Endocrine (2019) 66:249–253
https://doi.org/10.1007/s12020-019-01964-3

ORIGINAL ARTICLE

Check for updates

Thyroid nodules over 4 cm do not have higher malignancy or benign cytology false-negative rates

Muhammed Kizilgul[1,2] · Rupendra Shrestha[1] · Angela Radulescu[1] · Maria R. Evasovich[1] · Lynn A. Burmeister[1]

Received: 26 February 2019 / Accepted: 20 May 2019 / Published online: 29 May 2019

Abstract

Purpose Whether thyroid nodules 4 cm or larger with benign cytology carry a higher risk of malignancy, and should be managed differently than smaller nodules remains controversial. We aimed to evaluate the malignancy rate and benign cytology false-negative rate in thyroid nodules ≥4 cm compared with those <4 cm.

Methods All thyroidectomies between January 2010 and December 2014 were reviewed. Patient demographics, preoperative sonographic nodule size, fine needle aspiration cytology (FNAC), and final surgical pathology results were compared for index nodules ≥4 vs. <4 cm.

Results A total of 490 index nodules with preoperative FNAC were identified. A total of 137 nodules were ≥4 cm and 353 nodules were <4 cm. The prevalence of carcinoma was lower (23 vs. 53%) in nodules ≥4 vs. <4 cm ($p < 0.0001$). The false-negative rate of benign FNAC for ≥4 and <4 cm index nodule was 5.2% and 5.9%, respectively ($p = 1.000$).

Conclusions This study shows that thyroid nodules ≥4 cm do not have a higher malignancy rate at surgery nor higher benign cytology false-negative rate than smaller nodules. Thyroid nodules over 4 cm do not require resection, to rule out malignancy, based on size alone.

Keywords Thyroid nodule · ≥4 cm · Thyroid carcinoma · False-negative rate · Size

Introduction

The prevalence of thyroid nodule varies depending on the mode of examination, from about 6% by palpation to 65–68% by autopsy or high frequency (13 MHz) ultrasound (US) [1, 2]. Current thyroid cancer diagnostic guidelines recommend fine needle aspiration cytology (FNAC), to determine risk of malignancy, based on sonographic patterns combined with nodule size [3, 4]. Some studies have concluded that nodules ≥4 cm have a higher malignancy [5–9] and false-negative rate at FNAC [5, 6, 9–15]. Some authors recommend excision of nodules ≥4 cm regardless of cytology result [6, 10, 16–18]. Others state this is not necessary [5, 7–9, 11–15, 19–31]. Eleven of the past studies did not have a control group [6, 14, 16–19, 21, 24, 26, 27, 31]. The 2015 ATA Guidelines state "Based on the evidence, it is still unclear if patients with thyroid nodules ≥4 cm and benign cytology carry a higher risk of malignancy and should be managed differently than those with smaller nodules" [3]. Whether the presence of a thyroid nodule 4 cm or larger, regardless of FNAC, should be an independent indication for thyroidectomy remains controversial. In this regard, we aimed to evaluate the overall risk of malignancy and the performance of FNAC in patients with thyroid nodules ≥4 cm who underwent surgical resection.

Material and methods

This study was approved by the University of Minnesota Institutional Review Board. All thyroidectomies performed at the University of Minnesota Medical Center between January 2010 and December 2014 were retrospectively reviewed, identifying 590 subjects. Pediatric (age < 18 years), hyperthyroid cases (with solitary hot nodule or surgery performed only for treatment of hyperthyroidism due to lack of biopsied thyroid nodule), and nodules without well documented

✉ Lynn A. Burmeister
Burme008@umn.edu

[1] Department of Medicine, Surgery University of Minnesota, Minneapolis, MN, USA

[2] Department of Endocrinology and Metabolism, UHS Diskapi Training and Research Hospital, Ankara, Turkey

FNAC prior to thyroid surgery were excluded (100 patients). To be considered well documented, the actual FNAC report had to be available for review. We did not exclude subjects who had Graves' disease or toxic multinodular goiter but also had discovery of thyroid nodule which led to FNAC and surgery. Each individual is represented once, even if they had more than one thyroid operation.

