The Cost-effectiveness of Iodine 131 Scintigraphy, Ultrasonography, and Fine-Needle Aspiration Biopsy in the Initial Diagnosis of Solitary Thyroid Nodules

Ayesha N. Khalid, MD; Christopher S. Hollenbeak, PhD; Sadeq A. Quraishi, MD, MHA; Chris Y. Fan, MD; Brendan C. Stack, Jr, MD

Objective: To compare the cost-effectiveness of fine-needle aspiration biopsy, iodine 131 scintigraphy, and ultrasonography for the initial diagnostic workup of a solitary palpable thyroid nodule.

Design: A deterministic cost-effectiveness analysis was conducted using a decision tree to model the diagnostic strategies.

Setting: A single, mid-Atlantic academic medical center.

Main Outcome Measures: Expected costs, expected number of cases correctly diagnosed, and incremental cost per additional case correctly diagnosed.

Results: Relative to the routine use of fine-needle aspiration biopsy, the incremental cost per case correctly diagnosed is $24,554 for the iodine 131 scintigraphy strategy and $1,212 for the ultrasound strategy.

Conclusions: A diagnostic strategy using initial fine-needle aspiration biopsy for palpable thyroid nodules was found to be cost-effective compared with the other approaches as long as a payer’s willingness to pay for an additional correct diagnosis is less than $1,212. Prospective studies are needed to validate these findings in clinical practice.

Arch Otolaryngol Head Neck Surg. 2006;132:244-250

Thyroid nodules can be detected by palpation in approximately 4% to 7% of all patients. They are found 8 times more often in women than in men, and their prevalence increases with advancing age. In contrast to the high prevalence of nodular thyroid disease, the presence of thyroid carcinoma is rare, accounting for roughly 5% of all palpable nodules.

CME course available at www.archoto.com

In the United States, thyroid carcinoma represents approximately 1% to 2% of all malignancies, accounting for more than 25,000 new cancers each year and resulting in more than 1,400 deaths annually. Thyroid carcinoma occurs more frequently in females, who have a greater prevalence of nodular thyroid disease; however, nodular thyroid disease is twice as likely to be malignant in men than in women. Typically, the distribution of thyroid carcinoma across ages is bimodal, appearing most often in patients younger than 30 years or older than 60 years. Older individuals are also more likely to be diagnosed with more aggressive thyroid carcinomas, have more extensive disease, and have a higher likelihood of mortality than their younger counterparts.

Patients are usually asymptomatic when thyroid nodules are first identified, as most are discovered incidentally. Observation alone is often sufficient for the majority of benign lesions, but surgical excision is needed for malignant lesions or benign nodules large enough to cause symptoms of compression. Consequently, many clinicians favor an early diagnostic approach to avoid missing a potentially malignant condition, while others favor a more conservative approach because only a relatively small number of patients with nodules will ultimately be diagnosed as having malignant disease. Therefore, accurate diagnostic information may help to avoid the morbidity and/or mortality of prophylactic surgery in patients with benign disease and observation in patients with malignant disease. Despite studies supporting the cost-effectiveness of fine-needle aspiration biopsy (FNAB) as the diagnostic test of choice, iodine 131 (I-131) scintigraphy and ultrasonography (US) continue to be used by front-
line providers as primary diagnostic tools in the management of patients with nodular thyroid disease. Justifications for the continued use of these alternative diagnostic strategies usually range from historical practice patterns within institutions to relative ease of performing the test (in the case of US, which is proliferating in the outpatient setting), to faster turnaround time for results (when compared with waiting for FNAB pathology reports). Therefore, the purpose of this study was to develop a model to compare the cost-effectiveness of FNAB, I-131 scintigraphy, and US in the initial diagnostic strategy for solitary palpable thyroid nodules. We used decision analysis to model algorithms because this technique has been shown to provide valuable information when clinical decisions must be made under circumstances of uncertainty and/or limited resources. The uniqueness of this tool is that although it can be used to find meaningful patterns in large, complex quantitative databases, it also allows qualitative analyses.

