



Published in final edited form as:

Endocr Pract. 2017 April 02; 23(4): 442–450. doi:10.4158/EP161540.OR.

Gauging the Extent of Thyroidectomy for Indeterminate Thyroid Nodules: An Oncologic Perspective

David F. Schneider, MD, MS¹, Linda M. Cherney Stafford, MPH¹, Nicole Brys, MPH, Caprice C. Greenberg, MD, MPH¹, Courtney J. Balentine, MD, MPH¹, Dawn M. Elfenbein, MD, MPH², and Susan C. Pitt, MD, MPH¹

¹Department of Surgery and the Wisconsin Surgical Outcomes Research Program, University of Wisconsin, Madison, WI

²Department of Surgery, University of California, Irvine Medical Center, Orange, CA

Abstract

Introduction—Increasing emphasis is being placed on appropriateness of care and avoidance of over- and under-treatment. Indeterminate thyroid nodules (ITN), present a particular risk for this problem because cancer found via diagnostic lobectomy (DL) often requires a completion thyroidectomy (CT). However, initial total thyroidectomy (TT) for benign ITN results in lifelong thyroid hormone replacement. We sought to measure the accuracy and factors associated with the extent of initial thyroidectomy for ITN.

Methods—We queried a single institution thyroid surgery database for all adult patients undergoing an initial operation for ITN. Multivariate logistic regression identified factors associated with either oncologic under- or over-treatment at initial operation.

Results—There were 639 patients with ITN. The median age was 52 (range 18 – 93), 78.4% were female, and final pathology revealed a cancer > 1 cm in 24.7%. The most common cytology was follicular neoplasm (45.1%) followed by Hurthle cell neoplasm (20.2%). CT or initial oncologic under-treatment was required in 58 patients (9.3%). Excluding those with goiters, 19.0% were treated with total thyroidectomy for benign final pathology.

Multivariate analysis failed to identify any factor that independently predicted the need for CT. Female gender was associated with total thyroidectomy in benign disease (OR 2.1, 95% C.I. 1.0 – 4.5, $p = 0.05$). Age >45 predicted correct initial use of DL (OR 2.6, 95% C.I. 1.2 – 5.7, $p = 0.02$). Suspicious for PTC (OR 5.7, 95% C.I. 2.1 – 15.3, $p < 0.01$) and frozen section (OR 9.7, 95% C.I. 2.5 – 38.6, $p < 0.01$) were associated with oncologically appropriate initial TT.

The highest frequency of CT occurred in patients with follicular lesion of undetermined significance (11.6%). Total thyroidectomy for benign final pathology occurred most frequently in patients with a Hurthle cell neoplasm (24.8%).

Conclusions—In patients with ITN, nearly 30% received an inappropriate extent of initial thyroidectomy from an oncologic standpoint. Tools to preoperatively identify *both* benign and

malignant disease can assist in the complex decision-making to gauge the proper extent of initial surgery for ITN.

Introduction

Indeterminate thyroid nodules (ITN) represent a significant diagnostic and therapeutic challenge because up to 30% of thyroid nodules have indeterminate cytology but only 5–30% are found to be cancer (1). Since cytology alone cannot distinguish carcinomas from benign adenomas, surgery remains the only definitive means for distinguishing cancer from benign nodules (2). A diagnostic lobectomy remains the standard surgical treatment, and an estimated 75,000 lobectomies are performed annually in the U.S. (2–4). If the final pathology is carcinoma, then the patient must undergo a second surgery, or completion thyroidectomy, depending on the size and histopathologic features of the cancer. A second operation obviously carries additional risks, costs, and stress for the patient (3, 5, 6). Optimal oncologic treatment for thyroid cancer involves preoperative lymph node evaluation with ultrasound and compartment-oriented dissection of pathologically cancerous nodes. Hence, there are clear advantages to having a cancer diagnosis prior to proceeding with any surgery (5, 7–9).

Like many other conditions, increasing emphasis is being placed on appropriateness of care and avoidance of over- and under-treatment, for thyroid disease (10, 11), and thyroid surgery (12, 13). ITN, present a particular risk for this problem since an initial lobectomy for a cancer greater than 1 cm can be viewed as inadequate initial treatment, while an initial total thyroidectomy for benign disease might be considered excessive or unnecessary from an oncologic standpoint (4). In addition to higher risks of hypocalcemia and recurrent laryngeal nerve injury, patients also need for lifelong thyroid hormone replacement after total thyroidectomy.

Providers continue to use and evaluate several different tools to optimize the decision-making for the initial extent of surgery for ITN. Nomograms combine clinical, laboratory, imaging, and demographic data (14, 15). There also has been interest in analyzing ultrasound features to predict malignancy, but no single feature is sensitive or specific enough to preclude surgery (16–18). More recently, several molecular tests have been developed to classify ITN as either benign or malignant. These tests, while accurate in some settings, also carry a significant expense (19–21).

Ultimately, physicians and patients must synthesize the available information and decide on the appropriate extent of initial surgery when faced with an ITN. The advent of molecular testing prompted a plethora of studies examining how these tools impact the extent of surgery for ITN, but little data exists on how well physicians together with their patients decide on the extent of surgery without such adjuncts (3, 22–25). At our high volume thyroid center, we have historically not offered molecular testing to patients with ITN, providing an opportunity to examine how well treatment decisions are made without such testing. The purpose of this study is to measure the accuracy of and current decision-making associated with the extent of initial thyroidectomy for ITN.

