Accuracy of ultrasonography-guided fine needle aspiration cytology and significance of non-diagnostic cytology in the preoperative detection of thyroid malignancy

Julian Sau Lian Chieng, MD, FRCR, Chau Hung Lee, MBBS, FRCR, Amit Anand Karandikar, MBBS, FRCR, Julian Park Nam Goh, MBBS, FRCR, Susanna Soo See Tan, MBBS, FRCR

INTRODUCTION Thyroid fine-needle aspiration cytology (FNAC) is an established investigation for the preoperative evaluation of thyroid nodules and is often done under ultrasonography guidance. While its accuracy has been widely reported, there is little evidence in the literature on the approach to non-diagnostic cytology results. In our study, we aimed to determine the diagnostic performance of ultrasonography-guided thyroid FNAC for the preoperative diagnosis of thyroid cancer at our institution and evaluate the significance of a non-diagnostic thyroid FNAC.

METHODS We retrospectively reviewed the thyroid ultrasonography studies and medical records of all patients who underwent both thyroid FNAC and subsequent thyroid surgery at our institution from 2011 to 2013. FNAC results were correlated with the final histological diagnosis from surgery and the ultrasonography studies were reviewed for suspicious sonographic features.

RESULTS FNAC predicted malignancy with sensitivity, specificity, positive predictive value, negative predictive value, false positive rate, false negative rate and total accuracy of 90.7%, 53.6%, 43.3%, 93.7%, 46.4%, 9.3% and 64.1%, respectively. We found that only one of 26 nodules with non-diagnostic FNAC results was proven malignant on a second FNAC and subsequent thyroidectomy.

CONCLUSION The accuracy of ultrasonography-guided FNAC at our institution was comparable to that reported in the literature. There appears to be very low probability of malignancy in sonographically benign nodules with initial non-diagnostic FNAC results.

Keywords: fine-needle aspiration cytology, histology, thyroid neoplasms, thyroidectomy, ultrasonography

INTRODUCTION Thyroid nodules are very common: palpable thyroid nodules occur in about 4%–7% of the population, while incidentally discovered nodules, either on ultrasonography or computed tomography, can occur in up to 67%.[1] The overall incidence of thyroid malignancy has been reported to be 16% in thyroidectomy specimens.[2]

Fine-needle aspiration cytology (FNAC) is an established investigation of choice for the preoperative evaluation of thyroid nodules for malignancy. There are established guidelines for selecting thyroid nodules for FNAC based on sonographic features.[3] However, patient selection can be difficult due to the high prevalence of asymptomatic thyroid nodules and because indiscriminately performing FNAC puts patients at unnecessary procedural risk and discomfort. Severe pain, pain lasting several days and haematoma are the most common complications encountered, with complication rates in the range of 1%–8.6%.[4] Other rare examples of reported complications include recurrent laryngeal nerve palsy, vasovagal reaction and post-aspiration thyrotoxicosis. Ultrasonography is also operator-dependent and certain suspicious appearances, such as microcalcifications, can be confused with the reverberation artefact of colloid contents. Even when FNAC is performed under ultrasonography guidance, inadequate sampling does occur and cytological interpretation can be challenging.

We aimed to review the thyroid FNAC results of all patients who underwent both thyroid FNAC and subsequent thyroid surgery at our institution over three years (2011–2013) and compare them to those reported in the literature. We also aimed to determine the significance of FNAC that is categorised as non-diagnostic in terms of patient management.

METHODS Approval for this retrospective study was obtained from the institutional review board. We retrieved the results of all thyroid FNAC performed under ultrasonography guidance at the Department of Diagnostic Radiology, Tan Tock Seng Hospital, Singapore, from January 2011 to December 2013. All cases were captured and retrievable under the department’s Picture Archiving and Communication System (PACS).

The electronic medical records (EMRs) of the patients were then reviewed. Data from all patients who underwent subsequent thyroid surgery, whose histology reports were captured in the EMRs, was collected. Data recorded included age, gender, ethnicity, sonographic features of the sampled nodule, FNAC cytological category and histological reports from the surgical
specimen. The following were excluded: (a) patients who did not have any surgical or histological correlation in the EMRs; (b) patients who had surgery performed at a private external institution and thus were not captured in the EMRs, which is a government database; and (c) patients for whom the relationship between cancer in the surgical specimen and the FNAC-sampled nodule could not be established.

