OBJECTIVE. The purpose of this study was to determine the number of thyroid nodule workups that could be eliminated and the number of malignant tumors that would be missed if the Society of Radiologists in Ultrasound (SRU) recommendations and the three-tiered system were applied to incidental thyroid nodules (ITN) detected at imaging.

MATERIALS AND METHODS. This retrospective study included ITN in 390 consecutively registered patients who underwent ultrasound-guided fine-needle aspiration of one or more thyroid nodules from July 2010 to June 2011. Images were reviewed, and nodules were categorized according to two workup criteria: ITN seen on ultrasound images were categorized according to SRU recommendations, and those seen on CT, MR, or PET/CT images were classified according to the three-tiered risk-categorization system.

RESULTS. In this study 114 of 390 (29%) patients had nodules first detected incidentally during imaging studies, and 107 patients met the inclusion criteria. These patients had 47 ITN seen at ultrasound and 60 ITN seen at either CT, MRI, or PET/CT. If the SRU recommendations had been applied, 14 of 47 (30%) patients with ITN on ultrasound images would not have received fine-needle aspiration and one of four cases of cancer would have been missed. The missed malignant tumor was a 14-mm localized papillary carcinoma. If the three-tiered system had been applied, 21 of 60 (35%) patients with ITN on ultrasound images would not have received fine-needle aspiration, but none of the three malignancies would have been missed. Overall, 35 of 107 (33%) of patients with ITN did not meet the SRU recommendations or the three-tiered criteria.

CONCLUSION. Use of the SRU recommendations and three-tiered system can reduce the workup of ITN by one third compared with current practice without specific guidelines. One case of localized papillary carcinoma was missed when the SRU recommendations were used.

Keywords: CT, guidelines, incidental thyroid nodules, ultrasound, workup

DOI: 10.2214/AJR.13.10972

Received March 28, 2013; accepted after revision May 18, 2013.

Hasan A. Hobbs1
Manisha Bahl1
Rendon C. Nelson1
Peter G. Kranz1
Ramon M. Esclamado2
Nathan M. Wnuk3
Jenny K. Hoang1,4

thyroid nodules are common incidental findings at imaging. They are seen in 40–50% of ultrasound studies and 16% of CT and MRI studies that include the thyroid [1–5]. Despite a low malignancy rate of 0–9%, incidental thyroid nodules (ITN) are a management dilemma for radiologists and clinicians because of the risk of missing malignant tumors [5–7]. Because there are no established guidelines for workup, the current radiology practice for reporting and recommending workup of ITN varies widely.

The reporting and workup of ITN may be contributing to an increasing incidence of thyroid malignancy. New cases of thyroid carcinoma almost doubled in the 9 years 2000 to 2009, whereas previously (1973–2002) it took 30 years for the incidence to double [8, 9]. Although early diagnosis may be favorable in some malignancies, this may not be the case for thyroid cancer because most of the malignancies detected are small papillary carcinomas. Many experts believe that small thyroid cancers are pseudodisease and that most patients die with rather than of thyroid cancer [2, 8].

An ideal approach to selecting ITN for workup would be not to diagnose all cases of cancer but to “diagnose cancers that have reached clinical significance, while avoiding unnecessary tests and surgery in patients with benign nodules” [2]. There are two sets of criteria for selecting nodules, and they emphasize selection of clinically significant cancers and reduction of workup of benign nodules. The Society of Radiologists in Ultrasound (SRU) recommends fine-needle aspiration (FNA) for
solid nodules or coarsely calcified nodules 15 mm or larger, nodules with microcalcifications 10 mm or larger, solid nodules 15 mm or larger with coarse calcifications, and mixed solid-cystic nodules 20 mm or larger [2]. Another set of criteria is the three-tiered system proposed for workup of thyroid nodules seen at CT, MRI, or PET/CT [10]. The three-tiered system calls for workup of nodules with aggressive imaging features, including suspicious lymph nodes, local invasion, or focal metabolic activity at PET; presence in a patient younger than 35 years old; and solid nodule 15 mm or larger. Use of such criteria could reduce the workup of benign thyroid nodules, but to our knowledge the performance of the SRU criteria and the three-tiered system has not been investigated for a cohort of patients with ITN.

The aim of this study was to determine the number of nodule workups that could be eliminated and the number of malignant tumors that would be missed if the SRU recommendations and the three-tiered system were applied to ITN. Our hypothesis was that percentage of workups of ITN could be decreased through use of the SRU recommendations and the guidelines of the three-tiered system.