Methods

This is a retrospective study of adult patients identified within the University of Wisconsin Endocrine Surgery Database who underwent an initial operation for ITN between May 1994 and April 2015. Under an IRB-approved protocol, the Endocrine Surgery Database prospectively collects patient and tumor features for patients undergoing thyroid, parathyroid, and adrenal surgery at our institution. For this study, patients were excluded if their records showed they were < 18 years of age, had cytology that was not indeterminate, underwent previous thyroid/neck surgery, diagnosed with hyperthyroidism (TSH<0.35 uIU/mL), or had too much missing data to perform any meaningful analysis. Those patients with ITN, meeting inclusion criteria (n=639) underwent further chart abstraction to confirm their diagnosis and collect additional data. ITN were defined by the following cytologic diagnoses: follicular neoplasm, suspicious for follicular neoplasm, atypia, atypia of undetermined significance (AUS), follicular lesion of undetermined significance (FLUS), suspicious for Hurthle cell neoplasm, Hurthle cell neoplasm, or suspicious for papillary carcinoma.

Additional data collected during the chart abstraction included patient demographics, ultrasound imaging features and sizes, histopathology results, rationale for surgery, extent of initial surgery, laboratory data, and need for completion thyroidectomy.

Defining Oncologically Appropriate Treatment

For the purposes of this study, the extent of initial surgery was labeled as either oncologically inadequate if they required a completion thyroidectomy or oncologically excessive if a total thyroidectomy was performed when the final histopathology did not reveal a cancer ≥ 1 cm in size. The 2009 (and earlier) ATA guidelines on thyroid nodular disease and differentiated thyroid cancer informed clinical decision-making regarding appropriate extent of surgery during the study period. All patients in this cohort were treated prior to the publication of the 2015 ATA guidelines (4). After lobectomy, patients were offered completion thyroidectomy if the final pathology revealed a cancer ≥ 1 cm in size (26). Goiter was defined as a thyroid with at any two of the following features: largest lobe dimension of at least 7 cm, greater than 1 nodule in the contralateral lobe, or thyroid weight > 25 grams. When patients with thyroid glands meeting the criteria for goiter underwent total thyroidectomy, this was *not* considered oncologically excessive treatment even if final histopathology was benign. Similarly, we conducted a separate analysis where patients taking levothyroxine preoperatively were not considered in the oncologically excessive group despite benign final pathology.

Statistical Analysis

In addition to standard descriptive statistics, multivariate logistic regression was used to identify patient and thyroid features associated with under- and over-treatment. Variables with a p value ≤ 0.2 were selected for inclusion in the multivariate models. Statistical analyses were performed using STATA SE v. 12 (StataCorp LP, College Station, TX).

Results

Patient and Nodule Features

There were 639 patients with ITN who met inclusion criteria for this study. The mean age was 52.0 years old, and 78.5% of patients were female (Table 1). Very few patients had a history of head and neck radiation exposure (4.1%) or a family history of differentiated thyroid cancer (14.4%, Table 1). Of the patients with a family history, only 35 patients (5.5%) had a first degree relative with thyroid cancer. Most patients did not suffer from hypothyroidism, with only 14.1% of patients taking levothyroxine preoperatively, and the median TSH was 1.5 ± 3.6 mIU/L (Table 1).

The median nodule size was 2.3 ± 1.5 cm (Table 1). The most common indeterminate cytology diagnosis was a follicular neoplasm (45.1%), followed by Hurthle cell neoplasm (20.3%), AUS/FLUS (18.9%), and suspicious for papillary thyroid cancer (15.7%, Table 1).

The majority of patients were initially treated with thyroid lobectomy (58.8%) and 24.7% of all ITN were found to contain a cancer (≥ 1 cm) on final pathology (Table 1). There were also an additional 70 microcarcinomas (< 1 cm).

Extent of Initial Surgery

Overall, only 9.3% of patients required a completion thyroidectomy after initial lobectomy, whereas 19% of patients underwent an initial total thyroidectomy for benign disease (Figure 1). Forty-seven patients treated with total thyroidectomy who also met the definition of “goiter” were not considered as oncologically excessive despite benign pathology. Taken together, this means that 28.3% of patients did not receive appropriate initial extent of surgery from an oncologic standpoint.

In a separate analysis, we also excluded any patients taking levothyroxine preoperatively (90 patients) from the oncologically excessive group. Excluding these patients resulted in a considerable reduction in the proportion of patients considered to have received excessive extent of initial surgery (20.4%). Similarly, we excluded those with a family history of thyroid cancer or a history of radiation exposure who received a total thyroidectomy (61 patients), leaving only 10.3% of patients receiving an oncologically excessive initial extent of surgery.

Final pathology revealed cancers of at least one centimeter in size in 18.8%, 13.2%, 28.9%, 78.0% of nodules with cytologic diagnoses of follicular neoplasm, Hurthle cell neoplasm, AUS/FLUS, and suspicious for PTC, respectively.