Ultrasoundography studies for each patient were retrospectively reviewed by two head-and-neck radiologists, one with five years and the other with over ten years of experience. Both radiologists were blinded to the FNAC results and final histology. Horvath et al\(^{(5)}\) described ten ultrasonography features of thyroid nodules and related the rate of malignancy to each sonographic feature, in what is commonly known as a thyroid imaging reporting and data system. However, this system was deemed difficult to apply in clinical medicine because of subjectivity and complexity.\(^{(5)}\)

In our study, four sonographic features of each nodule were retrospectively recorded: borders; echogenicity; the presence of calcifications; and whether the nodule was taller than wide. These features have been shown to be significantly associated with malignancy and are easily applied clinically.\(^{(8)}\) While a solid nodule has a higher probability for malignancy compared to a mixed solid-cystic nodule,\(^{(46)}\) nodule composition per se is not considered to be an independent predictor of malignancy and therefore was not included as a criterion. The four established sonographic features that are independent predictors of malignancy are calcifications (microcalcifications and macrocalcifications), marked hypoechoicogenicity, taller-than-wide shape, and spiculated or ill-defined margins.\(^{(30)}\) Any one feature would classify the nodule as sonographically suspicious.\(^{(8-12)}\)

Borders were classified as circumscribed or ill-defined (depending on how clearly the margin was demarcated), and regular or irregular (depending on the presence of spiculations or microlobulations). Ill-defined nodules were defined as those with an indistinct demarcation between the tumour and the surrounding normal glandular tissue or nodules whose margins could not be obviously differentiated from the adjacent thyroid tissue.\(^{(8-9)}\) The border was deemed to be irregular or ill-defined if any one of the few representative ultrasound images showed a part of the border as irregular or unclear. Thyroid nodules were classified as hypoechoic, isoechoic or hyperechoic relative to thyroid parenchyma, with marked hypoechoicogenicity defined as being darker than the infrahyoid or strap muscles.\(^{(38)}\) Calcifications were classified as microcalcifications (punctate hyperechoic spots less than 1 mm without posterior acoustic shadowing) or coarse calcifications (larger echogenic foci with posterior shadowing).\(^{(8-9)}\) A nodule was taller than wide if its anteroposterior dimension was less than 1 mm without posterior acoustic shadowing) or coarse calcifications (larger echogenic foci with posterior shadowing).\(^{(8-9)}\)

Next, histologically proven malignant and benign nodules were compared to determine if there was any significant difference in terms of the presence of sonographically suspicious features between the two groups of nodules. Differences were compared using Fisher’s exact test, with statistical significance set at \(p < 0.05\).

First, histologically proven malignant and benign nodules were sampled, results for each nodule were recorded separately. For patients with more than one FNAC procedure performed for one nodule within the data collection period, the highest category was recorded. If separate nodules from one patient were sampled, results for each nodule were recorded separately. Raw data was collected and tabulated using a standard form. For patients with more than one FNAC procedure performed for one nodule within the data collection period, the highest category was recorded. If separate nodules from one patient were sampled, results for each nodule were recorded separately.

\[ \text{NPV} = \frac{TN}{TN + FN} \]
\[ \text{PPV} = \frac{TP}{TP + FP} \]
\[ \text{Specificity} = \frac{TN}{TN + FP} \]
\[ \text{Sensitivity} = \frac{TP}{TP + FN} \]
\[ \text{False positive rate} = \frac{FP}{FP + TN} \]
\[ \text{False negative rate} = \frac{FN}{FN + TP} \]
\[ \text{Accuracy} = \frac{(TP + TN)(TP + FN + FP + FN)}{TP + FN + FP + FN} \]

FNAC results classified as Category I (non-diagnostic) were then reviewed separately to determine if any cancers were missed in this group of nodules. The NPV of a Category I nodule was calculated in a similar way, as \( \text{NPV} = \frac{TN}{TN + FN} \), where \( TN \) and \( FN \) were defined as a Category I nodule proven to be benign and malignant, respectively, on histology.
RESULTS
A total of 179 patients, with 195 nodules, underwent both ultrasonography-guided FNAC and thyroid surgery over the three-year period. 14 microcarcinomas were excluded, as they were not within, or could not be established as being within, the sampled nodule. Two nodules that were histologically proven to be cancerous were also excluded, as they were not the nodules that were sampled for FNAC, based on the cytology report provided in the EMRs. Altogether, 26 nodules had non-diagnostic FNAC results and were analysed separately.

A total of 153 nodules were included for analysis of accuracy of ultrasonography-guided FNAC. There were 43 histologically proven malignant nodules and 110 histologically proven benign nodules. Of the malignant nodules, three follicular carcinomas did not show any suspicious sonographic features. However, they were relatively large, measuring 3.1 cm × 3.9 cm × 5.6 cm at maximum diameter. The other malignant nodules showed at least one suspicious sonographic feature. Overall, a significantly higher proportion of malignant nodules demonstrated at least one suspicious sonographic feature when compared to benign nodules (p < 0.05). These results are summarised in Table I.