### Materials and Methods

**Subjects**

The study was approved by our institutional review board and was compliant with HIPAA. We retrospectively reviewed the records of ITN in 390 consecutively registered patients who underwent ultrasound-guided thyroid FNA in the department of radiology during the 12-month period July 2010 through June 2011. A thyroid nodule was considered incidental if it was an unsuspected focal lesion discovered on images obtained for other reasons without clinical symptoms, examination findings, or suspicion of thyroid cancer [10]. Patients undergoing thyroid ultrasound for investigation of a thyroid abnormality were not included. Patients with active metastatic disease or strong risk factors for thyroid cancer as described by the American Thyroid Association were excluded because our criteria were not intended for patients at high risk for malignancy [11].

Medical records were also reviewed for patient age, sex, and pathologic diagnosis. The diagnosis was confirmed by surgical pathologic result when available ($n = 87$) and FNA cytopathologic result if the patient did not undergo surgery ($n = 303$). In patients with malignant disease, medical records were reviewed to determine the stage of disease, treatment response, and overall survival.

### Application of Society of Radiologists in Ultrasound and Three-Tiered Criteria

The imaging studies that led to detection of ITN were reviewed by two board-certified radiologists (13 and 7 years of experience). Patients without the original imaging studies were excluded. Ultrasound studies were evaluated for maximal dimension of the nodule, composition of the nodule (solid, mixed solid and cystic, or cystic), presence of microcalcifications, and presence of coarse calcifications. SRU criteria were met if at least one nodule had any of the following features: microcalcifications and size 10 mm or greater, solid composition or coarse calcifications and size 15 mm or greater, or a mixed solid-cystic composition and size 20 mm or greater. Nodules detected on ultrasound that did not meet...
the SRU criteria composed the subgroup for which workup with FNA could have been avoided.

CT, MRI, and PET/CT images were evaluated for maximal dimension of the largest nodule, PET metabolic uptake, and aggressive features. Aggressive features included local invasion and suspicious lymphadenopathy. Abnormal nodes were defined as larger than 10 mm in the short axis or with morphologic features of either calcifications or cystic composition. The three-tiered criteria were met if the nodule had characteristics of any of the following risk categories: Risk category 1 (highest risk) denotes patients with concerning findings such as local invasion, suspicious lymphadenopathy, and focal metabolic uptake at PET. Risk category 2 is age younger than 35 years and not meeting criteria for risk category 1. This group was selected because of their higher ratio of malignant to benign nodules [12–17]. Risk category 3 is nodule size 15 mm or greater and not meeting criteria for risk category 1 or 2. The 15-mm cutoff for risk category 3 is intended to reflect a higher size threshold for workup of nodules that lack aggressive imaging findings or demographic risk factors. Several groups have used this cutoff in ultrasound evaluation of thyroid nodules [2, 6].

Patients with nodules detected on CT, MRI, or PET meeting the criteria for any of the three risk categories were considered to be in need of further ultrasound workup. Nodules not meeting the three-tiered criteria comprised the subgroup for which workup could have been avoided. For simplicity, SRU criteria were not applied to nodules that met the three-tiered criteria to determine which nodules would be biopsied by FNA.

Statistical Analysis

The two measures of performance of the SRU recommendations and the three-tiered system were potential reduction in workup and number of missed malignant tumors in ITN that did not meet the criteria. Potential reduction in workup was the number of ITN that did not meet either the SRU or three-tiered criteria. Percentage reduction in workup was calculated as number of criteria-negative ITN divided by the total number of ITN multiplied by 100. The number of missed malignant tumors was the number of malignancies in ITN that did not meet the SRU or three-tiered criteria. This number was used to calculate the false-negative rate (number of missed malignant tumors by criteria divided by total number of malignant tumors in ITN). Sensitivity and specificity of the criteria were also calculated.

Results

Subjects

In this study 114 of 390 (29%) subjects (78 female patients, 36 male patients; mean age, 60 years; range, 14–91 years) had ITN (Table 1 and Fig. 1). The mean nodule size was 26 ± 17.3 (SD) mm. The patients with ITN detected with ultrasound did not have a palpable thyroid mass and were undergoing neck ultrasound for evaluation of lymphadenopathy, hyperparathyroidism, and carotid artery disease. The diagnosis was confirmed with surgical pathologic results for 29 patients and FNA cytopathologic results for 85 patients who did not undergo surgery. One patient was excluded because of active metastatic melanoma. This patient could not be considered at low risk of malignancy because the SRU recommendations and three-tiered system are intended for patients at low risk with ITN. The other 113 patients had seven malignant tumors, all localized papillary carcinomas. There were no regional or distant metastases. All patients were...
treated during the study period, and there were no cases of recurrent disease or death of other causes. Of 113 patients, six did not have images available for review, and these patients were excluded. In the other 107 patients, 47 (44%) nodules were incidentally detected at ultrasound, and the other 60 (56%) were incidentally seen at CT (n = 37), MRI (n = 11), or PET/CT (n = 12) (Table 1 and Fig. 1).

