Surgeon-performed ultrasound can predict benignity in thyroid nodules

Melanie Goldfarb, MD,a Steven Gondek, MD, MPH,b Carmen Solorzano, MD, FACS,c and John I. Lew, MD, FACS,a Miami, FL, Boston, MA, and Nashville, TN

Background. Surgeon-performed ultrasound (SUS) has become an extension of the physical examination in the evaluation of patients with thyroid nodules. Amid concerns that too many “diagnostic” thyroidectomies are performed, this study examines the usefulness of SUS in predicting benign thyroid nodules.

Methods. A retrospective review of 624 patients with solitary or dominant thyroid nodules and prospectively collected data of SUS nodule features before thyroidectomy (benign = 407; malignant = 217) was performed. A specific model for statistical analysis was created to predict benignity.

Results. Significant predictors of benignity by SUS after logistic regression included isoechogenicity, cystic component, no microcalcifications, regular borders, and size < 1 cm. The resultant specificity of this reduced model was 97.6% with a 10.6% sensitivity for predicting benign pathology. After excluding patients with lesions >4 cm or a history of thyroid cancer, thyroid nodules measuring 1–4 cm with the presence of the 4 other criteria by SUS was 98.5% specific for the detection of benign pathology (1.5% false-negative rate).

Conclusion. Approximately 10% of thyroidectomy patients had benign nodules that were 1–4 cm, isoechogenic, with regular borders, had a cystic component, and no microcalcifications by SUS. Patients without clinical risk factors and these SUS features may safely forego further fine needle aspiration and thyroid lobectomy, and can be monitored with serial ultrasound examinations. (Surgery 2011;150:436-41.)

UPTO 70% OF PEOPLE MAY Harbor thyroid nodules on routine screening. 1 This includes the increasing number of patients referred for nonpalpable thyroid nodules found incidentally on imaging studies for other unrelated health issues. 2-4 Although a small percentage of thyroid nodules actually harbor cancer, these incidentally discovered thyroid nodules are a cause of patient concern and represent a diagnostic dilemma for physicians.

Operative resection is the ultimate diagnostic approach in determining thyroid cancer, but it is expensive, requires general anesthesia, and can be anxiety provoking for the patient. Furthermore, only 20–25% of thyroid lobectomy specimens may have underlying cancer. 5-7 To date, fine needle aspiration (FNA) has been the “gold standard” preoperative test for the diagnosis of thyroid cancer, but its sensitivity ranges from 65–98%, with a specificity of 72–98%, and a false-negative rate of 1–11%. 8 In addition, it can be an invasive procedure with rare complications.

Surgeon-performed ultrasound (SUS) has become an extension of the physical examination in the evaluation of patients with thyroid nodules. This imaging modality is a proven, useful tool in the evaluation of patients with thyroid nodules. 9,10 Current portable ultrasound machines allow visualization of thyroid nodules as small as 1 mm, accurately determine characteristics such as microcalcifications, internal vasculature, and borders, and can detect suspicious appearing cervical lymph nodes. SUS has been shown to modify the extent of operation and aid considerably in the attainment of adequate specimens by FNA.11-13

There have been several studies in recent years that utilize ultrasound features in an attempt to predict and calculate the risk of malignancy of dominant thyroid nodules. However, no study has
examined this diagnostic dilemma from the opposite point of view; more specifically, can physicians comfortably advise a clinically significant group of patients with thyroid nodules, certain that their nodule was benign, not to undergo further invasive operative procedures for definitive diagnosis. This important question arises amid current concerns that too many “diagnostic” thyroidectomies and FNA biopsies are performed. The purpose of this study was to determine the usefulness of SUS in predicting benignity in thyroid nodules.

