

The Malignancy Risk Assessment of Cytologically Indeterminate Thyroid Nodules Improves Markedly by Using a Predictive Model

Bülent Öcal^a Mehmet Hakan Korkmaz^b Demet Yilmazer^c
Tuğba Taşkın Türkmenoğlu^c Ömer Bayır^a Güleser Saylam^a
Emel Çadallı Tatar^a Sevilay Karahan^d Erman Çakal^e

^aDepartment of Otolaryngology, Ministry of Health Dışkapı Yıldırım Beyazıt Training and Research Hospital, Ankara, Turkey; ^bDepartment of Otolaryngology, Yıldırım Beyazıt University Medical School, Ankara, Turkey; ^cDepartment of Pathology, Ministry of Health Dışkapı Yıldırım Beyazıt Training and Research Hospital, Ankara, Turkey; ^dDepartment of Biostatistics, Faculty of Medicine, Hacettepe University, Ankara, Turkey; ^eDepartment of Endocrinology, Ministry of Health Dışkapı Yıldırım Beyazıt Training and Research Hospital, Ankara, Turkey

Keywords

Thyroid nodule · Malignancy risk · Indeterminate cytology · Nomogram · Thyroid ultrasonography

Abstract

Objectives: The majority of thyroid nodules are discovered incidentally, and the management may be a challenge if the fine needle aspiration specimen yields indeterminate findings. Our aim was to develop an individualized risk prediction model to provide an accurate estimate of cancer risk in patients with cytologically indeterminate thyroid nodules.

Materials and Methods: Clinical records, ultrasound images, and cytopathology reports of patients who underwent thyroidectomy were retrospectively reviewed. Logistic regression analysis was used to identify the predictive ability of each variable for malignancy, and a nomogram was built by integrating patients' age, multiplicity of nodules, cytology results, and suspicious ultrasound features, such as microcalcifications and irregular margins. **Results:** For the 233 indeterminate nodules according to the Bethesda System for Reporting Thyroid Cytopathology, the malignancy rates of the subgroups "atypia of undetermined significance," "suspi-

cious follicular neoplasia," and "suspicious for malignancy" were 44.3, 47.7, and 88.0%, respectively. It was found that the Bethesda category "suspicious for malignancy," microcalcifications, and irregular margins were independent risk factors for malignancy. The area under the receiver operating characteristics curve was 0.784, which suggested that the presented nomogram had considerable discriminative performance. **Conclusions:** The nomogram developed in our study accurately predicts the malignancy risk of thyroid nodules with indeterminate cytology by using clinical, cytological, and ultrasonographic features.

© 2018 European Thyroid Association
Published by S. Karger AG, Basel

Introduction

The estimated prevalence of thyroid nodules in an adult population is 4–7% by palpation and 10–41% by ultrasound (US) scanning [1, 2]. The overall malignancy risk, regardless of nodule size, is <10%, and even smaller nodules (<1 cm) may harbor cancer. Therefore, every effort should be made either clinically or cytologically to

predict malignancy in order to perform the appropriate surgery with the correct indication.

It is critical to improve the diagnostic accuracy and to minimize false-negative diagnoses of thyroid nodules bearing cancer. A combination of clinical, cytological, and ultrasonographic features is expected to serve as a more specific prediction tool. Recently, some predictive systems and mathematical methods have been developed, which used a combination of nodule size, age at diagnosis, biochemical variables, fine needle aspiration (FNA) cytology, US features, and molecular analysis [3–8]. The most recent American Thyroid Association recommendations emphasized that each institution should independently define the risk of malignancy in each of the indeterminate cytology categories of the Bethesda System for Reporting Thyroid Cytopathology to guide clinicians and surgeons in choosing an appropriate treatment modality [4].

Our aim was to predict the probability of malignancy for each case of indeterminate cytology at our institution. Secondly, we aimed to develop a nomogram by integrating important clinical parameters with US and FNA features into a single numerical estimate system to generate individualized predictions of cancer in the nodules. This method finally enables us to identify patients who may derive greater benefit from thyroidectomy.