**Performance of the Society of Radiologists in Ultrasound Criteria**

Fourteen of 47 (30%) patients did not meet SRU recommendations and would not have undergone FNA if the recommendations had been applied. Thus the potential reduction in workup with SRU criteria was 30%. Four (9%) papillary carcinomas were found in the 47 patients. Use of the SRU recommendations would have resulted in missing one malignant tumor, which was a localized papillary thyroid carcinoma measuring 14 mm without microcalcifications or coarse calcifications. The false-negative rate was 25%; sensitivity, 75%; and specificity, 30%. Of the 30 SRU-positive patients with benign nodules, 20 were nodules 15 mm or larger that were solid or coarsely calcified, six were mixed solid cystic nodules 20 mm or larger, and four were nodules 10 mm or larger with microcalcifications. Of the three SRU-positive patients with malignant nodules, two were nodules 15 mm or larger that were solid or coarsely calcified, and one was a nodule measuring 10 mm with microcalcifications.

**Performance of the Three-Tiered Criteria**

Twenty-one of 60 (35%) patients did not meet the criteria for the three-tiered system and would not have undergone FNA if the criteria had been applied. Thus the potential reduction in workup with the three-tiered system was 35%. The potential reduction in workup by modality was 15 of 37 (41%) for CT, 5 of 11 (45%) for MRI, and 1 of 12 (8%) for PET/CT. The group of ITN evaluated with the three-tiered system had three (5%) papillary carcinomas, measuring 10 mm, 22 mm, and 32 mm. There were no missed malignant tumors in the 21 three-tiered system negative patients. The false-negative rate was 0%; sensitivity, 100%; and specificity, 100%. Of the 36 three-tiered system positive patients with benign nodules, six met criteria for Category 1 based on an FDG-avid nodule on PET/CT, and 21 were nodules 15 mm or larger on CT and MRI. There were no patients in Category 2 of the three-tiered system.

**Ultrasound of Thyroid Nodules**

Overall, at initial imaging 35 of 107 (33%) patients with ITN did not meet the SRU recommendations or three-tiered system criteria for workup. If these criteria had been applied before the decision to work up the ITN, the 34 patients with benign thyroid nodules would not have undergone FNA, and one localized papillary carcinoma would have been missed.

**Discussion**

ITN can be a challenge for radiologists and clinicians. This study showed that ITN represented almost one third of thyroid nodules biopsied with ultrasound-guided FNA and that use of guidelines could reduce the numbers of workups by one third. Implementation of such guidelines could result in enormous health care savings if multiplied over time and across other institutions in the United States.

The SRU recommendations were developed by an expert group of radiologists to assist physicians in deciding which nodules seen at ultrasound should be biopsied with FNA. One of the concerns of the society was the unknown effect of the guidelines on the workup rate of thyroid nodules and the possibility that following the guidelines could increase the rates of FNA and subsequent surgery. Our study showed that use of the SRU recommendations reduced by 30% the rate of workup of ITN seen initially at ultrasound compared with current practice in which these criteria are not used. Ahn et al. [18] also studied the SRU recommendations, applying the criteria to all nodules biopsied with ultrasound guidance. For nodules 10 mm or larger the reduction in workup would have been 40% (396 SRU-negative nodules among 996 total nodules). They found the sensitivity and specificity of the SRU criteria were 72% and 42% (28% false-negative rate). When the SRU recommendations are applied to our cohort of 47 ITN detected with ultrasound, we found a similar sensitivity of 75% but even lower specificity (30%). This result is likely due to differences in patient populations. Our study included only asymptomatic nodules (incidental) rather than all thyroid nodules biopsied by FNA, and so we had a larger number of benign cases.

The three-tiered system for selecting thyroid nodules for workup based on CT and MRI findings was first described in a review by Hong et al. [10] but has since been evaluated in a retrospective study that compared the three-tiered system with a 10-mm or greater size cutoff for ITN detected at neck CT [7]. When the investigators applied the three-tiered system to nodules 10 mm or larger, there was a 46% reduction in the workup rate. In our study the reduction in workup rate for ITN detected at CT, MRI, or PET/CT was 35% with the three-tiered system compared with current practice of no guidelines. When we excluded the PET cases and only considered ITN detected at CT or MRI, the reduction in workup rate was 42%, similar to the results of Nguyen et al. [7]. This is because most of the nodules detected incidentally at PET were metabolically active and met the criteria for workup. Our study confirmed that metabolic activity at PET/CT is an indication for workup of an ITN (category 1 of three-tiered system). Two of eight (25%) nodules with FDG avidity at PET were malignant. Our results are consistent with those of other studies that have shown higher rates of malignancy (27–47%) of nodules detected at PET/CT, particularly those with focal hypermetabolic activity [19–21].

