Clinical Radiology 74 (2019) 702–711

Contents lists available at ScienceDirect

Clinical Radiology

journal homepage: www.clinicalradiologyonline.net

Radiological–pathological correlation of the British Thyroid Association ultrasound classification of thyroid nodules: a real-world validation study

H. Al-Chalabi [a,b], S. Karthik [a], S. Vaidyanathan [a,*]

[a] Leeds Cancer Centre, St James's University Hospital, Beckett Street, Leeds, LS9 7TF, UK
[b] Department of Radiology, York Teaching Hospital NHS Foundation Trust, Wigginton Road, York, YO31 8HE, UK

ARTICLE INFORMATION

Article history:
Received 28 December 2018
Accepted 24 May 2019

AIM: To evaluate the real-world performance of the British Thyroid Association (BTA) U classification, specifically focusing on radiology–pathology correlation and to glean learning points.

METHODS AND MATERIALS: Adults undergoing a neck ultrasound for thyroid nodules were reviewed over a period of 1-year. Data including demographics, nodule characteristics, BTA grading, and cytology/histopathology were retrieved with a minimum 24-month follow-up.

RESULTS: Of 1,225 graded nodules in 964 patients, cytology and/or histology were available for 300 (24%). 57 cancers were detected. Of 24 (2%) U5 nodules, 14 were malignant, of 51 (4%) U4, 22 were malignant, of 256 (21%) U3, 20 were malignant, and from 894 (73%) U2 nodules, one cancer was discovered. BTA U grading with fine-needle aspiration (FNA)/core biopsy achieved 96.5% sensitivity, 93.7% specificity, and 93.9% accuracy compared to excision. There was no association between nodule size and rate of malignancy.

CONCLUSION: This is the first study to validate the use of the BTA U-grading system in UK clinical practice. The BTA U-grading system is a robust and reliable method of evaluating the risk of malignancy in thyroid nodules with a high negative predictive value. Key learning points gleaned from the study were accurate assessment of nodule echogenicity, careful evaluation of solid–cystic nodules, optimising ultrasound technique, and the low-risk nature of U3 nodules.

Introduction

There is a growing epidemic of incidentally detected thyroid nodules, in part due to the increased use of high-resolution ultrasound imaging.[1,2] Less than 10% of these are cancer, and most are clinically non-significant microcarcinoma, which have no impact on morbidity or mortality.[3,4] Indeed, this disconnect is evidenced by a two-to fifteen-fold increase in thyroid cancer incidence worldwide over the past three decades without a change in mortality.[4] The biggest increase in incidence by far is in the detection of clinically insignificant microcarcinoma, mainly attributable to ultrasound-guided sampling.

* Guarantor and correspondent: S. Vaidyanathan, Department of Nuclear Medicine, Level 1, Bexley Wing, St James's University Hospital, Beckett Street, Leeds, LS9 7TF, UK. Tel.: +44 0113 2068212; fax: +44 0113 2068228.
E-mail address: srivaidyanathan@gmail.com (S. Vaidyanathan).

https://doi.org/10.1016/j.crad.2019.05.026
0009-9260/

International guidelines have tried to rationalise the use of ultrasound-guided fine-needle aspiration (FNA) based on typical imaging discriminators of malignancy such as marked hypoechogenicity, lobulated margins, lack of perinodular halo, taller-than-wide morphology, intra-nodular vascularity, and microcalcifications.[5,6]

In July 2014, the British Thyroid Association (BTA) published clinical guidelines on the management of thyroid cancer, in which they recommended the use of a U1 to U5 grading system for assessing risk of malignancy and guiding FNA cytology (FNAC).[7] There is a paucity of data in the literature correlating the use of this U-grading system with pathology results in a clinical setting. The aim of the present study was to evaluate the real-world performance of the BTA U classification in a high-volume tertiary referral centre in the UK, specifically focusing on the radiology–pathology correlation of thyroid nodules. In addition, cases where there was a discrepancy between imaging and final histology were re-evaluated to glean learning points to inform future guidelines.

Materials and methods

Study population

The initial study population comprised 2,835 consecutive adult patients (age 16 and over) who underwent ultrasound of the neck at a large teaching hospital trust. Patients were identified retrospectively from the radiology information system (Computerised Radiology Information System, CRIS; Health Software Systems, Mansfield, UK) over a 12-month period between January 2015 and December 2015. This was after the implementation of the 2014 BTA guidelines and to allow a 2-year follow-up period. All patients had a minimum of 24-month clinicoradiological follow-up obtained by reviewing electronic hospital records (Patient Pathway Manager, PPM+, Leeds, UK). Patient Pathway Manager is an electronic system for patient clinic letters, operation notes, discharge summaries and pathology results. The electronic patient record and radiology information systems are automatically updated in the event of patient death so the incidence of patient mortality within the study period was readily available. The centre hosts the regional multidisciplinary team (MDT) and all thyroid cancer cases are discussed; however, individual patient feedback from the general practitioner (GP) was not obtained, and it is possible that some loss to follow-up may have occurred as a result. The authors' institution does not require formal ethics committee approval or written patient consent for anonymised retrospective studies such as this. The study was registered and added to the institutional clinical audit database.

A target nodule was defined as any nodule evaluated by the initial ultrasonologist, reviewing consultant radiologist, and/or at regional thyroid cancer MDT discussion. A total of 1,526 target nodules were identified in 964 patients. Nodule grading was determined from the ultrasound report and MDT documentation. Patients who had not had their thyroid gland scanned (n=801), had undergone a prior thyroidectomy (n=66), and those undergoing follow-up of a known thyroid nodule (n=10) were excluded. Patients who had ultrasound with no nodule detected were also excluded. The first index scan for those patients with multiple examinations during the study period was included for analysis. Subsequent follow-up scans were used in identification of management, but were not considered as new study participants. Cytology and histopathology results of nodules that underwent FNA (in the radiology department or clinic), core biopsy and excision were retrieved.

Nodule grading and analysis

Nodule grading was performed prospectively by the sonologist at the time of scanning. The scans were not read again retrospectively to avoid confirmation bias and to reflect clinical practice in a real-world thyroid service. If there was a discrepancy between the frontline scan and the MDT grading, the latter was chosen as correct and feedback offered to the sonologist.

The BTA guidelines categorise nodules based on their ultrasound features into a U1–U5 grading system in a similar way to the Breast Imaging-Reporting and Data System (BIRADS) classification in breast imaging (BTA thyroid guidelines). To summarise briefly, the categories of grading are as follows: U1, normal thyroid; U2, benign; U3, indeterminate; U4, suspicious; U5, malignant. This is meant to mirror the "Thy" pathology grading system as described below. As per local practice, nodules in a multinodular gland with no suspicious features (defined as U3 or above) were classified as BTA U2 for the purposes of this study.