The most common cytologic finding in patients who required completion thyroidectomy was AUS/FLUS (14.3%). As expected, completion thyroidectomy was least common when the cytology was suspicious for PTC (5.1%, Figure 1). Total thyroidectomy when final pathology was benign occurred most commonly in patients with AUS/FLUS (25.4%) or a Hurthle cell neoplasm (24.7%, Figure 1). The AUS/FLUS category contained the most number of patients (39.7%) receiving an inappropriate extent of initial treatment (Figure 1).

Multivariate Analysis: Oncologically Appropriate Treatment

Next, we investigated the factors associated with an oncologically inappropriate extent of surgery at the first operation. Multivariate analysis failed to identify any significant factors that independently predicted the need for completion thyroidectomy. We also investigated the factors associated with total thyroidectomy for benign disease (in the absence of goiter). Female gender was significantly associated with receipt of oncologically excessive extent of initial surgery (OR 2.10, 95% C. I. 1.00 – 4.45, $p = 0.05$, Table 2). Intraoperative frozen section protected against the need for completion thyroidectomy (OR 0.19, 95% C. I. 0.06 – 0.66, $p < 0.01$, Table 2). None of the cytologic diagnoses were independently associated with receipt of oncologically excessive extent of surgery.

Surgeon's rationale

When total thyroidectomy was the initial extent of surgery, the reason or rationale for this decision was collected from the preoperative note. The most common reason cited was the surgeon's assessment of the cancer risk (54.2%) given the cytologic diagnosis and other clinical and imaging factors. The next most common reason was the presence of bilateral nodules (23.0%), followed by preoperative levothyroxine use (7.7%). This trend in reasoning followed in all cytologic categories (data not shown).

Multivariate Analysis: Appropriate Treatment

Similarly, the factors associated with the oncologically appropriate extent of initial surgery were also examined. For appropriate initial lobectomy (benign final pathology), only age over 45 years old was significant (OR 2.58, 95% C. I. 1.16 – 5.74, $p = 0.02$, Table 3). No other factors independently predicted appropriate initial lobectomy, including gender and each of the cytologic diagnoses.

In analyzing the factors associated with oncologically correct initial total thyroidectomy, frozen section (OR 9.71, 95% C. I. 2.45 – 38.56, $p < 0.01$) and suspicious for papillary thyroid cancer cytology (OR 5.68, 95% C. I. 2.10 – 15.33, $p < 0.01$) were significant in multivariate analysis (Table 4). None of the other cytologic diagnoses independently predicted correct initial total thyroidectomy. Consistent with our analysis (Table 2), female gender was not associated with correct initial total thyroidectomy (OR 0.35, 95% C. I. 0.13 – 0.98, $p = 0.04$, Table 4).

Since frozen section was associated with the oncologically appropriate extent of initial surgery, we examined the frequency of its use further. Overall, frozen section was utilized in 190 cases (29.7%). The most frequent use of frozen section occurred with a cytologic diagnosis of suspicious for PTC (63.3%), followed by AUS/FLUS (38.1%).

Discussion

In this large series evaluating the initial extent of surgery for ITN, nearly 30% of patients received an inappropriate extent of surgery from a purely oncologic standpoint. At this single, high volume center oncologically excessive surgery (19%) was far more common than oncologically inadequate treatment (9.3%). The AUS/FLUS category proved the most

problematic for surgeons to determine the correct amount of initial surgery, with nearly 40% of patient receiving an inappropriate extent of initial surgery. The higher institutional rate of cancer in the AUS/FLUS category (28.9%) likely pushed surgeons toward more total thyroidectomies for patients with Bethesda III cytology. Similarly, the high cancer rate (78%) in the suspicious for PTC category also impacted decision-making, shifting the initial extent of treatment toward total thyroidectomy. As expected, the suspicious for papillary thyroid cancer carried the lowest rates of oncologically inappropriate treatment, likely due to the utility of frozen section for this category. The quality of frozen section varies by institution, and may not be appropriate in all settings.

We currently lack data on patient and provider decision-making for ITN. Many factors likely play into the decision-making regarding the initial extent of surgery for ITN. Contralateral nodules, overall gland size, and compressive symptoms are among the reasons for choosing total thyroidectomy over lobectomy. Perhaps the most important factor is the patient's preferences and priorities regarding the need for a second operation, surgical risk, or the need for lifelong thyroid hormone replacement. Therefore, some may find our application of "appropriateness" label unsuitable for this problem. Amidst the changing healthcare policy landscape and increasing interest in defining "appropriateness of care" for certain procedures, thyroid practitioners must begin to evaluate the accuracy of their decision-making because it impacts patients' quality of life and costs (27–29). This becomes especially important for ITN with the popularization of various adjuncts designed to assist in the decision-making.

A variety of tools to help surgeons determine the proper extent of surgery for ITN now exist. Many of these studies combine clinical and imaging features in order to predict malignancy (30–32). Most of these models rely heavily on ultrasound features such as calcifications, irregular borders, hypoechoogenicity, and nodule size. For example, Mendez et al. studied surgeon performed ultrasound, and found that at least two adverse nodule features correlated with the risk of malignancy, and could guide the extent of initial thyroidectomy in patients with ITN (33). In their series, even more patients (66%) were treated with a total thyroidectomy as the initial extent of surgery compared to only 42% in this study. However, they did not report on the overall accuracy of their treatment decisions (33).