**METHODS**

**DECISION ANALYSIS MODEL**

To study the economic implications of alternative front-line diagnostic methods for solitary palpable nodules, we developed a decision analysis tree to assess the cost-effectiveness of 3 diagnostic strategies in an effort to simulate clinical decision making. In our decision analysis model, the first decision node, depicted as a square, represents a solitary nodule palpated during a clinical encounter. The clinician then chooses one of 3 diagnostic methods: FNAB, I-131 scintigraphy, or US. If an FNAB strategy is pursued, the results of biopsy will either be “cystic” or “noncystic.” Inconclusive FNAB results will prompt the physician to proceed with a hemithyroidectomy to obtain a histopathologic diagnosis. Consequently, at the terminal node, the pathologic specimen will provide a diagnosis of either benign or malignant, which is assumed to be correct. Similarly, if the results of FNAB are conclusive, the biopsy specimen will either be malignant or benign, at which point the diagnosis will have been obtained.

If the clinician chooses an I-131 scintigraphy approach at the decision node, the nodule will be classified as hyperfunctioning (hot), isofunctioning (warm), or hypofunctioning (cold). If the result is hyperfunctioning, the nodule is likely benign and observed or otherwise treated medically. Alternatively, if the I-131 scintigraphy result classifies the nodule as isofunctional, the physician will proceed with the FNAB strategy described above. Similarly, if the I-131 scintigraphy result classifies the nodule as hypofunctional, the physician will also proceed with the FNAB strategy.

Finally, the clinician may use US as an initial diagnostic technique. The findings of US will classify the nodule as either “cystic” or “noncystic.” If a diagnosis of cystic is obtained, then the lesion is treated as benign. In the small proportion of patients whose cystic lesions are malignant, a false-negative result will have been obtained. On the other hand, if the findings of US classify the nodule as noncystic, it is assumed that the clinician will proceed with the FNAB strategy described previously, which will provide a conclusive cytopathologic diagnosis.

**MODEL PARAMETERS**

Model parameters, which included probability, sensitivity, specificity, and prevalence of malignancy among patients with solitary thyroid nodules, were estimated from a review of the literature (Table 1). Studies of FNAB represent a large and rapidly expanding body of literature, and only the subset related to the diagnosis of thyroid diseases was used in this study. A similar approach was taken for the US and I-131 scintigraphy strategies. The sensitivity and specificity were estimated using a weighted average of values reported in the literature, in which the weights were proportional to the number of patients in the study. The probabilities of true- and false-positive diagnoses and true- and false-negative diagnoses for each of the diagnostic strategies were computed using Bayes theorem.

Unit costs for FNAB, US, I-131 scintigraphy, and hemithyroidectomy were obtained from samples of patients receiving these tests and procedures at a single, mid-Atlantic academic medical center. Fully loaded operating costs were estimated by the hospital’s cost-accounting database using a ratio of costs-to-charges method. Consequently, these are estimates of the hospital’s cost for these tests and procedures and are not billed charges.

**BASELINE COST-EFFECTIVENESS ANALYSIS**

Cost-effectiveness analysis asks whether one intervention provides value for money compared with another intervention. It is used when alternative interventions have differing implications for costs and patient outcomes. The cost-effectiveness of 2 treatment strategies (A and B) would be summarized with an incremental cost-effectiveness ratio (ICER): (eA − eB) / (cA − cB), where cA is the cost of strategy A, eA is the cost of strategy B (usual care), eA is the effectiveness of strategy A, and eB is the effectiveness of strategy B. Effectiveness may be any relevant effectiveness measure, eg, quality-adjusted life-years or life expectancy. In this study, effectiveness is defined as correct classification of malignant disease. The ICER represents the incremental cost required to achieve 1 additional unit of benefit if strategy A is used instead of usual care (strategy B).

The numerator and the denominator of the ICER may be positive or negative. If the numerator is negative and the denominator is positive, then strategy A has lower expected costs and higher expected benefits. In such a case, strategy A would be said to dominate usual care. If strategy A is more costly but more effective, it still may be worthwhile if the extra costs do not exceed some threshold value determined by the decision maker(s).