U2 and U3 nodules with a subsequent histological diagnosis of malignancy were re-assessed by two consultant head and neck radiologists independently to look for any missed signs of cancer. The radiologists were blinded to the final histopathology of these lesions during nodule assessment. The same radiologists also subsequently reviewed the U5 nodules with an ultimate benign diagnosis to assess for common themes or learning points. Any differences in opinion were resolved through consensus.

Pathological grading

Pathological grading of FNA cytology was performed independently by pathologists, following the clinical or radiological-guided procedure, with a report issued and assignment of a grading between Thy1 and Thy5 as described by the Royal College of Pathologists guidance on the reporting of thyroid cytology specimens (2009). This document has most recently been updated in January 2016.[8] In summary, the numerical categories are as follows: Thy1/Thy1c, non-diagnostic for cytological diagnosis; Thy2/Thy2c, non-neoplastic; Thy3/Thy3a/Thy3f, neoplasm possible; Thy4, suspicious of malignancy; Thy5, malignant. For the purposes of analysis, a final histological diagnosis was defined as either malignant histology at core biopsy or excision, or a benign FNA with no progression on follow-up.

Outcomes

The primary outcome was to analyse and evaluate the correlation of the radiological U-grading system with final pathology obtained from surgical excision biopsy results. As a secondary outcome, the accuracy of the radiological U-grading on a per nodule basis was also evaluated.

Statistical analysis

Descriptive statistics were used for data from all patients. All patients were included in the analyses, which were performed on an intention-to-treat basis. Statistical significance in all analyses was set at $p<0.05$ (two-tailed). All statistical analyses were performed using SPSS version 17.0 (SPSS, Chicago, IL, USA) and vector graphics were created on GraphPad Prism 5.0 for Macintosh (GraphPad Software, San Diego, CA, USA). Inter-rater agreement between the initial practitioner BTA U-grade and the ultimate radiological U-grade determined at the thyroid MDT meeting was tested with Cohen's kappa coefficient. This was not a prospective blinded assessment. Although in the majority of cases, the radiology was reviewed independent to knowledge of the histopathology, the variance introduced in real-world execution means that true blinding to pathology results was not achievable. Any bias would be expected to negatively impact inter-rater agreement due to an unexpected histopathological finding influencing the MDT radiologist's interpretation in favour of different risk stratification.

Results

Study population

Nodule summary is illustrated in Fig 1. Following exclusion of patients apropos the methods section, 964 patients with 1,225 graded nodules were included in the study. Commonly patients had more than one graded nodule; the maximum number in any one patient was four. The median (range) age of this group was 55 (17–95) years; of which 82% were female. Demographics are noted in Table 1.

Median (IQR) follow-up was 35 (32–38) months. No unexcised U2 or U3 nodules were subsequently declared as malignant during this period; however, the aforementioned dependence of follow-up on the electronic patient record and radiology information systems is a limitation of the present study.

Primary outcome

Within the study population, 57 cancers were detected. This encompassed 14 (61%) in the U5 group, 22 (47%) of U4 nodules, 20 (8%) in the U3 group and one (0.001%) in the U2 group. Of 1,225 graded nodules identified in a total of 964 patients, cytology and/or histology were available for 300 (24%).

24 of the 1,225 (2%) nodules were graded as U5 and histology was available for 23. The patient who did not receive a histological confirmation died of lung cancer prior to biopsy/treatment. This case was excluded from analysis. Fourteen of the 23 (61%) U5 nodules had a final malignant diagnosis (Fig 2). A typical example of a U5 malignant nodule is demonstrated in Fig 3. The details of the patients with nine nodules with a U5 grade but a subsequent benign diagnosis, i.e. with a radiology–pathology discrepancy are summarised in Electronic Supplementary Material Table S1. On subsequent review, in almost all cases of U5 graded nodules with an ultimate benign diagnosis, colloid mimicked the appearance of microcalcification. All nine examples were re-graded from U5 on analysis; Fig 4a and b show an example of such a mislabelled nodule in contrast to a true U5 nodule as shown in Fig 4c and d.

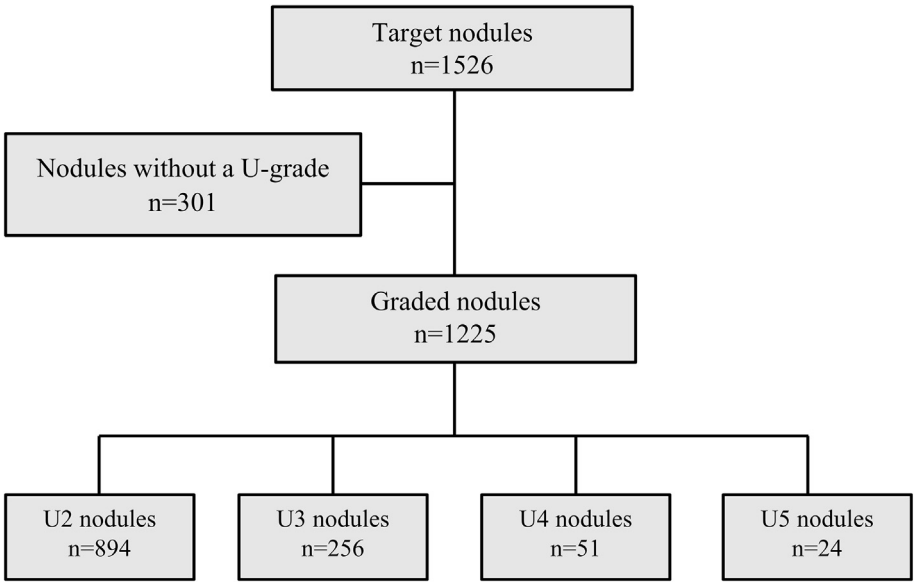


Figure 1 Illustration of target nodules. A total of 1,526 target nodules were identified, 1,225 of which had a recorded U-grade in the initial report, following specialist radiologist review or at MDT. N, number of patients; some patients had more than one target nodule; n, number of nodules; US, ultrasound.

Table 1
Analysed patient demographics and nodule characteristics.