Materials and Methods

A total of 233 patients having indeterminate nodules according to the Bethesda System underwent thyroidectomy in a single tertiary center between January 2009 and August 2015. The study protocol was approved by the institutional review board (20.10.2015-26/30). Patients with a history of previous thyroidectomy and with clinical findings suggestive of either central or lateral cervical lymph node metastasis were excluded. The data retrieved from an existing database included patient demographics, family history of thyroid cancer, preoperative US/US-guided FNA cytology findings, and final histopathology reports of thyroidectomy specimens.

Thyroid US examinations and FNA procedures were performed by endocrinologists who were experienced in thyroid US for at least 5 years. Since the current literature suggests that microcalcifications and irregular margins exhibit the highest specificities for the detection of thyroid cancer, these 2 features were used for statistical analysis [4]. Also, we compared the risk of malignancy between nodules classified as single or multiple.

FNA procedures were performed using a 27-gauge needle, a 10-mL plastic syringe, and a pistol handle (Comeco, Sweden). After the smearing procedure, most slides were air dried and stained with May-Grünwald-Giemsa stain, while the remaining slides were stained with Papanicolaou and hematoxylin and eosin stains after 95% ethanol fixation. The slides were interpreted by 2 cyto-

pathologists and were classified into 6 categories according to the Bethesda System. Thyroid nodules representing 3 indeterminate categories, including atypia of undetermined significance or follicular lesion of undetermined significance (AUS/FLUS), suspicious follicular neoplasm (sFN), and suspicious malignancy (sM), were analyzed in this study [9].

The final histopathology reports were correlated with preoperative FNA cytology, US findings, and patients' demographics to reveal whether these variables predicted malignancy. Those patients with multinodular disease, nodules >1 cm in size, and with malignancy-associated US features underwent biopsy. If more than 1 nodule had an indeterminate cytology result, the one with the highest risk of malignancy was selected and used for analysis. At our institution, the best management option for patients with thyroid nodules is discussed at multidisciplinary, weekly endocrine surgery board meetings on a case-by-case basis.

Statistical Analysis

Statistical analysis was carried out using the IBM SPSS for Windows, version 22.0, software package. Numeric and categorical variables were summed up with mean \pm standard deviation and with number of cases and percentage, respectively. The χ^2 test was used to assess the presence of significant differences in risk factors between malignant and benign cases. Multiple logistic regression analysis was performed to calculate the predictive value of variables. The performance of the nomogram was quantified based on discrimination and calibration. Discrimination (predictive accuracy) was measured via the area under the receiver operating characteristic (ROC) curve (AUC). The best cutoff point for the risk of malignancy was determined using the ROC curve. Diagnostic values of ROC curve analysis were given as sensitivity, specificity, positive predictive value, and negative predictive value. Calibration was assessed by a calibration curve which shows how closely the nomogram predicted risk to the actual risk. Statistical significance was set at $p < 0.05$.

Results

The presented cohort involved 233 patients who underwent thyroidectomy with indeterminate FNA cytology reports according to the Bethesda System. No patients reported a history of exposure to ionizing radiation in childhood or adolescence.

Cytologic subcategories were reported as AUS in 106 (45.5%), sFN in 44 (18.9%), and sM in 83 (35.6%) patients. The median age was 46 years, with 87.6% of patients being female. The median size of aspirated nodules was 16 mm (range, 4–81).

When US reports were reviewed retrospectively, 69 (24.9%) patients had malignancy-associated US features, including microcalcifications and/or irregular margins. Microcalcifications were found in 52 (22.3%) patients, whereas irregular margins were observed in 29 (12.4%) cases. Of all patients, 12 (5.1%) had more than 1 US feature associated with malignancy.

Table 1. Univariate analyses of malignant and benign nodules with clinicopathologic factors ($n = 233$)

Variables	Total, n	Benign nodule, n (%)	Malignant nodule, n (%)	p value
Age				
<45 years	106	30 (28.3)	76 (71.7)	0.001
≥45 years	127	62 (48.8)	65 (51.2)	
Indeterminate FNA cytology				
AUS	106	59 (55.7)	47 (44.3)	<0.001
sFN	44	23 (52.3)	21 (47.7)	
sM	83	10 (12)	73 (88)	
Microcalcification				
Absent	181	84 (46.4)	97 (53.6)	<0.001
Present	52	8 (15.4)	44 (84.6)	
Irregular margins				
Absent	204	90 (44.1)	114 (55.9)	<0.001
Present	29	2 (6.9)	27 (93.1)	
Solitary nodule				
Absent	181	79 (43.6)	102 (56.4)	0.015
Present	52	13 (25)	39 (75)	
Elastography scores				
≥3	54	12 (22.3)	42 (77.7)	0.019
<3	99	48 (48.5)	51 (51.5)	

FNA, fine needle aspiration; AUS, atypia of undetermined significance; sFN, suspicious for follicular neoplasm; sM, suspicious for malignancy.