Performance of the classification criteria is strongly influenced by the prevalence of malignancy. The true prevalence of malignancy in an ITN is difficult to determine accurately in retrospective studies because not all ITN are biopsied and the malignancy rate depends on the study population. We found that the malignancy rate in 113 low-risk ITN was 6%, lower than that in other studies. Nam-Goong et al. [22] found a malignancy rate of 12% in nonpalpable nodules biopsied with FNA, but only 10% of their cohort had truly incidental nodules and not selected nodules for which the patient was referred for biopsy from other primary care clinics. The only U.S. study, to our knowledge, evaluating the rate of malignancy in ITN biopsied with FNA showed a malignancy rate of 17% [23]. However, that was a small retrospective study of 21 ITN compared with 90 thyroid nodules not incidentally detected. The authors could not explain why the latter group had a much lower malignancy rate of 3%.

In acknowledging the wide range of the reported rates of malignancy of ITN, it is important to recognize that the true prevalence of malignancy in ITN seen at imaging is many times lower. For each ITN in our study, many more ITN were not biopsied because they might not have been reported at imaging, the patient did not undergo sonographic workup, or the patient underwent diagnostic ultrasound but not biopsy. This has been a limitation in previous retrospective studies that attempted to identify the malignancy rate in ITN seen at consecutive CT and MRI examinations. In two studies ITN were found in 16–18% of imaging examinations, but FNA was performed on only 10% of ITN [4, 7]. Thus the true prevalence of malignancy in ITN seen at imaging
is unknown but is certainly a small fraction of our malignancy rate of 6%.

The greatest drive to work up nodules is the fear of missing malignancy. In this study, one case of cancer was missed. It was a 14-mm papillary carcinoma that did not meet the SRU criteria for FNA. According to the 2009 SEER database [9], the 10-year survival rate for a 14-mm localized tumor is greater than 99.6%. A contentious question for future research is whether treatment of small papillary cancers improves survival [24]. An observational trial that provides a perspective on untreated small cancers followed 340 patients with papillary microcarcinomas who were selected to not receive treatment [25]. After 10 years, new nodal metastases were found in 3% of cases, and there had been no cancer deaths. The opinion that most patients die with, rather than of, thyroid cancer is also supported by a postmortem study that showed incidental small foci of thyroid cancer in 36% of patients who died of other causes [26]. On the basis of this body of evidence, the intentions of both the SRU recommendations and the three-tiered system are not to diagnose all cancers but to diagnose cancers that have reached clinical significance while avoiding unnecessary tests and surgery on patients with benign nodules. There were several limitations to this study. First, it was a retrospective study with a small number of malignant tumors. We would expect more malignancies to be missed with the SRU recommendations and the three-tiered system when a larger number of FNAs are evaluated. It is important to note, however, that the false-negative rate is low relative to the percentage reduction in benign nodules worked up. Second, this study was conducted over a limited period of 12 months at a single institution, and the result may not be generalizable to past and future referral patterns. Review of institutional data on ultrasound-guided neck FNA in the preceding 9 years shows that our patient population represents the expected volume of FNAs in the radiology department. Third, FNAs performed in the radiology department do not include all thyroid FNAs at our institution. FNA of thyroid nodules is performed under ultrasound guidance by surgeons, endocrinologists, and pathologists, and these were not included. The differences between nonradiology FNAs and our study cohort are unknown. The nodules biopsied by clinicians and pathologists might have a higher proportion of SRU-negative or three-tiered criteria–negative nodules because of more liberal biopsy practices due to convenience of biopsy in a clinic. Or the opposite could occur because clinic biopsies may be performed on larger, less technically challenging thyroid nodules. Finally, we applied the criteria on the basis of the imaging findings and clinical information available. There might have been other clinical information not available in the medical records that put the patients at higher risk of malignancy and warranted biopsy. The guidelines are intended for patients without a high-risk clinical history.

