



Published in final edited form as:

JAMA Intern Med. 2013 October 28; 173(19): 1788–1796. doi:10.1001/jamainternmed.2013.9245.

## Risk of Thyroid Cancer Based on Thyroid Ultrasound Imaging Characteristics: Results of a Population-Based Study

Rebecca Smith-Bindman, MD<sup>1,2,3</sup>, Paulette Lebda, MD<sup>4</sup>, Vickie A. Feldstein, MD<sup>1,3</sup>, Dorra Sellami, MD<sup>5</sup>, Ruth B. Goldstein, MD<sup>1,3</sup>, Natasha Brasic, MD<sup>1</sup>, Chengshi Jin, PhD<sup>2</sup>, and John Kornak, PhD<sup>2</sup>

<sup>1</sup>Department of Radiology and Biomedical Imaging, University of California, San Francisco, CA

<sup>2</sup>Department of Epidemiology and Biostatistics, University of California, San Francisco, CA

<sup>3</sup>Department of Obstetrics, Gynecology and Reproductive Sciences, University of California, San Francisco, CA

<sup>4</sup>Department of Radiology, Cleveland Clinic

<sup>5</sup>California Advanced Imaging Medical Associates, Mills-Peninsula Hospital, San Mateo California

### Abstract

**Importance**—There is wide variation in the management of thyroid nodules identified on ultrasound imaging.

**Objective**—To quantify the risk of thyroid cancer associated with thyroid nodules based on their ultrasound characteristics.

**Methods**—Retrospective case-control study of patients who underwent thyroid ultrasound between January 1<sup>st</sup>, 2000 and March 30<sup>th</sup>, 2005. Thyroid cancers were identified through linkage with the California Cancer Registry.

---

Correspondence: Rebecca Smith-Bindman, MD, Department of Radiology and Biomedical Imaging, 350 Parnassus Ave. Suite 307, University of California, San Francisco, CA, 94143-0336. Rebecca.smith-bindman@ucsf.edu. Phone: 415-353-4946. Fax: 415-353-2790.

Conflict of Interest Disclosure: All authors have completed and submitted the ICMJE Form for Disclosure of Potential Conflicts of Interest and none were reported.

Author Contributions: Smith-Bindman had full access to all the data in this study and takes responsibility for the integrity of the data and the accuracy of the data analysis.

*Study concept and design:* Smith-Bindman, Goldstein, Feldstein

*Acquisition of data:* Smith-Bindman

*Analysis and interpretation of data:* Smith-Bindman, Lebda, Feldstein, Sellami, Goldstein, Brasic, Jin, Kornak,

*Drafting of the manuscript:* Smith-Bindman

*Critical revision of the manuscript for important intellectual content:* Smith-Bindman, Lebda, Feldstein, Sellami Goldstein, Brasic, Kornak

*Statistical analysis:* Smith-Bindman, Jin, Kornak

*Obtained funding:* Smith-Bindman, Sellami

*Administrative, technical, or material support:* Smith-Bindman, Lebda, Sellami, Brasic, Feldstein, Goldstein, Jin, Kornak,

*Study supervision:* Smith-Bindman

**Additional Contributions:** We thank the following people for their valuable assistance in gathering data for this study: Phillip Chu, and the Northern California Cancer Registry. The collection of cancer incidence data used in this study was supported by the California Department of Public Health as part of the statewide cancer-reporting program.

**Results**—8 806 patients underwent 11 618 thyroid ultrasound examinations during the study period including 105 subsequently diagnosed with thyroid cancer. Thyroid nodules were common in patients diagnosed with cancer (97%) and patients not diagnosed with thyroid cancer (56%). Three ultrasound nodule characteristics—micro-calcifications (odds ratio [OR] 8.1 [95% CI 3.8, 17.3]), size greater than 2 cm (OR 3.6 [95% CI 1.7, 7.6]) and an entirely solid composition (OR 4.0 [95% CI 1.7, 9.2]) - were the only findings associated with the risk of thyroid cancer. If a single characteristic is used as an indication for biopsy, most patients with thyroid cancer would be detected (sensitivity .88 [95% CI .80, .94]) with a high false positive rate (.44 [95% CI .43, .45]) and a low likelihood ratio positive (2.0 [95% CI 1.8, 2.2]), and 56 biopsies will be performed per cancer diagnosed. If *two* characteristics were required for biopsy, the sensitivity and false positive rates would be lower (sensitivity 0.52 [95% CI 0.42, 0.62]; false positive rate 0.07 [95% CI 0.07, 0.08]), the likelihood ratio positive would be higher (7.1 [95% CI 6.2, 8.2]), and only 16 biopsies will be performed per cancer diagnosed. In comparison to performing biopsy of all thyroid nodules greater than 5 mm, adoption of this more stringent rule requiring two abnormal nodule characteristics to prompt biopsy would reduce unnecessary biopsies by 90%, while maintaining a low risk of cancer, 5 per 1000 patients, for whom biopsy is deferred.

**Conclusion**—Thyroid ultrasound could be used to identify patients who have a low risk of cancer for whom biopsy could be deferred. Based on these results, these findings should be validated in a large prospective cohort.

## INTRODUCTION

Ultrasound has replaced nuclear medicine as the most frequently used imaging test of the thyroid.<sup>1</sup> The growth in the use of thyroid ultrasound by radiologists, endocrinologists and head and neck surgeons has led to the discovery of large numbers of asymptomatic thyroid nodules, which may occur in 50% or more of adults.<sup>2,3</sup> as well as a rapid rise in the diagnosis of thyroid cancer.<sup>4</sup> In contrast, clinically apparent thyroid cancer is rare, affecting 1/10,000 people annually, and fewer than 1% of individuals over the course of their lives.<sup>4-6</sup> Because of the high prevalence of nodules, and the rarity of symptomatic cancer, only a minority of thyroid nodules is malignant. Uncertainty about which nodules may harbor cancer and lack of evidence-based management guidelines has resulted in a myriad of conflicting recommendations regarding which nodules warrant biopsy,<sup>6-17,18,19-21</sup> frequent thyroid biopsies, and the over-diagnosis of thyroid cancers that would otherwise likely have remained asymptomatic in the absence of detection.<sup>4,22,23</sup>

While many studies have analyzed the association between the ultrasound characteristics of thyroid nodules and the risk of thyroid cancer, most studies are small and all limited their analysis to patients who underwent biopsy, where the decision to biopsy was influenced by the ultrasound result.<sup>6-17,18,19-21</sup> This ascertainment bias will overestimate the risk of cancer associated with thyroid biopsy and the accuracy of ultrasound.<sup>24-26</sup> The information that is most important to patients and providers managing care includes quantifying the risk of cancer associated with a nodule with a particular imaging characteristic and no prior publication can accurately provide this information. This has hindered the development of an evidence-based strategy for determining which nodules should be biopsied because of an elevated cancer risk. The purpose of this study was to determine the ultrasound

characteristics that are associated with cancer, and to use this information for creating a standardized system for interpreting thyroid ultrasound.

