Diagnostic Utility of Molecular and Imaging Biomarkers in Cytological Indeterminate Thyroid Nodules

Elizabeth J. de Koster, MD[1], Lioe-Fee de Geus-Oei, MD Phd[2], Olaf M. Dekkers, MD PhD[3,4], Ilse van Engen-van Grunsven, MD Phd[5], Jaap Hamming, MD PhD[6], Eleonora P.M. Corssmit, MD PhD[3], Hans Morreau, MD PhD[7], Abbey Schepers, MD PhD[6], Jan Smit, MD PhD[8], Wim J.G. Oyen, MD PhD[9,1], Dennis Vriens, MD PhD[2]

1. Dept. of Radiology and Nuclear Medicine, Radboud University Medical Center, Nijmegen, the Netherlands.
2. Dept. of Radiology, section of Nuclear Medicine, Leiden University Medical Center, Leiden, the Netherlands.
3. Dept. of Endocrinology, Leiden University Medical Center, Leiden, the Netherlands.
4. Dept. of Epidemiology, Leiden University Medical Center, Leiden, the Netherlands.
5. Dept. of Pathology, Radboud University Medical Center, Nijmegen, the Netherlands.
6. Dept. of Surgery, Leiden University Medical Center, Leiden, the Netherlands.
7. Dept. of Pathology, Leiden University Medical Center, Leiden, the Netherlands.
8. Dept. of Endocrinology, Radboud University Medical Center, Nijmegen, the Netherlands.
9. Division of Radiotherapy and Imaging, Institute of Cancer Research, and Department of Nuclear Medicine, Royal Marsden Hospital, London, United Kingdom.

Short title: Diagnostic biomarkers in indeterminate thyroid nodules

Key words: indeterminate thyroid cytology, imaging, molecular, diagnostic

Corresponding author / address reprint requests to:

E.J. de Koster, M.D.

Radboudumc

Department of Radiology and Nuclear Medicine

P.O. Box 9101

6500 HB Nijmegen

The Netherlands

E-mail: Lisanne.deKoster@radboudumc.nl

Phone: +31 71 5297625 / +31 6 14607347

Fax: +31 71 5248256

Grants supporting the writing of the paper: the corresponding author receives a research grant from the Dutch Cancer Society (DCS).

Disclosures: none.

Abstract

Indeterminate thyroid cytology (Bethesda III and IV) corresponds to follicular-patterned benign and malignant lesions, which are particularly difficult to differentiate on cytology alone. As approximately 25% of these nodules harbor malignancy, diagnostic hemithyroidectomy is still custom. However, advanced preoperative diagnostics are rapidly evolving.

This review provides an overview of additional molecular and imaging diagnostics for indeterminate thyroid nodules in a pre-operative clinical setting, including considerations regarding cost-effectiveness, availability, and feasibility of combining techniques. Addressed diagnostics include gene mutation analysis, microRNA, immunocytochemistry, ultrasonography, elastosonography, CT, sestamibi scintigraphy, FDG-PET and diffusion-weighted MRI.

The best rule-out tests for malignancy were the Afirma® GEC and FDG-PET. The most accurate rule-in test was sole *BRAF* mutation analysis. No diagnostic had both near-perfect sensitivity and specificity, and estimated cost-effectiveness. Molecular techniques are rapidly advancing. However, given the currently available techniques a multimodality stepwise approach likely offers the most accurate diagnosis, sequentially applying one sensitive rule-out test and one specific rule-in test. Geographical variations in cytology (e.g. Hürthle cell neoplasms) and tumor genetics strongly influence local test performance and clinical utility. Multidisciplinary collaboration and implementation studies can aid the local decision for one or more eligible diagnostics.

Precis

This review discusses the value of additional molecular and imaging diagnostics for thyroid nodules with indeterminate cytology, including considerations regarding cost-effectiveness, availability, and feasibility of combining techniques. Addressed diagnostics include gene mutation analysis, microRNA, immunocytochemistry, ultrasonography, elastosonography, CT, sestamibi scintigraphy, FDG-PET and diffusion-weighted MRI.

Contents of this review

INTRODUCTION

Indeterminate thyroid cytology is an eyesore to physicians. It largely corresponds to histopathologically follicular-patterned lesions, both benign and malignant, including follicular adenoma, noninvasive follicular thyroid neoplasm with papillary-like nuclear features (NIFTP), (encapsulated) follicular variant of papillary thyroid carcinoma (FVPTC or EFVPTC) and follicular thyroid carcinoma (FTC). These neoplasms are particularly difficult to differentiate on fine needle aspiration cytology (FNAC). In the case of FTC, cytology lacks the insight into the tissue structure like histology does: it does not show the capsular and/or vascular invasion that distinguishes a FTC from a benign FA. In FVPTC, the growth pattern is follicular and clearly identifying nuclear features of PTC can usually not be identified cytologically (1-3). Nevertheless, FNAC currently has a most prominent place in the diagnostic work-up of thyroid nodules. The Bethesda System for the Reporting of Thyroid Cytology was adopted in its current form in 2009, recognizing six diagnostic categories with an incremental risk of malignancy and clinical management guidelines. Although the Bethesda system created a much-used handhold by standardizing the cytological diagnosis and consecutive management of thyroid nodules worldwide, the system does not provide a

clear answer for the heterogeneous group of nodules with indeterminate cytology (4,5). This includes cytology with atypia of undetermined significance or follicular lesion of undetermined significance (AUS/FLUS, Bethesda III), and cytology (suspicious for a) follicular neoplasm (SFN/FN) or (suspicious for a) Hürthle cell neoplasm (SHCN/HCN, Bethesda IV). Similar indeterminate cytological categories are found in the British Thyroid Association Thy system and Italian SIAPEC-IAP classification: Thy3a and Thy3f, and TIR3A and TIR3B, respectively **(Table 1)** (6,7).

Alongside a doubled incidence of thyroid carcinoma over the past two decades and a prevalence of thyroid nodules stretching far beyond the 5% for palpable nodules – explained by the incidental detection of nonpalpable nodules and clinically occult thyroid cancers on imaging studies – the need for a more accurate diagnostic procedure has grown (8). This urge was further emphasized when other research groups were unable to reproduce the prevalence of the cytological categories and corresponding malignancy risks proposed by Cibas *et al.*, especially those of the AUS category (4,9,10). Insuperable variations in the worldwide patient populations, and intra- and interobserver variation in the assessment of thyroid cytology were named as likely underlying causes (4,5,10,11). Yet, it raised questions concerning the overall approach of thyroid nodule diagnosis and whether cytology is the best starting ground. Cost-effectiveness is a major benefit of cytological examination, yet a more accurate test may eventually replace cytological examination completely (12,13). At present, however, a supplemental diagnostic procedure is specifically warranted for cytologically indeterminate thyroid nodules. Diagnostic hemithyroidectomies are still customarily performed to obtain a definite histological diagnosis. With a benign histopathological result in approximately three in four cases, surgery was not only unbeneficial but also exposed the patient to unnecessary surgical risks. In the case of malignant lesions, a second-stage completion thyroidectomy is often indicated, which is associated with additional costs and higher risks of surgical

complications (14-17). An additional preoperative test or combination of tests for thyroid nodules with indeterminate cytology should prevent unbeneficial diagnostic hemithyroidectomies for benign nodules, limit the number of two-stage surgeries for thyroid malignancies, or both. With rapidly advancing technology, the possibilities for additional diagnostic techniques seem endless: the applications of existing diagnostics such as ultrasound, PET/CT and immunocytochemistry are extended and more clearly demarcated for use in indeterminate thyroid nodules. High-tech molecular tests such as gene mutation panels, gene or microRNA expression profiles and sequencing techniques are hot-topic (4,18-22). Every currently known engagement point from the genotype to the phenotype of the tumor is being explored. Combined, the various research fields encompass an extensive range of investigative methods. Individually they usually focus on one or two methods only, making one-to-one comparison of these diagnostics difficult. The 2015 American Thyroid Association (ATA) guidelines suggested several additional tests, but a definitive answer or complete overview of all available tests is still lacking (23).

Alongside higher-level expert discussions and lobbying of med tech companies, clinical endocrinologists and thyroid surgeons ponder about the best solution for their individual patients. Their choices depend on the characteristics of their patient populations, availability and costs of a certain test, and personal preference. In any case, a useful additional test should be accurate, accessible, affordable and affect patient management.

This review aims to provide practical considerations for physicians involved in the management of patients with thyroid nodules. It gives an overview of the available literature on additional diagnostic tests for thyroid nodules with indeterminate cytology. We will work our way down from genotype to phenotype, discussing both anatomical and functional techniques, from the state-of-the-art molecular and imaging biomarkers as well as widely available conventional imaging techniques. The ability of a test to distinguish between

malignant and benign nodules in a preoperative setting is discussed, focusing on clinical validation and utility, and including the development phase, cost-effectiveness and availability of each technique, where appropriate. Table 2 provides a summarized overview of the discussed diagnostics and their main attributes.

1. MOLECULAR BIOMARKERS

1.1. Gene mutation analysis and Gene expression

In the last decades, researchers have unraveled important molecular mechanisms behind the thyroid tumorigenesis, and designated a great number of genetic alterations that are related to the various types of thyroid carcinoma. Several of these mutational markers have found their way to the preoperative diagnosis of indeterminate thyroid nodules. The most common markers are the somatic *BRAF* and *RAS* point mutations, and RET/PTC rearrangement, all of which involve the mitogen-activated protein kinase (MAPK) signaling pathway (24-26).
In the 2015 ATA guidelines the potentially strong diagnostic impact of molecular testing is explicitly unfolded, focusing on BRAF testing and the – at that date – two main commercially available tests: the seven-gene mutation panel miRInform® thyroid (Asuragen Inc., Austin, Texas) and the Afirma® gene expression classifier (Veracyte, Inc., South San Francisco, CA). The ATA recommends considerate application of one of these molecular tests for Bethesda III and IV nodules, provided that the result could change the treatment strategy (23).
In the following chapters, the diagnostic potential of mutation analysis in indeterminate thyroid nodules is discussed, including the tests mentioned in the guidelines as well as other individual molecular biomarkers and multi-gene panels addressed in literature.

1.1.1. BRAF mutation

B-type RAF kinase (BRAF) is a serine–threonine kinase belonging to the rapidly accelerated fibrosarcoma (RAF) family, and the most potent mitogen-activated protein kinase (MAPK) pathway activator. Point mutations in the *BRAF* proto-oncogene occur in various human cancers. The somatic $BRAF^{V600E}$ mutation is the most common activating mutation in many carcinomas, including thyroid carcinoma (24). This missense mutation consists of a thymine-to-adenine substitution at nucleotide 1799 (c.1799T>A), resulting in an amino acid substitution where valine is replaced with glutamate at codon 600 (hence V600E)(27,28). BRAF has an important function in cell proliferation, differentiation, and apoptosis. Upregulation of BRAF through the $BRAF^{V600E}$ activating mutation is associated with tumorigenesis (28). In differentiated thyroid cancer, the $BRAF^{V600E}$ mutation is exclusive to PTC, occurring in 50% to 80% of these tumors (24,25,29-39). The $BRAF^{V600E}$ mutation has been prognostically associated with poor clinicopathological outcomes, such as increased incidence of extrathyroidal invasion, recurrence of disease, and distant metastasis of the tumor (40-42).

BRAF mutation analysis has been extensively studied as a rule-in test for thyroid carcinoma. The *BRAF* mutation is superior to other mutations in its oftentimes 100% specificity – a positive mutation could prevent two-stage surgery for an indeterminate thyroid nodule (21,29-31,35-38,41-76). Even though the *BRAF* mutation was found in a majority of PTC in a number of studies, the prevalence of the *BRAF* mutation in indeterminate cytology ranged from 0% to 48% in individual studies (44,46,48,59,65,70). Reported sensitivities were therefore heterogeneous and generally poor, ranging from 0% to 83% (29,34,39,46). Other types of thyroid carcinoma occurring in indeterminate nodules, including FTC, FVPTC and Hürthle cell carcinoma (oncocytic variant of follicular thyroid carcinoma, FTC-OV), were

respectively never or infrequently *BRAF* mutation-positive (31,37,38,42,50,57,76). Predominated by follicular type carcinoma, the *BRAF* mutation rarely occurs in Bethesda IV cytology (29,31,37,41,50,52,54,57,59,60,63,65,66,70,73-80).

Likely contributors to the observed heterogeneity are known global variations in the occurrence rates of PTC and *BRAF* mutations. In South Korea, where iodine consumption is high, 90% to 95% of thyroid cancers are PTC. More specifically, the proportion of *BRAF*-mutated PTC is very high: rates of 80% to more than 90% are reported (34,46,77). Consequently, $BRAF^{V600E}$ mutation analysis might have both high specificity and high sensitivity in these populations. Studies with higher sensitivities were more often of South Korean origin and frequently demonstrated sensitivity above 40%, with the prevalence of *BRAF* mutations reported as high as 30% to 48%. (34,39,46,47,73,81-83). Conversely, the majority of studies with sensitivity below 10% were conducted in Western countries (USA, Europe or Canada), with some studies reporting no *BRAF* mutations at all (21,31,37,45,48,52,56-59,61,62,65,69,70,76,80).

Some South Korean studies based surgical decision-making on the result of the *BRAF* mutation analysis: surgery was relatively less often performed in *BRAF* mutation-negative indeterminate nodules (34,39,78,83). Such a surgical management strategy is not oncologically safe for Western countries (e.g. Europa or Northern America), where 80% to 90% of thyroid carcinomas are PTC and reported rates of *BRAF*-mutated PTC vary from 30% to 40% (34,46,77). Moreover, even though the true sensitivity of *BRAF* mutation analysis is presumably high in South Korea for the mentioned epidemiological reasons, the conservative management of *BRAF* mutation-negative nodules likely magnified test sensitivity by underestimating the rate of *BRAF*-negative malignant nodules in these studies. Altogether we estimate that approximately one in five South Korean patients would benefit from *BRAF* mutation analysis, opposite mere one in 25 patients from other countries.

BRAF mutation in papillary microcarcinoma

Papillary microcarcinoma (mPTC) have lower *BRAF* mutation rates (53,58,63,68,73,76,84). The ATA guidelines are reserved with regard to the recommended clinical management of positive *BRAF* mutation in mPTC, as its relation to extrathyroidal spread and positive lymph node metastases is not as clear as in larger thyroid carcinoma. Although there are studies that associate mPTC to factors of poorer prognosis, the 2015 guidelines recommend that *BRAF*-mutated mPTC are treated as low-risk malignancies (23,35).

$BRAF^{K601E}$ point mutation

A less common activating *BRAF* mutation is $BRAF^{K601E}$ (c.1801A>G), which occurs considerably less frequently than the $BRAF^{V600E}$ variant and is associated with FVPTC with high specificity (85). Clinically, the characterization of a small cohort of thyroid malignancies with a $BRAF^{K601E}$ mutation showed better outcomes than for $BRAF^{V600E}$ mutated tumors: no extrathyroidal tumor extension, recurrence, lymph node or distant metastasis were reported in indeterminate $BRAF^{K601E}$ positive tumors with a median follow-up of 20 months (range 4-47) (86).

Availability, cost-effectiveness and limitations of BRAF mutation analysis

Altogether, the consistent perfect specificity in a large number of studies supports the use of *BRAF* mutation analysis in obviating two-stage surgery. The technique is increasingly available in the clinical setting worldwide. A prior meta-analysis of eight studies questioned the cost-effectiveness of $BRAF^{V600E}$ mutation analysis in indeterminate thyroid nodules based on a mere 4.6% mean prevalence of the mutation (87). Cost-effectiveness studies concerning sole *BRAF* mutation analysis in indeterminate thyroid nodules are lacking. Regardless, cost-

effectiveness is generally presumed, as average costs for testing are relatively low and decreasing over time. Depending on the applied molecular technique, reported costs for *BRAF* mutation analysis ranged between €7.50 and $123 per tested sample (53,63,72,88).

Low sensitivity remains the main limitation of *BRAF* mutation analysis, irrespective of the type of indeterminate cytology. Proficiency of the test in preoperative patient management depends on the regional occurrence rate of *BRAF*-mutated PTC; in South Korea, more patients will benefit from *BRAF* mutation analysis, and the probability and extent of cost-effectiveness are likely to increase (66). In other health care systems, such as in the UK, cost-effectiveness is likely more constrained. Nonetheless, *BRAF* testing could still save approximately half the surgical costs in *BRAF* mutation-positive carcinoma (63,65). These global variations should be considered before local implementation of sole *BRAF* mutation analysis.

1.1.2. RAS point mutation

Point mutations in the gene family of retrovirus-associated DNA sequences (RAS) together constitute the second most frequently occurring genetic alteration in thyroid carcinoma. In indeterminate thyroid nodules, they are the most common genetic alteration, due to a strong association of *RAS* mutations with the follicular-patterned lesions that make up these cytological categories: follicular adenoma, FTC, FVPTC and noninvasive follicular thyroid neoplasms with papillary-like nuclear features (NIFTP) (1,3,31,59,89,90). Originally, two of the three homologous *RAS* genes were identified as viral genes of the oncogenic Harvey (*HRAS*) and Kirsten (*KRAS*) murine sarcoma virus; the third, *NRAS*, was first identified in neuroblastoma cells (91,92). The genes code for GTP-binding RAS proteins, which are

involved in intracellular signaling in the MAPK/ERK pathway. Mutation causes overactive RAS signaling and could ultimately induce malignant transition (26).

RAS mutation in thyroid carcinoma has been associated with favorable prognostic factors, such as encapsulation of the tumor and absence of lymph node metastases, but also with factors indicative of an adverse prognosis, such as poor cell differentiation (2). *RAS* mutations are not specific for carcinoma and found in both malignant and benign lesions (31,61,90). According to the 2015 ATA guidelines, Bethesda III or IV nodules with a *RAS* mutation should be treated similar to the Bethesda V category, as approximately 4 out of 5 are malignant (4,23). *HRAS, KRAS* and *NRAS* mutations are mutually exclusive. They are each associated with slightly different types of cytology and histology, and consequently a different clinical course. In general, point mutations in *NRAS* codon 61 and *HRAS* codon 61 are said to occur most frequently (3,64). *KRAS* is associated with oncocytic lesions and a lower malignancy rate than other *RAS* mutations (93).

A *RAS* point mutation is found in 0% to 38% of the indeterminate nodules (39,60). Moreover, approximately a third of all reported malignancies resulting from indeterminate thyroid cytology are *RAS* mutation positive, frequently FVPTC or FTC (31,37-39,76). Sporadic cases of *RAS* mutation-positive FTC-OV and MTC are reported (37,38). In individual studies, sensitivity and specificity of *RAS* mutation analysis ranged from 0% to 77% and from 75% to 100%, respectively (39,60,90). Test performance was similar for Bethesda III and IV categories, although the mutation occurred more frequently in Bethesda IV nodules (21,29,31,39,50,59,60,76,80,90). Histopathologically benign nodules carrying a *RAS* mutation are histopathological follicular adenoma in most cases, but also oncocytic variant of follicular adenoma (Hürthle cell adenoma) or hyperplastic nodules (29,31,38,50,90). There is an ongoing discussion regarding the interpretation of a false positive *RAS* mutation. It is presumed that an oncogenic *RAS* mutation predisposes a follicular adenoma for progression

into follicular carcinoma – a *RAS*-mutated follicular adenoma should be considered a premalignant pre-invasive follicular neoplasm. These assumptions put false-positives in a different light, as it would justify resection of such lesions through hemithyroidectomy. Consequently, the lesions could also be considered true-positives – improving the specificity of *RAS* mutation analysis (1,21,39,59,61,71). However, the exact mechanisms behind the malignant potential and transition for *RAS*-mutated follicular adenomas are not yet clarified and difficult to appreciate in a clinical setting.

Similar to *BRAF*, there was evident global variation in the distribution of *RAS* mutations. Many European and American studies reported a clear predominance of *RAS* mutations over *BRAF* mutations. Solely a Brazilian study of 116 Bethesda III and 20 Bethesda IV thyroid nodules reported only *BRAF* mutations and not a single *RAS* mutation (60). The previously described predominance of *BRAF* mutations in South Korean populations was confirmed in the sole study that investigated both point mutations in one population (39). Combined BRAF/RAS mutation analysis could be considered, although geographical differences in the distribution of the two genetic alterations strongly influence feasibility. A gene mutation panel consisting of more genetic alterations (discussed in a next chapter) is most likely more useful.

Sole *RAS* mutation analysis is not accurate in the preoperative setting. Although specificity is high, only two out of three *RAS* mutation positive indeterminate nodules are histopathologically malignant, evidently fewer than assumed and previously described in the ATA guidelines. Therefore, *RAS* mutation positive indeterminate thyroid nodules should be surgically managed with no more than hemithyroidectomy. Whether hemithyroidectomy is justified for *RAS*-mutated follicular adenomas as a precancerous lesion, is yet under debate.

1.1.3. RET/PTC rearrangement

Rearrangements of the *RET* proto-oncogene arise from the fusion of the 3' end of *RET* to the 5' regions of unrelated genes that are expressed in thyroid follicular cells. Proto-oncogene *RET* encodes for a transmembrane receptor with a tyrosine kinase domain; a RET/PTC rearrangement causes inappropriate overexpression of that domain. It activates the MAPK and PI3K/AKT pathways and stimulates malignant transition of the cell through BRAF (94,95). At least 12 different fusion variants have been detected until today, of which RET/PTC1 and RET/PTC3 are the most common. They have a well-known association with PTC. Cases of both rearrangements in a single lesion are also reported (2,94,96,97). RET/PTC rearrangements, especially RET/PTC3, occur more frequently in PTC in children or patients that were exposed to ionizing radiation and are clinically associated with the presence of lymph node metastases (2). Worldwide variations in frequency of RET/PTC rearrangements exist, dependent on demographics and ethnicity. The RET/PTC rearrangement is present in 42% of PTC in Western populations with a predominance of RET/PTC1, and in 37% of PTC in Asian populations with a predominance of RET/PTC3. Without radiation exposure, in female PTC patients RET/PTC1 is predominant (98). The rearrangements are also found in benign nodules, especially in patients that were exposed to ionizing irradiation (29,97). Alike *RAS* mutations, it is assumed to be an activating genetic alteration and it is argued that a histopathologically benign nodule with a RET/PTC rearrangement should be considered a precancerous lesion.

RET/PTC rearrangements are seldom found in indeterminate nodules. In many studies, no RET/PTC translocation was found at all. Most studies investigated RET/PTC in light of a gene mutation panel and paid it no specific attention (21,29,31,37,38,45,49,50,55,59,61,62,64,70,76,80,96). Only Guerra *et al*. solely investigated the RET/PTC rearrangement in 101 thyroid nodules of all cytological categories. In this

Italian study, RET/PTC rearrangements were found in 18 of the 50 PTC (36%) using RT-PCR and Southern-Blot. All these RET/PTC-positive carcinomas were Thy4 or Thy5 nodules on cytology. Among the 24 Thy3 nodules, two nodules with a RET/PTC3 rearrangement were histopathologically benign (96).

Noteworthy, Sapio *et al*. detected two *RET* mutations during their RET/PTC assessments. In contrast to the RET/PTC translocation, RET point mutations are related to sporadic and familial MTC (45,99). Surgery confirmed histopathological MTC in the *RET*-mutated nodules (45).

Even though previous histological studies undeniably associated RET/PTC1 and RET/PTC3 rearrangements to PTC, the low prevalence of the rearrangement in indeterminate cytology is a major downside. Testing exclusively for this genetic alteration in indeterminate nodules is not advantageous, even if issues regarding the number of tested variants and sensitivity of molecular techniques are overcome. The 2015 ATA guidelines only advise RET/PTC testing in context of a gene mutation panel (23).

1.1.4. PAX8/PPARγ rearrangement

The PAX8/PPARγ rearrangement arises from a fusion of the promoter and 5'-coding portion of the thyroid-specific transcription factor *PAX8* gene to the gene of the nuclear receptor peroxisome proliferator-activated receptor γ (PPARγ) (100,101). The role of the product of this translocation – the PAX8/PPARγ fusion protein – is not yet understood, as the DNA binding sites of both original proteins are uniquely preserved in the fusion (101). In the normal thyroid, transcription factor PAX8 is involved in differentiation of thyrocytes and regulation of the expression of thyroid-specific genes encoding thyroperoxidase,

thyroglobulin and the sodium/iodide symporter (102). Nuclear receptor PPARγ has multiple presumed functions, including involvement in the regulation of lipid metabolism, adipogenesis and insulin sensitivity (101,103).

The chromosomal translocation PAX8/PPARγ was first discovered in – and traditionally associated with – FTC and follicular adenoma (100). It is reported in 30% to 45% of FTC and in up to 33% of follicular adenoma (50,104-106). However, several studies have also uncovered varying amounts of FVPTC carrying the translocation, with published rates up to 38% (104,107,108). It has not been reported in benign or malignant Hürthle cell neoplasms (105,109).

PAX8/PPARγ is often related to well-differentiated malignancies with a relatively favorable prognosis. Capsular and vascular invasion are reported to a lesser extent in FTCs with a PAX8/PPARγ rearrangement than in *RAS*-mutated tumors (105). Widely invasive features are not reported. PAX8/PPARγ-mutated FVPTC are mostly encapsulated, following an indolent clinical course with minimal disease recurrence despite the presence of some capsular and vascular invasion at presentation (105-107). In contrast to the *BRAF*, *RAS* and RET/PTC genetic alterations in thyroid carcinoma, the PAX8/PPARγ rearrangement does not involve the RAS-RAF-MAPK pathway. Nikiforova *et al.* hypothesized that oncogenesis of follicular-type tumors likely takes place through two different molecular pathways: a *RAS*-mutation driven and PAX8/PPARγ rearrangement driven pathway (105).

Similar to the RET/PTC rearrangement, the PAX8/PPARγ rearrangement rarely occurred in indeterminate thyroid cytology. Approximately two-thirds of the indeterminate nodules carrying the rearrangement were histopathologically malignant, most often FVPTC or FTC (21,31,37,38,49,55,59,61,62,70,76,80,109). False-positive results corresponded to follicular adenomas (61,62,70). Similar to *RAS* mutations, histopathologically benign PAX8/PPARγ-mutated nodules are likely premalignant lesions, or pre-invasive FTC. Eszlinger *et al.*

observed a microfollicular morphological growth pattern in two of the PAX8/PPARγ-positive follicular adenoma, supporting this hypothesis (61).

Still, PAX8/PPARγ rearrangement is a rare rearrangement associated with (encapsulated) follicular tumors. Similar to RET/PTC rearrangements, the PAX8/PPARγ rearrangement should only be assessed in indeterminate thyroid nodules in combination with more frequently occurring genetic alterations in a gene mutation panel.

1.1.5. Other genetic alterations

hTERT

The enzyme human telomerase is involved in the maintenance of the chromosomes' telomeres, which are essential for cell life and proliferation. The catalytic subunit of telomerase is human telomerase reverse transcription (hTERT). In normal thyroid cells, it is inactive. Inappropriate reactivation is associated with malignancy and inflammatory thyroid disease (110). hTERT promotor mutations were previously observed in both PTC and FTC, sometimes together with a *BRAF* mutation. The mutation is strongly correlated to mortality in differentiated thyroid carcinoma (111). hTERT gene expression is potentially accurate in the preoperative differentiation of indeterminate nodules, with 57% to 88% sensitivity and 75% to 85% specificity demonstrated in two small clinical series of cytological follicular neoplasms (112,113).

TRK

The tyrosine receptor kinase (TRK) rearrangement arises from a translocation of the *NTRK1* gene, which is normally expressed in the central and peripheral nervous system and involved in cell differentiation. The TRK rearrangement is associated with PTC and presumably with

an adverse prognosis, although evidence is limited (114). In feasibility studies in indeterminate thyroid cytology, not a single TRK rearrangement has been detected – it is most likely not a useful marker (45,49,50).

HMGA2

Proteins high mobility group AT-hook (HMGA) 1 and 2 regulate the structure and function of chromatin. Normally only expressed during embryogenesis, the overexpression of HMGA in adult tissues is associated with malignancy (115). Lappinga *et al.* demonstrated that HMGA2 could be a promising additional biomarker. Using ROC curve analysis, a >5.9-fold HMGA2 overexpression had 76% sensitivity and 98% specificity in SFN/FN nodules (116). To date no other studies attempted to validate these results.

Galectin-3 and CD44v6

One Croatian study used RT-PCR to investigate the simultaneous expression of galectin-3 and CD44v6, two molecular biomarkers better known for their application in immunohistochemistry of their expression products (117). CD44v6 normally functions as the cell-surface receptor for hyaluronic acid. Overexpression is found in various human cancers, including thyroid (117,118). In indeterminate thyroid nodules, a positive test for either one of the two biomarkers resulted in 100% sensitivity and 60% specificity. It is presumed that similar results for these markers are achieved with the more economical immunohistochemistry techniques (118,119).

1.1.6. 7-Gene Mutation Panel

Ongoing research in the past years has demonstrated that assessment of individual oncogenic mutations generally has limited clinical utility in indeterminate thyroid cytology. Combining forces of individual genetic alterations into a gene mutation panel, however, likely improves diagnostic accuracy, especially as mutations are mutually exclusive in most cases. These gene mutation panels typically assess the seven genetic alterations – gene mutations as well as gene fusions – that occur most frequently in differentiated thyroid carcinoma, including $BRAF^{V600E}$, $BRAF^{K601E}$, *NRAS* codon 61, *HRAS* codon 61 and *KRAS* codon 12-13 point mutations and RET/PTC1, RET/PTC3 and PAX8/PPARγ gene rearrangements (31,55). The best known panel is the commercially available miRInform® thyroid (Asuragen Inc., Austin, Texas, USA), currently rebranded as the ThyGenX® Thyroid Oncogene Panel (Interpace Diagnostics, Parsippany, NJ, USA). The miRInform® thyroid tests 17 specific genetic alterations in these seven genes (59). It is marketed as a rule-in test for thyroid malignancy.

The first large clinical utility study to investigate the miRInform® thyroid test was published in 2011. Nikiforov *et al.* prospectively included 1,056 FNAC samples, 92% of which had sufficient epithelial cells and nucleic acids to pursue molecular testing. Residual FNAC material was used for mutation analysis – no additional aspirates were required. Unfortunately, surgery was performed for only 461 of 900 (51%) indeterminate thyroid nodules, independent of the test outcome; these operated cases were included in their final analysis. It is not reported whether nonsurgically managed nodules were mutation-positive or -negative. Sensitivity and specificity were 63% and 99% in the 247 Bethesda III nodules, and 57% and 97% in the 214 Bethesda IV nodules, respectively. The authors suggested that the high PPV of the miRInform® thyroid in these indeterminate thyroid nodules (88% and 87%, respectively) warrants a direct total thyroidectomy instead of two-step surgery in patients with a positive test (31).

None of the subsequent studies matched the initially reported excellent specificity. The next industry-sponsored prospective study by Beaudenon *et al.* reported 47% sensitivity and 88% specificity in 80 Bethesda III and IV nodules. Surprisingly, not a single BRAF mutation was detected (59). Valderrabano *et al.* reported not a single mutation in 47 included Bethesda III nodules. Moreover, only 1 of 18 nodules with Hürthle cell cytology in this study tested positive, suggesting that Hürthle cell nodules may carry different mutations than the ones investigated by the miRInform® thyroid (76). Ohori *et al.* demonstrated that genetic alterations less frequently occurred in the textbook colloid-poor Bethesda IV cytology compared to the less common colloid-rich variant. Differences in etiology are unknown, but the authors hypothesized that the two types have subtle histopathological differences. The colloid-rich thyroid carcinomas likely more often develop through the well-known mutations included in the miRInform® thyroid test, whereas mutations that elicit colloid-poor thyroid carcinoma are yet unknown (37).

Simultaneously with the American miRInform® studies, five European studies independently investigated whether a panel of the same 7 genes could reliably be assessed using different methods (38,55,61,62,70). In three separate studies, Eszlinger *et al.* demonstrated that testing was also feasible on routine air-dried FNAC samples from indeterminate thyroid nodules. Over the course of these studies, sensitivity of this method improved from 18% to 49% and specificity from 86% to 93%, respectively. The use of air-dried FNAC samples for mutation analysis could advance the implementation of mutation analysis in daily practice, as specific storage conditions of fresh FNAC samples for mutation analysis are no longer required (38,61,62). Mancini *et al.* showed that high-resolution melting (HRM) analysis is an accurate screening method for the seven genetic alterations, with 56% sensitivity and 90% specificity. HRM is a post-PCR procedure that does not require significant additional resources. This

519 could reserve the costlier direct sequencing procedures solely for samples with abnormal
520 HRM results, thereby reducing the overall costs of mutation analysis (55).
521
522 Overall, reported sensitivities and specificities of a 7-gene mutation panel in indeterminate
523 thyroid nodules ranges from 18% to 69% and 86% to 99%, respectively (22,37,61). It is an
524 adequate diagnostic tool with a high rule-in capacity in indeterminate nodules. Test
525 performance was similar in Bethesda III and Bethesda IV nodules, although the latter more
526 frequently had a positive test result based on the higher prevalence of *RAS* mutations (31,59).
527 Due to the common *RAS* mutations, PPV of the 7-gene mutation panel never exceeds 90% in
528 a range of realistic 15% to 40% prevalence of malignancy. As such it is debatable whether a
529 positive test warrants immediate single-stage total thyroidectomy. It translates into an
530 inappropriate overtreatment in a significant number of patients with a positive test but benign
531 final histology at higher risk of surgical complications and all requiring lifelong levothyroxine
532 supplementation. Deliberate surgical decision-making should consider the underlying positive
533 mutation rather than mere the positive test itself.
534 The limited size of the seven-gene gene mutation panel keeps the costs per test low compared
535 to other, larger molecular panels. Reported prices of the 7-gene mutation panel all concern the
536 commercial miRInform® thyroid (or ThyGenX® Thyroid Oncogene Panel) and range
537 between $425 and $1,700 (120,121). Implementation of miRInform® testing for
538 indeterminate nodules theoretically resulted in a 20% cost reduction in the USA: the
539 prevented two-step surgical procedures would outweigh the added expenses for miRInform®
540 testing and increased number of total thyroidectomies – including those for nodules with a
541 false-positive test (120). In a European setting, treatment and hospitalization costs are
542 generally lower and miRInform® would most likely not be cost-effective (14). However,

these cost-effectiveness studies both adopted the unequalled test performance from the initial key publication – true cost-effectiveness may be less optimistic (14,31,120,121).

1.1.7. Next Generation Sequencing

To improve the sensitivity of the miRInform® thyroid test, the existing 7-gene mutation panel was expanded to include additional gene mutations, fusions and translocations, and a microRNA gene expression panel. In addition, it adopted promising next generation sequencing (NGS) techniques. NGS enables the simultaneous targeted testing for multiple mutations in large gene panels and is faster, more sensitive and more cost-effective than traditional Sanger sequencing and other PCR-based methods (71,80,122). As NGS only requires a very small amount (5-10 ng) of nucleic acids, remainder material from regular FNAC passes suffices and no additional aspirates are required (80,122). The first thyroid-specific NGS-based gene panel was the ThyroSeq® v1, presented in 2013. It detected gene variations in 110 of 145 investigated thyroid cancer tissue samples and 5 of 83 benign specimens. Unfortunately, indeterminate FNAC samples were not analyzed separately in this study. Nonetheless, Nikiforova *et al*. demonstrated that NGS had a very high success rate and could be a promising molecular technique for thyroid FNAC samples (122).

Following the ThyroSeq® v1, the road was paved for further exploration of NGS-based diagnostics. Soon, the ThyroSeq® v2 (CBLPath, Ocala, FL, USA) was developed, with a number of primers for TERT promotor variants added to its panel. It simultaneously tested for point mutations in 13 genes and for 42 types of gene fusion products (80). The ThyroSeq® v2 was tested on 143 Bethesda IV thyroid nodules. Forty-two genetic alterations were found,

most frequently *NRAS*. Diagnostic accuracy of the ThyroSeq® v2 was 92%, with astonishing 90% sensitivity and 93% specificity (80).

More recently, Nikiforov *et al*. tested the ThyroSeq® v2.1 – including point mutations in 14 genes and 42 gene fusion transcripts – in 462 Bethesda III nodules. Based on the promising results of the previous study, surgery was withheld for 362 of 431 ThyroSeq®-negative patients. In the 95 patients with available histopathology, the ThyroSeq® v2.1 demonstrated 91% sensitivity and 92% specificity. Additionally, diagnostic accuracy was estimated for malignancy rates varying between 6% and 48%: PPV would range from 42% to 91%, NPV from 92% to 99%. Within reasonable limits the ThyroSeq® v2.1 is highly reliable to rule out malignancy (21).

Le Mercier *et al*. retrospectively tested a different commercially available 50-gene NGS panel, the Ampliseq™ Cancer Hotspot Panel v2 (ThermoFisher, San Diego, California, USA), which is a tumor-nonspecific NGS panel for detection of somatic tumor variants. This panel does not include thyroid-specific RET/PTC, PAX8/PPARγ and NTRK1 rearrangements. Albeit the study only assessed 34 FNAC samples, with a 71% sensitivity and 89% specificity in indeterminate thyroid nodules the Ampliseq™ panel seems less accurate than the ThyroSeq® (71).

The high diagnostic accuracy is also a downside to NGS. Highly sensitive, NGS is able to identify mutant alleles at very low levels (<10%). A low percentage of mutant alleles might reflect a subclone within the nodule, which is not histopathologically identified as carcinoma. This detection of germline or clinically insignificant low-level somatic mutations in benign nodules could decrease NGS specificity (22,80). Nikiforov *et al*. suggested that the next improvement of the NGS-related tests should therefore be to determine accurate threshold levels for the various gene variations (80).

NGS encompasses crucial technology that is rapidly advancing. The ThyroSeq® v3 was recently announced, promoting to encompass no less than ~95% of genetic alterations occurring in PTC. Extraordinary diagnostic accuracy above 90% is anticipated, including high accuracy in Hürthle cell lesions. Results of the prospective studies validating this new version will likely be published shortly. Nonetheless, NGS techniques currently have limited global availability, with the exception of some European countries and the USA. The ThyroSeq® is available for $3,200 per test (123). In contrast, the thyroid non-specific AmpliSeq™ panel can be ordered online for only €230 (124). Independent prospective studies are needed to validate its performance and predicted cost-utility in different patient populations, and confirm the superior position of the ThyroSeq® and other NGS techniques.

1.1.8. Afirma® Gene Expression Classifier

In molecular diagnostics, the chief competitor of the 7-gene mutation panel is the commercial Afirma® gene expression classifier (GEC) (Veracyte Inc., South San Francisco, CA, USA). The GEC uses quantification of the mRNA-expression of 167 genes and a proprietary classification algorithm to determine the probabilities of malignancy in the samples' expression patterns. The classification algorithm to discern a 'benign' (negative test) from a 'suspicious' (positive test) thyroid nodule results from a successful designer study that trained the GEC in both a tissue set and diverse FNAC sample sets with known histopathology (125). Alexander *et al.* performed the first prospective, blinded, industry-sponsored clinical study to validate this Afirma® GEC in patients with indeterminate thyroid nodules (126). From 49 hospitals 577 Bethesda III, IV and V FNAC samples were collected, obtained by two additional needle aspirates from thyroid nodules with a diameter of at least 1 cm. After exclusion of over half (312/577, 54%) of the samples for reasons such as nodules that were

not surgically resected, duplicate specimens from the same nodule, and issues with specimen shipments to Veracyte, finally 265 FNAC samples were included in the analysis. Sensitivity of the Afirma® GEC was 90% in the 129 Bethesda III as well as the 81 Bethesda IV nodules with a useful GEC-negative test result in 38% (100/265), but specificity was merely 53% and 49%, respectively (52% on average). Despite the relatively high malignancy rate in Bethesda III nodules and the high number of exclusions, this study is well conducted and recognized worldwide as the landmark study that demonstrated the strength of the Afirma® GEC (126). After the overwhelming results from this key-publication, popularity of the GEC took flight. It is marketed as a highly accurate rule-out test for malignancy in thyroid nodules with indeterminate cytology.

In 2014, the first multicentre study that retrospectively assessed the clinical utility of the Afirma® GEC was published. Only 6% of reported GEC-negative Bethesda III, IV and V nodules eventually underwent surgery, of which one resulted in a 6 mm mPTC. Unfortunately, data on GEC negative nodules were only reported on an aggregate level; exact test performance rates in Bethesda III and IV nodules cannot accurately be extracted from the publication. Less than half of the GEC-negative nodules without surgery (71/163, 44%) had clinical or radiological follow-up, ranging from 1 to 24 months (median 8 months) – a limited duration compared to the natural, indolent course of differentiated thyroid carcinoma. The published paper does not describe whether the remaining 92 patients with GEC-negative nodules received any follow-up at all. Despite evident limitations to the applied reference standards, Alexander *et al*. concluded that their results confirm both the accurate test performance from their prior study as well as the large impact that the Afirma® GEC has on clinical decision-making for cytologically indeterminate thyroid nodules (127).

Yet, physicians indeed seemed reassured by a negative GEC result based on the first studies alone (126,128). In many institutions in the USA the Afirma® GEC was immediately implemented in clinical practice. The retrospective studies that followed were mere post-implementation utility studies, and generally reported very high but moderately consistent sensitivities. GEC-negative nodules were largely managed without surgery and considered true-negative, resulting in possible overestimation of test sensitivity. Long-term follow-up is not yet available to endorse a benign diagnosis in these cases (129,130). The high degree of missing histology was recognized by most of these studies as a major limitation (127,129,131-134). This was confirmed by the 2015 ATA guidelines: recognizing the Afirma® GEC as a promising diagnostic tool, the guidelines stress that it is a major shortcoming that external clinical validation studies with full histological follow-up of Afirma® GEC-negative nodules are still lacking (23).

Not all studies were able to confirm the potential of the Afirma® GEC. Some struggled with a low benign call rate (i.e. useful negative test result that could lead to management change) (130,135). McIver *et al*. questioned the cost-effectiveness of the Afirma® GEC in their population, as the mere 22% (16/72) negative test rate was much lower than anticipated. Moreover, a quarter of these GEC-negative patients rejected the proposed conservative treatment of ultrasound-based follow-up and underwent surgery anyway; one of them was diagnosed with a 3.2 cm FTC with focal capsular and vascular invasion. Also, 84% of GEC-positive nodules proved histopathologically benign, overall resulting in a disappointing 83% sensitivity and 10% specificity (130).

Besides concerns regarding adequate clinical validation of test performance, the post-implementation influence of the GEC on surgical decision-making for individual patients was also questioned. In line with the results of their preliminary study, Noureldine *et al*. demonstrated that Afirma® GEC testing had not aided surgical decision-making (135,136). In

93% (206/222) of the included indeterminate nodules, a ‘benign’ or ‘suspicious’ GEC result did not affect management at all: the surgical strategy would have been identical had it been based merely on clinical, cytological or radiological suspicion. However, if management changes were based on the GEC result, they were more often wrong than right: 11 times GEC-positive results inappropriately tempted physicians into more aggressive surgery, and total thyroidectomy was performed instead of the initially recommended lobectomy for nodules that proved histopathologically benign. In contrast, in just four GEC-positive cases the more aggressive surgery was appropriate and the nodule was histopathologically malignant. Also, in just one patient surgery was withheld specifically due to a negative Afirma® GEC result. In the other unresected GEC-negative nodules surgery was not clinically indicated to begin with; the negative GEC-result merely endorsed conservative management (135). As the GEC was still a new technology when this study was conducted, it is possible that the involved physicians were unsure of the correct interpretation of the GEC results or hesitant to rely on a negative GEC result. However, clinical suspicions and physician and patient preference will always be considered when making surgical decisions. Yang *et al*. elegantly tried to solve the shortcoming (histological) follow-up by comparing their findings of GEC performance to a pre-GEC cohort of similar patients from their hospital in all of whom surgery was performed (11,131). The reported malignancy rates were comparable pre- and post-GEC implementation (18% versus 17%), and obviously relatively more surgeries were performed for benign nodules in the pre-GEC period. Assuming the true malignancy rates in the successively studied populations are indeed similar, the GEC only modestly reduced the number of futile surgeries for benign thyroid nodules from 66% to 52% (131). Altogether, the contribution of the Afirma® GEC to the surgical decision-making may be more limited than expected based on its diagnostic accuracy.

Availability, cost-effectiveness and limitations of the GEC

The Afirma® GEC is currently only available for routine use in the USA. There are high demands for the FNAC specimens regarding sample preservation and shipping. Cytology is revised by Veracyte cytologists and declined if not strictly Bethesda III or IV, with 14% to 17% discordancy between local assessment and central review, comparable to known interobserver rates for thyroid cytology (5,126,135). Reported rates of nondiagnostic GEC test results due to insufficient quantity or quality of the mRNA are substantial, varying from 1% to 17% (130,132). Insufficient mRNA quality was often caused by problems with long duration of the sample shipment to Veracyte (126,130). Fourth, Afirma® GEC testing is expensive and is currently marketed for $3,500 (range $1,750 to $7,000) per test (121,131,137). Testing for medullary carcinoma and *BRAF* mutation is not included in the Afirma® GEC, but can be performed by Veracyte at additional costs (131). Yet, ancillary *BRAF* mutation testing may not be relevant, as Kloos *et al*. found that it improved sensitivity nor specificity of the GEC (57).

Studies of cost-effectiveness yielded variable results, but most concluded that GEC testing would not be cost-effective over conventional surgical management or other diagnostic modalities in various clinical settings (14,121,137-140). The first of these studies proclaimed cost-effectiveness of the GEC even prior to publication of the first validation study by Alexander *et al.*, and has been criticized for several important methodological caveats. This study professedly overestimated test specificity at 75%, overestimated the rate of permanent complications from thyroid surgery, and did not consider the regularly reported GEC test failures (15-17,130,137,138,141). A recent study determined population-dependent thresholds for feasible cost-effectiveness by comparing GEC performance to conventional surgical management in a local Bethesda III/IV population. GEC-guided management was not cost-

effective, adding $1,197 to the $11,119 expenses for conventional treatment while hardly improving QALYs. Sensitivity analysis showed that the GEC would only become cost-effective if its specificity exceeds 71%, if it costs less than $2,640, or if the population malignancy rate decreases from the actual 24% to below 9.2%. This price threshold for cost-effectiveness decreases as the malignancy rate increases, as low as $2,023 per test at 35% cancer prevalence (137).