For the baseline cost-effectiveness analysis, we computed ICERs of I-131 scintigraphy compared with FNAB and US compared with FNAB. Costs were estimated from the perspective of the health care facility and represent a weighted average of values obtained from a cost-accounting database. Effectiveness was defined as correct classification of malignant or benign disease. The primary outcome measures included the expected costs and effectiveness of the 3 diagnostics strategies as well as the ICERs of the strategies. All analyses were performed using commercially available decision analysis software (DATA Version 4.0; TreeAge Software Inc, Williams-town, Mass).

**SENSITIVITY ANALYSES**

To study the robustness of our results, we performed several 1-way sensitivity analyses for key model parameters, including the cost of FNAB, US, I-131 scintigraphy, and hemithyroidectomy and the sensitivity and specificity of FNAB. To perform the sensitivity analyses, we varied the model parameters over a reasonable range and recomputed the ICERs. We also determined the threshold values at which the parameters gained or lost dominance as initial diagnostic strategies.

(RePRINTED) ARCH OTOLARYNGOL HEAD NECK SURG VOL 132, MAR 2006 WWW.ARCHOTO.COM

©2006 American Medical Association. All rights reserved.
RESULTS

EXPECTED COST AND BENEFITS

Baseline results for costs, effectiveness, and ICERs are presented in **Table 2**. The FNAB strategy had an expected cost of $715 per patient and could be expected to correctly classify nodules as malignant or benign in 85.4% of cases. The I-131 scintigraphy strategy was more costly than FNAB, with an expected cost of $897 per patient, and more effective, classifying 86.1% of cases correctly. At $737 per patient, the US strategy was slightly more costly than FNAB, and it was the most effective strategy, correctly classifying 87.2% of patients.

INCREMENTAL COST-EFFECTIVENESS

The ICERs for each of the strategies are presented in **Table 2**. Each ratio was computed relative to the FNAB...
strategy and therefore represents the incremental cost-effectiveness of the I-131 scintigraphy and the US strategies compared with the FNAB strategy. The ICER of the I-131 scintigraphy strategy is $24,554, which means that each additional case correctly diagnosed by the I-131 scintigraphy strategy costs $24,554. Similarly, each additional case correctly diagnosed under the US strategy costs an additional $1,212.

**SENSITIVITY ANALYSES**

A sensitivity analysis of the cost of FNAB and the cost of US is presented in Figure 2. In panel A of Figure 2, the solid line represents the ICER comparing the cost-effectiveness of the US strategy with that of the FNAB strategy. The downward slope of the line implies that as the cost of an FNAB increases, the cost-effectiveness of the US strategy increases. Note also that the solid line crosses the x-axis at $371. This means that if the cost of an FNAB is greater than $371, the US strategy dominates the FNAB strategy, meaning that the US strategy is less costly and more effective. If the cost of an FNAB is less than $371, the US strategy is more costly but more effective. In this case, the US strategy may still be cost-effective as long as the decision maker(s) is willing to pay the additional cost per additional correct diagnosis. The dashed line in panel A of Figure 2 shows how the ICER comparing the I-131 scintigraphy strategy with the FNAB strategy changes as the cost of the FNAB test increases. As indicated by the downward slope of the dashed line, as the cost of an FNAB test increases, the I-131 scintigraphy strategy becomes less costly than the FNAB strategy. The fact that the dashed line never crosses the x-axis implies that the FNAB strategy never dominates the I-131 scintigraphy strategy for any cost above 0.

As the cost of I-131 scintigraphy increases, the I-131 scintigraphy strategy becomes less cost-effective than the FNAB strategy (Figure 3A). Interestingly, in this same scenario, the US strategy becomes more cost-effective than either the FNAB or the I-131 scintigraphy strategy (Figure 3B). Also, as the cost of hemithyroidectomy increases, the I-131 scintigraphy strategy becomes more cost-effective, with a simultaneous decline in the cost-effectiveness of the US strategy.