METHODS

Clinical and pathologic data from 1356 consecutive patients who underwent thyroidectomy from 2002 to 2009 were collected prospectively in an Institutional Review Board–approved database at the University of Miami Health System. From this database, a retrospective review of 624 patients with solitary or dominant thyroid nodules who underwent SUS before thyroidectomy was performed. Patients were subdivided into 2 groups based on final pathology results: Patients with benign thyroid disease (n = 407) and those with thyroid malignancy (n = 217). Indications for thyroidectomy included FNA results that were consistent for cancer, suspicious for cancer or indeterminate, nodule size >4 cm, compressive symptoms, hyperthyroidism secondary to a toxic nodule, suspicious ultrasound characteristics for malignancy, and history of thyroid cancer and/or head/neck radiation. Demographics collected included age, gender, race, and body mass index.

Preoperative SUS was performed using high-frequency linear array transducers 7.5–13 MHz. All ultrasonography was performed by surgeons certified by the American College of Surgeons in basic and cervical ultrasonography. Prospectively collected SUS characteristics of thyroid nodules included size (width, length, height), echogenicity (hypoechoic, isoechoic, hyperechoic), borders (regular, irregular), calcifications (microcalcifications versus coarse/none), cystic component (versus solid), shape (taller versus wider on transverse view), number of nodules (single versus multiple), and location (unilobar versus bilobar).

For statistical analysis and modeling, the control group was defined as those patients with malignant findings on final pathology and the study group was defined as those patients with benign final pathology. Univariate analysis for predictors was performed using a Student’s t test for continuous data and Chi-square analysis for categorical data. Methodologically, this was performed in a case control fashion, with significance demonstrated in those categories where a greater number of patients with a given characteristic demonstrated a negative ultrasound.

After univariate analysis, logistic regression was performed for use in creation of the final model. All variables identified as nearly significant were included in logistic regression (P < .1) along with those variables previously identified as clinically important, regardless of their statistical significance. After definition of significance for all relevant variables in multivariate regression, the final model was created. This was performed using a main effects model and stepwise elimination on nonsignificant terms.

Variables remaining after stepwise elimination were used in creation of a scoring algorithm. A previously defined acceptable false-negative rate of 2% was used in creation of this algorithm (specificity of 98%) in prediction of a negative ultrasound. Initial models were based on odds-ratios defined in logistic regression and receiver-operator curves (ROCs) created to demonstrate potential cutoffs. The model was subsequently adapted for simplicity and usability while keeping below the accepted 2% false negative rate with patients receiving 1 point for each criteria demonstrated for a maximum score of 5 points. Statistical analyses of prospectively collected data were performed using SPSS 18.0 (IBM Co. Somers, NY).

RESULTS

From the overall group, 624 patients had solitary or dominant thyroid nodules who underwent SUS before thyroidectomy. On final pathology, 407 patients had benign disease and 217 had malignant neoplasms. Univariate analysis identified significant predictors of benignity by SUS ultrasound. These predictors included male gender, non-white race, multiple nodules, lack of microcalcifications, regular nodule borders, lesions not taller > wider, and isoechogenicity (Table I).

A complete multiple regression model including all significant or nearly significant variables with a P < .1 in univariate analysis was created; 410 patients had complete data to be included in this portion of analysis. Multiple regression results showing the odds ratio and significance of each variable after correction for covariance among significant variables are demonstrated in Table II. Stepwise elimination resulted in 5 significant predictors for use in modeling. These were isoechogenicity, cystic components, lack of microcalcifications, regular nodule borders, and size <1 cm (Table III); size <1 cm was the least significant contributor to this reduced model. The remaining variables were eliminated.
from the model because of covariance demonstrated in multivariate analysis. The ROC curve for this model before simplification is shown in Fig 1, and demonstrates a sensitivity of 9% and specificity of 99.4% in predicting a benign pathology.

The same analysis was repeated after elimination of all patients with lesions > 4 cm or a previous history of thyroid cancer. In the same fashion, a second complete multiple regression model including all significant or nearly significant variables with a P < .1 on univariate analysis was created. Multiple regression results showing the odds ratio and significance of each variable after correction for covariance among significant variables are demonstrated in Table IV. This did not lead to any major differences in the reduced model.