## METHODS

We conducted a retrospective, case-control study at the University of California San Francisco (UCSF), including consecutive patients who underwent thyroid ultrasound between January 1<sup>st</sup>, 2000 and March 30<sup>th</sup> 2005. A waiver of patient informed consent was obtained. Patients were excluded if they had a prior unilateral or bilateral thyroidectomy for benign or malignant disease.

### Linkage to Population Cancer Registry

Cancers identified in the cohort were identified through linkage with the California Cancer Registry (CCR), a population-based cancer registry collecting cancer incidence and mortality data for all of California.<sup>27</sup> The Registry is a collaboration between the Cancer Surveillance Section of California Department of Public Health, The Public Health Institute, and eight regional cancer registries, that by legislative mandate, have collected cancer incidence data from hospitals and other facilities across the state since 1988. The registry is certified by the North American Association of Central Cancer Registries (NAACCR) as meeting their highest standard for completeness of cancer ascertainment, reflecting capture of over 97% of cancers diagnosed in the state.<sup>28</sup> We included thyroid cancers diagnosed through March 30, 2007, allowing a minimum 2 years of follow-up after the last enrolled patient's ultrasound, during which a cancer could be diagnosed, and at least 2 years further follow-up to ensure reporting to the registry.<sup>27</sup> Patients diagnosed with non-thyroid malignant neoplasms (other than skin cancer) were excluded to prevent the inclusion of the rare, but theoretically possibly metastatic cancer to the thyroid, as these metastatic cancers would not be captured by the cancer registry. All patients diagnosed with cancer (cases, Table 1) and a sample of patients not diagnosed with cancer (controls), matched four to one to the cancer patients on age, sex and year of the ultrasound exam, were selected for detailed review of the sonogram.

### Characterizations of the Ultrasound Examinations

We retrieved and reviewed the ultrasound examinations on 96 (91%) of the cancer patients on the Radiology PACS (Picture Archiving and Communication System) and 369 controls. Each ultrasound examination was reviewed independently by two board-certified radiologists blinded to cancer status. Disagreement was resolved by consensus. For each patient, each reviewer independently recorded the number, size and characteristics of all nodules >5 mm. There was good to outstanding agreement (*kappas* 0.73 to 1.0) in the categorization of the specific ultrasound characteristics.

### Assignment of Cancer Status to Thyroid Nodules

In patients selected as controls, all nodules were considered benign. In 43 (44.8%) of cancer patients, a single nodule was identified and it was considered malignant. In 50 cancer patients (52.1%), multiple thyroid nodules were identified. To ensure correct attribution of cancer to the correct nodule, one of the authors was un-blinded, patient records (radiology,

pathology, surgery) were reviewed to determine which nodules were malignant. In the small number of cases that we were unable to determine which nodule harbored cancer, all nodules were considered malignant. Nodules in patients never diagnosed with thyroid cancer (n=428) and benign nodules in cancer patients (n=87) were combined to create our final control group of benign nodules (n=515). Note that 3 cancer patients (3.1%) did not have any nodules > 5 mm identified on their ultrasound imaging.

## Analysis

We compared mean age, age group, sex and year of study between patients diagnosed with cancer and controls. We used Chi-square to determine whether the number of nodules varied by age group. We performed single predictor modeling to assess the association between specific ultrasound characteristics and cancer status using Generalized Estimating Equations (GEE), with a compound symmetry (exchangeable) correlation structure to account for the correlated outcomes between multiple nodules within a patient. For variables that were statistically significant in single predictor model, we calculated diagnostic accuracy statistics (sensitivity, specificity, likelihood ratios, predictive values).

To build the GEE models, we added variables that were statistically significant in single predictor models one at a time in the order of the effect size. Variables were retained if the associated p-value after inclusion was < 0.10 for that variable. The ultrasound characteristics that were retained in the final multiple-predictor model (micro-calcifications, size  $\geq 2$  cm and solid composition, Figure 1) were combined in various ways, via logical “and/or” criteria, to define an “abnormal ultrasound interpretation.” The ultrasound characteristics that were retained in the final multiple-predictor model (micro-calcifications, size  $\geq 2$  cm and solid composition, Figure 1) were combined in various ways, via logical “and/or” criteria, to define an “abnormal ultrasound interpretation.” The risk of cancer (predictive values) associated with each definition of an abnormal ultrasound was calculated, accounting for the sampling strategy in the entire cohort. The positive predictive value (PPV) is the risk of cancer for a patient that is found to have an abnormal ultrasound interpretation, and the negative predictive (NPV) is the probability of being cancer free if the ultrasound is negative. For each definition of an abnormal ultrasound, we calculated the number of cancers missed per 1000 ultrasounds performed. The number of patients needed to undergo a biopsy (NNTB) to detect a single cancer was defined as the inverse of the PPV. We performed several sensitivity analyses to determine whether implicit assumptions in the primary analysis were reasonable. More details on the analysis are provided in the online Appendix.

## RESULTS

8 806 patients underwent 11 618 thyroid ultrasounds during the study period including 105 diagnosed with thyroid cancer (incidence 0.9 cancers per 100 ultrasound examinations). The cancers were diagnosed 1 day to 6.1 years following ultrasound imaging, and among control patients, there was a mean follow up of 4.2 years (range 2.0 – 10.9). There were no significant differences in the matching variables between cases and controls.

Thyroid nodules were common among patients diagnosed with thyroid cancer (96.9%) as well as patients not diagnosed with thyroid cancer (56.4%), Table 3. Among the 96 patients cases, 102 malignant nodules and 87 benign nodules were identified, with an increase in the number of nodules seen with advancing age. Among the 372 controls, 428 benign nodules were identified and the number of nodules did not vary with age.

