Balıkesir Medical Journal
e-ISSN: 2564-6664
B Med J 2024;8(1):26-34
DOI: 10.33716/bmedj.1424770

Malignancy Rates in Thyroid Nodules Classified as Benign According to the Nodule Size Threshold of 4 cm

Nodül Boyutu 4 cm. Üzerinde Benign Olarak Sınıflandırılan Tiroid Nodüllerinde Malignite Oranları

Yunushan Furkan AYDOĞDU[1*] Emre GÜLÇEK[2] Çağrı BÜYÜKKASAP[3] Kürşat DİKMEN[3]

[1] Bandırma Research and Training Hospital, Department of General Surgery, Balıkesir, Turkey.
[2] Polatlı Duatepe State Hospital, Department of General Surgery, Ankara, Turkey.
[3] Gazi University Faculty of Medicine, Department of General Surgery, Ankara, Turkey.

Received / Geliş Tarihi: **24.01.2024** ***Accepted / Kabul Tarihi:*** **21.02.2024**

Abstract

Aim: The false negative rate of fine needle aspiration biopsy (FNAB) in large thyroid nodules ranges from 7.7% to 53%. The treatment of nodules that are 4 cm or larger is controversial because of the potential for false-negative FNAB results. In order to make treatment recommendations for nodules ≥4 cm in size, we compared the results of preoperative FNAB with thyroid surgery specimens.

Materials and Methods: We evaluated patients who underwent thyroid surgery and had preoperative FNAB between 2017 and 2022. Patients were divided into two groups based on nodule size (<4 cm and ≥4 cm). FNAB and specimen pathology results of both groups were compared.

Results: 982 patients who underwent surgery for nodular disease were evaluated. Patients in the Bethesda 2 group (n=231) were divided into two categories based on nodule size: ≥4 cm (n=56) and <4 cm (n=175). Malignancy rates were higher in the group with tumors smaller than 4 cm. FNAB and specimen results were compared. A statistically significant difference was found between the groups (p = 0.039).

Conclusion: The benign fine-needle aspiration biopsy results were consistent with the specimen results in patients with nodules that were 4 cm or larger. Thyroid surgery should not be considered as the initial treatment for nodules ≥4 cm with benign cytology.

Keywords: Thyroid cancer; Nodular thyroid disease; Fine needle aspiration biopsy; Thyroidectomy; False negative rate.

Özet

Amaç: Büyük tiroid nodüllerinde ince iğne aspirasyon biyopsisinin (İİAB) yanlış negatiflik oranı %7,7 ile %53 arasında değişmektedir. Yanlış negatif İİAB sonuçları potansiyeli nedeniyle 4 cm. veya daha büyük nodüllerin tedavisi tartışmalıdır. Büyüklüğü 4 cm. ve üzerindeki nodüllere tedavi önerilerinde bulunmak amacıyla preoperatif İİAB sonuçlarını tiroid cerrahisi spesmenleri ile karşılaştırdık.

Gereç ve Yöntem: 2017-2022 yılları arasında tiroid ameliyatı geçiren ve ameliyat öncesi İİAB yapılan hastaları değerlendirdik. Hastalar nodül boyutlarına göre (<4 cm. ve ≥4 cm.) iki gruba ayrıldı. Her iki grubun İİAB ve spesmen patoloji sonuçları karşılaştırıldı.

Bulgular: Nodüler hastalık nedeniyle ameliyat edilen 982 hasta değerlendirildi. Bethesda 2 grubundaki hastalar (n=231) nodül boyutlarına göre ≥4 cm. (n=56) ve <4 cm. (n=175) olmak üzere iki kategoriye ayrıldı. Tümörü 4 cm.'den küçük olan grupta malignite oranları daha yüksekti. İİAB ve spesmen sonuçları karşılaştırıldı. Gruplar arasında istatistiksel olarak anlamlı fark bulundu (p=0,039).

Sonuç: Nodülü 4 cm. ve üzerinde olan hastalarda benign ince iğne aspirasyon biyopsisi sonuçları spesmen sonuçlarıyla uyumluydu. Benign sitolojiye sahip ≥4 cm. nodüllerde tiroid cerrahisi ilk tedavi seçeneği olarak düşünülmemelidir.

