Cost-Effectiveness of Immediate Specimen Adequacy Assessment of Thyroid Fine-Needle Aspirations

Christopher R. Eedes, MD, and Helen H. Wang, MD, DrPH

Key Words: Thyroid; Fine-needle aspiration; Adequacy; Assessment

Abstract

Pathologists and cytotechnologists often provide immediate specimen adequacy evaluation of thyroid fine-needle aspirations (FNAs) to ensure that diagnostic material is obtained. We assessed the cost-effectiveness of this practice. All patients who had a thyroid FNA specimen accessioned at the Beth Israel Deaconess Medical Center, Boston, MA, during a 6-month period were included and divided into 2 groups: (1) with or (2) without immediate adequacy assessment. Specimen adequacy from each group was compared. The time spent to perform the adequacy assessment was recorded. Compared with group 2, group 1 had more specimens with diagnostic cellular material (67.2% vs 47.0%) and fewer specimens with suboptimal (23.3% vs 38.1%) or nondiagnostic cellular material (9.5% vs 14.9%) (P = .002). At the time of adequacy assessment, 98% (60/61) of the adequate specimens were obtained with 3 or fewer passes. The improved rate of diagnostic material was achieved at a cost of 220 minutes of cytologists’ time per additional diagnostic specimen compared with group 2. It may be most cost-effective to routinely obtain 3 passes and to perform immediate adequacy assessment under special circumstances such as repeated procedures.

Materials and Methods

At the Beth Israel Deaconess Medical Center (BIDMC), Boston, MA, endocrinologists, surgeons, and radiologists perform thyroid FNA biopsy with or without the aid of ultrasound guidance and with or without the presence of a cytologist. In addition, an endocrinologist, a surgeon, a radiologist, and a pathologist attend the biweekly Thyroid Nodule Clinic to which clinicians refer their patients with thyroid nodules.
When a thyroid FNA biopsy is deemed necessary, the radiologist performs the procedure under ultrasound guidance and the pathologist provides an immediate adequacy assessment. A final diagnosis is made following evaluation of all submitted materials.

The reports of all thyroid FNA biopsy specimens obtained at the BIDMC from July to December 2002 were obtained. The BIDMC Internal Review Board granted approval to review clinical data for the study. Computerized medical records for all patients were reviewed to determine whether the thyroid FNA biopsy was performed under ultrasound guidance and the number of passes recorded by the radiologist or clinician. The presence of a cytologist for specimen adequacy assessment was determined from the reports. We then matched our patients’ information with the database of all patients who attended the Thyroid Nodule Clinic during the study period to determine whether the thyroid FNA biopsy specimen was obtained at the Thyroid Nodule Clinic.

When a cytologist (a cytotechnologist, a cytopathology fellow, or a cytopathologist) attends the procedure, 1 or 2 direct smears are made from each pass and are air dried and stained with the rapid Romanowsky method. The remainder of the aspirated material from each pass is placed in a container of CytoLyte solution (Cytyc, Boxborough, MA) to be later processed by the ThinPrep (Cytyc) method and stained with the Papanicolaou stain.

The attending cytologist reviews the direct smears immediately and determines the adequacy of the specimen based on the number of cells and the amount of colloid present. The criteria for an adequate specimen admittedly varies somewhat among the cytologists. Generally, a specimen is considered adequate when several groups of at least a dozen follicular cells are present. The requirement of cellularity may be reduced if abundant colloid is noted. Each pass is considered separately. If the smear is considered adequate, the procedure is terminated. Otherwise, another pass is obtained and a similar protocol is followed for each aspirate until the specimen is considered adequate or the radiologist or clinician decides to stop the procedure for other reasons.

The same or a different cytologist reviews the direct smears prepared at the time of the procedure in conjunction with the ThinPrep slide to make a final diagnosis. The procedure is the same whether the aspiration is done at the Thyroid Nodule Clinic or at other sites. For cases with specimen adequacy assessments done outside the Thyroid Nodule Clinic during the second half of the study period, the amount of time from the cytologist leaving the laboratory to returning to the laboratory was determined.

