



Risk of Malignancy in Thyroid Nodules 4 cm or Larger

Uchechukwu C. Megwalu^{1,2,3}

¹Department of Otolaryngology, Head and Neck Surgery, Stanford University School of Medicine, Stanford, CA; ²Department of Otolaryngology, Head and Neck Surgery, Icahn School of Medicine at Mount Sinai, New York, NY; ³Department of Otolaryngology, Queens Hospital Center, Jamaica, NY, USA

Background: Several authors have questioned the accuracy of fine-needle aspiration cytology (FNAC) in large nodules. Some surgeons recommend thyroidectomy for nodules ≥ 4 cm even in the setting of benign FNAC, due to increased risk of malignancy and increased false negative rates in large thyroid nodules. The goal of our study was to evaluate if thyroid nodule size is associated with risk of malignancy, and to evaluate the false negative rate of FNAC for thyroid nodules ≥ 4 cm in our patient population.

Methods: This is a retrospective study of 85 patients with 101 thyroid nodules, who underwent thyroidectomy for thyroid nodules measuring ≥ 4 cm.

Results: The overall risk of malignancy in nodules ≥ 4 cm was 9.9%. Nodule size was not associated with risk of malignancy (odds ratio, 1.02) after adjusting for nodule consistency, age, and sex ($P=0.6$). The false negative rate for FNAC was 0%.

Conclusion: Nodule size was not associated with risk of malignancy in nodules ≥ 4 cm in our patient population. FNAC had a false negative rate of 0. Patients with thyroid nodules ≥ 4 cm and benign cytology should not automatically undergo thyroidectomy.

Keywords: Thyroid nodule; Thyroid neoplasms; Biopsy, fine-needle; Diagnostic accuracy; Large nodules

INTRODUCTION

Thyroid nodules are very common, with an estimated prevalence of 1% to 5% for palpable nodules [1-3]. This prevalence is higher when considering nodules detected on imaging [4]. The incidence of thyroid malignancy is rising [5]. Thyroid cancer occurs in 5% to 15% of thyroid nodules [1,6].

Fine-needle aspiration biopsy (FNAB) is the most reliable diagnostic tool for detecting malignancy in thyroid nodules, with sensitivity and specificity $>90\%$ [1,7-9]. The false negative rate for thyroid FNAB is low. The expected incidence of malignancy in cytologically benign nodules is 0% to 3% according to the Bethesda system for reporting thyroid cytopathology [10,11]. However, several authors have questioned the accuracy of fine-

needle aspiration cytology (FNAC) in large nodules. Some surgeons recommend thyroidectomy for nodules ≥ 4 cm even in the setting of benign FNAC, due to increased risk of malignancy and increased false negative rates in large thyroid nodules [12-15]. Even more aggressive surgeons use a threshold of 3 cm [16]. This increase in false negative rates with large thyroid nodules is thought to be due to sampling error. On the other hand, several authors have demonstrated low false negative rates in large thyroid nodules, similar to those in small thyroid nodules [17-21].

It is our standard practice to offer thyroidectomy to patients with thyroid nodules measuring 4 cm or larger. The goals of our study were to evaluate the false negative rate of FNAC for thyroid nodules ≥ 4 cm in our patient population, and to evaluate if

Received: 12 October 2016, Revised: 17 December 2016,

Accepted: 26 December 2016

Corresponding author: Uchechukwu C. Megwalu

Department of Otolaryngology, Head and Neck Surgery, Stanford University School of Medicine, 801 Welch Road, Stanford, CA 94305, USA

Tel: +1-650-723-6880, Fax: +1-650-725-8502, E-mail: megwaluu@yahoo.com

Copyright © 2017 Korean Endocrine Society

This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/4.0/>) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

thyroid nodule size is associated with risk of malignancy in thyroid nodules ≥ 4 cm.

METHODS

This study was approved by the Icahn School of Medicine at Mount Sinai Institutional Review Board. Data were extracted from patients' charts. The study cohort included patients from our Otolaryngology, Head and Neck Surgery Department, who underwent thyroidectomy from August 2011 through September 2015 for thyroid nodules measuring ≥ 4 cm. It is our standard practice to offer thyroidectomy to patients with thyroid nodules measuring ≥ 4 cm. Patients were included if they had preoperative thyroid ultrasound demonstrating a nodule measuring ≥ 4 cm in at least one lobe. There was no age restriction for inclusion in the study. Data collection was performed by the author and an assistant. Patients without preoperative ultrasounds were excluded. Information on nodule size, texture (solid, complex, or cystic), and laterality was obtained from ultrasound reports. Nodule texture was determined at the discretion of the reading radiologist. Nodule texture was described as solid, cystic, or complex (containing both solid and cystic components). Nodule size was recorded as the largest of the three dimensions: length, width, and depth.