### Single Predictor Modeling Results

Several ultrasound findings were significantly associated with the odds of a nodule harboring cancer, Table 4. Micro-calcifications had the strongest association with cancer; 38% of cancer nodules vs. 5% of benign nodules had micro-calcifications, reflecting approximately a 7 fold increase in the likelihood of cancer if micro-calcifications were seen (likelihood ratio positive 7.0 [95% CI 6.0, 8.2]) and a 30% reduction in the likelihood of cancer if micro-calcifications were not seen (likelihood ratio negative 0.65 [95% CI 0.56, 0.76]). The corresponding odds ratio was 11.6 (95% CI 6.5, 20). Course calcifications, nodule composition, nodule echogenicity, central vascularity, margins and shape were also each significantly associated with cancer, but the magnitude of association was smaller, with odds ratios ranging from 1.6 to 2.8. Rim calcifications and comet tail artifacts, peripheral vascularity and the presence of a halo were not associated with the likelihood of cancer. The odds of cancer increased with nodule size, and the largest nodules had the greatest odds of cancer (likelihood ratio 1.8 [95% CI 1.5, 2.1], and OR 3.1, [95% CI 1.8, 5.2]) for nodules > 2 cm compared with nodules under 1 cm. Simple cysts never reflected cancer.

### Multiple Predictor Nodule Modeling Results

Only three nodule characteristics were significantly associated with the risk of cancer in the multiple predictor modeling; micro-calcifications (OR 8.1 [95% CI 3.8, 17.3]), size greater than 2 cm (OR 3.6 [95% CI 1.7, 7.6]), and an entirely solid composition (OR 4.0 [95% CI 1.7, 9.2]), Table 5. The inclusion of the remaining nodule characteristics were not significantly associated with the risk of cancer, and including them in the definition of an abnormal nodule added less than 2% cancer detection.

The accuracy of the several definitions of an abnormal ultrasound are provided in Table 6. If *any one* of the three characteristics is used to prompt biopsy, most patients with thyroid cancer would be detected (sensitivity 0.88 [95% CI 0.80, 0.94]) at a false positive rate of 0.44 (95% CI 0.43, 0.45). The high false positive rate of this approach is reflected in a low PPV (i.e., risk of cancer) of 1.8% (95% CI 1.5%, 2.2%) when a single characteristic is used to prompt biopsy, and 56 biopsies will be required per cancer diagnosed. If *two abnormal ultrasound characteristics* were required to prompt biopsy, the sensitivity and false positive rates would be lower (sensitivity 0.52 [95% CI 0.42, 0.62]; false positive rate 0.07 [95% CI 0.07, 0.08]) and the risk of cancer in those with a suspicious ultrasound would be higher (PPV 6.2% [95% CI 4.7%, 8.7%]) and fewer biopsies, 16, would be required per cancer diagnosed. In comparison to existing guidelines that suggest biopsy of all thyroid nodules greater than 5 mm<sup>7,8</sup> adoption of this more stringent rule requiring two abnormal characteristics to prompt biopsy would reduce unnecessary biopsies by 90%, while maintaining a low risk of cancer in patients in whom biopsy is deferred (i.e., 5 cancers per 1000 ultrasound examinations, 0.5%).

The most specific definition of an abnormal ultrasound is one requiring *all three* abnormal characteristics to prompt biopsy, however this definition would detect only a small proportion of cancers (sensitivity 0.07 [95% CI 0.03, 0.14]), but would have a high likelihood ratio positive of 28 (95% CI 23, 34).

The tradeoff between the different definitions of an abnormal ultrasound and test accuracy is shown in Figure 2. As the number of criteria required to prompt biopsy increases, the number of missed cancers (false negatives) increases, and the number of patients who will be biopsied in order to detect a cancer will decrease. For example, if two criteria instead of one are required to prompt biopsy, the rate of missed cancers among patients who do not undergo biopsy increases from 2 to 5 per 1000 ultrasound examinations, while the number of biopsies needed to detect a cancer decreases from 56 to 16.

### **Predictive Values: Patient-Level Analysis**

The risk of cancer based on the ultrasound appearance of the thyroid is shown in Table 7. The risk of cancer is low among patients with a homogeneous thyroid, where no nodules were identified (0.60 cancers per 1000). The risk of cancer is also low in patients where the only ultrasound characteristic is a simple cyst (0.32 cancers per 1000).

If the presence of a single abnormal characteristic is used to define an abnormal ultrasound, patients with a normal examination will have a risk of cancer of 2 per 1000, whereas patients with an abnormal exam will have a risk of cancer of 18 per 1000. If two or more characteristics are required to define an ultrasound exam as abnormal, patients with a negative exam will have a risk of cancer of 5 per 1000, and patients with an abnormal exam will have a risk of cancer of 62 per 1000, putting them in a moderate risk category. Microcalcifications are the most predictive characteristic and are associated with a cancer risk of 82 per 1000. If an abnormal ultrasound is defined as one where microcalcifications or a solid mass greater than 2 cm is seen, 58 cancers will be diagnosed per 1000 patients. When a solid mass > 2 cm with microcalcifications is seen, nearly all of these nodules harbor cancer. 960 per 1000.

### **Alternative Analysis**

The results were robust across all of the sensitivity analysis, and changed little when we varied our primary assumptions in the analysis.

## **DISCUSSION**

Thyroid nodules are extremely common. Even among patients selected as controls in our study, 56% had thyroid nodules greater than 5 mm, and nearly a third had multiple nodules. In contrast to previous reports that have suggested the prevalence of cancer in thyroid nodules as high as 23%, we found only 1.6% of patients who had one or more thyroid nodules 5 mm or greater harbored cancer. Thus while thyroid nodules are common, the vast majority, 98.5%, are benign, highlighting the importance of being prudent in deciding which nodules should be sampled to reduce unnecessary biopsies.<sup>22</sup> Unnecessary tissue sampling is not only invasive and costly, but leads to repeated sampling and unnecessary open surgical procedures, as up to one third of fine needle aspiration biopsies may be non-diagnostic,

requiring open surgical biopsy for diagnosis.<sup>8,9,29,30</sup> We found that only three ultrasound characteristics: micro-calcifications, size  $\geq 2$  cm, and entirely solid composition – were statistically significantly associated with the risk of cancer, and that when used in combination, these three characteristics could be used to help determine which nodules should be sampled. Simple cysts are essentially never malignant and should not be sampled.<sup>31</sup>