Anahtar Kelimeler: Tiroid kanseri; Nodüler tiroid hastalığı; İnce iğne aspirasyon biyopsisi; Tiroidektomi; Yanlış negatiflik oranı

***Corresponding Author / Sorumlu Yazar:** Yunushan Furkan Aydoğdu, Bandırma Research and Training Hospital, Department of General Surgery, Balıkesir, Turkey. **E-mail**: yfaydogdu92@gmail.com

INTRODUCTION

Thyroid nodules are clinically important because although more than 90% of identified nodules are clinically insignificant benign lesions (HaugenBryan et al., 2016; Popoveniuc & Jonklaas, 2012), they may indicate thyroid cancer in 4.0% to 6.5% of cases (Popoveniuc & Jonklaas, 2012). With an increasing incidence rate, it ranked ninth globally in 2020 (Siegel et al., 2018). Thyroid cancers have a favorable prognosis (Altiner et al., 2022).

The results of fine needle aspiration biopsy (FNAB) provide us, the surgeons, with an idea about the treatment (Hou et al., 2023). Fine needle aspiration biopsy prevents unnecessary surgeries and complications (Cibas et al., 2008; Wong et al., 2018). According to the Bethesda categorization system, thyroid FNAB cytology can define six main diagnostic groups (Cibas & Ali, 2017).

The 2015 ATA guidelines (HaugenBryan et al., 2016) (Table 1) state that estimated malignancy risks based on diagnostic Bethesda classification categories differ from estimated malignancy risks based on specimen results. In the diagnosis of thyroid cancer, FNAB has been reported to have a false negative rate of 2.5% (Amrikachi et al., 2001). Recently, the false-negative rate for FNAB has been evaluated in patients with large thyroid nodules, and multiple studies have found that these patients have higher false-negative rates for FNAB, ranging from 7.7% to 53% (Baser et al., 2022; Pinchot et al., 2009; Shi et al., 2017). The guidelines do not clearly state whether nodules with benign cytology below 4 cm have a higher chance of developing cancer or whether they should be treated differently from nodules above 4 cm (HaugenBryan et al., 2016; Russ et al., 2017). Therefore, unnecessary operations may be performed (HaugenBryan et al., 2016). In the literature, it has been reported that the treatment approach to large nodules is controversial due to the high false negative rate of FNAB in ≥4 cm nodules. (Baser et al., 2022; Carrillo et al., 2000; Giles et al., 2015; McCoy et al., 2007; Mehanna et al., 2013; Meko & Norton, 1995; Pinchot et al., 2009; Wharry et al., 2014).

We aimed to obtain evidence-based treatment recommendations for ≥4 cm nodules by comparing thyroid surgery specimens obtained from patients operated in our center with the results of preoperative FNAB. With this comparison, we aimed to investigate the compatibility of our results with the current literature and to determine the necessity of thyroidectomy in ≥4 cm thyroid nodules.

Table 1. Predicted malignancy risks according to the diagnostic categories of the Bethesda classification and malignancy risks according to the specimen result [1].

Category	Predicted Risk of Malignancy (%)	Risk of malignancy according to the specimen (%)
Bethesda 1	1-4	20 (9-32)
Bethesda 2	0-3	2.5 (1-10)
Bethesda 3	5-15	14 (6-48)
Bethesda 4	15-30	25 (14-34)
Bethesda 5	60-75	70 (53-97)
Bethesda 6	97-99	99 (94-100)

MATERIALS AND METHOD

January 1, 2017 and June 30, 2022, 982 patients who underwent preoperative FNAB and thyroidectomy at General Surgery clinic were retrospectively evaluated. Patients in whom preoperative FNAB was not performed in our center, patients in whom preoperative nodule size was unknown, and patients in whom postoperative specimen results were not evaluated in our center were excluded from the study (Figure 1). Patients (n=689) were divided into 6 groups using Bethesda classification. In our study, dominant nodules in thyroid tissue were considered as index nodules. In operated patients and pathology results evaluation belongs to the index nodule. Postoperative specimen pathology results were then compared for each group separately and false negative rates were compared with the existing literature. Demographic data, dominant nodule size and benign/malignant status of the patients were evaluated. In the Bethesda 2 group, a biopsy was taken from the index nodule. Patients who underwent procedures other than index nodule were not included in the study. The Bethesda 2 group (n=231), which differed from our data in the literature comparison, was divided into two subgroups as ≥4 cm and <4 cm and evaluated.

Study was approved by the Gazi University Clinical Researchs Ethics Committee (2023-076). The study complies with the Declaration of Helsinki.