Variable	Value
Age, years, median (range)	55 (17–95)
Female, *n* (%)	788/964 (82%)
Multinodular goitre, *n* (%)	456/964 (47%)
Nodule size, mm, median (range)	13 (1–80)
U2 nodules, *n* (%)	894/1225 (73%)
U3 nodules, *n* (%)	256/1225 (21%)
U4 nodules, *n* (%)	51/1225 (4%)
U5 nodules, *n* (%)	24/1225 (2%)
Nodules sampled at FNA, core biopsy, (hemi)thyroidectomy or a combination of techniques, *n* (%)	300/1225 (24%)

U4 nodules accounted for 51/1,225 (4%). Cytology and/or histology results were available for 47/51 (92%). 3 nodules that were not sampled at FNA, core biopsy, (hemi)thyroidectomy, or a combination of techniques were in patients who died of another advanced malignancy prior to further management, and the fourth was lost to follow-up; these nodules were excluded from subsequent analysis. 22 out of the sampled 47 (47%) U4 nodules were malignant.

Of 1,225 the nodules, 256 (21%) were U3. 196 were sampled with FNA, core biopsy, (hemi)thyroidectomy, or a combination of techniques. Of the 60 nodules not sampled,

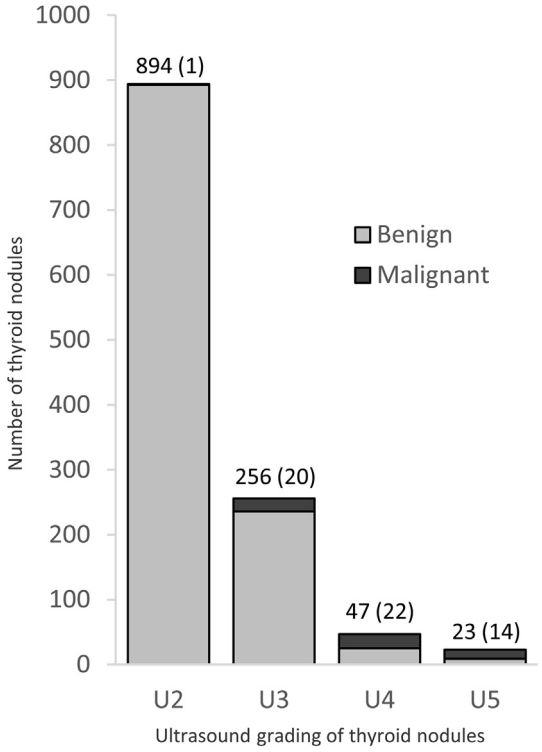
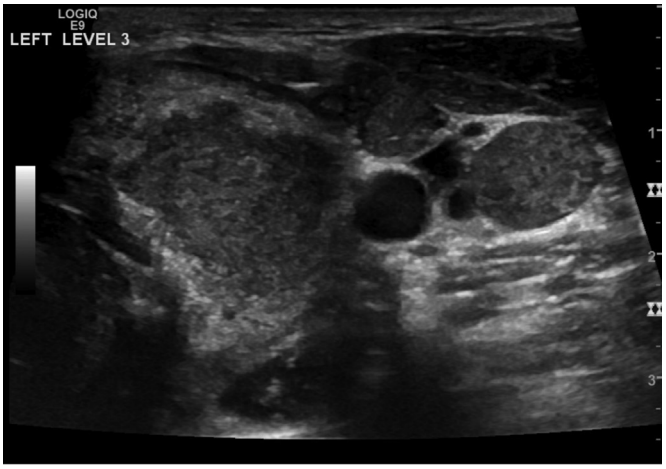
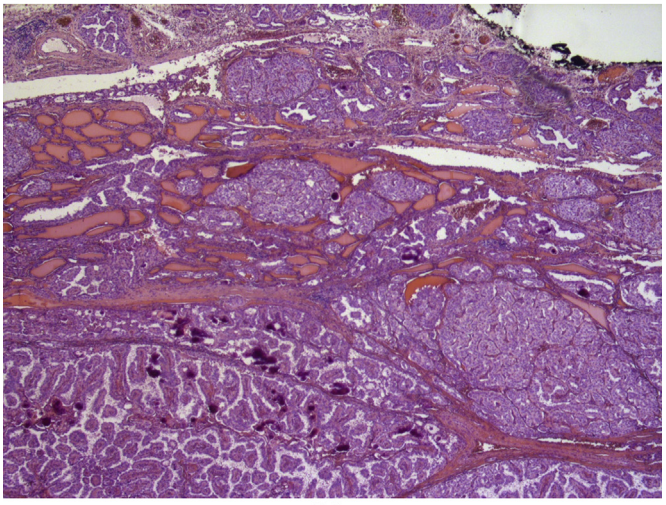


Figure 2 Number of benign and malignant nodules in each BTA U-grade. Unsampled U2 and U3 nodules and U3 nodules with Thy1 cytology/core biopsy considered "benign" provided no cancers declared themselves in the follow-up period. Four U4 nodules and one U5 nodule that did not undergo sampling with FNA, core biopsy, (hemi)thyroidectomy or a combination of techniques were excluded from analysis.



(a)



(b)

Figure 3 (a) Ultrasound image of a hypo-echoic, irregular, taller-than-wide, U5 left lobe of thyroid nodule with associated malignant lymphadenopathy. (b) The histopathology specimen shows solid pattern as well as classical papillary thyroid cancer pattern tumour components and a positive margin.

30 had a decision to opt for watch and wait with ultrasound follow-up, 12 were lost to follow-up, 8 were longstanding stable nodules, 6 patients declined sampling/subsequent surgery, and 4 had a concurrent cancer undergoing active management. None of these unsampled cases were declared as cancer in the follow-up period, and for the purposes of this study, were considered benign. This resulted in a U3 malignancy rate of 20/256 (8%). Imaging characteristics, pathology, potential reason for mis-categorisation, and learning points for U3 nodules with a subsequent malignant diagnosis are provided in Electronic Supplementary Material Table S2. An example of one of these U3 labelled nodules, which had a Thy3f result on FNAC and an ultimate diagnosis of minimally invasive follicular carcinoma, is shown in Fig 5.