All patients underwent US-guided FNAC of index thyroid nodules, chosen by the treating providers based on sonographic appearance or size, prior to thyroid surgery. All FNACs were classified according to 2008 Bethesda criteria [32]. All thyroidectomies and histologic interpretation occurred at the University of Minnesota Medical Center.

Data were recorded from preoperative thyroid US and FNAC results, correlating index nodule sonographic size, cytology, and surgical histology. If a patient had more than one biopsy of the same nodule, the first FNAC, was used in the analysis. If more than one nodule was biopsied in a given subject, the largest nodule greater than 4 cm was selected as the index nodule, or, if all nodules were under 4 cm, the nodule with most abnormal cytology was selected as the index nodule. All FNAC were classified by 2008 Bethesda categories: nondiagnostic, benign, atypia of undetermined significance/follicular lesion of undetermined significance, follicular neoplasm (FN), suspicious for malignancy and malignant [32].

In the analysis, only malignancy diagnosed in the index nodule subjected to FNAC was counted as thyroid cancer for an individual subject. Incidentally discovered occult malignancy was not counted as malignant since the focus of the study was to correlate cytology and histology. Benign histopathology included follicular adenoma and hurthle cell adenoma.

Thyroidectomy indications were determined by the treating physicians in all cases and may have been in response to FNAC results or for other less apparent reasons. It is not an institutional policy to operate on all nodules over 4 cm, regardless of cytology results. The reasons for surgery were tabulated and compared by 4-cm size category in patients with benign or nondiagnostic FNAC, using the following categories: large nodule size, interval growth or suspicious US features, compressive symptoms or appearance, cosmetic reasons or patient preference.

Patients were divided into two groups according to the size of the index nodule on ultrasound, ≥4 cm or under 4 cm. FNAC results were compared with final surgical histopathology, stratified by 4 cm nodule size cut-off at the time of the FNAC.

Statistical analysis

Statistical analysis was performed using JMP Pro v. 13 software (SAS Institute, Cary, NC). Continuous data were reported as the mean ± standard deviation and categorical data as count and proportions.

Chi square or Fisher exact test was used to compare categorical data, including malignancy rates across cytologic categories by index nodule maximum diameter size < or ≥4 cm category. All tests were two-sided. A p value of less than or equal to 0.05 was considered significant.

Results

During the study period, 590 patients underwent thyroidectomy in which 490 patients with preoperative FNAC were identified for analysis, after exclusion of 100. The study population had an average (SD) age of 49 ± 15 years at the time of operation. There were 391 (80%) females and 99 (20%) males. Totally, 45% (219/490) of the index nodules proved malignant on final histopathology, and 55% benign (271/490) (Table 1).

The demographics and results distribution of the whole-study population and the subpopulations defined by index nodule size less than or greater than 4 cm cut-off are further shown in Table 1. The index nodule size was <4 cm in 353 (72%) and ≥4 cm in 137 (28%) of the thyroidectomy patients. The ≥4-cm population had a significantly higher percentage of males and a higher average BMI than the <4-cm population. Significantly different distribution of both preoperative cytologic and final histopathologic diagnosis was found in the index nodules <4 cm compared with ≥4 cm ($p < 0.0001$). There was a higher benign, nondiagnostic and lower suspicious for malignancy cytology rate in the ≥4 vs. <4 cm nodule group. Malignancy was present in 53% (188/353) of the index nodules <4 cm and in 23% (31/137) of the index nodules ≥4 cm ($p < 0.0001$).

A higher percentage of benign nodules ≥4 cm (77/137, 56%) than <4 cm (68/353, 19%) nodules were removed. There was no difference in the distribution of histopathologic diagnosis by 4-cm size cut off group within each cytologic subcategory analyzed (Table 2). In the benign cytology group, eight false-negative results were identified based on final histopathology. Of these, 37.5% (3/8) proved to be follicular thyroid carcinoma and 25% (2/8) were follicular variant papillary thyroid carcinoma. The false-negative for malignancy rate did not differ between the <4 and ≥4 cm benign cytology groups (5.9 vs. 5.2%, $p = 1.0000$).