Furthermore, the results of a sensitivity analysis of the sensitivity and specificity of FNAB, which are shown in Figure 4, suggest that as long as the sensitivity of FNAB is not 0 and the specificity of FNAB is not 0, FNAB is not dominated by the other strategies. Finally, Figure 5 demonstrates that with a greater prevalence of malignancy in the population, the cost-effectiveness of the FNAB strategy is further increased compared with the US and I-131 scintigraphy strategies. In contrast, compared with the FNAB strategy, the greater the prevalence of cystic lesions in the population, the more cost-effective the US strategy. In fact, if the prevalence of cystic lesions exceeds 76.1%, the US strategy would dominate the FNAB strategy.

**COMMENT**

Thyroid nodules are a common problem in clinical practice.1,4,5,20 Because up to 5% of all nodules may harbor malignancy, it is crucial to correctly identify their benign or malignant status.1 Surgery (ie, hemithyroidectomy and total thyroidectomy) is not only the definitive diagnostic test but also the definitive treatment for thyroid carcinoma. However, routine hemithyroidectomy and/or total thyroidectomy for all nodules carries a small yet significant risk of morbidity and mortality relative to the likelihood of a positive cancer diagnosis. Consequently, more noninvasive diagnostic tests have gained wider acceptance in clinical practice over the years. While many authors have cited FNAB as the most important procedure for differentiating benign from malignant nodules, its role in the initial evaluation and management of nodular thyroid disease remains a topic of discussion. As a result, I-131 scintigraphy and US continue to be selected as primary diagnostic modalities in a variety of clinical settings, particularly by frontline providers.

**Table 1. Model Parameters Used in Decision Tree Analysis**

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Baseline Value</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cost of FNAB, $</td>
<td>253.50</td>
<td>NA</td>
</tr>
<tr>
<td>Cost of hemithyroidectomy, $</td>
<td>2,715.19</td>
<td>NA</td>
</tr>
<tr>
<td>Cost of I-131 scintigraphy, $</td>
<td>217.52</td>
<td>NA</td>
</tr>
<tr>
<td>Cost of ultrasonography, $</td>
<td>1,582.23</td>
<td>NA</td>
</tr>
<tr>
<td>Prevalence of malignancies</td>
<td>5.0</td>
<td>1, 5, and 6</td>
</tr>
<tr>
<td>Prevalence of cystic lesions</td>
<td>19.0</td>
<td>1, 5, and 6</td>
</tr>
<tr>
<td>Probability of “hot” node by I-131 scintigraphy</td>
<td>5.0</td>
<td>1, 5, 15, and 19</td>
</tr>
<tr>
<td>Probability of “warm” node by I-131 scintigraphy</td>
<td>10.0</td>
<td>1, 5, 15, and 19</td>
</tr>
<tr>
<td>Probability of “cold” node by I-131 scintigraphy</td>
<td>85.0</td>
<td>1, 5, 15, and 19</td>
</tr>
<tr>
<td>Prevalence of malignancy in hot nodules</td>
<td>4.0</td>
<td>1, 5, and 15</td>
</tr>
<tr>
<td>Prevalence of malignancy in warm nodules</td>
<td>7.0</td>
<td>1, 5, and 15</td>
</tr>
<tr>
<td>Prevalence of malignancy in cold nodules</td>
<td>16.0</td>
<td>1, 5, and 15</td>
</tr>
<tr>
<td>Probability of conclusive FNAB result</td>
<td>83.0</td>
<td>1, 5, 13, 14, 17, and 18</td>
</tr>
<tr>
<td>Probability of conclusive US result</td>
<td>56.0</td>
<td>1, 5, 8, and 16-18</td>
</tr>
<tr>
<td>Sensitivity of FNAB</td>
<td>0.8535</td>
<td>NA</td>
</tr>
<tr>
<td>Sensitivity of US</td>
<td>0.8719</td>
<td>NA</td>
</tr>
</tbody>
</table>