Furthermore, existing inter-institutional differences in test performance have consequences for local applicability and effectiveness (127,134). Marti *et al*. compared GEC performance in distinct populations of two large hospitals. The reproducibility of the tests' sensitivity and specificity was good, but utility strongly depended on the local prevalence of malignancy: as the population malignancy rate increased, a rarer negative GEC became less reliable to rule out malignancy. Oppositely, at low malignancy rates a negative GEC merely confirmed that the probability of cancer was low. In neither situation, the GEC changed the management strategy. GEC testing was most useful if the malignancy rate ranged between 15% and 21%, comparable to the prevalence reported by Alexander *et al*. (126,134).

Finally, the degree of missing histology is a major limitation to the performed studies. None of the studies following the key publication by Alexander *et al*. had complete histopathologic follow-up; histopathological confirmation ranged between 35% and 82% of specimens (126). Missing histology mainly comprised GEC negative nodules, likely resulting in overestimated sensitivity (i.e. missing some malignancies in the many unoperated GEC-negative nodules) and underestimated specificity (i.e. relatively more GEC-positive nodules with benign histology (false-positives) were operated on than GEC-negative nodules with benign histology (true-negatives)). The trend that studies with higher surgical rates for GEC-negative nodules showed more moderate results supports these hypotheses (126,130,142).

A recent meta-analysis by Santhanam *et al*. included seven studies and reported 96% pooled sensitivity and 31% pooled specificity for the GEC in Bethesda III, IV and V thyroid nodules with histopathological follow-up (143). The authors expected that more than 90% of patients with a negative test would be treated conservatively (143). However, in individual studies up to 25% of patients pursued surgery or conservative treatment despite GEC-based recommendation to do the opposite (127,130). This observation is crucial to cost-utility analyses. In addition, expensive rule-out tests such as the Afirma® GEC should not be performed in case surgery is considered for other reasons, such as cosmetic or mechanical complaints.

GEC in Hürthle cell cytology

Brauner *et al.* specifically validated the Afirma® GEC in 72 cytology samples suspicious for Hürthle cell neoplasm. They demonstrated that GEC testing could accurately have reduced the number of futile surgeries, although through a less profound reduction than in non-oncocytic indeterminate thyroid nodules (132). Similar results were noticed in other studies: despite a relatively low risk of malignancy, the majority of Hürthle cell nodules were GEC-positive. Regardless of good sensitivity, this unfavourable benign call rate in Hürthle cell cytology limits diagnostic efficacy in these nodules, increasing the number needed to test and negatively affecting possible cost-effectiveness (126,131,133,135,142,144). Diagnostic accuracy of the GEC would likely improve if Bethesda IV cytology suspicious for a Hürthle cell lesion was excluded from GEC testing. Otherwise, similar to the additional testing for medullary carcinoma, adaptations should be made to the Afirma® GEC to improve its clinical utility for Hürthle cell lesions.

In conclusion, it is generally assumed that the Afirma® GEC accurately reclassifies approximately two out of five indeterminate thyroid nodules as benign with published sensitivities ranging between 83% and 100% and similar test performance in Bethesda III and IV nodules. Withholding diagnostic surgery from these patients seems safe (130,131,134,144). However, the diagnostic strength and potential cost-utility of Afirma® GEC strongly rely on its NPV – thus on the prevalence of malignancy and benign call rate in the targeted population. There are important concerns regarding the currently insufficient number of clinical validation studies with adequate rates of histopathological confirmation or long-term clinical follow-up. Physicians are strongly advised to locally validate Afirma® GEC test performance before considering test implementation in daily practice. Nonetheless, further large validation studies on the Afirma® GEC may soon become obsolete, as an updated version of the test, the Gene Sequencing Classifier (Veracyte Inc., South San Francisco, CA, USA), is currently being put into operation. Improved diagnostic accuracy is anticipated, with specific attention to the differentiation of Hürthle cell nodules.

1.2. MicroRNA

First described in thyroid cytology in 2006, evaluation of the expression levels of microRNA (also called miRNA) is among the newer and more promising approaches to differentiate between benign and malignant thyroid neoplasms (145,146). MicroRNAs are small endogenous noncoding ribonucleic acids (RNAs) of approximately 22 nucleotides in length. As negative regulators (i.e. silencers) of protein synthesis at a post-transcriptional level, they are involved in many intracellular processes, including cell growth, differentiation and proliferation. Dysregulation of microRNA expression is found in almost all types of human

cancers (147). It reflects the deregulated expression of oncogenes and tumor suppressor genes (146,148-150). MicroRNA overexpression is present before morphological tissue changes are seen and therefore considered to be a part of premalignant changes in carcinogenesis (145). MicroRNA expression profiles are tissue-specific and can not only identify the tissue of origin, but also the histopathological subtype of the cancer and whether it concerns the primary tumor or a metastasis (148,151).

MicroRNA expression profiles are similar among the various types of thyroid carcinoma, even though expression levels are often distinctively different (148). In histopathological studies, PTC was associated with an up to 11- to 19-fold upregulation of miR-146b, miR-221, miR-222, miR-181b, miR-187, and a downregulation of miR-1 and miR-138 compared to healthy thyroid tissue and benign nodules. Upregulation of miR-221, miR-222 and miR-187 was also found in FTC, FTC-OV, poorly differentiated and anaplastic carcinoma (145,146,148,152,153). Overexpression of miR-146b-3p, miR-146b-5p and miR-375 was seen in both PTC and FVPTC (152,154). Furthermore, expression levels of miR-221 and miR-222 were reported about twice as high in FVPTC as compared to PTC or FTC (152). Only a few microRNAs were differently expressed between follicular neoplasm and FTC (155). Follicular adenoma was associated with the expression of miR-200a, whereas high expression of miR-31 was found in Hürthle cell adenoma (148). FTC is related to the differential expression of miR-146b, miR-7-5p, miR-346, miR-197 and miR-21, but results among studies are more heterogeneous (148,155,156). FTC-OV showed an expression pattern slightly similar to FTC, but also distinct overexpression of other microRNAs, such as miR-339, miR-183, miR-197 and miR-885-5p (148,153).

Accordingly, a diagnostic panel of a carefully selected combination of microRNAs and appropriate expression levels could aid in the preoperative distinction of indeterminate thyroid cytology (157). Recent meta-analyses struggled to reconcile the studies on microRNA

in FNAC, as the investigated set of microRNAs was never identical and individual microRNA performance was infrequently described. In unselected cytology, estimated sensitivity of microRNA expression analysis ranged from 75% to 78% regardless of the investigated set; estimated specificity from 73% to 81% (156-158).

In indeterminate thyroid cytology, different sets of microRNAs were evaluated; only several individual microRNAs were analyzed in more than one study. The selected microRNAs were first assessed in a test set of cytological and/or histopathological specimens and a cut-off for their expression level was determined. Subsequently, the significantly up- or downregulated microRNAs were validated in an independent set of (indeterminate) thyroid FNAC samples. Some studies developed a decision model for the validation step (149,154,159).

The most promising results were presented by Keutgen *et al.* (159). Of the six microRNAs investigated in their test set, miR-21, miR-146b, miR-181a and miR-222 were differentially expressed in malignant nodules with prior indeterminate cytology. The subsequently developed support vector machine model incorporated miR-21, miR-222 and the insignificantly expressed miR-197 and miR-328. Prospective validation in an independent set of 72 indeterminate FNAC samples resulted in 100% sensitivity and 86% specificity. Five of the seven false positives had Hürthle cell cytology; excluding these, raised specificity to 95% (159). Notably, even though overexpression of miR-146b is often related to thyroid carcinoma, it proved not useful to Keutgen *et al.* to include in their prediction model (159). In contrast, Agretti *et al.* and Shen *et al.* included miR-146b as the key differentiators in their models. Agretti *et al.* assessed a frequently quoted set of microRNAs consisting of miR-146b, miR-155, miR-187, miR-197, miR-221, miR-222 and miR-224 (148,149). Published in 2008, Nikiforova *et al.* had demonstrated that this 7-microRNA set in FNAC samples had 100% sensitivity and 94% specificity if one of the included microRNAs showed an at least two-fold overexpression (148). Analytic validation of this model by Agretti *et al.* showed differential

upregulation in PTC of all of these microRNAs except miR-197. In particular, miR-146b showed a >30-fold higher expression in PTC. A decision tree including miR-146b, miR-155 and miR-221 was 98% accurate in the test set, but validation in an independent set of indeterminate FNAC samples was unsuccessful, yielding mere 60% sensitivity and 58% specificity (149).

Vriens *et al.* used a microRNA array to detect 10 genes that were up- or downregulated by ≥5-fold in thyroid malignancies. Four microRNAs (miR-100, miR-125b, miR-138 and miR-768-3p) were significantly downregulated and accurately differentiated between benign and malignant follicular and Hürthle cell neoplasms in the test set. In their validation set of 125 indeterminate FNAC samples, only miR-138 was moderately distinctive with 81% NPV. For Hürthle cell carcinoma, miR-138 and miR-768-3p were both 98% accurate (160).

Finally, in a recent Italian study only miR-375 accurately differentiated between benign and malignant neoplasms. Subsequently, in TIR3 cytology excluding Hürthle cell lesions, a 12-fold or higher overexpression of miR-375 perfectly distinguished benign from malignant lesions with 100% accuracy. It was also significantly differently expressed between TIR3A and TIR3B categories and correlated with a different malignancy risk (161).

Availability and limitations of microRNA expression analysis

MicroRNA expression analysis has advantages over other techniques. MicroRNAs are more stable than mRNA at maintaining their expression in formalin-fixed paraffin-embedded (FFPE) tissue samples as well as FNAC specimens, irrespective of the preservation method (e.g. archived FNAC slides or nucleic acid preservation solutions) (148,161). Recently microRNA expression was even successfully measured in serum (162). Moreover, microRNA expression levels measured with generic methods (e.g. quantitative RT-PCR) correspond well

to their biological effect, as microRNAs affect biological processes without the additional step of protein synthesis (148).

However, general limitations of FNAC also translate to concerns with microRNA analysis: scant cellularity or low levels of malignant cells in FNAC specimens could cause a false-negative microRNA test result (149). Another limitation is the plurality of microRNAs associated with DTC in histopathological studies, causing vast heterogeneity between the limited number of studies in indeterminate cytology. Validation studies of the same microRNA set are lacking. Simultaneously, new microRNAs are still correlated to thyroid carcinoma. Ongoing research has yet to compose the optimal set of microRNAs. Recently, the first commercial test was marketed as the ThyraMIR™ (Interpace Diagnostics, Parsippany, NJ, USA). It evaluates the expression levels of miR-29b-1–5p, miR-31–5p, miR-138–1-3p, miR-139–5p, miR-146b-5p, miR-155, miR-204–5p, miR-222–3p, miR-375, and miR-551b-3p. The ThyraMIR™ demonstrated 57% sensitivity and 92% specificity in 109 Bethesda III and IV FNAC specimens (22). Prospective clinical validation of the ThyraMIR™ could affirm the diagnostic value of microRNA expression profiling in indeterminate thyroid nodules in the pre-operative setting.

1.3. Immunocytochemistry

Tissue characterization through selective staining of expressed proteins, i.e. immunohistochemistry (IHC), is a technique that combines histopathology and biochemistry. Exploiting basic antigen-antibody interactions, IHC is able to visualize the distribution and localization of specific cellular components within the cell and in the proper tissue context. This includes tissue biomarkers specific for e.g. infection or malignancy. IHC has been fully incorporated in the histopathological routine and is crucial to morphological and molecular

tissue characterization. When immunocytochemistry (ICC) – the application of this immunology-based technique in cytology – became available, the possibilities were extended to the preoperative setting, too. Specific immunomarkers have been developed to differentiate between benign and malignant thyroid nodules. The 2015 ATA guidelines acknowledge ICC as a technique under development with limited prospective validation studies in indeterminate cytology (23). In unselected thyroid cytology, the much-used immunomarkers galectin-3, Hector Battifora mesothelial-1 (HBME-1) and cytokeratin 19 (CK-19) demonstrated 85%, 83% and 80% sensitivity, and 90%, 79% and 79% specificity, respectively (163).

1.3.1. Galectin-3

Galectin-3 is a β-galactosyl-binding protein from the lectin group. It is involved in cell-cycle regulation, including cell migration and adhesion. Its exact function is still to be unraveled, but a role in the pathogenesis and progression of PTC is presumed (44,163-165). It is related to inhibition of apoptosis, induced by abnormal p53 expression (165). Galectin-3 can be present both in the intracellular as well as the extracellular matrix (166). Normal thyrocytes do not express galectin-3, but the physiological expression of galectin-3 in macrophages, neutrophils, mast cells and Langerhans cells provides an internal positive control of the investigated FNAC samples (167,168). Positive cytoplasmic staining – as opposed to nuclear staining – for galectin-3 is suspicious for malignancy and mainly associated with PTC (117,169,170). Galectin-3 expression has also been associated with the malignant transformation of follicular neoplasms, as it was present in follicular adenoma as well as FTC (166,171,172). Encapsulated FVPTC and minimally invasive FTC showed less frequent and weaker staining (172,173).

In 2001, Bartolazzi *et al*. argued that galectin-3 staining could accurately diagnose thyroid carcinoma in unselected thyroid cytology (117). Subsequent studies in indeterminate thyroid cytology mostly could not reproduce these promising results. With a positive stain in approximately a third of all nodules, sensitivity and specificity of galectin-3 ranged from 0% to 92% and from 68% to 100%, respectively (44,174-177). Merely Saggiorato *et al*. demonstrated that galectin-3 accurately differentiated follicular adenomas from FTC with 92% sensitivity and 94% specificity if a cytoplasmic stain in ≥10% of the cells was considered positive (174). The prospective multicenter clinical validation study by Bartolazzi *et al*. demonstrated 78% sensitivity and 93% specificity in Thy3 nodules if a cytoplasmic galectin-3 stain in >5% of the cells was considered positive. Nineteen of the 22 false-positive nodules were follicular adenoma. However, a group of 33 difficult-to-diagnose (follicular) tumor of unknown malignant potential lesions was disregarded, 22 of which were galectin-3 negative. If these neoplasms were considered malignant, sensitivity dropped to 69% (164).

1.3.2. HBME-1

HBME-1 is a monoclonal antibody targeting an unknown antigen on the microvilli of mesothelial cells. It is usually negative in normal thyroid follicular cells. Abnormal expression of HBME-1 shows cytoplasmic location with membrane accentuation. It is associated with, but does not necessarily indicate PTC (165,169,178,179). Its low detection limit enables assessment in liquid based cytology (180). Reported sensitivity and specificity of HBME-1 in indeterminate nodules ranged from 61% to 100% and from 75% to 96%, respectively (174,176,180,181). Approximately two out of five nodules showed positive staining. If only non-oncocytic follicular neoplasms were selected, Saggiorato *et al*.

demonstrated that HBME-1 had excellent 93% sensitivity and 98% specificity in indeterminate thyroid nodules (174).

1.3.3. Cytokeratin 19

Cytokeratin 19 (CK-19) is a type I keratin. It belongs to the group of intermediate filament proteins, which arrange the cell cytoskeleton and structural integrity. CK-19 is widely present in epithelial cells, but also found in basal cells layers of stratified epithelium (174,182). Strong and diffuse abnormal expression of CK-19 indicates PTC, including FVPTC. Expression in FTC is less intense and more variable, warranting nuanced interpretation of CK-19 staining intensity. CK-19 usually shows no or only focal expression in follicular neoplasms, hyperplastic nodules and adenomatous goiter (174,178,182,183). The reported sensitivities and specificities for CK-19 staining in indeterminate cytology ranged from 76% to 88% and 80% to 100%, respectively (174,180,182). Lacoste-Collin *et al*. demonstrated the importance of an accurate threshold. CK-19 staining in 31 Bethesda IV nodules accurately diagnosed five out of six malignancies, including a PTC, two FTC and two out of three FVPTC. At a threshold of ≥30% stained cells, 5 of 25 benign lesions tested false-positive; at a more sensitive threshold of ≥10% stained cells, 12 of 25 tested false positive (180).

1.3.4. Other immunocytochemistry markers

Immunohistochemistry studies identified more potential ICC markers. Some, like CD44v6, have not yet been investigated in indeterminate cytology (117). Other markers were sporadically investigated in preclinical studies, including Ki-67, TROP-2, emerin, keratan sulphate, thyroperoxidase, CD57 and GLUT-1.

962 Nuclear protein Ki-67 is expressed in nearly all cell cycle phases in proliferating cells. It is
963 associated with poor prognosis in PTC (184). The percentage of cells with Ki-67 expression is
964 considered the tissues' proliferative index. At a cutoff of ≥1% Ki-67 was 85% sensitive and
965 71% specific for thyroid carcinoma in Bethesda IV nodules. A combination of HBME-1, CK-
966 19 and Ki-67 immunomarkers was 91% accurate to diagnose malignancy (180). Ki-67
967 expression is likely only distinctive for follicular type carcinoma; expression in PTC is
968 generally low (180,184,185).
969 Glycoprotein human trophoblast cell surface marker (TROP-2) is overexpressed on the cell
970 surface of different epithelial carcinoma (e.g. breast, colon) and associated with tumor
971 aggressiveness and poor prognosis. In indeterminate thyroid cytology, it was only assessed in
972 one small subseries of Bethesda III samples, correctly diagnosing the three included
973 carcinoma and all but one of the nine benign nodules (186).
974 Emerin staining emphasizes features of the nuclear membrane often seen in PTC, such as
975 irregularities and invaginations. Consequently, the stain could facilitate the morphological
976 diagnosis of PTC and especially the more difficult-to-diagnose FVPTC (176,187). In 53 Thy3
977 nodules assessed by Asioli *et al*., positive emerin staining was highly specific for PTC
978 (including FVPTC), but misdiagnosed all FTCs (176).
979 Another immunomarker associated with PTC is keratan sulphate, an abnormal
980 glycosaminoglycan complex. It was 98% specific in indeterminate cytology, but correctly
981 predicted PTC only; its sensitivity was poor at 48% (174).
982 The expression of thyroid peroxidase (TPO) is related to benign follicular neoplasms. A
983 negative TPO stain was 80% sensitive and 86% specific for thyroid malignancy (174).
984 Finally, CD57 (Leu7) expression is associated with epithelial and nonepithelial malignancies,
985 including thyroid carcinoma. Cytological staining was only investigated in a small series of

indeterminate cytology, but seemed specific for PTC. In the same series, GLUT-1 was not a useful ICC marker – there were no positive stains (188).

Combined use of immunocytochemistry markers

Some research groups have suggested that evident single-marker galectin-3 positivity is sufficient to refer a patient for total thyroidectomy (164,170,177). The ATA guidelines did not adopt these suggestions, and many other researchers advocate that a panel of ICC markers should be applied to strengthen the suspicion of malignancy (23,165,174,179). Several panels were investigated in literature. Zhang *et al*. assessed a triple stain of galectin-3, HBME-1 and p27. P27 is a cyclin-dependent kinase inhibitor related to cell life span in normal thyroid cells. Downregulated in malignancy, positive P27 stain is related to benign histopathology. In a set of Bethesda III cytology samples, positive p27 staining with negative galectin-3 and HBME-1 staining was 100% predictive of a benign nodule and occurred in 38% of samples. Loss of p27 staining in combination with positive galectin-3 and/or HBME-1 staining was 100% sensitive and 86% specific (165). Another study investigated galectin-3 and HBME-1 in combination with a RET proto-oncogene stain, which reflects abnormal intracellular RET proto-oncogene activity and presence of the RET/PTC rearrangement. Unfortunately, RET staining was inaccurate in indeterminate thyroid nodules (181).

To find the most accurate combination of immunostains, Saggiorato *et al*. explored the expression of galectin-3, HBME-1, thyroperoxidase, CK-19 and keratan-sulphate in 125 cytological follicular neoplasms, 24 of which were Hürthle cell lesions. Galectin-3 was not only the most accurate marker individually, but also in combination with other stains. Sequential HBME-1 staining of galectin-3-negative cases reached 98% sensitivity and 98% specificity in non-oncocytic lesions. In oncocytic lesions, sequential CK-19 staining was more preferred with 100% sensitivity and 100% specificity (174).

The common denominator between all these studies is the combined use of galectin-3 and HBME-1. Unfortunately, clinical validation studies regarding this combination are limited. Its seemingly promising diagnostic accuracy warrants further assessment in future prospective studies.

Performance of immunocytochemistry in Hürthle cell cytology

Expression of ICC markers in Hürthle cell nodules differs from non-oncocytic indeterminate cytology. Hürthle cell carcinomas were distinguished in the cytological samples by typical overexpression of markers associated with a high degree of cell proliferation, disorganized tissue structure and intermediate differentiation, such as Ki-67, laminin, cyclin D1 and cyclin D3. Overexpression reflects the known more erratic behavior of Hürthle cell carcinoma (185,189). Moreover, markers that were highly diagnostic in indeterminate nodules in general, also seem differently expressed in Hürthle cell lesions. Saggiorato *et al*. demonstrated that two combinations of ICC markers were extraordinarily accurate: galectin-3 and CK-19 staining was 100% sensitive and 100% specific; galectin-3 and thyroperoxidase staining was 100% sensitive and 85% specific (174).

In a previous meta-analysis, inclusion of Hürthle cell lesions was related to between-study heterogeneity (163). Hürthle cell lesions require a biotin-free ICC method, as Hürthle cells themselves are rich in biotin. Thus, much-used biotin-based methods may consequently cause false positive and highly intensive staining in Hürthle cell neoplasms (166,172,181).

Availability, cost-effectiveness and limitations of immunocytochemistry

Current application of immunocytochemistry is limited. Clinical validation studies for all of the described immunomarkers are scarce, and no cost-effectiveness studies are available to date. Yet, the technique is widely available, relatively inexpensive and fast in comparison to

other (molecular) techniques. Costs per immunostain vary up to €20, partly depending on simultaneous local application of the technique and similar stains for immunohistochemistry. Immunocytochemistry is preferably performed on cell block FNAC specimens, but can be performed in all types of cytology, from direct smears to liquid-based cytology (179,181). ICC is impossible when the FNAC specimen has poor cellularity or too much obscuring blood (190). Also, immunostaining of cytology is technically more difficult than histological staining, especially in (destained) cytology smears. Technical inconsistency and interobserver variation likely lead to false-negative results (164,182). Stain intensity thresholds or percentage of stained cells necessary to raise suspicion of malignancy vary in the available literature. Consistent methodology and assessment thresholds should be determined to improve reproducibility of ICC results.

Clinical validation studies of existing ICC markers are ongoing. Meanwhile, new markers are also playing the field, searching for the interfaces between mutation analysis of highly specific oncogenic driver mutations and accessible ICC techniques. For example, Leslie *et al*. investigated ICC of the $BRAF^{V600E}$ mutation using the mutation specific antibody VE1 in a small series of thyroid FNAC samples. Concordance between ICC and conventional $BRAF^{V600E}$ mutation analysis was 85%. All samples that were $BRAF^{V600E}$ positive by either method were confirmed as $BRAF^{V600E}$ positive PTC on histopathology. Of the eight included indeterminate thyroid nodules, seven were histopathologically malignant and $BRAF^{V600E}$ mutation was detected in two nodules: one by both methods, one only by molecular analysis. The $BRAF^{V600E}$ specific antibody (VE1) stain was much weaker in cytology than in histology. Moreover, costs of the VE1 antibody are currently high and optimization of methodology is warranted. Yet, Leslie *et al*. demonstrated that $BRAF^{V600E}$ mutation analysis using ICC is a promising alternative to mutation analysis (79). If future studies could validate these results in larger cohorts of indeterminate thyroid nodules and detect reliable immunomarkers for other

oncogenic driver mutations, this technique unites the strengths of gene mutation analysis and immunocytochemistry in one technique, though likely at lower costs.

In general, ICC is a widely available and relatively inexpensive technique with a reasonable diagnostic accuracy. Many immunomarkers seem to have a pronounced association with PTC. Galectin-3 and HBME-1 were most frequently investigated, but their specificities and sensitivities seem to fall short of justifying ICC-based surgical decision making. Diagnostic accuracy of their combined use seems promising, yet current evidence is limited. Prospective validation trials are warranted to confirm the diagnostic potential of ICC, including validation of thresholds for stain positivity, panels of multiple immunostains and other methodology.

2. CONVENTIONAL IMAGING

2.1. Ultrasound

Ultrasound (US) is one of the principal steps in the initial work-up of thyroid nodules. It is cheap, fast, non-invasive and globally available, but accurate assessment strongly depends on operator experience (191). Multiple meta-analyses showed that well-known US features such as nodule hypoechogenicity, microcalcifications, irregular margins (including microlobulated or ill-defined margins), and a taller-than-wide shape raise the suspicion for thyroid malignancy and are mostly associated with PTC (191,192). Nonetheless, no single US feature is sufficiently sensitive nor specific to accurately identify a malignant nodule in an unselected population (191). Certain combinations of US features, however, may offer accurate closure. The current ATA guidelines now include a flowchart recommending FNAC dependent on nodule size and various combinations of US characteristics with an incremental risk of

malignancy (23). Despite the obvious importance of both ultrasound and cytology, the ATA guidelines do not provide recommendations regarding (re-)interpretation of US characteristics after FNAC has resulted in indeterminate cytology. Follicular-type malignancies typically have a different US appearance. More often FTC may be iso- to hyperechoic, with a spherical shape, smooth regular margins and no calcifications (193,194). FVPTC may also show FTC-like or benign features rather than the classic suspicious features, although microcalcifications may be distinctive (194-196). In the past years, Brito *et al*. and Remonti *et al*. performed meta-analyses on US assessment of unselected thyroid nodules. Both also briefly discussed its diagnostic value in indeterminate nodules, including a mere limited number of studies and also including cytology suspicious for malignancy. Increased central vascularization was most predictive of malignancy with reported 96% specificity (192). Yet, in general US seemed less accurate in indeterminate nodules than in unselected thyroid nodules (191,192).

In the dozens of available original ultrasound studies, individual US features generally demonstrated limited sensitivity in indeterminate thyroid nodules. Only the appearance of a solid thyroid nodule – as opposed to varying degrees of cystic content – had high sensitivity. Ranging between 46% and 100%, multiple studies demonstrated sensitivity above 90% (18,19,197-201).

A number of classic suspicious US characteristics, such as a taller-than-wide shape, presence of irregular margins and presence of microcalcifications, demonstrated valid specificity in indeterminate thyroid nodules. Specificities for each of these characteristics ranged from 72% to 99% (201-204), 65% to 100% (202,205,206) and 36% to 100%, respectively (207,208). Despite the wide range, presence of microcalcifications was more than 90% specific in many studies (197,198,200,202-204,206,207,209-211). Large nodule size (defined as a diameter

larger than 4 cm) was only investigated in a limited number of studies. Reported specificities ranged between 69% and 94% (207,212).

Other features, such as a solitary nodule, hypoechogeneity and absence of a hypoechoic halo were associated with thyroid malignancy, but less accurately differentiated between benign and malignant indeterminate thyroid nodules (198,201,206,212-217). Additionally, opposing the results from one of the mentioned meta-analyses, central vascularization also does not seem very accurate in indeterminate thyroid nodules. Specificity ranged from 0% to 100%, although multiple studies demonstrated extremely poor specificity (18,202,216-220).

Results regarding two US features are remarkably contradicting. First, the absence of a hypoechoic halo is typically considered suspicious for malignancy, but showed overall poor and very heterogeneous diagnostic potential in indeterminate thyroid nodules (191). Sensitivity and specificity ranged from 17% to 99% and 0% to 93%, respectively (200,201,205,221,222). Presence of a hypoechoic halo is typically considered a benign feature, but has also been associated with follicular types of thyroid carcinoma (223). Dogan *et al*. reported 88% specificity for presence of a halo in AUS/FLUS nodules and 78% in FN/SFN nodules (224). Second, the ultrasonographic nodule shape seems ambiguous. Similar to the unselected population, a typically suspicious taller-than-wide shape was generally specific for carcinoma, with reported specificities up to 99% (201,202,204). A spherical shape is generally considered benign, but has also been associated with FTC (191,193,225). In two studies in cytological follicular neoplasms, a spherical shape had an increased risk of malignancy, with 86% to 97% sensitivity and 19% to 26% specificity (226,227). Chin *et al*. even suggested that follicular neoplasms with a taller-than-wide shape could be treated conservatively (227). The uniquely balanced rates of PTC, FVPTC and FTC resulting from indeterminate cytology may explain why these and various other US characteristics have

different diagnostic accuracy than in the unselected population. Dependent on the local case mix, accurate differentiation of indeterminate nodules using the classical suspicious US features may or may not be feasible.

Combination of ultrasound characteristics

A combination of US characteristics likely provides more accurate differentiation than individual features. Different combinations were investigated in multiple studies (81,197,198,200,204-206,216,228-233). Yoo *et al*. reported 100% specificity for the combination of marked hypoechogenicity and taller-than-wide shape, a pattern that occurred in 9.6% (24/249) of the included Bethesda III nodules (201). In the elastosonography study by Rago *et al.*, absence of a hypoechoic halo in combination with presence of microcalcifications was 95% specific for thyroid malignancy, but only 6.4% sensitive (222). Maia *et al*. found 62% sensitivity and 89% specificity in Bethesda III and IV nodules if hypoechogenicity, microcalcifications, an irregular margin and increased intranodular vascularity were considered suspicious (230). Gulcelik *et al*. demonstrated that the US pattern of a solid, hypoechoic nodule with microcalcifications had 95% sensitivity and 99% specificity. The pattern was seen in 21% of cytological follicular neoplasms (234).

In multiple studies it was argued that cytological follicular neoplasms with a typically benign ultrasound pattern – a regular shape, isoechoic, homogeneous, with well-defined margins, cystic components or peripheral vascularity only, and not a single malignant feature - could be safely followed up clinically instead of undergoing diagnostic surgery (200,225,228). Consideration of more features generally increased the sensitivity of the US assessment at the cost of its specificity (198,199). The terms of their interpretation were crucial: Norlén *et al*. demonstrated that US was 95% sensitive and 48% specific if a Bethesda III nodule had either hypoechoic appearance, irregular margins or microcalcifications. If solely the simultaneous

presence of all three features was considered suspicious for malignancy, sensitivity dropped to 37% but specificity increased to 96% (231).

Altogether, diagnostic ultrasound scores or step-by-step algorithms could aid the classification of US patterns and consequent risk of malignancy (200,228,235). Best-known and most validated is the Thyroid Imaging Reporting and Data System (TIRADS), a classification to risk-stratify thyroid nodules, designed by Horvath *et al*. and modified by Kwak *et al*. following the example of the similar BIRADS classification for breast lesions (236,237). The TIRADS assigns nodules to a risk category based on five suspicious US features: solid appearance, (marked) hypoechogenicity, irregular margins, microcalcifications and a taller-than-wide shape. Nodules without any of these features are likely benign and categorized as TIRADS 3. Their risk of malignancy is ~1.7% in a cytologically unselected population. TIRADS 4 includes suspicious nodules, which are further classified according to an increasing malignancy risk into 4a (one suspicious US feature), 4b (two suspicious features) and 4c (three or four suspicious features). Nodules with all five suspicious US features are classified as TIRADS 5 and associated with a high 88% risk of cancer in an unselected population (236). Studies that validated the TIRADS specifically in indeterminate thyroid nodules, showed that diagnostic accuracy depended on the chosen cutoff score and type of cytology (202,235,238-241). Although TIRADS 5 scores were infrequently assigned in indeterminate nodules, a higher TIRADS score (4b/4c/5) was an accurate predictor of malignancy, especially in Bethesda IV cytology (202,238,240). In Bethesda III nodules, lower TIRADS scores (3/4a) could also rule out malignancy (238,241). Prospective validation studies applying the TIRADS in indeterminate cytology are warranted to assess its possible clinical utility in indeterminate nodules.

US performance in Hürthle cell nodules

Cytological Hürthle cell nodules expressed a large variation of US characteristics (203,208,212,223). Many malignant and most benign Hürthle cell nodules had a benign US appearance (208,223). Only three US features possibly predictive of malignancy were reported in individual studies: both hypoechogenicity and hyperechogenicity (as opposed to isoechogenicity) (208), large nodule size (223), and microcalcifications (203). Despite limited evidence, US evaluation does not seem reliable to differentiate Hürthle cell lesions.

Availability and limitations of ultrasonography

The major advantages of ultrasound over other additional diagnostics are its already permanent position in the workup of thyroid nodules, global availability and low costs. No additional resources nor hospital visits are needed to include US interpretations in preoperative management decisions and the investigation is noninvasive. Nonetheless, besides known limitations concerning interobserver variability and less reliable interpretation of small nodules, US feasibility in indeterminate thyroid nodules is limited by the presumed differences in US appearance of papillary and follicular thyroid malignancies, illustrated by the conflicting results for nodule shape and hypoechoic halo in indeterminate nodules. Consequently, local diagnostic accuracy likely follows variations in the local histopathological case mix.

In addition, many of the available ultrasound studies are retrospective, limiting the power of the evidence. As the decision to perform FNAC is customarily based on the results of the prior US, the prevalence of suspicious US features in indeterminate cytology in these studies is presumably overestimated.

Nonetheless, several individual US characteristics seem to have reasonable specificity in indeterminate nodules, although insufficient for accurate diagnosis. A combination of US

features is likely more accurate, although current evidence does not support US-based surgical decision-making. We propose that a future meta-analysis should use the individual patient data from the large number of available original studies to develop an ultrasound algorithm specifically for indeterminate thyroid nodules. The existing TIRADS needs prospective validation.

Even though more advanced and less operator-dependent techniques might be preferred, US features should always be assessed in current clinical practice. The presence of one or more suspicious US features in a Bethesda III or IV nodule increases the suspicion of malignancy and underpins the need for a definite diagnosis. Moreover, to centers or regions with limited access to other (molecular) diagnostics, ultrasound may definitely have clinical utility, pending local validation in the indeterminate population.

2.2. Elastosonography

Firm consistency of a thyroid nodule upon palpation is considered suspicious for malignancy – an established principle during physical examination (242). Ultrasound elastosonography (USE) is a dynamic ultrasound technique that is sometimes referred to as 'electronic palpation'. Tissue elasticity is evaluated by measuring tissue distortion while applying a standardized dosed external force by the US transducer. It was first applied to the thyroid gland by Lyshchik *et al*. in 2005 (243). Classic real-time qualitative USE is performed by free-hand compression and a sine-wave or numerical scale showing how much pressure the operator applies with the probe. A color-coded elastosonography image is superimposed on the grey-scale US images: red and orange visualizes high tissue elasticity (soft tissue), green represents intermediate elasticity and blue low elasticity (firm tissue). Several score systems are available. The original score was developed by Itoh *et al*. in 2006 for the evaluation of

breast tumors and considers scores 1-3 benign on a scale of 1 (highest elasticity) to 5 (no elasticity) (244). Rago *et al*. first applied it to thyroid tumors and modified it to a 3-point score (222,245). Asteria *et al*. derived a modified 4-point score (246).

The earliest studies in thyroid nodules reported opportune results of USE as an additional modality to B-mode ultrasound, but were heterogeneous in USE technique and study population (247). A recent meta-analysis by Nell *et al*. included twenty studies on qualitative USE prior to FNAC and concluded that qualitative USE is fit to diagnose benign nodules and safely dismiss FNAC, provided that the usual elasticity score cutoff is abandoned and only completely soft nodules (score 1 of all systems) are classified as benign. Pooled 99% sensitivity and 99% negative predictive value demonstrated the ability of USE to reliably rule-out malignancy in entirely soft thyroid nodules, composing 14% of their pooled study population (248).

In individual studies on USE in indeterminate nodules, sensitivity and specificity of qualitative USE ranged from 47% to 97% and from 6% to 100%, respectively (19,213,221,222,249). Results of several qualitative USE studies stand out. Lippolis *et al*. showed an aberrant 6.1% specificity, because they reported only eight nodules with high elasticity – 62 of 66 benign nodules were not elastic. The authors themselves suggest that a rather homogenous study population with predominantly small nodules with a solid US pattern, absence of cystic areas, and follicular histology with minimal colloid could be explanatory for the poor specificity rather than operator-dependent causes (213). Such possible relations remain undescribed in other studies. A meta-analysis on the value of USE in indeterminate thyroid nodules demonstrated meager pooled 69% sensitivity and 75% specificity (250).

As manually applied pressure is difficult to standardize, qualitative USE is strongly operator dependent (251). Different USE techniques have been developed to improve objectivity, such as semi-quantitative tissue-to-nodule strain ratio indices (also based on manual compression). Studies investigating semi-quantitative USE in indeterminate thyroid nodules reported sensitivity and specificity ranging from 82% to 100% and from 88% to 100%, respectively (215,217,219,249). Furthermore, quantitative shear wave USE measures the propagation velocity of focused acoustic pulses – shear waves – from the probe, which correlate to tissue stiffness (Young's modulus) (18,252). It had 82% sensitivity and 88% specificity in a recent prospective pilot study by Samir *et al*.(18). Performance of (semi-)quantitative USE seems better than qualitative USE, but results are subject to overfitting from the ROC analysis performed to determine the strain ratio cutoff value with the highest sensitivity and specificity. None of the studies applied a predefined cutoff or validated their own cutoff externally. Consequently, the resulting thresholds were hardly comparable (215,217,219,249).

Altogether, the results from currently available studies cannot support surgical decision-making in thyroid nodules with indeterminate cytology using elastosonography in any of its forms. Whereas color-coded qualitative USE has insufficient sensitivity and specificity, the semi-quantitative method lacks validation. The power of the available evidence is additionally limited by both methodological heterogeneity and the use of different USE techniques, image processing and elasticity scoring methods across studies. Nevertheless, the suggested promising rule-out capacity of qualitative USE when applying an alternative cutoff score of 1 in unselected nodules, deserves clinical validation in indeterminate thyroid nodules. Major advantages of the technique are the minor extra costs of USE, as it can be performed during regular thyroid US with the same equipment, and only adds approximately 5 minutes to the

procedure time per patient. Cost-effectiveness will largely depend on performance of USE, but no cost-effectiveness studies in indeterminate thyroid nodules are available to date.

2.3. Computed Tomography

There are no studies that investigated computed tomography (CT) scanning in thyroid nodules with indeterminate cytology. Prior studies indicated that CT cannot accurately differentiate thyroid carcinoma (253,254).

3. FUNCTIONAL AND MOLECULAR IMAGING

3.1. ^{99m}Tc-MIBI

Hexakis(2-methoxy-2-methylpropylisonitrile)technetium[^{99m}Tc] (^{99m}Tc-MIBI) is a Technetium-99m-labeled radiopharmaceutical, primarily known for its use in myocardial perfusion imaging since the 1980s and more recently the evaluation of hyperparathyroidism. Uptake of ^{99m}Tc-MIBI, a lipophilic cation, reflects both perfusion and the number of active mitochondria in the cells of the thyroid nodule and thus its oxidative burden (70,255). ^{99m}Tc-MIBI scintigraphy is more suitable for the differentiation between benign and malignant thyroid nodules than scintigraphy with ^{99m}Tc-pertechnetate ($^{99m}TcO_4^-$) or radioisotopes of iodide (often $^{131}I^-$,$^{123}I^-$ or $^{124}I^-$). These latter tracers interrogate the sodium-iodide symporter of the thyrocyte and are frequently used to assess thyroid nodule functioning to distinguish autonomous (“hot”) from hypofunctioning (“cold”) nodules. They are neither specific nor effective to detect malignancy: benign nodules can be anything from hyper- to hypofunctioning, and far outnumber the carcinomas. Still, thyroid malignancies are almost

always hypofunctioning: decrease of the sodium-iodide symporter or thyroid peroxidase are hallmarks of cell dedifferentiation and lead to loss of iodide-trapping function and thus ^{99m}Tc-pertechnetate or radioiodine uptake (23,255-257). ^{99m}Tc-MIBI uptake is independent of iodide trapping and organification in the thyrocytes. Nodules with increased uptake and late retention of ^{99m}Tc-MIBI are suspicious for malignancy (70,255). A 2013 meta-analysis by Treglia *et al*. demonstrated 82% sensitivity and 63% specificity for ^{99m}Tc-MIBI scintigraphy in clinically suspicious, hypofunctioning, cytologically unselected thyroid nodules. Hyperfunctioning benign adenomas can show false-positive increased uptake of ^{99m}Tc-MIBI due to their increased metabolic needs, thereby decreasing test specificity (255).

Only three studies investigated the role of ^{99m}Tc-MIBI in indeterminate thyroid nodules. In all studies, evaluation of thyroid nodules was performed by dual-time planar imaging: an early image was made ranging from 10-20 minutes after injection of the radiopharmaceutical and a delayed image 60-120 minutes post injection. The intensity of the ^{99m}Tc-MIBI uptake within the nodule, and possible increased uptake or denoting retention on delayed imaging were assessed and compared to the physiological washout of the tracer from normal thyroid tissue. A visual pattern of increased ^{99m}Tc-MIBI uptake on early images that persisted or further increased on the delayed images was generally considered suspicious for malignancy. The individual study sensitivity and specificity for this interpretation ranged from 56% to 79% and from 52% to 96%, respectively (19,70,258). Despite the limited number of available studies, the performance of ^{99m}Tc-MIBI in indeterminate thyroid nodules seems insufficient and less accurate than in cytologically unselected nodules (255).

Nonetheless, Giovanella *et al.* demonstrated that NPV for this method could increase from 88% to 100% if only the pattern of ^{99m}Tc MIBI uptake lower than or equal to the pertechnetate uptake within the nodule was considered benign. As few benign lesions expressed this uptake pattern, this would decrease the yield of this diagnostic (70).

Piccardo *et al.* did not preselect hypofunctioning lesions, but included all indeterminate thyroid nodules. As expected given the explanation above, the specificity of ^{99m}Tc-MIBI was poor: 52% (19).

Assessment of a retention index of the tracer based on semi-quantitative measurements of the lesion to non-lesion uptake ratios for early and delayed ^{99m}Tc-MIBI images yielded better accuracy. Optimal thresholds for the retention index were determined using ROC analysis and unfortunately not externally validated (70,258). As such, it is unclear whether semi-quantitative ^{99m}Tc-MIBI retention indices are truly more accurate than conventional visual assessment. Moreover, semi-quantitative analysis is still operator dependent, as it depends on the manual definition of ranges of interest (ROI)(19).

^{99m}Tc-MIBI in thyroid nodules with Hürthle cell cytology

Oncocytic cells are rich in mitochondria. Therefore, Hürthle cell lesions – malignant as well as benign – frequently show a more intense and persistent ^{99m}Tc-MIBI uptake (258-260). Boi *et al.* investigated ^{99m}Tc-MIBI in cold thyroid nodules with varying proportions of Hürthle cells in the cytology samples. A relation between ^{99m}Tc-MIBI uptake and increased tissue density of oncocytes was suggested (260). Subsequent studies also concluded that ^{99m}Tc-MIBI is not specific enough to differentiate indeterminate lesions with Hürthle cell cytology (70,255,258). Excluding Hürthle cell nodules from ^{99m}Tc-MIBI assessment likely excludes many false-positive tests while improving benign call rate, specificity and overall diagnostic accuracy in indeterminate thyroid nodules.

Availability, cost-effectiveness and limitations of ^{99m}Tc-MIBI

Imaging of ^{99m}Tc-MIBI requires conventional gamma cameras (with or without single-photon emission computed tomography (SPECT) and CT), which are more widely available than

PET, especially in non-Western countries. Furthermore, the tracer itself is more widely available due to relatively simple complexation using ^{99m}Tc-MIBI-kits together with the favorable half-life of ^{99m}Tc (~6 hours) obtained from on-site generators. The radiation burden of the recommended whole-body adult dose is 5-6 millisievert, but can be lowered by a factor 2-3 by partial-body imaging (261). However, the system resolution of state-of-art gamma cameras is a factor 3 lower than of PET/CT cameras. This decreases the measured signal of lesions smaller than 30 mm, increasingly limiting test sensitivity in smaller nodules. Average costs of ^{99m}Tc-MIBI scanning range from €119 to €500 in Europe and from $669 to $1,156 in the USA (139,262,263). From a German perspective, ^{99m}Tc-MIBI-based management was cost-effective over Afirma® GEC-testing and conventional management. However, this study inappropriately extrapolated auspicious performance parameters of ^{99m}Tc-MIBI in unselected thyroid nodules (96% sensitivity and 46% specificity) to the indeterminate population, and likely underestimated modelled costs for ^{99m}Tc-MIBI scanning and thyroid surgery (14,120,138,139,262-264). Therefore, these assumptions regarding cost-effectiveness in indeterminate thyroid nodules are decidedly questionable and require careful re-evaluation.

Altogether, there is an increased risk of malignancy in thyroid nodules that show increased ^{99m}Tc-MIBI uptake, provided that hypofunctioning nodules are preselected. Nonetheless, test performance in indeterminate thyroid nodules seems insufficient. Excluding Hürthle cell lesions suggests high specificity, but does not resolve the reported poor sensitivity. However, the number of studies currently available for indeterminate thyroid nodules is limited. We believe prospective validation studies in non-oncocytic indeterminate thyroid nodules should be performed. Future studies should also focus on external threshold validation for retention indices to reduce operator dependency and increase accuracy and objectivity of ^{99m}Tc-MIBI. Based on the current evidence, we recommend that ^{99m}Tc-MIBI scanning is not used in

surgical management decisions in indeterminate thyroid nodules without another adjunctive test.

3.2. FDG-PET

Positron emission tomography (PET) using [^{18}F]-2-fluoro-2-deoxy-D-glucose (fluorodeoxyglucose or ^{18}F-FDG), also known as FDG-PET, is an imaging modality that exploits the basic principle that (malignant) tumours and inflammatory tissues are much more metabolically active than normal tissues. Whereas normal tissues predominantly produce energy by low rates of aerobic glycolysis followed by the citric acid cycle in mitochondria, glycolytic rates of rapidly growing cancers can be up to 200 times higher. Subsequent lactic acid fermentation takes place even if oxygen is plentiful (the Warburg effect) (265). Similar to regular glucose, the glucose analogue ^{18}F-FDG is internalized by transmembranous GLUT transporters and converted by hexokinase to ^{18}F-FDG-6-phosphate. However, unlike the 6-phosphorylation product of regular glucose, ^{18}F-FDG-6-phosphate cannot be metabolized further. It is trapped intracellularly and thus accumulates in the tissue. Subsequently, PET scanning can visualize the increased glucose metabolism of the (abnormal) tissue (266). Nowadays, FDG-PET is generally performed in combination with computed tomography (FDG-PET/CT), mainly to correlate metabolically active regions to their anatomic substrates and to correct for tissue-attenuation of the radioactive signal. It is increasingly applied in the diagnostic work-up, staging and therapeutic response monitoring of various malignancies. For thyroid cancer, FDG-PET is frequently used to characterize recurrent disease, especially if dedifferentiation is expected in thyroid carcinomas that lost the capacity to concentrate radioiodide, yet still have measurable serum values of the tumour marker thyroglobulin. It may also be considered in the initial staging of poorly differentiated or invasive Hürthle cell

carcinoma. Moreover, FDG-avid thyroid incidentalomas require additional work-up by FNAC when >1 cm (20,23). In the current ATA guidelines FDG-PET is not routinely recommended for the diagnostic workup of indeterminate thyroid nodules due to limited clinical validation, despite a 2011 meta-analysis by Vriens *et al.* that demonstrated 95% sensitivity and 96% NPV in indeterminate thyroid nodules larger than 15mm (23,267).

