic findings of the patients were examined, the median nodule size was 9.0 mm in the GD group, while it was 16.0 mm in the MNG group ($p = 0.094$). No statistically significant difference was found between groups in terms of the presence of hypoechogenicity, microcalcification, and irregular margins ($p > 0.05$). However, solid composition of nodules was observed at a significantly higher rate in the MNG group (78.8%) compared to the GD group (33.3%) ($p = 0.015$) (Table 2).

When tumor characteristics between groups were compared, the median tumor size in the GD group was 6.0 mm (5.5 - 7.5), while it was 10.0 mm (8.0 - 17.0) in the MNG group, and this difference was statistically significant ($p = 0.007$). Although multifocality and multicentricity rates were lower in the GD group (18.2% and 27.3%, respectively) compared to the MNG group (42.4% and 45.5%), this difference did not reach statistical significance ($p > 0.05$). A similar distribution was observed between groups in terms of tumor location and concomitant background pathologies (Table 3).

In the immunohistochemical evaluation performed on tumor foci, HBME-1 expression was observed positive in 72.7% (n=8) of patients in the GD group and in 51.5% (n=17) of patients in the MNG group (Table 4). Although HBME-1 positivity was numerically higher in PTC cases arising in the background of GD, this difference did not reach the level of statistical significance ($p = 0.301$) (Figure 2). No recurrence was detected in any patient during the median follow-up period of 44 months.

Table 2

Preoperative Ultrasonographic Nodule Characteristics of the Groups

Variable	GD Group (n=11)	MNG Group (n=33)	p-Value
Nodule Size (mm), Median (IQR)	9.0 (7.0-20.0)	16.0 (11.0-20.0)	0.094
Hypoechogenicity, n (%)	3 (33.3%)	20 (60.6%)	0.257
Microcalcification, n (%)	1 (11.1%)	17 (51.5%)	0.055
Solid Structure, n (%)	3 (33.3%)	26 (78.8%)	0.015
Irregular Border, n (%)	2 (22.2%)	18 (54.5%)	0.135

IQR: Interquartile range (25th–75th percentiles). n: Number of patients. For statistical analyses, the Mann-Whitney U test was used for continuous variables, and Fisher's exact test was used for categorical variables. GD: Graves Disease. MNG: Multinodular Goiter

Table 3

Histopathologic Characteristics of the GD and MNG Groups

Variable	GD Group (n=11)	MNG Group (n=33)	p-Value
Tumor Size (mm), Median (IQR)	6.0 (5.5 - 7.5)	10.0 (8.0 - 17.0)	0.007
Multifocality, n (%)	2 (18.2%)	14 (42.4%)	0.278
Multicentricity, n (%)	3 (27.3%)	15 (45.5%)	0.309
Tumor Location, n (%)			0.641
Right Lobe	5 (45.5%)	19 (57.6%)	
Left Lobe	3 (27.3%)	9 (27.3%)	
Isthmus	3 (27.3%)	5 (15.2%)	

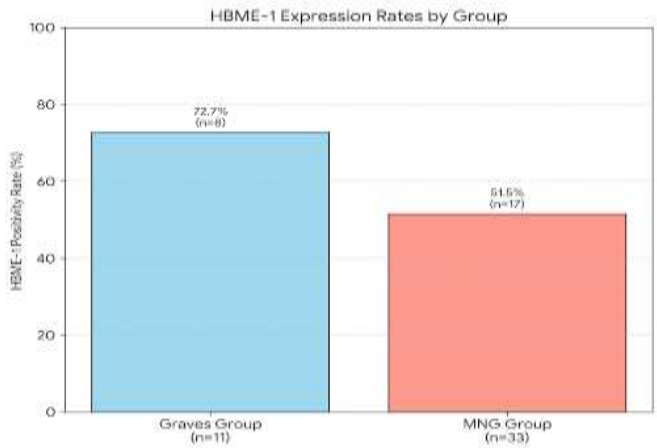
IQR: Interquartile range (25th–75th percentiles). n: Number of patients. For statistical analyses, the Mann-Whitney U test was used for continuous variables, and Fisher's exact test was used for categorical variables. GD: Graves Disease. MNG: Multinodular Goiter