Subjective symptoms and/or imaged mass effect was the most common reason for surgery and was more frequent with nodules ≥4 cm (67%, 52/77) than in 43% (29/68) <4 cm, $p = 0.0042$. Other reasons for surgery were not different between the 4 cm size cut-off groups, including imaging demonstrated growth, patient anxiety, repeat FNA

Table 1 Comparison of characteristics according to index nodule size

	Whole population	<4 cm	≥4 cm	*p*
No. of total population	490	353 (72%)	137 (28%)	
Age, mean (SEM), years	49	48 (0.8)	51 (1.3)	0.0953
Sex (male), *n* (%)	99 (20%)	55 (16%)	44 (32%)	<0.001
BMI (kg/m^2) (SEM)	29.3	28.7 (0.4)	30.9 (0.7)	0.0025
Cytology result				
Nondiagnostic	16	8 (2.3%)	8 (5.8%)	<0.0001
Benign	145	68 (19%)	77 (56%)	
Atypia of undetermined significance	101	78 (22 %)	23 (17%)	
Follicular neoplasm	62	50 (14%)	12 (8.8%)	
Suspicious	59	54 (15%)	5 (3.7%)	
Malignant	107	95 (27%)	12 (8.8%)	
Pathology result				
Benign	271 (55%)	165 (47%)	106 (77%)	<0.0001
Malignant	219 (45%)	188 (53%)	31 (23%)	

Nodule size >4 cm was associated with male sex, higher BMI, different FNAC distribution, and lower malignancy rate on final histopathology. For sex, cytology, and pathology result analysis, the *p* statistic represents chi square Fisher's exact test (two-tailed) comparing subcategory distribution per <4 vs. ≥4 cm nodule size.

BMI body mass index

Table 2 Distribution of surgical histopathology by cytology and size cut-off groups

Cytologic category	Surgical histopathology *n* (% of cytologic size category)				
	Benign		Malignant		
	Benign < 4 cm	Benign ≥ 4 cm	<4 cm	≥4 cm	*p*
Benign $n = 145$	64/68 (94%)	73/77 (95%)	4/68 (5.9%)	4/77 (5.2%)	1.000
AUS and FN $n = 163$	87/128 (68%)	25/35 (71%)	41/128 (32%)	10/35 (29%)	0.8376
Suspicious for malignancy $n = 59$	6/54 (11%)	0/5 (0%)	48/54 (89%)	5/5 (100%)	1.000

Histopathology distribution was not different by 4 cm size cut-off across within a given cytology category. The false-negative for malignancy rate did not differ between the <4 and ≥4 cm benign cytology groups (5.9% vs. 5.2%, $p = 1.0000$).

The *p* statistic represents chi square Fisher's exact test (two-tailed) comparing malignancy rate by size <4 vs. ≥4 cm across the cytologic category row.

AUS atypia of undetermined significance, *FN* follicular neoplasm

with different cytologic diagnosis, other changes in sonographic features (Table 3).

Discussion

In this retrospective surgical series of 490 consecutive thyroidectomies over a 5-year period comparing index nodules smaller or larger than 4 cm in size, a higher (53 vs. 22%) surgical prevalence of thyroid cancer in the <4-cm group, and a comparable benign cytology false-negative rate (5.9 vs. 5.2%) was found. More nodules ≥4 cm were benign than malignant, compared to the distribution in the <4 cm group.