The histological subtypes of the 43 histologically proven malignant nodules are summarised in Table II. There were 39 TP nodules, 59 TN nodules, 51 FP nodules and 4 FN nodules. The sensitivity, specificity, PPV, NPV, FP rate, FN rate and total accuracy of ultrasonography-guided thyroid FNAC for preoperative detection of thyroid cancer were 90.7%, 53.6%, 43.3%, 93.7%, 46.4%, 9.3% and 64.1%, respectively.

There were 26 nodules with non-diagnostic FNAC results because of inadequate cellular material. Of these, one was entirely cystic, 15 were solid-cystic and ten were solid. These nodules had a mean size of 4 (range 0.8–7.1) cm. The 1 (3.8%) nodule that was eventually proven to be malignant was small (size 0.8 cm) but had irregular borders and was deemed suspicious (Fig. 1). Repeat FNAC was performed, which showed suspicious cytology, and papillary carcinoma was confirmed on histology. Thyroidectomy was performed for two nodules that showed calcifications and were deemed suspicious (Figs. 2 & 3). Both were goitre nodules on histology. The remaining 23 nodules did not show suspicious sonographic features and were histologically benign: 20 were goitre nodules and three were adenomas. In our study, the overall NPV of a non-diagnostic FNAC was 96.2%.

---

**Table I. Presence of suspicious sonographic features in histologically benign and malignant thyroid nodules.**

<table>
<thead>
<tr>
<th>Feature</th>
<th>No. (%)</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Benign (n = 110)</td>
<td>Malignant (n = 43)</td>
</tr>
<tr>
<td>Calcification*</td>
<td>21 (19.1)</td>
<td>35 (81.4)</td>
</tr>
<tr>
<td>Irregular/ill-defined border</td>
<td>34 (30.9)</td>
<td>26 (60.5)</td>
</tr>
<tr>
<td>Markedly hypoechoic</td>
<td>21 (19.1)</td>
<td>34 (79.1)</td>
</tr>
<tr>
<td>Taller-than-wide shape</td>
<td>6 (5.5)</td>
<td>12 (27.9)</td>
</tr>
</tbody>
</table>

*Coarse calcification or microcalcification.

**Table II. Histopathological subtypes of malignant thyroid nodules (n = 43) with corresponding cytological category.**

<table>
<thead>
<tr>
<th>Subtype/cytological diagnosis</th>
<th>No. (%)</th>
<th>Cytology</th>
</tr>
</thead>
<tbody>
<tr>
<td>Papillary carcinoma</td>
<td>36 (83.7)</td>
<td>TP</td>
</tr>
<tr>
<td>Atypical/papillary carcinoma</td>
<td>33</td>
<td>TP</td>
</tr>
<tr>
<td>Follicular lesion</td>
<td>2</td>
<td>TP</td>
</tr>
<tr>
<td>Benign cyst content</td>
<td>1</td>
<td>FN</td>
</tr>
<tr>
<td>Follicular carcinoma</td>
<td>4 (9.3)</td>
<td>TP</td>
</tr>
<tr>
<td>Follicular carcinoma</td>
<td>1</td>
<td>TP</td>
</tr>
<tr>
<td>Nodular goitre</td>
<td>3</td>
<td>FN</td>
</tr>
<tr>
<td>Medullary carcinoma</td>
<td>2 (4.7)</td>
<td>TP</td>
</tr>
<tr>
<td>Medullary carcinoma</td>
<td>2</td>
<td>TP</td>
</tr>
<tr>
<td>Oncocytic tumour with atypical nuclear feature</td>
<td>1 (2.3)</td>
<td>TP</td>
</tr>
<tr>
<td>Atypical</td>
<td>1</td>
<td>TP</td>
</tr>
</tbody>
</table>

FN: false negative; TP: true positive

---

**Fig. 1** US image shows a markedly hypoechoic nodule (asterisk) with irregular borders (arrows). Although the initial fine-needle aspiration cytology (FNAC) result was non-diagnostic, the sonographic features were deemed suspicious and a repeat FNAC yielded a suspicious result (Category IV). Papillary carcinoma was confirmed after total thyroidectomy.

**Fig. 2** US image shows a well-circumscribed nodule with coarse calcifications (arrow) and microcalcifications (arrowhead). Fine-needle aspiration cytology was non-diagnostic. Hemithyroidectomy was performed as the nodule was deemed suspicious and final histology was a benign goitre nodule.
while the NPV of a non-diagnostic FNAC in a nodule with no suspicious sonographic feature was 100.0%.