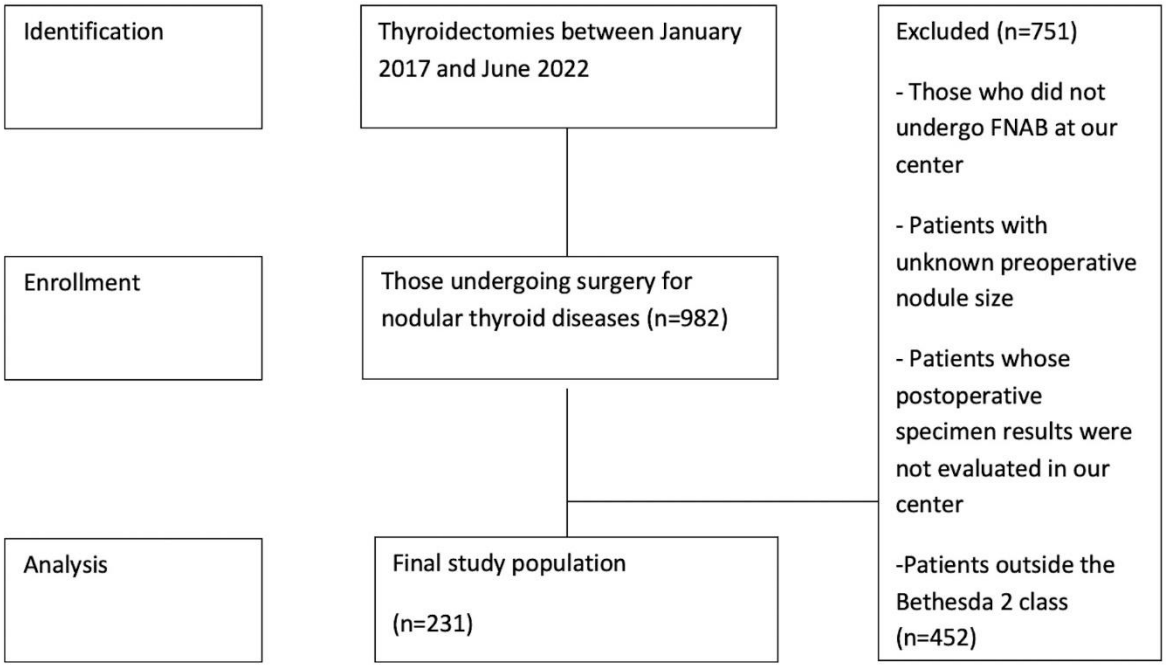

Figure 1. The sample collection scheme.

Statistical Analysis

The Statistical Package for the Social Sciences (SPSS) version 26.0, developed by SPSS Inc. in Chicago, IL, USA, was used. The median (min-max) or mean ± standard deviation (SD) was used to express numerical parameters. Percentages were used to evaluate categorical variables. For parameters that did not exhibit normal distribution, the Mann-Whitney U test was used to compare two independent groups. The association between binary categorical groups was assessed using the chi-square test. A p-value of <0.05 was considered statistically significant.

RESULTS

Of the study population, 527 (76.5%) were female and 162 (23.5%) were male. When subdivided according to the Bethesda classification, there were 4 patients (0.6%) in the Bethesda 1 group, 231 patients (33.5%) in the Bethesda 2 group, 138 patients (20%) in the Bethesda 3 group, 93 patients (13.5%) in the Bethesda 4 group, 69 patients (10%) in the Bethesda 5 group and 154 patients (22.4%) in the Bethesda 6 group. Malignant specimens were observed 0% in Bethesda 1 group, 18.2% in Bethesda 2 group, 44.9% in Bethesda 3 group, 38.7% in Bethesda 4 group, 59.4% in Bethesda 5 group and 95.5% in Bethesda 6 group. The specimen malignancy rate was as high as 18.2% in the Bethesda 2 group (Table 2).

Table 2. Distribution and clinicopathological characteristics of the groups according to Bethesda classification