There are many ways to characterize the accuracy of ultrasound. We believe the risk of cancer, PPV, is the most relevant to patients and physicians and ours is the first study that permits estimating this risk. A patient's risk of harboring cancer ranges from 2 per 1000 among patients whose thyroid ultrasound has none of the three characteristics identified; 18 per 1000 if a patient has a nodule with a single characteristic; 62 per 1000 if a patient has a nodule with two abnormal characteristics; and 960 per 1000 if a patient has a nodule with all three characteristics. While there is growing concern regarding over-diagnosis and over treatment across all areas of medicine,<sup>22,32–34</sup> there are no well-established guidelines of what risk is low enough that an imaging finding can be ignored. In other areas of diagnostic testing, for example when assessing patients at risk for acute coronary syndrome, or breast cancer (diseases with higher morbidity and mortality than thyroid cancer), often a risk of less than 1% or 0.5% is considered sufficiently low that further evaluation is deemed unnecessary. If a thyroid cancer risk  $< 0.5\%$  is considered acceptable for those in whom biopsy is deferred, using micro-calcifications or the combined observations of a large ( $\geq 2$  cm) solid nodule as the only features to prompt biopsy reflects a good choice. In comparison with various guidelines that recommend biopsies in a larger number of patients<sup>13</sup> limiting biopsy to nodules that fulfill this definition would reduce the number of biopsies by as much as 90%, while maintaining a low cancer rate of 5/1000 among individuals who do not undergo thyroid sampling. Most thyroid cancers have a favorable prognosis, with a 20-year survival greater than 97% seen even among patients who do not receive immediate treatment.<sup>10,23,34,35</sup> Thus, given the favorable prognosis of most thyroid cancer even without treatment, a risk of cancer of 0.5% among those with a negative examination seems to balance between detection and unnecessary tissue sampling. Ongoing ultrasound surveillance of patients with nodules who do not meet the criteria for biopsy, is unlikely to prove beneficial given our results ascribe these patients a low risk of cancer for as long as 10 years following imaging.

Our study was designed to determine how to reduce unnecessary and excessive thyroid surveillance and biopsy. Our study does not provide evidence as to whether the detection of thyroid cancers will lead to improved patient outcomes. There has been a recent rise in the observed incidence of small and micro thyroid cancer<sup>4,5,35</sup> without a corresponding change in thyroid cancer mortality rate, raising the question as to whether there is benefit to the earlier diagnosis or treatment of incidental thyroid cancer.<sup>22,23,36,37</sup>

A large number of previous studies have assessed the risk of cancer associated with the ultrasound appearance of the thyroid.<sup>6,7,8,10,11–17,18,19–21,38–43</sup> All previous studies will have inflated the association between nodule characteristics and cancer risk because they limited their analysis to nodules that underwent biopsy. For example, Ahn et al. compared various existing guidelines for prompting fine needle aspiration in a sample of 1398 patients

who had undergone biopsy.<sup>13</sup> In this sample, 20% of the included patients had cancer, contrasting with the 1.5% cancer rate in our study. He reports that the PPV value for cancer if a patient has micro-calcifications is 85.1%, whereas using our population-based approach without ascertainment bias, we found a PPV of 5.8%. We considered a large number of nodule characteristics endorsed by other authors<sup>5, 7–26</sup> but when put into the multiple predictor models, most of the characteristics were not significantly associated with cancer risk.

It is widely reported that the number of benign thyroid nodules increases with age. We observed this relationship among patients diagnosed with cancer, but not among patients without cancer.

The main strength of our study is the large sample size and the linkage of the cohort with data from a comprehensive cancer registry, which allows accurate assessment of the true underlying prevalence of cancer. The analysis has several limitations. We did not have accurate information about why patients underwent imaging – and the risk of cancer may vary by why patients were sent for sonograms. We did not stratify the results by the histological type of cancer, although the majority of included cancers were papillary cancer, as is the case for thyroid cancer in general. There are several ultrasound features that we did not assess, but these are rare, such as extra-capsular growth, or abnormal lymph nodes.<sup>11</sup> We did not include the theoretical metastatic cancer to the thyroid as these would not be captured in the cancer registry data. However, we also linked to the local pathology database, and no cases of metastatic cancer were identified in the thyroid biopsies included.

The increased utilization<sup>1</sup> and improved technical quality of ultrasound has given rise to the detection of multiple morphologic characteristics, without clear criteria for what needs further evaluation,<sup>22</sup> resulting in greater tissue sampling and excessive treatment.<sup>23,35</sup> In mammography, the adoption of uniform interpretation standards through BI-RADS (Breast Imaging Reporting and Data System) has been useful in allowing comparative effectiveness work in breast imaging, and efforts to standardize the interpretation of mammograms. Similar adoption of uniform standards for interpreting of thyroid sonograms, would be a first step toward standardizing the diagnosis and treatment of thyroid cancer, and limiting unnecessary diagnostic testing and treatment.

## Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

## Acknowledgments

Funding/Support: This study was supported by the National Cancer Institute R21CA131698 and K24 CA125036, and a University of California San Francisco Department of Radiology and Biomedical Engineering SEED grant. The content is solely the responsibility of the authors and does not represent the official views of the National Cancer Institute or the National Institutes of Health. The funding organizations had no role in the design and conduct of the study, data collection, management, analysis, or interpretation of the data; or preparation, review, and approval of the manuscript.