The 22% malignancy rate prevalence for the over 4 cm nodule group reported here is similar to that reported by

Table 3 Reasons for surgery in the benign cytology index nodule group

	<4 cm ($n = 68$)	≥4 cm ($n = 77$)	*p*
Subjective symptoms or mass effect of goiter	29/68 (43%)	52/77 (67%)	0.0042
Other	19/68 (28%)	13/77 (17%)	0.1595
Growth by US	13/68 (19%)	14/77 (18%)	1.0000
Second FNAC with other diagnosis	10/68 (15%)	4/77 (5.2%)	0.0882
Anxiety	9/68 (13%)	4/77 (5.2%)	0.1436
US change (new nodules, change in nonindex nodule)	4/68 (5.9%)	1/77 (1.3%)	0.1865
Cosmetic	0	0	

The "other" group includes asymptomatic nodules that were removed due to size alone

others [6, 8, 9, 15, 18, 19, 28, 30, 31]. Rates as high as 28.6–58.5% [22, 24, 29, 33] and as low as 2.1–17.3% have also been reported [11, 13, 14, 16, 17, 20, 21, 23, 25, 26, 34]. Likewise, the rate of malignancy in the nodules ≥4 cm was lower than that of smaller nodules. This pattern of higher malignancy rate in smaller vs. ≥4 cm nodules was also reported in some surgical series [7, 11, 13, 20, 22, 28, 29], and a recent meta-analysis [15], while others showed similar [25, 33] or lower rates [8, 9, 30].

A similar benign cytology false-negative rate was demonstrated in nodules <4 cm (5.9%) compared with ≥4 cm (5.2%). Several surgical series reported false-negative range between 3.6 and 7.7% for nodules >4 cm [8, 11, 13, 16, 19, 20, 31, 33], a range similar to our study. Others reported higher false-negative rates (9.7–20%) [5, 6, 9, 10, 12, 14, 18, 22, 24]. Five studies reported rates <2.1% [21, 23, 25–27]. Five surgical series including <4 cm nodule control groups, like our study, did not find significant differences in the false-negative rates of benign cytology by 4 cm nodule size cut-off [11, 12, 20, 22, 25]. A meta-analysis, not including the present study, concluded higher false-negative rate of 6.7% for nodules >4 cm and 4.5% for nodules <4 cm [15].

Five of our eight benign cytology false-negative for malignancy index nodules harbored either follicular carcinoma or follicular variant papillary thyroid carcinoma, tumor types which lack features essential for the cytologic diagnosis of malignancy [32]. This pattern has also been observed by Wharry and Rosario in their >4 cm benign cytology false-negative groups [18, 19].

This study has some limitations. First, it is a single center retrospective surgical series which included analysis of only those index nodules selected for surgery, not all nodules. Factors beyond what we analyzed may have gone into the decision-making for surgery. This is a limitation for all such studies. Second, we cannot exclude selection bias to send larger benign nodules to surgery, increasing the benign denominator for this group. Specifically, 56% (76/145) of operated nodules with benign cytology were ≥4 cm, whereas only 22% (69/145) were under 4 cm. In addition, the ≥4 cm group comprised 28% of the study, a higher rate than observed in reports of larger nodule populations [23]. This bias would tend to reduce the false-negative for malignancy rate unless the rate of malignancy in nodules >4 cm is actually higher than those <4 cm. Still, we believe we have a comparable population to those previously reported.

The study findings add to the number of reports showing that FNAC results in nodules ≥4 cm can be used to guide the management of asymptomatic patients with thyroid nodules just as they do for smaller nodule size. The rate of malignancy in nodules and the benign cytology false-negative rate for malignancy in ≥4 cm nodules is not higher than that of smaller nodules. The decision for surgery in large nodules should be individualized to the patient and the size of nodule should not be the only criteria for such recommendation.

Funding This study was funded by the National Institutes of Health's National Center for Advancing Translational Sciences, grant UL1TR002494. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health's National Center for Advancing Translational Sciences. Muhammed Kizilgul was supported by a grant from The Scientific and Technical Research Council of Turkey (TUBITAK).

Compliance with ethical standards

Conflict of interest The authors delare that they have no conflict of interest.

Ethical approval All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards. This study was approved by the University of Minnesota Institutional Review Board.