Of the 1,225 nodules, 894 (73%) nodules were graded as U2. As per the BTA guidelines, U2 nodules were not routinely sampled at the authors' institution; however,

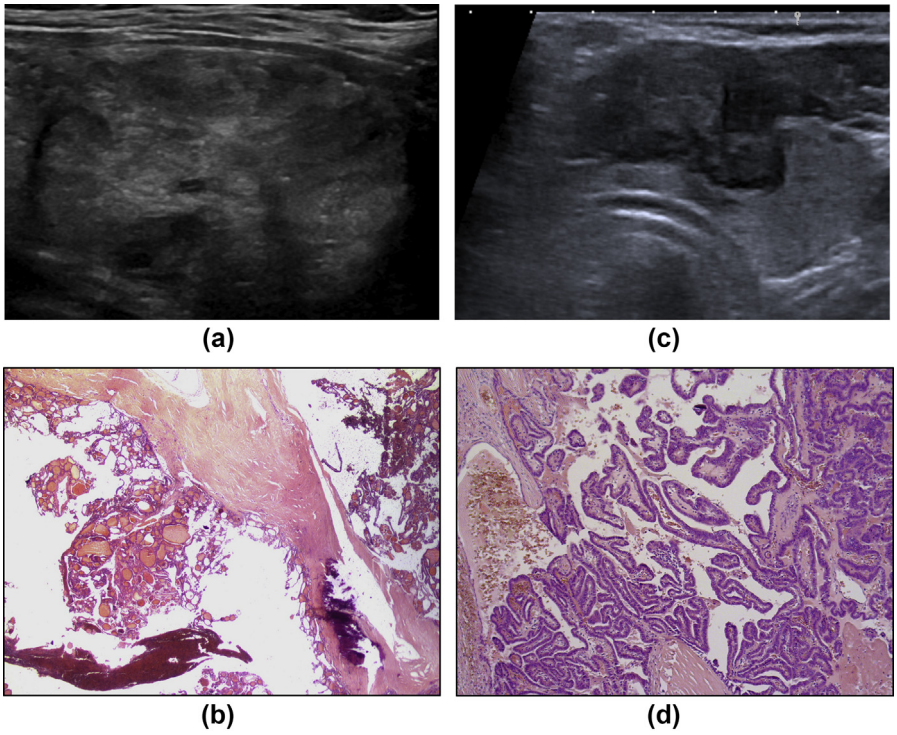
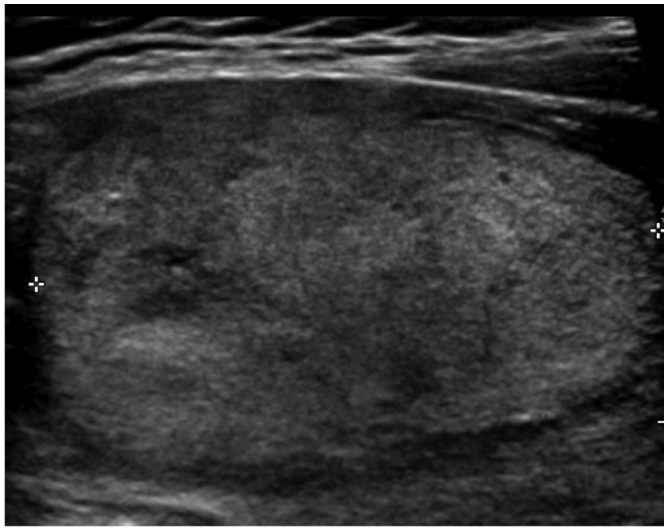


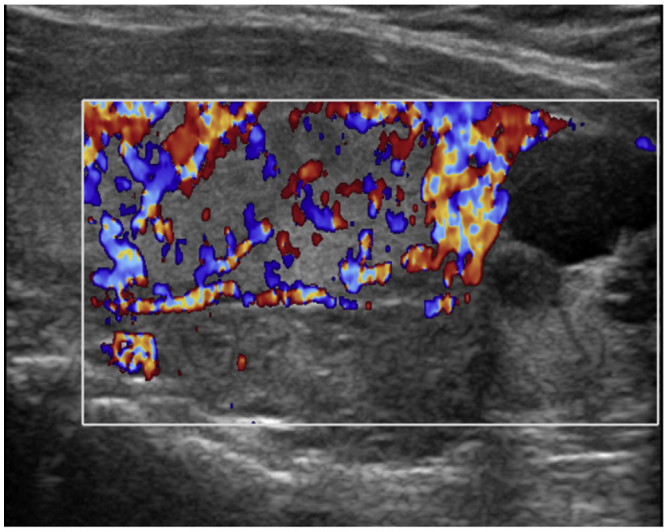
Figure 4 (a) B mode ultrasound image displaying a left lobe of thyroid solid nodule with incomplete halo incorrectly labelled as U5 on initial scanning due to the misinterpretation of colloid as microcalcification. This was proven to be a benign hyperplastic nodule on histological review following diagnostic hemithyroidectomy. Feedback was provided to the department sonographers with guidance on turning off spatial compounding in cases of equivocal hyperechogenic foci. (b) Histopathology specimen of the above nodule showing no evidence of malignancy. (c) In contrast, a typical BTA U5 solid, hypo-echoic nodule with an irregular outline consistent with malignancy. This lesion demonstrates extra-thyroidal extension. (d) Histological assessment following thyroidectomy of this true U5 nodule confirmed the presence of papillary thyroid carcinoma.

histology was available for 34 U2 nodules due to the following reasons: 29 underwent thyroidectomy for pressure symptoms, three due to other risk factors (i.e., MEN2, family history, "age"), one hyperfunctioning nodule, and one at the patient's request. Incidental microcarcinomas measuring no more than 3 mm were found in three of the hemithyroidectomy/thyroidectomy specimens and no further treatment was recommended for these incidental findings; the graded nodules were therefore considered benign. One of the 894 (0.1%) nodules graded as U2 was found to be a papillary carcinoma (Fig 6). In retrospect, the authors thought that this nodule was mislabelled and should have been classified as at least U4 if not U5, likely occurring due to not appreciating the hypoechoic appearance and microcalcification being misinterpreted as colloid. This could have also been partially as a result of the spatial compounding feature available on the ultrasound machine, which is enabled by default in the neck preset, which reduces post-acoustic shadowing. Feedback was given to the sonographer through usual departmental channels and the case highlighted in the learning from discrepancy meeting.