Table 4

Extrathyroidal Extension and Invasion Characteristics of GD and MNG Groups

Variable	GD Group (n=11)	MNG Group (n=33)	p-Value
Capsular Invasion, n (%)	4 (36.4%)	17 (51.5%)	0.494
Vascular Invasion, n (%)	1 (9.1%)	5 (15.2%)	1.000
Lymphatic Invasion, n (%)	0 (0.0%)	11 (33.3%)	0.041
Extrathyroidal Extension (ETE), n (%)	0 (0.0%)	4 (12.1%)	0.558
Perineural Invasion (PNI), n (%)	0 (0.0%)	6 (18.8%)	0.312

Figure 2
HBME-1 Expression in GD and MNG Groups



The bar chart compares HBME-1 expression between the GD (n=11) and MNG (n=33) groups. Although the positivity rate was numerically higher in the GD group (72.7%) than in the MNG group (51.5%), this difference was not statistically significant (p=0.301).

4. Discussion

The analyses performed in this study, balanced by PSM for age and sex, demonstrated that CV-PTCs arising in GD background had statistically significantly smaller tumor size compared to those arising in MNG background, and histopathologic aggressiveness parameters such as lymphatic invasion, vascular invasion, ETE, and PNI were present at higher rates in the MNG group. These findings suggest that background pathology may have a potential modulatory effect on the biological behavior of tumors.

In our study, the median tumor size of PTC foci detected in the GD group (6.0 mm) was found to be significantly smaller than that in the MNG group (10.0 mm) (p=0.007). This finding is consistent with the smaller tumor size (mean 1.0 cm, p=0.002) reported by Gopinath et al. in their study of GD-associated differentiated thyroid carcinomas.[11] This may be explained by an early detection effect due to closer follow-up of GD patients, as well as the possibility that intense lymphocytic infiltration and autoimmune activation in GD create a microenvironment that suppresses tumor progression.[12] However, whether the small tumor size in the GD group truly reflects a more indolent biological behavior or is solely due to early detection effect could not be distinguished with the current retrospective design.

When histopathologic features were evaluated, lymphatic invasion was not detected in any patient in the GD group, while it was observed at a rate of 33.3% in the MNG group (p=0.041). Although there was no lymph node metastasis in any case in our study, this finding suggests that PTCs in GD background may have less aggressive biological behavior, as the presence of lymphatic invasion is an independent risk factor that negatively affects disease-free survival.[13] In the study by Kwon and Moon using PSM (114 GD-PTC vs 570 controls), the rate of lymphatic invasion in PTC cases in GD background was reported to be significantly lower than in the non-GD group (1.8% vs 6.7%, p=0.037).[6]

While ETE and PNI were not detected in the GD group, vascular invasion was observed in only a single case. In contrast, all invasion parameters were observed at higher rates in the MNG group (ETE 12.1%, PNI 18.8%, vascular invasion 15.2%, capsular invasion 51.5%). These findings are consistent with the study by Güden et al., which reported that ETE and PNI were not observed in PTCs in GD background, and the rate of vascular invasion was lower.[14] Although the underlying mechanism for the low invasion parameters in the GD group is not clear, the potential protective effect of intense lymphocytic infiltration and the autoimmune microenvironment has been proposed. However, these findings should be interpreted with caution due to the limited sample size.

Conflicting results have been reported in the literature regarding multifocality/multicentricity in PTC in GD background; in the study by Mekraksakit et al., an increased risk of multifocality/multicentricity was found in PTCs in GD background (OR 1.45, 95% CI 1.04–2.02).[8] In contrast, Sayiner et al. reported that multifocal involvement in PTC cases arising in GD background was lower compared to other PTC cases.[15] In our study, although multifocality (18.2% vs 42.4%, p=0.278) and multicentricity (27.3% vs 45.5%, p=0.309) rates were numerically lower in the GD group, they did not reach statistical significance. This result can be explained by low statistical power due to insufficient sample size and carries a risk of Type II error. Validation studies in larger cohorts are needed.