When no specimen adequacy assessment is done, the aspirated material from each separately designated site is placed in a container of CytoLyte solution and submitted to the BIDMC Cytology Laboratory to be processed by the ThinPrep method and stained with the Papanicolaou stain. Some clinicians might choose to make direct smears by themselves and submit them with the material preserved in CytoLyte. Otherwise, the final diagnosis is made on the basis of a single ThinPrep slide unless ancillary studies are deemed necessary.

We use 7 diagnostic categories in the laboratory, and each specimen is assigned a category with further description. The 7 categories are: (1) positive for malignancy (papillary carcinoma, medullary carcinoma, lymphoma, and anaplastic carcinoma), (2) “suspicious for” malignancy, (3) follicular lesion with indeterminate potential of malignancy (microfollicular neoplasm, Hürthle cell neoplasm, and follicular lesion with features suggestive of papillary carcinoma), (4) most probably benign follicular lesions (mixed microfollicular and macrofollicular lesion and macrofollicular lesion), (5) negative for malignancy, (6) suboptimal cellularity, and (7) nondiagnostic owing to virtual absence of follicular cells. We usually attempt to subcategorize specimens in the suboptimal cellularity category based on the limited material to suggest whether they represent a microfollicular lesion or a mixed microfollicular and macrofollicular lesion. The specimens with cyst contents only also are included in this category. For the purpose of the present study, diagnoses in the categories other than the nondiagnostic and suboptimal cellularity categories were combined and considered adequate for diagnosis.

Specimens were divided into 2 groups: (1) those that had an immediate specimen adequacy assessment performed by a cytologist and (2) those that did not. The final diagnoses, number of passes to the extent we could determine, and clinical information for each group were compared. Comparisons also were made between cases with specimen adequacy determined at the Thyroid Nodule Clinic and those with specimen adequacy determined elsewhere. The statistical significance of any difference was determined by the Fisher exact test or $\chi^2$ for categorical data and by the Mann-Whitney test for numeric data.

Results

During the study period, 311 patients (261 women; 50 men; mean age, 49.4 years) underwent thyroid FNA biopsy at BIDMC. Of the 311 patients, 291 underwent a single procedure and 20 underwent 2 procedures, for a total of 331 specimens. Of the 331 specimens, 116 (35.0%) specimens had an immediate specimen adequacy assessment, and 215 specimens (65.0%) did not. Of the 116 with a determination of adequacy assessment, 71 were done at the Thyroid Nodule Clinic on 12 different occasions. The number of thyroid FNA
66
Am J Clin Pathol 2004;121:64–69
DOI: 10.1309/XLNDTE289WAQYK0Y
© American Society for Clinical Pathology

Outside the Thyroid Nodule Clinic. However, an adequate adequacy category whether the assessment was done at or more likely than those that did not to have a diagnosis in the adequate category at final diagnosis. Up to 20% of specimens with an adequate verdict at the immediate adequacy assessment had suboptimal cellularity on the final reports. Since we routinely make at most 2 direct smears and place the remainder of the aspirate into CytoLyt, the cytologist who attends the adequacy assessment expects to see more material on the ThinPrep slide. Consequently, the cytologists tend to give an adequate verdict at the immediate assessment even when the cellularity appears borderline adequate. Such expectations may or may not materialize. Nevertheless, none of the specimens with an adequate verdict on adequacy assessment was nondiagnostic.

Table 1 shows the comparison of adequacy categories between FNA biopsies done with ultrasound guidance and those done without. Although the nondiagnostic rate was similar between these groups, the specimens obtained under ultrasound guidance were much more likely to have a diagnosis in the adequate category (57.1% [121/212] vs 42% [40/95]), and the difference was statistically significant ($P = .02$). However, specimens from procedures done under ultrasound guidance also were more likely to have a determination adequacy assessment than those without ultrasound guidance (43.4% [92/212] vs 3% [3/95]; $P < .001$). Table 2 includes only specimens without adequacy assessment and compares the distribution of adequacy categories between biopsies done with ultrasound guidance and those done without. Biopsies done with ultrasound guidance had higher nondiagnostic and adequate rates (and a lower suboptimal cellularity rate), but the difference was not statistically significant.