Results of available individual studies were mutually consistent despite limited sample sizes. Especially the first studies showed extremely promising results, each reporting 100% sensitivity (266,268-270). De Geus-Oei *et al.* argued that implementation of FDG-PET could reduce the number of futile hemithyroidectomies for benign nodules by 66%, likely outweighing the costs of the extra scans and suggesting cost-effectiveness of this technique in the preoperative setting (269). A subsequent study suggested a less optimistic 39% reduction in futile surgeries, following a lower benign call rate (270). More recent studies demonstrated more modest performance of FDG-PET(/CT) (19,20,271-273). Overall, reported sensitivity and specificity of FDG-PET(/CT) to detect thyroid carcinoma in indeterminate thyroid nodules ranged from 77% to 100% and from 33% to 64%, respectively. A negative index test was reported in approximately 40% of patients (19,20,266,268-274).

Several reasons for false-negativity were proposed, foremost small nodule size. It is how Traugott *et al*. explained their 20% false negative FDG-PET scans: eight lesions were histopathologically smaller than 1 cm. Excluding these, sensitivity and NPV increased to 100% (271). FDG-avidity in very small nodules may be missed on FDG-PET due to the low volume of malignant cells and due to the partial volume effect: the detected FDG-concentration is underestimated dependent on nodule size in relation to the (limited) spatial resolution of the scanner. In larger nodules, this effect is negligible (20,269). Although the improving resolution of state-of-the-art PET scanners pushes the detection limit towards 10 mm, PET is less sensitive in lesions smaller than 15 mm on US. It less reliable to rule-out

microcarcinomas (267). Theoretically, the improving spatial resolution could also become a limitation of the technique: not only will there be less false-negatives, but likely also more false-positive results - leading to a decrease in the already limited specificity over time. In the currently available literature no such downward trend is noted, but future studies should monitor this possibility.

Semi-quantitative FDG-PET

Semi-quantitative analysis of FDG-PET is performed using the maximum standardized uptake value (SUV_{max}): the ratio between the maximum radioactivity concentration measured within a region of interest on the PET image (the 'hottest' voxel) and the decay-corrected amount injected radiotracer per unit of body mass. It reflects the FDG-concentration factor compared to a homogenous distribution of the radiotracer (275). The SUV_{max} is generally significantly higher in malignant than in benign lesions (20,269,270,272,273,275,276). There is a possible correlation between higher SUV_{max} values and increasing size in nodules, insufficiently explained by the abovementioned partial volume effect (20,276). Also, in FTC a higher SUV was associated with capsular or vascular invasion (274). Nonetheless, even though Kresnik *et al.* demonstrated that all carcinoma and Hürthle cell adenoma had an $SUV_{max} \geq 2$ and all other benign lesions an $SUV_{max} < 2$, in multiple other studies the SUV_{max} of benign and malignant indeterminate thyroid nodules overlapped. No threshold could accurately tell them apart (20,269,270,272,273,275,276). Moreover, as SUV_{max} calculations strongly depend on image acquisition and reconstruction methods, type of PET-scanner and other variable methodology, reported absolute SUV_{max} thresholds are not simply valid for other institutions (20). Standardized optimized FDG-PET protocols are required for inter-institution comparison of study results and advancement of PET research (277,278).

FDG-PET in thyroid nodules with Hürthle cell cytology

Multiple studies observed aberrant FDG-PET characteristics in indeterminate nodules with Hürthle cell cytology: both benign and malignant lesions are mostly FDG-positive. Twenty-nine Hürthle cell lesions were reported by Deandreis *et al.*, consisting 52% of their study population and providing an explanation for their limited sensitivity (273). Moreover, Hürthle cell adenoma generally demonstrated a significantly higher SUV_{max} than other benign lesions (20,266,268,269,273,279). The proportion of Hürthle cell cytology in individual studies is relatively small, but overall FDG-PET seems inadequate in these neoplasms.

Availability, cost-effectiveness and limitations of FDG-PET

PET systems are less widely available than conventional gamma cameras. Moreover, ^{18}F used for ^{18}F-FDG synthesis is produced in cyclotrons, and transport distances are limited due to the short half-life of this isotope (~110 min). In Europe, FDG-PET/CT is approximately 1.5-2 times more expensive than ^{99m}Tc-MIBI SPECT/CT. The radiation exposure of FDG-PET/CT is largely accounted for by the FDG dosage at approximately 19 μSv/MBq, i.e. about 3-4 mSv for a typical activity of 185 MBq administered to an average adult (280). Insights regarding common practice total-body FDG-PET/CT imaging are changing (271,273). The CT radiation dose greatly varies, and can be less than 0.5 mSv for a low-dose CT of the neck region only. When scanning the thyroid region only, a longer imaging time can compensate for a reduction in FDG dose, which would lower the radiation burden as well as the costs. Such solutions may counter prevailing reservations regarding ionizing radiation exposure. Additionally, partial-body acquisition could limit the number of coincidental PET-positive findings. Much of the criticism on FDG-PET focuses on these potential incidental findings, which require additional diagnostics, are not always clinically relevant and may negatively impact potential cost-effectiveness (281,282). Malignant ipsi- or contralateral thyroid incidentalomas are

reported while the nodule under investigation was histopathologically benign (271,272). PET-positive incidentalomas are histopathologically malignant in about 20% of patients (282). Cost-effectiveness of FDG-PET/CT was modelled by Vriens *et al.* (14). From a Dutch health care perspective, FDG-PET/CT driven treatment would decrease the rate of unbeneficial diagnostic hemithyroidectomies for benign thyroid nodules by 35% and reduce the costs per patient by €822 compared to the €8,804 expenses for conventional surgical treatment. Also, FDG-PET/CT was favoured over the miRInform® and Afirma® GEC (14).

Contrasting the generally strong sensitivity, specificity of FDG-PET is consistently poor. The underlying mechanism is not yet fully elucidated. The negative influence of Hürthle cell cytology may be partly responsible. It could also be explained by cellular atypia, which was significantly and independently related to FDG uptake, and found in both benign and malignant lesions. Atypia was also related to the presence of Hürthle cells (273). Sebastianes *et al.* hypothesized that FDG uptake is related to variations in gene expression patterns. They suggested that genetic variations between populations may also explain the varying diagnostic accuracy of FDG-PET between studies (270).

In conclusion, FDG-PET(/CT) has the potential to accurately rule-out malignancy in all indeterminate nodules except Hürthle cell lesions. It could prevent unnecessary diagnostic surgery for a significant number of benign thyroid nodules. Sample sizes of existing studies are small, but larger prospective trials are currently ongoing to settle the diagnostic value of this technique and its utility in clinical practice. We recommend that these studies also focus on identifying (genetic) causes for the occasional false-negativity and generally low specificity of this technique.

3.3. DW-MRI

Diffusion-weighted magnetic resonance imaging (DW-MRI) is a functional nuclear magnetic resonance imaging technique that evaluates the rate of random (Brownian) motion of water in tissue, also called diffusivity. By applying diffusion-sensitizing magnetic gradients (the strength and duration of which are expressed as b-values) different levels of diffusion-weighting are obtained: from non-diffusion images (b-value = 0 s/mm^2) to highly diffusion weighted images (i.e. b-value >800 s/mm^2)(283). Lesions that show high signal intensity on DW-MRI images with a high b-value thus show restricted diffusion. The apparent diffusion coefficient (ADC, in mm^2/s) is calculated based on the exponential relationship between signal intensity and the corresponding b-value according to $S(b)=S(0)*e^{-b*ADC}$. A high ADC represents a high degree of diffusion; a low ADC represents diffusion restriction (283,284). DW-MRI thus allows noninvasive quantification of tissue properties without ionizing radiation exposure for the patient. Differentiation between benign and malignant tissues by DW-MRI is based on the assumption that increased cell proliferation, cellular-density and disorganized structures in malignant tissue restrict random motion and thus diffusion of water: a lower ADC-value, together with high signal intensity at high b-values, is more suspicious for malignancy (283,284). Oppositely, increased ADC-values suggest free movement of water molecules in the tissue. It is found in for example edema, colloid follicles, fibrous tissue, hemorrhage and calcification, all of which associated with benign tissues (285). Prior application of DW-MRI in i.e. neuroradiology, breast and lymph nodes showed high diagnostic accuracy (286,287).

Recent exploratory studies in small cohorts of thyroid nodules found distinctively higher ADC values for benign than malignant nodules (283-285,288-292). A recent meta-analysis in 765 cytologically unselected thyroid nodules estimated that DW-MRI had 90% sensitivity and 95% specificity to distinguish thyroid carcinoma (293). Among the individual studies,

however, presented optimum ADC thresholds varied and were not externally validated (283-285,288-292).

Only one small study had assessed DW-MRI in indeterminate thyroid nodules to date. Nakahira *et al.* reported a mean ADC value of $1.27 \pm 0.29 *10^{-3}$ mm^2/s in malignancies opposite $1.95 \pm 0.24 *10^{-3}$ mm^2/s in benign nodules with indeterminate cytology. These results were similar to those of their entire study population (n=42), in which a cutoff ADC value of $1.95 \pm 0.24 *10^{-3}$ mm^2/s was 95% sensitive and 83% specific (283).

Availability and limitations of DW-MRI

DW-MRI is infrequently and only experimentally used in the workup of thyroid nodules. Nonetheless, the worldwide availability and application of MRI is growing. As it uses no ionizing but only radiofrequency radiation, the associated risk to the patients is limited, provided that specific measures are taken for patients with MRI-incompatible implanted devices or metal. No MRI-contrast is necessary for DW-MRI, thus avoiding gadolinium-associated toxicity. As the spatial resolution of MRI-scanners is still improving, technical limitations of DW-MRI with regard to minimal lesion size are becoming less relevant compared to SPECT and probably also PET. Still, spatial resolution of DW-MRI sequences is less than that of conventional anatomical MRI-sequences.

There are several major limitations to DW-MRI. MRI is still a rather costly technique; additional sequences such as DW-MRI adds scanner time (~5-10 min) per patient and thus further increases costs. DW-MRI methodology is not standardized yet and its optimal settings still unsettled, leading to varying ADC and b-values (283,289,292). Suboptimal methodology or artifacts cause poor image quality, impede accurate interpretation and caused undesirable exclusions from already small-sized studies, with reported exclusion rates up to 28% (283,284,292). Image artifacts are often caused by inhomogeneity in pathologic tissues or by

their vicinity to interfaces between soft-tissues and bone or air, a source of MRI-artifacts specifically in the thyroid region. Besides viable tumor tissue, malignant tumors partly exist of components with high diffusivity, such necrosis, cystic components or intratumoral hemorrhage (283,285). For accurate ADC measurement, such macroscopic areas should be manually avoided when drawing a region-of-interest. However, avoiding microscopic areas of similar origin, invisible to the human eye, is an impossible task (283). Furthermore, it is hypothesized that the substantial amounts of follicular or Hürthle cells limit the diagnostic accuracy of DW-MRI, specifically in indeterminate thyroid neoplasms. Follicular and Hürthle cell neoplasms are known for their varying colloid tissue involvement. Histologically they contain more fluid. Thus, DW-MRI would inaccurately provide a more benign image (283,292). These hypotheses are currently based on very limited evidence. Further prospective validation studies are desired to determine the possible diagnostic value of DW-MRI in indeterminate thyroid nodules. Future prospects also include improvements of the technique, including consensus on methodology and standardization of acquisition techniques.

4. COMBINED AND MULTISTEP DIAGNOSTICS

The previous chapters of this review addressed the large number of available diagnostic tools to assess indeterminate thyroid nodules. Most studies focused on a single diagnostic technique only. The elimination of between-study population-level differences is a major advantage when comparing the performance of multiple diagnostics independently in one study, optimally in a prospective, independent and blinded fashion. Moreover, assessment of multiple techniques in one study allows investigation of the complementary value of multiple techniques as a *diagnostic* tool by means of simultaneous or sequential testing while at the

same time aiding to further unravel tumor biology as a *research* tool, especially in the current multidisciplinary in-hospital working environment. For example, the question how the presence of a certain oncogenic mutation relates to the (positive) result of an FDG-PET scan could be addressed.

Piccardo *et al*. compared ^{99m}Tc-MIBI, FDG-PET/CT and US plus USE in 87 indeterminate TIR3 nodules with a 21% malignancy rate. FDG-PET/CT was the superior technique with 94% sensitivity and 58% specificity. Following a non-specific positive FDG-PET result, review of ultrasound characteristics offered slight further differentiation; it improved specificity to 77%. However, an additional negative ^{99m}Tc-MIBI scan increased specificity to 94%; this combination was found in 13% of patients (19).

Giovanella *et al.* performed both ^{99m}Tc-MIBI and a 7-gene mutation panel in cold indeterminate thyroid nodules. Combined testing did not improve diagnostic accuracy. Performance of the gene mutation panel was inferior to ^{99m}Tc-MIBI imaging. Of the seven (11%) mutation-positive nodules (four *RAS* mutations and three PAX8/PPARγ rearrangements), only four were malignant. It is unclear whether the low sensitivity of the gene mutation panel in this study can be explained by the selected population of hypofunctioning nodules (70).

Elastosonography and Ultrasonography

USE is superior to ultrasound in indeterminate thyroid nodules – both individual US characteristics as well as combined US patterns described in various articles (210,213,215,217,219,221,222). Two recent prospective studies demonstrated that additional USE evidently improved the diagnostic accuracy of US. Garino *et al.* included nodule stiffness as additional characteristic into a panel of US characteristics and demonstrated that USE identified eight additional malignancies that would have been missed by US assessment

alone. Presence of one or more suspicious US/USE features was 100% sensitive; two or more 88% sensitive and 77% specific. Benign test results were found in 57% of patients. The authors suggested that the 6.4% remaining risk of malignancy– similar to the benign cytology category – would justify follow-up instead of diagnostic hemithyroidectomy in this group (210). In another study of 315 Thy3 nodules, semi-quantitative USE correctly diagnosed 75% of the histopathologically benign lesions that were considered suspicious for malignancy on US, and 83% of the malignancies that were misdiagnosed as benign on US (217). These results suggest that the existing TIRADS classification could be extended with tissue elasticity features. In unselected thyroid nodules this improved TIRADS sensitivity, but not specificity (240,294). The combination is a suitable topic for future research in indeterminate thyroid nodules. Major benefit is that the two techniques are individually inexpensive and obviously easily combined during one diagnostic procedure. Cost-effectiveness can be anticipated.

US and Mutation Analysis

US assessment was also reported in various studies on gene mutation analysis, presumably because US data were usually readily available in clinical studies at no additional costs and thus easily combined with results of more experimental techniques. Even though US assessment improved the diagnostic accuracy of both FDG-PET and elastosonography, combined use of ultrasound with the sensitive Afirma® GEC or specific *BRAF* mutation analysis demonstrated little additional diagnostic value (57). Suspicious US features such as hypoechogenicity, presence of calcifications and hypervascularity were not predictors of malignancy in Afirma® GEC-positive nodules (144). Also, as expected by their individual association to classic PTC, a positive $BRAF^{V600E}$ mutation was correlated to the presence of suspicious US features in unselected nodules, including hypoechogeneity and the presence of

microcalcifications (29,66,78). *BRAF* mutation less frequently occurred in thyroid nodules without suspicious US features (66,78). In Bethesda III and IV thyroid without suspicious US features the prevalence of the *BRAF* mutation was only 1.5% (1/67) in the study by Seo *et al.* – very low, particularly for a South Korean population – all while the malignancy rate was still 18% (12/67)(66). Considering the negligible yield at additional costs, *BRAF* mutation analysis might not be contributory in indeterminate nodules without suspicious US features. An even lower yield from *BRAF* mutation analysis in US-unsuspicious nodules is presumed in populations with a lower general prevalence of *BRAF* mutations. Additionally, these results suggest a different US appearance of *BRAF* mutation-negative malignancies – or a different molecular profile of thyroid carcinoma without suspicious US features.

RAS mutation analysis and assessment of the typical suspicious US features could be complementary in the differentiation of indeterminate thyroid nodules, as follicular-type thyroid carcinomas are associated with *RAS* mutations and infrequently showed the typically suspicious US features (1,29,193-196). Combined assessment could improve diagnostic accuracy of either technique in indeterminate thyroid nodules, identifying papillary thyroid malignancies through classic suspicious US features and follicular-type carcinoma by *RAS* mutation analysis. However, challenges for clinical practice continue to exist in the imperfect specificity of *RAS* mutation analysis, and the interobserver variability and ambiguity of certain US features.

Immunocytochemistry and Mutation Analysis

In histopathology samples, certain genetic alterations were correlated to positive staining for specific immunomarkers: PAX8/PPARγ rearrangement was associated with galectin-3 reactivity, and *RAS* point mutation with HBME-1 (105). Only one study investigated this combination of techniques in indeterminate thyroid cytology. Although no significant

correlation was demonstrated between positive $BRAF^{V600E}$ mutation and galectin-3 overexpression – benefitting possible complementary use – no additional diagnostic value was demonstrated either (44).

MicroRNA and Mutation Analysis

Combined microRNA expression profiling and mutation analysis could accurately aid diagnosis and prognosis of thyroid malignancy. Distinct microRNAs have been related to oncogenic mutations. For example, miR-221, miR-222 and miR-146b were more overexpressed in *BRAF*- and *RAS*-mutated PTC. High expression of miR-187 was associated with RET/PTC rearrangement (148,295). The first step towards diagnostic integration of the two techniques was taken by Labourier *et al.*, who tested the commercially developed 10-microRNA thyroid classifier ThyraMIR™ simultaneously with the miRInform® thyroid (22). The ThyraMIR™ was designed to increase the sensitivity of the miRInform® without affecting its specificity. Combined use demonstrated 89% sensitivity and 85% specificity (22). A recent decision analytics model for Bethesda III and IV nodules estimated that combined miRInform® and ThyraMIR™ testing was cost-effective, reducing the rate of unnecessary surgery (diagnostic hemithyroidectomy as well as two-step thyroidectomies) from 88% to 20% and saving $1,384 per patient in the first year of treatment or $3,170 per avoided surgery. However, it is not described how the economic consequences of the 15% missed malignancies are accounted for in this model (140). The economic as well as medical-ethical consequences of such a high number of missed malignancies question the current clinical utility of this combination of expensive techniques.

In brief, the combined or sequential use of multiple diagnostics in indeterminate thyroid nodules was infrequently studied. Regrettably, the available studies also mostly remained

within their own field of expertise: comparing tests either within the domain of pathological (molecular) techniques or within the domain of imaging. Although a sequential combination of a sensitive and an uncorrelated specific test might bring the solution that this clinical issue has been waiting for, the most accurate combination of tests cannot reliably be determined yet.

5. RECENT DEVELOPMENTS AND FUTURE PROSPECTS

5.1. The Cancer Genome Atlas

Papillary thyroid cancer was one of the cancers targeted by the cancer genome atlas (TCGA) research network, a large collaborative project by the National Cancer Institute (NCI) and National Human Genome Research Institute (NHGRI). The incentive of the project is to map genomic alterations occurring in 33 types of cancer in 11.000 patients and improve the understanding, classification and extending possibilities for targeted therapy of these cancers (296). Genetic alterations of all kinds were detected in nearly five hundred clinically non-aggressive PTCs (classical, follicular and tall cell variants) using one proteomic and six genomic platforms. PTC harbored fewer somatic mutations than other human cancer types, but if they were present, driver mutations were detected in the majority of the cancer cells. As expected, the known driver mutations in the MAPK/ERK pathway were dominant, confirming the mutually exclusive relation for *BRAF* and *RAS* point mutations and RET/PTC rearrangements. Other detected genetic alterations included genetic variations of the TERT promoter, PI3K and PPARγ pathways, as well as new alterations of known and new drivers, such as *EIF1AX*, *PPM1D* and *CHEK2*. Moreover, molecular subtypes of for example *BRAF*-mutated PTC were identified and linked to different clinical subtypes. The role of microRNA

in determining cancer phenotype was elaborated, allowing better understanding of clinical behavior of various genetic variants of PTC. Somatic copy number alterations were mostly linked to FVPTC. Ultimately, the TCGA Research Network envisions a reclassification of thyroid carcinoma, abandoning the discrimination between PTC and FTC, and classifying according to molecular subtypes instead of by histopathological subtype first (297). The identified markers may not just have an application in the diagnosis of thyroid carcinoma, but also in better risk-stratification of the different cancers and in targeted therapies. The plurality of applications is best known for the $BRAF^{V600E}$ mutation, which has an association with clinically more aggressive tumor behavior on several fronts. Also, non-thyroid malignancies carrying a *BRAF* mutation are now (experimentally) treated with RAF inhibitors (298,299). There is little doubt that molecular classification systems are the future of oncology diagnostics in all types of human cancers. The position of histopathological assessment is changing, but cannot be renounced. With the current knowledge of thyroid genomics, the need to distinguish the mutated malignant from the mutated benign – premalignant – neoplasms remains, with all due consequences for the surgical and postoperative treatment strategy.

Cytological application of the TCGA set was also already investigated in a recent study. Pagan *et al*. validated a panel containing the genomic alterations identified by the TCGA in 88 FNAC samples selected from a previous cohort study, including 22 indeterminate thyroid nodules (126,300). In the latter, 33% sensitivity and 84% specificity were demonstrated. In the same set of patients, Pagan *et al*. also performed the Afirma® GEC. The GEC yielded less false negatives and a much higher sensitivity. Even though technical limitations of the applied sequencing techniques could leave RNA transcriptions with low expression levels undetected and thus negatively influence sensitivity of the TCGA set, the scopes of the TCGA and GEC

most likely explain their difference in performance. The TCGA was developed using PTC only. It did not include follicular lesions and their distinctive genetic alterations. Moreover, in contrast to the GEC, the TCGA set was not optimized for preoperative diagnostic application in indeterminate thyroid nodules (300). Consequently, the comparison performed by this Veracyte-sponsored study seems unjust: it is obvious that the Afirma® GEC yielded better diagnostic performance in this specific clinical setting. Yet, the results of this study did prove that a large panel of genetic alterations such as the TCGA was not useful in clinical practice without further expansion of the scope of the panel towards follicular thyroid neoplasms. Still, the genetic alterations and their relations detected by TCGA are groundbreaking for the progression of research. From these comprehensive sets of biomarkers, we may select new combinations of genetic alterations for future clinical research to develop an accurate rule-in or rule-out molecular test for indeterminate thyroid nodules.

5.2. Proteomics

Other molecular advances include protein expression diagnostics, or proteomic profiling. These techniques allow for more detailed insight in the molecular biology and protein expression of thyroid neoplasms. For example, matrix-assisted laser desorption ionization / mass spectrometry imaging (MALDI-MSI) is able to simultaneously visualize the spatial distribution of proteins and profile up- and downregulated protein expression in relation to the morphological features of the thyroid specimen. These and related proteomic techniques could identify new biomarkers for preoperative cytological diagnosis, but require high levels of expertise. Application to thyroid cytology has so far been investigated by few studies (301,302). Ex-vivo cytology studies show accurate and reproducible differentiation between various lesions, including the currently difficult to diagnose Hürthle cell neoplasms (302). No

studies investigated the diagnostic value of proteomics in in-vivo indeterminate thyroid cytology yet.

DISCUSSION

This review provides a comprehensive overview of the available literature on molecular and imaging biomarkers as additional diagnostics for thyroid nodules with indeterminate cytology (Bethesda III and IV) and their application in a clinical preoperative setting. Clinical utility requires more from a diagnostic than mere well-validated test performance and high rule-in or rule-out capacity. The 2015 ATA guidelines suggested that the ideal rule-out diagnostic for thyroid carcinoma should have a NPV similar to a benign cytological diagnosis (~96.3%) and the ideal rule-in test a PPV that is at least similar to a malignant cytological diagnosis (~98.6%) (10,23). The balance between test sensitivity and specificity – and their prevalence-dependent derivatives PPV and NPV – directly reflects on feasibility and cost-effectiveness estimates. A diagnostic with (near) perfect sensitivity but limited specificity is inefficient and unlikely cost-effective: the NPV will be close to 100%, but the majority of nodules will test positive. Therefore, instead of focusing on the reproducible highest sensitivity or specificity, a diagnostic is better appreciated by end points such as desired minimal rates of accurately prevented unbeneficial surgeries or accurately diagnosed carcinomas. More importantly, clinical utility demands that implementation of the ancillary test leads to changes in patient management and overall health benefits (303). All these requirements directly depend on a plurality of epidemiological and economic factors within the tested population, such as the local test availability, professional expertise and case mix – prevalence of malignancy as well as the balance of various subtypes of indeterminate cytology including especially Hürthle-cell neoplasms and *BRAF*-mutation. Additionally, clinical utility considerations should include

less tangible factors such as physician and patient preference, multidisciplinary decision making and compatibility with everyday clinical routine and logistics in endocrine practice. All things considered, global perspectives regarding the preferred diagnostic for indeterminate thyroid nodules likely greatly differ.

Recommendation for clinical use of rule-out tests

The most accurate currently available rule-out tests are the Afirma® GEC and FDG-PET(/CT) imaging. The Afirma® GEC had strikingly high sensitivity in nearly all studies (127,129,131-134). However, there are concerns regarding the lack of strong validation studies. With a high degree of missing histology, especially in GEC negative nodules, there is a potentially strong diminution of the tests' sensitivity if unresected GEC-negative lesions were less often benign than presumed. In the USA, physicians should locally validate the tests' utility prior to implementation. However, with its limited global availability, high costs and low probability of cost-effectiveness, clinical implementation of the Afirma® GEC outside the USA is currently not favored (14,121,137-140).

FDG-PET/CT may be the preferred rule-out test for indeterminate thyroid nodules in a European setting. With sufficient validation studies with complete histopathological follow-up, it demonstrated consistent high sensitivity and a benign test result in 40% of the patients, although the number of currently published patients is moderate. Cost-effectiveness of FDG-PET over other diagnostics is presumed (14). Its popularity in the USA is more limited, although the efficacy of this molecular imaging technique could likely compete with molecular biomarkers panels, even if the costs per scan are somewhat higher than in Europe. The main drawback of FDG-PET/CT is its – admitted minor – risk to the patient by using a limited dose of ionizing radiation.

The recently announced version 3 of the ThyroSeq® may become a prime contender. Dependent on the case mix, the ThyroSeq® v2.1 anticipated high negative predictive value (21). However, the number of studies to confirm test performance and clinical utility in different patient populations is limited. Clinical results for the ThyroSeq® v3 are eagerly awaited.

Semi-quantitative elastosonography could be a suitable alternative, in particular in case a more economic test is required. However, overfitting and lack of external cut-off validation likely overestimated the performance of this technique in the limited number of available studies. If future prospective studies can confirm its performance and thresholds of this operator-dependent but globally accessible method, USE could become a more important diagnostic in this field.

None of the diagnostic techniques under investigation in this review has a perfect NPV or fulfills the threshold proposed by the ATA. A number of malignant nodules will be misdiagnosed as benign on first assessment. Considering the typical indolent clinical course of differentiated thyroid cancer, follow-up of these initially false-negative nodules will most likely still result in timely diagnosis without relevant treatment delay and dismal prognostic consequences.

Recommendation for clinical use of rule-in tests

The best rule-in performance was unmistakably demonstrated by *BRAF* mutation analysis, which showed perfect 100% specificity in an abundance of studies. Yet, strong regional differences in prevalence of *BRAF* mutations have a major impact on its clinical utility, especially when comparing South Korea to other countries. Moreover, the analysis most likely has very low yield in Bethesda IV nodules, in which the mostly follicular type malignancies are more frequently *RAS*-mutated (31,50,70,76). Testing for individual genetic

alterations other than the $BRAF^{V600E}$ point mutation is not useful. In American and European settings, a gene mutation panel is likely preferred over any individual mutation analysis. Promising rule-in capacity was also demonstrated for Galectin-3 immunocytochemistry. An infrequently applied technique with limited validation studies, further prospective studies are warranted to validate its performance in indeterminate thyroid nodules and endorse its possible clinical use.

Besides *BRAF* mutation analysis, none of diagnostics meet the 2015 ATA requirements of an ideal rule-in test. Compared to ruling-out tests, ruling-in tests face an additional challenge. With a generally low frequency of thyroid carcinoma in indeterminate thyroid nodules, achieving a reliable PPV – higher than 95% – can be a major challenge despite adequate test specificity. Such high demands to a ruling-in test advocate the use of a ruling-out test in populations with a limited pre-test probability of malignancy.

Clinical recommendation for a step-wise approach

Most of the diagnostic modalities are optimized for either ruling in or ruling out malignancy. No single diagnostic addressed in the current review currently has it all: both a near-perfect sensitivity and a near-perfect specificity, and (proven) cost-effectiveness. It is extremely challenging to develop such test performance parameters in a single diagnostic. Even promising new diagnostics, such as the ThyroSeq® and ThyraMir™, require significant further optimization to get near this diagnostic utopia.

With the diagnostics currently available in the clinical setting, a multimodality stepwise approach could offer a conclusive diagnosis for indeterminate thyroid nodules, sequentially combining one sensitive rule-out and one specific rule-in test. Unfortunately, thus far few studies investigated this approach (19,70). Combinations of (molecular) imaging and somatic genetics were especially scarce. There is currently insufficient evidence to accommodate

reliable interpretation of sequentially used tests, as performance of the second test is unknown in a population preselected by the first. Besides choosing two accurate and uncorrelated tests to achieve maximum diagnostic accuracy, the sequence of testing, local availability and costs of the selected diagnostics are crucial. Costs of two or more additional tests may compromise cost-utility estimates. Available cost-effectiveness studies for individual diagnostic modalities were additionally greatly susceptible to global variations in population-dependent factors such as pre-test probability of thyroid carcinoma and local test performance, and varying health care costs including the surgical reimbursement rates (14,121,138,139). Reported surgical and hospitalization costs range from $4,628 to $6,549 for hemithyroidectomy, $5,272 to $7,068 for completion thyroidectomy and $5,680 to $11,265 for initial total thyroidectomy. Secondary expenses following surgery should be considered as well, including postoperative observation, thyroid hormone replacement (approximately $150 per patient per year), treatment for hypoparathyroidism (approximately $860 per patient per year), and resolution of rare but potentially serious surgical complications (14,120,138,264). Secondary endpoints such as quality of life and survival are of minor importance to cost-effectiveness, due to the generally indolent course of differentiated thyroid cancer, adequate treatment options and overall low disease-related mortality (14,138,139).

Recent discussions in thyroid histopathology

Histopathology is classically based on microscopic assessment of tumor phenotype, aided by immunohistochemistry. However, this ‘gold standard test’ is also subject to advancing insights regarding tumor phenotype, increasingly aided by knowledge regarding tumor genotype. Mutation-negative malignancies resulting from indeterminate cytology were frequently identified as encapsulated follicular variants of papillary thyroid carcinoma without histologic features of aggressive behavior (21,22,31,59,80). Also, several studies defined a

separate intermediate *histopathological* category called '(follicular) tumor of uncertain malignant potential' for encapsulated, well-differentiated follicular tumors with questionable PTC-type nuclear changes (71,164,177,273). These examples illustrate one of the important ongoing discussions in thyroid histopathology. In 2016, Nikiforov *et al*. proposed an official downscaling of the classification of proven noninvasive encapsulated FVPTCs, renaming them 'noninvasive follicular neoplasm with papillary-like nuclear features' (NIFTP). The behavior of these neoplasms is benign unlike other thyroid carcinoma subtypes, showing no evidence of recurrent disease after a median 13-year follow-up. About one in four of the neoplasms in the retrospective cohort were mutated, most frequently carrying *RAS* (*NRAS*) or PAX8/PPARγ alterations. Presence of a mutation likely predisposes the NIFTP to progress into an invasive encapsulated FVPTC, justifying surgical resection. Treatment of NIFTP should most likely be limited to hemithyroidectomy, waiving totalizing thyroidectomy and radioiodine ablation (304). Although revolutionizing, this new nomenclature complicates mutation-based preoperative decision-making (21,31,80). The justification to skip two-stage surgery and perform a total thyroidectomy at once for mutation-positive nodules is the driving force of the 7-gene mutation panel and similar tests, but would be overkill for the subgroup of NIFTP (31). Nonetheless, most of the undesirable possible overtreatment for NIFTP is likely resolved if *RAS*-mutated indeterminate nodules are treated with hemi- instead of total thyroidectomy, as previously suggested. No comprehensive diagnostic test is currently available to diagnose mutation-positive NIFTP preoperatively, as follicular tumor invasiveness and encapsulation cannot be distinguished on cytology.

Hürthle cell cytology

The Achilles heel of many diagnostics investigated in this review is cytology suspicious for a Hürthle cell neoplasm (Bethesda IV SHCN/HCN). Hürthle cells are oxyphilic cells with

abundant cytoplasm and an enlarged nucleus with a prominent nucleolus. They are found in benign thyroid diseases such as Hashimoto's thyroiditis, but also occur in the notorious Hürthle cell adenoma and carcinoma, the oncocytic variant of follicular adenoma and carcinoma (4,185). Although Hürthle cell carcinomas (FTC-OV) are rare, their aberrant clinical course and association with invasive features justifies the special attention given to Hürthle cell cytology by the Bethesda and other classification systems. An accurate additional diagnostic is desired. Disappointingly, several studies concluded that the investigated test was accurate in all except Hürthle cell lesions (70,132,258,273). Immunocytochemistry handed some solutions, although promising results of combined galectin-3 and CK-19 staining have not yet been validated (174). Besides that, *BRAF*, *RAS*, RET/PTC or PAX8/PPARγ alterations are only occasionally found (70,76). These findings support previous presumptions that oncocytic thyroid nodules are a completely separate entity with a unique molecular and phenotypic profile (305-308). Malignant transition in Hürthle cell nodules most likely involves the PIK3CA-Akt-mTOR and Wnt/bèta-catenin pathways rather than the MAPK/ERK pathway (305,308). Rare *TP53* mutations, usually associated with poorly differentiated and anaplastic carcinoma, were recently also identified in well-differentiated Hürthle cell nodules (306). Also, recurrent FTC-OV have shown genome haploidisation, a rare phenomenon in other types of differentiated thyroid carcinoma (309). Specific markers for the preoperative molecular differentiation of Hürthle cell nodules should be developed. Adaptation of existing tests to additionally suit Hürthle cell nodules (e.g. the Afirma® GEC) is a strategy being explored, for example by the ThyroSeq® v3 and the Afirma® Gene Sequence Classifier. Caution should be taken that these adaptations do not decrease the diagnostic accuracy for non-oncocytic lesions. MicroRNA expression profiling of these lesions is currently also under investigation (148,152).

Strengths and limitations of the current review

There are several important strengths and limitations to this comprehensive review. This review provides a complete overview of the available additional diagnostics for indeterminate thyroid nodules, resulting from a careful and systematic literature selection and quality appraisal. Different types of clinical data of various levels of evidence were considerately presented. Nonetheless, this review is generally prone to inaccuracies from low study quality, study heterogeneity and different types of bias. For some of the assessed diagnostics, the limited number of available publications and small study cohorts contribute to heterogeneity of data and loss of applicability. This mainly concerns studies on non-routine imaging techniques. By nature, these clinical studies need to prospectively include subjects to voluntarily undergo an extra investigation with – at least in the clinical validation phase – no implications for individual patient management. These types of studies require more resources than 'further use' tissue biobank studies. Consequently, the number of studies is more limited and published series often are small. In contrast, cytological biomarker research gratefully profits from available large tissue biobanks for initial validation studies. We believe consistent results from properly designed imaging studies should not be disregarded due to mere their sample size, but be appreciated by the quality of their study design and statistics. Population-level study differences were often observed, not only related to test performance but also strongly varying malignancy rates that were oftentimes much lower or higher than expected from indeterminate thyroid nodules. Besides insuperable epidemiological variations, the selection of indeterminate cytology, and the retrospective nature of many studies may have contributed to these discrepancies.

The type of indeterminate cytology included by individual studies varied, likely leading to between-study heterogeneity. Besides global variations and known intra- and interobserver discordance, diverse definitions of indeterminate cytology were adhered (5). Nowadays, the

Bethesda system differentiates indeterminate from benign and suspicious cytology in a more standardized manner in both literature and clinic. Bethesda III and/or IV and similar categories from other classification systems were frequently applied. Unfortunately, some studies also included small numbers of Bethesda V nodules without presenting results for individual categories separately (127). Many other studies adhered to their own definition of indeterminate cytology. This especially, but not exclusively, concerns studies published before the introduction of the Bethesda system in 2009.

Retrospective study designs and subsequent selection bias – only including indeterminate thyroid nodules that had undergone both thyroid surgery and (routine) pre-operative testing – likely also caused overestimation of the true efficacy of certain techniques (e.g. *BRAF* mutation analysis or ultrasound).

CONCLUSION AND RECOMMENDATIONS

In current-day practice, there are numerous additional diagnostics available to further assess thyroid nodules with indeterminate cytology, all with advantages and disadvantages. This review provided a comprehensive overview of the available literature on these techniques, addressing both molecular and imaging biomarkers, aiming to provide an objective and nuanced comparison of their performance and cost-effectiveness with regard to rightful surgical decision-making. Many of these diagnostics have either an adequate rule-in or rule-out capacity, but no single currently available test seems to serve both purposes well. Diagnostics from the different research fields likely complement each other in a multimodality stepwise diagnostic approach towards. Notwithstanding, test performance is always population-dependent. To correctly interpret the results, the prevalence of malignancy and the performance, costs and feasibility of the desired diagnostic in the local patient

population should be known beforehand. Local implementation studies are strongly recommended to confirm clinical utility. Most importantly, the local decision favoring or opposing a certain diagnostic should be a deliberate and multidisciplinary one. Cooperation between clinical endocrinologists, endocrine surgeons, pathologists, radiologists and nuclear medicine physicians is crucial.

References

1. Gupta N, Dasyam AK, Carty SE, Nikiforova MN, Ohori NP, Armstrong M, Yip L, LeBeau SO, McCoy KL, Coyne C, Stang MT, Johnson J, Ferris RL, Seethala R, Nikiforov YE, Hodak SP. RAS mutations in thyroid FNA specimens are highly predictive of predominantly low-risk follicular-pattern cancers. The Journal of clinical endocrinology and metabolism 2013; 98:E914-922

2. Zhu Z, Gandhi M, Nikiforova MN, Fischer AH, Nikiforov YE. Molecular profile and clinical-pathologic features of the follicular variant of papillary thyroid carcinoma. An unusually high prevalence of ras mutations. American journal of clinical pathology 2003; 120:71-77

3. Lee SR, Jung CK, Kim TE, Bae JS, Jung SL, Choi YJ, Kang CS. Molecular genotyping of follicular variant of papillary thyroid carcinoma correlates with diagnostic category of fine-needle aspiration cytology: values of RAS mutation testing. Thyroid 2013; 23:1416-1422

4. Cibas ES, Ali SZ, Conference NCITFSotS. The Bethesda System For Reporting Thyroid Cytopathology. American journal of clinical pathology 2009; 132:658-665

5. Cibas ES, Baloch ZW, Fellegara G, LiVolsi VA, Raab SS, Rosai J, Diggans J, Friedman L, Kennedy GC, Kloos RT, Lanman RB, Mandel SJ, Sindy N, Steward DL,

Zeiger MA, Haugen BR, Alexander EK. A prospective assessment defining the limitations of thyroid nodule pathologic evaluation. Annals of internal medicine 2013; 159:325-332

6. Mitchell AL, Gandhi A, Scott-Coombes D, Perros P. Management of thyroid cancer: United Kingdom National Multidisciplinary Guidelines. The Journal of laryngology and otology 2016; 130:S150-S160

7. Nardi F, Basolo F, Crescenzi A, Fadda G, Frasoldati A, Orlandi F, Palombini L, Papini E, Zini M, Pontecorvi A, Vitti P. Italian consensus for the classification and reporting of thyroid cytology. Journal of endocrinological investigation 2014; 37:593-599

8. Brito JP, Al Nofal A, Montori VM, Hay ID, Morris JC. The Impact of Subclinical Disease and Mechanism of Detection on the Rise in Thyroid Cancer Incidence: A Population-Based Study in Olmsted County, Minnesota During 1935 Through 2012. Thyroid 2015; 25:999-1007

9. Fazeli R, Schneider EB, Ali SZ, Zeiger MA, Olson MT. Diagnostic Frequency Ratios Are Insufficient to Measure Laboratory Precision with The Bethesda System for Reporting Thyroid Cytopathology. Acta cytologica 2015; 59:225-232

10. Bongiovanni M, Spitale A, Faquin WC, Mazzucchelli L, Baloch ZW. The Bethesda System for Reporting Thyroid Cytopathology: a meta-analysis. Acta cytologica 2012; 56:333-339

11. Sullivan PS, Hirschowitz SL, Fung PC, Apple SK. The impact of atypia/follicular lesion of undetermined significance and repeat fine-needle aspiration: 5 years before and after implementation of the Bethesda System. Cancer cytopathology 2014; 122:866-872

12. Caplan RH, Wester SM, Lambert PJ, Rooney BL. Efficient evaluation of thyroid nodules by primary care providers and thyroid specialists. The American journal of managed care 2000; 6:1134-1140

13. Khalid AN, Quraishi SA, Hollenbeak CS, Stack BC, Jr. Fine-needle aspiration biopsy versus ultrasound-guided fine-needle aspiration biopsy: cost-effectiveness as a frontline diagnostic modality for solitary thyroid nodules. Head & neck 2008; 30:1035-1039

14. Vriens D, Adang EM, Netea-Maier RT, Smit JW, de Wilt JH, Oyen WJ, de Geus-Oei LF. Cost-effectiveness of FDG-PET/CT for cytologically indeterminate thyroid nodules: a decision analytic approach. The Journal of clinical endocrinology and metabolism 2014; 99:3263-3274

15. McHenry CR, Slusarczyk SJ. Hypothyroidisim following hemithyroidectomy: incidence, risk factors, and management. Surgery 2000; 128:994-998

16. Rosato L, Avenia N, Bernante P, De Palma M, Gulino G, Nasi PG, Pelizzo MR, Pezzullo L. Complications of thyroid surgery: analysis of a multicentric study on 14,934 patients operated on in Italy over 5 years. World journal of surgery 2004; 28:271-276

17. Jeannon JP, Orabi AA, Bruch GA, Abdalsalam HA, Simo R. Diagnosis of recurrent laryngeal nerve palsy after thyroidectomy: a systematic review. International journal of clinical practice 2009; 63:624-629

18. Samir AE, Dhyani M, Anvari A, Prescott J, Halpern EF, Faquin WC, Stephen A. Shear-Wave Elastography for the Preoperative Risk Stratification of Follicular-patterned Lesions of the Thyroid: Diagnostic Accuracy and Optimal Measurement Plane. Radiology 2015:141627

19. Piccardo A, Puntoni M, Treglia G, Foppiani L, Bertagna F, Paparo F, Massollo M, Dib B, Paone G, Arlandini A, Catrambone U, Casazza S, Pastorino A, Cabria M, Giovanella L. Thyroid nodules with indeterminate cytology: prospective comparison between 18F-FDG-PET/CT, multiparametric neck ultrasonography, 99mTc-MIBI scintigraphy and histology. European journal of endocrinology / European Federation of Endocrine Societies 2016;

20. Munoz Perez N, Villar del Moral JM, Muros Fuentes MA, Lopez de la Torre M, Arcelus Martinez JI, Becerra Massare P, Esteva Martinez D, Canadas Garre M, Coll Del Rey E, Bueno Larano P, Ferron Orihuela JA. Could 18F-FDG-PET/CT avoid unnecessary thyroidectomies in patients with cytological diagnosis of follicular neoplasm? Langenbeck's archives of surgery / Deutsche Gesellschaft fur Chirurgie 2013; 398:709-716

21. Nikiforov YE, Carty SE, Chiosea SI, Coyne C, Duvvuri U, Ferris RL, Gooding WE, LeBeau SO, Ohori NP, Seethala RR, Tublin ME, Yip L, Nikiforova MN. Impact of the Multi-Gene ThyroSeq Next-Generation Sequencing Assay on Cancer Diagnosis in Thyroid Nodules with Atypia of Undetermined Significance/Follicular Lesion of Undetermined Significance Cytology. Thyroid 2015; 25:1217-1223

22. Labourier E, Shifrin A, Busseniers AE, Lupo MA, Manganelli ML, Andruss B, Wylie D, Beaudenon-Huibregtse S. Molecular Testing for miRNA, mRNA, and DNA on Fine-Needle Aspiration Improves the Preoperative Diagnosis of Thyroid Nodules With Indeterminate Cytology. The Journal of clinical endocrinology and metabolism 2015; 100:2743-2750

23. Haugen BR, Alexander EK, Bible KC, Doherty GM, Mandel SJ, Nikiforov YE, Pacini F, Randolph GW, Sawka AM, Schlumberger M, Schuff KG, Sherman SI, Sosa JA, Steward DL, Tuttle RM, Wartofsky L. 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; 26:1-133

24. Kimura ET, Nikiforova MN, Zhu Z, Knauf JA, Nikiforov YE, Fagin JA. High prevalence of BRAF mutations in thyroid cancer: genetic evidence for constitutive activation of the RET/PTC-RAS-BRAF signaling pathway in papillary thyroid carcinoma. Cancer research 2003; 63:1454-1457

25. Adeniran AJ, Zhu Z, Gandhi M, Steward DL, Fidler JP, Giordano TJ, Biddinger PW, Nikiforov YE. Correlation between genetic alterations and microscopic features, clinical manifestations, and prognostic characteristics of thyroid papillary carcinomas. The American journal of surgical pathology 2006; 30:216-222