Information was obtained on whether FNAB was performed. Thyroid FNAB is performed by endocrinologists under ultrasound guidance in our institution. Cytology results were recorded for the nodules in question according to the Bethesda system: nondiagnostic (category 1), benign (category 2), atypia of undetermined significance (AUS)/follicular lesion of undetermined significance (FLUS) (category 3), suspicious for follicular neoplasm/follicular neoplasm (category 4), suspicious for malignant disease (category 5), and malignant (category 6) [10]. In situations where discrepant cytology results were obtained from the same nodule, final cytology was taken as the more high-risk result.

Thyroidectomy was performed by the senior author (U.C.M.) in all patients. All specimen were submitted to pathology for permanent section with sutures identifying laterality. Information on final pathological diagnosis was obtained from the surgical pathology reports. Final diagnoses were classified as benign, papillary thyroid carcinoma, follicular thyroid carcinoma, medullary thyroid carcinoma, and other malignancy. Papillary thyroid cancer was further sub-classified as classical, follicular variant, or tall cell variant. Final results were determined by comparing nodule size and laterality from pathology reports to

the thyroid ultrasound and FNAC reports. Final pathology was classified as benign if there was incidental microcarcinoma identified outside of the nodule of interest, as long as the nodule of interest was benign. Data were also collected on demographic and clinical factors such as age and sex.

IBM SPSS version 20 (IBM Co., Armonk, NY, USA) was used for statistical analysis. Pearson chi-square was used to evaluate the impact of Bethesda category, sex, and age on the risk of malignancy. Independent sample *t* test was used to compare means for continuous variables (age and nodule size). Logistic regression was used for multivariable analysis. Nodule size, nodule consistency, age, and sex were entered *a priori* into the model. $P < 0.05$ was considered statistically significant. NCSS statistical software (Number Cruncher Statistical Systems, Kaysville, UT, USA) was used to analyze diagnostic characteristics of FNAC. Nodules with nondiagnostic cytology were excluded when calculating sensitivity, specificity, and false negative rates. Nodules with benign FNAC were classified as negative. All other sufficient FNA results were classified as positive. Post-test odds were calculated by multiplying the likelihood ratio (LR) with the pre-test odds of malignancy (the odds of malignancy in a thyroid nodule ≥ 4 cm) [22]. LR is the likelihood of a particular FNAC result in a nodule with malignancy compared with the likelihood of the same result in a nodule without malignancy.

RESULTS

Patient and nodule characteristics

Chart review identified a total of 85 patients with 101 thyroid nodules meeting the inclusion criteria. Seventy-two patients (84.7%) were female. The mean age was 55 (95% confidence interval [CI], 52.3 to 57.6). Patient ages ranged from 23 to 76, with a median age of 55. Fifty-four patients underwent total thyroidectomy, 27 underwent hemithyroidectomy, and four underwent completion thyroidectomy. Fifteen patients (17.6%) had thyroid cancer on final pathology (including incidental microcarcinomas).

Forty-nine nodules (48.5%) were located in the left lobe, 51 (50.5%) in the right lobe, and one in the isthmus. The mean nodule size was 53.6 mm (95% CI, 51.6 to 55.6). The median nodule size was 52 mm (range, 40 to 90). Information on nodule consistency was available for 94 nodules. Of these, 40 nodules (42.6%) were solid, 53 (56.4%) were complex, and one was cystic. FNAB was performed on 90 patients (89.1%). The distribution of the FNAC was 7.8% nondiagnostic (category 1),

68.9% benign (category 2), 10% AUS/FLUS (category 3), 10% suspicious for follicular neoplasm/follicular neoplasm (category 4), 2.2% suspicious for malignant disease (category 5), and 1.1% malignant (category 6).