References

1. D.S. Dean, H. Gharib, Epidemiology of thyroid nodules. Best Pract. Res Clin. Endocrinol. Metab. **22**(6), 901–911 (2008)
2. S. Guth, U. Theune, J. Aberle, A. Galach, C.M. Bamberger, Very high prevalence of thyroid nodules detected by high frequency (13 MHz) ultrasound examination. Eur. J. Clin. Invest. **39**(8), 699–706 (2009)
3. B.R. Haugen, E.K. Alexander, K.C. Bible 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 **26**(1), 1–133 (2016)
4. H. Gharib, E. Papini, R. Paschke et al. American Association of Clinical Endocrinologists, Associazione Medici Endocrinologi, and European Thyroid Association Medical Guidelines for Clinical Practice for the Diagnosis and Management of Thyroid Nodules. J. Endocrinol. Invest. **33**(5 Suppl), 1–50 (2010)
5. J.F. Carrillo, M. Frias-Mendivil, F.J. Ochoa-Carrillo, M. Ibarra, Accuracy of fine-needle aspiration biopsy of the thyroid combined with an evaluation of clinical and radiologic factors. Otolaryngol. Head Neck Surg. **122**(6), 917–921 (2000)
6. K.L. McCoy, N. Jabbour, J.B. Ogilvie, N.P. Ohori, S.E. Carty, J. H. Yim, 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 (2007). discussion 844.e1–e3
7. C.R. McHenry, E.S. Huh, R.N. Machekano, Is nodule size an independent predictor of thyroid malignancy? Surgery **144**(6), 1062–1068 (2008)
8. B. Kuru, N.E. Gulcelik, M.A. Gulcelik, H. Dincer, The false-negative rate of fine-needle aspiration cytology for diagnosing