Similarly, numerous imaging classifiers exist to distinguish benign from malignant nodules (16, 34–37). The majority of these studies utilize ultrasound features and statistical models to select nodules with the highest cancer risk for FNA biopsy. These studies examine all nodules, not just ITN. Since 95% of all nodules are benign, identifying malignancy among the indeterminate subset is likely a more difficult task. In the present study, all surgeons used ultrasound features in their decision-making about the initial extent of surgery. However, surgeons rarely cited ultrasound features (6.1%) as the overriding reason for the initial extent of surgery. Instead, this may have been documented as the overall cancer risk, with ultrasound features not explicitly stated. This reflects the challenge of identifying the surgeons' rationale by retrospectively coding clinic notes, particularly for such a complex decision-making process.

More recently, clinicians have turned to molecular markers for determining the extent of initial surgery for ITN. One strategy uses genetic mutations that, when present, identify malignant ITN thereby prompting total thyroidectomy up front and avoiding the need for completion thyroidectomies. In an analysis of one such mutational panel, those undergoing cytogenetic testing preoperatively were 2.5 times more likely to undergo total thyroidectomy when a clinically significant cancer was present on pathology compared to the group not undergoing mutational testing (38). This rule-in strategy may not prove as helpful for high volume surgeons – in the current study less than 10% of patients required completion thyroidectomy.

Another strategy for using molecular markers in ITN is to identify a genetic signature for benign nodules. The commercially available Afirma gene expression classifier rules out malignancy so that benign nodules can simply undergo surveillance rather than any type of thyroid surgery. This strategy addresses the issue of oncologically excessive treatment as reported here. The initial study reported a negative (benign) predictive value of 85–95% depending on the exact cytology (19). However, this gene expression classifier is expensive, and the cost-effectiveness depends on the institutional rates of malignancy for each cytologic category (21). Nonetheless, this rule-out strategy may prevent an oncologically excessive extent of initial surgery as we found in this study.

The latest version of the American Thyroid Association's (ATA) Guidelines on Thyroid Nodules and Differentiated Thyroid Cancer stress the importance of counseling the patients about the limitations of molecular testing for ITN (4). In discussing the principles of employing molecular testing, these guidelines also underscore the pretest probability of malignancy, feasibility of performing molecular testing, and patient preferences (4). Both the Disease State Commentary from the American Association of Clinical Endocrinologist's Thyroid Scientific Committee and a consensus statement from the ATA Surgical Affairs Committee indicated that molecular testing should not replace clinical judgment (39, 40). The results reported here represent the clinical judgment of high volume clinicians in the absence of adjuncts like molecular testing. This judgment incorporates the pre-test probability of malignancy based on the cytology, history, physical exam, and ultrasound features. The assumption implicit in many of the molecular testing strategies for ITN, is that the surgeon will perform a diagnostic lobectomy based on the cytology result alone. As our results demonstrate, this is not always the case.

Here, we report that for high volume surgeons, completion thyroidectomy is less frequent, and strategies to rule-out malignancy may become more useful in that setting. There is a paucity of data on the frequency of completion thyroidectomy, or the accuracy of decision-making surrounding the initial extent of surgery for ITN. Here we report a large series of indeterminate nodules, and the oncologic accuracy of decision-making around the extent of initial surgery. Certainly, these data are important as hospital systems decide if and how to employ molecular testing or other forms of clinical decision-support for ITN. Each of these adjuncts carries a unique accuracy and cost profile. Therefore, one must understand the clinician's baseline accuracy before superimposing additional decision-support. For example, at our institution, tools to identify benign nodules among those with AUS/FLUS cytology would be of greatest value given our high degree of oncologically excessive

treatment. For this reason, we have begun utilizing frozen section for selected AUS/FLUS nodules.

While the cohort of patients presented here were all treated under the 2009 ATA guidelines, the most recent version of these guidelines indicate that lobectomy is adequate treatment for low risk, T2 (4 cm) clinically N0 patients. Under these recommendations, this would likely increase the degree of oncologically excessive treatment seen in this cohort(4). Regardless of how the categories change, we present a framework for objectively evaluating extent of initial thyroidectomy from a purely oncologic standpoint and stress the importance of knowing individual or institutional rates of oncologically inappropriate treatment to tailor the use of decision support tools or diagnostic adjuncts.

Patient preference clearly plays into the decision-making on the extent of initial surgery for ITN, and the 2015 ATA guidelines also stress this point (4). Patients have varying priorities regarding the need for thyroid hormone replacement, surgical risk, and the need for a second operation. Patients potentially drove much of the decision-making reported here, but it was seldom cited (0.47%) as the overriding reason for choosing total thyroidectomy over lobectomy. This finding likely reflects clinical documentation and the challenges of capturing this type of data retrospectively.

Aside from its retrospective nature, this study has other limitations. This is a single institution analysis reporting on the accuracy of decision-making for high volume surgeons, so the data are not necessarily generalizable. However, the literature lacks data on accuracy of the initial extent of surgery for ITN, and these results provide a useful reference point for designing decision-support tools or employing adjuncts such as molecular testing.