Role of the Sponsor: None



## References

1. Smith-Bindman R, Miglioretti DL, Johnson E, et al. Use of diagnostic imaging studies and associated radiation exposure for patients enrolled in large integrated health care systems, 1996–2010. *JAMA*. Jun 13; 2012 307(22):2400–2409. [PubMed: 22692172]
2. Ezzat S, Sarti DA, Cain DR, Braunstein GD. Thyroid incidentalomas. Prevalence by palpation and ultrasonography. *Arch Intern Med*. Aug 22; 1994 154(16):1838–1840. [PubMed: 8053752]
3. Harach HR, Franssila KO, Wasenius VM. Occult papillary carcinoma of the thyroid. A “normal” finding in Finland. A systematic autopsy study. *Cancer*. Aug 1; 1985 56(3):531–538. [PubMed: 2408737]
4. Davies L, Welch HG. Increasing incidence of thyroid cancer in the United States, 1973–2002. *JAMA*. May 10; 2006 295(18):2164–2167. [PubMed: 16684987]
5. Howlader, NNA.; Krapcho, M.; Neyman, N.; Aminou, R.; Waldron, W.; Altekruse, SF.; Kosary, CL.; Ruhl, J.; Tatalovich, Z.; Cho, H.; Mariotto, A.; Eisner, MP.; Lewis, DR.; Chen, HS.; Feuer, EJ.; Cronin, KA.; Edwards, BK., editors. SEER Cancer Statistics Review, 1975–2008. National Cancer Institute; Bethesda, MD: 2011. [http://seer.cancer.gov/csr/1975\\_2008/](http://seer.cancer.gov/csr/1975_2008/), based on November 2010 SEER data submission, posted to the SEER web site
6. Aschebrook-Kilfoy B, Ward MH, Sabra MM, Devesa SS. Thyroid cancer incidence patterns in the United States by histologic type, 1992–2006. *Thyroid*. Feb; 2011 21(2):125–134. [PubMed: 21186939]
7. Cooper DS, Doherty GM, Haugen BR, et al. Revised American Thyroid Association management guidelines for patients with thyroid nodules and differentiated thyroid cancer. *Thyroid*. Nov; 2009 19(11):1167–1214. [PubMed: 19860577]
8. Cooper DS, Doherty GM, Haugen BR, et al. Management guidelines for patients with thyroid nodules and differentiated thyroid cancer. *Thyroid*. Feb; 2006 16(2):109–142. [PubMed: 16420177]
9. Castro MR, Gharib H. Continuing controversies in the management of thyroid nodules. *Annals of internal medicine*. Jun 7; 2005 142(11):926–931. [PubMed: 15941700]
10. Tan GH, Gharib H. Thyroid incidentalomas: management approaches to nonpalpable nodules discovered incidentally on thyroid imaging. *Annals of internal medicine*. Feb 1; 1997 126(3):226–231. [PubMed: 9027275]
11. Gharib H, Papini E, Paschke R, 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: executive summary of recommendations. *J Endocrinol Invest*. 2010; 33(5 Suppl):51–56. [PubMed: 20543551]
12. Ito Y, Amino N, Yokozawa T, et al. Ultrasonographic evaluation of thyroid nodules in 900 patients: Comparison among ultrasonographic, cytological, and histological findings. *Thyroid*. Dec; 2007 17(12):1269–1276. [PubMed: 17988196]
13. Ahn SS, Kim EK, Kang DR, Lim SK, Kwak JY, Kim MJ. Biopsy of thyroid nodules: comparison of three sets of guidelines. *AJR Am J Roentgenol*. Jan; 2010 194(1):31–37. [PubMed: 20028902]
14. Frates MC, Benson CB, Charboneau JW, et al. Management of thyroid nodules detected at US: Society of Radiologists in Ultrasound consensus conference statement. *Radiology*. Jan; 2005 237(3):794–800. [PubMed: 16304103]
15. Maia FF, Zantut-Wittmann DE. Thyroid nodule management: clinical, ultrasound and cytopathological parameters for predicting malignancy. *Clinics (Sao Paulo)*. Aug; 2012 67(8):945–954. [PubMed: 22948464]
16. Gonzalez-Gonzalez A, Mate Valdezate A, Parra Arroyo A, Tenias Burillo JM. Diagnostic efficiency of sonographic findings of thyroid nodules in the detection of malignancy. *Endocrinol Nutr*. Jun-Jul; 2010 57(6):240–244. [PubMed: 20447882]
17. Butros R, Boyvat F, Ozyer U, et al. Management of infracentimetric thyroid nodules with respect to ultrasonographic features. *Eur Radiol*. May; 2007 17(5):1358–1364. [PubMed: 17021705]
18. Cappelli C, Castellano M, Pirola I, et al. The predictive value of ultrasound findings in the management of thyroid nodules. *QJM*. Jan; 2007 100(1):29–35. [PubMed: 17178736]
19. Moon WJ, Jung SL, Lee JH, et al. Benign and malignant thyroid nodules: US differentiation - Multicenter retrospective study. *Radiology*. Jun; 2008 247(3):762–770. [PubMed: 18403624]