On final histopathology, 141 (60.6%) patients had thyroid carcinoma, consisting of 132 (93.6%) with papillary cancer, 7 (4.9%) with follicular cancer, and 2 (1.4%) with undifferentiated cancer. Ninety-two patients had a diagnosis of a benign process.

Malignancy Rates of Variables

The risk of malignancy in the AUS group was 44.3% (47 of 106 patients). In the sFN group, 47.7% (21 of 44 patients) were diagnosed with malignancy on the final histopathology. Thyroid cancer was diagnosed in 88% (73 of 83 patients) of cases in the sM group ($p < 0.001$).

Forty-four of 52 (84.6%) nodules containing microcalcifications had thyroid cancer on final histopathology, while nodules without microcalcifications showed malignancy in 97 of 181 (53.6%) patients ($p < 0.001$). The incidence of malignancy was 93.1% (27 of 29 patients) for nodules with irregular margins and was 55.9% (114 of 204 patients) in those without irregular margins ($p < 0.001$).

Among 233 patients, 181 (77.7%) had a single nodule and 52 (22.3%) presented 2 or more nodules. The prevalence of thyroid cancer was higher in patients with a solitary thyroid nodule (39 of 52 patients, 75.0%) than in

those with multiple nodules (102 of 181 patients, 56.4%) ($p = 0.015$).

Younger age (<45 years) was found to be associated with an increased risk of malignancy in our cohort. Overall, 76 of 106 patients (71.7%) <45 years had thyroid malignancy, while 65 of 127 patients (51.2%) >45 years had thyroid malignancy ($p = 0.001$).

Elastographic examination was also available for only 153 suspicious nodules, 54 of which (35.2%) had ≥3 elastographic scores. The risk of malignancy was 77.7% in nodules with ≥3 elastographic scores, while malignancy risk was much lower (51.5%) for nodules having <3 elastographic scores ($p = 0.019$).

On univariate analysis, the Bethesda System sM subgroup, microcalcifications, irregular margins, solitary thyroid nodule, ≥3 elastographic scores, and younger age (<45 years) were associated with a high risk for malignancy (Table 1).

On multivariate analysis, sM cytology (OR: 6.734, 95% CI: 3.007–15.080, $p < 0.001$), microcalcifications (OR: 3.260, 95% CI: 1.336–7.952, $p = 0.009$), and irregular margins (OR: 6.201, 95% CI: 1.303–29.523, $p = 0.022$) were independent risk factors for malignancy (Table 2).

Table 2. Multivariate logistic regression analysis for predicting thyroid malignancy

Variable	Category	Odds ratio (95% CI)	p value
Age	<45 years	0.676 (0.354–1.291)	0.235
Indeterminate FNA	sFN	1.096 (0.518–2.321)	0.810
	sM	6.734 (3.007–15.080)	<0.001
Microcalcification	present	3.260 (1.336–7.952)	0.009
Irregular borders	present	6.201 (1.303–29.523)	0.022
Solitary nodule	present	1.941 (0.875–4.303)	0.103

FNA, fine needle aspiration; sFN, suspicious for follicular neoplasm; sM, suspicious for malignancy; CI, confidence interval.

The Performance of a Predictive Model

To develop a nomogram, regression coefficients for each variable were calculated by logistic regression analysis. Each predictive variable was proportionally assigned to point ranges from 0 to 10 on a scale in the nomogram. Risk of malignancy was calculated for each patient according to the constructed predictive model. For example, the risk of thyroid malignancy in a 42-year-old (+2.1 points) patient having a solitary nodule (+3.6 points) with AUS cytology (+0 points) and microcalcifications (+6.3 points = total 12 points) was approximately 80–85%.