The simplified scoring system represented by the reduced model is shown in Table V. Each of the 5 variables is given 1 point, with a total of 4 out of 5 points predicting benign disease. Reduction to this model resulted in a sensitivity of 10.6% and a specificity of 97.6% for predicting benign thyroid pathology. After elimination of lesions > 4 cm or a previous history of thyroid cancer, sensitivity was 9.6% and specificity 98.5% for predicting benign thyroid pathology.

**DISCUSSION**

Clinic-based ultrasonography has become an integral part of every office visit for surgeons performing a high volume of thyroid surgery. SUS is easy to use, cost-effective, and importantly, an accurate noninvasive method for evaluating thyroid nodules. SUS provides the most thorough
assessment of nodule characteristics with simultaneous appraisal of the contralateral thyroid lobe and cervical lymph nodes. This information allows the surgeon to not only direct operative planning including its extent, but also the ability to have a detailed, informative discussion with the patient preoperatively. When SUS features are highly suggestive of malignancy, surgeons can confidently plan for total thyroidectomy with or without central neck dissection, or if a small nodule, advise the patient to undergo an FNA biopsy or diagnostic thyroid lobectomy. Conversely, this current study shows that if preoperative SUS fits the authors’ model for benignity, clinicians can confidently advise patients that watchful waiting is an acceptable option; if the patient is truly concerned, thyroid lobectomy can be recommended.

The current study is novel in that previous reports have focused on predicting thyroid malignancy using various combinations of ultrasound features with or without FNA biopsy results and demographics. Ultrasound features have also been studied specifically in indeterminate nodules. Irregular borders, microcalcifications, and hypoechochogenicity have been shown to have the strongest association with well-differentiated thyroid cancer (30-fold greater risk). Additionally, solid composition, central vascularity, greater stiffness, and taller greater than wider in a transverse dimension are ultrasound characteristics having predictive value for malignancy on a lesser scale.

The incidence of thyroid cancer has been on the rise for the past 30 years, as well as the frequency of incidentally discovered nonpalpable thyroid nodules. Not only is the accurate diagnosis of cancer essential, but as important, the ability to confidently reassure patients of the benign nature of their thyroid nodules. Reliable clinic-based ultrasound by trained surgeons may spare many patients unnecessary FNA biopsies and operations, and possibly decrease the reliance on FNA biopsies if SUS becomes common practice.

Patients with thyroid nodules between 1 and 4 cm pose the greatest diagnostic dilemma. The current practice among many clinicians is to recommend removal of all thyroid nodules >4 cm, mainly owing to diagnostic uncertainty and secondarily because of potential for causing

Table IV. Complete multivariate logistic regression of all significant or nearly significant SUS characteristics and demographics that predict a benign thyroid nodule after elimination of lesions >4 cm and patients with previous history of thyroid cancer

<table>
<thead>
<tr>
<th>Variable</th>
<th>Odds ratio of being benign (confidence interval)</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Demographics</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gender</td>
<td>0.688 (0.333–1.423)</td>
<td>.314</td>
</tr>
<tr>
<td>White race</td>
<td>0.656 (0.392–1.097)</td>
<td>.108</td>
</tr>
<tr>
<td>Nodule characteristics</td>
<td></td>
<td></td>
</tr>
<tr>
<td>US single Lesion</td>
<td>1.612 (0.763–1.620)</td>
<td>.482</td>
</tr>
<tr>
<td>Unilobar Lesion</td>
<td>0.763 (0.360–1.620)</td>
<td>.482</td>
</tr>
<tr>
<td>Size &lt; 1cm</td>
<td>2.70 (0.68–10.75)</td>
<td>.159</td>
</tr>
<tr>
<td>No microcalcifications</td>
<td>3.047 (1.798–5.164)</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Isoechoic lesion</td>
<td>2.896 (1.694–4.951)</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Mostly cystic nodule</td>
<td>5.353 (1.338–21.416)</td>
<td>.018</td>
</tr>
<tr>
<td>Irregular borders</td>
<td>0.513 (0.300–0.877)</td>
<td>.015</td>
</tr>
<tr>
<td>Mixed cystic and solid</td>
<td>2.070 (1.155–3.712)</td>
<td>.015</td>
</tr>
<tr>
<td>Nodule not taller &gt;</td>
<td>0.898 (0.348–2.316)</td>
<td>.824</td>
</tr>
</tbody>
</table>