20. Hong YJ, Son EJ, Kim EK, Kwak JY, Hong SW, Chang HS. Positive predictive values of sonographic features of solid thyroid nodule. *Clin Imaging*. 2010; 34(2):127–133. [PubMed: 20189077]
21. Alexander EK, Marqusee E, Orcutt J, et al. Thyroid nodule shape and prediction of malignancy. *Thyroid*. Nov; 2004 14(11):953–958. [PubMed: 15671774]
22. Moynihan R, Doust J, Henry D. Preventing overdiagnosis: how to stop harming the healthy. *BMJ*. 2012; 344:e3502. [PubMed: 22645185]
23. Davies L, Welch HG. Thyroid cancer survival in the United States: observational data from 1973 to 2005. *Arch Otolaryngol Head Neck Surg*. May; 2010 136(5):440–444. [PubMed: 20479371]
24. Lijmer JG, Mol BW, Heisterkamp S, et al. Empirical evidence of design-related bias in studies of diagnostic tests. *Jama*. Sep 15; 1999 282(11):1061–1066. [PubMed: 10493205]
25. Bossuyt PM, Reitsma JB, Bruns DE, et al. The STARD statement for reporting studies of diagnostic accuracy: explanation and elaboration. *Annals of internal medicine*. Jan 7; 2003 138(1):W1–12. [PubMed: 12513067]
26. Mol BW, Lijmer JG, Evers JL, Bossuyt PM. Characteristics of good diagnostic studies. *Semin Reprod Med*. Feb; 2003 21(1):17–25. [PubMed: 12806556]
27. CDPH. Incidence and mortality counts and age-adjusted (2000 U.S. Population) Rates per 100,000 persons by year, race/ethnicity, and sex, California, 1988–2009. *Thyroid Gland Cancer*. 2011. <http://www.ccrca.org/6/2011>
28. Das B, Clegg LX, Feuer EJ, Pickle LW. A new method to evaluate the completeness of case ascertainment by a cancer registry. *Cancer causes & control : CCC*. Jun; 2008 19(5):515–525. [PubMed: 18270798]
29. Baloch ZW, Hendreen S, Gupta PK, et al. Interinstitutional review of thyroid fine-needle aspirations: impact on clinical management of thyroid nodules. *Diagnostic cytopathology*. Oct; 2001 25(4):231–234. [PubMed: 11599106]
30. Alexander EK, Heering JP, Benson CB, et al. Assessment of nondiagnostic ultrasound-guided fine needle aspirations of thyroid nodules. *The Journal of clinical endocrinology and metabolism*. Nov; 2002 87(11):4924–4927. [PubMed: 12414851]
31. Frates MC, Benson CB, Doubilet PM. Likelihood of thyroid cancer based on sonographic assessment of nodule size and composition. *Radiological Society of North America*. 2004
32. Medicine ABoI. [Accessed Last accessed on: 5/15/2013, 2012] U.S. Physician Groups Identify Commonly Used Tests or Procedures They Say are Often Not Necessary. 2012. <http://www.abimfoundation.org/News/ABIM-Foundation-News/2012/Choosing-Wisely.aspx>
33. Grady D, Redberg RF. Less Is More How Less Health Care Can Result in Better Health. *Arch Intern Med*. May 10; 2010 170(9):749–750. [PubMed: 20458080]
34. Welch KC, McHenry CR. Total thyroidectomy: is morbidity higher for Graves' disease than nontoxic goiter? *J Surg Res*. Sep; 2011 170(1):96–99. [PubMed: 21550063]
35. Davies L, Ouellette M, Hunter M, Welch HG. The increasing incidence of small thyroid cancers: where are the cases coming from? *Laryngoscope*. 2010; 120(12):2446–2451. [PubMed: 21108428]
36. Shiller SM, Konduri K, Harshman LK, Welch BJ, O'Brien JC. Recurrent thyroid cancer with changing histologic features. *Proc (Bayl Univ Med Cent)*. Jul; 2010 23(3):304–310. [PubMed: 21240322]
37. Black P, Straaten A, Gutjahr P. Secondary thyroid carcinoma after treatment for childhood cancer. *Med Pediatr Oncol*. Aug; 1998 31(2):91–95. [PubMed: 9680933]
38. Baier NH, PF, Gervais DA, Samir A, Halpern EF, Mueller PR, Harisinghani MG. Fine-needle aspiration biopsy of thyroid nodules: experience in a cohort of 944 patients. *American Journal of Roentgenology*. 2009; 194(3):1175–1179. [PubMed: 19770344]
39. Kim EK, Park CS, Chung WY, et al. New sonographic criteria for recommending fine-needle aspiration biopsy of nonpalpable solid nodules of the thyroid. *AJR Am J Roentgenol*. Mar; 2002 178(3):687–691. [PubMed: 11856699]
40. Papini E, Guglielmi R, Bianchini A, et al. Risk of malignancy in nonpalpable thyroid nodules: predictive value of ultrasound and color-Doppler features. *The Journal of clinical endocrinology and metabolism*. May; 2002 87(5):1941–1946. [PubMed: 11994321]

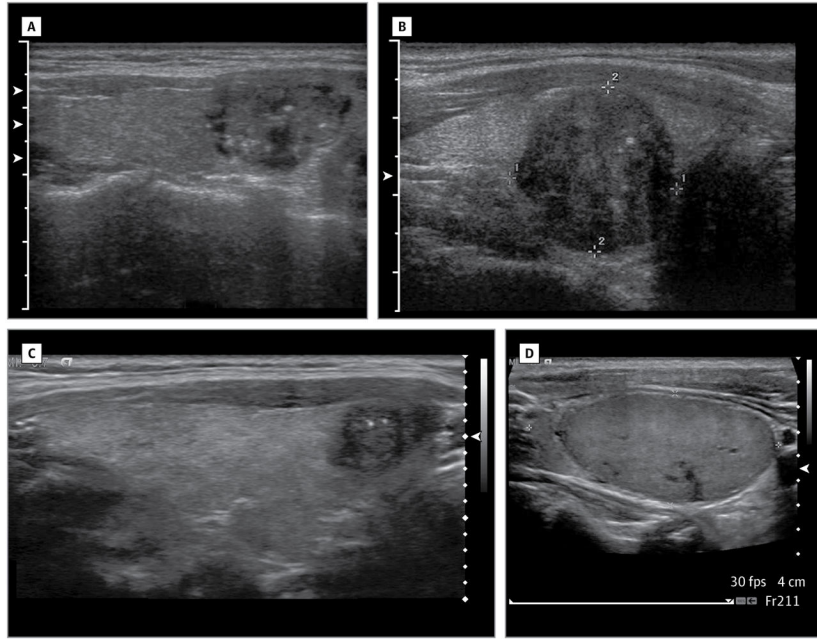
41. Gharib H, Papini E, Paschke R. Thyroid nodules: a review of current guidelines, practices, and prospects. *Eur J Endocrinol*. Nov; 2008 159(5):493–505. [PubMed: 18728120]
42. Cibas ES, Alexander EK, Benson CB, et al. Indications for thyroid FNA and Pre-FNA requirements: A Synopsis of the National Cancer Institute Thyroid Fine-Needle Aspiration State of the Science Conference. *Diagnostic cytopathology*. Jun; 2008 36(6):390–399. [PubMed: 18478607]
43. Stojadinovic A, Peoples GE, Libutti SK, et al. Development of a clinical decision model for thyroid nodules. *BMC Surg*. 2009; 9:12. [PubMed: 19664278]

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript



**Figure 1.** Images of thyroid nodules that are entirely solid (Figures 1A–1D), and that demonstrate micro-calcifications (Figures 1A–1C)

Table 1

## Histology of included cancers

Type of Thyroid Cancer	Cancers		Size (cm)	
	No. (%) (n=105)	Mean [Range]	No. (%) (n=105)	Mean [Range]
Papillary	85 (.81)	1.7 [1.10–7.0]		
Follicular				
Minimally invasive	6 (.06)	3.8 [1.7–8.5]		
Poorly Differentiated	1 (.01)	3.0		
Not otherwise specified	3 (.03)	2.6 [1.3–5.0]		
Medullary	6 (.06)	2.6 [1.6–3.8]		
Anaplastic	1 (.01)	<i>a</i>		
Not specified	3 (.03)	<i>b</i>		

<sup>a</sup>Extensive involving thyroid, invading esophagus and carotid<sup>b</sup>Not available