The best probability cutoff point was determined as 48.5% using the ROC curve. At this cutoff point, sensitivity, specificity, positive predictive value, and negative predictive value for this established model were calculated as 74.5, 76.1, 82.7, and 66.0%, respectively. The AUC was found to be 0.784 ($p < 0.001$) (Fig. 1).

Because the curve was close to the 45° line, the calibration plot for predicting malignancy risk indicated good agreement between the predicted and observed probabilities of malignancy in indeterminate thyroid nodules according to the Bethesda System (Fig. 2).

Discussion

The interpretation and clinical decision-making concerning indeterminate cytology reports according to the Bethesda System have been considered challenging by both clinicians and patients. Indeterminate cytology has been reported in 15–30% of FNA results, 66% of these eventually proving to have benign disease after diagnostic

thyroid surgery [10]. In this study, the malignancy rate of each indeterminate category was found to be much higher than that originally recommended by the Bethesda System. The Bethesda System proposed that the anticipated rate of malignancy for each subgroup in the indeterminate category was in the range of 5–15% for AUS, 15–30% for sFN, and 60–75% for sM [11]. Although this system provides the opportunity for clinicians to use the same terminology, several recent studies have found that great variability exists among institutions when reporting malignancy risk for each indeterminate Bethesda category. Our findings also concur with those of previous studies. In a recent meta-analysis, Wang et al. [12] reported that based on 11 studies from the US, the malignancy risk of sM nodules ranged from 65 to 85%. There existed broad variation in thyroid cancer risk among AUS/FLUS and sFN categories, ranging from 0 to 48% and from 14 to 49%, respectively [12–15].

Nodules with sM that carry a higher risk of malignancy warrant thyroidectomy in almost all instances. As for patients with sFN nodules, thyroid lobectomy is a common procedure regardless of their malignancy risk. This is basically why the presence of capsular and/or vascular invasion has to be determined to reach a differential diagnosis of follicular cell neoplasms, since it is impossible to be detected by cytology. Estimating the risk of malignancy of the AUS/FLUS category represents difficulty mainly stemming from the fact that not all patients undergo surgery. Decision-making in this category to select patients for surgery is not straightforward and necessitates consideration of clinical and US features together with FNA findings. The rate of malignancy for cases undergoing surgery with AUS cytology has been estimated to be not lower than 30% in most of the recent cohorts [3, 4, 16]. A number of factors detailed below may contribute to improving the estimation of malignancy. The FLUS/AUS category is highly associated with heterogeneity in interpretation such that morphologic distinction is often made subjectively between intermediate thyroid cytology subgroups. Additionally, different cytological subdefinitions of the AUS category have also been subject to debate with respect to estimation of malignancy. The nodules which are defined as AUS mainly based on nuclear atypia carry a higher malignancy risk than those in which a predominance of the microfollicular pattern is the leading feature for AUS diagnosis. The prevalence of malignancy in patients undergoing thyroidectomy with an AUS diagnosis is reported to be higher in a tertiary referral center like ours. Low-volume cytopathologists may be conservative in their interpretations to minimize false-