thyroid carcinoma in thyroid nodules. Langenbeck's. Arch. Surg. **395**(2), 127–132 (2010)
9. D.H. Koo, K. Song, H. Kwon et al. Does tumor size influence the diagnostic accuracy of ultrasound-guided fine-needle aspiration cytology for thyroid nodules? Int. J. Endocrinol. **2016**, 1–6 (2016)
10. W.H. Giles, R.A. Maclellan, A.A. Gawande et al. False negative cytology in large thyroid nodules. Ann. Surg. Oncol. **22**(1), 152–157 (2015)
11. G. Godazandeh, Z. Kashi, S. Zargarnataj, M. Fazli, R. Ebadi, E.H. Kerdabadi, Evaluation the relationship between thyroid nodule size with malignancy and accuracy of fine needle aspiration biopsy (FNAB). Acta Inf. Med. **24**(5), 347–351 (2016)
12. S. Khalife, S. Bouhabel, V.I. Forest et al. The McGill Thyroid Nodule Score's (MTNS +) role in the investigation of thyroid nodules with benign ultrasound guided fine needle aspiration biopsies: a retrospective review. J. Otolaryngol. **45**(1), 1–7 (2016)
13. N. Bestepe, D. Ozdemir, A.A. Tam et al. Malignancy risk and false-negative rate of fine needle aspiration cytology in thyroid nodules ≥ 4.0 cm. Surgery **160**(2), 405–412 (2016)
14. O. Bozbiyik, S. Ozturk, M. Unver, V. Erol, U. Bayol, C. Aydin, Reliability of fine needle aspiration biopsy in large thyroid nodules. Turk. J. Surg. **33**(1), 10–13 (2017)
15. N.A. Cipriani, M.G. White, P. Angelos, R.H. Grogan, Large cytologically benign thyroid nodules do not have high rates of malignancy or false-negative rates and clinical observation should be considered: a meta-analysis. Thyroid **28**(12), 1595–1608 (2018)
16. S.N. Pinchot, H. Al-Wagih, S. Schaefer, R. Sippel, H. Chen, Accuracy of fine-needle aspiration biopsy for predicting neoplasm or carcinoma in thyroid nodules 4 cm or larger. Arch. Surg. **144** (7), 649–655 (2009)
17. P.P. Parikh, B.J. Allan, J.I. Lew, Sex variability of fine-needle aspiration reliability in the diagnosis of malignancy in thyroid nodules ≥ 4 cm. Am. J. Surg. **206**(5), 778–782 (2013)
18. L.I. Wharry, K.L. McCoy, M.T. Stang et al. Thyroid nodules (≥4 cm): can ultrasound and cytology reliably exclude cancer? World J. Surg. **38**(3), 614–621 (2014)
19. P.W. Rosario, D.S. Salles, B. Bessa, S. Purisch, Low false-negative rate of cytology in thyroid nodules > or = 4 cm. Arq. Bras. Endocrinol. Metabol. **53**(9), 1143–1145 (2009)
20. M. Shrestha, B.A. Crothers, H.B. Burch, The impact of thyroid nodule size on the risk of malignancy and accuracy of fine-needle aspiration: a 10-year study from a single institution. Thyroid **22** (12), 1251–1256 (2012)
21. M.D. Raj, S. Grodski, S. Woodruff, M. Yeung, E. Paul, J.W. Serpell, Diagnostic lobectomy is not routinely required to exclude malignancy in thyroid nodules greater than four centimetres. ANZ J. Surg. **82**(1-2), 73–77 (2012)
22. M.B. Albuja-Cruz, M. Goldfarb, S.S. Gondek, B.J. Allan, J.I. Lew, Reliability of fine-needle aspiration for thyroid nodules greater than or equal to 4 cm. J. Surg. Res. **181**(1), 6–10 (2013)
23. S.C. Kamran, E. Marqusee, M.I. Kim et al. Thyroid nodule size and prediction of cancer. J. Clin. Endocrinol. Metab. **98**(2), 564–570 (2013)
24. J.H. Kim, N.K. Kim, Y.L. Oh, et al., The validity of ultrasonography-guided fine needle aspiration biopsy in thyroid nodules 4 cm or larger depends on ultrasonography characteristics. Endocrinol. Metab. **29**, 545–552 (2014).
25. H. Shi, I. Bobanga, C.R. McHenry, Are large thyroid nodules classified as benign on fine needle aspiration more likely to harbor cancer? Am. J. Surg. **213**(3), 464–466 (2017)
26. U.C. Megwalu, Risk of malignancy in thyroid nodules 4 cm or larger. Endocrinol. Metab. **32**(1), 77–82 (2017)
27. A. Ozpek, C.C. Goret, C.S. Topal, Comparison of FNAC and histopathological results for 4 cm oR Larger Benign Thyroid Nodules. J. Coll. Physicians Surg. Pak. **28**(9), 703–706 (2018)
28. M.A. Hatem, S. Farheen, Thyroid nodule size as an indicator for surgery. Int Surg. J. **5**(7), 2401–2405 (2018)
29. R. Varshney, V.-I. Forest, F. Zawawi et al. Ultrasound-guided fine-needle aspiration of thyroid nodules: does size matter? Am. J. Otolaryngol. **35**(3), 373–376 (2014)
30. A.E. Uçar, S.M. Sarikaya, Ö. Parlak, A. Yalçin, Effect of nodule size on the reliability of fine-needle aspiration biopsy in thyroid nodules. Turk. J. Med. Sci. **44**(6), 1002–1009 (2014)
31. R. Kulstad, Do all thyroid nodules >4 cm need to be removed? an evaluation of thyroid fine-needle aspiration biopsy in large thyroid nodules. Endocr. Pract. **22**(7), 791–798 (2016)
32. E.S. Cibas, S.Z. Ali, N.C.I. Thyroid, FNA state of the science conference. The bethesda system for reporting thyroid cytopathology. Am. J. Clin. Pathol. **132**(5), 658–665 (2009)
33. L. Bohacek, M. Milas, J. Mitchell, A. Siperstein, E. Berber, Diagnostic accuracy of surgeon-performed ultrasound-guided fine-needle aspiration of thyroid nodules. Ann. Surg. Oncol. **19**(1), 45–51 (2012)
34. M.C. Frates, C.B. Benson, P.M. Doubilet et al. Prevalence and distribution of carcinoma in patients with solitary and multiple thyroid nodules on sonography. J. Clin. Endocrinol. Metab. **91**(9), 3411–3417 (2006)