CrossMark

HEAD AND NECK

A retrospective study of ultrasound and FNA cytology investigation of thyroid nodules: working towards combined risk stratification

Zi Wei Liu[1] · **Richard Fox[1]** · **Samit Unadkat[1]** · **Roy Farrell[1]**

Abstract The British Thyroid Association recommended in new guidelines on thyroid cancer treatment [Kwak et al. (Korean J Radiol 14:110–117, 2013)] that ultrasound grading of thyroid nodules should be incorporated into MDT management. A retrospective study was carried out to determine that the impact of US grading has had on MDT decision making in practice. The design used in the study is a retrospective review of case notes. The study was carried out in the hub hospital for thyroid cancer in the North west London Cancer network. We included consecutive patients referred to the regional thyroid multidisciplinary meeting between August 2014 and May 2015 for investigation of thyroid nodules. Data were collected on patient demographics, co-morbidity, thy grading, ultrasound grading, surgery, post-operative histology, and radioactive iodine treatment details. Accuracy of cytology and ultrasound in diagnosing malignancy was correlated to definitive histology. 99 patients with thyroid nodules were included in the study. 97% of patients had at least one fine needle aspiration and 75% had ultrasound grading. Thy3f (Bethesda IV) nodules were more likely to be carcinoma if associated with a U4 grade rather than U3 (67 vs 18%, $p = 0.028$). Ultrasound grading has recently been introduced to the standard practice in investigation of thyroid nodules. Further assessment of the accuracy of ultrasound grading in clinical practice may allow us to risk-stratify thy3a/thy3f (Bethesda III/IV) lesions and personalise treatment.

Keywords Thyroid cancer · Thyroid surgery · Fine needle cytology

✉ Zi Wei Liu
  zwl20cam@gmail.com

[1] Northwick Park Hospital, Watford Rd, Harrow HA1 3UJ, UK

Introduction

Ultrasound assessment of thyroid nodules provides a reliable and reproducible assessment of whether a nodule is likely to be malignant. It is only recently that a validated ultrasound classification of nodules (U1–U5) has made comparisons of such assessments possible [1, 2].

Approximately 40% of the population have one or more thyroid nodules, but thyroid cancer is rare, representing 1% of all malignancies with 2700 new cases and 250 deaths recorded in the UK in 2011 [3]. The incidence of thyroid cancer is rising over time mainly due to better diagnostic modalities and incidental findings of micropapillary thyroid cancer [4]. Surgery remains the mainstay of treatment for thyroid cancer. Over-investigation and over-treatment of early thyroid cancer remain a significant concern, and molecular genetic studies are underway to distinguish aggressive thyroid neoplasms from the ones that exhibit indolent behaviour [5]. Current clinical risk stratification tools in use include consideration of age, size of tumour, multifocality, and nodal status [1, 6, 7].

Thyroid cytopathology has been the cornerstone of thyroid nodule assessment for years. Reported sensitivity and specificity range from 60 to 95%. Multiple validated classification systems are available, the most commonly used in the UK being the Thy classification 2007 [1, 3]. Worldwide, the Bethesda System for reporting thyroid pathology is commonly used [8]. The classifications are interchangeable in clinical practice. The reported risk of malignancy associated with each Thy grade is as follows: Thy 1 (Bethesda I), non-diagnostic (4%), thy 2 (Bethesda II), benign (0–3%), Thy3a (Bethesda III, atypical, 5–15%), Thy3f (Bethesda IV, suspicious of follicular neoplasm, 15–30%), Thy 4 (Bethesda V, suspicious of malignancy, 60–75%), and Thy 5 (Bethesda VI, diagnostic of malignancy, 97–99%) [3, 8].

Consequently, suggested management for each of these categories is as follows, taking into account patient and local MDT preferences: Thy 1/Bethesda I, repeat FNA; Thy2/Bethesda II, clinical follow-up/repeat FNA at interval; Thy3a/Bethesda III, repeat FNA. Thy3f/Bethesda IV, surgical lobectomy; Thy4/Bethesda V, lobectomy or total thyroidectomy; and Thy5/Bethesda VI, total or near-total thyroidectomy [1].