26. Malumbres M, Barbacid M. RAS oncogenes: the first 30 years. Nat Rev Cancer 2003; 3:459-465

27. Davies H, Bignell GR, Cox C, Stephens P, Edkins S, Clegg S, Teague J, Woffendin H, Garnett MJ, Bottomley W, Davis N, Dicks E, Ewing R, Floyd Y, Gray K, Hall S, Hawes R, Hughes J, Kosmidou V, Menzies A, Mould C, Parker A, Stevens C, Watt S, Hooper S, Wilson R, Jayatilake H, Gusterson BA, Cooper C, Shipley J, Hargrave D, Pritchard-Jones K, Maitland N, Chenevix-Trench G, Riggins GJ, Bigner DD, Palmieri G, Cossu A, Flanagan A, Nicholson A, Ho JW, Leung SY, Yuen ST, Weber BL, Seigler HF, Darrow TL, Paterson H, Marais R, Marshall CJ, Wooster R, Stratton MR, Futreal PA. Mutations of the BRAF gene in human cancer. Nature 2002; 417:949-954

28. Xing M. BRAF mutation in papillary thyroid cancer: pathogenic role, molecular bases, and clinical implications. Endocrine reviews 2007; 28:742-762

29. Rossi M, Buratto M, Tagliati F, Rossi R, Lupo S, Trasforini G, Lanza G, Franceschetti P, Bruni S, Degli Uberti E, Zatelli MC. Relevance of BRAF(V600E) mutation testing

versus RAS point mutations and RET/PTC rearrangements evaluation in the diagnosis of thyroid cancer. Thyroid 2015; 25:221-228

30. Carr R, Ustun B, Chhieng D, Schofield K, Theoharis C, Hammers L, Adeniran AJ. Radiologic and clinical predictors of malignancy in the follicular lesion of undetermined significance of the thyroid. Endocrine pathology 2013; 24:62-68

31. Nikiforov YE, Ohori NP, Hodak SP, Carty SE, LeBeau SO, Ferris RL, Yip L, Seethala RR, Tublin ME, Stang MT, Coyne C, Johnson JT, Stewart AF, Nikiforova MN. Impact of mutational testing on the diagnosis and management of patients with cytologically indeterminate thyroid nodules: a prospective analysis of 1056 FNA samples. The Journal of clinical endocrinology and metabolism 2011; 96:3390-3397

32. Nikiforova MN, Kimura ET, Gandhi M, Biddinger PW, Knauf JA, Basolo F, Zhu Z, Giannini R, Salvatore G, Fusco A, Santoro M, Fagin JA, Nikiforov YE. BRAF mutations in thyroid tumors are restricted to papillary carcinomas and anaplastic or poorly differentiated carcinomas arising from papillary carcinomas. The Journal of clinical endocrinology and metabolism 2003; 88:5399-5404

33. Xing M. BRAF mutation in thyroid cancer. Endocrine-related cancer 2005; 12:245-262

34. Kim SK, Hwang TS, Yoo YB, Han HS, Kim DL, Song KH, Lim SD, Kim WS, Paik NS. Surgical results of thyroid nodules according to a management guideline based on the BRAF(V600E) mutation status. The Journal of clinical endocrinology and metabolism 2011; 96:658-664

35. Rossi M, Buratto M, Bruni S, Filieri C, Tagliati F, Trasforini G, Rossi R, Beccati MD, Degli Uberti EC, Zatelli MC. Role of ultrasonographic/clinical profile, cytology, and BRAF V600E mutation evaluation in thyroid nodule screening for malignancy: a

prospective study. The Journal of clinical endocrinology and metabolism 2012; 97:2354-2361

36. Adeniran AJ, Hui P, Chhieng DC, Prasad ML, Schofield K, Theoharis C. BRAF mutation testing of thyroid fine-needle aspiration specimens enhances the predictability of malignancy in thyroid follicular lesions of undetermined significance. Acta cytologica 2011; 55:570-575

37. Ohori NP, Wolfe J, Hodak SP, LeBeau SO, Yip L, Carty SE, Duvvuri U, Schoedel KE, Nikiforova MN, Nikiforov YE. "Colloid-rich" follicular neoplasm/suspicious for follicular neoplasm thyroid fine-needle aspiration specimens: cytologic, histologic, and molecular basis for considering an alternate view. Cancer cytopathology 2013; 121:718-728

38. Eszlinger M, Piana S, Moll A, Bosenberg E, Bisagni A, Ciarrocchi A, Ragazzi M, Paschke R. Molecular testing of thyroid fine-needle aspirations improves presurgical diagnosis and supports the histologic identification of minimally invasive follicular thyroid carcinomas. Thyroid 2015; 25:401-409

39. Hwang TS, Kim WY, Han HS, Lim SD, Kim WS, Yoo YB, Park KS, Oh SY, Kim SK, Yang JH. Preoperative RAS mutational analysis is of great value in predicting follicular variant of papillary thyroid carcinoma. BioMed research international 2015; 2015:697068

40. Zhang Q, Liu SZ, Zhang Q, Guan YX, Chen QJ, Zhu QY. Meta-Analyses of Association Between BRAF(V600E) Mutation and Clinicopathological Features of Papillary Thyroid Carcinoma. Cellular physiology and biochemistry : international journal of experimental cellular physiology, biochemistry, and pharmacology 2016; 38:763-776

41. Yeo MK, Liang ZL, Oh T, Moon Y, An S, Kim MK, Kim KS, Shong M, Kim JM, Jo YS. Pyrosequencing cut-off value identifying BRAFV600E mutation in fine needle aspiration samples of thyroid nodules. Clinical endocrinology 2011; 75:555-560

42. Pelizzo MR, Boschin IM, Barollo S, Pennelli G, Toniato A, Zambonin L, Vianello F, Piotto A, Ide EC, Pagetta C, Sorgato N, Torresan F, Girelli ME, Nacamulli D, Mantero F, Mian C. BRAF analysis by fine needle aspiration biopsy of thyroid nodules improves preoperative identification of papillary thyroid carcinoma and represents a prognostic factor. A mono-institutional experience. Clinical chemistry and laboratory medicine : CCLM / FESCC 2011; 49:325-329

43. Xing M, Tufano RP, Tufaro AP, Basaria S, Ewertz M, Rosenbaum E, Byrne PJ, Wang J, Sidransky D, Ladenson PW. Detection of BRAF mutation on fine needle aspiration biopsy specimens: a new diagnostic tool for papillary thyroid cancer. The Journal of clinical endocrinology and metabolism 2004; 89:2867-2872

44. Sapio MR, Guerra A, Posca D, Limone PP, Deandrea M, Motta M, Troncone G, Caleo A, Vallefuoco P, Rossi G, Fenzi G, Vitale M. Combined analysis of galectin-3 and BRAFV600E improves the accuracy of fine-needle aspiration biopsy with cytological findings suspicious for papillary thyroid carcinoma. Endocrine-related cancer 2007; 14:1089-1097

45. Sapio MR, Posca D, Raggioli A, Guerra A, Marotta V, Deandrea M, Motta M, Limone PP, Troncone G, Caleo A, Rossi G, Fenzi G, Vitale M. Detection of RET/PTC, TRK and BRAF mutations in preoperative diagnosis of thyroid nodules with indeterminate cytological findings. Clinical endocrinology 2007; 66:678-683

46. Kim SK, Kim DL, Han HS, Kim WS, Kim SJ, Moon WJ, Oh SY, Hwang TS. Pyrosequencing analysis for detection of a BRAFV600E mutation in an FNAB

specimen of thyroid nodules. Diagnostic molecular pathology : the American journal of surgical pathology, part B 2008; 17:118-125

47. Jo YS, Huang S, Kim YJ, Lee IS, Kim SS, Kim JR, Oh T, Moon Y, An S, Ro HK, Kim JM, Shong M. Diagnostic value of pyrosequencing for the BRAF V600E mutation in ultrasound-guided fine-needle aspiration biopsy samples of thyroid incidentalomas. Clinical endocrinology 2009; 70:139-144

48. Marchetti I, Lessi F, Mazzanti CM, Bertacca G, Elisei R, Coscio GD, Pinchera A, Bevilacqua G. A morpho-molecular diagnosis of papillary thyroid carcinoma: BRAF V600E detection as an important tool in preoperative evaluation of fine-needle aspirates. Thyroid 2009; 19:837-842

49. Cantara S, Capezzone M, Marchisotta S, Capuano S, Busonero G, Toti P, Di Santo A, Caruso G, Carli AF, Brilli L, Montanaro A, Pacini F. Impact of proto-oncogene mutation detection in cytological specimens from thyroid nodules improves the diagnostic accuracy of cytology. The Journal of clinical endocrinology and metabolism 2010; 95:1365-1369

50. Moses W, Weng J, Sansano I, Peng M, Khanafshar E, Ljung BM, Duh QY, Clark OH, Kebebew E. Molecular testing for somatic mutations improves the accuracy of thyroid fine-needle aspiration biopsy. World journal of surgery 2010; 34:2589-2594

51. Adeniran AJ, Theoharis C, Hui P, Prasad ML, Hammers L, Carling T, Udelsman R, Chhieng DC. Reflex BRAF testing in thyroid fine-needle aspiration biopsy with equivocal and positive interpretation: a prospective study. Thyroid 2011; 21:717-723

52. Patel A, Klubo-Gwiezdzinska J, Hoperia V, Larin A, Jensen K, Bauer A, Vasko V. BRAF(V600E) mutation analysis from May-Grunwald Giemsa-stained cytological samples as an adjunct in identification of high-risk papillary thyroid carcinoma. Endocrine pathology 2011; 22:195-199

53. Canadas-Garre M, Becerra-Massare P, Lopez de la Torre-Casares M, Villar-del Moral J, Cespedes-Mas S, Vilchez-Joya R, Muros-de Fuentes T, Garcia-Calvente C, Piedrola-Maroto G, Lopez-Nevot MA, Montes-Ramirez R, Llamas-Elvira JM. Reduction of false-negative papillary thyroid carcinomas by the routine analysis of BRAF(T1799A) mutation on fine-needle aspiration biopsy specimens: a prospective study of 814 thyroid FNAB patients. Annals of surgery 2012; 255:986-992

54. Kang G, Cho EY, Shin JH, Chung JH, Kim JW, Oh YL. Role of BRAFV600E mutation analysis and second cytologic review of fine-needle aspiration for evaluating thyroid nodule. Cancer cytopathology 2012; 120:44-51

55. Mancini I, Pinzani P, Pupilli C, Petrone L, De Feo ML, Bencini L, Pazzagli M, Forti G, Orlando C. A high-resolution melting protocol for rapid and accurate differential diagnosis of thyroid nodules. The Journal of molecular diagnostics : JMD 2012; 14:501-509

56. Tonacchera M, Agretti P, Rago T, De Marco G, Niccolai F, Molinaro A, Scutari M, Candelieri A, Conforti D, Musmanno R, Di Coscio G, Basolo F, Iacconi P, Miccoli P, Pinchera A, Vitti P. Genetic markers to discriminate benign and malignant thyroid nodules with undetermined cytology in an area of borderline iodine deficiency. Journal of endocrinological investigation 2012; 35:754-759

57. Kloos RT, Reynolds JD, Walsh PS, Wilde JI, Tom EY, Pagan M, Barbacioru C, Chudova DI, Wong M, Friedman L, LiVolsi VA, Rosai J, Lanman RB, Kennedy GC. Does addition of BRAF V600E mutation testing modify sensitivity or specificity of the Afirma Gene Expression Classifier in cytologically indeterminate thyroid nodules? The Journal of clinical endocrinology and metabolism 2013; 98:E761-768

58. Agretti P, Niccolai F, Rago T, De Marco G, Molinaro A, Scutari M, Di Cosmo C, Di Coscio G, Vitale M, Maccheroni M, Vitti P, Tonacchera M. BRAF mutation analysis

in thyroid nodules with indeterminate cytology: our experience on surgical management of patients with thyroid nodules from an area of borderline iodine deficiency. Journal of endocrinological investigation 2014; 37:1009-1014

59. Beaudenon-Huibregtse S, Alexander EK, Guttler RB, Hershman JM, Babu V, Blevins TC, Moore P, Andruss B, Labourier E. Centralized molecular testing for oncogenic gene mutations complements the local cytopathologic diagnosis of thyroid nodules. Thyroid 2014; 24:1479-1487

60. Danilovic DL, Lima EU, Domingues RB, Brandao LG, Hoff AO, Marui S. Pre-operative role of BRAF in the guidance of the surgical approach and prognosis of differentiated thyroid carcinoma. European journal of endocrinology / European Federation of Endocrine Societies 2014; 170:619-625

61. Eszlinger M, Krogdahl A, Munz S, Rehfeld C, Precht Jensen EM, Ferraz C, Bosenberg E, Drieschner N, Scholz M, Hegedus L, Paschke R. Impact of molecular screening for point mutations and rearrangements in routine air-dried fine-needle aspiration samples of thyroid nodules. Thyroid 2014; 24:305-313

62. Eszlinger M, Neustadt M, Ruschenburg I, Neumann A, Franzius C, Adam S, Bacher K, Hach A, Hammoser R, Langvogt C, Molwitz T, Paschke R. [Fine-needle aspiration cytology of thyroid nodules: molecular diagnostics in a routine diagnostic setting]. Deutsche medizinische Wochenschrift (1946) 2014; 139:476-480

63. Johnson SJ, Hardy SA, Roberts C, Bourn D, Mallick U, Perros P. Pilot of BRAF mutation analysis in indeterminate, suspicious and malignant thyroid FNA cytology. Cytopathology 2014; 25:146-154

64. Liu S, Gao A, Zhang B, Zhang Z, Zhao Y, Chen P, Ji M, Hou P, Shi B. Assessment of molecular testing in fine-needle aspiration biopsy samples: an experience in a Chinese population. Experimental and molecular pathology 2014; 97:292-297

65. Poller DN, Glaysher S, Agrawal A, Caldera S, Kim D, Yiangou C. BRAF V600 co-testing in thyroid FNA cytology: short-term experience in a large cancer centre in the UK. Journal of clinical pathology 2014; 67:684-689

66. Seo JY, Kim EK, Kwak JY. Additional BRAF mutation analysis may have additional diagnostic value in thyroid nodules with "suspicious for malignant" cytology alone even when the nodules do not show suspicious US features. Endocrine 2014; 47:283-289

67. Borrelli N, Ugolini C, Giannini R, Antonelli A, Giordano M, Sensi E, Torregrossa L, Fallahi P, Miccoli P, Basolo F. Role of gene expression profiling in defining indeterminate thyroid nodules in addition to BRAF analysis. Cancer cytopathology 2015;

68. Capelli L, Marfisi C, Puccetti M, Saragoni L, De Paola F, Zaccaroni A, Chiadini E, Gagliardi L, Ferretti G, Zoli W, Ulivi P. Role of BRAF molecular analysis in the management of papillary thyroid carcinoma: analysis of cytological and histological samples. Cytopathology 2015; 26:297-302

69. Gill MS, Nayan S, Kocovski L, Cutz JC, Archibald SD, Jackson BS, Young JE, Gupta MK. Local molecular analysis of indeterminate thyroid nodules. Journal of otolaryngology - head & neck surgery = Le Journal d'oto-rhino-laryngologie et de chirurgie cervico-faciale 2015; 44:52

70. Giovanella L, Campenni A, Treglia G, Verburg FA, Trimboli P, Ceriani L, Bongiovanni M. Molecular imaging with Tc-MIBI and molecular testing for mutations in differentiating benign from malignant follicular neoplasm: a prospective comparison. European journal of nuclear medicine and molecular imaging 2015;

71. Le Mercier M, D'Haene N, De Neve N, Blanchard O, Degand C, Rorive S, Salmon I. Next-generation sequencing improves the diagnosis of thyroid FNA specimens with indeterminate cytology. Histopathology 2015; 66:215-224

72. Marino M, Monzani ML, Brigante G, Cioni K, Madeo B, Santi D, Maiorana A, Bettelli S, Moriondo V, Pignatti E, Bonacini L, Carani C, Rochira V, Simoni M. High-Resolution Melting Is a Sensitive, Cost-Effective, Time-Saving Technique for BRAF V600E Detection in Thyroid FNAB Washing Liquid: A Prospective Cohort Study. European Thyroid Journal 2015; 4:73-81

73. Park KS, Oh YL, Ki CS, Kim JW. Evaluation of the Real-Q BRAF V600E Detection Assay in Fine-Needle Aspiration Samples of Thyroid Nodules. The Journal of molecular diagnostics : JMD 2015; 17:431-437

74. Beisa A, Beisa V, Stoskus M, Ostaneviciute E, Griskevicius L, Strupas K. The value of the repeated examination of BRAF V600E mutation status in diagnostics of papillary thyroid cancer. Endokrynologia Polska 2016; 67:35-40

75. Kowalska A, Kowalik A, Palyga I, Walczyk A, Gasior-Perczak D, Kopczynski J, Lizis-Kolus K, Szyska-Skrobot D, Hurej S, Radowicz-Chil A, Chodurska R, Wypiorkiewicz E, Chlopek M, Nowak E, Niemyska K, Gozdz S. The usefulness of determining the presence of BRAF V600E mutation in fine-needle aspiration cytology in indeterminate cytological results. Endokrynologia Polska 2016; 67:41-47

76. Valderrabano P, Leon ME, Centeno BA, Otto KJ, Khazai L, McCaffrey JC, Russell JS, McIver B. Institutional prevalence of malignancy of indeterminate thyroid cytology is necessary but insufficient to accurately interpret molecular marker tests. European journal of endocrinology / European Federation of Endocrine Societies 2016;

77. Kim SW, Lee JI, Kim JW, Ki CS, Oh YL, Choi YL, Shin JH, Kim HK, Jang HW, Chung JH. BRAFV600E mutation analysis in fine-needle aspiration cytology specimens for evaluation of thyroid nodule: a large series in a BRAFV600E-prevalent population. The Journal of clinical endocrinology and metabolism 2010; 95:3693-3700

78. Koh J, Choi JR, Han KH, Kim EK, Yoon JH, Moon HJ, Kwak JY. Proper indication of BRAF(V600E) mutation testing in fine-needle aspirates of thyroid nodules. PloS one 2013; 8:e64505

79. Leslie C, Grieu-Iacopetta F, Richter A, Platten M, Murray J, Frost FA, Amanuel B, Kumarasinghe MP. BRAF p.Val600Glu (V600E) mutation detection in thyroid fine needle aspiration cell block samples: a feasibility study. Pathology 2015; 47:432-438

80. Nikiforov YE, Carty SE, Chiosea SI, Coyne C, Duvvuri U, Ferris RL, Gooding WE, Hodak SP, LeBeau SO, Ohori NP, Seethala RR, Tublin ME, Yip L, Nikiforova MN. Highly accurate diagnosis of cancer in thyroid nodules with follicular neoplasm/suspicious for a follicular neoplasm cytology by ThyroSeq v2 next-generation sequencing assay. Cancer 2014; 120:3627-3634

81. Jeong SH, Hong HS, Lee EH, Cha JG, Park JS, Kwak JJ. Outcome of thyroid nodules characterized as atypia of undetermined significance or follicular lesion of undetermined significance and correlation with Ultrasound features and BRAF(V600E) mutation analysis. AJR American journal of roentgenology 2013; 201:W854-860

82. Hyeon J, Ahn S, Shin JH, Oh YL. The prediction of malignant risk in the category "atypia of undetermined significance/follicular lesion of undetermined significance" of the Bethesda System for Reporting Thyroid Cytopathology using subcategorization and BRAF mutation results. Cancer cytopathology 2014; 122:368-376

83. Park HJ, Moon JH, Yom CK, Kim KH, Choi JY, Choi SI, Ahn SH, Jeong WJ, Lee WW, Park SY. Thyroid "atypia of undetermined significance" with nuclear atypia has high rates of malignancy and BRAF mutation. Cancer cytopathology 2014; 122:512-520

84. Lee ST, Kim SW, Ki CS, Jang JH, Shin JH, Oh YL, Kim JW, Chung JH. Clinical implication of highly sensitive detection of the BRAF V600E mutation in fine-needle aspirations of thyroid nodules: a comparative analysis of three molecular assays in 4585 consecutive cases in a BRAF V600E mutation-prevalent area. The Journal of clinical endocrinology and metabolism 2012; 97:2299-2306

85. Ohori NP, Singhal R, Nikiforova MN, Yip L, Schoedel KE, Coyne C, McCoy KL, LeBeau SO, Hodak SP, Carty SE, Nikiforov YE. BRAF mutation detection in indeterminate thyroid cytology specimens: underlying cytologic, molecular, and pathologic characteristics of papillary thyroid carcinoma. Cancer cytopathology 2013; 121:197-205

86. Afkhami M, Karunamurthy A, Chiosea S, Nikiforova MN, Seethala R, Nikiforov YE, Coyne C. Histopathologic and Clinical Characterization of Thyroid Tumors Carrying the BRAF(K601E) Mutation. Thyroid 2016; 26:242-247

87. Trimboli P, Treglia G, Condorelli E, Romanelli F, Crescenzi A, Bongiovanni M, Giovanella L. BRAF-mutated carcinomas among thyroid nodules with prior indeterminate FNA report: a systematic review and meta-analysis. Clinical endocrinology 2016; 84:315-320

88. Lee WS, Palmer BJ, Garcia A, Chong VE, Liu TH. BRAF mutation in papillary thyroid cancer: A cost-utility analysis of preoperative testing. Surgery 2014; 156:1569-1577; discussion 1577-1568

89. An JH, Song KH, Kim SK, Park KS, Yoo YB, Yang JH, Hwang TS, Kim DL. RAS mutations in indeterminate thyroid nodules are predictive of the follicular variant of papillary thyroid carcinoma. Clinical endocrinology 2015; 82:760-766

90. Yoon JH, Kwon HJ, Lee HS, Kim EK, Moon HJ, Kwak JY. RAS Mutations in AUS/FLUS Cytology: Does it Have an Additional Role in BRAFV600E Mutation-Negative Nodules? Medicine 2015; 94:e1084

91. Marshall CJ, Hall A, Weiss RA. A transforming gene present in human sarcoma cell lines. Nature 1982; 299:171-173

92. Shimizu K, Goldfarb M, Perucho M, Wigler M. Isolation and preliminary characterization of the transforming gene of a human neuroblastoma cell line. Proceedings of the National Academy of Sciences of the United States of America 1983; 80:383-387

93. Radkay LA, Chiosea SI, Seethala RR, Hodak SP, LeBeau SO, Yip L, McCoy KL, Carty SE, Schoedel KE, Nikiforova MN, Nikiforov YE, Ohori NP. Thyroid nodules with KRAS mutations are different from nodules with NRAS and HRAS mutations with regard to cytopathologic and histopathologic outcome characteristics. Cancer cytopathology 2014; 122:873-882

94. Kunavisarut T. Diagnostic biomarkers of differentiated thyroid cancer. Endocrine 2013; 44:616-622

95. Mitsutake N, Miyagishi M, Mitsutake S, Akeno N, Mesa C, Jr., Knauf JA, Zhang L, Taira K, Fagin JA. BRAF mediates RET/PTC-induced mitogen-activated protein kinase activation in thyroid cells: functional support for requirement of the RET/PTC-RAS-BRAF pathway in papillary thyroid carcinogenesis. Endocrinology 2006; 147:1014-1019

96. Guerra A, Sapio MR, Marotta V, Campanile E, Moretti MI, Deandrea M, Motta M, Limone PP, Fenzi G, Rossi G, Vitale M. Prevalence of RET/PTC rearrangement in benign and malignant thyroid nodules and its clinical application. Endocrine journal 2011; 58:31-38

97. Elisei R, Romei C, Vorontsova T, Cosci B, Veremeychik V, Kuchinskaya E, Basolo F, Demidchik EP, Miccoli P, Pinchera A, Pacini F. RET/PTC rearrangements in thyroid nodules: studies in irradiated and not irradiated, malignant and benign thyroid lesions in children and adults. The Journal of clinical endocrinology and metabolism 2001; 86:3211-3216

98. Su X, Li Z, He C, Chen W, Fu X, Yang A. Radiation exposure, young age, and female gender are associated with high prevalence of RET/PTC1 and RET/PTC3 in papillary thyroid cancer: a meta-analysis. Oncotarget 2016; 7:16716-16730

99. de Groot JW, Links TP, Plukker JT, Lips CJ, Hofstra RM. RET as a diagnostic and therapeutic target in sporadic and hereditary endocrine tumors. Endocrine reviews 2006; 27:535-560

100. Kroll TG, Sarraf P, Pecciarini L, Chen CJ, Mueller E, Spiegelman BM, Fletcher JA. PAX8-PPARgamma1 fusion oncogene in human thyroid carcinoma [corrected]. Science (New York, NY) 2000; 289:1357-1360

101. Zhang Y, Yu J, Lee C, Xu B, Sartor MA, Koenig RJ. Genomic binding and regulation of gene expression by the thyroid carcinoma-associated PAX8-PPARG fusion protein. Oncotarget 2015; 6:40418-40432

102. Pasca di Magliano M, Di Lauro R, Zannini M. Pax8 has a key role in thyroid cell differentiation. Proceedings of the National Academy of Sciences of the United States of America 2000; 97:13144-13149

103. Raman P, Koenig RJ. Pax-8-PPAR-gamma fusion protein in thyroid carcinoma. Nature reviews Endocrinology 2014; 10:616-623

104. Castro P, Rebocho AP, Soares RJ, Magalhaes J, Roque L, Trovisco V, Vieira de Castro I, Cardoso-de-Oliveira M, Fonseca E, Soares P, Sobrinho-Simoes M. PAX8-PPARgamma rearrangement is frequently detected in the follicular variant of papillary thyroid carcinoma. The Journal of clinical endocrinology and metabolism 2006; 91:213-220

105. Nikiforova MN, Lynch RA, Biddinger PW, Alexander EK, Dorn GW, 2nd, Tallini G, Kroll TG, Nikiforov YE. RAS point mutations and PAX8-PPAR gamma rearrangement in thyroid tumors: evidence for distinct molecular pathways in thyroid follicular carcinoma. The Journal of clinical endocrinology and metabolism 2003; 88:2318-2326

106. Jeong SH, Hong HS, Kwak JJ, Lee EH. Analysis of RAS mutation and PAX8/PPARgamma rearrangements in follicular-derived thyroid neoplasms in a Korean population: frequency and ultrasound findings. Journal of endocrinological investigation 2015; 38:849-857

107. Armstrong MJ, Yang H, Yip L, Ohori NP, McCoy KL, Stang MT, Hodak SP, Nikiforova MN, Carty SE, Nikiforov YE. PAX8/PPARgamma rearrangement in thyroid nodules predicts follicular-pattern carcinomas, in particular the encapsulated follicular variant of papillary carcinoma. Thyroid 2014; 24:1369-1374

108. French CA, Alexander EK, Cibas ES, Nose V, Laguette J, Faquin W, Garber J, Moore F, Jr., Fletcher JA, Larsen PR, Kroll TG. Genetic and biological subgroups of low-stage follicular thyroid cancer. The American journal of pathology 2003; 162:1053-1060

109. French CA, Fletcher JA, Cibas ES, Caulfield C, Allard P, Kroll TG. Molecular detection of PPAR gamma rearrangements and thyroid carcinoma in preoperative fine-needle aspiration biopsies. Endocrine pathology 2008; 19:166-174

110. Collins K, Mitchell JR. Telomerase in the human organism. Oncogene 2002; 21:564-579

111. Melo M, da Rocha AG, Vinagre J, Batista R, Peixoto J, Tavares C, Celestino R, Almeida A, Salgado C, Eloy C, Castro P, Prazeres H, Lima J, Amaro T, Lobo C, Martins MJ, Moura M, Cavaco B, Leite V, Cameselle-Teijeiro JM, Carrilho F, Carvalheiro M, Maximo V, Sobrinho-Simoes M, Soares P. TERT promoter mutations are a major indicator of poor outcome in differentiated thyroid carcinomas. The Journal of clinical endocrinology and metabolism 2014; 99:E754-765

112. Umbricht CB, Conrad GT, Clark DP, Westra WH, Smith DC, Zahurak M, Saji M, Smallridge RC, Goodman S, Zeiger MA. Human telomerase reverse transcriptase gene expression and the surgical management of suspicious thyroid tumors. Clinical cancer research : an official journal of the American Association for Cancer Research 2004; 10:5762-5768

113. Liou MJ, Chan EC, Lin JD, Liu FH, Chao TC. Human telomerase reverse transcriptase (hTERT) gene expression in FNA samples from thyroid neoplasms. Cancer letters 2003; 191:223-227

114. Greco A, Miranda C, Pierotti MA. Rearrangements of NTRK1 gene in papillary thyroid carcinoma. Molecular and cellular endocrinology 2010; 321:44-49

115. Berlingieri MT, Pierantoni GM, Giancotti V, Santoro M, Fusco A. Thyroid cell transformation requires the expression of the HMGA1 proteins. Oncogene 2002; 21:2971-2980

116. Lappinga PJ, Kip NS, Jin L, Lloyd RV, Henry MR, Zhang J, Nassar A. HMGA2 gene expression analysis performed on cytologic smears to distinguish benign from malignant thyroid nodules. Cancer cytopathology 2010; 118:287-297

117. Bartolazzi A, Gasbarri A, Papotti M, Bussolati G, Lucante T, Khan A, Inohara H, Marandino F, Orlandi F, Nardi F, Vecchione A, Tecce R, Larsson O, Thyroid Cancer Study G. Application of an immunodiagnostic method for improving preoperative diagnosis of nodular thyroid lesions. Lancet 2001; 357:1644-1650

118. Maruta J, Hashimoto H, Yamashita H, Yamashita H, Noguchi S. Immunostaining of galectin-3 and CD44v6 using fine-needle aspiration for distinguishing follicular carcinoma from adenoma. Diagnostic cytopathology 2004; 31:392-396

119. Matesa N, Samija I, Kusic Z. Accuracy of fine needle aspiration biopsy with and without the use of tumor markers in cytologically indeterminate thyroid lesions. Collegium antropologicum 2010; 34:53-57

120. Yip L, Farris C, Kabaker AS, Hodak SP, Nikiforova MN, McCoy KL, Stang MT, Smith KJ, Nikiforov YE, Carty SE. Cost impact of molecular testing for indeterminate thyroid nodule fine-needle aspiration biopsies. The Journal of clinical endocrinology and metabolism 2012; 97:1905-1912

121. Lee L, How J, Tabah RJ, Mitmaker EJ. Cost-effectiveness of molecular testing for thyroid nodules with atypia of undetermined significance cytology. The Journal of clinical endocrinology and metabolism 2014; 99:2674-2682

122. Nikiforova MN, Wald AI, Roy S, Durso MB, Nikiforov YE. Targeted next-generation sequencing panel (ThyroSeq) for detection of mutations in thyroid cancer. The Journal of clinical endocrinology and metabolism 2013; 98:E1852-1860

123. Nishino M. Molecular cytopathology for thyroid nodules: A review of methodology and test performance. Cancer cytopathology 2016; 124:14-27

124. ThermoFisher Scientific. Ion AmpliSeq™ Cancer Hotspot Panel v2. 2107; https://www.thermofisher.com/order/catalog/product/4475346. Accessed 25-03-2017.

125. Chudova D, Wilde JI, Wang ET, Wang H, Rabbee N, Egidio CM, Reynolds J, Tom E, Pagan M, Rigl CT, Friedman L, Wang CC, Lanman RB, Zeiger M, Kebebew E, Rosai J, Fellegara G, LiVolsi VA, Kennedy GC. Molecular classification of thyroid nodules using high-dimensionality genomic data. The Journal of clinical endocrinology and metabolism 2010; 95:5296-5304

126. Alexander EK, Kennedy GC, Baloch ZW, Cibas ES, Chudova D, Diggans J, Friedman L, Kloos RT, LiVolsi VA, Mandel SJ, Raab SS, Rosai J, Steward DL, Walsh PS, Wilde JI, Zeiger MA, Lanman RB, Haugen BR. Preoperative diagnosis of benign thyroid nodules with indeterminate cytology. The New England journal of medicine 2012; 367:705-715

127. Alexander EK, Schorr M, Klopper J, Kim C, Sipos J, Nabhan F, Parker C, Steward DL, Mandel SJ, Haugen BR. Multicenter clinical experience with the Afirma gene expression classifier. The Journal of clinical endocrinology and metabolism 2014; 99:119-125

128. Duick DS, Klopper JP, Diggans JC, Friedman L, Kennedy GC, Lanman RB, McIver B. The impact of benign gene expression classifier test results on the endocrinologist-patient decision to operate on patients with thyroid nodules with indeterminate fine-needle aspiration cytopathology. Thyroid 2012; 22:996-1001

129. Witt RL. Outcome of thyroid gene expression classifier testing in clinical practice. Laryngoscope 2016; 126:524-527

130. McIver B, Castro MR, Morris JC, Bernet V, Smallridge R, Henry M, Kosok L, Reddi H. An independent study of a gene expression classifier (Afirma) in the evaluation of

cytologically indeterminate thyroid nodules. The Journal of clinical endocrinology and metabolism 2014; 99:4069-4077

131. Yang SE, Sullivan PS, Zhang J, Govind R, Levin MR, Rao JY, Moatamed NA. Has Afirma gene expression classifier testing refined the indeterminate thyroid category in cytology? Cancer cytopathology 2016; 124:100-109

132. Brauner E, Holmes BJ, Krane JF, Nishino M, Zurakowski D, Hennessey JV, Faquin WC, Parangi S. Performance of the Afirma Gene Expression Classifier in Hurthle Cell Thyroid Nodules Differs from Other Indeterminate Thyroid Nodules. Thyroid 2015; 25:789-796

133. Celik B, Whetsell CR, Nassar A. Afirma GEC and thyroid lesions: An institutional experience. Diagnostic cytopathology 2015; 43:966-970

134. Marti JL, Avadhani V, Donatelli LA, Niyogi S, Wang B, Wong RJ, Shaha AR, Ghossein RA, Lin O, Morris LG, Ho AS. Wide Inter-institutional Variation in Performance of a Molecular Classifier for Indeterminate Thyroid Nodules. Annals of surgical oncology 2015;

135. Noureldine SI, Olson MT, Agrawal N, Prescott JD, Zeiger MA, Tufano RP. Effect of Gene Expression Classifier Molecular Testing on the Surgical Decision-Making Process for Patients With Thyroid Nodules. JAMA otolaryngology-- head & neck surgery 2015; 141:1082-1088

136. Aragon Han P, Olson MT, Fazeli R, Prescott JD, Pai SI, Schneider EB, Tufano RP, Zeiger MA. The impact of molecular testing on the surgical management of patients with thyroid nodules. Annals of surgical oncology 2014; 21:1862-1869

137. Wu JX, Lam R, Levin M, Rao J, Sullivan PS, Yeh MW. Effect of malignancy rates on cost-effectiveness of routine gene expression classifier testing for indeterminate thyroid nodules. Surgery 2016; 159:118-129

138. Li H, Robinson KA, Anton B, Saldanha IJ, Ladenson PW. Cost-effectiveness of a novel molecular test for cytologically indeterminate thyroid nodules. The Journal of clinical endocrinology and metabolism 2011; 96:E1719-1726

139. Heinzel A, Muller D, Behrendt FF, Giovanella L, Mottaghy FM, Verburg FA. Thyroid nodules with indeterminate cytology: molecular imaging with (9)(9)mTc-methoxyisobutylisonitrile (MIBI) is more cost-effective than the Afirma gene expression classifier. European journal of nuclear medicine and molecular imaging 2014; 41:1497-1500

140. Labourier E. Utility and cost-effectiveness of molecular testing in thyroid nodules with indeterminate cytology. Clinical endocrinology 2016; 85:624-631

141. Macias AA, Eappen S, Malikin I, Goldfarb J, Kujawa S, Konowitz PM, Kamani D, Randolph GW. Successful intraoperative electrophysiologic monitoring of the recurrent laryngeal nerve, a multidisciplinary approach: The Massachusetts Eye and Ear Infirmary monitoring collaborative protocol with experience in over 3000 cases. Head & neck 2016; 38:1487-1494

142. Harrell RM, Bimston DN. Surgical utility of Afirma: effects of high cancer prevalence and oncocytic cell types in patients with indeterminate thyroid cytology. Endocrine practice : official journal of the American College of Endocrinology and the American Association of Clinical Endocrinologists 2014; 20:364-369

143. Santhanam P, Khthir R, Gress T, Elkadry A, Olajide O, Yaqub A, Driscoll H. Gene expression classifier for the diagnosis of indeterminate thyroid nodules: a meta-analysis. Medical oncology (Northwood, London, England) 2016; 33:14

144. Lastra RR, Pramick MR, Crammer CJ, LiVolsi VA, Baloch ZW. Implications of a suspicious afirma test result in thyroid fine-needle aspiration cytology: an institutional experience. Cancer cytopathology 2014; 122:737-744

145. He H, Jazdzewski K, Li W, Liyanarachchi S, Nagy R, Volinia S, Calin GA, Liu CG, Franssila K, Suster S, Kloos RT, Croce CM, de la Chapelle A. The role of microRNA genes in papillary thyroid carcinoma. Proceedings of the National Academy of Sciences of the United States of America 2005; 102:19075-19080

146. Pallante P, Visone R, Ferracin M, Ferraro A, Berlingieri MT, Troncone G, Chiappetta G, Liu CG, Santoro M, Negrini M, Croce CM, Fusco A. MicroRNA deregulation in human thyroid papillary carcinomas. Endocrine-related cancer 2006; 13:497-508

147. Nikiforova MN, Nikiforov YE. Molecular diagnostics and predictors in thyroid cancer. Thyroid 2009; 19:1351-1361

148. Nikiforova MN, Tseng GC, Steward D, Diorio D, Nikiforov YE. MicroRNA expression profiling of thyroid tumors: biological significance and diagnostic utility. The Journal of clinical endocrinology and metabolism 2008; 93:1600-1608

149. Agretti P, Ferrarini E, Rago T, Candelieri A, De Marco G, Dimida A, Niccolai F, Molinaro A, Di Coscio G, Pinchera A, Vitti P, Tonacchera M. MicroRNA expression profile helps to distinguish benign nodules from papillary thyroid carcinomas starting from cells of fine-needle aspiration. European journal of endocrinology / European Federation of Endocrine Societies 2012; 167:393-400

150. Aragon Han P, Weng CH, Khawaja HT, Nagarajan N, Schneider EB, Umbricht CB, Witwer KW, Zeiger MA. MicroRNA Expression and Association with Clinicopathologic Features in Papillary Thyroid Cancer: A Systematic Review. Thyroid 2015; 25:1322-1329

151. Li X, Abdel-Mageed AB, Mondal D, Kandil E. MicroRNA expression profiles in differentiated thyroid cancer, a review. International journal of clinical and experimental medicine 2013; 6:74-80

152. Dettmer M, Perren A, Moch H, Komminoth P, Nikiforov YE, Nikiforova MN. Comprehensive MicroRNA expression profiling identifies novel markers in follicular variant of papillary thyroid carcinoma. Thyroid 2013; 23:1383-1389

153. Dettmer M, Vogetseder A, Durso MB, Moch H, Komminoth P, Perren A, Nikiforov YE, Nikiforova MN. MicroRNA expression array identifies novel diagnostic markers for conventional and oncocytic follicular thyroid carcinomas. The Journal of clinical endocrinology and metabolism 2013; 98:E1-7

154. Shen R, Liyanarachchi S, Li W, Wakely PE, Jr., Saji M, Huang J, Nagy R, Farrell T, Ringel MD, de la Chapelle A, Kloos RT, He H. MicroRNA signature in thyroid fine needle aspiration cytology applied to "atypia of undetermined significance" cases. Thyroid 2012; 22:9-16

155. Weber F, Teresi RE, Broelsch CE, Frilling A, Eng C. A limited set of human MicroRNA is deregulated in follicular thyroid carcinoma. The Journal of clinical endocrinology and metabolism 2006; 91:3584-3591

156. Stokowy T, Wojtas B, Fujarewicz K, Jarzab B, Eszlinger M, Paschke R. miRNAs with the potential to distinguish follicular thyroid carcinomas from benign follicular thyroid tumors: results of a meta-analysis. Hormone and metabolic research = Hormon- und Stoffwechselforschung = Hormones et metabolisme 2014; 46:171-180

157. Zhang Y, Zhong Q, Chen X, Fang J, Huang Z. Diagnostic value of microRNAs in discriminating malignant thyroid nodules from benign ones on fine-needle aspiration samples. Tumour biology : the journal of the International Society for Oncodevelopmental Biology and Medicine 2014; 35:9343-9353

158. Wei WJ, Shen CT, Song HJ, Qiu ZL, Luo QY. MicroRNAs as a potential tool in the differential diagnosis of thyroid cancer: a systematic review and meta-analysis. Clinical endocrinology 2016; 84:127-133

159. Keutgen XM, Filicori F, Crowley MJ, Wang Y, Scognamiglio T, Hoda R, Buitrago D, Cooper D, Zeiger MA, Zarnegar R, Elemento O, Fahey TJ, 3rd. A panel of four miRNAs accurately differentiates malignant from benign indeterminate thyroid lesions on fine needle aspiration. Clinical cancer research : an official journal of the American Association for Cancer Research 2012; 18:2032-2038

160. Vriens MR, Weng J, Suh I, Huynh N, Guerrero MA, Shen WT, Duh QY, Clark OH, Kebebew E. MicroRNA expression profiling is a potential diagnostic tool for thyroid cancer. Cancer 2012; 118:3426-3432

161. Rossi ED, Bizzarro T, Martini M, Capodimonti S, Sarti D, Cenci T, Bilotta M, Fadda G, Larocca LM. The evaluation of miRNAs on thyroid FNAC: the promising role of miR-375 in follicular neoplasms. Endocrine 2016;

162. Igci YZ, Ozkaya M, Korkmaz H, Bozgeyik E, Bayraktar R, Ulasli M, Erkilic S, Eraydin A, Oztuzcu S. Expression Levels of miR-30a-5p in Papillary Thyroid Carcinoma: A Comparison Between Serum and Fine Needle Aspiration Biopsy Samples. Genetic testing and molecular biomarkers 2015;

163. de Matos LL, Del Giglio AB, Matsubayashi CO, de Lima Farah M, Del Giglio A, da Silva Pinhal MA. Expression of CK-19, galectin-3 and HBME-1 in the differentiation of thyroid lesions: systematic review and diagnostic meta-analysis. Diagnostic pathology 2012; 7:97

164. Bartolazzi A, Orlandi F, Saggiorato E, Volante M, Arecco F, Rossetto R, Palestini N, Ghigo E, Papotti M, Bussolati G, Martegani MP, Pantellini F, Carpi A, Giovagnoli MR, Monti S, Toscano V, Sciacchitano S, Pennelli GM, Mian C, Pelizzo MR, Rugge M, Troncone G, Palombini L, Chiappetta G, Botti G, Vecchione A, Bellocco R. Galectin-3-expression analysis in the surgical selection of follicular thyroid nodules

with indeterminate fine-needle aspiration cytology: a prospective multicentre study. The Lancet Oncology 2008; 9:543-549

165. Zhang L, Krausz T, DeMay RM. A Pilot Study of Galectin-3, HBME-1, and p27 Triple Immunostaining Pattern for Diagnosis of Indeterminate Thyroid Nodules in Cytology With Correlation to Histology. Applied immunohistochemistry & molecular morphology : AIMM / official publication of the Society for Applied Immunohistochemistry 2015; 23:481-490

166. Collet JF, Hurbain I, Prengel C, Utzmann O, Scetbon F, Bernaudin JF, Fajac A. Galectin-3 immunodetection in follicular thyroid neoplasms: a prospective study on fine-needle aspiration samples. British journal of cancer 2005; 93:1175-1181

167. Coli A, Bigotti G, Parente P, Federico F, Castri F, Massi G. Atypical thyroid nodules express both HBME-1 and Galectin-3, two phenotypic markers of papillary thyroid carcinoma. Journal of experimental & clinical cancer research : CR 2007; 26:221-227

168. Bartolazzi A, Bellotti C, Sciacchitano S. Methodology and technical requirements of the galectin-3 test for the preoperative characterization of thyroid nodules. Applied immunohistochemistry & molecular morphology : AIMM / official publication of the Society for Applied Immunohistochemistry 2012; 20:2-7

169. Rossi ED, Raffaelli M, Mule A, Miraglia A, Lombardi CP, Vecchio FM, Fadda G. Simultaneous immunohistochemical expression of HBME-1 and galectin-3 differentiates papillary carcinomas from hyperfunctioning lesions of the thyroid. Histopathology 2006; 48:795-800

170. Pennelli G, Mian C, Pelizzo MR, Naccamulli D, Piotto A, Girelli ME, Mescoli C, Rugge M. Galectin-3 cytotest in thyroid follicular neoplasia: a prospective, monoinstitutional study. Acta cytologica 2009; 53:533-539

171. Ito Y, Yoshida H, Tomoda C, Miya A, Kobayashi K, Matsuzuka F, Yasuoka H, Kakudo K, Inohara H, Kuma K, Miyauchi A. Galectin-3 expression in follicular tumours: an immunohistochemical study of its use as a marker of follicular carcinoma. Pathology 2005; 37:296-298

172. Oestreicher-Kedem Y, Halpern M, Roizman P, Hardy B, Sulkes J, Feinmesser R, Stern Y. Diagnostic value of galectin-3 as a marker for malignancy in follicular patterned thyroid lesions. Head & neck 2004; 26:960-966

173. Torregrossa L, Faviana P, Camacci T, Materazzi G, Berti P, Minuto M, Elisei R, Vitti P, Miccoli P, Basolo F. Galectin-3 is highly expressed in nonencapsulated papillary thyroid carcinoma but weakly expressed in encapsulated type; comparison with Hector Battifora mesothelial cell 1 immunoreactivity. Human pathology 2007; 38:1482-1488

174. Saggiorato E, De Pompa R, Volante M, Cappia S, Arecco F, Dei Tos AP, Orlandi F, Papotti M. Characterization of thyroid 'follicular neoplasms' in fine-needle aspiration cytological specimens using a panel of immunohistochemical markers: a proposal for clinical application. Endocrine-related cancer 2005; 12:305-317

175. Aiad HA, Kandil MA, Asaad NY, El-Kased AM, El-Goday SF. Galectin-3 immunostaining in cytological and histopathological diagnosis of thyroid lesions. Journal of the Egyptian National Cancer Institute 2008; 20:36-46

176. Asioli S, Maletta F, Pacchioni D, Lupo R, Bussolati G. Cytological detection of papillary thyroid carcinomas by nuclear membrane decoration with emerin staining. Virchows Archiv : an international journal of pathology 2010; 457:43-51