**Table 2**

Selected characteristics of patients included in the study

	Patients Diagnosed with Thyroid		Patients not Diagnosed with Thyroid		P Value <sup>b</sup>
	No. (%) (n = 11 618)	Cancers per 100 <sup>a</sup>	Cases (%) (n = 96)	Controls (%) (n = 369)	
Age, Mean [SD]	51.6, SD [15.5]		44.9, SD [15.3]	45.6, SD [14.9]	.61
<b>Age Category</b>					
<40 yrs	3021 (26.0)	1.5	41 (42.7)	164 (44.1)	.71
41–60yrs	5158 (44.4)	0.8	36 (37.5)	132 (35.5)	
>60 yrs	3439 (29.6)	0.6	19 (19.8)	76 (20.4)	
<b>Sex</b>					
Female	9341 (80.4)	1.1	71 (74.0)	286 (77.5)	.46
Male	2277 (19.6)	1.5	25 (26.0)	83 (22.5)	
<b>Number of Nodules</b>					
0			3 <sup>c</sup> (3.1)	161 <sup>d</sup> (43.6)	< .001
1			43 <sup>c</sup> (44.8)	83 <sup>d</sup> (22.5)	
2			21 <sup>c</sup> (21.9)	63 <sup>d</sup> (17.1)	
3			12 <sup>c</sup> (12.5)	29 <sup>d</sup> (7.9)	
4			17 <sup>c</sup> (17.7)	33 <sup>d</sup> (8.9)	

<sup>a</sup>To calculate the cancer rates, we included all cancers (n=105) not just the 96 reviewed

<sup>b</sup>Difference between Cases / Controls.

<sup>c</sup>There were a total of 189 nodules diagnosed in the 96 cancer patients, including 102 cancer nodules and 87 benign nodules. There were 3 patients diagnosed with cancer in whom no nodules were identified.

<sup>d</sup>There were 428 benign nodules diagnosed in patients selected as controls.

**Table 3**

Distribution in the number of thyroid nodules by age, among cases and controls<sup>a</sup>

Number of Nodules	Patient Age						P Value
	< 20 years No. (%) (n = 2)	20–40 years No. (%) (n = 42)	41–60 years No. (%) (n = 33)	> 60 years No. (%) (n = 19)	All Ages No. (%) (n=96)		
(0)	1 (50.0)	0 (0.0)	1 (3.0)	1 (5.3)	3 (3.1)		
(1)	1 (50.0)	23 (54.8)	16 (48.5)	3 (15.8)	43 (44.8)	p = .001	
(2)	0 (0.0)	11 (26.2)	7 (21.2)	3 (15.8)	21 (21.9)		
(3)	0 (0.0)	3 (7.1)	5 (15.2)	4 (21.1)	12 (12.5)		
(4)	0 (0.0)	5 (11.9)	4 (12.1)	8 (42.1)	17 (17.7)		
Total number of nodules							189 (100)

Number of Nodules	Patient Age						P Value
	< 20 years No. (%) (n = 7)	20–40 years No. (%) (n = 155)	41–60 years No. (%) (n = 132)	> 60 years No. (%) (n = 75)	All Ages No. (%) (n=369)		
(0)	4 (57.1)	76 (49.0)	56 (42.4)	25 (33.3)	161 (43.6)	p = .22	
(1)	3 (42.9)	37 (23.9)	25 (18.9)	18 (24.0)	83 (22.5)		
(2)	0 (0.0)	24 (15.5)	21 (15.9)	18 (24.0)	63 (17.1)		
(3)	0 (0.0)	10 (6.5)	13 (9.8)	6 (8.0)	29 (7.9)		
(4)	0 (0.0)	8 (5.2)	17 (12.9)	8 (10.7)	33 (8.9)		
Total number of nodules							428 (100)

<sup>a</sup>The number in each cell reflects the number of patients with that number of thyroid nodules. Thus among the Cases, there were 23 patients age 20–40 that had a single thyroid nodule, and 11 patients in this age group with 2 thyroid nodules.

**Table 4**

Accuracy of individual ultrasound characteristics.

Characteristic Level	Cancers		Controls		Likelihood Ratio	
	Sensitivity <sup>d</sup> No. (%) (n = 102)	False Positive Rate <sup>d</sup> No. (%) (n = 515)	Characteristic Present LR (95% CI)	Characteristic Absent LR (95% CI)	Odds Ratio (95% CI)	
Microcalcifications	39 (.38)	28 (.05)	7.0 (6.0–8.2)	0.65 (0.56–0.76)	11.6 (6.5–20) <sup>b</sup>	
Coarse calcifications	13 (.13)	34 (.07)	1.9 (1.6–2.3)	0.93 (0.78–1.1)	2.1 (1.1–4.1) <sup>c</sup>	
Rim Calcifications	5 (.05)	23 (.04)	1.1 (0.91–1.3)	1.0 (0.82–1.2)	1.0 (0.41–2.6)	
Comet Tail Artifact	17 (.17)	77 (.15)	1.1 (0.93–1.3)	0.98 (0.82–1.2)	1.1 (0.60–2.0)	
Composition						
Entirely Solid	68 (.67)	220 (.43)	1.6 (1.4–1.8)	0.58 (0.51–0.66)	2.2 (1.4–3.5) <sup>b</sup>	
Entirely Cystic <sup>e</sup>	0 (.00)	37 (.07)	0.034 (0.028–0.041)	1.1 (0.88–1.3)	<i>e</i>	
Mixed Solid and Cystic	34 (.33)	248 (.48)	0.69 (0.58–0.82)	1.3 (1.1–1.5)	1	
Echogenicity						
Hypochoic to strap	16 (.16)	33 (.06)	2.4 (2.0–2.9)	0.89 (0.75–1.1)	2.9 (1.4–6.0) <sup>d</sup>	
Isochoic/hyper to strap	51 (.50)	198 (.38)	1.3 (1.1–1.5)	0.81 (0.70–0.94)	1.8 (1.1–2.9) <sup>c</sup>	
Isochoic to thyroid	30 (.29)	209 (.41)	0.72 (0.61–0.86)	1.2 (1.0–1.4)	1	
Hyperchoic to thyroid	4 (.04)	38 (.07)	0.53 (0.44–0.64)	1.0 (0.86–1.3)	0.75 (0.25–2.2)	
Central vascular flow	40 (.39)	136 (.26)	1.5 (1.3–1.7)	0.83 (0.71–0.97)	1.6 (1.0–2.6) <sup>c</sup>	
Peripheral vascular flow	49 (.48)	201 (.39)	1.2 (1.1–1.4)	0.85 (0.73–0.99)	1.3 (0.82–2.1)	
Halo	28 (.27)	116 (.23)	1.2 (1.0–1.4)	0.94 (0.79–1.1)	2.0 (1.3–3.1)	
Margins (ill defined/lobulated)	61 (.60)	212 (.41)	1.5 (1.3–1.7)	0.68 (0.60–0.78)	2.0 (1.3–3.1) <sup>c</sup>	
Shape (Taller than wide)	18 (.18)	42 (.08)	2.2 (1.8–2.6)	0.89 (0.75–1.1)	2.3 (1.2–4.3) <sup>c</sup>	