Recent BTA guidelines suggest that all patients being investigated for possible thyroid cancer should undergo ultrasound guided FNA of the neck [1], and our thyroid cancer MDT has recently introduced routine review of U-grading of all ultrasound neck scans performed for thyroid nodules. However, the clinical significance and prognostic value of intermediate ultrasound grades (U3–U4) remain uncertain, and decision making becomes difficult when conflicting FNA cytology is present (e.g., U4 grade with Thy2/Bethesda II cytology). Currently, our MDT uses cytology as the primary decision-making tool and a suspicious ultrasound grade would warrant a repeat FNA, although there is good evidence that ultrasound alone is highly sensitive and specific for detecting malignancy [9, 10]. This retrospective study aims to evaluate the utility of ultrasound grading in the context of investigating thyroid nodules and to produce an estimate of risk of malignancy when U-grade and Thy classification of nodules are considered in combination, which will enable a better risk stratification and informed patient choice.

Methods

Ethical considerations

This is a retrospective study involving review of anonymised case records. No patient identifiable data are included, and ethical approval is not required.

The study included 99 patients, whose case was discussed at the regional thyroid cancer MDT between August 2014 and May 2015. The case notes were reviewed by three independent clinicians and data collected on patient demographics, FNA cytology, ultrasound grading, tumour staging (if applicable), surgical procedures, operative histology (if available), and MDT recommendation.

US FNA were carried out by experienced head and neck radiologists who classified nodules according to U-criteria [1]. Suspicious US features for thyroid nodules were 'taller than wide', ill-defined margins, hypoechogenicity, and microcalcification.

FNA was carried out, where there is clinical suspicion. Nodules with a benign ultrasound appearance (U2) were not subject to FNA unless there is additional clinical indication as per BTA guidelines.

FNAc results were reported by one of three thyroid cytopathologists according to the thy classification. All cases with thy3/4/5 (Bethesda III, IV, V, and VI) cytology had slides reviewed centrally. If there is disagreement between the pathologists, cytology, and histology, slides were sent out for an external opinion.

US and FNA grades were correlated with post-operative histology where available. Incidental finding of micropapillary cancer where this is unrelated to the nodule being investigated is excluded. The Chi-square and Fisher's exact tests were used to compare cancer rates between two groups. All statistical analyses were carried out on R [11].

Results

A total of 99 patients were included in the final analysis. The age at diagnosis ranged from 24 to 88, with a mean age of 52. 63 patients underwent surgery and had a histologically confirmed diagnosis. The rest remain under clinical follow-up. No patients died during the 8 month study period. There were 35 cancer diagnoses in the final sample, of which 14/35 (40%) were papillary, 16/35 (46%) were follicular, and 5/35 (14%) were anaplastic, medullary, or lymphoma. The relatively high rate of cancers in comparison with benign histology is due to the study being conducted in a cancer MDT setting, where only cases with cytology thy3f (Bethesda IV) and above is routinely reviewed. Incidental findings of micropapillary carcinomas were not routinely reviewed at MDT and this accounts for the relatively high proportion of follicular carcinomas in our study.

Fine needle aspiration findings

In our study, 96/99 patients had at least one graded fine needle aspiration cytology. 13% were thy 1/Bethesda I, 10% were thy 2/Bethesda II, 25% were thy3a/Bethesda III, 38% were thy3f/Bethesda IV, and 10% were thy4/5 (Bethesda V and VI). In the thy3a category, 7/25 repeat FNAs were thy3a (Bethesda III), whereas 3/25 were downgraded to thy2/Bethesda II on second aspirate. For thy3f/Bethesda IV lesions, five repeat aspirates were thy3a/Bethesda III and one was downgraded to thy2/Bethesda II. No thy4/5 (Bethesda V/VI) nodules had repeated aspiration.

The cytology results were correlated to surgical histology. 100% of thy4/5 (Bethesda V/VI) lesions were malignant. 20% of Thy3a (Bethesda III) lesions were malignant, and 34% of thy3f (Bethesda IV) lesions were malignant. The thy3a/thy3f (Bethesda III/IV) malignancies are mostly follicular carcinomas, and the thy4/5 (Bethesda V/VI) lesions are mostly papillary/medullary cancers. These figures are broadly in line with published data.

US grading findings

74 of 99 patients had a U-grading on their ultrasound report. Of these, 10 were U2, 41 U3, 19 U4, and 3 in the U5 category (Table 1). In the U3 category, 19/41 patients had histological correlation and 4 of 19 were malignant. In the U4 category, 14/19 had histology and 10 of the 14 were malignant. 100% of the U5 lesions were malignant. The MDT recommendations for U3, U4, and U5 lesions were as follows: U3: 73% hemithyroidectomy, 27% repeat US/FNAC and U4/U5: 100% hemithryoidectomy or total thyroidectomy.