On final pathology, 10 nodules (9.9%) were malignant. Of the malignant nodules, the most common histologic type was papillary thyroid cancer, accounting for 70% of malignancies. There were seven cases of papillary thyroid cancer of the following subtypes: two classical, and five follicular variant. There were two cases of follicular thyroid cancer, and one case of anaplastic thyroid cancer. The FNA cytology results for these malignant nodules are shown in Table 1. Fifteen incidental malignancies were identified on final pathology: eight in the same lobe as the nodule of interest, and seven in the contralateral lobe.

Risk of malignancy

After excluding the cystic nodule, the overall risk of malignancy in nodules ≥ 4 cm was 10%. When including incidental malignancy within the same lobe, the risk of malignancy was 16%. The risk of malignancy for patients who did not undergo FNAB was 9.1%. The mean nodule size was 54.8 mm for malignant nodules and 53.5 mm for benign nodules. The difference was not statistically significant (mean difference, 1.3 mm; 95% CI, -5.5 to 8.1 ; $P=0.7$). The mean age was 53 years for patients

with malignant nodules and 55.2 years for patients with benign nodules. The difference was not statistically significant (mean difference, -2.2 years; 95% CI, -10.2 to 5.8 ; $P=0.6$). The results of univariable analysis of the impact of Bethesda class, nodule consistency, and sex on risk of malignancy are shown in Table 2. The risk of malignancy was 0% for classes 1 and 2, and increased with increasing Bethesda diagnostic category thereafter ($P<0.001$). Nodule consistency and sex had no impact on risk of malignancy. The results of multivariable analysis are shown in Table 3. Nodule size was not associated with risk of malignancy (odds ratio [OR], 1.02) after adjusting for nodule consistency, age, and sex.

Diagnostic characteristics of FNA cytology for nodules ≥ 4 cm

FNAC had a sensitivity of 100% (95% CI, 70% to 100%), and a specificity of 84% (95% CI, 74% to 90%) in our cohort. The positive predictive value was 43% (95% CI, 25% to 64%), and the negative predictive value (NPV) was 100% (95% CI, 94%

Table 1. Bethesda Class for Tumors Found to Be Malignant on Final Pathology

Final histology	FNA cytology (Bethesda class)					
	Class 1	Class 2	Class 3	Class 4	Class 5	Class 6
Overall	0	0	1	5	2	1
Classical papillary	0	0	1	0	0	1
Follicular variant papillary	0	0	0	3	2	0
Follicular	0	0	0	1	0	0
Anaplastic	0	0	0	1	0	0

FNA, fine-needle aspiration.

Table 2. Impact of Bethesda Class, Nodule Consistency, and Sex on Risk of Malignancy (101 Nodules from 85 Patients)

Variable	Number	Risk of malignancy, %	P value
Bethesda class			<0.001
Class 1	7	0	
Class 2	62	0	
Class 3	9	11.1	
Class 4	9	55.6	
Class 5	2	100	
Class 6	1	100	
Consistency			0.175
Solid	40	17.5	
Complex	53	5.7	
Cystic	1	0	
Sex			0.641
Male	86	13.3	
Female	15	9.3	

Table 3. Impact of Nodule Size on Risk of Malignancy (Logistic Regression)

Variable	Coefficient (SE)	Wald χ^2	Odds ratio	95% CI	P value
Nodule size	0.02 (0.03)	0.28	1.02	0.96–1.08	0.597
Solid nodule	1.41 (0.74)	3.59	4.09	0.95–17.52	0.058
Age	-0.03 (0.03)	0.83	0.97	0.92–1.03	0.363
Female sex	-0.22 (0.93)	0.05	0.81	0.13–4.98	0.816

SE, standard error; CI, confidence interval.

to 100%). The false negative rate was 0%. The LR for benign FNAC was 0. For a nodule ≥ 4 cm with benign FNA cytology, the post-test odds of malignancy was 0.

DISCUSSION

The results of our study show that nodule size was not associated with risk of malignancy in nodules ≥ 4 cm in our patient population. Furthermore, FNA cytology had a false negative rate of 0%, a sensitivity of 100%, and a NPV of 100%. This suggests that the diagnostic accuracy of FNA cytology is not limited by large nodule size.