Table V. Scoring system to predict thyroid nodule benignity

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Points</th>
</tr>
</thead>
<tbody>
<tr>
<td>Size &lt;1 cm</td>
<td>1</td>
</tr>
<tr>
<td>Cystic components</td>
<td>1</td>
</tr>
<tr>
<td>No microcalcifications</td>
<td>1</td>
</tr>
<tr>
<td>Regular borders</td>
<td>1</td>
</tr>
<tr>
<td>Isoechoic</td>
<td>1</td>
</tr>
</tbody>
</table>

Four points predicts nodule benignity.
obstructive symptoms. Thyroid nodules < 1 cm are generally observed unless they have worrisome features on ultrasonography, which are then subject to FNA biopsy. On initial statistical analysis, close to 11% of patients could forego an invasive diagnostic procedure with a 2.4% false-negative rate of nonbenign pathology if they met the criteria of the authors’ scoring system. When the analysis was repeated excluding patients with lesions >4 cm and a previous history of thyroid cancer, the reduced model allowed 10.6% of patients to forego an invasive diagnostic procedure with a 1.5% false-negative rate of nonbenign pathology.

Because of the large number of patients in this series, this study has strong statistical power in evaluating combinations of previously reported ultrasound characteristics associated with benign disease to truly predict benignity. Ultrasound features that were significant on univariate analysis confirm isoechogeticity, lack of calcifications, regular borders, cystic components, shape not taller > wider and small size (<1 cm) are characteristic of benign disease. The main limitation of this study, however, is a selection bias of thyroid nodules that were actually surgically resected. Because patient selection was from a surgeon’s office, most patients were referred after undergoing evaluation for their thyroid nodule(s) by their primary care physician or endocrinologist, and therefore, were either large or possibly worrisome for thyroid cancer by FNA and/or ultrasonography. Second, patients who were either poor candidates for operative intervention or had benign lesions and elected for close observation were not seen for an operation, and therefore were not included in the analysis.

The real clinical significance of this report, however, lies in the statistical model that was developed with logistic regression multivariate analysis. The proposed model of utilizing 4 ultrasound characteristics to recommend with confidence more than ten percent of patients with thyroid nodules >1 cm to simply monitor their nodule with interval ultrasound can be expeditiously implemented in the everyday clinic setting by surgeons with ultrasound experience. When thyroid nodules are >1 cm, the presence of all 4 criteria on SUS (regular borders, cystic component, no microcalcifications, and isoechogeticity) was 97.6% specific for the detection of benign pathology with a 2.4% false-negative rate. Specifically for lesions between 1 and 4 cm, the presence of all 4 criteria on SUS was 98.5% specific for the detection of benign pathology with a 1.5% false-negative rate. These outcomes are comparable or more accurate than FNA. Based on this information, the authors propose an office-based algorithm for clinicians to use when evaluating their patients with thyroid nodules (Fig 2).

In conclusion, SUS is highly specific in the determination of benign thyroid nodules. In this study, approximately 10% of all patients undergoing

Fig 2. Office-based algorithm for evaluation of thyroid nodules with surgeon performed ultrasound.
thyroidectomy had an isoechoic thyroid nodule between 1 and 4 cm associated with regular borders, a cystic component, and no microcalcifications by SUS that was benign on final pathology. In the absence of clinical risk factors, patients with these SUS features may safely forego further FNA and diagnostic thyroid lobectomy, and can be monitored reliably with serial ultrasound exams.

REFERENCES

4. Davies L, Ouellette M, Hunter M, Welch GH. The increasing incidence of small thyroid cancers: where are the cases coming from? Laryngoscope 2010;120:2446-51.