177. Raggio E, Camandona M, Solerio D, Martino P, Franchello A, Orlandi F, Gasparri G. The diagnostic accuracy of the immunocytochemical markers in the pre-operative evaluation of follicular thyroid lesions. Journal of endocrinological investigation 2010; 33:378-381

178. Cheung CC, Ezzat S, Freeman JL, Rosen IB, Asa SL. Immunohistochemical diagnosis of papillary thyroid carcinoma. Modern pathology : an official journal of the United States and Canadian Academy of Pathology, Inc 2001; 14:338-342

179. Fadda G, Rossi ED, Raffaelli M, Pontecorvi A, Sioletic S, Morassi F, Lombardi CP, Zannoni GF, Rindi G. Follicular thyroid neoplasms can be classified as low- and high-risk according to HBME-1 and Galectin-3 expression on liquid-based fine-needle cytology. European journal of endocrinology / European Federation of Endocrine Societies 2011; 165:447-453

180. Lacoste-Collin L, d'Aure D, Berard E, Rouquette I, Delisle MB, Courtade-Saidi M. Improvement of the cytological diagnostic accuracy of follicular thyroid lesions by the use of the Ki-67 proliferative index in addition to cytokeratin-19 and HBME-1 immunomarkers: a study of 61 cases of liquid-based FNA cytology with histological controls. Cytopathology 2014; 25:160-169

181. Rossi ED, Raffaelli M, Minimo C, Mule A, Lombardi CP, Vecchio FM, Fadda G. Immunocytochemical evaluation of thyroid neoplasms on thin-layer smears from fine-needle aspiration biopsies. Cancer 2005; 105:87-95

182. Khurana KK, Truong LD, LiVolsi VA, Baloch ZW. Cytokeratin 19 immunolocalization in cell block preparation of thyroid aspirates. An adjunct to fine-needle aspiration diagnosis of papillary thyroid carcinoma. Archives of pathology & laboratory medicine 2003; 127:579-583

183. Baloch ZW, Abraham S, Roberts S, LiVolsi VA. Differential expression of cytokeratins in follicular variant of papillary carcinoma: an immunohistochemical study and its diagnostic utility. Human pathology 1999; 30:1166-1171

184. Kjellman P, Wallin G, Hoog A, Auer G, Larsson C, Zedenius J. MIB-1 index in thyroid tumors: a predictor of the clinical course in papillary thyroid carcinoma. Thyroid 2003; 13:371-380

185. Pisani T, Pantellini F, Centanni M, Vecchione A, Giovagnoli MR. Immunocytochemical expression of Ki67 and laminin in Hurthle cell adenomas and carcinomas. Anticancer research 2003; 23:3323-3326

186. Addati T, Achille G, Centrone M, Petroni S, Popescu O, Russo S, Grammatica L, Simone G. TROP-2 expression in papillary thyroid cancer: a preliminary cyto-histological study. Cytopathology 2015;

187. Coban I, Cakir A, Unal TD, Bassullu N, Karpuz V, Dogusoy GB, Alper M. Emerin expression in well differentiated epithelial lesions of thyroid: implications in papillary thyroid carcinoma diagnosis and predicting malignant behavior. Pathology oncology research : POR 2015; 21:357-366

188. Chandan VS, Faquin WC, Wilbur DC, Khurana KK. The role of immunolocalization of CD57 and GLUT-1 in cell blocks in fine-needle aspiration diagnosis of papillary thyroid carcinoma. Cancer 2006; 108:331-336

189. Troncone G, Volante M, Iaccarino A, Zeppa P, Cozzolino I, Malapelle U, Palmieri EA, Conzo G, Papotti M, Palombini L. Cyclin D1 and D3 overexpression predicts malignant behavior in thyroid fine-needle aspirates suspicious for Hurthle cell neoplasms. Cancer 2009; 117:522-529

190. Mills LJ, Poller DN, Yiangou C. Galectin-3 is not useful in thyroid FNA. Cytopathology 2005; 16:132-138

191. Brito JP, Gionfriddo MR, Al Nofal A, Boehmer KR, Leppin AL, Reading C, Callstrom M, Elraiyah TA, Prokop LJ, Stan MN, Murad MH, Morris JC, Montori VM. The accuracy of thyroid nodule ultrasound to predict thyroid cancer: systematic

review and meta-analysis. The Journal of clinical endocrinology and metabolism 2014; 99:1253-1263

192. Remonti LR, Kramer CK, Leitao CB, Pinto LC, Gross JL. Thyroid ultrasound features and risk of carcinoma: a systematic review and meta-analysis of observational studies. Thyroid 2015; 25:538-550

193. Jeh SK, Jung SL, Kim BS, Lee YS. Evaluating the degree of conformity of papillary carcinoma and follicular carcinoma to the reported ultrasonographic findings of malignant thyroid tumor. Korean journal of radiology 2007; 8:192-197

194. Lee SH, Baek JS, Lee JY, Lim JA, Cho SY, Lee TH, Ku YH, Kim HI, Kim MJ. Predictive factors of malignancy in thyroid nodules with a cytological diagnosis of follicular neoplasm. Endocrine pathology 2013; 24:177-183

195. Kim DS, Kim JH, Na DG, Park SH, Kim E, Chang KH, Sohn CH, Choi YH. Sonographic features of follicular variant papillary thyroid carcinomas in comparison with conventional papillary thyroid carcinomas. Journal of ultrasound in medicine : official journal of the American Institute of Ultrasound in Medicine 2009; 28:1685-1692

196. Ozdemir D, Ersoy R, Cuhaci N, Arpaci D, Ersoy EP, Korukluoglu B, Guler G, Cakir B. Classical and follicular variant papillary thyroid carcinoma: comparison of clinical, ultrasonographical, cytological, and histopathological features in 444 patients. Endocrine pathology 2011; 22:58-65

197. Brophy C, Stewart J, O'Donovan N, McCarthy J, Murphy M, Sheahan P. Impact of Microcalcifications on Risk of Malignancy in Thyroid Nodules with Indeterminate or Benign Cytology. Otolaryngology--head and neck surgery : official journal of American Academy of Otolaryngology-Head and Neck Surgery 2016; 154:46-51

198. Gweon HM, Son EJ, Youk JH, Kim JA. Thyroid nodules with Bethesda system III cytology: can ultrasonography guide the next step? Annals of surgical oncology 2013; 20:3083-3088

199. Lee KH, Shin JH, Oh YL, Hahn SY. Atypia of undetermined significance in thyroid fine-needle aspiration cytology: prediction of malignancy by US and comparison of methods for further management. Annals of surgical oncology 2014; 21:2326-2331

200. Pompili G, Tresoldi S, Primolevo A, De Pasquale L, Di Leo G, Cornalba G. Management of thyroid follicular proliferation: an ultrasound-based malignancy score to opt for surgical or conservative treatment. Ultrasound in medicine & biology 2013; 39:1350-1355

201. Yoo WS, Choi HS, Cho SW, Moon JH, Kim KW, Park HJ, Park SY, Choi SI, Choi SH, Lim S, Yi KH, Park do J, Jang HC, Park YJ. The role of ultrasound findings in the management of thyroid nodules with atypia or follicular lesions of undetermined significance. Clinical endocrinology 2014; 80:735-742

202. Chng CL, Kurzawinski TR, Beale T. Value of sonographic features in predicting malignancy in thyroid nodules diagnosed as follicular neoplasm on cytology. Clinical endocrinology 2014;

203. Tutuncu Y, Berker D, Isik S, Akbaba G, Ozuguz U, Kucukler FK, Gocmen E, Yalcin Y, Aydin Y, Guler S. The frequency of malignancy and the relationship between malignancy and ultrasonographic features of thyroid nodules with indeterminate cytology. Endocrine 2014; 45:37-45

204. Yoon JH, Kwak JY, Kim EK, Moon HJ, Kim MJ, Kim JY, Koo HR, Kim MH. How to approach thyroid nodules with indeterminate cytology. Annals of surgical oncology 2010; 17:2147-2155

205. Khoncarly SM, Tamarkin SW, McHenry CR. Can ultrasound be used to predict malignancy in patients with a thyroid nodule and an indeterminate fine-needle aspiration biopsy? Surgery 2014; 156:967-970

206. Yoon JH, Lee HS, Kim EK, Moon HJ, Kwak JY. A nomogram for predicting malignancy in thyroid nodules diagnosed as atypia of undetermined significance/follicular lesions of undetermined significance on fine needle aspiration. Surgery 2014; 155:1006-1013

207. Dutta S, Thaha MA, Smith DM. Do sonographic and cytological features predict malignancy in cytologically indeterminate thyroid nodules? Annals of the Royal College of Surgeons of England 2011; 93:361-364

208. Parikh PP, Allan BJ, Lew JI. Surgeon-performed ultrasound predictors of malignancy in patients with Hurthle cell neoplasms of the thyroid. The Journal of surgical research 2013; 184:247-252

209. Batawil N, Alkordy T. Ultrasonographic features associated with malignancy in cytologically indeterminate thyroid nodules. European journal of surgical oncology : the journal of the European Society of Surgical Oncology and the British Association of Surgical Oncology 2014; 40:182-186

210. Garino F, Deandrea M, Motta M, Mormile A, Ragazzoni F, Palestini N, Freddi M, Gasparri G, Sgotto E, Pacchioni D, Limone PP. Diagnostic performance of elastography in cytologically indeterminate thyroid nodules. Endocrine 2015; 49:175-183

211. Macias CA, Arumugam D, Arlow RL, Eng OS, Lu SE, Javidian P, Davidov T, Trooskin SZ. A risk model to determine surgical treatment in patients with thyroid nodules with indeterminate cytology. Annals of surgical oncology 2015; 22:1527-1532

212. Turanli S, Pirhan Y, Ozcelik CK, Cetin A. Predictors of malignancy in patients with a thyroid nodule that contains Hurthle cells. Otolaryngology--head and neck surgery : official journal of American Academy of Otolaryngology-Head and Neck Surgery 2011; 144:514-517

213. Lippolis PV, Tognini S, Materazzi G, Polini A, Mancini R, Ambrosini CE, Dardano A, Basolo F, Seccia M, Miccoli P, Monzani F. Is elastography actually useful in the presurgical selection of thyroid nodules with indeterminate cytology? The Journal of clinical endocrinology and metabolism 2011; 96:E1826-1830

214. Kayilioglu SI, Dinc T, Sozen I, Senol K, Katar K, Karabeyoglu M, Tez M, Coskun F. Thyroid nodules with atypia or follicular lesions of undetermined significance (AUS/FLUS): analysis of variables associated with outcome. Asian Pacific journal of cancer prevention : APJCP 2014; 15:10307-10311

215. Cakir B, Ersoy R, Cuhaci FN, Aydin C, Polat B, Kilic M, Yazgan A. Elastosonographic strain index in thyroid nodules with atypia of undetermined significance. Journal of endocrinological investigation 2014; 37:127-133

216. Calo PG, Medas F, Santa Cruz R, Podda F, Erdas E, Pisano G, Nicolosi A. Follicular nodules (Thy3) of the thyroid: is total thyroidectomy the best option? BMC surgery 2014; 14:12

217. Cantisani V, Maceroni P, D'Andrea V, Patrizi G, Di Segni M, De Vito C, Grazhdani H, Isidori AM, Giannetta E, Redler A, Frattaroli F, Giacomelli L, Di Rocco G, Catalano C, D'Ambrosio F. Strain ratio ultrasound elastography increases the accuracy of colour-Doppler ultrasound in the evaluation of Thy-3 nodules. A bi-centre university experience. European radiology 2015;

218. Rago T, Di Coscio G, Basolo F, Scutari M, Elisei R, Berti P, Miccoli P, Romani R, Faviana P, Pinchera A, Vitti P. Combined clinical, thyroid ultrasound and cytological

features help to predict thyroid malignancy in follicular and Hupsilonrthle cell thyroid lesions: results from a series of 505 consecutive patients. Clinical endocrinology 2007; 66:13-20

219. Cantisani V, Ulisse S, Guaitoli E, De Vito C, Caruso R, Mocini R, D'Andrea V, Ascoli V, Antonaci A, Catalano C, Nardi F, Redler A, Ricci P, De Antoni E, Sorrenti S. Q-elastography in the presurgical diagnosis of thyroid nodules with indeterminate cytology. PloS one 2012; 7:e50725

220. Matthey-Gie ML, Walsh SM, O'Neill AC, Lowery A, Evoy D, Gibbons D, Prichard RS, Skehan S, McDermott EW. Ultrasound predictors of malignancy in indeterminate thyroid nodules. Irish journal of medical science 2014; 183:633-637

221. Mohey N, Hassan TA, Abdel-Baki S. Role of combined grey scale US and US tissue elastography in differentiating solid thyroid nodules. Egyptian Journal of Radiology and Nuclear Medicine 2013; 44:505-512

222. Rago T, Scutari M, Santini F, Loiacono V, Piaggi P, Di Coscio G, Basolo F, Berti P, Pinchera A, Vitti P. Real-time elastosonography: useful tool for refining the presurgical diagnosis in thyroid nodules with indeterminate or nondiagnostic cytology. The Journal of clinical endocrinology and metabolism 2010; 95:5274-5280

223. Lee KH, Shin JH, Ko ES, Hahn SY, Kim JS, Kim JH, Oh YL. Predictive factors of malignancy in patients with cytologically suspicious for Hurthle cell neoplasm of thyroid nodules. International journal of surgery (London, England) 2013; 11:898-902

224. Dogan L, Gulcelik MA, Karaman N, Baskan E, Kahraman YS, Aksel B, Ozaslan C. The contribution of clinical and radiological features to the diagnosis in AUS/FLUS and FN/SFN thyroid nodules. Acta Medica Mediterranea 2016; 32:151-155

225. Moon HG, Jung EJ, Park ST, Ha WS, Choi SK, Hong SC, Lee YJ, Joo YT, Jeong CY, Choi DS, Ryoo JW. Role of ultrasonography in predicting malignancy in patients with thyroid nodules. World journal of surgery 2007; 31:1410-1416

226. Alexander EK, Marqusee E, Orcutt J, Benson CB, Frates MC, Doubilet PM, Cibas ES, Atri A. Thyroid nodule shape and prediction of malignancy. Thyroid 2004; 14:953-958

227. Chin C, Franklin J, Sowerby L, Fung K, Yoo J. Stratification of intermediate-risk fine-needle aspiration biopsies. Journal of otolaryngology - head & neck surgery = Le Journal d'oto-rhino-laryngologie et de chirurgie cervico-faciale 2010; 39:393-396

228. Kim DW, Lee EJ, Jung SJ, Ryu JH, Kim YM. Role of sonographic diagnosis in managing Bethesda class III nodules. AJNR American journal of neuroradiology 2011; 32:2136-2141

229. Lee MJ, Hong SW, Chung WY, Kwak JY, Kim MJ, Kim EK. Cytological results of ultrasound-guided fine-needle aspiration cytology for thyroid nodules: emphasis on correlation with sonographic findings. Yonsei medical journal 2011; 52:838-844

230. Maia FF, Matos PS, Pavin EJ, Vassallo J, Zantut-Wittmann DE. Value of ultrasound and cytological classification system to predict the malignancy of thyroid nodules with indeterminate cytology. Endocrine pathology 2011; 22:66-73

231. Norlen O, Popadich A, Kruijff S, Gill AJ, Sarkis LM, Delbridge L, Sywak M, Sidhu S. Bethesda III thyroid nodules: the role of ultrasound in clinical decision making. Annals of surgical oncology 2014; 21:3528-3533

232. Rosario PW, Salles DS, Bessa B, Purisch S. Contribution of scintigraphy and ultrasonography to the prediction of malignancy in thyroid nodules with indeterminate cytology. Arquivos brasileiros de endocrinologia e metabologia 2010; 54:56-59

233. Rosario PW. Thyroid nodules with atypia or follicular lesions of undetermined significance (Bethesda Category III): importance of ultrasonography and cytological subcategory. Thyroid 2014; 24:1115-1120

234. Gulcelik NE, Gulcelik MA, Kuru B. Risk of malignancy in patients with follicular neoplasm: predictive value of clinical and ultrasonographic features. Archives of otolaryngology--head & neck surgery 2008; 134:1312-1315

235. Kamaya A, Lewis GH, Liu Y, Akatsu H, Kong C, Desser TS. Atypia of undetermined significance and follicular lesions of undetermined significance: sonographic assessment for prediction of the final diagnosis. Journal of ultrasound in medicine : official journal of the American Institute of Ultrasound in Medicine 2015; 34:767-774

236. Kwak JY, Han KH, Yoon JH, Moon HJ, Son EJ, Park SH, Jung HK, Choi JS, Kim BM, Kim EK. Thyroid imaging reporting and data system for US features of nodules: a step in establishing better stratification of cancer risk. Radiology 2011; 260:892-899

237. Horvath E, Majlis S, Rossi R, Franco C, Niedmann JP, Castro A, Dominguez M. An ultrasonogram reporting system for thyroid nodules stratifying cancer risk for clinical management. The Journal of clinical endocrinology and metabolism 2009; 94:1748-1751

238. Maia FF, Matos PS, Pavin EJ, Zantut-Wittmann DE. Thyroid imaging reporting and data system score combined with Bethesda system for malignancy risk stratification in thyroid nodules with indeterminate results on cytology. Clinical endocrinology 2015; 82:439-444

239. Park VY, Kim EK, Kwak JY, Yoon JH, Moon HJ. Malignancy risk and characteristics of thyroid nodules with two consecutive results of atypia of undetermined significance or follicular lesion of undetermined significance on cytology. European radiology 2015; 25:2601-2607

240. Russ G, Royer B, Bigorgne C, Rouxel A, Bienvenu-Perrard M, Leenhardt L. Prospective evaluation of thyroid imaging reporting and data system on 4550 nodules with and without elastography. European journal of endocrinology / European Federation of Endocrine Societies 2013; 168:649-655

241. Yoon JH, Kwon HJ, Kim EK, Moon HJ, Kwak JY. Subcategorization of atypia of undetermined significance/follicular lesion of undetermined significance (AUS/FLUS): a study applying Thyroid Imaging Reporting and Data System (TIRADS). Clinical endocrinology 2016;

242. Raber W, Kaserer K, Niederle B, Vierhapper H. Risk factors for malignancy of thyroid nodules initially identified as follicular neoplasia by fine-needle aspiration: results of a prospective study of one hundred twenty patients. Thyroid 2000; 10:709-712

243. Lyshchik A, Higashi T, Asato R, Tanaka S, Ito J, Mai JJ, Pellot-Barakat C, Insana MF, Brill AB, Saga T, Hiraoka M, Togashi K. Thyroid gland tumor diagnosis at US elastography. Radiology 2005; 237:202-211

244. Itoh A, Ueno E, Tohno E, Kamma H, Takahashi H, Shiina T, Yamakawa M, Matsumura T. Breast disease: clinical application of US elastography for diagnosis. Radiology 2006; 239:341-350

245. Rago T, Santini F, Scutari M, Pinchera A, Vitti P. Elastography: new developments in ultrasound for predicting malignancy in thyroid nodules. The Journal of clinical endocrinology and metabolism 2007; 92:2917-2922

246. Asteria C, Giovanardi A, Pizzocaro A, Cozzaglio L, Morabito A, Somalvico F, Zoppo A. US-elastography in the differential diagnosis of benign and malignant thyroid nodules. Thyroid 2008; 18:523-531

247. Veer V, Puttagunta S. The role of elastography in evaluating thyroid nodules: a literature review and meta-analysis. European archives of oto-rhino-laryngology : official journal of the European Federation of Oto-Rhino-Laryngological Societies (EUFOS) : affiliated with the German Society for Oto-Rhino-Laryngology - Head and Neck Surgery 2015; 272:1845-1855

248. Nell S, Kist JW, Debray TP, de Keizer B, van Oostenbrugge TJ, Borel Rinkes IH, Valk GD, Vriens MR. Qualitative elastography can replace thyroid nodule fine-needle aspiration in patients with soft thyroid nodules. A systematic review and meta-analysis. European journal of radiology 2015; 84:652-661

249. Abdelrahman SF, Ali FH, El-Sayed Khalil M, El Masry MR. Ultrasound elastography in the diagnostic evaluation of indeterminate thyroid nodule s. Egyptian Journal of Radiology and Nuclear Medicine 2015; 46:639-648

250. Trimboli P, Treglia G, Sadeghi R, Romanelli F, Giovanella L. Reliability of real-time elastography to diagnose thyroid nodules previously read at FNAC as indeterminate: a meta-analysis. Endocrine 2014;

251. Park SH, Kim SJ, Kim EK, Kim MJ, Son EJ, Kwak JY. Interobserver agreement in assessing the sonographic and elastographic features of malignant thyroid nodules. AJR American journal of roentgenology 2009; 193:W416-423

252. Sebag F, Vaillant-Lombard J, Berbis J, Griset V, Henry JF, Petit P, Oliver C. Shear wave elastography: a new ultrasound imaging mode for the differential diagnosis of benign and malignant thyroid nodules. The Journal of clinical endocrinology and metabolism 2010; 95:5281-5288

253. Kim DW. Computed tomography features of papillary thyroid carcinomas. Journal of computer assisted tomography 2014; 38:936-940

254. Lee C, Chalmers B, Treister D, Adhya S, Godwin B, Ji L, Groshen S, Grant E. Thyroid lesions visualized on CT: sonographic and pathologic correlation. Academic radiology 2015; 22:203-209

255. Treglia G, Caldarella C, Saggiorato E, Ceriani L, Orlandi F, Salvatori M, Giovanella L. Diagnostic performance of (99m)Tc-MIBI scan in predicting the malignancy of thyroid nodules: a meta-analysis. Endocrine 2013; 44:70-78

256. Gharib H, Papini E. Thyroid nodules: clinical importance, assessment, and treatment. Endocrinology and metabolism clinics of North America 2007; 36:707-735, vi

257. Lumachi F, Varotto L, Borsato S, Tregnaghi A, Zucchetta P, Marzola MC, Cecchin D, Bui F. Usefulness of 99mTc-pertechnetate scintigraphy and fine-needle aspiration cytology in patients with solitary thyroid nodules and thyroid cancer. Anticancer research 2004; 24:2531-2534

258. Saggiorato E, Angusti T, Rosas R, Martinese M, Finessi M, Arecco F, Trevisiol E, Bergero N, Puligheddu B, Volante M, Podio V, Papotti M, Orlandi F. 99mTc-MIBI Imaging in the presurgical characterization of thyroid follicular neoplasms: relationship to multidrug resistance protein expression. Journal of nuclear medicine : official publication, Society of Nuclear Medicine 2009; 50:1785-1793

259. Vattimo A, Bertelli P, Cintorino M, Burroni L, Volterrani D, Vella A. Identification of Hurthle cell tumor by single-injection, double-phase scintigraphy with technetium-99m-sestamibi. Journal of nuclear medicine : official publication, Society of Nuclear Medicine 1995; 36:778-782

260. Boi F, Lai ML, Deias C, Piga M, Serra A, Uccheddu A, Faa G, Mariotti S. The usefulness of 99mTc-SestaMIBI scan in the diagnostic evaluation of thyroid nodules with oncocytic cytology. European journal of endocrinology / European Federation of Endocrine Societies 2003; 149:493-498

261. Hindie E, Ugur O, Fuster D, O'Doherty M, Grassetto G, Urena P, Kettle A, Gulec SA, Pons F, Rubello D, Parathyroid Task Group of the E. 2009 EANM parathyroid guidelines. European journal of nuclear medicine and molecular imaging 2009; 36:1201-1216

262. Wale A, Miles KA, Young B, Zammit C, Williams A, Quin J, Dizdarevic S. Combined (99m)Tc-methoxyisobutylisonitrile scintigraphy and fine-needle aspiration cytology offers an accurate and potentially cost-effective investigative strategy for the assessment of solitary or dominant thyroid nodules. European journal of nuclear medicine and molecular imaging 2014; 41:105-115

263. Madorin CA, Owen R, Coakley B, Lowe H, Nam KH, Weber K, Kushnir L, Rios J, Genden E, Pawha PS, Inabnet WB, 3rd. Comparison of radiation exposure and cost between dynamic computed tomography and sestamibi scintigraphy for preoperative localization of parathyroid lesions. JAMA Surg 2013; 148:500-503

264. Heller M, Zanocco K, Zydowicz S, Elaraj D, Nayar R, Sturgeon C. Cost-effectiveness analysis of repeat fine-needle aspiration for thyroid biopsies read as atypia of undetermined significance. Surgery 2012; 152:423-430

265. Upadhyay M, Samal J, Kandpal M, Singh OV, Vivekanandan P. The Warburg effect: insights from the past decade. Pharmacol Ther 2013; 137:318-330

266. Mitchell JC, Grant F, Evenson AR, Parker JA, Hasselgren PO, Parangi S. Preoperative evaluation of thyroid nodules with 18FDG-PET/CT. Surgery 2005; 138:1166-1174; discussion 1174-1165

267. Vriens D, de Wilt JH, van der Wilt GJ, Netea-Maier RT, Oyen WJ, de Geus-Oei LF. The role of [18F]-2-fluoro-2-deoxy-d-glucose-positron emission tomography in thyroid nodules with indeterminate fine-needle aspiration biopsy: systematic review and meta-analysis of the literature. Cancer 2011; 117:4582-4594

268. Kresnik E, Gallowitsch HJ, Mikosch P, Stettner H, Igerc I, Gomez I, Kumnig G, Lind P. Fluorine-18-fluorodeoxyglucose positron emission tomography in the preoperative assessment of thyroid nodules in an endemic goiter area. Surgery 2003; 133:294-299

269. de Geus-Oei LF, Pieters GF, Bonenkamp JJ, Mudde AH, Bleeker-Rovers CP, Corstens FH, Oyen WJ. 18F-FDG PET reduces unnecessary hemithyroidectomies for thyroid nodules with inconclusive cytologic results. Journal of nuclear medicine : official publication, Society of Nuclear Medicine 2006; 47:770-775

270. Sebastianes FM, Cerci JJ, Zanoni PH, Soares J, Jr., Chibana LK, Tomimori EK, de Camargo RY, Izaki M, Giorgi MC, Eluf-Neto J, Meneghetti JC, Pereira MA. Role of 18F-fluorodeoxyglucose positron emission tomography in preoperative assessment of cytologically indeterminate thyroid nodules. The Journal of clinical endocrinology and metabolism 2007; 92:4485-4488

271. Traugott AL, Dehdashti F, Trinkaus K, Cohen M, Fialkowski E, Quayle F, Hussain H, Davila R, Ylagan L, Moley JF. Exclusion of malignancy in thyroid nodules with indeterminate fine-needle aspiration cytology after negative 18F-fluorodeoxyglucose positron emission tomography: interim analysis. World journal of surgery 2010; 34:1247-1253

272. Hales NW, Krempl GA, Medina JE. Is there a role for fluorodeoxyglucose positron emission tomography/computed tomography in cytologically indeterminate thyroid nodules? American journal of otolaryngology 2008; 29:113-118

273. Deandreis D, Al Ghuzlan A, Auperin A, Vielh P, Caillou B, Chami L, Lumbroso J, Travagli JP, Hartl D, Baudin E, Schlumberger M, Leboulleux S. Is (18)F-fluorodeoxyglucose-PET/CT useful for the presurgical characterization of thyroid nodules with indeterminate fine needle aspiration cytology? Thyroid 2012; 22:165-172

274. Smith RB, Robinson RA, Hoffman HT, Graham MM. Preoperative FDG-PET imaging to assess the malignant potential of follicular neoplasms of the thyroid. Otolaryngology--head and neck surgery : official journal of American Academy of Otolaryngology-Head and Neck Surgery 2008; 138:101-106

275. Valderrabano P, Montilla-Soler J, Mifsud M, Leon M, Centeno B, Khazai L, Padhya T, McCaffrey T, Russell J, McIver B, Otto K. Hypermetabolism on (18)F-Fluorodeoxyglucose Positron Emission Tomography Scan Does Not Influence the Interpretation of Thyroid Cytopathology, and Nodules with a SUVmax <2.5 Are Not at Increased Risk for Malignancy. Thyroid 2016; 26:1300-1307

276. Kim JM, Ryu JS, Kim TY, Kim WB, Kwon GY, Gong G, Moon DH, Kim SC, Hong SJ, Shong YK. 18F-fluorodeoxyglucose positron emission tomography does not predict malignancy in thyroid nodules cytologically diagnosed as follicular neoplasm. The Journal of clinical endocrinology and metabolism 2007; 92:1630-1634

277. Boellaard R, Delgado-Bolton R, Oyen WJ, Giammarile F, Tatsch K, Eschner W, Verzijlbergen FJ, Barrington SF, Pike LC, Weber WA, Stroobants S, Delbeke D, Donohoe KJ, Holbrook S, Graham MM, Testanera G, Hoekstra OS, Zijlstra J, Visser E, Hoekstra CJ, Pruim J, Willemsen A, Arends B, Kotzerke J, Bockisch A, Beyer T, Chiti A, Krause BJ. FDG PET/CT: EANM procedure guidelines for tumour imaging: version 2.0. European journal of nuclear medicine and molecular imaging 2015; 42:328-354

278. Vriens D, Visser EP, de Geus-Oei LF, Oyen WJ. Methodological considerations in quantification of oncological FDG PET studies. European journal of nuclear medicine and molecular imaging 2010; 37:1408-1425

279. Pathak KA, Klonisch T, Nason RW, Leslie WD. FDG-PET characteristics of Hurthle cell and follicular adenomas. Annals of nuclear medicine 2016; 30:506-509

280. Boellaard R, O'Doherty MJ, Weber WA, Mottaghy FM, Lonsdale MN, Stroobants SG, Oyen WJ, Kotzerke J, Hoekstra OS, Pruim J, Marsden PK, Tatsch K, Hoekstra CJ, Visser EP, Arends B, Verzijlbergen FJ, Zijlstra JM, Comans EF, Lammertsma AA, Paans AM, Willemsen AT, Beyer T, Bockisch A, Schaefer-Prokop C, Delbeke D, Baum RP, Chiti A, Krause BJ. FDG PET and PET/CT: EANM procedure guidelines for tumour PET imaging: version 1.0. European journal of nuclear medicine and molecular imaging 2010; 37:181-200

281. Treglia G, Taralli S, Salsano M, Muoio B, Sadeghi R, Giovanella L. Prevalence and malignancy risk of focal colorectal incidental uptake detected by (18)F-FDG-PET or PET/CT: a meta-analysis. Radiology and oncology 2014; 48:99-104

282. Nayan S, Ramakrishna J, Gupta MK. The Proportion of Malignancy in Incidental Thyroid Lesions on 18-FDG PET Study: A Systematic Review and Meta-analysis. Otolaryngology--head and neck surgery : official journal of American Academy of Otolaryngology-Head and Neck Surgery 2014; 151:190-200

283. Nakahira M, Saito N, Murata S, Sugasawa M, Shimamura Y, Morita K, Takajyo F, Omura G, Matsumura S. Quantitative diffusion-weighted magnetic resonance imaging as a powerful adjunct to fine needle aspiration cytology for assessment of thyroid nodules. American journal of otolaryngology 2012; 33:408-416

284. Brown AM, Nagala S, McLean MA, Lu Y, Scoffings D, Apte A, Gonen M, Stambuk HE, Shaha AR, Tuttle RM, Deasy JO, Priest AN, Jani P, Shukla-Dave A, Griffiths J. Multi-institutional validation of a novel textural analysis tool for preoperative stratification of suspected thyroid tumors on diffusion-weighted MRI. Magnetic resonance in medicine 2015;

285. Razek AA, Sadek AG, Kombar OR, Elmahdy TE, Nada N. Role of apparent diffusion coefficient values in differentiation between malignant and benign solitary thyroid nodules. AJNR American journal of neuroradiology 2008; 29:563-568

286. Chen X, Li WL, Zhang YL, Wu Q, Guo YM, Bai ZL. Meta-analysis of quantitative diffusion-weighted MR imaging in the differential diagnosis of breast lesions. BMC cancer 2010; 10:693

287. Wang J, Takashima S, Takayama F, Kawakami S, Saito A, Matsushita T, Momose M, Ishiyama T. Head and neck lesions: characterization with diffusion-weighted echo-planar MR imaging. Radiology 2001; 220:621-630

288. Bozgeyik Z, Coskun S, Dagli AF, Ozkan Y, Sahpaz F, Ogur E. Diffusion-weighted MR imaging of thyroid nodules. Neuroradiology 2009; 51:193-198

289. Dilli A, Ayaz UY, Cakir E, Cakal E, Gultekin SS, Hekimoglu B. The efficacy of apparent diffusion coefficient value calculation in differentiation between malignant and benign thyroid nodules. Clinical imaging 2012; 36:316-322

290. Erdem G, Erdem T, Muammer H, Mutlu DY, Firat AK, Sahin I, Alkan A. Diffusion-weighted images differentiate benign from malignant thyroid nodules. Journal of magnetic resonance imaging : JMRI 2010; 31:94-100

291. Mutlu H, Sivrioglu AK, Sonmez G, Velioglu M, Sildiroglu HO, Basekim CC, Kizilkaya E. Role of apparent diffusion coefficient values and diffusion-weighted magnetic resonance imaging in differentiation between benign and malignant thyroid nodules. Clinical imaging 2012; 36:1-7

292. Schueller-Weidekamm C, Kaserer K, Schueller G, Scheuba C, Ringl H, Weber M, Czerny C, Herneth AM. Can quantitative diffusion-weighted MR imaging differentiate benign and malignant cold thyroid nodules? Initial results in 25 patients. AJNR American journal of neuroradiology 2009; 30:417-422

293. Chen L, Xu J, Bao J, Huang X, Hu X, Xia Y, Wang J. Diffusion-weighted MRI in differentiating malignant from benign thyroid nodules: a meta-analysis. BMJ Open 2016; 6:e008413

294. Friedrich-Rust M, Meyer G, Dauth N, Berner C, Bogdanou D, Herrmann E, Zeuzem S, Bojunga J. Interobserver agreement of Thyroid Imaging Reporting and Data System (TIRADS) and strain elastography for the assessment of thyroid nodules. PloS one 2013; 8:e77927

295. Chou CK, Chen RF, Chou FF, Chang HW, Chen YJ, Lee YF, Yang KD, Cheng JT, Huang CC, Liu RT. miR-146b is highly expressed in adult papillary thyroid carcinomas with high risk features including extrathyroidal invasion and the BRAF(V600E) mutation. Thyroid 2010; 20:489-494

296. National Cancer Institute. The website of the National Cancer Institute. https://www.cancer.gov, 2017.

297. Cancer Genome Atlas Research N. Integrated genomic characterization of papillary thyroid carcinoma. Cell 2014; 159:676-690

298. Hyman DM, Puzanov I, Subbiah V, Faris JE, Chau I, Blay JY, Wolf J, Raje NS, Diamond EL, Hollebecque A, Gervais R, Elez-Fernandez ME, Italiano A, Hofheinz RD, Hidalgo M, Chan E, Schuler M, Lasserre SF, Makrutzki M, Sirzen F, Veronese ML, Tabernero J, Baselga J. Vemurafenib in Multiple Nonmelanoma Cancers with BRAF V600 Mutations. The New England journal of medicine 2015; 373:726-736

299. Chapman PB, Hauschild A, Robert C, Haanen JB, Ascierto P, Larkin J, Dummer R, Garbe C, Testori A, Maio M, Hogg D, Lorigan P, Lebbe C, Jouary T, Schadendorf D, Ribas A, O'Day SJ, Sosman JA, Kirkwood JM, Eggermont AM, Dreno B, Nolop K, Li J, Nelson B, Hou J, Lee RJ, Flaherty KT, McArthur GA, Group B-S. Improved

survival with vemurafenib in melanoma with BRAF V600E mutation. The New England journal of medicine 2011; 364:2507-2516

300. Pagan M, Kloos RT, Lin CF, Travers KJ, Matsuzaki H, Tom EY, Kim SY, Wong MG, Stewart AC, Huang J, Walsh PS, Monroe RJ, Kennedy GC. The diagnostic application of RNA sequencing in patients with thyroid cancer: an analysis of 851 variants and 133 fusions in 524 genes. BMC bioinformatics 2016; 17 Suppl 1:6

301. Ciregia F, Giusti L, Molinaro A, Niccolai F, Agretti P, Rago T, Di Coscio G, Vitti P, Basolo F, Iacconi P, Tonacchera M, Lucacchini A. Presence in the pre-surgical fine-needle aspiration of potential thyroid biomarkers previously identified in the post-surgical one. PloS one 2013; 8:e72911

302. Pagni F, Mainini V, Garancini M, Bono F, Vanzati A, Giardini V, Scardilli M, Goffredo P, Smith AJ, Galli M, De Sio G, Magni F. Proteomics for the diagnosis of thyroid lesions: preliminary report. Cytopathology 2014;

303. Bossuyt PM, Reitsma JB, Linnet K, Moons KG. Beyond diagnostic accuracy: the clinical utility of diagnostic tests. Clinical chemistry 2012; 58:1636-1643

304. Nikiforov YE, Seethala RR, Tallini G, Baloch ZW, Basolo F, Thompson LD, Barletta JA, Wenig BM, Al Ghuzlan A, Kakudo K, Giordano TJ, Alves VA, Khanafshar E, Asa SL, El-Naggar AK, Gooding WE, Hodak SP, Lloyd RV, Maytal G, Mete O, Nikiforova MN, Nose V, Papotti M, Poller DN, Sadow PM, Tischler AS, Tuttle RM, Wall KB, LiVolsi VA, Randolph GW, Ghossein RA. Nomenclature Revision for Encapsulated Follicular Variant of Papillary Thyroid Carcinoma: A Paradigm Shift to Reduce Overtreatment of Indolent Tumors. JAMA Oncol 2016; 2:1023-1029

305. Cerutti JM, Oler G, Delcelo R, Gerardt R, Michaluart P, Jr., de Souza SJ, Galante PA, Huang P, Riggins GJ. PVALB, a new Hurthle adenoma diagnostic marker identified

through gene expression. The Journal of clinical endocrinology and metabolism 2011; 96:E151-160

306. Evangelisti C, de Biase D, Kurelac I, Ceccarelli C, Prokisch H, Meitinger T, Caria P, Vanni R, Romeo G, Tallini G, Gasparre G, Bonora E. A mutation screening of oncogenes, tumor suppressor gene TP53 and nuclear encoded mitochondrial complex I genes in oncocytic thyroid tumors. BMC cancer 2015; 15:157

307. Finley DJ, Zhu B, Fahey TJ, 3rd. Molecular analysis of Hurthle cell neoplasms by gene profiling. Surgery 2004; 136:1160-1168

308. Ganly I, Ricarte Filho J, Eng S, Ghossein R, Morris LG, Liang Y, Socci N, Kannan K, Mo Q, Fagin JA, Chan TA. Genomic dissection of Hurthle cell carcinoma reveals a unique class of thyroid malignancy. The Journal of clinical endocrinology and metabolism 2013; 98:E962-972

309. Corver WE, Ruano D, Weijers K, den Hartog WC, van Nieuwenhuizen MP, de Miranda N, van Eijk R, Middeldorp A, Jordanova ES, Oosting J, Kapiteijn E, Hovens G, Smit J, van Wezel T, Morreau H. Genome haploidisation with chromosome 7 retention in oncocytic follicular thyroid carcinoma. PloS one 2012; 7:e38287

Diagnostic Utility of Molecular and Imaging Biomarkers in Cytological Indeterminate Thyroid Nodules

Elizabeth J. de Koster, MD[1], Lioe-Fee de Geus-Oei, MD Phd[2], Olaf M. Dekkers, MD PhD[3,4], Ilse van Engen-van Grunsven, MD Phd[5], Jaap Hamming, MD PhD[6], Eleonora P.M. Corssmit, MD PhD[3], Hans Morreau, MD PhD[7], Abbey Schepers, MD PhD[6], Jan Smit, MD PhD[8], Wim J.G. Oyen, MD PhD[9,1], Dennis Vriens, MD PhD[2]

1. Dept. of Radiology and Nuclear Medicine, Radboud University Medical Center, Nijmegen, the Netherlands.

2. Dept. of Radiology, section of Nuclear Medicine, Leiden University Medical Center, Leiden, the Netherlands.

3. Dept. of Endocrinology, Leiden University Medical Center, Leiden, the Netherlands.

4. Dept. of Epidemiology, Leiden University Medical Center, Leiden, the Netherlands.

5. Dept. of Pathology, Radboud University Medical Center, Nijmegen, the Netherlands.

6. Dept. of Surgery, Leiden University Medical Center, Leiden, the Netherlands.

7. Dept. of Pathology, Leiden University Medical Center, Leiden, the Netherlands.

8. Dept. of Endocrinology, Radboud University Medical Center, Nijmegen, the Netherlands.

9. Division of Radiotherapy and Imaging, Institute of Cancer Research, and Department of Nuclear Medicine, Royal Marsden Hospital, London, United Kingdom.

Short title: Diagnostic biomarkers in indeterminate thyroid nodules

Key words: indeterminate thyroid cytology, imaging, molecular, diagnostic

Corresponding author / address reprint requests to:

E.J. de Koster, M.D.

Radboudumc

Department of Radiology and Nuclear Medicine

P.O. Box 9101

6500 HB Nijmegen

The Netherlands

E-mail: Lisanne.deKoster@radboudumc.nl

Phone: +31 71 5297625 / +31 6 14607347

Fax: +31 71 5248256

Grants supporting the writing of the paper: the corresponding author receives a research grant from the Dutch Cancer Society (DCS).

Disclosures: none.

Abstract

Indeterminate thyroid cytology (Bethesda III and IV) corresponds to follicular-patterned benign and malignant lesions, which are particularly difficult to differentiate on cytology alone. As approximately 25% of these nodules harbor malignancy, diagnostic hemithyroidectomy is still custom. However, advanced preoperative diagnostics are rapidly evolving.

This review provides an overview of additional molecular and imaging diagnostics for indeterminate thyroid nodules in a pre-operative clinical setting, including considerations regarding cost-effectiveness, availability, and feasibility of combining techniques. Addressed diagnostics include gene mutation analysis, microRNA, immunocytochemistry, ultrasonography, elastosonography, CT, sestamibi scintigraphy, FDG-PET and diffusion-weighted MRI.

The best rule-out tests for malignancy were the Afirma® GEC and FDG-PET. The most accurate rule-in test was sole *BRAF* mutation analysis. No diagnostic had both near-perfect sensitivity and specificity, and estimated cost-effectiveness. Molecular techniques are rapidly advancing. However, given the currently available techniques a multimodality stepwise approach likely offers the most accurate diagnosis, sequentially applying one sensitive rule-out test and one specific rule-in test. Geographical variations in cytology (e.g. Hürthle cell neoplasms) and tumor genetics strongly influence local test performance and clinical utility. Multidisciplinary collaboration and implementation studies can aid the local decision for one or more eligible diagnostics.

Precis

This review discusses the value of additional molecular and imaging diagnostics for thyroid nodules with indeterminate cytology, including considerations regarding cost-effectiveness, availability, and feasibility of combining techniques. Addressed diagnostics include gene mutation analysis, microRNA, immunocytochemistry, ultrasonography, elastosonography, CT, sestamibi scintigraphy, FDG-PET and diffusion-weighted MRI.

Contents of this review

INTRODUCTION

Indeterminate thyroid cytology is an eyesore to physicians. It largely corresponds to histopathologically follicular-patterned lesions, both benign and malignant, including follicular adenoma, noninvasive follicular thyroid neoplasm with papillary-like nuclear features (NIFTP), (encapsulated) follicular variant of papillary thyroid carcinoma (FVPTC or EFVPTC) and follicular thyroid carcinoma (FTC). These neoplasms are particularly difficult to differentiate on fine needle aspiration cytology (FNAC). In the case of FTC, cytology lacks the insight into the tissue structure like histology does: it does not show the capsular and/or vascular invasion that distinguishes a FTC from a benign FA. In FVPTC, the growth pattern is follicular and clearly identifying nuclear features of PTC can usually not be identified cytologically (1-3). Nevertheless, FNAC currently has a most prominent place in the diagnostic work-up of thyroid nodules. The Bethesda System for the Reporting of Thyroid Cytology was adopted in its current form in 2009, recognizing six diagnostic categories with an incremental risk of malignancy and clinical management guidelines. Although the Bethesda system created a much-used handhold by standardizing the cytological diagnosis and consecutive management of thyroid nodules worldwide, the system does not provide a

clear answer for the heterogeneous group of nodules with indeterminate cytology (4,5). This includes cytology with atypia of undetermined significance or follicular lesion of undetermined significance (AUS/FLUS, Bethesda III), and cytology (suspicious for a) follicular neoplasm (SFN/FN) or (suspicious for a) Hürthle cell neoplasm (SHCN/HCN, Bethesda IV). Similar indeterminate cytological categories are found in the British Thyroid Association Thy system and Italian SIAPEC-IAP classification: Thy3a and Thy3f, and TIR3A and TIR3B, respectively **(Table 1)** (6,7).

Alongside a doubled incidence of thyroid carcinoma over the past two decades and a prevalence of thyroid nodules stretching far beyond the 5% for palpable nodules – explained by the incidental detection of nonpalpable nodules and clinically occult thyroid cancers on imaging studies – the need for a more accurate diagnostic procedure has grown (8). This urge was further emphasized when other research groups were unable to reproduce the prevalence of the cytological categories and corresponding malignancy risks proposed by Cibas *et al.*, especially those of the AUS category (4,9,10). Insuperable variations in the worldwide patient populations, and intra- and interobserver variation in the assessment of thyroid cytology were named as likely underlying causes (4,5,10,11). Yet, it raised questions concerning the overall approach of thyroid nodule diagnosis and whether cytology is the best starting ground. Cost-effectiveness is a major benefit of cytological examination, yet a more accurate test may eventually replace cytological examination completely (12,13). At present, however, a supplemental diagnostic procedure is specifically warranted for cytologically indeterminate thyroid nodules. Diagnostic hemithyroidectomies are still customarily performed to obtain a definite histological diagnosis. With a benign histopathological result in approximately three in four cases, surgery was not only unbeneficial but also exposed the patient to unnecessary surgical risks. In the case of malignant lesions, a second-stage completion thyroidectomy is often indicated, which is associated with additional costs and higher risks of surgical

complications (14-17). An additional preoperative test or combination of tests for thyroid nodules with indeterminate cytology should prevent unbeneficial diagnostic hemithyroidectomies for benign nodules, limit the number of two-stage surgeries for thyroid malignancies, or both. With rapidly advancing technology, the possibilities for additional diagnostic techniques seem endless: the applications of existing diagnostics such as ultrasound, PET/CT and immunocytochemistry are extended and more clearly demarcated for use in indeterminate thyroid nodules. High-tech molecular tests such as gene mutation panels, gene or microRNA expression profiles and sequencing techniques are hot-topic (4,18-22). Every currently known engagement point from the genotype to the phenotype of the tumor is being explored. Combined, the various research fields encompass an extensive range of investigative methods. Individually they usually focus on one or two methods only, making one-to-one comparison of these diagnostics difficult. The 2015 American Thyroid Association (ATA) guidelines suggested several additional tests, but a definitive answer or complete overview of all available tests is still lacking (23).