	Bethesda 1	Bethesda 2	Bethesda 3	Bethesda 4	Bethesda 5	Bethesda 6	Total
Number of patients (n)	4	231	138	93	69	154	689
Age (year) (median), range	53.0 (50.0-60.0)	52.0 (20.0-74.0)	50.0 (21.0-78.0)	52.0 (19.0-81.0)	49.0 (18.0-75.0)	46.0 (18.0-76.0)	50.0 (18.0-81.0)
Gender							
Female n (%)	3 (75%)	179 (77.5%)	101 (73.2%)	71 (76.3%)	55 (79.7%)	118 (76.6%)	527 (76.5%)
Male n (%)	1 (25%)	52 (22.5%)	37 (26.8%)	22 (23.7%)	14 (20.3%)	36 (23.4%)	162 (23.5%)
Malignity							
Benign n (%)	4 (100%)	189 (81.8%)	76 (55.1%)	57 (61.3%)	28 (40.6%)	7 (4.5%)	361 (52.4%)
Malign n (%)	0 (0%)	42 (18.2%)	62 (44.9%)	36 (38.7%)	41 (59.4%)	147 (95.5%)	328 (47.6%)
Nodule size (cm), (median), range	1.60 (0.40-3.50)	2.50 (0.30-9.50)	2.30 (0.40-9.0)	2.0 (0.20-2.0)	1.50 (0.30-6.0)	1.50 (0.30-6.0)	2.0 (0.20-9.50)

In the continuation of the study, 231 patients in the Bethesda 2 group were evaluated. Demographic data showed that there were 179 female (77.5%) and 52 male (22.5%) patients. The median age was 52.0 years. When the final pathology results of the Bethesda 2 group were evaluated according to the specimen results, there were 189 benign (81.8%) and 42 malignant (18.2%) patients. The Bethesda 2 group was divided into groups by nodule size. There were 56 (24.2%) patients with ≥4 cm nodules and 175 (75.8%) patients with <4 cm nodules. There were 5 (8.9%) malignant patients in the ≥4 cm group and 37 (21.1%) in the <4 cm group. A significant difference was found when the results of FNAB and final specimen were compared (p=0.039) (Table 3).

Table 3. Sociodemographic, clinical and pathological characteristics of the Bethesda 2 group

Bethesda 2 group (n=231)	**Benign (n=189 (81.8%))**	**Malign (n=42 (18.2%))**	***p* value**
Age (Year, Median)	52.0	50.0	0.9
Gender (Female/Male)	146/43	33/9	0.85
<4 cm n (%)	138 (78.9%)	37 (21.1%)	**0.039**
≥ 4 cm n (%)	51 (91.1%)	5 (8.9%)	

DISCUSSION

Current literature in thyroid diseases is determined and updated according to ATA data (HaugenBryan et al., 2016). Based on the latest updated guidelines, evaluation of our data according to the Bethesda classification showed that our results were mostly consistent with the

current literature (HaugenBryan et al., 2016). FNAB results were evaluated and compared with specimen pathology. False negative rates of FNAB were similar in Bethesda 1, Bethesda 3, Bethesda 5 and Bethesda 6 classes. However, a higher malignancy rate was seen in the Bethesda 2 class with 18.2% and in the Bethesda 4 class with 38.7% malignant patients compared to the literature (Table 1) (HaugenBryan et al., 2016; Popoveniuc & Jonklaas, 2012).

ATA guidelines remain controversial as they do not provide a clear answer as to whether ≥4 cm nodules with benign cytology carry a greater risk of malignancy and whether there should be a different approach from <4 cm nodules. This uncertainty may lead to unnecessary operations (HaugenBryan et al., 2016). Most studies report that the treatment approach for large nodules is controversial due to the higher rate of false negatives of FNAB in ≥4 cm nodules (Baser et al., 2022; Carrillo et al., 2000; Giles et al., 2015; McCoy et al., 2007; Mehanna et al., 2013; Meko & Norton, 1995; Pinchot et al., 2009; Wharry et al., 2014). Moreover, despite these studies, endocrinologists and surgeons disagree on the treatment approach for large nodules (Pinchot et al., 2009).

McCoy et al. (McCoy et al., 2007) evaluated 233 patients with nodules ≥4 cm in size and reported a false negative rate of approximately 16% for FNAB when all Bethesda classifications were included. Thyroid nodules measuring 4 cm or larger were more likely to develop carcinoma. They recommended diagnostic lobectomy for thyroid nodules ≥4 cm in size, regardless of the FNAB result. Pinchot et al. (Pinchot et al., 2009) reported a high rate of inaccurate FNAB results in patients with nodules ≥4 cm. They found that half of the patients reported as benign on FNAB were misclassified. Both preoperative fine-needle aspiration biopsy (FNAB) cytopathology and specimen pathology results were reported in 97 out of 155 patients (62.5%) with nodules <4cm. Although they supported the use of FNAB in thyroid nodules, they recommended diagnostic lobectomy in patients with nodules ≥4 cm regardless of FNAB cytology. Wharry et al. (Wharry et al., 2014) found that the incidence of thyroid cancer in nodules ≥4 cm was 22% (n=83/382). The study reported that the presence of suspicious ultrasonographic features did not differentiate malignant nodules from benign nodules. Furthermore, the study found that the risk of thyroid cancer was 20% in 86 nodules that were ≥4 cm and lacked suspicious ultrasonographic features. They recommended a primary thyroid lobectomy for thyroid nodules ≥4 cm in size. Giles et al. (Giles et al., 2015) reported a high rate of false negative results in large thyroid nodules and recommended lobectomy for such nodules, even when the FNAB cytology indicates benignity. Carrillo et al. (Carrillo et al., 2000) suggested in a study of 159 patients that close follow-up and lobectomy should be considered