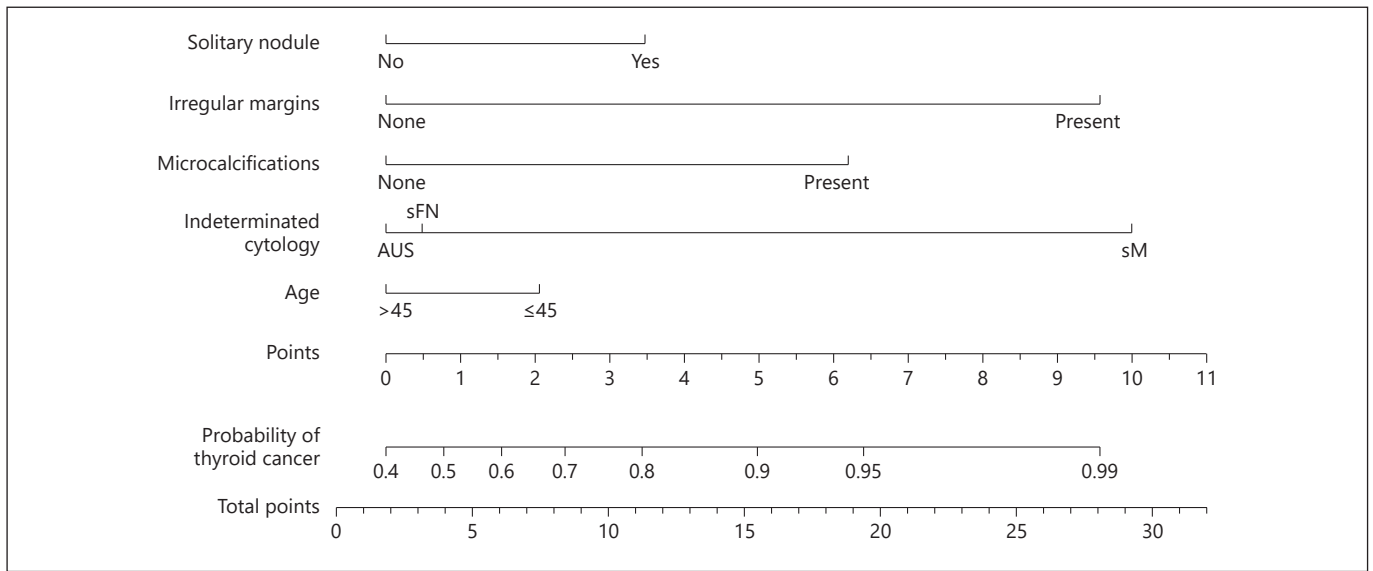


Fig. 1. A nomogram for predicting the probability of thyroid carcinoma. To use the nomogram, locate the patient variable on each axis. Draw a vertical line to the point axis and get the arithmetical sum of the points for all variables. Locate this addition score on the total point line to assess the individual probability of papillary thyroid cancer.

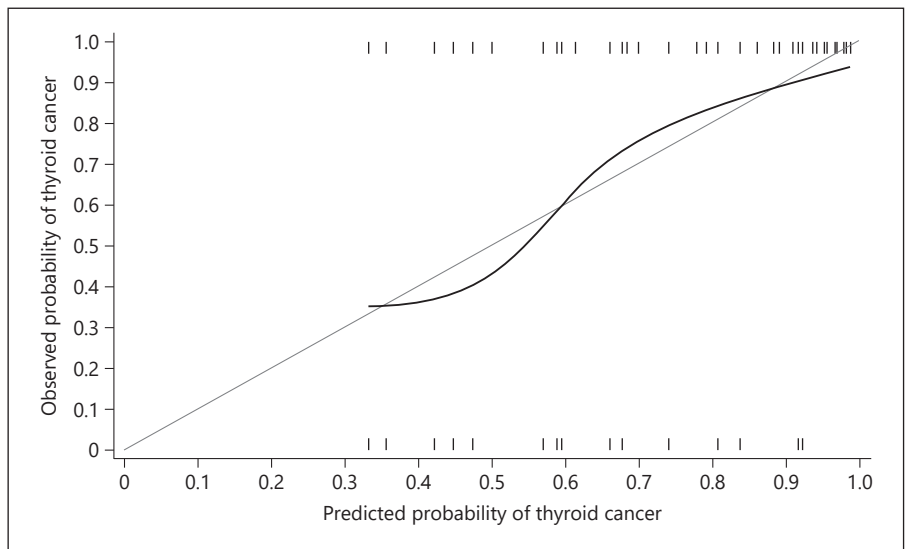


Fig. 2. Calibration plot (curve): closer distance to the 45° straight line indicates a higher accuracy of the nomogram.

positive malignant results, which may explain the wide variations of malignancy risk in AUS [17, 18]. The high rate of malignancy in indeterminate categories means that a decision for thyroidectomy is given meticulously at this particular institution. At our institution, the decision for thyroidectomy is taken by the multidisciplinary endocrine board, which involves experienced cytopathologists, pathologists, endocrinologists, and surgeons.

Therefore, our malignancy rate is 44% in the AUS category, and a comment that unnecessary surgery should be avoided as much as possible can be added.