Although the “accuracy” of decision-making was analyzed from a purely oncologic standpoint, we recognize that other factors contributed to treatment decisions. Contralateral nodules, the size of other nodules, overall gland size, and symptoms likely influence patients and providers toward up front total thyroidectomy.

Although this reasoning is difficult to capture in a retrospective study, we accounted for this by excluding goiters from the “over-treatment” designation. The literature lacks standard definitions of goiter, and it becomes difficult to retrospectively apply any potential definitions of goiter. Additionally, we performed separate analyses, excluding patients taking levothyroxine preoperatively or those with a family history or history of radiation exposure. Excluding these types of patients did reduce the portion of patients receiving total thyroidectomy for benign final pathology. However, some degree of mismatch still remained. Regardless of the absolute numbers, it is important for providers to understand their own performance as pre-test probability for decision support or additional diagnostic testing.

Conclusions

Of patients with ITN, 30% received an inappropriate extent of initial surgery from an oncologic standpoint. At a high volume center, total thyroidectomy for benign final pathology was more frequent than the need for completion thyroidectomy. Tools to

preoperatively identify *both* malignant *and* benign disease can potentially assist in clinical decision-making to gauge the initial extent of thyroidectomy for ITN, depending on the institution-specific needs. Prospective examination of patients' preferences and decisional needs for ITN treatment will compliment these tools and expand our decision-making capacity.

Acknowledgments

This work was supported by NIH UL1TR000427 and NIH KL2TR000428.

References

1. Yassa L, Cibas ES, Benson CB, et al. Long-term assessment of a multidisciplinary approach to thyroid nodule diagnostic evaluation. *Cancer*. 2007; 111:508–516. [PubMed: 17999413]
2. Baloch ZW, LiVolsi VA, Asa SL, et al. Diagnostic terminology and morphologic criteria for cytologic diagnosis of thyroid lesions: a synopsis of the National Cancer Institute Thyroid Fine-Needle Aspiration State of the Science Conference. *Diagn Cytopathol*. 2008; 36:425–437. [PubMed: 18478609]
3. Yip L, Farris C, Kabaker AS, et al. Cost impact of molecular testing for indeterminate thyroid nodule fine-needle aspiration biopsies. *The Journal of clinical endocrinology and metabolism*. 2012; 97:1905–1912. [PubMed: 22419727]
4. Haugen BRM, Alexander EK, Bible KC, et al. 2015 American Thyroid Association Management Guidelines for Adult Patients with Thyroid Nodules and Differentiated Thyroid Cancer. *Thyroid*. 2015
5. Starmer H, Noureldine SI, Ozgursoy OB, Tufano RP. Voice outcomes following reoperative central neck dissection for recurrent/persistent thyroid cancer. *Laryngoscope*. 2015; 125:2621–2625. [PubMed: 26153252]
6. Pai SI, Tufano RP. Reoperation for recurrent/persistent well-differentiated thyroid cancer. *Otolaryngol Clin North Am*. 2010; 43:353–363. ix. [PubMed: 20510718]
7. Oltmann SC, Schneider DF, Levenson G, Sivashanmugam T, Chen H, Sippel RS. Radioactive Iodine Remnant Uptake After Completion Thyroidectomy: Not Such a Complete Cancer Operation. *Annals of surgical oncology*. 2013
8. Roy R, Kouniavsky G, Venkat R, et al. The role of preoperative neck ultrasounds to assess lymph nodes in patients with suspicious or indeterminate thyroid nodules. *J Surg Oncol*. 2011
9. Oltmann SC, Schneider DF, Chen H, Sippel RS. All thyroid ultrasound evaluations are not equal: sonographers specialized in thyroid cancer correctly label clinical N0 disease in well differentiated thyroid cancer. *Annals of surgical oncology*. 2015; 22:422–428. [PubMed: 25234019]
10. Toubert ME, Chevret S, Cassinat B, Schlageter MH, Beressi JP, Rain JD. From guidelines to hospital practice: reducing inappropriate ordering of thyroid hormone and antibody tests. *European journal of endocrinology/European Federation of Endocrine Societies*. 2000; 142:605–610.
11. Liel Y, Fraenkel N. Brief report: Use and misuse of thyroid ultrasound in the initial workup of patients with suspected thyroid problems referred by primary care physicians to an endocrine clinic. *J Gen Intern Med*. 2005; 20:766–768. [PubMed: 16050890]
12. Tufano RP, Noureldine SI, Angelos P. Incidental thyroid nodules and thyroid cancer: considerations before determining management. *JAMA Otolaryngol Head Neck Surg*. 2015; 141:566–572. [PubMed: 25928353]
13. Adam MA, Pura J, Gu L, et al. Extent of surgery for papillary thyroid cancer is not associated with survival: an analysis of 61,775 patients. *Annals of surgery*. 2014; 260:601–605. discussion 605–607. [PubMed: 25203876]
14. 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; 156:967–970. [PubMed: 25074360]