when negative cytology is present with a nodule size of ≥4 cm. Meko et al. (Meko & Norton, 1995) recommended thyroid lobectomy as the diagnostic method due to the high false negative rate of ≥3 cm fine-needle aspiration biopsy (FNAB). Kim et al. (Kim et al., 2022) evaluated 123 patients with nodules ≥4 cm and benign cytology. According to their results, they recommended primarily diagnostic lobectomy for these nodules due to the high rate of false negative results from preoperative FNAB. Lee et al. (Lee et al., 2016) demonstrated that performing routine thyroid lobectomy leads to lower costs for thyroid nodules ≥4 cm. Baser et al. (Baser et al., 2022) recommended performing diagnostic lobectomy for nodules <4 cm due to the high likelihood of false-negative results from preoperative cytology.

Shi et al. (Shi et al., 2017) analyzed 337 patients and reported that 99 patients had nodules <4 cm, while 238 patients had nodules ≥4 cm. In their study, they found no difference in malignancy rates for nodules less than 4 cm or 4 cm or larger. In addition to these studies, in our extensive series of 982 patients, we categorized 231 patients based on nodule size using the Bethesda 2 classification (benign). Consistent with the findings of Shi et al. (Shi et al., 2017), our study also demonstrated that nodules larger than or equal to 4 cm did not elevate the risk of malignancy. However, the risk of malignancy had a higher false-negative rate compared to the sample results in nodules <4 cm.

Our study included 56 patients (24.2%) with ≥4 cm nodules and 175 patients (75.8%) with <4 cm nodules. There were 5 (8.9%) malignant patients in the ≥4 cm group and 37 (21.1%) malignant patients in the <4 cm group. When the FNAB and specimen results were compared, it was determined that there was a significant difference. (p=0.039).

CONCLUSION

Contrary to the existing literature, benign FNAB results were compatible with specimen results in patients with nodules ≥4 cm. However, the malignancy rates predicted by the specimen pathology results in <4 cm nodules were well above the literature. The fact that we had similar rates with the current literature in the other classes except for the Bethesda 2 class shows that we do not have any limitations in the evaluation of FNAB and postoperative specimen pathology. Thyroid surgery should not be considered as the initial option when benign cytology is present in nodules ≥4 cm.

Funding: None

Conflicts of Interest: The authors declare no conflict of interest.

Authors Contributions: YF.A., E.G. and Ç.B. wrote the main manuscript text. YF.A., K.D. prepared figure and tables. All authors reviewed the manuscript.

Ethics Approval: The study protocol was approved by the ethics committee of the Gazi University (2023-076)

REFERENCES

Altiner, S., Kozan, R., Emral, A. C., Taneri, F., & Karamercan, A. (2022). Effects of patient and tumor characteristics on central lymph node metastasis in papillary thyroid cancer: a guide for selective node dissection. *Archives of Iranian Medicine*, *25*(11), 730-736.

Amrikachi, M., Ramzy, I., Rubenfeld, S., & Wheeler, T. M. (2001). Accuracy of fine-needle aspiration of thyroid: a review of 6226 cases and correlation with surgical or clinical outcome. *Archives of pathology & laboratory medicine*, *125*(4), 484-488.

Baser, O. O., Koseoglu, D., Cetin, Z., Catak, M., & Kizilkaya, H. (2022). Benign cytology does not rule out malignancy in thyroid nodules larger than 4 cm. *Diagnostic cytopathology*, *50*(11), 508-512.

Carrillo, J. F., Frias-Mendivil, M., Ochoa-Carrillo, F. J., & Ibarra, M. (2000). Accuracy of fine-needle aspiration biopsy of the thyroid combined with an evaluation of clinical and radiologic factors. *Otolaryngology—Head and Neck Surgery*, *122*(6), 917-921.