Our findings are similar to those of several other studies. Raj et al. [17] evaluated the prevalence of thyroid cancer in 223 patients who underwent thyroidectomy for thyroid nodules 4 cm or larger. Sixteen patients (7.2%) had malignancy on final pathology. There was no association between nodule size and malignancy rate. Similar to our study, FNAC had a low false negative rate of 0.85%. Mehanna et al. [18] evaluated the impact of large nodule size and follicular variant of papillary thyroid cancer on FNA false negative rates. The study included 569 thyroid nodules, 262 of which were subjected to thyroidectomy. When all nodules (surgical and nonsurgical) were included in the analysis, nodules ≥ 3 cm were significantly more likely to be malignant. However, among nodules that were subjected to thyroidectomy, there was no difference in size between benign and malignant nodules. The false negative rates were 10.9% for nodules ≥ 3 cm, and 6.1% for nodules < 3 cm, but this difference was not statistically significant. Of note, the false negative rate in this study was higher than the expected 0% to 3% according to the Bethesda system for reporting thyroid cytopathology [10,11]. Another retrospective study of 1,068 patients, who underwent thyroidectomy, investigated the impact of nodule size on accuracy of FNAC [19]. This included 75 patients with nodules ≥ 4 cm. The overall false negative rate was 13%. The false negative rate for nodules ≥ 4 cm was 15%. There was no significant difference in false negative rates between nodules ≥ 4 cm and nodules < 4 cm. Similar to the study by Mehanna et al. [18], the false negative rate in this study was higher than the expected 0% to 3%.

Shrestha et al. [20] performed a retrospective study of 540 patients with 695 nodules, who underwent FNAB and subsequent thyroid surgery. There was no set size criteria for surgical referral. The overall malignancy rate was 18.6%, and did not vary significantly based on size. The overall false negative rate was 7%, and did not differ significantly according to size. The au-

thors noted that the accuracy of FNAC increased with increasing nodule size. Similarly, Magister et al. [21] performed a retrospective study of 297 patients with 326 thyroid nodules, who underwent FNAB and subsequent thyroidectomy. The overall rate of malignancy on final pathology was 43.9%. The false negative rate was 6% for all nodules combined, and 3.8% for nodules 3 cm or greater. They found that smaller nodules had higher probability of malignancy. The finding of increased risk of malignancy and decreased diagnostic accuracy in smaller nodules in these studies could be due to selection bias. The studies included nodules of all sizes, with no set size criteria for recommending surgery. Consequently, the smaller nodules included in the study may have undergone thyroidectomy due to the presence of unmeasured risk factors for malignancy.

On the other hand, several studies have demonstrated increased risk of malignancy and increased false negative rates in large thyroid nodules. In a prospective study of 571 patients, who underwent thyroidectomy, Kuru et al. [12] found that nodule size ≥ 4 cm was associated with increased risk of malignancy compared with nodule size < 4 cm. However, no difference in risk of malignancy was noted when a cut-off of 3 cm was used. Giles et al. [16] evaluated the accuracy of benign FNAC in patients who underwent surgery for nodules 3 cm or larger. The patient cohort included patients with thyroid nodules measuring 3 cm or larger, with benign FNAC. Additional nodules measuring < 3 cm within the specimen, with benign cytology, were identified for comparison. Two-hundred and forty nodules measuring 3 cm or larger were compared with 83 nodules measuring < 3 cm. The false negative rates were significantly higher for nodules measuring 3 cm or larger (4.8% for nodules < 3 cm, 12.8% for nodules 3 to 3.9 cm, and 11% for nodules 4 cm or larger). The authors recommended that thyroidectomy should be considered for nodule ≥ 3 cm regardless of FNA cytology. Kim et al. [13] performed a retrospective study of 263 patients, who underwent thyroidectomy for nodules ≥ 4 cm. A significant proportion of the nodules (58.6%) were malignant on final histopathology. They noted a high false negative rate of 11.9%. Another study of 155 patients, who underwent thyroidectomy for thyroid nodules 4 cm or larger, found that 7.7% of nodules reported as benign on FNAB were found to be malignant on final pathology [14]. Contrary to our study, the risk of malignancy in their study was 27.3% among patients with nondiagnostic FNA cytology. Similar to our study, no differences in nodule size were noted between benign and malignant nodules. The authors recommended diagnostic thyroid lobectomy for nodules 4 cm or larger due to the high false negative rate of FNAC. A re-

cent prospective study of 361 patients with 382 nodules ≥ 4 cm identified clinically significant thyroid cancer (defined by the authors as ≥ 1 cm, or associated with lymph node metastasis) in 22% of the nodules [15]. The authors also noted a high false negative rate of 10.4% for FNAC.