Several reports emphasized that well-established sonographic features together with other clinical and cytological parameters may assist in improving the prediction of malignancy of indeterminate thyroid nodules [3, 16]. Microcalcifications and irregular margins have the high-

est specificities with a median value of >90%, but the sensitivities for any single US feature are significantly lower [4]. Our multivariate analysis showed that microcalcifications (OR: 3.260, $p = 0.009$) and irregular margins (OR: 6.201, $p = 0.022$) were independent risk factors for malignancy on US examination. In a recent meta-analysis pooling data from 31 separate studies, Brito et al. [19] found that the diagnostic odds ratios of internal calcifications and irregular margins for indicating malignancy were 6.78 (95% CI: 4.48–10.24) and 6.89 (95% CI: 3.35–14.1), respectively.

Current guidelines recommend that, when evaluating thyroid nodules, one should assume a comparable risk of thyroid cancer in patients with a solitary nodule and in patients with multinodular disease. However, Frates et al. [2] found that a solitary nodule had a higher risk of malignancy than a nonsolitary nodule, but the risk of malignancy per patient remains the same regardless of the number of nodules. In a recent systematic review and meta-analysis, Brito et al. [19] concluded that the risk of thyroid cancer was more frequent in a solitary nodule than in multinodular disease based on the data stemming from mostly iodine-deficient populations. In our study, the likelihood of being diagnosed with thyroid cancer was slightly higher in patients with a solitary nodule than in those with multiple nodules (75–56.4%, $p = 0.015$).

In this study, a nomogram system which makes it possible to calculate the probability of a specific clinical outcome for each patient using univariate and multivariate logistic regression analysis was designed. These analyses yielded the following predictive features (clinical, US, and cytological) to be used in this design: age, microcalcifications, irregular margins, multiplicity of nodules, and indeterminate subgroups of the Bethesda System. These variables were used to establish our predictive nomogram model. This is a very practical, efficient, and widely applicable model, as the nomogram can easily be constructed based on the aforementioned parameters.

The presented study has some limitations. Although the performance of our model is good in the determination of surgical need in individual cases, results with only internal validation were reported. This model should be

validated in multiple, independent datasets of different patient groups to obtain unbiased estimates of model performance and so that its applicability to other populations can be determined.

Conclusions

A predictive model or nomogram presented herein may help clinicians and surgeons to more reliably select patients with indeterminate thyroid cytology for considering thyroidectomy. This statistical model had good discrimination (AUC of 0.784) and accurate calibration in our patient cohort. Therefore, the rate of unnecessary thyroidectomies in patients with thyroid nodules diagnosed cytologically as “intermediate subcategories of the Bethesda System” can be decreased.

Statement of Ethics

Subjects have given their written informed consent. The study protocol has been approved by the institutional review board (20.10.2015-26/30).

Disclosure Statement

The authors declare that they have no conflict of interest.

Funding Sources

The authors have no funding to report.

Author Contributions

B.Ö. and M.H.K. conceived and designed the study. B.Ö. and M.H.K. took the lead in writing the manuscript. B.Ö., M.H.K., G.S., E.Ç.T., Ö.B., and S.K. contributed to the interpretation of the results. D.Y. and T.T.T. carried out all cytological examinations. S.K. performed the analysis and designed the figures. E.Ç. developed the theoretical framework and supervised the project. All authors discussed the results and provided critical feedback.

References

- 1 Banks ND, Kowalski J, Tsai HL, Somervell H, Tufano R, Dackiw AP, et al. A diagnostic predictor model for indeterminate or suspicious thyroid FNA samples. *Thyroid*. 2008 Sep; 18(9):933–41.
- 2 Frates MC, Benson CB, Doubilet PM, Kunreuther E, Contreras M, Cibas ES, et al. Prevalence and distribution of carcinoma in patients with solitary and multiple thyroid nodules on sonography. *J Clin Endocrinol Metab*. 2006 Sep;91(9):3411–7.
- 3 Remonti LR, Kramer CK, Leitão CB, Pinto LC, Gross JL. Thyroid ultrasound features and risk of carcinoma: a systematic review and meta-analysis of observational studies. *Thyroid*. 2015 May;25(5):538–50.