Characteristic Level	Cancers		Controls		Likelihood Ratio	
	Sensitivity <sup>d</sup> No. (%) (n = 102)	False Positive Rate <sup>d</sup> No. (%) (n = 515)	Characteristic Present LR (95% CI)	Characteristic Absent LR (95% CI)	Odds Ratio (95% CI)	
Size						
<=1cm	30 (.29)	248 (.48)	0.61 (0.51-0.73)	1.4	1	
1.01-2cm	38 (.37)	169 (.33)	1.1 (0.97-1.3)	0.94	1.9 (1.2-3.1) <sup>d</sup>	
>2cm	34 (.33)	97 (.19)	1.8 (1.5-2.1)	0.82	3.1 (1.8-5.2) <sup>b</sup>	

<sup>a</sup> Each characteristic was considered individually and was not adjusted for the presence of the other findings. The odds ratios were adjusted for multiple nodules within subjects, while the accuracy statistics were not adjusted for multiple nodules within subjects. The numbers may not sum as nodules in which the ultrasound characteristic was missing are included in the calculations of sensitivity and specificity but are not shown in the table.

<sup>b</sup> p < .001

<sup>c</sup> p < .05

<sup>d</sup> p < .01

<sup>e</sup> Entirely cystic includes cystic honeycomb. To calculate the likelihood ratio, 0.25 was added to zero cells. Because entirely cystic nodules had no cancer cases, an odds ratio could not be calculated.

Multivariate association between ultrasound characteristics and the likelihood of cancer, chosen using a stepwise logistic regression<sup>a</sup>

Table 5

	Multivariate Odds Ratios (95% CI)	P-Value
Microcalcifications	8.1 (3.8–17.3)	<.001
Nodule Size >2cm	3.6 (1.7–7.6)	.001
Composition <sup>b</sup> Entirely Solid	4.0 (1.7–9.2)	.001

<sup>a</sup> Only three specific ultrasound characteristics remained significant in the multivariate models.

**Table 6**

Accuracy of several definitions of an abnormal ultrasound

	Sensitivity (95%CI)		False Positive Rate (95%CI)		Likelihood Ratio (Predictive Values) (95% CI)		Risk of Cancer (Predictive Values) (95% CI)		Number Needed to Biopsy to Diagnose a Cancer (95%CI)
					If Characteristic(s)	If Characteristic(s)	If Characteristic(s)	If Characteristic(s)	
Any 1 Suspicious US Characteristic (Microcales or Solid or Size >2 cm)	0.88 (0.80-0.94)	0.44 (0.45-0.43)	2.0 (1.8-2.2)	0.21 (0.19-0.23)	0.018 (0.015-0.022)	0.002 (0.001-0.003)	0.002 (0.001-0.003)	56 (45-67)	
Solid	0.77 (0.68-0.85)	0.32 (0.31-0.33)	2.4 (2.1-2.7)	0.34 (0.30-0.38)	0.021 (0.017-0.027)	0.003 (0.002-0.005)	0.003 (0.002-0.005)	48 (37-59)	
Size>2cm	0.39 (0.30-0.49)	0.21 (0.20-0.21)	1.9 (1.6-2.3)	0.76 (0.65-0.90)	0.017 (0.012-0.024)	0.007 (0.005-0.009)	0.007 (0.005-0.009)	59 (42-83)	
(Microcalcification) or (Solid and >2 cm)	0.54 (0.44-0.64)	0.08 (0.08-0.09)	6.7 (5.8-7.7)	0.50 (0.43-0.57)	0.058 (0.044-0.075)	0.005 (0.003-0.006)	0.005 (0.003-0.006)	17 (13-23)	
Any 2 Suspicious US Characteristics (Microcales or Solid or Size >2 cm)	0.52 (0.42-0.62)	0.07 (0.07-0.08)	7.1 (6.2-8.2)	0.52 (0.45-0.60)	0.062 (0.047-0.080)	0.005 (0.004-0.006)	0.005 (0.004-0.006)	16 (13-21)	
Microcalcification	0.39 (0.30-0.49)	0.04 (0.04-0.05)	9.7 (8.3-11.4)	0.63 (0.54-0.74)	0.082 (0.059-0.110)	0.006 (0.004-0.007)	0.006 (0.004-0.007)	12 (9-17)	
All 3 Suspicious US Characteristics (Microcales and Solid and Size >2 cm)	0.07 (0.03-0.14)	0.00 (0.00-0.00)	28 <sup>a</sup> (23.0-34.0)	0.93 (0.77-1.10)	1.0 (0.826-1.0)	0.008 (0.007-0.010)	0.008 (0.007-0.010)	1 (1-1.2)	

<sup>a</sup> 0.25 was added to zero cells to calculate accuracy statistics

A patient's risk of thyroid cancer based on the ultrasound appearance of the thyroid gland and the characteristics of any nodules identified.

**Table 7**

Individual or Combination of Ultrasound	Cancers per 1000 patients (95% CI)
<b>Lowest Risk</b>	
Thyroid has only nodule(s) that are entirely cystic	0.32 (0.1–1.8)
Homogenous Thyroid	0.63 (0.1–1.9)
<b>Very low</b>	
Thyroid has no solid nodule(s), no nodule > 2 cm, and no nodule with Microcalcification	2 (1–3)
Thyroid has no nodule(s) with 2 or more suspicious characteristic	5 (4–6)
<b>Low</b>	
Thyroid has at least one nodule that is entirely solid or > 2cm or has microcalcification	18 (15–22)
<b>Moderate</b>	
Thyroid has at least one nodule with 2 suspicious characteristics	62 (47–80)
Thyroid has at least one nodule with Microcalcifications	82 (59–110)
Thyroid has at least one nodule with Microcalcifications or is both Solid and > 2cm	58 (44–75)
<b>High</b>	
Thyroid has at least one entirely solid nodule with microcalcification and a size greater than 2 cm	960 (803–991)