The implication of this study is that simple criteria could be introduced to reduce the rate of unnecessary biopsy of benign nodules and thus reduce the rising financial and nonmone-
tary costs associated with workup and subsequent surgery. However, this study represents the biopsy practices in only one academic radiology practice. Practices may differ at other institutions depending on how thyroid nodules are reported at imaging and the practice patterns of the referring clinicians. To understand the effect of implementing guidelines nationally, it would be useful to evaluate the criteria across different practice types or assess practice by performing a survey of how ITNs are reported by radiologists. If reduction in unnecessary workup is possible at other institutions, these criteria could also be used to develop evidence-based practice guidelines for ITN. The challenge beyond creating guidelines is having them adopted by the clinicians who refer patients for biopsy. There are other existing guidelines for nodules seen at ultrasound, such as those of the National Comprehensive Cancer Network and the American Thyroid Association [27, 28]. The National Comprehensive Cancer Network guidelines are similar to the SRU recommenda-
tions in use of a 1.5-cm cutoff for nodules without other suspicious findings. We agree that the larger size cutoff of 1.5 cm used in the SRU criteria and the three-tiered system will help to reduce unnecessary workups.

Conclusion

This study showed that applying the SRU recommendations and the three-tiered system to ITN can reduce the workup rate almost one third compared with current practice without specific guidelines. The cost of reducing workup would be one missed papillary carcinoma. This malignant tumor and all other incidental thyroid malignancies in this study were low stage without regional or distant metastatic disease.

References

11. American Thyroid Association (ATA) Guidelines Taskforce on Thyroid Nodules and Differentiated Thyroid Cancer; Cooper DS, Doherty GM, et al. Revised American Thyroid Association management guidelines for patients with thyroid nodules and differentiated thyroid cancer. Thyroid 2009; 19:1167–1214
16. Rago T, Fiore E, Scutari M, et al. Male sex, single nodularity, and young age are associated with the
risk of finding a papillary thyroid cancer on fine-needle aspiration cytology in a large series of patients with nodular thyroid disease. Eur J Endocrinol 2010; 162:763–770
17. Shetty SK, Maher MM, Hahn PF, Halpern EF, Aquino SL. Significance of incidental thyroid lesions detected on CT: correlation among CT, sonography, and pathology. AJR 2006; 187:1349–1356

FOR YOUR INFORMATION

This article is available for CME and Self-Assessment (SA-CME) credit that satisfies Part II requirements for maintenance of certification (MOC). To access the examination for this article, follow the prompts associated with the online version of the article.

This article has been selected for AJR Journal Club activity. The accompanying Journal Club study guide can be found on the following page.
Study Guide

Incidental Thyroid Nodules Detected at Imaging: Can Diagnostic Workup Be Reduced by Use of the Society of Radiologists in Ultrasound Recommendations and the Three-Tiered System?

Joseph J. Budovec, Margaret Mulligan, Alan Mautz*
Medical College of Wisconsin, Milwaukee, WI
jbudovec@mcw.edu, mmulliga@mcw.edu, amautz@mcw.edu

Introduction
1. Is this study timely and relevant? Do the authors provide an appropriate rationale for performing the study? Is the study based on an appropriate review of the medical literature?

Methods
2. What were the inclusion criteria for the study? What were the exclusion criteria? Was the study institutional review board and HIPAA compliant?
3. Did the authors justify or explain their choice of study design? How did the authors decide to use the Society of Radiologists in Ultrasound (SRU) consensus criteria? What is the three-tiered system and why did the authors decide to use those criteria as well?
4. What are the limitations of this study? Are these limitations adequately discussed?
5. What statistical methods were used to analyze the data? Did the authors’ study design use methods that permitted their hypothesis to be tested?

Results
6. Was the research question answered? Was the hypothesis verified or disproved?
7. Are any of the results surprising? Do the results mirror empirical evidence from practice?

Study Design
8. What is the definition of disease prevalence? How does prevalence differ from incidence? Why is the knowledge of the true prevalence of malignancy in an incidental thyroid nodule important? How does prevalence influence diagnostic accuracy?

Discussion
9. How does this study compare with similar studies?
10. Does the patient population in this study match the patient population at your institution or practice? Can the data from this study be extrapolated to other patient populations? Why or why not?
11. What are the implications of the study results? To what degree, if any, do the results of this study influence how you practice? Would you adopt the SRU and three-tiered system criteria to incidental thyroid nodules? Why or why not?
12. What criteria do clinicians use when requesting fine-needle aspiration of thyroid nodules? Are the criteria in the SRU recommendations and the three-tiered system aggressive enough to minimize the number of missed malignancies?
13. How would you improve on this study? Should the authors suggest additional criteria for evaluating incidental thyroid nodules or does the discussion in the article adequately relay why no such alterations are offered?

Background Reading

*Please note that the authors of the Study Guide are distinct from those of the companion article.