Combined FNA/US findings

The most interesting sub-categories are thy3a/U3, thy3f/U3, and thy3f/U4, as highlighted in Table 1.

Thy3a (Bethesda III)/U3 is a low clinical suspicion category. Only 3/13 underwent surgery, and all three had benign lesions. Of the 16 patients with a Thy3f (Bethesda IV)/U3 nodule, 11/16 underwent surgical resection, of which 2/11 had a thyroid malignancy (18%). The Thy3f (Bethesda IV)/U4 category carried much higher risk of malignancy. 9/11 had surgery, of which 6/9 had cancer (67%), see Fig. 1. The vast majority was follicular carcinoma. There is nearly a threefold difference in risk of malignancy between the thy3f (Bethesda IV)/U3 and thy3f (Bethesda IV)/U4 categories (p value = 0.028, significance level <0.05). There is no significant difference in rates of malignancy between the thy3a (Bethesda III) and thy3f (Bethesda IV)/U3 groups ($p = 0.73$, significance level <0.05).

Discussion

Thyroid nodules with intermediate gradings are difficult to manage in clinical practice. A majority of thy3a/thy3f (Bethesda III/IV) nodules will be benign, and large numbers of patients may have been subjected to surgery for investigation of benign disease. Recurrent laryngeal nerve injury remains a small but significant cause of long-term morbidity post-thyroid surgery.
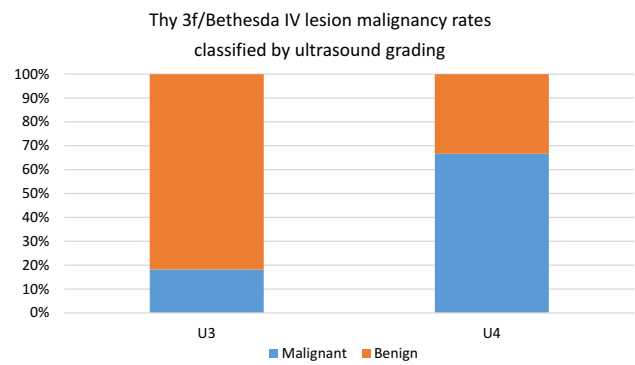

Fig. 1 Thy 3f/Bethesda IV lesion malignancy rates classified by ultrasound grading

Currently, thy3a/Bethesda III nodules (which carry a 10–15% risk of malignancy) are managed with repeat US FNA [1]. Only a small number of thy3a/Bethesda III nodules undergo hemithyroidectomy usually on grounds of compressive effect, nodule growth, or clinical suspicion. Hemithyroidectomy is generally recommended in thy 3f/Bethesda IV nodules which are thought to have a ~30% risk of malignancy. Our study shows that thy3f (Bethesda IV)/U3 nodules have a similar rate of malignancy as the lower risk thy3a (Bethesda III) group, and arguably may be managed expectantly. Thy3f (Bethesda IV)/U4 nodules, however, have a 67% risk of malignancy, and may warrant more prompt investigation and treatment. Thy3f (Bethesda IV)/U3 nodules constitute 52% of our thy3f (Bethesda IV) group. If these results are replicated by larger studies, it may reduce the number of hemithroidectomies performed by 50%.

The strengths of this study include use of consecutive patients discussed at MDT, and high data completion rates for cytology (97%) and ultrasound grading (75%). The cytology and ultrasound results are independently reviewed by expert thyroid pathologists and radiologists. Our cancer rates based on cytology and radiology confirm and are consistent with findings from the previous studies [1, 3].

Weaknesses of this study include small numbers of benign pathology (low numbers of thy2/Bethesda II and U2 nodules) as these do not need to be discussed at our

Table 1 Number of patients with combined FNA/US categories

US grading	Thy grading					
	Thy1/Bethesda I	Thy2/Bethesda II	Thy3a/Bethesda III	Thy3f/Bethesda IV	Thy4/Bethesda V	Thy5/Bethesda VI
U2		1	4	4		
U3	7	6	13	16		
U4	2		3	11		2
U5					2	1

MDT unless there is additional clinical indication. However, the management of low-risk thyroid nodules (thy2/U2) is well established and this is unlikely to change with wider use of ultrasound grading. There are also relatively few thy3a/Bethesda III nodules with histological confirmation of cytology/ultrasound findings, but this will be a consistent feature in all retrospective studies into thyroid nodules as thy3a/Bethesda III nodules are usually not excised. Our study is preliminary in nature and larger numbers are required in this category to determine whether our results are representative.