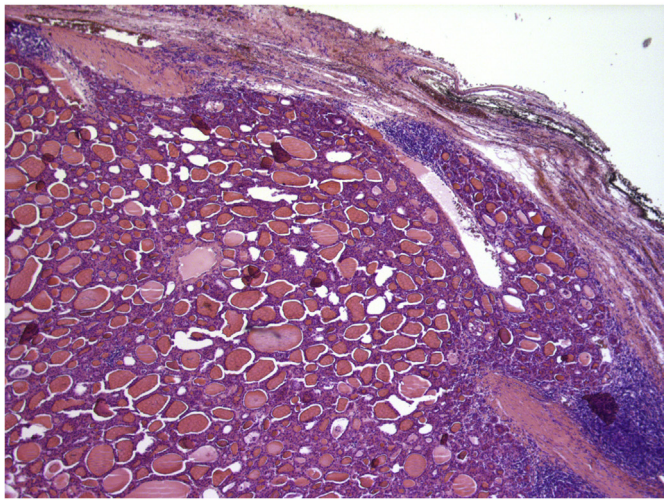
To determine the radiology–pathology correlation with final histological diagnosis, we considered the following nodules which would routinely undergo excision/hemithyroidectomy at St James's University Hospital as "radiology–pathology positive": U5 nodules, U4 nodules with Thy1, Thy3, Thy4, or Thy5 FNA/core biopsy results, U4 nodules with a Thy2 FNA/core biopsy on a single sample only, and U3 nodules with a Thy3, Thy4, or Thy5 FNA/core biopsy. "Radiology–pathology negative" nodules were considered as: all U2 nodules, U3 nodules with a Thy2 result on FNA/core biopsy, U3 nodules with Thy1 FNA/core biopsy results with no cancer on follow-up, and U4 nodules with Thy2 confirmed on two FNA/core biopsy specimens. Of the 1,092 radiology–pathology negative nodules, 117 (11%) had a conclusive cytohistological diagnosis; 34 U2 nodules with final histology, six U2–Thy2 nodules, 73 U3–Thy2 nodules, and four U4 nodules with two Thy2 results. There were no U5 nodules with Thy2 results confirmed on FNA/core biopsy on two specimens. Of the 1,092 radiology–pathology negative nodules, 975 (89%) were considered benign based on the absence of a malignant presentation during the follow-up interval. Using this method, 2/1092 radiology–pathology negative nodules and 55/128 radiology–pathology positive nodules were ultimately malignant and 1090/1092 radiology–pathology negative nodules and 73/128 radiology–pathology positive nodules were found to be benign (Table 2; sensitivity 96.5%, specificity 93.7%, positive predictive value [PPV] 43%, negative predictive value [NPV] 99.8%, accuracy 93.9%).



(a)



(b)



(c)

Figure 5 A U3 labelled nodule summarised as Thy3f following FNAC with an ultimate diagnosis of pT2 pNx pR0 pV0 minimally invasive follicular carcinoma. (a) B mode ultrasound image showing a predominantly iso/hyper-echoic left lobe of thyroid nodule with small areas of low echogenicity raising suspicion for cystic degeneration. (b) Colour Doppler image demonstrating marked central vascularity. (c) Histopathology specimen showing the capsular invasion focus of a follicular carcinoma.

Secondary outcomes

As mentioned above, 57 malignancies were found in total. Of these, 6/57 (11%) were in nodules <10 mm, 14/57 (25%) in nodules between 10–19 mm, 16/57 (28%) in nodules between 20–29 mm, 6/57 (11%) in nodules between 30–39 mm, and 15/57 (26%) in nodules ≥40 mm (Fig 7). Thirty-seven U3 nodules measuring <1.5 cm underwent FNA and three of these had a subsequent malignant diagnosis on histology. Nine U4 nodules measuring <1 cm underwent FNA and six of these had a subsequent malignant diagnosis on histology (Fig 8).

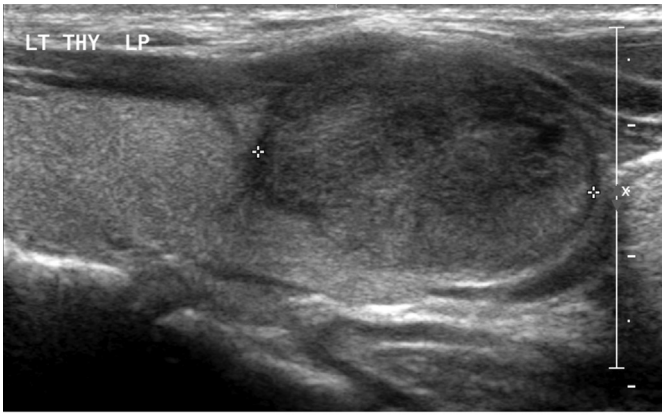
Inter-rater agreement between the initial BTA U grading and the final consensus grading done at the thyroid MDT was deemed excellent with a kappa coefficient of 0.95 (95% confidence interval [CI}: 0.93 to 0.97, $p<0.001$).

Discussion

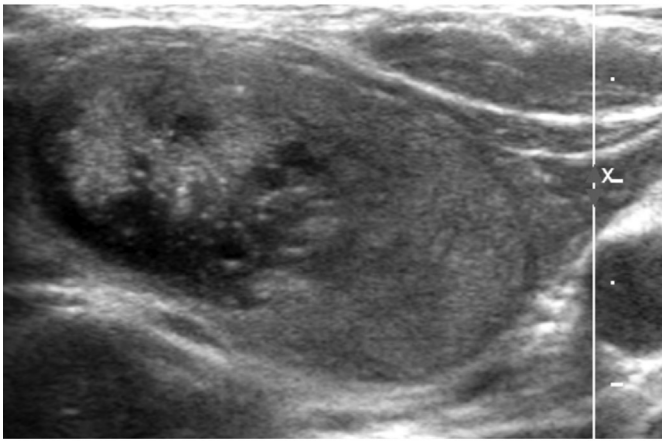
The results of the present study demonstrate a good correlation between the BTA U1–5 ultrasound classification and final histology in a real-world setting. To the authors' knowledge, this is the first study to evaluate and validate the U-grading system in a non-screening clinically referred patient population.

The overall radiology–pathology agreement was very good and the test sensitivity, specificity, and accuracy were 96.5%, 93.7% and 93.9%, respectively (Table 2) with a high NPV (99.8%). The present findings are also generalisable to a real-world practice, due to the varied personnel performing thyroid ultrasound including otorhinolaryngologists, sonographers, radiology trainees, and radiologists. The BTA guidelines[7] were published in July 2014, and a steep learning curve was experienced during the initial 6 months after their implementation when information transmission and dissemination was at its peak. Therefore the current study period beginning in 2015 was chosen to obviate this effect, but also, to have a suitable length of follow-up of at least 24 months.