DISCUSSION

Thyroid FNAC is important for the preoperative evaluation of thyroid nodules, helping in risk stratification and surgical planning. Ultrasonography guidance allows targeting of non-palpable nodules and the most suspicious sites in larger nodules. This has been shown to increase cellular yield for adequate cytological analysis. The presence of a cytologist on site to screen for adequacy of samples also reduces the likelihood of inadequate sampling. At our institution, ultrasonography-guided FNAC is usually performed by a radiologist with an on-site cytologist.

A major dilemma in the management of thyroid nodules for most head-and-neck surgeons and radiologists is deciding which nodule should undergo FNAC. Guidelines are in place to aid decision-making, for example, those from the Society of Radiologists in Ultrasound, American Thyroid Association (ATA), American Association of Clinical Endocrinologists (AACE)/Associazione Medici Endocrinologi (AME) (AACE/AME), European Thyroid Association (ETA) and Korean Society of Thyroid Radiology (KSThR).

Differences between guidelines are mainly regarding the size criteria for FNAC. The ATA recommends FNAC for nodules over 1 cm in size if there are suspicious features and over 0.5 cm if the patient has high-risk factors, while the KSThR recommends FNAC for all nodules more than 0.5 cm if there are suspicious features, regardless of risk factors. The KSThR also recommends selective FNAC of nodules even smaller than 0.5 cm, if there are high-risk factors, while the AACE/AME recommends FNAC in high-risk patients regardless of size. In the absence of risk factors and suspicious sonographic features, the KSThR, AACE/AME and ETA recommend FNAC for all solid nodules over 1 cm and any complex solid-cystic nodule, while the ATA recommends FNAC for solid nodules that are more than 1.5 cm and solid-cystic nodules more than 2 cm.

The head-and-neck surgeons and radiologists at our institution generally follow the ATA guidelines for thyroid FNAC, which allow a greater size cut-off for FNAC, particularly for nodules without any sonographically suspicious features. This also implies that FNAC was performed on a proportion of nodules without any of the four established sonographically suspicious features. While chaotic intranodal vascularity is classified as a suspicious sonographic feature in accordance with the KSThR guidelines, we do not utilise this criterion at our institution, in line with the ATA recommendations.

The high sensitivity of thyroid FNAC in our study (90.7%) was compatible with reported sensitivities in the literature. However, specificity and overall accuracy appeared to be relatively low (53.6% and 64.1%, respectively). We believe this was because Category III nodules (atypical or undetermined significance) were classified as cytologically positive for malignancy. This resulted in a significant FP rate (51/153), as cytological atypia can be seen in a variety of benign conditions such as thyroiditis, adenomatous nodules, Hurthle cell hyperplasia or even a recent FNAC procedure. However, as risk of malignancy in this category ranges from 5% to 15%, a Category III cytology result should not be overlooked. The decision for thyroidectomy in this group of patients therefore depends on risk stratification in terms of the patient’s age and gender as well as sonographic features of the nodule. A similar approach was used in other studies, with similar specificities. For example, Tabaqchali et al reported a specificity and accuracy of 51.2% and 57.4%, respectively, while Muratli et al reported specificity and accuracy of 64.6% and 77.6%, respectively, for FNAC cytology of thyroid nodules.

There were four nodules with FN cytology in our study, with an FN rate of 9.3%. This was consistent with reports that suggested FN rates of 2%–10.5%. Therefore, careful clinical follow-up is required to prevent the potential miss of malignancy. Three nodules were cytologically classified as nodular goitre (benign, Category II). They were deemed sonographically benign but were large (sizes 3.1 cm, 3.9 cm and 4.1 cm) and were found to be follicular carcinomas. This illustrates the need for a high index of suspicion for solitary large thyroid nodules or a dominant large nodule in a multinodular goitre. The other nodule...
with FN cytology was classified as a benign cyst. In this case, a young woman underwent FNAC of a solitary 4.8-cm solid-cystic nodule. Cytology showed benign cyst contents (Category II). However, microliths were present within the solid component on ultrasonography (Fig. 4) and the FNAC result was deemed discordant. On thyroidectomy, histology showed papillary carcinoma with cystic degeneration. While cystic and solid-cystic nodules are generally accepted as benign features, one case series has reported a malignancy rate of 5% in solid-cystic nodules, particularly in cases where the solid component is over 50%.\(^{13}\)