15. Rago T, Di Coscio G, Basolo F, et al. Combined clinical, thyroid ultrasound and cytological features help to predict thyroid malignancy in follicular and Hupsilonnrthle cell thyroid lesions: results from a series of 505 consecutive patients. *Clin Endocrinol (Oxf)*. 2007; 66:13–20. [PubMed: 17201796]
16. Liu YI, Kamaya A, Desser TS, Rubin DL. A Bayesian classifier for differentiating benign versus malignant thyroid nodules using sonographic features. *AMIA Annual Symposium proceedings/ AMIA Symposium AMIA Symposium*. 2008:419–423. [PubMed: 18999209]
17. Frates MC, Benson CB, Doubilet PM, et al. Prevalence and distribution of carcinoma in patients with solitary and multiple thyroid nodules on sonography. *The Journal of clinical endocrinology and metabolism*. 2006; 91:3411–3417. [PubMed: 16835280]
18. Seiberling KA, Dutra JC, Grant T, Bajramovic S. Role of intrathyroidal calcifications detected on ultrasound as a marker of malignancy. *Laryngoscope*. 2004; 114:1753–1757. [PubMed: 15454766]
19. Alexander EK, Kennedy GC, Baloch ZW, et al. Preoperative diagnosis of benign thyroid nodules with indeterminate cytology. *N Engl J Med*. 2012; 367:705–715. [PubMed: 22731672]
20. Nikiforov YE, Ohori NP, Hodak SP, et al. Impact of mutational testing on the diagnosis and management of patients with cytologically indeterminate thyroid nodules: a prospective analysis of 1056 FNA samples. *The Journal of clinical endocrinology and metabolism*. 2011; 96:3390–3397. [PubMed: 21880806]
21. Wu JX, Lam R, Levin M, Rao J, Sullivan PS, Yeh MW. Effect of malignancy rates on cost-effectiveness of routine gene expression classifier testing for indeterminate thyroid nodules. *Surgery*. 2016; 159:118–129. [PubMed: 26435428]
22. Gill MS, Nayan S, Kocovski L, et al. Local molecular analysis of indeterminate thyroid nodules. *J Otolaryngol Head Neck Surg*. 2015; 44:52. [PubMed: 26621130]
23. Noureldine SI, Olson MT, Agrawal N, Prescott JD, Zeiger MA, Tufano RP. Effect of Gene Expression Classifier Molecular Testing on the Surgical Decision-Making Process for Patients With Thyroid Nodules. *JAMA Otolaryngol Head Neck Surg*. 2015; 141:1082–1088. [PubMed: 26606459]
24. Panebianco F, Mazzanti C, Tomei S, et al. The combination of four molecular markers improves thyroid cancer cytologic diagnosis and patient management. *BMC Cancer*. 2015; 15:918. [PubMed: 26581891]
25. Nikiforov YE, Carty SE, Chiosea SI, et al. Impact of the Multi-Gene ThyroSeq Next-Generation Sequencing Assay on Cancer Diagnosis in Thyroid Nodules with Atypia of Undetermined Significance/Follicular Lesion of Undetermined Significance Cytology. *Thyroid*. 2015; 25:1217–1223. [PubMed: 26356635]
26. Cooper DS, Doherty GM, Haugen BR, et al. Revised American Thyroid Association management guidelines for patients with thyroid nodules and differentiated thyroid cancer. *Thyroid*. 2009; 19:1167–1214. [PubMed: 19860577]
27. Bernstein SJ, Brorsson B, Aberg T, Emanuelsson H, Brook RH, Werko L. Appropriateness of referral of coronary angiography patients in Sweden. *SECOR/SBU Project Group. Heart*. 1999; 81:470–477. [PubMed: 10212163]
28. Lawson EH, Gibbons MM, Ko CY, Shekelle PG. The appropriateness method has acceptable reliability and validity for assessing overuse and underuse of surgical procedures. *J Clin Epidemiol*. 2012; 65:1133–1143. [PubMed: 23017632]
29. Morden NE, Colla CH, Sequist TD, Rosenthal MB. Choosing wisely—the politics and economics of labeling low-value services. *N Engl J Med*. 2014; 370:589–592. [PubMed: 24450859]
30. Maia FF, Matos PS, Pavin EJ, Vassallo J, Zantut-Wittmann DE. Value of ultrasound and cytological classification system to predict the malignancy of thyroid nodules with indeterminate cytology. *Endocrine pathology*. 2011; 22:66–73. [PubMed: 21547508]
31. Macias CA, Arumugam D, Arlow RL, et al. A risk model to determine surgical treatment in patients with thyroid nodules with indeterminate cytology. *Annals of surgical oncology*. 2015; 22:1527–1532. [PubMed: 25388058]
32. Calo PG, Medas F, Santa Cruz R, et al. Follicular nodules (Thy3) of the thyroid: is total thyroidectomy the best option? *BMC Surg*. 2014; 14:12. [PubMed: 24597765]