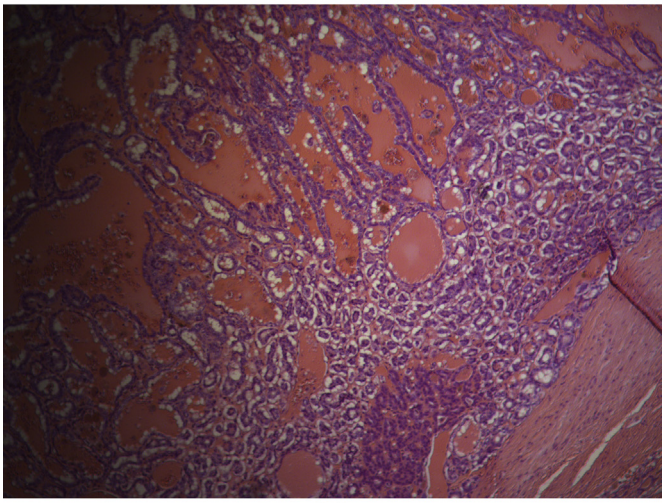
The present data showed that 8% of U3 nodules were found to be malignant. These data are similar to those published in the literature.[9] The American Thyroid Association (ATA) guidelines have labelled nodules with U3 characteristics as "low-risk", which the present authors feel is a more accurate descriptor as opposed to "indeterminate". This is because the term "indeterminate" usually has a connotation of equipoise, i.e., a roughly equal probability of being benign or malignant. In the present study, U4 nodules had an approximately 50% cancer positivity rate. Again this prevalence is concordant with the literature, and as in the ATA guidelines, such nodules are more accurately labelled as "intermediate risk". In the authors day-to-day experience, U3 nodules with cancer and U4 nodules with



(a)



(b)



(c)

Figure 6 A single nodule within the cohort labelled as U2 declared itself as malignant. (a,b) Ultrasound images of a single nodule within the cohort labelled as U2, which was subsequently declared as malignant. Microcalcification was misinterpreted as colloid. (c) Histopathology specimen of the same nodule demonstrating nuclear features of papillary thyroid carcinoma.

a benign FNA/core biopsy can often be viewed in the MDT as discrepancies due to the labelling of these as indeterminate and suspicious, respectively. From the surgeon's perspective, counselling patients for surgery in these cases (U4 Thy2 or U3 Thy3f) is challenging given the 50–90% benign rate and attendant risks with the operation, including life-altering ones such as recurrent laryngeal nerve injury. Therefore relabelling the descriptor of U3 nodules from "indeterminate" to "low risk" and U4 from "suspicious" to "intermediate risk" similar to the ATA guidelines would be more appropriate.

Table 2
Summary of radiology–pathology correlation with final histological diagnosis.

	Benign[a]	Malignant[b]	
Radiology–pathology negative[c]	1,090	2	1,092
Radiology–pathology positive[d]	73	55	128
	1,163	57	1,220

Sensitivity 96.5%, Specificity 93.7%, PPV 43.0%, NPV 99.8%, Accuracy 93.9%.
[a] Benign: nodules benign on histology and unsampled U2 and U3 nodules with no declaration of malignancy in the ≥24 month follow-up period.
[b] Malignant: nodules malignant on histology.
[c] Radiology–pathology positive: U5 nodules, U4 nodules with Thy1, Thy3, Thy4 or Thy5 FNA/core biopsy results, U4 nodules with a Thy2 FNA/core biopsy on a single sample only and U3 nodules with a Thy3, Thy4 or Thy5 FNA/core biopsy.
[d] Radiology–pathology negative: U2 nodules, U3 nodules with a Thy2 result on FNA/core biopsy, U3 nodules with Thy1 FNA/core biopsy results with no cancer on follow-up and U4/U5 nodules with Thy2 confirmed on two FNA/core biopsy specimens.

As shown in Electronic Supplementary Material Table S2, a large proportion of U2 and U3 nodules with subsequent malignant histology were labelled incorrectly. In the majority of cases the nodules were mildly hypoechoic and more in keeping with a U4 grade. This has subsequently informed the authors' clinical practice since the study period.

Another diagnostic conundrum is the mixed cystic–solid nodule. In the vast majority of cases, these are benign; however, occasionally these turn out to be a cystic papillary carcinoma, which is a variant of papillary cancer (e.g. Fig. 9). In the present study, two cystic papillary carcinomas were labelled as U3 nodules (Electronic Supplementary Material Table S2). Although the BTA guidelines mention the features of cystic papillary carcinoma in the text, this is not sufficiently emphasised in the final graphic representation. In the authors' experience, it is not common to find U5 features in the solid component of such lesions, i.e., it is rare to find marked hypoechogenicity and microcalcification in the solid element. More commonly, the solid component is iso-echoic without macrocalcifications. In such an event, certain features can help raise suspicion for a cystic papillary cancer. These include a non-dependent mural nodule, solid element comprising >50% of the lesion, lobulation within the solid component, and hypervascularity of the solid element.[10,11] Such nodules should be excised in spite of a Thy1 FNA or histology result. This highlights the need for a greater emphasis on this diagnosis in subsequent versions of the guideline.

In the present cohort, only 1/894 (0.001%) thyroid nodule with a U2 grade was subsequently diagnosed as malignant, but this was felt to have been mislabelled due to the misinterpretation of colloid as microcalcification,

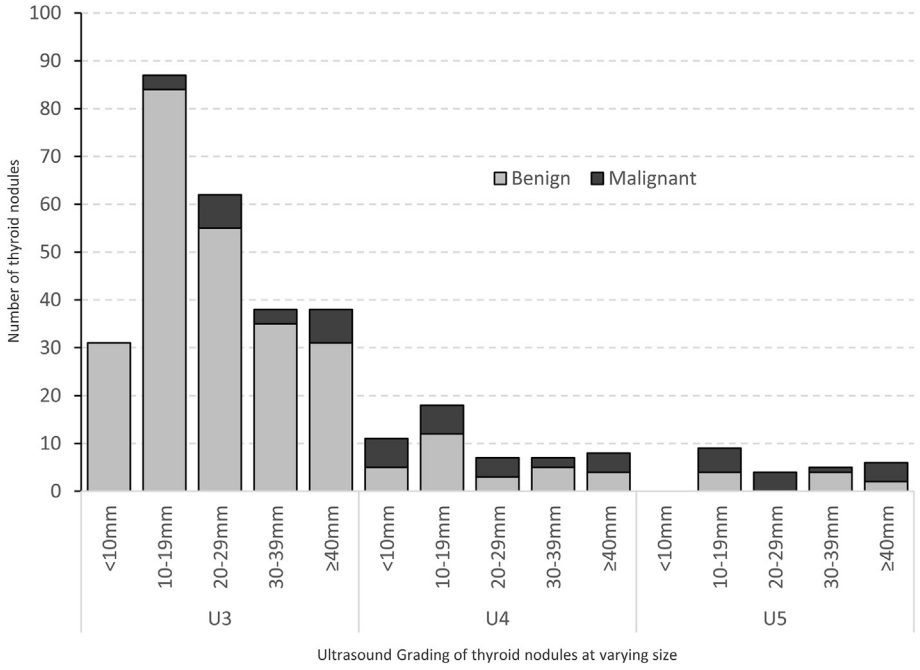


Figure 7 Distribution of benign and malignant U2, U3, and U4 nodules per nodule size range.

which is sometimes a diagnostic challenge in clinicoradiological practice, but should be a rare occurrence if spatial compounding is turned off and other features are taken into account. In the present cohort, after at least 24 months of follow-up, no unsampled U2 nodule assessed as benign resulted in a subsequent diagnosed malignancy. Evaluation of U5 nodules with a final benign histology (Electronic Supplementary Material Table S1) revealed misinterpretation of colloid as microcalcification in the majority of cases. The findings of the present study indicate that the BTA U-grading system provides a robust method of excluding malignancy in thyroid nodules with the absence of sinister features. Moreover, analysis of radiology–pathology correlation reveals a NPV of 99.8% when combining the U-grading system with results from FNA ± core biopsy.