- 4 Haugen BR, Alexander EK, Bible KC, Doherty GM, Mandel SJ, Nikiforov YE, et al. 2015 American Thyroid Association Management Guidelines for Adult Patients with Thyroid Nodules and Differentiated Thyroid Cancer: The American Thyroid Association Guidelines Task Force on Thyroid Nodules and Differentiated Thyroid Cancer. *Thyroid*. 2016 Jan;26(1):1–133.
- 5 Nixon IJ, Ganly I, Hann LE, Yu C, Palmer FL, Whitcher MM, et al. Nomogram for selecting thyroid nodules for ultrasound-guided fine-needle aspiration biopsy based on a quantification of risk of malignancy. *Head Neck*. 2013 Jul;35(7):1022–5.
- 6 Yoon JH, Lee HS, Kim EK, Moon HJ, Kwak JY. A nomogram for predicting malignancy in thyroid nodules diagnosed as atypia of undetermined significance/follicular lesions of undetermined significance on fine needle aspiration. *Surgery*. 2014 Jun;155(6):1006–13.
- 7 Kim SK, Lee JH, Woo JW, Park I, Choe JH, Kim JH, et al. Prediction Table and Nomogram as Tools for Diagnosis of Papillary Thyroid Carcinoma: Combined Analysis of Ultrasonography, Fine-Needle Aspiration Biopsy, and BRAF V600E Mutation. *Medicine (Baltimore)*. 2015 May;94(21):e760.
- 8 Raparia K, Min SK, Mody DR, Anton R, Amrikachi M. Clinical outcomes for “suspicious” category in thyroid fine-needle aspiration biopsy: patient’s sex and nodule size are possible predictors of malignancy. *Arch Pathol Lab Med*. 2009 May;133(5):787–90.
- 9 Cibas ES, Ali SZ. The Bethesda System for Reporting Thyroid Cytopathology. *Thyroid*. 2009 Nov;19(11):1159–65.
- 10 Yoon JH, Kwon HJ, Lee HS, Kim EK, Moon HJ, Kwak JY. RAS Mutations in AUS/FLUS Cytology: Does it Have an Additional Role in BRAFV600E Mutation-Negative Nodules? *Medicine (Baltimore)*. 2015 Jul;94(27):e1084.
- 11 Wu HH, Rose C, Elsheikh TM. The Bethesda system for reporting thyroid cytopathology: an experience of 1,382 cases in a community practice setting with the implication for risk of neoplasm and risk of malignancy. *Diagn Cytopathol*. 2012 May;40(5):399–403.
- 12 Wang CC, Friedman L, Kennedy GC, Wang H, Kebebew E, Steward DL, et al. A large multicenter correlation study of thyroid nodule cytopathology and histopathology. *Thyroid*. 2011 Mar;21(3):243–51.
- 13 Sullivan PS, Hirschowitz SL, Fung PC, Apple SK. The impact of atypia/follicular lesion of undetermined significance and repeat fine-needle aspiration: 5 years before and after implementation of the Bethesda System. *Cancer Cytopathol*. 2014 Dec;122(12):866–72.
- 14 Theoharis CG, Schofield KM, Hammers L, Udelsman R, Chhieng DC. The Bethesda thyroid fine-needle aspiration classification system: year 1 at an academic institution. *Thyroid*. 2009 Nov;19(11):1215–23.
- 15 Deandrea M, Ragazzoni F, Motta M, Torchio B, Mormile A, Garino F, et al. Diagnostic value of a cytomorphological subclassification of follicular patterned thyroid lesions: a study of 927 consecutive cases with histological correlation. *Thyroid*. 2010 Oct;20(10):1077–83.
- 16 Mathur A, Najafian A, Schneider EB, Zeiger MA, Olson MT. Malignancy risk and reproducibility associated with atypia of undetermined significance on thyroid cytology. *Surgery*. 2014 Dec;156(6):1471–6.
- 17 Ho AS, Sarti EE, Jain KS, Wang H, Nixon IJ, Shaha AR, et al. Malignancy rate in thyroid nodules classified as Bethesda category III (AUS/FLUS). *Thyroid*. 2014 May;24(5):832–9.
- 18 Khoncarly SM, Tamarkin SW, McHenry CR. Can ultrasound be used to predict malignancy in patients with a thyroid nodule and an indeterminate fine-needle aspiration biopsy? *Surgery*. 2014 Oct;156(4):967–70.
- 19 Brito JP, Gionfriddo MR, Al Nofal A, Boehmer KR, Leppin AL, Reading C, et al. The accuracy of thyroid nodule ultrasound to predict thyroid cancer: systematic review and meta-analysis. *J Clin Endocrinol Metab*. 2014 Apr;99(4):1253–63.