In an attempt to resolve the inconsistencies in the literature, Shin et al. [23] performed a systematic review to determine if thyroid nodule size >3 to 4 cm was associated with increased prevalence of thyroid cancer, and worse diagnostic accuracy. The review included 15 studies, including 13,180 participants. The majority of the studies showed higher prevalence of malignancy in larger nodules compared with smaller nodules. All seven studies that reported on diagnostic accuracy showed higher false negative rates, and lower sensitivity for larger nodules. Few studies reported on post-test odds or probabilities of malignancy. All five articles that reported on post-test odds or probability of malignancy showed that the post-test probability of malignancy in cytologically benign nodules was higher for nodules larger than 3 to 4 cm compared with smaller nodules. Benign FNAC had a post-test probability of malignancy of 0.8% to 3.6%, and nondiagnostic FNAC had a post-test probability of 16.7% to 27.3%. In contrast, our study showed a post-test probability of 0% for benign FNAC in patients with thyroid ≥ 4 cm.

The discrepancy in the risk of malignancy and accuracy of FNAC reported in the literature for large thyroid nodules may be due to variability in sampling and interpretation of cytopathologic findings. The accuracy of FNAC depends on adequate sampling and correct interpretation. Furthermore, the post-test probability of malignancy is significantly influenced by the overall prevalence of malignancy in the population, which varies among institutions. These factors need to be accounted for when counseling patients about the need for thyroidectomy for large nodules.

We limited our analysis to thyroid nodules measuring ≥ 4 cm. This reduces selection bias because it was our standard practice to offer thyroidectomy for this subset of patients regardless of FNAC results. Including patients with smaller nodules would increase the risk of selection bias because these patients are more likely to have other risk factors influencing the decision to undergo thyroidectomy. We excluded FNA cytology in our logistic regression model evaluating the impact of nodule and patient factors on risk of malignancy because FNA cytology had a false negative rate of 0% in our study. Consequently, an OR could not be calculated for the logistic regression due to “complete separation” of patients with malignancy from those with benign pathology based on FNA cytology [24].

The main strength of our study lies in the systematic collection and analysis of data. Our study is primarily limited by its retrospective nature. Also, this represents our experience at a single institution, which may not be generalizable to other institutions. Finally, in spite of standardized reporting of cytopathology results, it is impossible to completely eliminate variability in interpretation and reporting of cytopathology results.

In conclusion, nodule size was not associated with risk of malignancy in nodules ≥ 4 cm in our patient population. Furthermore, FNAC had a false negative rate of 0%, a NPV of 100%, and a post-test probability of malignancy of 0% for nodules ≥ 4 cm with benign FNAC at our institution. These findings suggest that patients with thyroid nodules ≥ 4 cm and benign FNAC should not automatically undergo thyroidectomy. FNAC as well as the presence of symptoms are important factors to consider when recommending thyroidectomy for these patients. Selection of the appropriated treatment option (thyroidectomy vs. observation) should involve shared decision making between the patient and the healthcare provider. The overall accuracy of FNAC at the particular facility, and the prevalence of malignancy in the population need to be considered.

CONFLICTS OF INTEREST

No potential conflict of interest relevant to this article was reported.

ACKNOWLEDGMENTS

The author wishes to acknowledge Jason Herel (Mount Sinai Services Department of Otolaryngology, Queens Hospital Center) for his assistance in data collection.

REFERENCES

1. American Thyroid Association (ATA) Guidelines Taskforce on Thyroid Nodules and Differentiated Thyroid Cancer, Cooper DS, Doherty GM, Haugen BR, Kloos RT, Lee SL, et al. Revised American Thyroid Association management guidelines for patients with thyroid nodules and differentiated thyroid cancer. *Thyroid* 2009;19:1167-214.
2. Tunbridge WM, Evered DC, Hall R, Appleton D, Brewis M, Clark F, et al. The spectrum of thyroid disease in a community: the Wickham survey. *Clin Endocrinol (Oxf)* 1977;7:481-93.
3. Vander JB, Gaston EA, Dawber TR. The significance of nontoxic thyroid nodules. Final report of a 15-year study of