In preoperative ultrasonographic evaluation, solid nodule composition was detected at a significantly higher rate in the MNG group compared to the GD group (78.8% vs 33.3%, p=0.015). Although hypoechogenicity, microcalcification, and irregular margin features were also numerically higher in the MNG group, they did not reach statistical significance. These findings can be explained by the more prominent suspicious ultrasonographic features for malignancy due to the larger nodule size in the MNG group (median 16 mm vs 9 mm). The failure to detect nodules on preoperative ultrasonography in two patients (18.2%) in the GD group may be related to radiologic masking of very small foci in the background of diffuse hypoechoic/heterogeneous parenchyma and prominent hypervascularity characteristic of GD.[16]

Literature data regarding the prognosis of thyroid carcinomas detected in GD background are heterogeneous. While some studies report higher rates of multifocality and distant metastasis in GD-associated PTCs,[8] others report that these tumors are indolent foci and do not adversely affect survival outcomes.[15,17] It has particularly been emphasized that prognosis in tumors <1 cm may be independent of GD presence.[9,18] In the PSM analysis by Kwon and Moon, recurrence-free survival was found to be similar after matching, and GD was shown not to be an independent risk factor for recurrence.[6]

Although the findings of our study support the indolent phenotype hypothesis, the median 44-month follow-up period is insufficient for evaluating recurrence and long-term oncologic outcomes.

In immunohistochemical analysis, HBME-1 positivity was found to be numerically higher in tumors in the GD group (72.7% vs 51.5%); however, this difference did not reach statistical significance (p=0.301). In the literature, HBME-1 is accepted as a diagnostic marker with high sensitivity in distinguishing PTC from benign lesions.[19,20] In our study, despite higher HBME-1 positivity in the GD group, the smaller size of tumors and low invasion parameters suggest that the prognostic value of this marker is limited. Further studies analyzing HBME-1 expression profiles in different background pathologies are needed.

Limitations

Our study has several limitations. First, some clinical and laboratory data could not be accessed due to the retrospective design. The limited sample size increases the risk of Type II error. Although homogeneity was achieved in terms of age, sex, and histopathologic variant through PSM, other potential confounding variables such as body mass index, smoking status, and iodine intake could not be controlled. Our study reflects single-center experience, and the generalizability of the findings is limited. Additionally, only CV-PTCs were included in the study, and other histopathologic subtypes were not evaluated. Finally, the median follow-up period of 44 months was insufficient for evaluating long-term oncologic outcomes.

5. Conclusions

This PSM-controlled cohort study demonstrates that CV-PTC foci incidentally detected in GD background are smaller in size (p=0.007) and histopathologically more indolent compared to the MNG group. While our results suggest that the autoimmune microenvironment in GD may exert a potential protective effect against tumor invasiveness, these findings should be considered hypothesis-generating due to the study's retrospective design and sample size. Further large-scale, molecularly-based prospective studies are required to elucidate the underlying biological mechanisms.

Statement of ethics

Approval for the study protocol was granted by the Clinical Research Ethics Committee of Çukurova University Faculty of Medicine under decision number 25, dated January 9, 2026.

genAI

No artificial intelligence-based tools or generative AI technologies were used in this study. The entire content of the manuscript was originally prepared, reviewed, and approved by both authors.

Funding

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

Conflict of interest statement

The authors declare that they have no conflict of interest.

Availability of data and materials

The datasets used and/or analyzed during the current study are available from the corresponding author upon reasonable request.