Alongside higher-level expert discussions and lobbying of med tech companies, clinical endocrinologists and thyroid surgeons ponder about the best solution for their individual patients. Their choices depend on the characteristics of their patient populations, availability and costs of a certain test, and personal preference. In any case, a useful additional test should be accurate, accessible, affordable and affect patient management.

This review aims to provide practical considerations for physicians involved in the management of patients with thyroid nodules. It gives an overview of the available literature on additional diagnostic tests for thyroid nodules with indeterminate cytology. We will work our way down from genotype to phenotype, discussing both anatomical and functional techniques, from the state-of-the-art molecular and imaging biomarkers as well as widely available conventional imaging techniques. The ability of a test to distinguish between

malignant and benign nodules in a preoperative setting is discussed, focusing on clinical validation and utility, and including the development phase, cost-effectiveness and availability of each technique, where appropriate. Table 2 provides a summarized overview of the discussed diagnostics and their main attributes.

1. MOLECULAR BIOMARKERS

1.1. Gene mutation analysis and Gene expression

In the last decades, researchers have unraveled important molecular mechanisms behind the thyroid tumorigenesis, and designated a great number of genetic alterations that are related to the various types of thyroid carcinoma. Several of these mutational markers have found their way to the preoperative diagnosis of indeterminate thyroid nodules. The most common markers are the somatic *BRAF* and *RAS* point mutations, and RET/PTC rearrangement, all of which involve the mitogen-activated protein kinase (MAPK) signaling pathway (24-26).

In the 2015 ATA guidelines the potentially strong diagnostic impact of molecular testing is explicitly unfolded, focusing on BRAF testing and the – at that date – two main commercially available tests: the seven-gene mutation panel miRInform® thyroid (Asuragen Inc., Austin, Texas) and the Afirma® gene expression classifier (Veracyte, Inc., South San Francisco, CA). The ATA recommends considerate application of one of these molecular tests for Bethesda III and IV nodules, provided that the result could change the treatment strategy (23).

In the following chapters, the diagnostic potential of mutation analysis in indeterminate thyroid nodules is discussed, including the tests mentioned in the guidelines as well as other individual molecular biomarkers and multi-gene panels addressed in literature.

1.1.1. BRAF mutation

B-type RAF kinase (BRAF) is a serine–threonine kinase belonging to the rapidly accelerated fibrosarcoma (RAF) family, and the most potent mitogen-activated protein kinase (MAPK) pathway activator. Point mutations in the *BRAF* proto-oncogene occur in various human cancers. The somatic $BRAF^{V600E}$ mutation is the most common activating mutation in many carcinomas, including thyroid carcinoma (24). This missense mutation consists of a thymine-to-adenine substitution at nucleotide 1799 (c.1799T>A), resulting in an amino acid substitution where valine is replaced with glutamate at codon 600 (hence V600E)(27,28). BRAF has an important function in cell proliferation, differentiation, and apoptosis. Upregulation of BRAF through the $BRAF^{V600E}$ activating mutation is associated with tumorigenesis (28). In differentiated thyroid cancer, the $BRAF^{V600E}$ mutation is exclusive to PTC, occurring in 50% to 80% of these tumors (24,25,29-39). The $BRAF^{V600E}$ mutation has been prognostically associated with poor clinicopathological outcomes, such as increased incidence of extrathyroidal invasion, recurrence of disease, and distant metastasis of the tumor (40-42).

BRAF mutation analysis has been extensively studied as a rule-in test for thyroid carcinoma. The *BRAF* mutation is superior to other mutations in its oftentimes 100% specificity – a positive mutation could prevent two-stage surgery for an indeterminate thyroid nodule (21,29-31,35-38,41-76). Even though the *BRAF* mutation was found in a majority of PTC in a number of studies, the prevalence of the *BRAF* mutation in indeterminate cytology ranged from 0% to 48% in individual studies (44,46,48,59,65,70). Reported sensitivities were therefore heterogeneous and generally poor, ranging from 0% to 83% (29,34,39,46). Other types of thyroid carcinoma occurring in indeterminate nodules, including FTC, FVPTC and Hürthle cell carcinoma (oncocytic variant of follicular thyroid carcinoma, FTC-OV), were

respectively never or infrequently *BRAF* mutation-positive (31,37,38,42,50,57,76). Predominated by follicular type carcinoma, the *BRAF* mutation rarely occurs in Bethesda IV cytology (29,31,37,41,50,52,54,57,59,60,63,65,66,70,73-80).

Likely contributors to the observed heterogeneity are known global variations in the occurrence rates of PTC and *BRAF* mutations. In South Korea, where iodine consumption is high, 90% to 95% of thyroid cancers are PTC. More specifically, the proportion of *BRAF* -mutated PTC is very high: rates of 80% to more than 90% are reported (34,46,77). Consequently, $BRAF^{V600E}$ mutation analysis might have both high specificity and high sensitivity in these populations. Studies with higher sensitivities were more often of South Korean origin and frequently demonstrated sensitivity above 40%, with the prevalence of *BRAF* mutations reported as high as 30% to 48%. (34,39,46,47,73,81-83). Conversely, the majority of studies with sensitivity below 10% were conducted in Western countries (USA, Europe or Canada), with some studies reporting no *BRAF* mutations at all (21,31,37,45,48,52,56-59,61,62,65,69,70,76,80).

Some South Korean studies based surgical decision-making on the result of the *BRAF* mutation analysis: surgery was relatively less often performed in *BRAF* mutation-negative indeterminate nodules (34,39,78,83). Such a surgical management strategy is not oncologically safe for Western countries (e.g. Europa or Northern America), where 80% to 90% of thyroid carcinomas are PTC and reported rates of *BRAF*-mutated PTC vary from 30% to 40% (34,46,77). Moreover, even though the true sensitivity of *BRAF* mutation analysis is presumably high in South Korea for the mentioned epidemiological reasons, the conservative management of *BRAF* mutation-negative nodules likely magnified test sensitivity by underestimating the rate of *BRAF*-negative malignant nodules in these studies. Altogether we estimate that approximately one in five South Korean patients would benefit from *BRAF* mutation analysis, opposite mere one in 25 patients from other countries.

BRAF mutation in papillary microcarcinoma

Papillary microcarcinoma (mPTC) have lower *BRAF* mutation rates (53,58,63,68,73,76,84). The ATA guidelines are reserved with regard to the recommended clinical management of positive *BRAF* mutation in mPTC, as its relation to extrathyroidal spread and positive lymph node metastases is not as clear as in larger thyroid carcinoma. Although there are studies that associate mPTC to factors of poorer prognosis, the 2015 guidelines recommend that *BRAF*-mutated mPTC are treated as low-risk malignancies (23,35).

$BRAF^{K601E}$ point mutation

A less common activating *BRAF* mutation is $BRAF^{K601E}$ (c.1801A>G), which occurs considerably less frequently than the $BRAF^{V600E}$ variant and is associated with FVPTC with high specificity (85). Clinically, the characterization of a small cohort of thyroid malignancies with a $BRAF^{K601E}$ mutation showed better outcomes than for $BRAF^{V600E}$ mutated tumors: no extrathyroidal tumor extension, recurrence, lymph node or distant metastasis were reported in indeterminate $BRAF^{K601E}$ positive tumors with a median follow-up of 20 months (range 4-47) (86).

Availability, cost-effectiveness and limitations of BRAF mutation analysis

Altogether, the consistent perfect specificity in a large number of studies supports the use of *BRAF* mutation analysis in obviating two-stage surgery. The technique is increasingly available in the clinical setting worldwide. A prior meta-analysis of eight studies questioned the cost-effectiveness of $BRAF^{V600E}$ mutation analysis in indeterminate thyroid nodules based on a mere 4.6% mean prevalence of the mutation (87). Cost-effectiveness studies concerning sole *BRAF* mutation analysis in indeterminate thyroid nodules are lacking. Regardless, cost-

effectiveness is generally presumed, as average costs for testing are relatively low and decreasing over time. Depending on the applied molecular technique, reported costs for *BRAF* mutation analysis ranged between €7.50 and $123 per tested sample (53,63,72,88).

Low sensitivity remains the main limitation of *BRAF* mutation analysis, irrespective of the type of indeterminate cytology. Proficiency of the test in preoperative patient management depends on the regional occurrence rate of *BRAF*-mutated PTC; in South Korea, more patients will benefit from *BRAF* mutation analysis, and the probability and extent of cost-effectiveness are likely to increase (66). In other health care systems, such as in the UK, cost-effectiveness is likely more constrained. Nonetheless, *BRAF* testing could still save approximately half the surgical costs in *BRAF* mutation-positive carcinoma (63,65). These global variations should be considered before local implementation of sole *BRAF* mutation analysis.

1.1.2. RAS point mutation

Point mutations in the gene family of retrovirus-associated DNA sequences (RAS) together constitute the second most frequently occurring genetic alteration in thyroid carcinoma. In indeterminate thyroid nodules, they are the most common genetic alteration, due to a strong association of *RAS* mutations with the follicular-patterned lesions that make up these cytological categories: follicular adenoma, FTC, FVPTC and noninvasive follicular thyroid neoplasms with papillary-like nuclear features (NIFTP) (1,3,31,59,89,90). Originally, two of the three homologous *RAS* genes were identified as viral genes of the oncogenic Harvey (*HRAS*) and Kirsten (*KRAS*) murine sarcoma virus; the third, *NRAS*, was first identified in neuroblastoma cells (91,92). The genes code for GTP-binding RAS proteins, which are

involved in intracellular signaling in the MAPK/ERK pathway. Mutation causes overactive RAS signaling and could ultimately induce malignant transition (26).

RAS mutation in thyroid carcinoma has been associated with favorable prognostic factors, such as encapsulation of the tumor and absence of lymph node metastases, but also with factors indicative of an adverse prognosis, such as poor cell differentiation (2). *RAS* mutations are not specific for carcinoma and found in both malignant and benign lesions (31,61,90). According to the 2015 ATA guidelines, Bethesda III or IV nodules with a *RAS* mutation should be treated similar to the Bethesda V category, as approximately 4 out of 5 are malignant (4,23). *HRAS, KRAS* and *NRAS* mutations are mutually exclusive. They are each associated with slightly different types of cytology and histology, and consequently a different clinical course. In general, point mutations in *NRAS* codon 61 and *HRAS* codon 61 are said to occur most frequently (3,64). *KRAS* is associated with oncocytic lesions and a lower malignancy rate than other *RAS* mutations (93).

A *RAS* point mutation is found in 0% to 38% of the indeterminate nodules (39,60). Moreover, approximately a third of all reported malignancies resulting from indeterminate thyroid cytology are *RAS* mutation positive, frequently FVPTC or FTC (31,37-39,76). Sporadic cases of *RAS* mutation-positive FTC-OV and MTC are reported (37,38). In individual studies, sensitivity and specificity of *RAS* mutation analysis ranged from 0% to 77% and from 75% to 100%, respectively (39,60,90). Test performance was similar for Bethesda III and IV categories, although the mutation occurred more frequently in Bethesda IV nodules (21,29,31,39,50,59,60,76,80,90). Histopathologically benign nodules carrying a *RAS* mutation are histopathological follicular adenoma in most cases, but also oncocytic variant of follicular adenoma (Hürthle cell adenoma) or hyperplastic nodules (29,31,38,50,90). There is an ongoing discussion regarding the interpretation of a false positive *RAS* mutation. It is presumed that an oncogenic *RAS* mutation predisposes a follicular adenoma for progression

into follicular carcinoma – a *RAS*-mutated follicular adenoma should be considered a premalignant pre-invasive follicular neoplasm. These assumptions put false-positives in a different light, as it would justify resection of such lesions through hemithyroidectomy. Consequently, the lesions could also be considered true-positives – improving the specificity of *RAS* mutation analysis (1,21,39,59,61,71). However, the exact mechanisms behind the malignant potential and transition for *RAS*-mutated follicular adenomas are not yet clarified and difficult to appreciate in a clinical setting.

Similar to *BRAF*, there was evident global variation in the distribution of *RAS* mutations. Many European and American studies reported a clear predominance of *RAS* mutations over *BRAF* mutations. Solely a Brazilian study of 116 Bethesda III and 20 Bethesda IV thyroid nodules reported only *BRAF* mutations and not a single *RAS* mutation (60). The previously described predominance of *BRAF* mutations in South Korean populations was confirmed in the sole study that investigated both point mutations in one population (39). Combined BRAF/RAS mutation analysis could be considered, although geographical differences in the distribution of the two genetic alterations strongly influence feasibility. A gene mutation panel consisting of more genetic alterations (discussed in a next chapter) is most likely more useful.

Sole *RAS* mutation analysis is not accurate in the preoperative setting. Although specificity is high, only two out of three *RAS* mutation positive indeterminate nodules are histopathologically malignant, evidently fewer than assumed and previously described in the ATA guidelines. Therefore, *RAS* mutation positive indeterminate thyroid nodules should be surgically managed with no more than hemithyroidectomy. Whether hemithyroidectomy is justified for *RAS*-mutated follicular adenomas as a precancerous lesion, is yet under debate.

1.1.3. RET/PTC rearrangement

Rearrangements of the *RET* proto-oncogene arise from the fusion of the 3' end of *RET* to the 5' regions of unrelated genes that are expressed in thyroid follicular cells. Proto-oncogene *RET* encodes for a transmembrane receptor with a tyrosine kinase domain; a RET/PTC rearrangement causes inappropriate overexpression of that domain. It activates the MAPK and PI3K/AKT pathways and stimulates malignant transition of the cell through BRAF (94,95). At least 12 different fusion variants have been detected until today, of which RET/PTC1 and RET/PTC3 are the most common. They have a well-known association with PTC. Cases of both rearrangements in a single lesion are also reported (2,94,96,97). RET/PTC rearrangements, especially RET/PTC3, occur more frequently in PTC in children or patients that were exposed to ionizing radiation and are clinically associated with the presence of lymph node metastases (2). Worldwide variations in frequency of RET/PTC rearrangements exist, dependent on demographics and ethnicity. The RET/PTC rearrangement is present in 42% of PTC in Western populations with a predominance of RET/PTC1, and in 37% of PTC in Asian populations with a predominance of RET/PTC3. Without radiation exposure, in female PTC patients RET/PTC1 is predominant (98). The rearrangements are also found in benign nodules, especially in patients that were exposed to ionizing irradiation (29,97). Alike *RAS* mutations, it is assumed to be an activating genetic alteration and it is argued that a histopathologically benign nodule with a RET/PTC rearrangement should be considered a precancerous lesion.

RET/PTC rearrangements are seldom found in indeterminate nodules. In many studies, no RET/PTC translocation was found at all. Most studies investigated RET/PTC in light of a gene mutation panel and paid it no specific attention (21,29,31,37,38,45,49,50,55,59,61,62,64,70,76,80,96). Only Guerra *et al*. solely investigated the RET/PTC rearrangement in 101 thyroid nodules of all cytological categories. In this

Italian study, RET/PTC rearrangements were found in 18 of the 50 PTC (36%) using RT-PCR and Southern-Blot. All these RET/PTC-positive carcinomas were Thy4 or Thy5 nodules on cytology. Among the 24 Thy3 nodules, two nodules with a RET/PTC3 rearrangement were histopathologically benign (96).

Noteworthy, Sapio *et al*. detected two *RET* mutations during their RET/PTC assessments. In contrast to the RET/PTC translocation, RET point mutations are related to sporadic and familial MTC (45,99). Surgery confirmed histopathological MTC in the *RET*-mutated nodules (45).

Even though previous histological studies undeniably associated RET/PTC1 and RET/PTC3 rearrangements to PTC, the low prevalence of the rearrangement in indeterminate cytology is a major downside. Testing exclusively for this genetic alteration in indeterminate nodules is not advantageous, even if issues regarding the number of tested variants and sensitivity of molecular techniques are overcome. The 2015 ATA guidelines only advise RET/PTC testing in context of a gene mutation panel (23).

1.1.4. PAX8/PPARγ rearrangement

The PAX8/PPARγ rearrangement arises from a fusion of the promoter and 5'-coding portion of the thyroid-specific transcription factor *PAX8* gene to the gene of the nuclear receptor peroxisome proliferator-activated receptor γ (PPARγ) (100,101). The role of the product of this translocation – the PAX8/PPARγ fusion protein – is not yet understood, as the DNA binding sites of both original proteins are uniquely preserved in the fusion (101). In the normal thyroid, transcription factor PAX8 is involved in differentiation of thyrocytes and regulation of the expression of thyroid-specific genes encoding thyroperoxidase,

thyroglobulin and the sodium/iodide symporter (102). Nuclear receptor PPARγ has multiple presumed functions, including involvement in the regulation of lipid metabolism, adipogenesis and insulin sensitivity (101,103).

The chromosomal translocation PAX8/PPARγ was first discovered in – and traditionally associated with – FTC and follicular adenoma (100). It is reported in 30% to 45% of FTC and in up to 33% of follicular adenoma (50,104-106). However, several studies have also uncovered varying amounts of FVPTC carrying the translocation, with published rates up to 38% (104,107,108). It has not been reported in benign or malignant Hürthle cell neoplasms (105,109).

PAX8/PPARγ is often related to well-differentiated malignancies with a relatively favorable prognosis. Capsular and vascular invasion are reported to a lesser extent in FTCs with a PAX8/PPARγ rearrangement than in *RAS*-mutated tumors (105). Widely invasive features are not reported. PAX8/PPARγ-mutated FVPTC are mostly encapsulated, following an indolent clinical course with minimal disease recurrence despite the presence of some capsular and vascular invasion at presentation (105-107). In contrast to the *BRAF*, *RAS* and RET/PTC genetic alterations in thyroid carcinoma, the PAX8/PPARγ rearrangement does not involve the RAS-RAF-MAPK pathway. Nikiforova *et al.* hypothesized that oncogenesis of follicular-type tumors likely takes place through two different molecular pathways: a *RAS*-mutation driven and PAX8/PPARγ rearrangement driven pathway (105).

Similar to the RET/PTC rearrangement, the PAX8/PPARγ rearrangement rarely occurred in indeterminate thyroid cytology. Approximately two-thirds of the indeterminate nodules carrying the rearrangement were histopathologically malignant, most often FVPTC or FTC (21,31,37,38,49,55,59,61,62,70,76,80,109). False-positive results corresponded to follicular adenomas (61,62,70). Similar to *RAS* mutations, histopathologically benign PAX8/PPARγ-mutated nodules are likely premalignant lesions, or pre-invasive FTC. Eszlinger *et al.*

observed a microfollicular morphological growth pattern in two of the PAX8/PPARγ-positive follicular adenoma, supporting this hypothesis (61).

Still, PAX8/PPARγ rearrangement is a rare rearrangement associated with (encapsulated) follicular tumors. Similar to RET/PTC rearrangements, the PAX8/PPARγ rearrangement should only be assessed in indeterminate thyroid nodules in combination with more frequently occurring genetic alterations in a gene mutation panel.

1.1.5. Other genetic alterations

hTERT

The enzyme human telomerase is involved in the maintenance of the chromosomes' telomeres, which are essential for cell life and proliferation. The catalytic subunit of telomerase is human telomerase reverse transcription (hTERT). In normal thyroid cells, it is inactive. Inappropriate reactivation is associated with malignancy and inflammatory thyroid disease (110). hTERT promotor mutations were previously observed in both PTC and FTC, sometimes together with a *BRAF* mutation. The mutation is strongly correlated to mortality in differentiated thyroid carcinoma (111). hTERT gene expression is potentially accurate in the preoperative differentiation of indeterminate nodules, with 57% to 88% sensitivity and 75% to 85% specificity demonstrated in two small clinical series of cytological follicular neoplasms (112,113).

TRK

The tyrosine receptor kinase (TRK) rearrangement arises from a translocation of the *NTRK1* gene, which is normally expressed in the central and peripheral nervous system and involved in cell differentiation. The TRK rearrangement is associated with PTC and presumably with

an adverse prognosis, although evidence is limited (114). In feasibility studies in indeterminate thyroid cytology, not a single TRK rearrangement has been detected – it is most likely not a useful marker (45,49,50).

HMGA2

Proteins high mobility group AT-hook (HMGA) 1 and 2 regulate the structure and function of chromatin. Normally only expressed during embryogenesis, the overexpression of HMGA in adult tissues is associated with malignancy (115). Lappinga *et al.* demonstrated that HMGA2 could be a promising additional biomarker. Using ROC curve analysis, a >5.9-fold HMGA2 overexpression had 76% sensitivity and 98% specificity in SFN/FN nodules (116). To date no other studies attempted to validate these results.

Galectin-3 and CD44v6

One Croatian study used RT-PCR to investigate the simultaneous expression of galectin-3 and CD44v6, two molecular biomarkers better known for their application in immunohistochemistry of their expression products (117). CD44v6 normally functions as the cell-surface receptor for hyaluronic acid. Overexpression is found in various human cancers, including thyroid (117,118). In indeterminate thyroid nodules, a positive test for either one of the two biomarkers resulted in 100% sensitivity and 60% specificity. It is presumed that similar results for these markers are achieved with the more economical immunohistochemistry techniques (118,119).

1.1.6. 7-Gene Mutation Panel

Ongoing research in the past years has demonstrated that assessment of individual oncogenic mutations generally has limited clinical utility in indeterminate thyroid cytology. Combining forces of individual genetic alterations into a gene mutation panel, however, likely improves diagnostic accuracy, especially as mutations are mutually exclusive in most cases. These gene mutation panels typically assess the seven genetic alterations – gene mutations as well as gene fusions – that occur most frequently in differentiated thyroid carcinoma, including $BRAF^{V600E}$, $BRAF^{K601E}$, *NRAS* codon 61, *HRAS* codon 61 and *KRAS* codon 12-13 point mutations and RET/PTC1, RET/PTC3 and PAX8/PPARγ gene rearrangements (31,55). The best known panel is the commercially available miRInform® thyroid (Asuragen Inc., Austin, Texas, USA), currently rebranded as the ThyGenX® Thyroid Oncogene Panel (Interpace Diagnostics, Parsippany, NJ, USA). The miRInform® thyroid tests 17 specific genetic alterations in these seven genes (59). It is marketed as a rule-in test for thyroid malignancy.

The first large clinical utility study to investigate the miRInform® thyroid test was published in 2011. Nikiforov *et al.* prospectively included 1,056 FNAC samples, 92% of which had sufficient epithelial cells and nucleic acids to pursue molecular testing. Residual FNAC material was used for mutation analysis – no additional aspirates were required. Unfortunately, surgery was performed for only 461 of 900 (51%) indeterminate thyroid nodules, independent of the test outcome; these operated cases were included in their final analysis. It is not reported whether nonsurgically managed nodules were mutation-positive or -negative. Sensitivity and specificity were 63% and 99% in the 247 Bethesda III nodules, and 57% and 97% in the 214 Bethesda IV nodules, respectively. The authors suggested that the high PPV of the miRInform® thyroid in these indeterminate thyroid nodules (88% and 87%, respectively) warrants a direct total thyroidectomy instead of two-step surgery in patients with a positive test (31).

None of the subsequent studies matched the initially reported excellent specificity. The next industry-sponsored prospective study by Beaudenon *et al.* reported 47% sensitivity and 88% specificity in 80 Bethesda III and IV nodules. Surprisingly, not a single BRAF mutation was detected (59). Valderrabano *et al.* reported not a single mutation in 47 included Bethesda III nodules. Moreover, only 1 of 18 nodules with Hürthle cell cytology in this study tested positive, suggesting that Hürthle cell nodules may carry different mutations than the ones investigated by the miRInform® thyroid (76). Ohori *et al.* demonstrated that genetic alterations less frequently occurred in the textbook colloid-poor Bethesda IV cytology compared to the less common colloid-rich variant. Differences in etiology are unknown, but the authors hypothesized that the two types have subtle histopathological differences. The colloid-rich thyroid carcinomas likely more often develop through the well-known mutations included in the miRInform® thyroid test, whereas mutations that elicit colloid-poor thyroid carcinoma are yet unknown (37).

Simultaneously with the American miRInform® studies, five European studies independently investigated whether a panel of the same 7 genes could reliably be assessed using different methods (38,55,61,62,70). In three separate studies, Eszlinger *et al.* demonstrated that testing was also feasible on routine air-dried FNAC samples from indeterminate thyroid nodules. Over the course of these studies, sensitivity of this method improved from 18% to 49% and specificity from 86% to 93%, respectively. The use of air-dried FNAC samples for mutation analysis could advance the implementation of mutation analysis in daily practice, as specific storage conditions of fresh FNAC samples for mutation analysis are no longer required (38,61,62). Mancini *et al.* showed that high-resolution melting (HRM) analysis is an accurate screening method for the seven genetic alterations, with 56% sensitivity and 90% specificity. HRM is a post-PCR procedure that does not require significant additional resources. This

could reserve the costlier direct sequencing procedures solely for samples with abnormal HRM results, thereby reducing the overall costs of mutation analysis (55).

Overall, reported sensitivities and specificities of a 7-gene mutation panel in indeterminate thyroid nodules ranges from 18% to 69% and 86% to 99%, respectively (22,37,61). It is an adequate diagnostic tool with a high rule-in capacity in indeterminate nodules. Test performance was similar in Bethesda III and Bethesda IV nodules, although the latter more frequently had a positive test result based on the higher prevalence of *RAS* mutations (31,59). Due to the common *RAS* mutations, PPV of the 7-gene mutation panel never exceeds 90% in a range of realistic 15% to 40% prevalence of malignancy. As such it is debatable whether a positive test warrants immediate single-stage total thyroidectomy. It translates into an inappropriate overtreatment in a significant number of patients with a positive test but benign final histology at higher risk of surgical complications and all requiring lifelong levothyroxine supplementation. Deliberate surgical decision-making should consider the underlying positive mutation rather than mere the positive test itself.

The limited size of the seven-gene gene mutation panel keeps the costs per test low compared to other, larger molecular panels. Reported prices of the 7-gene mutation panel all concern the commercial miRInform® thyroid (or ThyGenX® Thyroid Oncogene Panel) and range between $425 and $1,700 (120,121). Implementation of miRInform® testing for indeterminate nodules theoretically resulted in a 20% cost reduction in the USA: the prevented two-step surgical procedures would outweigh the added expenses for miRInform® testing and increased number of total thyroidectomies – including those for nodules with a false-positive test (120). In a European setting, treatment and hospitalization costs are generally lower and miRInform® would most likely not be cost-effective (14). However,

these cost-effectiveness studies both adopted the unequalled test performance from the initial key publication – true cost-effectiveness may be less optimistic (14,31,120,121).

1.1.7. Next Generation Sequencing

To improve the sensitivity of the miRInform® thyroid test, the existing 7-gene mutation panel was expanded to include additional gene mutations, fusions and translocations, and a microRNA gene expression panel. In addition, it adopted promising next generation sequencing (NGS) techniques. NGS enables the simultaneous targeted testing for multiple mutations in large gene panels and is faster, more sensitive and more cost-effective than traditional Sanger sequencing and other PCR-based methods (71,80,122). As NGS only requires a very small amount (5-10 ng) of nucleic acids, remainder material from regular FNAC passes suffices and no additional aspirates are required (80,122). The first thyroid-specific NGS-based gene panel was the ThyroSeq® v1, presented in 2013. It detected gene variations in 110 of 145 investigated thyroid cancer tissue samples and 5 of 83 benign specimens. Unfortunately, indeterminate FNAC samples were not analyzed separately in this study. Nonetheless, Nikiforova *et al*. demonstrated that NGS had a very high success rate and could be a promising molecular technique for thyroid FNAC samples (122).

Following the ThyroSeq® v1, the road was paved for further exploration of NGS-based diagnostics. Soon, the ThyroSeq® v2 (CBLPath, Ocala, FL, USA) was developed, with a number of primers for TERT promotor variants added to its panel. It simultaneously tested for point mutations in 13 genes and for 42 types of gene fusion products (80). The ThyroSeq® v2 was tested on 143 Bethesda IV thyroid nodules. Forty-two genetic alterations were found,

most frequently *NRAS*. Diagnostic accuracy of the ThyroSeq® v2 was 92%, with astonishing 90% sensitivity and 93% specificity (80).

More recently, Nikiforov *et al*. tested the ThyroSeq® v2.1 – including point mutations in 14 genes and 42 gene fusion transcripts – in 462 Bethesda III nodules. Based on the promising results of the previous study, surgery was withheld for 362 of 431 ThyroSeq®-negative patients. In the 95 patients with available histopathology, the ThyroSeq® v2.1 demonstrated 91% sensitivity and 92% specificity. Additionally, diagnostic accuracy was estimated for malignancy rates varying between 6% and 48%: PPV would range from 42% to 91%, NPV from 92% to 99%. Within reasonable limits the ThyroSeq® v2.1 is highly reliable to rule out malignancy (21).

Le Mercier *et al*. retrospectively tested a different commercially available 50-gene NGS panel, the Ampliseq™ Cancer Hotspot Panel v2 (ThermoFisher, San Diego, California, USA), which is a tumor-nonspecific NGS panel for detection of somatic tumor variants. This panel does not include thyroid-specific RET/PTC, PAX8/PPARγ and NTRK1 rearrangements. Albeit the study only assessed 34 FNAC samples, with a 71% sensitivity and 89% specificity in indeterminate thyroid nodules the Ampliseq™ panel seems less accurate than the ThyroSeq® (71).

The high diagnostic accuracy is also a downside to NGS. Highly sensitive, NGS is able to identify mutant alleles at very low levels (<10%). A low percentage of mutant alleles might reflect a subclone within the nodule, which is not histopathologically identified as carcinoma. This detection of germline or clinically insignificant low-level somatic mutations in benign nodules could decrease NGS specificity (22,80). Nikiforov *et al*. suggested that the next improvement of the NGS-related tests should therefore be to determine accurate threshold levels for the various gene variations (80).

NGS encompasses crucial technology that is rapidly advancing. The ThyroSeq® v3 was recently announced, promoting to encompass no less than ~95% of genetic alterations occurring in PTC. Extraordinary diagnostic accuracy above 90% is anticipated, including high accuracy in Hürthle cell lesions. Results of the prospective studies validating this new version will likely be published shortly. Nonetheless, NGS techniques currently have limited global availability, with the exception of some European countries and the USA. The ThyroSeq® is available for $3,200 per test (123). In contrast, the thyroid non-specific AmpliSeq™ panel can be ordered online for only €230 (124). Independent prospective studies are needed to validate its performance and predicted cost-utility in different patient populations, and confirm the superior position of the ThyroSeq® and other NGS techniques.

1.1.8. Afirma® Gene Expression Classifier

In molecular diagnostics, the chief competitor of the 7-gene mutation panel is the commercial Afirma® gene expression classifier (GEC) (Veracyte Inc., South San Francisco, CA, USA). The GEC uses quantification of the mRNA-expression of 167 genes and a proprietary classification algorithm to determine the probabilities of malignancy in the samples' expression patterns. The classification algorithm to discern a 'benign' (negative test) from a 'suspicious' (positive test) thyroid nodule results from a successful designer study that trained the GEC in both a tissue set and diverse FNAC sample sets with known histopathology (125). Alexander *et al.* performed the first prospective, blinded, industry-sponsored clinical study to validate this Afirma® GEC in patients with indeterminate thyroid nodules (126). From 49 hospitals 577 Bethesda III, IV and V FNAC samples were collected, obtained by two additional needle aspirates from thyroid nodules with a diameter of at least 1 cm. After exclusion of over half (312/577, 54%) of the samples for reasons such as nodules that were

not surgically resected, duplicate specimens from the same nodule, and issues with specimen shipments to Veracyte, finally 265 FNAC samples were included in the analysis. Sensitivity of the Afirma® GEC was 90% in the 129 Bethesda III as well as the 81 Bethesda IV nodules with a useful GEC-negative test result in 38% (100/265), but specificity was merely 53% and 49%, respectively (52% on average). Despite the relatively high malignancy rate in Bethesda III nodules and the high number of exclusions, this study is well conducted and recognized worldwide as the landmark study that demonstrated the strength of the Afirma® GEC (126). After the overwhelming results from this key-publication, popularity of the GEC took flight. It is marketed as a highly accurate rule-out test for malignancy in thyroid nodules with indeterminate cytology.

In 2014, the first multicentre study that retrospectively assessed the clinical utility of the Afirma® GEC was published. Only 6% of reported GEC-negative Bethesda III, IV and V nodules eventually underwent surgery, of which one resulted in a 6 mm mPTC. Unfortunately, data on GEC negative nodules were only reported on an aggregate level; exact test performance rates in Bethesda III and IV nodules cannot accurately be extracted from the publication. Less than half of the GEC-negative nodules without surgery (71/163, 44%) had clinical or radiological follow-up, ranging from 1 to 24 months (median 8 months) – a limited duration compared to the natural, indolent course of differentiated thyroid carcinoma. The published paper does not describe whether the remaining 92 patients with GEC-negative nodules received any follow-up at all. Despite evident limitations to the applied reference standards, Alexander *et al*. concluded that their results confirm both the accurate test performance from their prior study as well as the large impact that the Afirma® GEC has on clinical decision-making for cytologically indeterminate thyroid nodules (127).

Yet, physicians indeed seemed reassured by a negative GEC result based on the first studies alone (126,128). In many institutions in the USA the Afirma® GEC was immediately implemented in clinical practice. The retrospective studies that followed were mere post-implementation utility studies, and generally reported very high but moderately consistent sensitivities. GEC-negative nodules were largely managed without surgery and considered true-negative, resulting in possible overestimation of test sensitivity. Long-term follow-up is not yet available to endorse a benign diagnosis in these cases (129,130). The high degree of missing histology was recognized by most of these studies as a major limitation (127,129,131-134). This was confirmed by the 2015 ATA guidelines: recognizing the Afirma® GEC as a promising diagnostic tool, the guidelines stress that it is a major shortcoming that external clinical validation studies with full histological follow-up of Afirma® GEC-negative nodules are still lacking (23).

Not all studies were able to confirm the potential of the Afirma® GEC. Some struggled with a low benign call rate (i.e. useful negative test result that could lead to management change) (130,135). McIver *et al*. questioned the cost-effectiveness of the Afirma® GEC in their population, as the mere 22% (16/72) negative test rate was much lower than anticipated. Moreover, a quarter of these GEC-negative patients rejected the proposed conservative treatment of ultrasound-based follow-up and underwent surgery anyway; one of them was diagnosed with a 3.2 cm FTC with focal capsular and vascular invasion. Also, 84% of GEC-positive nodules proved histopathologically benign, overall resulting in a disappointing 83% sensitivity and 10% specificity (130).

Besides concerns regarding adequate clinical validation of test performance, the post-implementation influence of the GEC on surgical decision-making for individual patients was also questioned. In line with the results of their preliminary study, Noureldine *et al*. demonstrated that Afirma® GEC testing had not aided surgical decision-making (135,136). In

93% (206/222) of the included indeterminate nodules, a ‘benign’ or ‘suspicious’ GEC result did not affect management at all: the surgical strategy would have been identical had it been based merely on clinical, cytological or radiological suspicion. However, if management changes were based on the GEC result, they were more often wrong than right: 11 times GEC-positive results inappropriately tempted physicians into more aggressive surgery, and total thyroidectomy was performed instead of the initially recommended lobectomy for nodules that proved histopathologically benign. In contrast, in just four GEC-positive cases the more aggressive surgery was appropriate and the nodule was histopathologically malignant. Also, in just one patient surgery was withheld specifically due to a negative Afirma® GEC result. In the other unresected GEC-negative nodules surgery was not clinically indicated to begin with; the negative GEC-result merely endorsed conservative management (135). As the GEC was still a new technology when this study was conducted, it is possible that the involved physicians were unsure of the correct interpretation of the GEC results or hesitant to rely on a negative GEC result. However, clinical suspicions and physician and patient preference will always be considered when making surgical decisions. Yang *et al*. elegantly tried to solve the shortcoming (histological) follow-up by comparing their findings of GEC performance to a pre-GEC cohort of similar patients from their hospital in all of whom surgery was performed (11,131). The reported malignancy rates were comparable pre- and post-GEC implementation (18% versus 17%), and obviously relatively more surgeries were performed for benign nodules in the pre-GEC period. Assuming the true malignancy rates in the successively studied populations are indeed similar, the GEC only modestly reduced the number of futile surgeries for benign thyroid nodules from 66% to 52% (131). Altogether, the contribution of the Afirma® GEC to the surgical decision-making may be more limited than expected based on its diagnostic accuracy.

Availability, cost-effectiveness and limitations of the GEC

The Afirma® GEC is currently only available for routine use in the USA. There are high demands for the FNAC specimens regarding sample preservation and shipping. Cytology is revised by Veracyte cytologists and declined if not strictly Bethesda III or IV, with 14% to 17% discordancy between local assessment and central review, comparable to known interobserver rates for thyroid cytology (5,126,135). Reported rates of nondiagnostic GEC test results due to insufficient quantity or quality of the mRNA are substantial, varying from 1% to 17% (130,132). Insufficient mRNA quality was often caused by problems with long duration of the sample shipment to Veracyte (126,130). Fourth, Afirma® GEC testing is expensive and is currently marketed for $3,500 (range $1,750 to $7,000) per test (121,131,137). Testing for medullary carcinoma and *BRAF* mutation is not included in the Afirma® GEC, but can be performed by Veracyte at additional costs (131). Yet, ancillary *BRAF* mutation testing may not be relevant, as Kloos *et al*. found that it improved sensitivity nor specificity of the GEC (57).

Studies of cost-effectiveness yielded variable results, but most concluded that GEC testing would not be cost-effective over conventional surgical management or other diagnostic modalities in various clinical settings (14,121,137-140). The first of these studies proclaimed cost-effectiveness of the GEC even prior to publication of the first validation study by Alexander *et al.*, and has been criticized for several important methodological caveats. This study professedly overestimated test specificity at 75%, overestimated the rate of permanent complications from thyroid surgery, and did not consider the regularly reported GEC test failures (15-17,130,137,138,141). A recent study determined population-dependent thresholds for feasible cost-effectiveness by comparing GEC performance to conventional surgical management in a local Bethesda III/IV population. GEC-guided management was not cost-

effective, adding $1,197 to the $11,119 expenses for conventional treatment while hardly improving QALYs. Sensitivity analysis showed that the GEC would only become cost-effective if its specificity exceeds 71%, if it costs less than $2,640, or if the population malignancy rate decreases from the actual 24% to below 9.2%. This price threshold for cost-effectiveness decreases as the malignancy rate increases, as low as $2,023 per test at 35% cancer prevalence (137).

Furthermore, existing inter-institutional differences in test performance have consequences for local applicability and effectiveness (127,134). Marti *et al*. compared GEC performance in distinct populations of two large hospitals. The reproducibility of the tests' sensitivity and specificity was good, but utility strongly depended on the local prevalence of malignancy: as the population malignancy rate increased, a rarer negative GEC became less reliable to rule out malignancy. Oppositely, at low malignancy rates a negative GEC merely confirmed that the probability of cancer was low. In neither situation, the GEC changed the management strategy. GEC testing was most useful if the malignancy rate ranged between 15% and 21%, comparable to the prevalence reported by Alexander *et al*. (126,134).

Finally, the degree of missing histology is a major limitation to the performed studies. None of the studies following the key publication by Alexander *et al*. had complete histopathologic follow-up; histopathological confirmation ranged between 35% and 82% of specimens (126). Missing histology mainly comprised GEC negative nodules, likely resulting in overestimated sensitivity (i.e. missing some malignancies in the many unoperated GEC-negative nodules) and underestimated specificity (i.e. relatively more GEC-positive nodules with benign histology (false-positives) were operated on than GEC-negative nodules with benign histology (true-negatives)). The trend that studies with higher surgical rates for GEC-negative nodules showed more moderate results supports these hypotheses (126,130,142).

A recent meta-analysis by Santhanam *et al*. included seven studies and reported 96% pooled sensitivity and 31% pooled specificity for the GEC in Bethesda III, IV and V thyroid nodules with histopathological follow-up (143). The authors expected that more than 90% of patients with a negative test would be treated conservatively (143). However, in individual studies up to 25% of patients pursued surgery or conservative treatment despite GEC-based recommendation to do the opposite (127,130). This observation is crucial to cost-utility analyses. In addition, expensive rule-out tests such as the Afirma® GEC should not be performed in case surgery is considered for other reasons, such as cosmetic or mechanical complaints.

GEC in Hürthle cell cytology

Brauner *et al.* specifically validated the Afirma® GEC in 72 cytology samples suspicious for Hürthle cell neoplasm. They demonstrated that GEC testing could accurately have reduced the number of futile surgeries, although through a less profound reduction than in non-oncocytic indeterminate thyroid nodules (132). Similar results were noticed in other studies: despite a relatively low risk of malignancy, the majority of Hürthle cell nodules were GEC-positive. Regardless of good sensitivity, this unfavourable benign call rate in Hürthle cell cytology limits diagnostic efficacy in these nodules, increasing the number needed to test and negatively affecting possible cost-effectiveness (126,131,133,135,142,144). Diagnostic accuracy of the GEC would likely improve if Bethesda IV cytology suspicious for a Hürthle cell lesion was excluded from GEC testing. Otherwise, similar to the additional testing for medullary carcinoma, adaptations should be made to the Afirma® GEC to improve its clinical utility for Hürthle cell lesions.

In conclusion, it is generally assumed that the Afirma® GEC accurately reclassifies approximately two out of five indeterminate thyroid nodules as benign with published sensitivities ranging between 83% and 100% and similar test performance in Bethesda III and IV nodules. Withholding diagnostic surgery from these patients seems safe (130,131,134,144). However, the diagnostic strength and potential cost-utility of Afirma® GEC strongly rely on its NPV – thus on the prevalence of malignancy and benign call rate in the targeted population. There are important concerns regarding the currently insufficient number of clinical validation studies with adequate rates of histopathological confirmation or long-term clinical follow-up. Physicians are strongly advised to locally validate Afirma® GEC test performance before considering test implementation in daily practice. Nonetheless, further large validation studies on the Afirma® GEC may soon become obsolete, as an updated version of the test, the Gene Sequencing Classifier (Veracyte Inc., South San Francisco, CA, USA), is currently being put into operation. Improved diagnostic accuracy is anticipated, with specific attention to the differentiation of Hürthle cell nodules.

1.2. MicroRNA

First described in thyroid cytology in 2006, evaluation of the expression levels of microRNA (also called miRNA) is among the newer and more promising approaches to differentiate between benign and malignant thyroid neoplasms (145,146). MicroRNAs are small endogenous noncoding ribonucleic acids (RNAs) of approximately 22 nucleotides in length. As negative regulators (i.e. silencers) of protein synthesis at a post-transcriptional level, they are involved in many intracellular processes, including cell growth, differentiation and proliferation. Dysregulation of microRNA expression is found in almost all types of human

cancers (147). It reflects the deregulated expression of oncogenes and tumor suppressor genes (146,148-150). MicroRNA overexpression is present before morphological tissue changes are seen and therefore considered to be a part of premalignant changes in carcinogenesis (145). MicroRNA expression profiles are tissue-specific and can not only identify the tissue of origin, but also the histopathological subtype of the cancer and whether it concerns the primary tumor or a metastasis (148,151).

MicroRNA expression profiles are similar among the various types of thyroid carcinoma, even though expression levels are often distinctively different (148). In histopathological studies, PTC was associated with an up to 11- to 19-fold upregulation of miR-146b, miR-221, miR-222, miR-181b, miR-187, and a downregulation of miR-1 and miR-138 compared to healthy thyroid tissue and benign nodules. Upregulation of miR-221, miR-222 and miR-187 was also found in FTC, FTC-OV, poorly differentiated and anaplastic carcinoma (145,146,148,152,153). Overexpression of miR-146b-3p, miR-146b-5p and miR-375 was seen in both PTC and FVPTC (152,154). Furthermore, expression levels of miR-221 and miR-222 were reported about twice as high in FVPTC as compared to PTC or FTC (152). Only a few microRNAs were differently expressed between follicular neoplasm and FTC (155). Follicular adenoma was associated with the expression of miR-200a, whereas high expression of miR-31 was found in Hürthle cell adenoma (148). FTC is related to the differential expression of miR-146b, miR-7-5p, miR-346, miR-197 and miR-21, but results among studies are more heterogeneous (148,155,156). FTC-OV showed an expression pattern slightly similar to FTC, but also distinct overexpression of other microRNAs, such as miR-339, miR-183, miR-197 and miR-885-5p (148,153).

Accordingly, a diagnostic panel of a carefully selected combination of microRNAs and appropriate expression levels could aid in the preoperative distinction of indeterminate thyroid cytology (157). Recent meta-analyses struggled to reconcile the studies on microRNA

in FNAC, as the investigated set of microRNAs was never identical and individual microRNA performance was infrequently described. In unselected cytology, estimated sensitivity of microRNA expression analysis ranged from 75% to 78% regardless of the investigated set; estimated specificity from 73% to 81% (156-158).

In indeterminate thyroid cytology, different sets of microRNAs were evaluated; only several individual microRNAs were analyzed in more than one study. The selected microRNAs were first assessed in a test set of cytological and/or histopathological specimens and a cut-off for their expression level was determined. Subsequently, the significantly up- or downregulated microRNAs were validated in an independent set of (indeterminate) thyroid FNAC samples. Some studies developed a decision model for the validation step (149,154,159).

The most promising results were presented by Keutgen *et al.* (159). Of the six microRNAs investigated in their test set, miR-21, miR-146b, miR-181a and miR-222 were differentially expressed in malignant nodules with prior indeterminate cytology. The subsequently developed support vector machine model incorporated miR-21, miR-222 and the insignificantly expressed miR-197 and miR-328. Prospective validation in an independent set of 72 indeterminate FNAC samples resulted in 100% sensitivity and 86% specificity. Five of the seven false positives had Hürthle cell cytology; excluding these, raised specificity to 95% (159). Notably, even though overexpression of miR-146b is often related to thyroid carcinoma, it proved not useful to Keutgen *et al.* to include in their prediction model (159). In contrast, Agretti *et al.* and Shen *et al.* included miR-146b as the key differentiators in their models. Agretti *et al.* assessed a frequently quoted set of microRNAs consisting of miR-146b, miR-155, miR-187, miR-197, miR-221, miR-222 and miR-224 (148,149). Published in 2008, Nikiforova *et al.* had demonstrated that this 7-microRNA set in FNAC samples had 100% sensitivity and 94% specificity if one of the included microRNAs showed an at least two-fold overexpression (148). Analytic validation of this model by Agretti *et al.* showed differential

upregulation in PTC of all of these microRNAs except miR-197. In particular, miR-146b showed a >30-fold higher expression in PTC. A decision tree including miR-146b, miR-155 and miR-221 was 98% accurate in the test set, but validation in an independent set of indeterminate FNAC samples was unsuccessful, yielding mere 60% sensitivity and 58% specificity (149).