The results not only show a positive correlation between U-grade and resultant rate of malignancy, but also

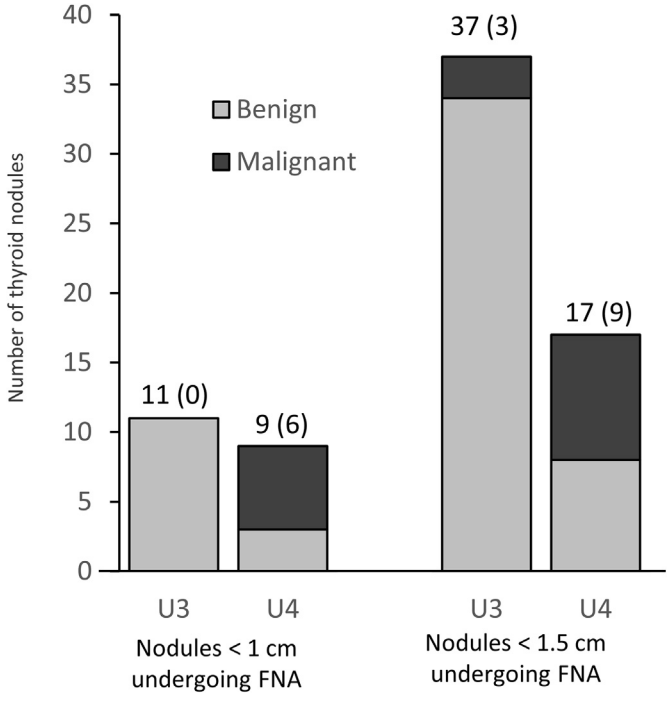


Figure 8 Number of U3 and U4 nodules measuring <1 and 1.5 cm, which underwent FNA (number of those malignant in brackets).

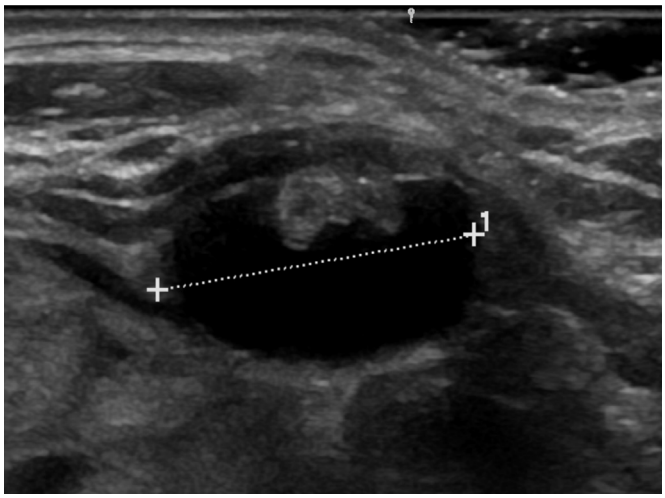
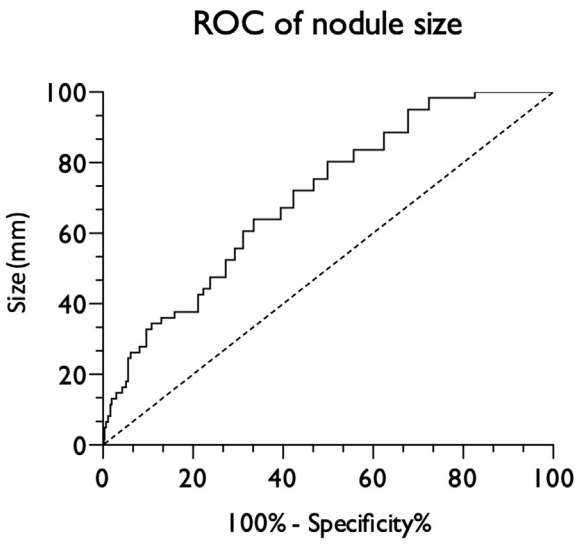


Figure 9 Ultrasound image of a histologically confirmed cystic papillary thyroid cancer demonstrating features of a non-dependent mural nodule and lobulation within the solid component. Such nodules should be graded at least as U3 and undergo appropriate evaluation; they must not be dismissed as benign.



	Results
Area under the ROC curve	
Area	0.7038
Std. Error	0.03134
95% confidence interval	0.6424 to 0.7653
P value	<0.0001

Figure 10 An ROC curve demonstrates a poor correlation between nodule size and rate of malignancy across all U-grades.

demonstrate the importance of judicious sampling of these nodules to maximise detection of cancers whilst minimising unnecessary sampling or surgery. Some centres have a one-stop clinic with a cytopathologist and radiologist/sonographer in clinic, but this is not commonplace. A multidisciplinary approach is vital in ensuring adequate use of limited resources and minimising patient anxiety due to over-diagnosing clinically insignificant cancers.

Certain international guidelines, such as the ATA and American College of Radiologists (ACR) Thyroid Imaging-Reporting And Data System (TI-RADS), have an inbuilt size-based threshold for sampling unlike the BTA guidelines.[9,12] For example, the ATA guidelines advise that FNA is not required for nodules <1.5 cm with a low suspicion pattern (U3 equivalent) or for nodules <2 cm with a very low suspicion pattern (U2 equivalent).[12] The ACR TI-RADS recommendation for TR3 (mildly suspicious) nodules is FNA if ≥2.5 cm and follow if ≥1.5 cm.[9] These nodules approximately equate to a U3 category, e.g., category U3(b) of the BTA guidelines: ?hypo-echoic nodules (TI-RADS 2 points) with cystic change (TI-RADS 1 point). In the present cohort, 37 U3 nodules measuring <1.5 cm underwent FNA; three of these had a subsequent malignant diagnosis on histology. The ATA guidelines also recommend that, in general, only nodules >1 cm should be evaluated, as they have a greater potential to be clinically significant cancers. In the present cohort, nine U4 nodules measuring <1 cm underwent FNA and six of these had a subsequent malignant diagnosis on histology. It is, of course, unclear what the natural history of the disease would have been if the lesions had not been excised. This is outwith the scope of the present study.