**Table 2. Expected Cost, Effect, and Incremental Cost-effectiveness Ratio (ICER) for Diagnostic Strategies**

<table>
<thead>
<tr>
<th>Strategy</th>
<th>Expected Cost, $</th>
<th>Incremental Cost, $</th>
<th>Expected Effect</th>
<th>Incremental Effect</th>
<th>ICER, $</th>
</tr>
</thead>
<tbody>
<tr>
<td>FNAB</td>
<td>715</td>
<td>NA</td>
<td>0.8535</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>I-131</td>
<td>897</td>
<td>182</td>
<td>0.8609</td>
<td>0.0074</td>
<td>24,554</td>
</tr>
<tr>
<td>scintigraphy</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>US</td>
<td>737</td>
<td>22</td>
<td>0.8719</td>
<td>0.0184</td>
<td>1,212</td>
</tr>
</tbody>
</table>

Abbreviations: FNAB, fine-needle aspiration biopsy; I-131, iodine 131; NA, not applicable; US, ultrasonography.
The results of our analysis confirm that FNAB is the most cost-effective test for patients presenting with palpable solitary thyroid nodules given our parameterization and assumptions. To illustrate this point, let us consider a hypothetical scenario of 1000 consecutive patients who present with palpable thyroid nodules (Table 3).

If all patients were given FNAB as the initial diagnostic strategy, we would expect to correctly diagnose 854 cases correctly at a total cost of $715,100. The I-131 scintigraphy strategy would be expected to correctly diagnose 861 cases at a total cost of $896,800. Therefore, the I-131 scintigraphy strategy produces 7 (861−854) additional correctly diagnosed cases at a cost of $181,700 ($896,800−$715,100), or $24,554 per additional correctly diagnosed case. Although 872 of 1000 cases would be correctly diagnosed using the US strategy, the total cost would be $737,400. Therefore, the US strategy correctly diagnoses 18 cases more than the FNAB strategy, at a cost of $22,300, or $1212 per additional correctly diagnosed case. It must be recognized, however, that the effectiveness (and higher cost) of the US strategy results from having to continue to the FNAB strategy when diagnosis is not readily apparent from the results of US alone. Similarly, I-131 scintigraphy appears more effective than FNAB because of its default to FNAB in cases of isofunctioning and hypofunctioning nodules.

A number of assumptions were used to create the model parameters for our decision tree. We assumed that an evaluation of a solitary palpable nodule and a proper medical workup of any underlying functional thyroid condition, ie, hyperthyroidism or hypothyroidism, had been obtained before entry into the algorithm. We modeled the practice of many frontline providers as observed from the perspective of a referral center specializing in the treatment of thyroid lesions
and receiving these "evaluated" patients. We also assumed that hemithyroidectomy provides the definitive "gold standard" pathologic diagnosis. Furthermore, in determining the cost of each diagnostic strategy, we took an average of reported costs from a large tertiary care center and assumed that they were in the same range as costs of these procedures elsewhere. Finally, we assumed that the experience or expertise of the cytopathologist for FNAB, the radiologist for I-131 scintigraphy, the sonographer for US, and the surgeon for hemithyroidectomy did not affect test outcomes. However, recognizing that in clinical practice these parameters may vary significantly across different institutions, we performed sensitivity analyses to assess how sensitive our results were to changes in these model parameters and demonstrated that these results are robust over a reasonable range of parameters.

![Figure 4](https://example.com/figure4.png)

**Figure 4.** Sensitivity analysis of the sensitivity and specificity of fine-needle aspiration biopsy (FNAB). As long as the sensitivity of FNAB (panel A) is not 0 and the specificity of FNAB (panel B) is not 0, FNAB is not dominated by the other strategies. This result is demonstrated by the lack of intersection of the sensitivity and specificity of the FNAB strategies with the other strategies. ICER indicates incremental cost-effectiveness ratio; US, ultrasonography; and I-131, iodine 131.