- the incidence of thyroid malignancy. *Ann Intern Med* 1968; 69:537-40.
4. Tan GH, Gharib H. Thyroid incidentalomas: management approaches to nonpalpable nodules discovered incidentally on thyroid imaging. *Ann Intern Med* 1997;126:226-31.
 5. Hughes DT, Haymart MR, Miller BS, Gauger PG, Doherty GM. The most commonly occurring papillary thyroid cancer in the United States is now a microcarcinoma in a patient older than 45 years. *Thyroid* 2011;21:231-6.
 6. Hegedus L. Clinical practice. The thyroid nodule. *N Engl J Med* 2004;351:1764-71.
 7. Bouvet M, Feldman JI, Gill GN, Dillmann WH, Nahum AM, Russack V, et al. Surgical management of the thyroid nodule: patient selection based on the results of fine-needle aspiration cytology. *Laryngoscope* 1992;102(12 Pt 1):1353-6.
 8. Lansford CD, Teknos TN. Evaluation of the thyroid nodule. *Cancer Control* 2006;13:89-98.
 9. Amrikachi M, Ramzy I, Rubinfeld S, Wheeler TM. Accuracy of fine-needle aspiration of thyroid. *Arch Pathol Lab Med* 2001;125:484-8.
 10. Cibas ES, Ali SZ. The Bethesda system for reporting thyroid cytopathology. *Thyroid* 2009;19:1159-65.
 11. Jo VY, Stelow EB, Dustin SM, Hanley KZ. Malignancy risk for fine-needle aspiration of thyroid lesions according to the Bethesda system for reporting thyroid cytopathology. *Am J Clin Pathol* 2010;134:450-6.
 12. Kuru B, Gulcelik NE, Gulcelik MA, Dincer H. Predictive index for carcinoma of thyroid nodules and its integration with fine-needle aspiration cytology. *Head Neck* 2009;31:856-66.
 13. Kim JH, Kim NK, Oh YL, Kim HJ, Kim SY, Chung JH, et al. The validity of ultrasonography-guided fine needle aspiration biopsy in thyroid nodules 4 cm or larger depends on ultrasonography characteristics. *Endocrinol Metab (Seoul)* 2014;29:545-52.
 14. Pinchot SN, Al-Wagih H, Schaefer S, Sippel R, Chen H. Accuracy of fine-needle aspiration biopsy for predicting neoplasm or carcinoma in thyroid nodules 4 cm or larger. *Arch Surg* 2009;144:649-55.
 15. Wharry LI, McCoy KL, Stang MT, Armstrong MJ, LeBeau SO, Tublin ME, et al. Thyroid nodules (≥ 4 cm): can ultrasound and cytology reliably exclude cancer? *World J Surg* 2014;38:614-21.
 16. Giles WH, Maclellan RA, Gawande AA, Ruan DT, Alexander EK, Moore FD Jr, et al. False negative cytology in large thyroid nodules. *Ann Surg Oncol* 2015;22:152-7.
 17. Raj MD, Grodski S, Woodruff S, Yeung M, Paul E, Serpell JW. Diagnostic lobectomy is not routinely required to exclude malignancy in thyroid nodules greater than four centimetres. *ANZ J Surg* 2012;82:73-7.
 18. Mehanna R, Murphy M, McCarthy J, O'Leary G, Tuthill A, Murphy MS, et al. False negatives in thyroid cytology: impact of large nodule size and follicular variant of papillary carcinoma. *Laryngoscope* 2013;123:1305-9.
 19. Albuja-Cruz MB, Goldfarb M, Gondek SS, Allan BJ, Lew JI. Reliability of fine-needle aspiration for thyroid nodules greater than or equal to 4 cm. *J Surg Res* 2013;181:6-10.
 20. Shrestha M, Crothers BA, Burch HB. 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* 2012;22:1251-6.
 21. Magister MJ, Chaikhoutdinov I, Schaefer E, Williams N, Saunders B, Goldenberg D. Association of thyroid nodule size and Bethesda class with rate of malignant disease. *JAMA Otolaryngol Head Neck Surg* 2015;141:1089-95.
 22. Shin JJ, Stinnett S, Page J, Randolph GW. Evidence-based medicine in otolaryngology, part 3: everyday probabilities: diagnostic tests with binary results. *Otolaryngol Head Neck Surg* 2012;147:185-92.
 23. Shin JJ, Caragacianu D, Randolph GW. Impact of thyroid nodule size on prevalence and post-test probability of malignancy: a systematic review. *Laryngoscope* 2015;125:263-72.
 24. de Irala J, Navahas RF, del Castillo AS. Abnormally wide confidence intervals in logistic regression: interpretation of statistical program results. *Pan Am J Public Health* 1997; 2:268-71.