33. Mendez W, Rodgers SE, Lew JI, Montano R, Solorzano CC. Role of surgeon-performed ultrasound in predicting malignancy in patients with indeterminate thyroid nodules. *Annals of surgical oncology*. 2008; 15:2487–2492. [PubMed: 18622644]
34. Lim KJ, Choi CS, Yoon DY, et al. Computer-aided diagnosis for the differentiation of malignant from benign thyroid nodules on ultrasonography. *Acad Radiol*. 2008; 15:853–858. [PubMed: 18572120]
35. Dutta S, Thaha MA, Smith DM. Do sonographic and cytological features predict malignancy in cytologically indeterminate thyroid nodules? *Ann R Coll Surg Engl*. 2011; 93:361–364. [PubMed: 21943458]
36. Smith-Bindman R, Lebda P, Feldstein VA, et al. Risk of thyroid cancer based on thyroid ultrasound imaging characteristics: results of a population-based study. *JAMA Intern Med*. 2013; 173:1788–1796. [PubMed: 23978950]
37. Iakovidis DK, Keramidas EG, Maroulis D. Fusion of fuzzy statistical distributions for classification of thyroid ultrasound patterns. *Artif Intell Med*. 2010; 50:33–41. [PubMed: 20427164]
38. Yip L, Wharry LI, Armstrong MJ, et al. A clinical algorithm for fine-needle aspiration molecular testing effectively guides the appropriate extent of initial thyroidectomy. *Annals of surgery*. 2014; 260:163–168. [PubMed: 24901361]
39. Bernet V, Hupart KH, Parangi S, Woeber KA. AACE/ACE disease state commentary: molecular diagnostic testing of thyroid nodules with indeterminate cytopathology. *Endocr Pract*. 2014; 20:360–363. [PubMed: 24727662]
40. Ferris RL, Baloch Z, Bernet V, et al. American Thyroid Association Statement on Surgical Application of Molecular Profiling for Thyroid Nodules: Current Impact on Perioperative Decision Making. *Thyroid*. 2015; 25:760–768. [PubMed: 26058403]