Vriens *et al.* used a microRNA array to detect 10 genes that were up- or downregulated by ≥5-fold in thyroid malignancies. Four microRNAs (miR-100, miR-125b, miR-138 and miR-768-3p) were significantly downregulated and accurately differentiated between benign and malignant follicular and Hürthle cell neoplasms in the test set. In their validation set of 125 indeterminate FNAC samples, only miR-138 was moderately distinctive with 81% NPV. For Hürthle cell carcinoma, miR-138 and miR-768-3p were both 98% accurate (160).

Finally, in a recent Italian study only miR-375 accurately differentiated between benign and malignant neoplasms. Subsequently, in TIR3 cytology excluding Hürthle cell lesions, a 12-fold or higher overexpression of miR-375 perfectly distinguished benign from malignant lesions with 100% accuracy. It was also significantly differently expressed between TIR3A and TIR3B categories and correlated with a different malignancy risk (161).

Availability and limitations of microRNA expression analysis

MicroRNA expression analysis has advantages over other techniques. MicroRNAs are more stable than mRNA at maintaining their expression in formalin-fixed paraffin-embedded (FFPE) tissue samples as well as FNAC specimens, irrespective of the preservation method (e.g. archived FNAC slides or nucleic acid preservation solutions) (148,161). Recently microRNA expression was even successfully measured in serum (162). Moreover, microRNA expression levels measured with generic methods (e.g. quantitative RT-PCR) correspond well

to their biological effect, as microRNAs affect biological processes without the additional step of protein synthesis (148).

However, general limitations of FNAC also translate to concerns with microRNA analysis: scant cellularity or low levels of malignant cells in FNAC specimens could cause a false-negative microRNA test result (149). Another limitation is the plurality of microRNAs associated with DTC in histopathological studies, causing vast heterogeneity between the limited number of studies in indeterminate cytology. Validation studies of the same microRNA set are lacking. Simultaneously, new microRNAs are still correlated to thyroid carcinoma. Ongoing research has yet to compose the optimal set of microRNAs. Recently, the first commercial test was marketed as the ThyraMIR™ (Interpace Diagnostics, Parsippany, NJ, USA). It evaluates the expression levels of miR-29b-1–5p, miR-31–5p, miR-138–1-3p, miR-139–5p, miR-146b-5p, miR-155, miR-204–5p, miR-222–3p, miR-375, and miR-551b-3p. The ThyraMIR™ demonstrated 57% sensitivity and 92% specificity in 109 Bethesda III and IV FNAC specimens (22). Prospective clinical validation of the ThyraMIR™ could affirm the diagnostic value of microRNA expression profiling in indeterminate thyroid nodules in the pre-operative setting.

1.3. Immunocytochemistry

Tissue characterization through selective staining of expressed proteins, i.e. immunohistochemistry (IHC), is a technique that combines histopathology and biochemistry. Exploiting basic antigen-antibody interactions, IHC is able to visualize the distribution and localization of specific cellular components within the cell and in the proper tissue context. This includes tissue biomarkers specific for e.g. infection or malignancy. IHC has been fully incorporated in the histopathological routine and is crucial to morphological and molecular

tissue characterization. When immunocytochemistry (ICC) – the application of this immunology-based technique in cytology – became available, the possibilities were extended to the preoperative setting, too. Specific immunomarkers have been developed to differentiate between benign and malignant thyroid nodules. The 2015 ATA guidelines acknowledge ICC as a technique under development with limited prospective validation studies in indeterminate cytology (23). In unselected thyroid cytology, the much-used immunomarkers galectin-3, Hector Battifora mesothelial-1 (HBME-1) and cytokeratin 19 (CK-19) demonstrated 85%, 83% and 80% sensitivity, and 90%, 79% and 79% specificity, respectively (163).

1.3.1. Galectin-3

Galectin-3 is a β-galactosyl-binding protein from the lectin group. It is involved in cell-cycle regulation, including cell migration and adhesion. Its exact function is still to be unraveled, but a role in the pathogenesis and progression of PTC is presumed (44,163-165). It is related to inhibition of apoptosis, induced by abnormal p53 expression (165). Galectin-3 can be present both in the intracellular as well as the extracellular matrix (166). Normal thyrocytes do not express galectin-3, but the physiological expression of galectin-3 in macrophages, neutrophils, mast cells and Langerhans cells provides an internal positive control of the investigated FNAC samples (167,168). Positive cytoplasmic staining – as opposed to nuclear staining – for galectin-3 is suspicious for malignancy and mainly associated with PTC (117,169,170). Galectin-3 expression has also been associated with the malignant transformation of follicular neoplasms, as it was present in follicular adenoma as well as FTC (166,171,172). Encapsulated FVPTC and minimally invasive FTC showed less frequent and weaker staining (172,173).

In 2001, Bartolazzi *et al*. argued that galectin-3 staining could accurately diagnose thyroid carcinoma in unselected thyroid cytology (117). Subsequent studies in indeterminate thyroid cytology mostly could not reproduce these promising results. With a positive stain in approximately a third of all nodules, sensitivity and specificity of galectin-3 ranged from 0% to 92% and from 68% to 100%, respectively (44,174-177). Merely Saggiorato *et al*. demonstrated that galectin-3 accurately differentiated follicular adenomas from FTC with 92% sensitivity and 94% specificity if a cytoplasmic stain in ≥10% of the cells was considered positive (174). The prospective multicenter clinical validation study by Bartolazzi *et al*. demonstrated 78% sensitivity and 93% specificity in Thy3 nodules if a cytoplasmic galectin-3 stain in >5% of the cells was considered positive. Nineteen of the 22 false-positive nodules were follicular adenoma. However, a group of 33 difficult-to-diagnose (follicular) tumor of unknown malignant potential lesions was disregarded, 22 of which were galectin-3 negative. If these neoplasms were considered malignant, sensitivity dropped to 69% (164).

1.3.2. HBME-1

HBME-1 is a monoclonal antibody targeting an unknown antigen on the microvilli of mesothelial cells. It is usually negative in normal thyroid follicular cells. Abnormal expression of HBME-1 shows cytoplasmic location with membrane accentuation. It is associated with, but does not necessarily indicate PTC (165,169,178,179). Its low detection limit enables assessment in liquid based cytology (180). Reported sensitivity and specificity of HBME-1 in indeterminate nodules ranged from 61% to 100% and from 75% to 96%, respectively (174,176,180,181). Approximately two out of five nodules showed positive staining. If only non-oncocytic follicular neoplasms were selected, Saggiorato *et al*.

demonstrated that HBME-1 had excellent 93% sensitivity and 98% specificity in indeterminate thyroid nodules (174).

1.3.3. Cytokeratin 19

Cytokeratin 19 (CK-19) is a type I keratin. It belongs to the group of intermediate filament proteins, which arrange the cell cytoskeleton and structural integrity. CK-19 is widely present in epithelial cells, but also found in basal cells layers of stratified epithelium (174,182). Strong and diffuse abnormal expression of CK-19 indicates PTC, including FVPTC. Expression in FTC is less intense and more variable, warranting nuanced interpretation of CK-19 staining intensity. CK-19 usually shows no or only focal expression in follicular neoplasms, hyperplastic nodules and adenomatous goiter (174,178,182,183). The reported sensitivities and specificities for CK-19 staining in indeterminate cytology ranged from 76% to 88% and 80% to 100%, respectively (174,180,182). Lacoste-Collin *et al*. demonstrated the importance of an accurate threshold. CK-19 staining in 31 Bethesda IV nodules accurately diagnosed five out of six malignancies, including a PTC, two FTC and two out of three FVPTC. At a threshold of ≥30% stained cells, 5 of 25 benign lesions tested false-positive; at a more sensitive threshold of ≥10% stained cells, 12 of 25 tested false positive (180).

1.3.4. Other immunocytochemistry markers

Immunohistochemistry studies identified more potential ICC markers. Some, like CD44v6, have not yet been investigated in indeterminate cytology (117). Other markers were sporadically investigated in preclinical studies, including Ki-67, TROP-2, emerin, keratan sulphate, thyroperoxidase, CD57 and GLUT-1.

Nuclear protein Ki-67 is expressed in nearly all cell cycle phases in proliferating cells. It is associated with poor prognosis in PTC (184). The percentage of cells with Ki-67 expression is considered the tissues' proliferative index. At a cutoff of ≥1% Ki-67 was 85% sensitive and 71% specific for thyroid carcinoma in Bethesda IV nodules. A combination of HBME-1, CK-19 and Ki-67 immunomarkers was 91% accurate to diagnose malignancy (180). Ki-67 expression is likely only distinctive for follicular type carcinoma; expression in PTC is generally low (180,184,185).

Glycoprotein human trophoblast cell surface marker (TROP-2) is overexpressed on the cell surface of different epithelial carcinoma (e.g. breast, colon) and associated with tumor aggressiveness and poor prognosis. In indeterminate thyroid cytology, it was only assessed in one small subseries of Bethesda III samples, correctly diagnosing the three included carcinoma and all but one of the nine benign nodules (186).

Emerin staining emphasizes features of the nuclear membrane often seen in PTC, such as irregularities and invaginations. Consequently, the stain could facilitate the morphological diagnosis of PTC and especially the more difficult-to-diagnose FVPTC (176,187). In 53 Thy3 nodules assessed by Asioli *et al*., positive emerin staining was highly specific for PTC (including FVPTC), but misdiagnosed all FTCs (176).

Another immunomarker associated with PTC is keratan sulphate, an abnormal glycosaminoglycan complex. It was 98% specific in indeterminate cytology, but correctly predicted PTC only; its sensitivity was poor at 48% (174).

The expression of thyroid peroxidase (TPO) is related to benign follicular neoplasms. A negative TPO stain was 80% sensitive and 86% specific for thyroid malignancy (174).

Finally, CD57 (Leu7) expression is associated with epithelial and nonepithelial malignancies, including thyroid carcinoma. Cytological staining was only investigated in a small series of

indeterminate cytology, but seemed specific for PTC. In the same series, GLUT-1 was not a useful ICC marker – there were no positive stains (188).

Combined use of immunocytochemistry markers

Some research groups have suggested that evident single-marker galectin-3 positivity is sufficient to refer a patient for total thyroidectomy (164,170,177). The ATA guidelines did not adopt these suggestions, and many other researchers advocate that a panel of ICC markers should be applied to strengthen the suspicion of malignancy (23,165,174,179). Several panels were investigated in literature. Zhang *et al*. assessed a triple stain of galectin-3, HBME-1 and p27. P27 is a cyclin-dependent kinase inhibitor related to cell life span in normal thyroid cells. Downregulated in malignancy, positive P27 stain is related to benign histopathology. In a set of Bethesda III cytology samples, positive p27 staining with negative galectin-3 and HBME-1 staining was 100% predictive of a benign nodule and occurred in 38% of samples. Loss of p27 staining in combination with positive galectin-3 and/or HBME-1 staining was 100% sensitive and 86% specific (165). Another study investigated galectin-3 and HBME-1 in combination with a RET proto-oncogene stain, which reflects abnormal intracellular RET proto-oncogene activity and presence of the RET/PTC rearrangement. Unfortunately, RET staining was inaccurate in indeterminate thyroid nodules (181).

To find the most accurate combination of immunostains, Saggiorato *et al*. explored the expression of galectin-3, HBME-1, thyroperoxidase, CK-19 and keratan-sulphate in 125 cytological follicular neoplasms, 24 of which were Hürthle cell lesions. Galectin-3 was not only the most accurate marker individually, but also in combination with other stains. Sequential HBME-1 staining of galectin-3-negative cases reached 98% sensitivity and 98% specificity in non-oncocytic lesions. In oncocytic lesions, sequential CK-19 staining was more preferred with 100% sensitivity and 100% specificity (174).

The common denominator between all these studies is the combined use of galectin-3 and HBME-1. Unfortunately, clinical validation studies regarding this combination are limited. Its seemingly promising diagnostic accuracy warrants further assessment in future prospective studies.

Performance of immunocytochemistry in Hürthle cell cytology

Expression of ICC markers in Hürthle cell nodules differs from non-oncocytic indeterminate cytology. Hürthle cell carcinomas were distinguished in the cytological samples by typical overexpression of markers associated with a high degree of cell proliferation, disorganized tissue structure and intermediate differentiation, such as Ki-67, laminin, cyclin D1 and cyclin D3. Overexpression reflects the known more erratic behavior of Hürthle cell carcinoma (185,189). Moreover, markers that were highly diagnostic in indeterminate nodules in general, also seem differently expressed in Hürthle cell lesions. Saggiorato *et al.* demonstrated that two combinations of ICC markers were extraordinarily accurate: galectin-3 and CK-19 staining was 100% sensitive and 100% specific; galectin-3 and thyroperoxidase staining was 100% sensitive and 85% specific (174).

In a previous meta-analysis, inclusion of Hürthle cell lesions was related to between-study heterogeneity (163). Hürthle cell lesions require a biotin-free ICC method, as Hürthle cells themselves are rich in biotin. Thus, much-used biotin-based methods may consequently cause false positive and highly intensive staining in Hürthle cell neoplasms (166,172,181).

Availability, cost-effectiveness and limitations of immunocytochemistry

Current application of immunocytochemistry is limited. Clinical validation studies for all of the described immunomarkers are scarce, and no cost-effectiveness studies are available to date. Yet, the technique is widely available, relatively inexpensive and fast in comparison to

other (molecular) techniques. Costs per immunostain vary up to €20, partly depending on simultaneous local application of the technique and similar stains for immunohistochemistry. Immunocytochemistry is preferably performed on cell block FNAC specimens, but can be performed in all types of cytology, from direct smears to liquid-based cytology (179,181). ICC is impossible when the FNAC specimen has poor cellularity or too much obscuring blood (190). Also, immunostaining of cytology is technically more difficult than histological staining, especially in (destained) cytology smears. Technical inconsistency and interobserver variation likely lead to false-negative results (164,182). Stain intensity thresholds or percentage of stained cells necessary to raise suspicion of malignancy vary in the available literature. Consistent methodology and assessment thresholds should be determined to improve reproducibility of ICC results.

Clinical validation studies of existing ICC markers are ongoing. Meanwhile, new markers are also playing the field, searching for the interfaces between mutation analysis of highly specific oncogenic driver mutations and accessible ICC techniques. For example, Leslie *et al*. investigated ICC of the $BRAF^{V600E}$ mutation using the mutation specific antibody VE1 in a small series of thyroid FNAC samples. Concordance between ICC and conventional $BRAF^{V600E}$ mutation analysis was 85%. All samples that were $BRAF^{V600E}$ positive by either method were confirmed as $BRAF^{V600E}$ positive PTC on histopathology. Of the eight included indeterminate thyroid nodules, seven were histopathologically malignant and $BRAF^{V600E}$ mutation was detected in two nodules: one by both methods, one only by molecular analysis. The $BRAF^{V600E}$ specific antibody (VE1) stain was much weaker in cytology than in histology. Moreover, costs of the VE1 antibody are currently high and optimization of methodology is warranted. Yet, Leslie *et al*. demonstrated that $BRAF^{V600E}$ mutation analysis using ICC is a promising alternative to mutation analysis (79). If future studies could validate these results in larger cohorts of indeterminate thyroid nodules and detect reliable immunomarkers for other

oncogenic driver mutations, this technique unites the strengths of gene mutation analysis and immunocytochemistry in one technique, though likely at lower costs.

In general, ICC is a widely available and relatively inexpensive technique with a reasonable diagnostic accuracy. Many immunomarkers seem to have a pronounced association with PTC. Galectin-3 and HBME-1 were most frequently investigated, but their specificities and sensitivities seem to fall short of justifying ICC-based surgical decision making. Diagnostic accuracy of their combined use seems promising, yet current evidence is limited. Prospective validation trials are warranted to confirm the diagnostic potential of ICC, including validation of thresholds for stain positivity, panels of multiple immunostains and other methodology.

2. CONVENTIONAL IMAGING

2.1. Ultrasound

Ultrasound (US) is one of the principal steps in the initial work-up of thyroid nodules. It is cheap, fast, non-invasive and globally available, but accurate assessment strongly depends on operator experience (191). Multiple meta-analyses showed that well-known US features such as nodule hypoechogenicity, microcalcifications, irregular margins (including microlobulated or ill-defined margins), and a taller-than-wide shape raise the suspicion for thyroid malignancy and are mostly associated with PTC (191,192). Nonetheless, no single US feature is sufficiently sensitive nor specific to accurately identify a malignant nodule in an unselected population (191). Certain combinations of US features, however, may offer accurate closure. The current ATA guidelines now include a flowchart recommending FNAC dependent on nodule size and various combinations of US characteristics with an incremental risk of

malignancy (23). Despite the obvious importance of both ultrasound and cytology, the ATA guidelines do not provide recommendations regarding (re-)interpretation of US characteristics after FNAC has resulted in indeterminate cytology. Follicular-type malignancies typically have a different US appearance. More often FTC may be iso- to hyperechoic, with a spherical shape, smooth regular margins and no calcifications (193,194). FVPTC may also show FTC-like or benign features rather than the classic suspicious features, although microcalcifications may be distinctive (194-196). In the past years, Brito *et al*. and Remonti *et al*. performed meta-analyses on US assessment of unselected thyroid nodules. Both also briefly discussed its diagnostic value in indeterminate nodules, including a mere limited number of studies and also including cytology suspicious for malignancy. Increased central vascularization was most predictive of malignancy with reported 96% specificity (192). Yet, in general US seemed less accurate in indeterminate nodules than in unselected thyroid nodules (191,192).

In the dozens of available original ultrasound studies, individual US features generally demonstrated limited sensitivity in indeterminate thyroid nodules. Only the appearance of a solid thyroid nodule – as opposed to varying degrees of cystic content – had high sensitivity. Ranging between 46% and 100%, multiple studies demonstrated sensitivity above 90% (18,19,197-201).

A number of classic suspicious US characteristics, such as a taller-than-wide shape, presence of irregular margins and presence of microcalcifications, demonstrated valid specificity in indeterminate thyroid nodules. Specificities for each of these characteristics ranged from 72% to 99% (201-204), 65% to 100% (202,205,206) and 36% to 100%, respectively (207,208). Despite the wide range, presence of microcalcifications was more than 90% specific in many studies (197,198,200,202-204,206,207,209-211). Large nodule size (defined as a diameter

larger than 4 cm) was only investigated in a limited number of studies. Reported specificities ranged between 69% and 94% (207,212).

Other features, such as a solitary nodule, hypoechogeneity and absence of a hypoechoic halo were associated with thyroid malignancy, but less accurately differentiated between benign and malignant indeterminate thyroid nodules (198,201,206,212-217). Additionally, opposing the results from one of the mentioned meta-analyses, central vascularization also does not seem very accurate in indeterminate thyroid nodules. Specificity ranged from 0% to 100%, although multiple studies demonstrated extremely poor specificity (18,202,216-220).

Results regarding two US features are remarkably contradicting. First, the absence of a hypoechoic halo is typically considered suspicious for malignancy, but showed overall poor and very heterogeneous diagnostic potential in indeterminate thyroid nodules (191). Sensitivity and specificity ranged from 17% to 99% and 0% to 93%, respectively (200,201,205,221,222). Presence of a hypoechoic halo is typically considered a benign feature, but has also been associated with follicular types of thyroid carcinoma (223). Dogan *et al*. reported 88% specificity for presence of a halo in AUS/FLUS nodules and 78% in FN/SFN nodules (224). Second, the ultrasonographic nodule shape seems ambiguous. Similar to the unselected population, a typically suspicious taller-than-wide shape was generally specific for carcinoma, with reported specificities up to 99% (201,202,204). A spherical shape is generally considered benign, but has also been associated with FTC (191,193,225). In two studies in cytological follicular neoplasms, a spherical shape had an increased risk of malignancy, with 86% to 97% sensitivity and 19% to 26% specificity (226,227). Chin *et al*. even suggested that follicular neoplasms with a taller-than-wide shape could be treated conservatively (227). The uniquely balanced rates of PTC, FVPTC and FTC resulting from indeterminate cytology may explain why these and various other US characteristics have

different diagnostic accuracy than in the unselected population. Dependent on the local case mix, accurate differentiation of indeterminate nodules using the classical suspicious US features may or may not be feasible.

Combination of ultrasound characteristics

A combination of US characteristics likely provides more accurate differentiation than individual features. Different combinations were investigated in multiple studies (81,197,198,200,204-206,216,228-233). Yoo *et al*. reported 100% specificity for the combination of marked hypoechogenicity and taller-than-wide shape, a pattern that occurred in 9.6% (24/249) of the included Bethesda III nodules (201). In the elastosonography study by Rago *et al*., absence of a hypoechoic halo in combination with presence of microcalcifications was 95% specific for thyroid malignancy, but only 6.4% sensitive (222). Maia *et al*. found 62% sensitivity and 89% specificity in Bethesda III and IV nodules if hypoechogenicity, microcalcifications, an irregular margin and increased intranodular vascularity were considered suspicious (230). Gulcelik *et al*. demonstrated that the US pattern of a solid, hypoechoic nodule with microcalcifications had 95% sensitivity and 99% specificity. The pattern was seen in 21% of cytological follicular neoplasms (234).

In multiple studies it was argued that cytological follicular neoplasms with a typically benign ultrasound pattern – a regular shape, isoechoic, homogeneous, with well-defined margins, cystic components or peripheral vascularity only, and not a single malignant feature - could be safely followed up clinically instead of undergoing diagnostic surgery (200,225,228). Consideration of more features generally increased the sensitivity of the US assessment at the cost of its specificity (198,199). The terms of their interpretation were crucial: Norlén *et al*. demonstrated that US was 95% sensitive and 48% specific if a Bethesda III nodule had either hypoechoic appearance, irregular margins or microcalcifications. If solely the simultaneous

presence of all three features was considered suspicious for malignancy, sensitivity dropped to 37% but specificity increased to 96% (231).

Altogether, diagnostic ultrasound scores or step-by-step algorithms could aid the classification of US patterns and consequent risk of malignancy (200,228,235). Best-known and most validated is the Thyroid Imaging Reporting and Data System (TIRADS), a classification to risk-stratify thyroid nodules, designed by Horvath *et al*. and modified by Kwak *et al*. following the example of the similar BIRADS classification for breast lesions (236,237). The TIRADS assigns nodules to a risk category based on five suspicious US features: solid appearance, (marked) hypoechogenicity, irregular margins, microcalcifications and a taller-than-wide shape. Nodules without any of these features are likely benign and categorized as TIRADS 3. Their risk of malignancy is ~1.7% in a cytologically unselected population. TIRADS 4 includes suspicious nodules, which are further classified according to an increasing malignancy risk into 4a (one suspicious US feature), 4b (two suspicious features) and 4c (three or four suspicious features). Nodules with all five suspicious US features are classified as TIRADS 5 and associated with a high 88% risk of cancer in an unselected population (236). Studies that validated the TIRADS specifically in indeterminate thyroid nodules, showed that diagnostic accuracy depended on the chosen cutoff score and type of cytology (202,235,238-241). Although TIRADS 5 scores were infrequently assigned in indeterminate nodules, a higher TIRADS score (4b/4c/5) was an accurate predictor of malignancy, especially in Bethesda IV cytology (202,238,240). In Bethesda III nodules, lower TIRADS scores (3/4a) could also rule out malignancy (238,241). Prospective validation studies applying the TIRADS in indeterminate cytology are warranted to assess its possible clinical utility in indeterminate nodules.

US performance in Hürthle cell nodules

Cytological Hürthle cell nodules expressed a large variation of US characteristics (203,208,212,223). Many malignant and most benign Hürthle cell nodules had a benign US appearance (208,223). Only three US features possibly predictive of malignancy were reported in individual studies: both hypoechogenicity and hyperechogenicity (as opposed to isoechogenicity) (208), large nodule size (223), and microcalcifications (203). Despite limited evidence, US evaluation does not seem reliable to differentiate Hürthle cell lesions.

Availability and limitations of ultrasonography

The major advantages of ultrasound over other additional diagnostics are its already permanent position in the workup of thyroid nodules, global availability and low costs. No additional resources nor hospital visits are needed to include US interpretations in preoperative management decisions and the investigation is noninvasive. Nonetheless, besides known limitations concerning interobserver variability and less reliable interpretation of small nodules, US feasibility in indeterminate thyroid nodules is limited by the presumed differences in US appearance of papillary and follicular thyroid malignancies, illustrated by the conflicting results for nodule shape and hypoechoic halo in indeterminate nodules. Consequently, local diagnostic accuracy likely follows variations in the local histopathological case mix.

In addition, many of the available ultrasound studies are retrospective, limiting the power of the evidence. As the decision to perform FNAC is customarily based on the results of the prior US, the prevalence of suspicious US features in indeterminate cytology in these studies is presumably overestimated.

Nonetheless, several individual US characteristics seem to have reasonable specificity in indeterminate nodules, although insufficient for accurate diagnosis. A combination of US

features is likely more accurate, although current evidence does not support US-based surgical decision-making. We propose that a future meta-analysis should use the individual patient data from the large number of available original studies to develop an ultrasound algorithm specifically for indeterminate thyroid nodules. The existing TIRADS needs prospective validation.

Even though more advanced and less operator-dependent techniques might be preferred, US features should always be assessed in current clinical practice. The presence of one or more suspicious US features in a Bethesda III or IV nodule increases the suspicion of malignancy and underpins the need for a definite diagnosis. Moreover, to centers or regions with limited access to other (molecular) diagnostics, ultrasound may definitely have clinical utility, pending local validation in the indeterminate population.

2.2. Elastosonography

Firm consistency of a thyroid nodule upon palpation is considered suspicious for malignancy – an established principle during physical examination (242). Ultrasound elastosonography (USE) is a dynamic ultrasound technique that is sometimes referred to as 'electronic palpation'. Tissue elasticity is evaluated by measuring tissue distortion while applying a standardized dosed external force by the US transducer. It was first applied to the thyroid gland by Lyshchik *et al*. in 2005 (243). Classic real-time qualitative USE is performed by free-hand compression and a sine-wave or numerical scale showing how much pressure the operator applies with the probe. A color-coded elastosonography image is superimposed on the grey-scale US images: red and orange visualizes high tissue elasticity (soft tissue), green represents intermediate elasticity and blue low elasticity (firm tissue). Several score systems are available. The original score was developed by Itoh *et al*. in 2006 for the evaluation of

breast tumors and considers scores 1-3 benign on a scale of 1 (highest elasticity) to 5 (no elasticity) (244). Rago *et al*. first applied it to thyroid tumors and modified it to a 3-point score (222,245). Asteria *et al*. derived a modified 4-point score (246).

The earliest studies in thyroid nodules reported opportune results of USE as an additional modality to B-mode ultrasound, but were heterogeneous in USE technique and study population (247). A recent meta-analysis by Nell *et al*. included twenty studies on qualitative USE prior to FNAC and concluded that qualitative USE is fit to diagnose benign nodules and safely dismiss FNAC, provided that the usual elasticity score cutoff is abandoned and only completely soft nodules (score 1 of all systems) are classified as benign. Pooled 99% sensitivity and 99% negative predictive value demonstrated the ability of USE to reliably rule-out malignancy in entirely soft thyroid nodules, composing 14% of their pooled study population (248).

In individual studies on USE in indeterminate nodules, sensitivity and specificity of qualitative USE ranged from 47% to 97% and from 6% to 100%, respectively (19,213,221,222,249). Results of several qualitative USE studies stand out. Lippolis *et al*. showed an aberrant 6.1% specificity, because they reported only eight nodules with high elasticity – 62 of 66 benign nodules were not elastic. The authors themselves suggest that a rather homogenous study population with predominantly small nodules with a solid US pattern, absence of cystic areas, and follicular histology with minimal colloid could be explanatory for the poor specificity rather than operator-dependent causes (213). Such possible relations remain undescribed in other studies. A meta-analysis on the value of USE in indeterminate thyroid nodules demonstrated meager pooled 69% sensitivity and 75% specificity (250).

As manually applied pressure is difficult to standardize, qualitative USE is strongly operator dependent (251). Different USE techniques have been developed to improve objectivity, such as semi-quantitative tissue-to-nodule strain ratio indices (also based on manual compression). Studies investigating semi-quantitative USE in indeterminate thyroid nodules reported sensitivity and specificity ranging from 82% to 100% and from 88% to 100%, respectively (215,217,219,249). Furthermore, quantitative shear wave USE measures the propagation velocity of focused acoustic pulses – shear waves – from the probe, which correlate to tissue stiffness (Young's modulus) (18,252). It had 82% sensitivity and 88% specificity in a recent prospective pilot study by Samir *et al*.(18). Performance of (semi-)quantitative USE seems better than qualitative USE, but results are subject to overfitting from the ROC analysis performed to determine the strain ratio cutoff value with the highest sensitivity and specificity. None of the studies applied a predefined cutoff or validated their own cutoff externally. Consequently, the resulting thresholds were hardly comparable (215,217,219,249).

Altogether, the results from currently available studies cannot support surgical decision-making in thyroid nodules with indeterminate cytology using elastosonography in any of its forms. Whereas color-coded qualitative USE has insufficient sensitivity and specificity, the semi-quantitative method lacks validation. The power of the available evidence is additionally limited by both methodological heterogeneity and the use of different USE techniques, image processing and elasticity scoring methods across studies. Nevertheless, the suggested promising rule-out capacity of qualitative USE when applying an alternative cutoff score of 1 in unselected nodules, deserves clinical validation in indeterminate thyroid nodules. Major advantages of the technique are the minor extra costs of USE, as it can be performed during regular thyroid US with the same equipment, and only adds approximately 5 minutes to the

procedure time per patient. Cost-effectiveness will largely depend on performance of USE, but no cost-effectiveness studies in indeterminate thyroid nodules are available to date.

2.3. Computed Tomography

There are no studies that investigated computed tomography (CT) scanning in thyroid nodules with indeterminate cytology. Prior studies indicated that CT cannot accurately differentiate thyroid carcinoma (253,254).

3. FUNCTIONAL AND MOLECULAR IMAGING

3.1. ^{99m}Tc-MIBI

Hexakis(2-methoxy-2-methylpropylisonitrile)technetium[^{99m}Tc] (^{99m}Tc-MIBI) is a Technetium-99m-labeled radiopharmaceutical, primarily known for its use in myocardial perfusion imaging since the 1980s and more recently the evaluation of hyperparathyroidism. Uptake of ^{99m}Tc-MIBI, a lipophilic cation, reflects both perfusion and the number of active mitochondria in the cells of the thyroid nodule and thus its oxidative burden (70,255). ^{99m}Tc-MIBI scintigraphy is more suitable for the differentiation between benign and malignant thyroid nodules than scintigraphy with ^{99m}Tc-pertechnetate ($^{99m}TcO_4^-$) or radioisotopes of iodide (often $^{131}I^-$,$^{123}I^-$ or $^{124}I^-$). These latter tracers interrogate the sodium-iodide symporter of the thyrocyte and are frequently used to assess thyroid nodule functioning to distinguish autonomous (“hot”) from hypofunctioning (“cold”) nodules. They are neither specific nor effective to detect malignancy: benign nodules can be anything from hyper- to hypofunctioning, and far outnumber the carcinomas. Still, thyroid malignancies are almost

always hypofunctioning: decrease of the sodium-iodide symporter or thyroid peroxidase are hallmarks of cell dedifferentiation and lead to loss of iodide-trapping function and thus ^{99m}Tc-pertechnetate or radioiodine uptake (23,255-257). ^{99m}Tc-MIBI uptake is independent of iodide trapping and organification in the thyrocytes. Nodules with increased uptake and late retention of ^{99m}Tc-MIBI are suspicious for malignancy (70,255). A 2013 meta-analysis by Treglia *et al*. demonstrated 82% sensitivity and 63% specificity for ^{99m}Tc-MIBI scintigraphy in clinically suspicious, hypofunctioning, cytologically unselected thyroid nodules. Hyperfunctioning benign adenomas can show false-positive increased uptake of ^{99m}Tc-MIBI due to their increased metabolic needs, thereby decreasing test specificity (255).

Only three studies investigated the role of ^{99m}Tc-MIBI in indeterminate thyroid nodules. In all studies, evaluation of thyroid nodules was performed by dual-time planar imaging: an early image was made ranging from 10-20 minutes after injection of the radiopharmaceutical and a delayed image 60-120 minutes post injection. The intensity of the ^{99m}Tc-MIBI uptake within the nodule, and possible increased uptake or denoting retention on delayed imaging were assessed and compared to the physiological washout of the tracer from normal thyroid tissue.

A visual pattern of increased ^{99m}Tc-MIBI uptake on early images that persisted or further increased on the delayed images was generally considered suspicious for malignancy. The individual study sensitivity and specificity for this interpretation ranged from 56% to 79% and from 52% to 96%, respectively (19,70,258). Despite the limited number of available studies, the performance of ^{99m}Tc-MIBI in indeterminate thyroid nodules seems insufficient and less accurate than in cytologically unselected nodules (255).

Nonetheless, Giovanella *et al.* demonstrated that NPV for this method could increase from 88% to 100% if only the pattern of ^{99m}Tc MIBI uptake lower than or equal to the pertechnetate uptake within the nodule was considered benign. As few benign lesions expressed this uptake pattern, this would decrease the yield of this diagnostic (70).

Piccardo *et al.* did not preselect hypofunctioning lesions, but included all indeterminate thyroid nodules. As expected given the explanation above, the specificity of ^{99m}Tc-MIBI was poor: 52% (19).

Assessment of a retention index of the tracer based on semi-quantitative measurements of the lesion to non-lesion uptake ratios for early and delayed ^{99m}Tc-MIBI images yielded better accuracy. Optimal thresholds for the retention index were determined using ROC analysis and unfortunately not externally validated (70,258). As such, it is unclear whether semi-quantitative ^{99m}Tc-MIBI retention indices are truly more accurate than conventional visual assessment. Moreover, semi-quantitative analysis is still operator dependent, as it depends on the manual definition of ranges of interest (ROI)(19).

^{99m}Tc-MIBI in thyroid nodules with Hürthle cell cytology

Oncocytic cells are rich in mitochondria. Therefore, Hürthle cell lesions – malignant as well as benign – frequently show a more intense and persistent ^{99m}Tc-MIBI uptake (258-260). Boi *et al.* investigated ^{99m}Tc-MIBI in cold thyroid nodules with varying proportions of Hürthle cells in the cytology samples. A relation between ^{99m}Tc-MIBI uptake and increased tissue density of oncocytes was suggested (260). Subsequent studies also concluded that ^{99m}Tc-MIBI is not specific enough to differentiate indeterminate lesions with Hürthle cell cytology (70,255,258). Excluding Hürthle cell nodules from ^{99m}Tc-MIBI assessment likely excludes many false-positive tests while improving benign call rate, specificity and overall diagnostic accuracy in indeterminate thyroid nodules.

Availability, cost-effectiveness and limitations of ^{99m}Tc-MIBI

Imaging of ^{99m}Tc-MIBI requires conventional gamma cameras (with or without single-photon emission computed tomography (SPECT) and CT), which are more widely available than

PET, especially in non-Western countries. Furthermore, the tracer itself is more widely available due to relatively simple complexation using ^{99m}Tc-MIBI-kits together with the favorable half-life of ^{99m}Tc (~6 hours) obtained from on-site generators. The radiation burden of the recommended whole-body adult dose is 5-6 millisievert, but can be lowered by a factor 2-3 by partial-body imaging (261). However, the system resolution of state-of-art gamma cameras is a factor 3 lower than of PET/CT cameras. This decreases the measured signal of lesions smaller than 30 mm, increasingly limiting test sensitivity in smaller nodules. Average costs of ^{99m}Tc-MIBI scanning range from €119 to €500 in Europe and from $669 to $1,156 in the USA (139,262,263). From a German perspective, ^{99m}Tc-MIBI-based management was cost-effective over Afirma® GEC-testing and conventional management. However, this study inappropriately extrapolated auspicious performance parameters of ^{99m}Tc-MIBI in unselected thyroid nodules (96% sensitivity and 46% specificity) to the indeterminate population, and likely underestimated modelled costs for ^{99m}Tc-MIBI scanning and thyroid surgery (14,120,138,139,262-264). Therefore, these assumptions regarding cost-effectiveness in indeterminate thyroid nodules are decidedly questionable and require careful re-evaluation.

Altogether, there is an increased risk of malignancy in thyroid nodules that show increased ^{99m}Tc-MIBI uptake, provided that hypofunctioning nodules are preselected. Nonetheless, test performance in indeterminate thyroid nodules seems insufficient. Excluding Hürthle cell lesions suggests high specificity, but does not resolve the reported poor sensitivity. However, the number of studies currently available for indeterminate thyroid nodules is limited. We believe prospective validation studies in non-oncocytic indeterminate thyroid nodules should be performed. Future studies should also focus on external threshold validation for retention indices to reduce operator dependency and increase accuracy and objectivity of ^{99m}Tc-MIBI. Based on the current evidence, we recommend that ^{99m}Tc-MIBI scanning is not used in

surgical management decisions in indeterminate thyroid nodules without another adjunctive test.

3.2. FDG-PET

Positron emission tomography (PET) using [^{18}F]-2-fluoro-2-deoxy-D-glucose (fluorodeoxyglucose or ^{18}F-FDG), also known as FDG-PET, is an imaging modality that exploits the basic principle that (malignant) tumours and inflammatory tissues are much more metabolically active than normal tissues. Whereas normal tissues predominantly produce energy by low rates of aerobic glycolysis followed by the citric acid cycle in mitochondria, glycolytic rates of rapidly growing cancers can be up to 200 times higher. Subsequent lactic acid fermentation takes place even if oxygen is plentiful (the Warburg effect) (265). Similar to regular glucose, the glucose analogue ^{18}F-FDG is internalized by transmembranous GLUT transporters and converted by hexokinase to ^{18}F-FDG-6-phosphate. However, unlike the 6-phosphorylation product of regular glucose, ^{18}F-FDG-6-phosphate cannot be metabolized further. It is trapped intracellularly and thus accumulates in the tissue. Subsequently, PET scanning can visualize the increased glucose metabolism of the (abnormal) tissue (266). Nowadays, FDG-PET is generally performed in combination with computed tomography (FDG-PET/CT), mainly to correlate metabolically active regions to their anatomic substrates and to correct for tissue-attenuation of the radioactive signal. It is increasingly applied in the diagnostic work-up, staging and therapeutic response monitoring of various malignancies. For thyroid cancer, FDG-PET is frequently used to characterize recurrent disease, especially if dedifferentiation is expected in thyroid carcinomas that lost the capacity to concentrate radioiodide, yet still have measurable serum values of the tumour marker thyroglobulin. It may also be considered in the initial staging of poorly differentiated or invasive Hürthle cell

carcinoma. Moreover, FDG-avid thyroid incidentalomas require additional work-up by FNAC when >1 cm (20,23). In the current ATA guidelines FDG-PET is not routinely recommended for the diagnostic workup of indeterminate thyroid nodules due to limited clinical validation, despite a 2011 meta-analysis by Vriens *et al.* that demonstrated 95% sensitivity and 96% NPV in indeterminate thyroid nodules larger than 15mm (23,267).

Results of available individual studies were mutually consistent despite limited sample sizes. Especially the first studies showed extremely promising results, each reporting 100% sensitivity (266,268-270). De Geus-Oei *et al.* argued that implementation of FDG-PET could reduce the number of futile hemithyroidectomies for benign nodules by 66%, likely outweighing the costs of the extra scans and suggesting cost-effectiveness of this technique in the preoperative setting (269). A subsequent study suggested a less optimistic 39% reduction in futile surgeries, following a lower benign call rate (270). More recent studies demonstrated more modest performance of FDG-PET(/CT) (19,20,271-273). Overall, reported sensitivity and specificity of FDG-PET(/CT) to detect thyroid carcinoma in indeterminate thyroid nodules ranged from 77% to 100% and from 33% to 64%, respectively. A negative index test was reported in approximately 40% of patients (19,20,266,268-274).

Several reasons for false-negativity were proposed, foremost small nodule size. It is how Traugott *et al*. explained their 20% false negative FDG-PET scans: eight lesions were histopathologically smaller than 1 cm. Excluding these, sensitivity and NPV increased to 100% (271). FDG-avidity in very small nodules may be missed on FDG-PET due to the low volume of malignant cells and due to the partial volume effect: the detected FDG-concentration is underestimated dependent on nodule size in relation to the (limited) spatial resolution of the scanner. In larger nodules, this effect is negligible (20,269). Although the improving resolution of state-of-the-art PET scanners pushes the detection limit towards 10 mm, PET is less sensitive in lesions smaller than 15 mm on US. It less reliable to rule-out

microcarcinomas (267). Theoretically, the improving spatial resolution could also become a limitation of the technique: not only will there be less false-negatives, but likely also more false-positive results - leading to a decrease in the already limited specificity over time. In the currently available literature no such downward trend is noted, but future studies should monitor this possibility.

Semi-quantitative FDG-PET

Semi-quantitative analysis of FDG-PET is performed using the maximum standardized uptake value (SUV_{max}): the ratio between the maximum radioactivity concentration measured within a region of interest on the PET image (the 'hottest' voxel) and the decay-corrected amount injected radiotracer per unit of body mass. It reflects the FDG-concentration factor compared to a homogenous distribution of the radiotracer (275). The SUV_{max} is generally significantly higher in malignant than in benign lesions (20,269,270,272,273,275,276). There is a possible correlation between higher SUV_{max} values and increasing size in nodules, insufficiently explained by the abovementioned partial volume effect (20,276). Also, in FTC a higher SUV was associated with capsular or vascular invasion (274). Nonetheless, even though Kresnik *et al*. demonstrated that all carcinoma and Hürthle cell adenoma had an $SUV_{max} \geq 2$ and all other benign lesions an $SUV_{max} < 2$, in multiple other studies the SUV_{max} of benign and malignant indeterminate thyroid nodules overlapped. No threshold could accurately tell them apart (20,269,270,272,273,275,276). Moreover, as SUV_{max} calculations strongly depend on image acquisition and reconstruction methods, type of PET-scanner and other variable methodology, reported absolute SUV_{max} thresholds are not simply valid for other institutions (20). Standardized optimized FDG-PET protocols are required for inter-institution comparison of study results and advancement of PET research (277,278).

FDG-PET in thyroid nodules with Hürthle cell cytology

Multiple studies observed aberrant FDG-PET characteristics in indeterminate nodules with Hürthle cell cytology: both benign and malignant lesions are mostly FDG-positive. Twenty-nine Hürthle cell lesions were reported by Deandreis *et al.*, consisting 52% of their study population and providing an explanation for their limited sensitivity (273). Moreover, Hürthle cell adenoma generally demonstrated a significantly higher SUV_{max} than other benign lesions (20,266,268,269,273,279). The proportion of Hürthle cell cytology in individual studies is relatively small, but overall FDG-PET seems inadequate in these neoplasms.

Availability, cost-effectiveness and limitations of FDG-PET

PET systems are less widely available than conventional gamma cameras. Moreover, ^{18}F used for ^{18}F-FDG synthesis is produced in cyclotrons, and transport distances are limited due to the short half-life of this isotope (~110 min). In Europe, FDG-PET/CT is approximately 1.5-2 times more expensive than ^{99m}Tc-MIBI SPECT/CT. The radiation exposure of FDG-PET/CT is largely accounted for by the FDG dosage at approximately 19 µSv/MBq, i.e. about 3-4 mSv for a typical activity of 185 MBq administered to an average adult (280). Insights regarding common practice total-body FDG-PET/CT imaging are changing (271,273). The CT radiation dose greatly varies, and can be less than 0.5 mSv for a low-dose CT of the neck region only. When scanning the thyroid region only, a longer imaging time can compensate for a reduction in FDG dose, which would lower the radiation burden as well as the costs. Such solutions may counter prevailing reservations regarding ionizing radiation exposure. Additionally, partial-body acquisition could limit the number of coincidental PET-positive findings. Much of the criticism on FDG-PET focuses on these potential incidental findings, which require additional diagnostics, are not always clinically relevant and may negatively impact potential cost-effectiveness (281,282). Malignant ipsi- or contralateral thyroid incidentalomas are

reported while the nodule under investigation was histopathologically benign (271,272). PET-positive incidentalomas are histopathologically malignant in about 20% of patients (282). Cost-effectiveness of FDG-PET/CT was modelled by Vriens *et al.* (14). From a Dutch health care perspective, FDG-PET/CT driven treatment would decrease the rate of unbeneficial diagnostic hemithyroidectomies for benign thyroid nodules by 35% and reduce the costs per patient by €822 compared to the €8,804 expenses for conventional surgical treatment. Also, FDG-PET/CT was favoured over the miRInform® and Afirma® GEC (14).

Contrasting the generally strong sensitivity, specificity of FDG-PET is consistently poor. The underlying mechanism is not yet fully elucidated. The negative influence of Hürthle cell cytology may be partly responsible. It could also be explained by cellular atypia, which was significantly and independently related to FDG uptake, and found in both benign and malignant lesions. Atypia was also related to the presence of Hürthle cells (273). Sebastianes *et al.* hypothesized that FDG uptake is related to variations in gene expression patterns. They suggested that genetic variations between populations may also explain the varying diagnostic accuracy of FDG-PET between studies (270).

In conclusion, FDG-PET(/CT) has the potential to accurately rule-out malignancy in all indeterminate nodules except Hürthle cell lesions. It could prevent unnecessary diagnostic surgery for a significant number of benign thyroid nodules. Sample sizes of existing studies are small, but larger prospective trials are currently ongoing to settle the diagnostic value of this technique and its utility in clinical practice. We recommend that these studies also focus on identifying (genetic) causes for the occasional false-negativity and generally low specificity of this technique.