Table 5 compares the number of passes according to the outcome of the immediate adequacy assessment. Specimens that were determined to be inadequate in the immediate assessment had significantly more passes than those judged adequate. This was particularly true for cases done outside the Thyroid Nodule Clinic. The radiologist who attends the Thyroid Nodule Clinic has a policy not to take more than 4 passes. For specimens that had an adequate verdict at adequacy assessment, 66% (40/61) had an adequate specimen biopsy procedures performed at each Thyroid Nodule Clinic varied from 1 to 11, with an average of 5.9 per occasion.

Table 11 gives the diagnostic categories for specimens without a determination of adequacy assessment, specimens with adequacy assessment done at the Thyroid Nodule Clinic, and specimens with adequacy assessment done outside the Thyroid Nodule Clinic. No significant difference was noted between specimens with assessment done at the Thyroid Nodule Clinic and those with assessment done outside the Thyroid Nodule Clinic, although for patients for whom adequacy assessment was done at the Thyroid Nodule Clinic, the rate of nondiagnostic specimens seemed lower. When the specimens in the suboptimal cellularity and adequate categories were combined into 1 category, there was no statistically significant difference in the nondiagnostic rate between specimens that had an adequate assessment and those that did not ($P = .09$; Fisher exact test).

Table 2 shows the obvious: specimens that had an adequate verdict at the adequacy assessment were much more likely than those that did not to have a diagnosis in the adequate category whether the assessment was done at or outside the Thyroid Nodule Clinic. However, an adequate verdict at the adequacy assessment did not guarantee a diagnosis in the adequate category at final diagnosis. Up to 20% of specimens with an adequate verdict at the immediate adequacy assessment had suboptimal cellularity on the final reports. Since we routinely make at most 2 direct smears and place the remainder of the aspirate into CytoLyt, the cytologist who attends the adequacy assessment expects to see more material on the ThinPrep slide. Consequently, the cytologists tend to give an adequate verdict at the immediate assessment even when the cellularity appears borderline adequate. Such expectations may or may not materialize. Nevertheless, none of the specimens with an adequate verdict on adequacy assessment was nondiagnostic.

Table 3 shows the comparison of adequacy categories between FNA biopsies done with ultrasound guidance and those done without. Although the nondiagnostic rate was similar between these groups, the specimens obtained under ultrasound guidance were much more likely to have a diagnosis in the adequate category (57.1% [121/212] vs 42% [40/95]), and the difference was statistically significant ($P = .02$). However, specimens from procedures done under ultrasound guidance also were more likely to have a determination adequacy assessment than those without ultrasound guidance (43.4% [92/212] vs 3% [3/95]; $P < .001$). Table 4 includes only specimens without adequacy assessment and compares the distribution of adequacy categories between biopsies done with ultrasound guidance and those done without. Biopsies done with ultrasound guidance had higher nondiagnostic and adequate rates (and a lower suboptimal cellularity rate), but the difference was not statistically significant.

Table 5 compares the number of passes according to the outcome of the immediate adequacy assessment. Specimens that were determined to be inadequate in the immediate assessment had significantly more passes than those judged adequate. This was particularly true for cases done outside the Thyroid Nodule Clinic. The radiologist who attends the Thyroid Nodule Clinic has a policy not to take more than 4 passes. For specimens that had an adequate verdict at adequacy assessment, 66% (40/61) had an adequate specimen
with the first 2 passes and 98% (60/61) had an adequate specimen with 3 or fewer passes. More passes did not always pay off. Table 6 shows the number of passes according to the final diagnosis. Similarly, nondiagnostic specimens were more likely to have more passes, but to no avail.