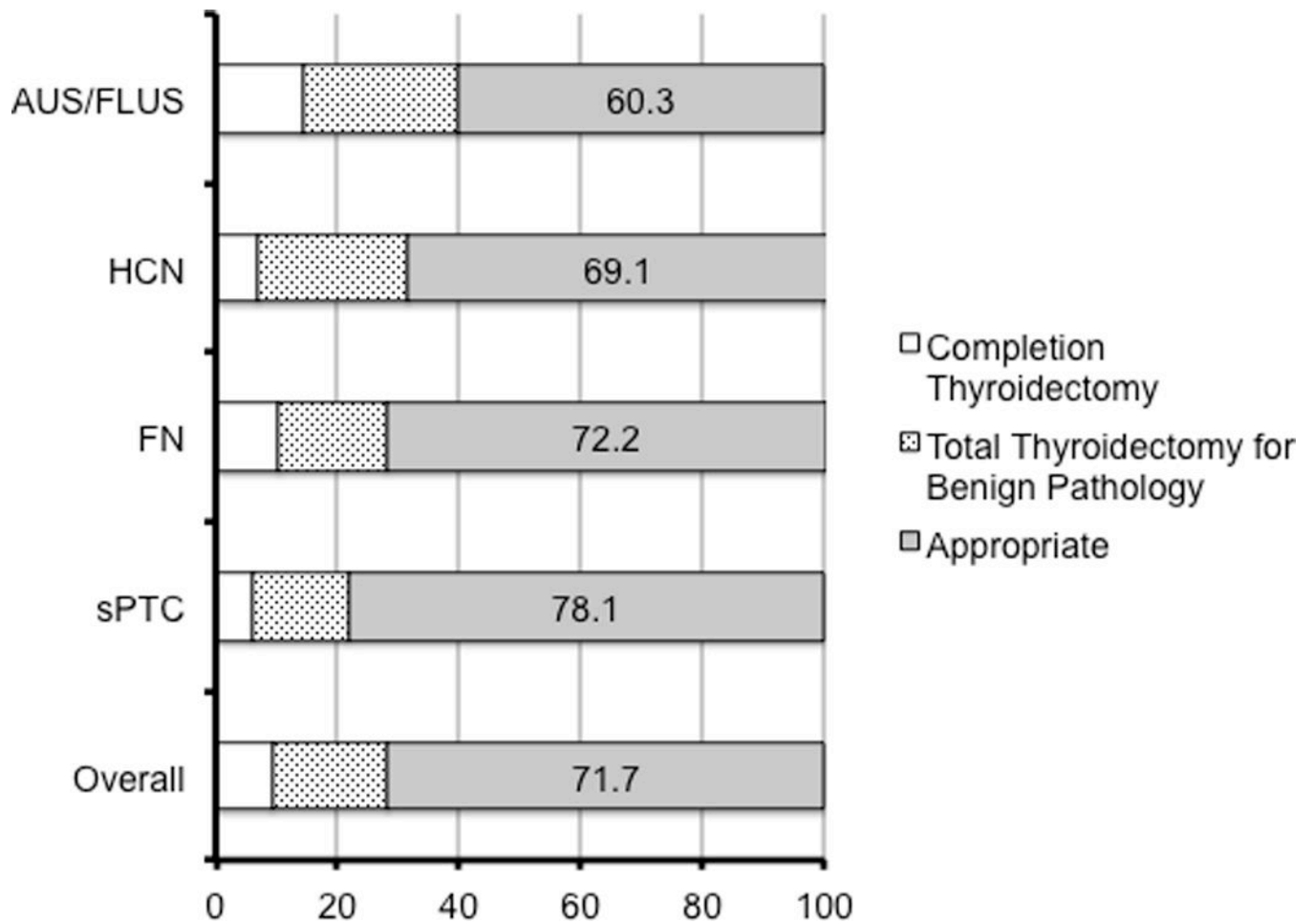


Figure 1. Accuracy and the Extent of Initial Thyroidectomy

Bar charts show the percentage of patients who required completion thyroidectomy needed, were treated with total thyroidectomy for benign disease, or the appropriate extent of initial surgery are shown for each cytologic category. Numbers indicate the percentage of patients receiving the appropriate extent of initial treatment for each category. Patients with goiter receiving total thyroidectomy were not considered to be over-treated, and cancers ≤ 1 cm were not considered clinically significant cancers. *AUS* = atypia of undetermined significance; *FLUS* = follicular lesion of undetermined significance; *HCN* = Hurthle cell neoplasm; *FN* = follicular neoplasm; *sPTC* = suspicious for papillary thyroid cancer.

TABLE 1**Patient and Nodule Characteristics (n = 639)**

Variable	Number (%)
Age (years)	52.0 ± 14.7
Female gender	501 (78.5)
Radiation history	26 (4.1)
Family history of thyroid cancer	92 (14.4)
Median TSH (U/mL)	1.5 ± 3.6
Taking levothyroxine	90 (14.1)
Median nodule size (cm)	2.3 ± 1.5
Cytology	
Follicular neoplasm	288 (45.1)
Hurthle cell neoplasm	130 (20.3)
AUS/FLUS	121 (18.9)
Suspicious for PTC	100 (15.7)
Extent of Surgery	
Lobectomy	376 (58.8)
Total thyroidectomy	263 (41.2)
Final Pathology	
Total Cancers	228 (35.7)
Cancer ≥ 1 cm	158 (24.7)
Cancer < 1 cm	70 (10.9)

Data are represented as either the mean or median value ± the standard deviation for continuous variables or the number of patients with the corresponding percentage for categorical variables. *U* = units, *mL* = milliliters, *cm* = centimeters, *AUS* = atypia of undetermined significance, *FLUS* = follicular lesion of undetermined significance, *PTC* = papillary thyroid cancer

TABLE 2

Multivariate Analysis: Total Thyroidectomy for Benign Pathology

Variable	OR	95% C. I.	p value
Age > 45 years old	0.91	0.53 – 1.59	0.75
Female gender	2.10	1.00 – 4.45	0.05
Family history of thyroid cancer	0.51	0.25 – 1.07	0.07
TSH < 2.5 (U/mL)	0.83	0.46 – 1.52	0.55
Size > 3 cm	0.65	0.37 – 1.15	0.16
Frozen Section	0.19	0.06 – 0.66	<0.01

Multivariate logistic regression was performed with variables carrying a p value < 0.2 from univariate analysis. Variables shown in **bold** were significantly associated with over-treatment. Also included in the model was each cytologic diagnosis, except for AUS/FLUS which was dropped due to collinearity. None of the cytologic diagnoses were significant. *U* = units, *mL* = milliliters, *cm* = centimeters.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

TABLE 3

Multivariate Analysis: Oncologically Appropriate Initial Lobectomy

Variable	OR	95% C. I.	p value
Age > 45 years old	2.58	1.16 – 5.74	0.02
Family history of thyroid cancer	0.49	0.21 – 1.15	0.10
History of radiation exposure	0.23	0.04 – 1.34	0.13
Female gender	1.44	0.58 – 3.57	0.43
TSH < 2.5 (U/mL)	0.48	0.21 – 1.10	0.08
Size < 3 cm	0.96	0.44 – 2.10	0.91
Surgeon performed ultrasound	0.81	0.35 – 1.85	0.62

Multivariate logistic regression was performed with variables carrying a p value < 0.2 from univariate analysis. Variables shown in bold were significantly associated with oncologically correct initial lobectomy. Also included in the model was each cytologic diagnosis, except for AUS/FLUS which was dropped due to collinearity. None of the cytologic diagnoses were significant. *U*= units, *mL* = milliliters, *cm* = centimeters.

TABLE 4

Multivariate Analysis: Oncologically Appropriate Initial Total Thyroidectomy

Variable	OR	95% C. I.	p value
Age > 45 years old	1.44	0.63 – 3.29	0.38
Female gender	0.35	0.13 – 0.98	0.04
Suspicious for papillary thyroid cancer	5.68	2.10 – 15.33	<0.01
Family history of thyroid cancer	2.66	0.99 – 7.10	0.06
TSH < 2.5 (U/mL)	1.82	0.74 – 4.48	0.19
Frozen section	9.71	2.45 – 38.56	<0.01

Multivariate logistic regression was performed with variables carrying a p value < 0.2 from univariate analysis. Variables shown in bold were significantly associated with oncologically correct initial total thyroidectomy. In addition to suspicious for papillary thyroid cancer, the other cytologic diagnosis were also included in the model except for AUS/FLUS which was dropped due to collinearity. None of the other cytologic diagnoses were significant. *U* = units, *mL* = millimeters.