The rate of malignancy at varying nodule size and receiver operating characteristic (ROC) curve analysis across all U-grades emphasises that size is not an accurate criterion for assessing risk of malignancy (Figs 7 and 10). This concurs with the assertion in the BTA guidelines that thyroid nodule size is a poor indicator of malignancy and validates the move to base the guidelines on key red-flag imaging parameters as opposed to size.

The study has limitations. This was a retrospective study, but initial U-grade at report or MDT were used so that any bias was minimised and the study reflects real-world practice as scans are often performed by a variety of personnel. Agreement between the initial ultrasound grading and the final MDT grading was excellent in the present study. This was not a prospective blinded assessment, and due to the inherent limitations of real-world practice, blinding of the thyroid MDT radiologist to pathology results at the time of final consensus grading could not be guaranteed in all cases. Any bias would be expected to negatively impact inter-rater agreement due to an unexpected histopathological finding influencing the MDT radiologist's interpretation in favour of different risk stratification. As recommended in the BTA guidelines, the vast majority of U2 nodules did not undergo sampling and were considered benign; this resulted in a small number of U2 nodules with cytology/histology available. There was, therefore, a potential bias towards a benign diagnosis without cytological or histological "proof" of benign disease; however, no unsampled U2 nodule re-presented as a malignancy within the 24-month follow-up period.

Previous studies, for example, a paper by Kwak *et al.*, 2010 have shown that repeat FNAC should be performed for thyroid nodules that have suspicious ultrasound features, even if the initial cytological results indicate it to be a benign lesion.[13] This methodology is also adopted by the BTA guidelines and a Thy2 cytology result in a U4 or U5 nodule necessitates repeat sampling to ensure benignity. For this reason, there is the potential inherent bias for suspicious nodules to undergo more thorough pathological investigation than their innocent appearing counterparts, and for U3 nodules to be less well interrogated in comparison. Despite a large patient population in the present tertiary referral centre, the number of U4 and U5 nodules is relatively small. It is well-recognised that there are variable confounding factors that increase the possibility of a non-diagnostic pathology result following FNAC,[14] and although the reasons for this were not investigated in the current study, multiple U3 nodules that underwent FNA and/or core biopsy resulted in a non-diagnostic sample. This resulted in a reduction of nodules available for correlation of a combination of ultrasound and cytology results with final histology.

In conclusion, the present study is the first to validate the use of the BTA guidelines ultrasound grading system in routine clinical practice in the UK. We demonstrated a high inter-rater reliability, the ability of the guidelines to classify thyroid nodules as benign with a high NPV and the

importance of a pattern-based rather than a size-based approach in diagnosing thyroid cancer. Important learning points gleaned from the study were accurate assessment of nodule echogenicity, careful evaluation of solid-cystic nodules, optimising ultrasound technique, and the low-risk nature of U3 nodules.

Conflicts of interest

The authors declare no conflict of interest.

Acknowledgements

The authors thank Dr Preetha Chengot, Consultant Histopathologist, from Leeds Teaching Hospitals NHS Trust for the pathology images provided.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.crad.2019.05.026.

References

1. Pellegriti G, Frasca F, Regalbuto C, *et al*. Worldwide increasing incidence of thyroid cancer : update on epidemiology and risk factors. *J Cancer Epidemiol* 2013;**2013** https://doi.org/10.1155/2013/965212.
2. Davies L, Welch HG. Increasing incidence of thyroid cancer in the United States, 1973–2002. *JAMA* 2006;**295**(18):2164–7 https://doi.org/10.1001/jama.295.18.2164.
3. Brito JP, Morris JC, Montori VM. Thyroid cancer: zealous imaging has increased detection and treatment of low risk tumours. *BMJ* 2013;**347**:f4706 https://doi.org/10.1136/bmj.f4706.
4. Davies L. Overdiagnosis of thyroid cancer. *BMJ* 2016;**355**:i6312 https://doi.org/10.1136/bmj.i6312.
5. Yoon JH, Han K, Kim E-K, *et al*. Diagnosis and management of small thyroid nodules: a comparative study with six guidelines for thyroid nodules. *Radiology* 2017;**283**:560–9 https://doi.org/10.1148/radiol.2016160641.
6. Moon WJ, Jung SL, Lee JH, *et al*. Benign and malignant thyroid nodules: US differentiation—multicenter retrospective study. *Radiology* 2008;**247**(3):762–70 https://doi.org/10.1148/radiol.2473070944.
7. Perros P, Colley S, Boelaert K, *et al*. Guidelines for the management of thyroid cancer. *Clin Endocrinol (Oxf)* 2014;**81**(Suppl):1–122 https://doi.org/10.1111/cen.12515.
8. Cross P, Chandra A, Giles T, *et al*. *Guidance on the reporting of thyroid cytology specimens January 2016*. The Royal College of Pathologists; 2016 January. Available at: https://www.rcpath.org/uploads/assets/uploaded/9ddf3c1d-c58f-4b8c-b89b63e0704f5a50.pdf. [Accessed 18 December 2018].
9. Tessler FN, Middleton WD, Grant EG, *et al*. ACR thyroid imaging-reporting and data system (TI-RADS): white paper of the ACR TI-RADS committee. *J Am Coll Radiol* 2017;**14**(5):587–95 https://doi.org/10.1016/j.jacr.2017.01.046.
10. Perros P, Colley S, Boelaert K, *et al*. Guidelines for the management of thyroid cancer. *Clin Endocrinol (Oxf)* 2014;**81**:1–136 https://doi.org/10.1111/cen.12515.
11. Hoang JK, Lee WK, Lee M, *et al*. US Features of thyroid malignancy: pearls and pitfalls. *RadioGraphics* 2007;**27**(3):847–65 https://doi.org/10.1148/rg.273065038.
12. Haugen BR, Alexander EK, Bible KC, *et al*. 2015 American thyroid association management guidelines for adult patients with thyroid nodules and differentiated thyroid cancer: the American thyroid association guidelines task force on thyroid nodules and differentiated thyroid cancer. *Thyroid* 2016;**26**(1):1–133 https://doi.org/10.1089/thy.2015.0020.
13. Kwak JY, Koo H, Youk JH, *et al*. Value of US correlation of a thyroid nodule with initially benign cytologic results. *Radiology* 2010;**254**(1):292–300 https://doi.org/10.1148/radiol.2541090460.
14. Kavanagh J, McVeigh N, McCarthy E, *et al*. Ultrasound-guided fine needle aspiration of thyroid nodules: factors affecting diagnostic outcomes and confounding variables. *Acta Radiol* 2017;**58**(3):301–6 https://doi.org/10.1177/0284185116654331.