Non-diagnostic FNAC also poses a management dilemma. Non-diagnostic smears result from hypocellular specimens usually caused by cystic fluid, bloody smears or suboptimal preparation, and account for 2%–20% of specimens.\(^{14-16}\) Several factors influence non-diagnostic rates for FNAC results, including skill of the operator, vascularity of the nodule, criteria used to judge the adequacy of the specimen, and the cystic component of a nodule.\(^{14,16}\) Several studies have shown that a cystic or predominantly cystic nodule is strongly associated with a higher likelihood of non-diagnostic cytology.\(^{10,15,17}\) This was reflected in our study, where 61.5% (n = 16/26) of nodules with non-diagnostic FNAC were either cystic or solid-cystic. It is recommended that for a nodule with non-diagnostic FNAC, repeat FNAC should be performed, as a second FNAC, even a non-diagnostic one, significantly lowers the risk of malignancy as opposed to having only one non-diagnostic sample.\(^{18}\) Occasionally, a second FNAC may yield a different result.

In our study, the overall malignancy rate for nodules with single non-diagnostic cytology was 3.8% (n = 1/26), which is comparable to that reported in the literature (4.5%–7%).\(^{19,40}\) For this malignant nodule, repeat FNAC yielded suspicious cytology (Category IV) and papillary carcinoma was confirmed on thyroidectomy. Repeat FNAC was performed because on ultrasonography, the nodule was hypoechoic with irregular borders and was deemed sonographically suspicious.

Such cases illustrate the need for further management, such as repeat FNAC or surgery, if initial FNAC results are discordant with sonographic features. Our study also showed that nodules with non-diagnostic cytology and none of the four sonographically suspicious features (disregarding size, which was a criterion for FNAC but was not considered a sonographically suspicious feature) were all confirmed to be benign lesions on histology. We suggest that a second FNAC for such nodules is not required. However, it is important that the whole nodule be interrogated in detail, which can be time-consuming in larger nodules and multinodular goitre. While size is not an established predictor of malignancy, there is evidence that nodules larger than 4 cm show an increased risk of malignancy, and FNAC is recommended in this group of nodules even without the presence of sonographically suspicious features.\(^{41}\)

A few limitations were identified in our study. First, for convenience of data collection, we excluded patients for whom FNAC was performed outside of the radiology department, as they were not captured in our department’s PACS. Second, it can be difficult to determine if the site of malignancy is within the nodule sampled during FNAC, even for the examining pathologist, particularly for patients with multiple thyroid nodules or in cases of microcarcinoma, where the cancer may not even be sonographically visible. We also did not specifically analyse the possibility of sampling error within the nodule, which could have contributed to non-diagnostic or FN cytology. Third, it can be difficult to retrospectively analyse ultrasonography images, as certain features such as calcifications and borders are best analysed real-time. The actual extent of an irregular or ill-defined border cannot be accurately determined on a few representative ultrasonography images. Moreover, to the best of our knowledge, there is no absolute value or extent beyond which a nodule is considered to have irregular or ill-defined borders. For the ease of image interpretation in our study, as long as any one of the ultrasonography images showed a part of the border as irregular or not clear, the nodule was taken to be ill-defined or poorly demarcated. Fourth, a majority of patients who underwent ultrasonography-guided FNAC did not have histological correlation. Therefore, our retrospective study may not reflect the true accuracy of ultrasonography-guided thyroid FNAC. Nonetheless, our results seem consistent with findings in the literature. Fifth, the number of nodules with non-diagnostic cytology with histological correlation was rather small. A larger sample size and prospective study design would be required to more definitively conclude that a nodule with a single non-diagnostic cytology and no sonographically suspicious feature is almost certainly benign. Lastly, the malignancy rate for non-diagnostic FNAC in our study may be higher than expected compared with the general population, as this retrospective review of patients based on surgical correlates potetially selected patients with higher clinical risk, given that 43 out of 195 nodules were proven malignant. Nonetheless, our results were similar to the published literature.

In conclusion, ultrasonography-guided thyroid FNAC has high sensitivity for detection of thyroid malignancy, and the accuracy at our institution was comparable to that reported in the literature. The selection criteria for FNAC for these nodules were generally based on the ATA guidelines. We therefore suggest that prospective head-to-head studies be done to determine any differences between guidelines in the eventual detection of thyroid malignancy. Cytologically inadequate (non-diagnostic) FNAC results should be managed cautiously. Risk of malignancy appears very low if there are none of the four established sonographically suspicious features and follow-up can be recommended without a need for a second FNAC. However, if there are sonographically suspicious features, repeat FNAC or surgery should then be considered. Further prospective studies with a larger sample size are required to consolidate our recommendation.

REFERENCES