Cibas, E. S., Alexander, E. K., Benson, C. B., De Agustín, P. P., Doherty, G. M., Faquin, W. C., Middleton, W. D., Miller, T., Raab, S. S., & White, M. L. (2008). Indications for thyroid FNA and pre-FNA requirements: a synopsis of the National Cancer Institute Thyroid Fine-Needle Aspiration State of the Science Conference. *Diagnostic cytopathology*, *36*(6), 390-399.

Cibas, E. S., & Ali, S. Z. (2017). The 2017 Bethesda system for reporting thyroid cytopathology. *Thyroid*, *27*(11), 1341-1346.

Giles, W. H., Maclellan, R. A., Gawande, A. A., Ruan, D. T., Alexander, E. K., Moore, F. D., & Cho, N. L. (2015). False negative cytology in large thyroid nodules. *Annals of surgical oncology*, *22*, 152-157.

HaugenBryan, R., AlexanderErik, K., BibleKeith, C., DohertyGerard, M., MandelSusan, J., NikiforovYuri, E., RandolphGregory, W., SawkaAnna, M., SchuffKathryn, G., & ShermanSteven, I. (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*.

Hou, Y., Gao, Y., Guo, S., Zhang, Z., Chen, R., & Zhang, X. (2023). Applications of spatially resolved omics in the field of endocrine tumors. *Frontiers in Endocrinology*, *13*, 993081.

Kim, H. K., Kim, S. Y., Lee, Y. S., Soh, E. Y., Chang, H.-S., & Park, C. S. (2022). Suspicious thyroid nodules 4 cm require a diagnostic lobectomy regardless of their benign fine needle aspiration results. *Asian Journal of Surgery*, *45*(5), 1113-1116.

Lee, L., Mitmaker, E. J., Chabot, J. A., Lee, J. A., & Kuo, J. H. (2016). Cost-effectiveness of diagnostic lobectomy versus observation for thyroid nodules> 4 cm. *Thyroid*, *26*(2), 271-279.

McCoy, K. L., Jabbour, N., Ogilvie, J. B., Ohori, N. P., Carty, S. E., & Yim, J. H. (2007). The incidence of cancer and rate of false-negative cytology in thyroid nodules greater than or equal to 4 cm in size. *Surgery*, *142*(6), 837-844. e833.

Mehanna, R., Murphy, M., McCarthy, J., O'Leary, G., Tuthill, A., Murphy, M. S., & Sheahan, P. (2013). False negatives in thyroid cytology: impact of large nodule size and follicular variant of papillary carcinoma. *The Laryngoscope*, *123*(5), 1305-1309.

Meko, J. B., & Norton, J. A. (1995). Large cystic/solid thyroid nodules: a potential false-negative fine-needle aspiration. *Surgery*, *118*(6), 996-1004.

Pinchot, S. N., Al-Wagih, H., Schaefer, S., Sippel, R., & Chen, H. (2009). Accuracy of fine-needle aspiration biopsy for predicting neoplasm or carcinoma in thyroid nodules 4 cm or larger. *Archives of surgery*, *144*(7), 649-655.

Popoveniuc, G., & Jonklaas, J. (2012). Thyroid nodules. *Medical Clinics of North America*, *96*(2), 329-349.

Russ, G., Bonnema, S. J., Erdogan, M. F., Durante, C., Ngu, R., & Leenhardt, L. (2017). European Thyroid Association guidelines for ultrasound malignancy risk stratification of thyroid nodules in adults: the EU-TIRADS. *European thyroid journal*, *6*(5), 225-237.

Shi, H., Bobanga, I., & McHenry, C. R. (2017). Are large thyroid nodules classified as benign on fine needle aspiration more likely to harbor cancer? *The American Journal of Surgery*, *213*(3), 464-466.

Siegel, R. L., Miller, K. D., & Jemal, A. (2018). Cancer statistics, 2018. *CA: a cancer journal for clinicians*, *68*(1), 7-30.

Wharry, L. I., McCoy, K. L., Stang, M. T., Armstrong, M. J., LeBeau, S. O., Tublin, M. E., Sholosh, B., Silbermann, A., Ohori, N. P., & Nikiforov, Y. E. (2014). Thyroid nodules (≥ 4 cm): can ultrasound and cytology reliably exclude cancer? *World journal of surgery*, *38*, 614-621.

Wong, R., Farrell, S. G., & Grossmann, M. (2018). Thyroid nodules: diagnosis and management. *Medical Journal of Australia*, *209*(2), 92-98.