Another potentially interesting clinical question would be whether ultrasound grading alone would be sufficient to make a clinical decision in patients with recurrent thy1/Bethesda I FNAs. These are currently managed with repeat US FNA. None of our thy1/Bethesda I patients went on to have surgical excision of their nodule, and our study is too small to answer this question, but this could be another instance, where ultrasound would complement cytology in a tricky clinical scenario.

In this study, we confirm the utility of cytological and ultrasound grading of thyroid nodules. We also attempt to combine the thy- and U-grades to establish better risk stratification. We suggest that the thy3f (Bethesda IV) category should be split into thy3f (Bethesda IV)/U3 and thy3f (Bethesda IV)/U4, as these two groups have significantly different rates of malignancy, and the lower risk thy3f (Bethesda IV)/U3 group may be suitable for clinical observation or repeat US/FNAc, depending on patient and clinician preference.

Compliance with ethical standards

Conflict of interest All authors declare that they have no conflict of interest.

Ethical approval This article does not contain any studies with human participants or animals performed by any of the authors.

References

1. Perros P, Boelaert K, Colley S, Evans C, Evans RM, Gerrard Ba G, Gilbert J, Harrison B, Johnson SJ, Giles TE, Moss L, Lewington V, Newbold K, Taylor J, Thakker RV, Watkinson J, Williams GR, British Thyroid Association (2014) Guidelines for the management of thyroid cancer. Clin Endocrinol (Oxf). 81(Suppl 1):1–122
2. Kwak JY, Jung I, Baek JH, Baek SM, Choi N, Choi YJ, Jung SL, Kim EK, Kim JA, Kim JH, Kim KS, Lee JH, Lee JH, Moon HJ, Moon WJ, Park JS, Ryu JH, Shin JH, Son EJ, Sung JY, Na DG, Korean Society of Thyroid Radiology (KSThR), Korean Society of Radiology (2013) Image reporting and characterization system for ultrasound features of thyroid nodules: multicentric Korean retrospective study. Korean J Radiol 14(1):110–117
3. Guidelines on the reporting of thyroid cytology specimens (2016) The Royal College of Pathologists. http://ukeps.com/docs/thyroidfna.pdf Accessed 15 June 2016
4. Konturek A, Barczyński M, Stopa M, Nowak W (2016) Trends in prevalence of thyroid cancer over three decades: a retrospective cohort study of 17,526 surgical patients. World J Surg 40(3):538–544
5. Xing M, Clark D, Guan H, Ji M, Dackiw A, Carson KA, Kim M, Tufaro A, Ladenson P, Zeiger M, Tufano R (2009) BRAF mutation testing of thyroid fine-needle aspiration biopsy specimens for preoperative risk stratification in papillary thyroid cancer. J Clin Oncol 27(18):2977–2982
6. Nixon IJ, Wang LY, Migliacci JC, Eskander A, Campbell MJ, Aniss A, Morris L, Vaisman F, Corbo R, Momesso D, Vaisman M, Carvalho A, Learoyd D, Leslie WD, Nason RW, Kuk D, Wreesmann V, Morris L, Palmer FL, Ganly I, Patel SG, Singh B, Tuttle RM, Shaha AR, Gönen M, Pathak KA, Shen WT, Sywak M, Kowalski L, Freeman J, Perrier N, Shah JP (2016) An international multi-institutional validation of age 55 years as a cutoff for risk stratification in the AJCC/UICC staging system for well-differentiated thyroid cancer. Thyroid 26(3):373–380
7. Omry-Orbach G (2016) Risk stratification in differentiated thyroid cancer: an ongoing process. Rambam Maimonides Med J 7(1). doi:10.5041/RMMJ.10230
8. Cibas ES, Ali SZ (2009 Nov) The bethesda system for reporting thyroid cytopathology. Thyroid 19(11):1159–1165
9. Lee YH, Kim DW, In HS (2011) Differentiation between benign and malignant thyroid solid thyroid nodules using an ultra-sound classification system. Korean J Radiol 12:559–567
10. Hambly NM, Gonen M, Gerst SR (2011) Implementation of evidence-based guidelines for thyroid nodule biopsy: a model for establishment of practice standards. Am J Roentgenol 196:655–660
11. R Core Team (2016) R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/. Accessed 12 Feb 2016