![Figure 5](https://example.com/figure5.png)

**Figure 5.** Sensitivity analysis of prevalence of malignancy and the probability that fine-needle aspiration biopsy (FNAB) yields a conclusive result. The greater the prevalence of malignancy in the population, the greater the increase in incremental cost-effectiveness ratio (ICER) of the FNAB strategy compared with the ultrasonography (US) strategy, meaning that the US strategy becomes more costly and less effective (panel A). However, the ICER of the FNAB strategy decreases with respect to the I-131 scintigraphy strategy, meaning as prevalence increases, I-131 scintigraphy strategy becomes less costly and more effective (panel A). In contrast, in panel B, the lower the probability of a conclusive FNAB result, there is a decreased ICER of the FNAB strategy compared with the US strategy. This would make the US strategy less costly and more effective. If the probability of the conclusive FNAB result is below 78%, the US strategy would dominate the FNAB strategy.

### Table 3. Hypothetical Results for Diagnosing 1000 Patients With Each Diagnostic Strategy

<table>
<thead>
<tr>
<th></th>
<th>Correct Diagnosis</th>
<th>Incorrect Diagnosis</th>
<th>Total Cost, $</th>
<th>Additional Correct Diagnosis</th>
<th>Additional Cost, $</th>
<th>Cost per Additional Correct Diagnosis, $</th>
</tr>
</thead>
<tbody>
<tr>
<td>FNAB</td>
<td>854</td>
<td>146</td>
<td>715,100</td>
<td>851</td>
<td>7.4</td>
<td>24,554</td>
</tr>
<tr>
<td>I-131 scintigraphy</td>
<td>861</td>
<td>139</td>
<td>896,800</td>
<td>867</td>
<td>7.5</td>
<td>22,300</td>
</tr>
<tr>
<td>US</td>
<td>872</td>
<td>128</td>
<td>737,400</td>
<td>873</td>
<td>18.4</td>
<td>12,12</td>
</tr>
</tbody>
</table>

Abbreviations: FNAB, fine-needle aspiration biopsy; I-131, iodine 131; NA, not applicable; US, ultrasonography.
It is important to note that other diagnostic techniques for assessing palpable thyroid nodules are available but were not included in this analysis. Principally, US-guided FNAB was not modeled in this study because it is not used routinely in front-line clinical practice as an initial diagnostic strategy for solitary nodules. The use of US-guided FNAB in a referral setting has proliferated in recent years. Ultrasound-guided FNAB is typically reserved for cases involving multiple nodules, for cases with inconclusive results on conventional fine-needle aspiration, or for cases in which a lesion has been detected on conventional US that would require US localization for precise needle placement. It is acknowledged that a greater use of US-guided FNAB is being realized in conjunction with more specific sonographic criteria for biopsy to increase the likelihood that malignant lesions are being adequately scrutinized.

As our health care system continues to evolve toward a more patient-centered delivery model, cost-effectiveness analyses must attempt to account for patient preferences and effects on quality of life. Future work needs to identify the individual trade-offs that patients are willing to make to achieve various levels of “risk” for harboring thyroid carcinoma, to ascertain characteristics that make particular tests (and strategies that depend on it) more attractive than other methods, and to undertake prospective analyses to determine whether diagnostic strategies affect short-term and/or long-term patient quality of life.

In summary, our model supports the notion that FNAB is the most cost-effective initial test currently available to distinguish between benign and malignant solitary palpable thyroid nodules as long as the acceptable cost per additional diagnosis remains below $1212. As such, we encourage its use by primary care providers when indicated and prospective clinical studies to validate these findings.

Submitted for Publication: June 20, 2005; accepted August 17, 2005.

Correspondence: Brendan C. Stack, Jr, MD, Department of Otolaryngology–Head and Neck Surgery, 4301 W Markham St, Suite 543, Little Rock, AR 72205-7199 (stackbrendanc@uams.edu).

Financial Disclosure: None.

Funding/Support: This study was supported in part by a 2002 Alpha Omega Alpha student research fellowship grant, Menlo Park, Calif.

Previous Presentation: This study was presented in part at the 74th Annual Meeting of the American Thyroid Association; October 10-13, 2002; Los Angeles, Calif.

REFERENCES