References

1. Zhou J, Wang WR, Zhang HF, et al. Molecular and clinical features of papillary thyroid cancer in adult patients with a non-classical phenotype. Front Endocrinol (Lausanne). 2023;14:1138100. **Crossref**
2. Tez M. Letter to "the Prevalence of Incidental Thyroid Carcinoma in Patients Who Underwent Thyroidectomy for Multinodular Goiter". Indian J Surg. 2018;80(1):103. **Crossref**
3. May Mohammed Saleem AlQaraghuli ARA. The prevalence of micro papillary thyroid carcinoma and its correlation with age, gender and the histological background. World J Adv Healthc Res. 2022;6(11):1-4.
4. Belfiore A, Russo D, Vigneri R, Filetti S. Graves' disease, thyroid nodules and thyroid cancer. Clin Endocrinol (Oxf). 2001;55(6):711-718. **Crossref**
5. Yildirim Simsir I, Cetinkalp S, Kabalak T. Review of Factors Contributing to Nodular Goiter and Thyroid Carcinoma. Med Princ Pract. 2020;29(1):1-5. **Crossref**
6. Kwon H, Moon BI. Prognosis of papillary thyroid cancer in patients with Graves' disease: a propensity score-matched analysis. World J Surg Oncol. 2020;18(1):266. **Crossref**
7. Pellegriti G, Mannarino C, Russo M, et al. Increased mortality in patients with differentiated thyroid cancer associated with Graves' disease. J Clin Endocrinol Metab. 2013;98(3):1014-1021. **Crossref**
8. Mekraksakit P, Rattanawong P, Karnchanasorn R, et al. Prognosis of differentiated thyroid carcinoma in patients with Graves disease: A systematic review and meta-analysis. Endocr Pract. 2019;25(12):1323-1337. **Crossref**
9. Zhang D, Zhang G, Li Y, Liu H, Shan Z, Teng W. Prognosis of differentiated thyroid cancer in patients with Graves' disease: a meta-analysis. Thyroid Res. 2025;18(1):51. **Crossref**
10. Yano Y, Shibuya H, Kitagawa W, et al. Recent outcome of Graves' disease patients with papillary thyroid cancer. Eur J Endocrinol. 2007;157(3):325-329. **Crossref**
11. Gopinath C, Crow H, Panthi S, Bantis L, Burman KD, Choudhary C. Characteristics, staging and outcomes of differentiated thyroid cancer in patients with and without Graves' disease. J Clin Transl Endocrinol. 2023;33:100321. **Crossref**
12. Jaume JC. Thyroid Cancer—The Tumor Immune Microenvironment (TIME) over Time and Space. Cancers (Basel). 2025;17(5):794. **Crossref**
13. Nguyen VPT, Kurohama H, Akazawa Y, et al. Associations of clincopathologic features and mutation status with lateral lymph node metastasis as a predictor of disease-free survival in papillary thyroid carcinoma. Pathol Res Pract. 2025;269:155951. **Crossref**
14. Güden İ, Karatas I, Iscan Y, et al. Comparison of clinical and histopathological characteristics of differentiated thyroid cancers in euthyroid nodular goiter and Graves' disease. Endocr Connect. 2025;14(8). **Crossref**
15. Abidin Sayiner Z, Yatkin Keles Y, Ozturk S, Akarsu E. An Analytical Comparison of Papillary Thyroid Carcinoma Patients Manifested with or without Graves' Disease. Sisli Etfal Hastan Tip Bul. 2024;58(3):354-358. **Crossref**
16. Yuen HY, Wong KT, Ahuja AT. Sonography of diffuse thyroid disease. Australas J Ultrasound Med. 2016;19(1):13-29. **Crossref**
17. González-Sánchez-Migallón E, Flores-Pastor B, Pérez-Guarinos CV, et al. Incidental versus non-incidental thyroid carcinoma: Clinical presentation, surgical management and prognosis. Endocrinol Nutr. 2016;63(9):475-481. **Crossref**
18. Song Y, Fu L, Wang P, et al. Effect of Graves' disease on the prognosis of differentiated thyroid carcinoma: a meta-analysis. Endocrine. 2020;67(3):516-525. **Crossref**
19. Mohan U, Sunny SP, Mendonca P, Kuriakose MA, Kannan S, Suresh A. Systematic Review and Meta-Analysis to Identify the Immunocytochemical Markers Effective in Delineating Benign from Malignant Thyroid Lesions in FNAC Samples. Endocr Pathol. 2022;33(2):243-256. **Crossref**
20. de Matos LL, Del Giglio AB, Matsubayashi CO, de Lima Farah M, Del Giglio A, da Silva Pinhal MA. Expression of CK-19, galectin-3 and HBME-1 in the differentiation of thyroid lesions: systematic review and diagnostic meta-analysis. Diagn Pathol. 2012;7:97. **Crossref**