Table 7 shows that an adequacy assessment was much more likely to be done on repeated procedures than on first procedures (70% [14/20] vs 32.8% [102/311]; \( P = .001 \)). All repeated procedures were done owing to a less-than-diagnostic diagnosis on the first procedure (50% [10/20] each nondiagnostic and suboptimal cellularity categories). Of the repeated procedures, 60% (12/20) yielded a diagnostic diagnosis, while 35% (7/20) and 5% (1/20) remained suboptimal and nondiagnostic, respectively.

The average time for a cytologist to be away from the laboratory to attend an adequacy assessment was 57 minutes per procedure for those done outside the Thyroid Nodule Clinic. The average time cytologists spent at the Thyroid Nodule Clinic was 3.5 hours each session. During the study period, 71 specimens were obtained in 12 sessions. Therefore, the average time for each specimen at the Thyroid Nodule Clinic was 35 minutes. The total time spent on adequacy assessment for the 116 cases (57 minutes each for 45 cases outside the Thyroid Nodule Clinic and 35 minutes each for 71 cases at the Thyroid Nodule Clinic) was 5,050 minutes. Based on an adequate rate of 47% in cases without an adequacy assessment, we would have had 55 specimens in the adequate category if no immediate adequacy assessment had been done for the 116 cases. However, 78 of 116 cases with adequacy assessment done had a diagnosis in the adequate category. Therefore, the additional 23 adequate specimens (78 – 55) were obtained at a “cost” of 220 minutes of cytologist time per specimen (5,050/23). If all immediate adequacy assessments had been done at Thyroid Nodule Clinics, the cost would have been 177 minutes per additional adequate specimen.

**Discussion**

Our results demonstrate that immediate assessment of specimen adequacy increases (by approximately 20%) but does not guarantee the diagnostic yield of thyroid FNA biopsies. However, the increase was achieved at a stiffer price of 220 minutes of cytologist time per additional adequate case. The cost would be lower (177 minutes) if all adequacy assessments could be done in a designated location and at a designated time, as in the Thyroid Nodule Clinic. Thyroid FNA biopsies done with ultrasound guidance but without immediate adequacy assessment had only slightly higher diagnostic yields than those done without ultrasound guidance. Most of the specimens judged adequate at the immediate adequacy assessment and the specimens judged adequate on final diagnosis were obtained with 3 or fewer passes.

The literature on this issue is controversial. The only study that specifically addresses the effectiveness of on-site evaluation by a cytologist of ultrasound-guided thyroid FNA biopsy reported significantly prolonged procedure time (44.4 vs 12.5 minutes) when a cytologist provided immediate assessment. However, the distribution among the adequacy...
categories (satisfactory, evaluation limited, and unsatisfactory) was similar between procedures with immediate assessment and those without the assessment. Approximately 50% of cases in both groups had satisfactory adequacy. The remaining 50% were about equally divided between the evaluation limited and unsatisfactory categories. These results are similar to our results. The average time our cytologist spent on an adequacy assessment outside the Thyroid Nodule Clinic was longer than the average procedure time spent on an adequacy assessment in the Thyroid Nodule Clinic. Similarly, in our study, cases with adequacy assessment performed did not differ significantly from those without adequacy assessment performed in the ultrasound suite. Therefore, the time cytologists spent on an adequacy assessment outside the Thyroid Nodule Laboratory was similar between procedures with immediate on-site interpretation of FNA biopsy specimens. 

Two of these studies primarily showed good correlation between the immediate interpretation and the final diagnosis. Although Baloch et al reported a nondiagnostic rate of 11% for thyroid FNA biopsy specimens without specimen adequacy assessments vs a nondiagnostic rate of 5% for thyroid FNA biopsy specimens with adequacy assessments, the procedures for all specimens without adequacy assessments were done by palpation, and procedures for all specimens with adequacy assessments were done under ultrasound guidance. It was not possible to separate the effect of ultrasound guidance from that of adequacy assessment. Furthermore, these 2 groups of specimens were not from the same period, although they were from the same laboratory. Therefore, the comparability of nondiagnostic rates for the 2 groups in the 2 studies by Baloch et al was questionable. Austin and Cohen reported a diagnostic rate of 100% with on-site evaluation by a cytopathologist in 25 cases of proven lung malignancy. This was in contrast with a diagnostic rate of 80% without on-site evaluation in 30 cases of proven lung malignancy. However, the authors did not state the time spent to perform these on-site evaluations.

To our knowledge, only 1 study analyzed the cost and compensation of immediate on-site interpretation of FNA biopsy specimens. The authors’ conclusion was: “intraprocedural consultations by cytopathologists for CT [computed tomography]–guided, ultrasound-guided, bronchoscopic, or endoscopic procedures are compensated insufficiently by current Medicare compensation.” The time expenditure by cytopathologists for on-site evaluations of ultrasound-guided procedures in the study by Layfield et al was identical to that reported by O’Malley et al, ie, 44.4 minutes, although Tambouret et al reported a longer time (70 minutes) for cytotechnologists to perform ultrasound-guided thyroid FNA biopsy specimen evaluations.

Since compensation varies by insurance carrier and the credentials of the person performing the assessment, we used cytologist time as our measurement of cost. Although each case took on average 35 to 57 minutes, depending on the circumstance, most of the cases would have had adequate material without on-site evaluation. Only about one third of the cases with adequate material could have been attributed to on-site evaluations. Therefore, the time cytologists spent to obtain each additional adequate case was 177 to 220 minutes. In fact, the cases considered to have suboptimal cellularity were not exactly nondiagnostic or noninformative. Tulecke and Wang have shown that the cytohistologic
correlation in these cases was still statistically significant, although the significance was not as high as for the cases in the adequate category. If we use on-site evaluation to avoid nondiagnostic diagnoses only, then only 6 nondiagnostic cases were avoided in our study at a cost of 842 minutes of cytologists’ time per specimen. The time a cytologist spends on each procedure is somewhat less than the time the ultrasonographer or clinician has to spend to perform the procedure. However, this cost needs to be considered against the inconvenience and cost involved in repeating the procedures. Fortunately, a delay in diagnosis of thyroid lesions usually does not have serious consequences.

One potential confounding factor of our results is the effect of different operators performing the FNA biopsies. The adequacy rate of thyroid FNA biopsy is known to be operator-dependent. We receive our thyroid FNA biopsy specimens from clinicians who practice within our medical center and from clinicians whose offices are outside the medical center. We provide specimen adequacy assessments only for operators who are within the medical center. These operators, including endocrinologists and ultrasonographers, are highly experienced operators who see referred patients and perform frequent FNA biopsies. If our results had been attributed to the effect of operators, the bias would have been in favor of the impact of adequacy assessment and ultrasound guidance. Therefore, the cost we estimated for adequacy assessment probably represents a minimum cost owing to the proficiency of the operators who used specimen adequacy assessments and ultrasound guidance in our study.

Another potential confounding factor is selection bias of patients for immediate specimen adequacy assessment. As our data indicated (Table 7), ultrasonographers and endocrinologists tend to request immediate adequacy assessment of specimens from patients undergoing repeated aspirations owing to previous less-than-diagnostic FNA biopsy results. These patients probably had lesions that were difficult to sample. Other factors, such as size of the lesion, that we did not examine but that had an impact on the diagnostic yield also might have influenced the aspirator’s decision to request immediate assessment of specimen adequacy. This potential bias would have underestimated the effectiveness of specimen adequacy assessment and, thus, overestimated the cost.

Immediate assessment of specimen adequacy increases the diagnostic yield of thyroid FNA biopsies but at a tremendous expense in cytologist time and, perhaps, ultrasonographer and clinician time. This expense has to be weighed against the cost involved in repeating the procedure, such as time spent by patients and ultrasonographers. It might be worthwhile to design a strategy for judicious use of immediate specimen adequacy assessments to increase cost-effectiveness. For example, all initial thyroid FNA biopsy specimens can be obtained without immediate adequacy assessment. Then, thyroid FNA biopsies repeated owing to specimen inadequacy or the need for ancillary studies would be done with immediate specimen adequacy assessments in a designated location and at a designated time, such as in the weekly Thyroid Nodule Clinic.

References