3.3. DW-MRI

Diffusion-weighted magnetic resonance imaging (DW-MRI) is a functional nuclear magnetic resonance imaging technique that evaluates the rate of random (Brownian) motion of water in tissue, also called diffusivity. By applying diffusion-sensitizing magnetic gradients (the strength and duration of which are expressed as b-values) different levels of diffusion-weighting are obtained: from non-diffusion images (b-value = 0 s/mm^2) to highly diffusion weighted images (i.e. b-value >800 s/mm^2)(283). Lesions that show high signal intensity on DW-MRI images with a high b-value thus show restricted diffusion. The apparent diffusion coefficient (ADC, in mm^2/s) is calculated based on the exponential relationship between signal intensity and the corresponding b-value according to $S(b)=S(0)*e^{-b*ADC}$. A high ADC represents a high degree of diffusion; a low ADC represents diffusion restriction (283,284). DW-MRI thus allows noninvasive quantification of tissue properties without ionizing radiation exposure for the patient. Differentiation between benign and malignant tissues by DW-MRI is based on the assumption that increased cell proliferation, cellular-density and disorganized structures in malignant tissue restrict random motion and thus diffusion of water: a lower ADC-value, together with high signal intensity at high b-values, is more suspicious for malignancy (283,284). Oppositely, increased ADC-values suggest free movement of water molecules in the tissue. It is found in for example edema, colloid follicles, fibrous tissue, hemorrhage and calcification, all of which associated with benign tissues (285). Prior application of DW-MRI in i.e. neuroradiology, breast and lymph nodes showed high diagnostic accuracy (286,287).

Recent exploratory studies in small cohorts of thyroid nodules found distinctively higher ADC values for benign than malignant nodules (283-285,288-292). A recent meta-analysis in 765 cytologically unselected thyroid nodules estimated that DW-MRI had 90% sensitivity and 95% specificity to distinguish thyroid carcinoma (293). Among the individual studies,

however, presented optimum ADC thresholds varied and were not externally validated (283-285,288-292).

Only one small study had assessed DW-MRI in indeterminate thyroid nodules to date. Nakahira *et al.* reported a mean ADC value of 1.27 ± 0.29 $*10^{-3}$ mm^2/s in malignancies opposite 1.95 ± 0.24 $*10^{-3}$ mm^2/s in benign nodules with indeterminate cytology. These results were similar to those of their entire study population (n=42), in which a cutoff ADC value of 1.95 ±0.24 $*10^{-3}$ mm^2/s was 95% sensitive and 83% specific (283).

Availability and limitations of DW-MRI

DW-MRI is infrequently and only experimentally used in the workup of thyroid nodules. Nonetheless, the worldwide availability and application of MRI is growing. As it uses no ionizing but only radiofrequency radiation, the associated risk to the patients is limited, provided that specific measures are taken for patients with MRI-incompatible implanted devices or metal. No MRI-contrast is necessary for DW-MRI, thus avoiding gadolinium-associated toxicity. As the spatial resolution of MRI-scanners is still improving, technical limitations of DW-MRI with regard to minimal lesion size are becoming less relevant compared to SPECT and probably also PET. Still, spatial resolution of DW-MRI sequences is less than that of conventional anatomical MRI-sequences.

There are several major limitations to DW-MRI. MRI is still a rather costly technique; additional sequences such as DW-MRI adds scanner time (~5-10 min) per patient and thus further increases costs. DW-MRI methodology is not standardized yet and its optimal settings still unsettled, leading to varying ADC and b-values (283,289,292). Suboptimal methodology or artifacts cause poor image quality, impede accurate interpretation and caused undesirable exclusions from already small-sized studies, with reported exclusion rates up to 28% (283,284,292). Image artifacts are often caused by inhomogeneity in pathologic tissues or by

their vicinity to interfaces between soft-tissues and bone or air, a source of MRI-artifacts specifically in the thyroid region. Besides viable tumor tissue, malignant tumors partly exist of components with high diffusivity, such necrosis, cystic components or intratumoral hemorrhage (283,285). For accurate ADC measurement, such macroscopic areas should be manually avoided when drawing a region-of-interest. However, avoiding microscopic areas of similar origin, invisible to the human eye, is an impossible task (283). Furthermore, it is hypothesized that the substantial amounts of follicular or Hürthle cells limit the diagnostic accuracy of DW-MRI, specifically in indeterminate thyroid neoplasms. Follicular and Hürthle cell neoplasms are known for their varying colloid tissue involvement. Histologically they contain more fluid. Thus, DW-MRI would inaccurately provide a more benign image (283,292). These hypotheses are currently based on very limited evidence. Further prospective validation studies are desired to determine the possible diagnostic value of DW-MRI in indeterminate thyroid nodules. Future prospects also include improvements of the technique, including consensus on methodology and standardization of acquisition techniques.

4. COMBINED AND MULTISTEP DIAGNOSTICS

The previous chapters of this review addressed the large number of available diagnostic tools to assess indeterminate thyroid nodules. Most studies focused on a single diagnostic technique only. The elimination of between-study population-level differences is a major advantage when comparing the performance of multiple diagnostics independently in one study, optimally in a prospective, independent and blinded fashion. Moreover, assessment of multiple techniques in one study allows investigation of the complementary value of multiple techniques as a *diagnostic* tool by means of simultaneous or sequential testing while at the

same time aiding to further unravel tumor biology as a *research* tool, especially in the current multidisciplinary in-hospital working environment. For example, the question how the presence of a certain oncogenic mutation relates to the (positive) result of an FDG-PET scan could be addressed.

Piccardo *et al*. compared ^{99m}Tc-MIBI, FDG-PET/CT and US plus USE in 87 indeterminate TIR3 nodules with a 21% malignancy rate. FDG-PET/CT was the superior technique with 94% sensitivity and 58% specificity. Following a non-specific positive FDG-PET result, review of ultrasound characteristics offered slight further differentiation; it improved specificity to 77%. However, an additional negative ^{99m}Tc-MIBI scan increased specificity to 94%; this combination was found in 13% of patients (19).

Giovanella *et al.* performed both ^{99m}Tc-MIBI and a 7-gene mutation panel in cold indeterminate thyroid nodules. Combined testing did not improve diagnostic accuracy. Performance of the gene mutation panel was inferior to ^{99m}Tc-MIBI imaging. Of the seven (11%) mutation-positive nodules (four *RAS* mutations and three PAX8/PPARγ rearrangements), only four were malignant. It is unclear whether the low sensitivity of the gene mutation panel in this study can be explained by the selected population of hypofunctioning nodules (70).

Elastosonography and Ultrasonography

USE is superior to ultrasound in indeterminate thyroid nodules – both individual US characteristics as well as combined US patterns described in various articles (210,213,215,217,219,221,222). Two recent prospective studies demonstrated that additional USE evidently improved the diagnostic accuracy of US. Garino *et al.* included nodule stiffness as additional characteristic into a panel of US characteristics and demonstrated that USE identified eight additional malignancies that would have been missed by US assessment

alone. Presence of one or more suspicious US/USE features was 100% sensitive; two or more 88% sensitive and 77% specific. Benign test results were found in 57% of patients. The authors suggested that the 6.4% remaining risk of malignancy– similar to the benign cytology category – would justify follow-up instead of diagnostic hemithyroidectomy in this group (210). In another study of 315 Thy3 nodules, semi-quantitative USE correctly diagnosed 75% of the histopathologically benign lesions that were considered suspicious for malignancy on US, and 83% of the malignancies that were misdiagnosed as benign on US (217). These results suggest that the existing TIRADS classification could be extended with tissue elasticity features. In unselected thyroid nodules this improved TIRADS sensitivity, but not specificity (240,294). The combination is a suitable topic for future research in indeterminate thyroid nodules. Major benefit is that the two techniques are individually inexpensive and obviously easily combined during one diagnostic procedure. Cost-effectiveness can be anticipated.

US and Mutation Analysis

US assessment was also reported in various studies on gene mutation analysis, presumably because US data were usually readily available in clinical studies at no additional costs and thus easily combined with results of more experimental techniques. Even though US assessment improved the diagnostic accuracy of both FDG-PET and elastosonography, combined use of ultrasound with the sensitive Afirma® GEC or specific *BRAF* mutation analysis demonstrated little additional diagnostic value (57). Suspicious US features such as hypoechogenicity, presence of calcifications and hypervascularity were not predictors of malignancy in Afirma® GEC-positive nodules (144). Also, as expected by their individual association to classic PTC, a positive $BRAF^{V600E}$ mutation was correlated to the presence of suspicious US features in unselected nodules, including hypoechogeneity and the presence of

microcalcifications (29,66,78). *BRAF* mutation less frequently occurred in thyroid nodules without suspicious US features (66,78). In Bethesda III and IV thyroid without suspicious US features the prevalence of the *BRAF* mutation was only 1.5% (1/67) in the study by Seo *et al.* – very low, particularly for a South Korean population – all while the malignancy rate was still 18% (12/67)(66). Considering the negligible yield at additional costs, *BRAF* mutation analysis might not be contributory in indeterminate nodules without suspicious US features. An even lower yield from *BRAF* mutation analysis in US-unsuspicious nodules is presumed in populations with a lower general prevalence of *BRAF* mutations. Additionally, these results suggest a different US appearance of *BRAF* mutation-negative malignancies – or a different molecular profile of thyroid carcinoma without suspicious US features.

RAS mutation analysis and assessment of the typical suspicious US features could be complementary in the differentiation of indeterminate thyroid nodules, as follicular-type thyroid carcinomas are associated with *RAS* mutations and infrequently showed the typically suspicious US features (1,29,193-196). Combined assessment could improve diagnostic accuracy of either technique in indeterminate thyroid nodules, identifying papillary thyroid malignancies through classic suspicious US features and follicular-type carcinoma by *RAS* mutation analysis. However, challenges for clinical practice continue to exist in the imperfect specificity of *RAS* mutation analysis, and the interobserver variability and ambiguity of certain US features.

Immunocytochemistry and Mutation Analysis

In histopathology samples, certain genetic alterations were correlated to positive staining for specific immunomarkers: PAX8/PPARγ rearrangement was associated with galectin-3 reactivity, and *RAS* point mutation with HBME-1 (105). Only one study investigated this combination of techniques in indeterminate thyroid cytology. Although no significant

correlation was demonstrated between positive $BRAF^{V600E}$ mutation and galectin-3 overexpression – benefitting possible complementary use – no additional diagnostic value was demonstrated either (44).

MicroRNA and Mutation Analysis

Combined microRNA expression profiling and mutation analysis could accurately aid diagnosis and prognosis of thyroid malignancy. Distinct microRNAs have been related to oncogenic mutations. For example, miR-221, miR-222 and miR-146b were more overexpressed in *BRAF*- and *RAS*-mutated PTC. High expression of miR-187 was associated with RET/PTC rearrangement (148,295). The first step towards diagnostic integration of the two techniques was taken by Labourier *et al.*, who tested the commercially developed 10-microRNA thyroid classifier ThyraMIR™ simultaneously with the miRInform® thyroid (22). The ThyraMIR™ was designed to increase the sensitivity of the miRInform® without affecting its specificity. Combined use demonstrated 89% sensitivity and 85% specificity (22). A recent decision analytics model for Bethesda III and IV nodules estimated that combined miRInform® and ThyraMIR™ testing was cost-effective, reducing the rate of unnecessary surgery (diagnostic hemithyroidectomy as well as two-step thyroidectomies) from 88% to 20% and saving $1,384 per patient in the first year of treatment or $3,170 per avoided surgery. However, it is not described how the economic consequences of the 15% missed malignancies are accounted for in this model (140). The economic as well as medical-ethical consequences of such a high number of missed malignancies question the current clinical utility of this combination of expensive techniques.

In brief, the combined or sequential use of multiple diagnostics in indeterminate thyroid nodules was infrequently studied. Regrettably, the available studies also mostly remained

within their own field of expertise: comparing tests either within the domain of pathological (molecular) techniques or within the domain of imaging. Although a sequential combination of a sensitive and an uncorrelated specific test might bring the solution that this clinical issue has been waiting for, the most accurate combination of tests cannot reliably be determined yet.

5. RECENT DEVELOPMENTS AND FUTURE PROSPECTS

5.1. The Cancer Genome Atlas

Papillary thyroid cancer was one of the cancers targeted by the cancer genome atlas (TCGA) research network, a large collaborative project by the National Cancer Institute (NCI) and National Human Genome Research Institute (NHGRI). The incentive of the project is to map genomic alterations occurring in 33 types of cancer in 11.000 patients and improve the understanding, classification and extending possibilities for targeted therapy of these cancers (296). Genetic alterations of all kinds were detected in nearly five hundred clinically non-aggressive PTCs (classical, follicular and tall cell variants) using one proteomic and six genomic platforms. PTC harbored fewer somatic mutations than other human cancer types, but if they were present, driver mutations were detected in the majority of the cancer cells. As expected, the known driver mutations in the MAPK/ERK pathway were dominant, confirming the mutually exclusive relation for *BRAF* and *RAS* point mutations and RET/PTC rearrangements. Other detected genetic alterations included genetic variations of the TERT promoter, PI3K and PPARγ pathways, as well as new alterations of known and new drivers, such as *EIF1AX*, *PPM1D* and *CHEK2*. Moreover, molecular subtypes of for example *BRAF*-mutated PTC were identified and linked to different clinical subtypes. The role of microRNA

in determining cancer phenotype was elaborated, allowing better understanding of clinical behavior of various genetic variants of PTC. Somatic copy number alterations were mostly linked to FVPTC. Ultimately, the TCGA Research Network envisions a reclassification of thyroid carcinoma, abandoning the discrimination between PTC and FTC, and classifying according to molecular subtypes instead of by histopathological subtype first (297). The identified markers may not just have an application in the diagnosis of thyroid carcinoma, but also in better risk-stratification of the different cancers and in targeted therapies. The plurality of applications is best known for the $BRAF^{V600E}$ mutation, which has an association with clinically more aggressive tumor behavior on several fronts. Also, non-thyroid malignancies carrying a *BRAF* mutation are now (experimentally) treated with RAF inhibitors (298,299). There is little doubt that molecular classification systems are the future of oncology diagnostics in all types of human cancers. The position of histopathological assessment is changing, but cannot be renounced. With the current knowledge of thyroid genomics, the need to distinguish the mutated malignant from the mutated benign – premalignant – neoplasms remains, with all due consequences for the surgical and postoperative treatment strategy.

Cytological application of the TCGA set was also already investigated in a recent study. Pagan *et al*. validated a panel containing the genomic alterations identified by the TCGA in 88 FNAC samples selected from a previous cohort study, including 22 indeterminate thyroid nodules (126,300). In the latter, 33% sensitivity and 84% specificity were demonstrated. In the same set of patients, Pagan *et al*. also performed the Afirma® GEC. The GEC yielded less false negatives and a much higher sensitivity. Even though technical limitations of the applied sequencing techniques could leave RNA transcriptions with low expression levels undetected and thus negatively influence sensitivity of the TCGA set, the scopes of the TCGA and GEC

most likely explain their difference in performance. The TCGA was developed using PTC only. It did not include follicular lesions and their distinctive genetic alterations. Moreover, in contrast to the GEC, the TCGA set was not optimized for preoperative diagnostic application in indeterminate thyroid nodules (300). Consequently, the comparison performed by this Veracyte-sponsored study seems unjust: it is obvious that the Afirma® GEC yielded better diagnostic performance in this specific clinical setting. Yet, the results of this study did prove that a large panel of genetic alterations such as the TCGA was not useful in clinical practice without further expansion of the scope of the panel towards follicular thyroid neoplasms. Still, the genetic alterations and their relations detected by TCGA are groundbreaking for the progression of research. From these comprehensive sets of biomarkers, we may select new combinations of genetic alterations for future clinical research to develop an accurate rule-in or rule-out molecular test for indeterminate thyroid nodules.

5.2. Proteomics

Other molecular advances include protein expression diagnostics, or proteomic profiling. These techniques allow for more detailed insight in the molecular biology and protein expression of thyroid neoplasms. For example, matrix-assisted laser desorption ionization / mass spectrometry imaging (MALDI-MSI) is able to simultaneously visualize the spatial distribution of proteins and profile up- and downregulated protein expression in relation to the morphological features of the thyroid specimen. These and related proteomic techniques could identify new biomarkers for preoperative cytological diagnosis, but require high levels of expertise. Application to thyroid cytology has so far been investigated by few studies (301,302). Ex-vivo cytology studies show accurate and reproducible differentiation between various lesions, including the currently difficult to diagnose Hürthle cell neoplasms (302). No

studies investigated the diagnostic value of proteomics in in-vivo indeterminate thyroid cytology yet.

DISCUSSION

This review provides a comprehensive overview of the available literature on molecular and imaging biomarkers as additional diagnostics for thyroid nodules with indeterminate cytology (Bethesda III and IV) and their application in a clinical preoperative setting. Clinical utility requires more from a diagnostic than mere well-validated test performance and high rule-in or rule-out capacity. The 2015 ATA guidelines suggested that the ideal rule-out diagnostic for thyroid carcinoma should have a NPV similar to a benign cytological diagnosis (~96.3%) and the ideal rule-in test a PPV that is at least similar to a malignant cytological diagnosis (~98.6%) (10,23). The balance between test sensitivity and specificity – and their prevalence-dependent derivatives PPV and NPV – directly reflects on feasibility and cost-effectiveness estimates. A diagnostic with (near) perfect sensitivity but limited specificity is inefficient and unlikely cost-effective: the NPV will be close to 100%, but the majority of nodules will test positive. Therefore, instead of focusing on the reproducible highest sensitivity or specificity, a diagnostic is better appreciated by end points such as desired minimal rates of accurately prevented unbeneficial surgeries or accurately diagnosed carcinomas. More importantly, clinical utility demands that implementation of the ancillary test leads to changes in patient management and overall health benefits (303). All these requirements directly depend on a plurality of epidemiological and economic factors within the tested population, such as the local test availability, professional expertise and case mix – prevalence of malignancy as well as the balance of various subtypes of indeterminate cytology including especially Hürthle-cell neoplasms and *BRAF*-mutation. Additionally, clinical utility considerations should include

less tangible factors such as physician and patient preference, multidisciplinary decision making and compatibility with everyday clinical routine and logistics in endocrine practice. All things considered, global perspectives regarding the preferred diagnostic for indeterminate thyroid nodules likely greatly differ.

Recommendation for clinical use of rule-out tests

The most accurate currently available rule-out tests are the Afirma® GEC and FDG-PET(/CT) imaging. The Afirma® GEC had strikingly high sensitivity in nearly all studies (127,129,131-134). However, there are concerns regarding the lack of strong validation studies. With a high degree of missing histology, especially in GEC negative nodules, there is a potentially strong diminution of the tests' sensitivity if unresected GEC-negative lesions were less often benign than presumed. In the USA, physicians should locally validate the tests' utility prior to implementation. However, with its limited global availability, high costs and low probability of cost-effectiveness, clinical implementation of the Afirma® GEC outside the USA is currently not favored (14,121,137-140).

FDG-PET/CT may be the preferred rule-out test for indeterminate thyroid nodules in a European setting. With sufficient validation studies with complete histopathological follow-up, it demonstrated consistent high sensitivity and a benign test result in 40% of the patients, although the number of currently published patients is moderate. Cost-effectiveness of FDG-PET over other diagnostics is presumed (14). Its popularity in the USA is more limited, although the efficacy of this molecular imaging technique could likely compete with molecular biomarkers panels, even if the costs per scan are somewhat higher than in Europe. The main drawback of FDG-PET/CT is its – admitted minor – risk to the patient by using a limited dose of ionizing radiation.

The recently announced version 3 of the ThyroSeq® may become a prime contender. Dependent on the case mix, the ThyroSeq® v2.1 anticipated high negative predictive value (21). However, the number of studies to confirm test performance and clinical utility in different patient populations is limited. Clinical results for the ThyroSeq® v3 are eagerly awaited.

Semi-quantitative elastosonography could be a suitable alternative, in particular in case a more economic test is required. However, overfitting and lack of external cut-off validation likely overestimated the performance of this technique in the limited number of available studies. If future prospective studies can confirm its performance and thresholds of this operator-dependent but globally accessible method, USE could become a more important diagnostic in this field.

None of the diagnostic techniques under investigation in this review has a perfect NPV or fulfills the threshold proposed by the ATA. A number of malignant nodules will be misdiagnosed as benign on first assessment. Considering the typical indolent clinical course of differentiated thyroid cancer, follow-up of these initially false-negative nodules will most likely still result in timely diagnosis without relevant treatment delay and dismal prognostic consequences.

Recommendation for clinical use of rule-in tests

The best rule-in performance was unmistakably demonstrated by *BRAF* mutation analysis, which showed perfect 100% specificity in an abundance of studies. Yet, strong regional differences in prevalence of *BRAF* mutations have a major impact on its clinical utility, especially when comparing South Korea to other countries. Moreover, the analysis most likely has very low yield in Bethesda IV nodules, in which the mostly follicular type malignancies are more frequently *RAS*-mutated (31,50,70,76). Testing for individual genetic

alterations other than the $BRAF^{V600E}$ point mutation is not useful. In American and European settings, a gene mutation panel is likely preferred over any individual mutation analysis. Promising rule-in capacity was also demonstrated for Galectin-3 immunocytochemistry. An infrequently applied technique with limited validation studies, further prospective studies are warranted to validate its performance in indeterminate thyroid nodules and endorse its possible clinical use.

Besides *BRAF* mutation analysis, none of diagnostics meet the 2015 ATA requirements of an ideal rule-in test. Compared to ruling-out tests, ruling-in tests face an additional challenge. With a generally low frequency of thyroid carcinoma in indeterminate thyroid nodules, achieving a reliable PPV – higher than 95% – can be a major challenge despite adequate test specificity. Such high demands to a ruling-in test advocate the use of a ruling-out test in populations with a limited pre-test probability of malignancy.

Clinical recommendation for a step-wise approach

Most of the diagnostic modalities are optimized for either ruling in or ruling out malignancy. No single diagnostic addressed in the current review currently has it all: both a near-perfect sensitivity and a near-perfect specificity, and (proven) cost-effectiveness. It is extremely challenging to develop such test performance parameters in a single diagnostic. Even promising new diagnostics, such as the ThyroSeq® and ThyraMir™, require significant further optimization to get near this diagnostic utopia.

With the diagnostics currently available in the clinical setting, a multimodality stepwise approach could offer a conclusive diagnosis for indeterminate thyroid nodules, sequentially combining one sensitive rule-out and one specific rule-in test. Unfortunately, thus far few studies investigated this approach (19,70). Combinations of (molecular) imaging and somatic genetics were especially scarce. There is currently insufficient evidence to accommodate

reliable interpretation of sequentially used tests, as performance of the second test is unknown in a population preselected by the first. Besides choosing two accurate and uncorrelated tests to achieve maximum diagnostic accuracy, the sequence of testing, local availability and costs of the selected diagnostics are crucial. Costs of two or more additional tests may compromise cost-utility estimates. Available cost-effectiveness studies for individual diagnostic modalities were additionally greatly susceptible to global variations in population-dependent factors such as pre-test probability of thyroid carcinoma and local test performance, and varying health care costs including the surgical reimbursement rates (14,121,138,139). Reported surgical and hospitalization costs range from $4,628 to $6,549 for hemithyroidectomy, $5,272 to $7,068 for completion thyroidectomy and $5,680 to $11,265 for initial total thyroidectomy. Secondary expenses following surgery should be considered as well, including postoperative observation, thyroid hormone replacement (approximately $150 per patient per year), treatment for hypoparathyroidism (approximately $860 per patient per year), and resolution of rare but potentially serious surgical complications (14,120,138,264). Secondary endpoints such as quality of life and survival are of minor importance to cost-effectiveness, due to the generally indolent course of differentiated thyroid cancer, adequate treatment options and overall low disease-related mortality (14,138,139).

Recent discussions in thyroid histopathology

Histopathology is classically based on microscopic assessment of tumor phenotype, aided by immunohistochemistry. However, this ‘gold standard test’ is also subject to advancing insights regarding tumor phenotype, increasingly aided by knowledge regarding tumor genotype. Mutation-negative malignancies resulting from indeterminate cytology were frequently identified as encapsulated follicular variants of papillary thyroid carcinoma without histologic features of aggressive behavior (21,22,31,59,80). Also, several studies defined a

separate intermediate *histopathological* category called '(follicular) tumor of uncertain malignant potential' for encapsulated, well-differentiated follicular tumors with questionable PTC-type nuclear changes (71,164,177,273). These examples illustrate one of the important ongoing discussions in thyroid histopathology. In 2016, Nikiforov *et al*. proposed an official downscaling of the classification of proven noninvasive encapsulated FVPTCs, renaming them 'noninvasive follicular neoplasm with papillary-like nuclear features' (NIFTP). The behavior of these neoplasms is benign unlike other thyroid carcinoma subtypes, showing no evidence of recurrent disease after a median 13-year follow-up. About one in four of the neoplasms in the retrospective cohort were mutated, most frequently carrying *RAS* (*NRAS*) or PAX8/PPARγ alterations. Presence of a mutation likely predisposes the NIFTP to progress into an invasive encapsulated FVPTC, justifying surgical resection. Treatment of NIFTP should most likely be limited to hemithyroidectomy, waiving totalizing thyroidectomy and radioiodine ablation (304). Although revolutionizing, this new nomenclature complicates mutation-based preoperative decision-making (21,31,80). The justification to skip two-stage surgery and perform a total thyroidectomy at once for mutation-positive nodules is the driving force of the 7-gene mutation panel and similar tests, but would be overkill for the subgroup of NIFTP (31). Nonetheless, most of the undesirable possible overtreatment for NIFTP is likely resolved if *RAS*-mutated indeterminate nodules are treated with hemi- instead of total thyroidectomy, as previously suggested. No comprehensive diagnostic test is currently available to diagnose mutation-positive NIFTP preoperatively, as follicular tumor invasiveness and encapsulation cannot be distinguished on cytology.

Hürthle cell cytology

The Achilles heel of many diagnostics investigated in this review is cytology suspicious for a Hürthle cell neoplasm (Bethesda IV SHCN/HCN). Hürthle cells are oxyphilic cells with

abundant cytoplasm and an enlarged nucleus with a prominent nucleolus. They are found in benign thyroid diseases such as Hashimoto's thyroiditis, but also occur in the notorious Hürthle cell adenoma and carcinoma, the oncocytic variant of follicular adenoma and carcinoma (4,185). Although Hürthle cell carcinomas (FTC-OV) are rare, their aberrant clinical course and association with invasive features justifies the special attention given to Hürthle cell cytology by the Bethesda and other classification systems. An accurate additional diagnostic is desired. Disappointingly, several studies concluded that the investigated test was accurate in all except Hürthle cell lesions (70,132,258,273). Immunocytochemistry handed some solutions, although promising results of combined galectin-3 and CK-19 staining have not yet been validated (174). Besides that, *BRAF*, *RAS*, RET/PTC or PAX8/PPARγ alterations are only occasionally found (70,76). These findings support previous presumptions that oncocytic thyroid nodules are a completely separate entity with a unique molecular and phenotypic profile (305-308). Malignant transition in Hürthle cell nodules most likely involves the PIK3CA-Akt-mTOR and Wnt/bèta-catenin pathways rather than the MAPK/ERK pathway (305,308). Rare *TP53* mutations, usually associated with poorly differentiated and anaplastic carcinoma, were recently also identified in well-differentiated Hürthle cell nodules (306). Also, recurrent FTC-OV have shown genome haploidisation, a rare phenomenon in other types of differentiated thyroid carcinoma (309). Specific markers for the preoperative molecular differentiation of Hürthle cell nodules should be developed. Adaptation of existing tests to additionally suit Hürthle cell nodules (e.g. the Afirma® GEC) is a strategy being explored, for example by the ThyroSeq® v3 and the Afirma® Gene Sequence Classifier. Caution should be taken that these adaptations do not decrease the diagnostic accuracy for non-oncocytic lesions. MicroRNA expression profiling of these lesions is currently also under investigation (148,152).

Strengths and limitations of the current review

There are several important strengths and limitations to this comprehensive review. This review provides a complete overview of the available additional diagnostics for indeterminate thyroid nodules, resulting from a careful and systematic literature selection and quality appraisal. Different types of clinical data of various levels of evidence were considerately presented. Nonetheless, this review is generally prone to inaccuracies from low study quality, study heterogeneity and different types of bias. For some of the assessed diagnostics, the limited number of available publications and small study cohorts contribute to heterogeneity of data and loss of applicability. This mainly concerns studies on non-routine imaging techniques. By nature, these clinical studies need to prospectively include subjects to voluntarily undergo an extra investigation with – at least in the clinical validation phase – no implications for individual patient management. These types of studies require more resources than 'further use' tissue biobank studies. Consequently, the number of studies is more limited and published series often are small. In contrast, cytological biomarker research gratefully profits from available large tissue biobanks for initial validation studies. We believe consistent results from properly designed imaging studies should not be disregarded due to mere their sample size, but be appreciated by the quality of their study design and statistics. Population-level study differences were often observed, not only related to test performance but also strongly varying malignancy rates that were oftentimes much lower or higher than expected from indeterminate thyroid nodules. Besides insuperable epidemiological variations, the selection of indeterminate cytology, and the retrospective nature of many studies may have contributed to these discrepancies.

The type of indeterminate cytology included by individual studies varied, likely leading to between-study heterogeneity. Besides global variations and known intra- and interobserver discordance, diverse definitions of indeterminate cytology were adhered (5). Nowadays, the

Bethesda system differentiates indeterminate from benign and suspicious cytology in a more standardized manner in both literature and clinic. Bethesda III and/or IV and similar categories from other classification systems were frequently applied. Unfortunately, some studies also included small numbers of Bethesda V nodules without presenting results for individual categories separately (127). Many other studies adhered to their own definition of indeterminate cytology. This especially, but not exclusively, concerns studies published before the introduction of the Bethesda system in 2009.

Retrospective study designs and subsequent selection bias – only including indeterminate thyroid nodules that had undergone both thyroid surgery and (routine) pre-operative testing – likely also caused overestimation of the true efficacy of certain techniques (e.g. *BRAF* mutation analysis or ultrasound).

CONCLUSION AND RECOMMENDATIONS

In current-day practice, there are numerous additional diagnostics available to further assess thyroid nodules with indeterminate cytology, all with advantages and disadvantages. This review provided a comprehensive overview of the available literature on these techniques, addressing both molecular and imaging biomarkers, aiming to provide an objective and nuanced comparison of their performance and cost-effectiveness with regard to rightful surgical decision-making. Many of these diagnostics have either an adequate rule-in or rule-out capacity, but no single currently available test seems to serve both purposes well. Diagnostics from the different research fields likely complement each other in a multimodality stepwise diagnostic approach towards. Notwithstanding, test performance is always population-dependent. To correctly interpret the results, the prevalence of malignancy and the performance, costs and feasibility of the desired diagnostic in the local patient

population should be known beforehand. Local implementation studies are strongly recommended to confirm clinical utility. Most importantly, the local decision favoring or opposing a certain diagnostic should be a deliberate and multidisciplinary one. Cooperation between clinical endocrinologists, endocrine surgeons, pathologists, radiologists and nuclear medicine physicians is crucial.

References

1. Gupta N, Dasyam AK, Carty SE, Nikiforova MN, Ohori NP, Armstrong M, Yip L, LeBeau SO, McCoy KL, Coyne C, Stang MT, Johnson J, Ferris RL, Seethala R, Nikiforov YE, Hodak SP. RAS mutations in thyroid FNA specimens are highly predictive of predominantly low-risk follicular-pattern cancers. The Journal of clinical endocrinology and metabolism 2013; 98:E914-922

2. Zhu Z, Gandhi M, Nikiforova MN, Fischer AH, Nikiforov YE. Molecular profile and clinical-pathologic features of the follicular variant of papillary thyroid carcinoma. An unusually high prevalence of ras mutations. American journal of clinical pathology 2003; 120:71-77

3. Lee SR, Jung CK, Kim TE, Bae JS, Jung SL, Choi YJ, Kang CS. Molecular genotyping of follicular variant of papillary thyroid carcinoma correlates with diagnostic category of fine-needle aspiration cytology: values of RAS mutation testing. Thyroid 2013; 23:1416-1422

4. Cibas ES, Ali SZ, Conference NCITFSotS. The Bethesda System For Reporting Thyroid Cytopathology. American journal of clinical pathology 2009; 132:658-665

5. Cibas ES, Baloch ZW, Fellegara G, LiVolsi VA, Raab SS, Rosai J, Diggans J, Friedman L, Kennedy GC, Kloos RT, Lanman RB, Mandel SJ, Sindy N, Steward DL,

Zeiger MA, Haugen BR, Alexander EK. A prospective assessment defining the limitations of thyroid nodule pathologic evaluation. Annals of internal medicine 2013; 159:325-332

6. Mitchell AL, Gandhi A, Scott-Coombes D, Perros P. Management of thyroid cancer: United Kingdom National Multidisciplinary Guidelines. The Journal of laryngology and otology 2016; 130:S150-S160

7. Nardi F, Basolo F, Crescenzi A, Fadda G, Frasoldati A, Orlandi F, Palombini L, Papini E, Zini M, Pontecorvi A, Vitti P. Italian consensus for the classification and reporting of thyroid cytology. Journal of endocrinological investigation 2014; 37:593-599

8. Brito JP, Al Nofal A, Montori VM, Hay ID, Morris JC. The Impact of Subclinical Disease and Mechanism of Detection on the Rise in Thyroid Cancer Incidence: A Population-Based Study in Olmsted County, Minnesota During 1935 Through 2012. Thyroid 2015; 25:999-1007

9. Fazeli R, Schneider EB, Ali SZ, Zeiger MA, Olson MT. Diagnostic Frequency Ratios Are Insufficient to Measure Laboratory Precision with The Bethesda System for Reporting Thyroid Cytopathology. Acta cytologica 2015; 59:225-232

10. Bongiovanni M, Spitale A, Faquin WC, Mazzucchelli L, Baloch ZW. The Bethesda System for Reporting Thyroid Cytopathology: a meta-analysis. Acta cytologica 2012; 56:333-339

11. Sullivan PS, Hirschowitz SL, Fung PC, Apple SK. The impact of atypia/follicular lesion of undetermined significance and repeat fine-needle aspiration: 5 years before and after implementation of the Bethesda System. Cancer cytopathology 2014; 122:866-872

12. Caplan RH, Wester SM, Lambert PJ, Rooney BL. Efficient evaluation of thyroid nodules by primary care providers and thyroid specialists. The American journal of managed care 2000; 6:1134-1140

13. Khalid AN, Quraishi SA, Hollenbeak CS, Stack BC, Jr. Fine-needle aspiration biopsy versus ultrasound-guided fine-needle aspiration biopsy: cost-effectiveness as a frontline diagnostic modality for solitary thyroid nodules. Head & neck 2008; 30:1035-1039

14. Vriens D, Adang EM, Netea-Maier RT, Smit JW, de Wilt JH, Oyen WJ, de Geus-Oei LF. Cost-effectiveness of FDG-PET/CT for cytologically indeterminate thyroid nodules: a decision analytic approach. The Journal of clinical endocrinology and metabolism 2014; 99:3263-3274

15. McHenry CR, Slusarczyk SJ. Hypothyroidisim following hemithyroidectomy: incidence, risk factors, and management. Surgery 2000; 128:994-998

16. Rosato L, Avenia N, Bernante P, De Palma M, Gulino G, Nasi PG, Pelizzo MR, Pezzullo L. Complications of thyroid surgery: analysis of a multicentric study on 14,934 patients operated on in Italy over 5 years. World journal of surgery 2004; 28:271-276

17. Jeannon JP, Orabi AA, Bruch GA, Abdalsalam HA, Simo R. Diagnosis of recurrent laryngeal nerve palsy after thyroidectomy: a systematic review. International journal of clinical practice 2009; 63:624-629

18. Samir AE, Dhyani M, Anvari A, Prescott J, Halpern EF, Faquin WC, Stephen A. Shear-Wave Elastography for the Preoperative Risk Stratification of Follicular-patterned Lesions of the Thyroid: Diagnostic Accuracy and Optimal Measurement Plane. Radiology 2015:141627

19. Piccardo A, Puntoni M, Treglia G, Foppiani L, Bertagna F, Paparo F, Massollo M, Dib B, Paone G, Arlandini A, Catrambone U, Casazza S, Pastorino A, Cabria M, Giovanella L. Thyroid nodules with indeterminate cytology: prospective comparison between 18F-FDG-PET/CT, multiparametric neck ultrasonography, 99mTc-MIBI scintigraphy and histology. European journal of endocrinology / European Federation of Endocrine Societies 2016;

20. Munoz Perez N, Villar del Moral JM, Muros Fuentes MA, Lopez de la Torre M, Arcelus Martinez JI, Becerra Massare P, Esteva Martinez D, Canadas Garre M, Coll Del Rey E, Bueno Larano P, Ferron Orihuela JA. Could 18F-FDG-PET/CT avoid unnecessary thyroidectomies in patients with cytological diagnosis of follicular neoplasm? Langenbeck's archives of surgery / Deutsche Gesellschaft fur Chirurgie 2013; 398:709-716

21. Nikiforov YE, Carty SE, Chiosea SI, Coyne C, Duvvuri U, Ferris RL, Gooding WE, LeBeau SO, Ohori NP, Seethala RR, Tublin ME, Yip L, Nikiforova MN. Impact of the Multi-Gene ThyroSeq Next-Generation Sequencing Assay on Cancer Diagnosis in Thyroid Nodules with Atypia of Undetermined Significance/Follicular Lesion of Undetermined Significance Cytology. Thyroid 2015; 25:1217-1223

22. Labourier E, Shifrin A, Busseniers AE, Lupo MA, Manganelli ML, Andruss B, Wylie D, Beaudenon-Huibregtse S. Molecular Testing for miRNA, mRNA, and DNA on Fine-Needle Aspiration Improves the Preoperative Diagnosis of Thyroid Nodules With Indeterminate Cytology. The Journal of clinical endocrinology and metabolism 2015; 100:2743-2750

23. Haugen BR, Alexander EK, Bible KC, Doherty GM, Mandel SJ, Nikiforov YE, Pacini F, Randolph GW, Sawka AM, Schlumberger M, Schuff KG, Sherman SI, Sosa JA, Steward DL, Tuttle RM, Wartofsky L. 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; 26:1-133

24. Kimura ET, Nikiforova MN, Zhu Z, Knauf JA, Nikiforov YE, Fagin JA. High prevalence of BRAF mutations in thyroid cancer: genetic evidence for constitutive activation of the RET/PTC-RAS-BRAF signaling pathway in papillary thyroid carcinoma. Cancer research 2003; 63:1454-1457

25. Adeniran AJ, Zhu Z, Gandhi M, Steward DL, Fidler JP, Giordano TJ, Biddinger PW, Nikiforov YE. Correlation between genetic alterations and microscopic features, clinical manifestations, and prognostic characteristics of thyroid papillary carcinomas. The American journal of surgical pathology 2006; 30:216-222

26. Malumbres M, Barbacid M. RAS oncogenes: the first 30 years. Nat Rev Cancer 2003; 3:459-465

27. Davies H, Bignell GR, Cox C, Stephens P, Edkins S, Clegg S, Teague J, Woffendin H, Garnett MJ, Bottomley W, Davis N, Dicks E, Ewing R, Floyd Y, Gray K, Hall S, Hawes R, Hughes J, Kosmidou V, Menzies A, Mould C, Parker A, Stevens C, Watt S, Hooper S, Wilson R, Jayatilake H, Gusterson BA, Cooper C, Shipley J, Hargrave D, Pritchard-Jones K, Maitland N, Chenevix-Trench G, Riggins GJ, Bigner DD, Palmieri G, Cossu A, Flanagan A, Nicholson A, Ho JW, Leung SY, Yuen ST, Weber BL, Seigler HF, Darrow TL, Paterson H, Marais R, Marshall CJ, Wooster R, Stratton MR, Futreal PA. Mutations of the BRAF gene in human cancer. Nature 2002; 417:949-954

28. Xing M. BRAF mutation in papillary thyroid cancer: pathogenic role, molecular bases, and clinical implications. Endocrine reviews 2007; 28:742-762

29. Rossi M, Buratto M, Tagliati F, Rossi R, Lupo S, Trasforini G, Lanza G, Franceschetti P, Bruni S, Degli Uberti E, Zatelli MC. Relevance of BRAF(V600E) mutation testing

versus RAS point mutations and RET/PTC rearrangements evaluation in the diagnosis of thyroid cancer. Thyroid 2015; 25:221-228

30. Carr R, Ustun B, Chhieng D, Schofield K, Theoharis C, Hammers L, Adeniran AJ. Radiologic and clinical predictors of malignancy in the follicular lesion of undetermined significance of the thyroid. Endocrine pathology 2013; 24:62-68

31. Nikiforov YE, Ohori NP, Hodak SP, Carty SE, LeBeau SO, Ferris RL, Yip L, Seethala RR, Tublin ME, Stang MT, Coyne C, Johnson JT, Stewart AF, Nikiforova MN. Impact of mutational testing on the diagnosis and management of patients with cytologically indeterminate thyroid nodules: a prospective analysis of 1056 FNA samples. The Journal of clinical endocrinology and metabolism 2011; 96:3390-3397

32. Nikiforova MN, Kimura ET, Gandhi M, Biddinger PW, Knauf JA, Basolo F, Zhu Z, Giannini R, Salvatore G, Fusco A, Santoro M, Fagin JA, Nikiforov YE. BRAF mutations in thyroid tumors are restricted to papillary carcinomas and anaplastic or poorly differentiated carcinomas arising from papillary carcinomas. The Journal of clinical endocrinology and metabolism 2003; 88:5399-5404

33. Xing M. BRAF mutation in thyroid cancer. Endocrine-related cancer 2005; 12:245-262

34. Kim SK, Hwang TS, Yoo YB, Han HS, Kim DL, Song KH, Lim SD, Kim WS, Paik NS. Surgical results of thyroid nodules according to a management guideline based on the BRAF(V600E) mutation status. The Journal of clinical endocrinology and metabolism 2011; 96:658-664

35. Rossi M, Buratto M, Bruni S, Filieri C, Tagliati F, Trasforini G, Rossi R, Beccati MD, Degli Uberti EC, Zatelli MC. Role of ultrasonographic/clinical profile, cytology, and BRAF V600E mutation evaluation in thyroid nodule screening for malignancy: a

prospective study. The Journal of clinical endocrinology and metabolism 2012; 97:2354-2361

36. Adeniran AJ, Hui P, Chhieng DC, Prasad ML, Schofield K, Theoharis C. BRAF mutation testing of thyroid fine-needle aspiration specimens enhances the predictability of malignancy in thyroid follicular lesions of undetermined significance. Acta cytologica 2011; 55:570-575

37. Ohori NP, Wolfe J, Hodak SP, LeBeau SO, Yip L, Carty SE, Duvvuri U, Schoedel KE, Nikiforova MN, Nikiforov YE. "Colloid-rich" follicular neoplasm/suspicious for follicular neoplasm thyroid fine-needle aspiration specimens: cytologic, histologic, and molecular basis for considering an alternate view. Cancer cytopathology 2013; 121:718-728

38. Eszlinger M, Piana S, Moll A, Bosenberg E, Bisagni A, Ciarrocchi A, Ragazzi M, Paschke R. Molecular testing of thyroid fine-needle aspirations improves presurgical diagnosis and supports the histologic identification of minimally invasive follicular thyroid carcinomas. Thyroid 2015; 25:401-409

39. Hwang TS, Kim WY, Han HS, Lim SD, Kim WS, Yoo YB, Park KS, Oh SY, Kim SK, Yang JH. Preoperative RAS mutational analysis is of great value in predicting follicular variant of papillary thyroid carcinoma. BioMed research international 2015; 2015:697068

40. Zhang Q, Liu SZ, Zhang Q, Guan YX, Chen QJ, Zhu QY. Meta-Analyses of Association Between BRAF(V600E) Mutation and Clinicopathological Features of Papillary Thyroid Carcinoma. Cellular physiology and biochemistry : international journal of experimental cellular physiology, biochemistry, and pharmacology 2016; 38:763-776

41. Yeo MK, Liang ZL, Oh T, Moon Y, An S, Kim MK, Kim KS, Shong M, Kim JM, Jo YS. Pyrosequencing cut-off value identifying BRAFV600E mutation in fine needle aspiration samples of thyroid nodules. Clinical endocrinology 2011; 75:555-560

42. Pelizzo MR, Boschin IM, Barollo S, Pennelli G, Toniato A, Zambonin L, Vianello F, Piotto A, Ide EC, Pagetta C, Sorgato N, Torresan F, Girelli ME, Nacamulli D, Mantero F, Mian C. BRAF analysis by fine needle aspiration biopsy of thyroid nodules improves preoperative identification of papillary thyroid carcinoma and represents a prognostic factor. A mono-institutional experience. Clinical chemistry and laboratory medicine : CCLM / FESCC 2011; 49:325-329

43. Xing M, Tufano RP, Tufaro AP, Basaria S, Ewertz M, Rosenbaum E, Byrne PJ, Wang J, Sidransky D, Ladenson PW. Detection of BRAF mutation on fine needle aspiration biopsy specimens: a new diagnostic tool for papillary thyroid cancer. The Journal of clinical endocrinology and metabolism 2004; 89:2867-2872

44. Sapio MR, Guerra A, Posca D, Limone PP, Deandrea M, Motta M, Troncone G, Caleo A, Vallefuoco P, Rossi G, Fenzi G, Vitale M. Combined analysis of galectin-3 and BRAFV600E improves the accuracy of fine-needle aspiration biopsy with cytological findings suspicious for papillary thyroid carcinoma. Endocrine-related cancer 2007; 14:1089-1097

45. Sapio MR, Posca D, Raggioli A, Guerra A, Marotta V, Deandrea M, Motta M, Limone PP, Troncone G, Caleo A, Rossi G, Fenzi G, Vitale M. Detection of RET/PTC, TRK and BRAF mutations in preoperative diagnosis of thyroid nodules with indeterminate cytological findings. Clinical endocrinology 2007; 66:678-683

46. Kim SK, Kim DL, Han HS, Kim WS, Kim SJ, Moon WJ, Oh SY, Hwang TS. Pyrosequencing analysis for detection of a BRAFV600E mutation in an FNAB

specimen of thyroid nodules. Diagnostic molecular pathology : the American journal of surgical pathology, part B 2008; 17:118-125

47. Jo YS, Huang S, Kim YJ, Lee IS, Kim SS, Kim JR, Oh T, Moon Y, An S, Ro HK, Kim JM, Shong M. Diagnostic value of pyrosequencing for the BRAF V600E mutation in ultrasound-guided fine-needle aspiration biopsy samples of thyroid incidentalomas. Clinical endocrinology 2009; 70:139-144

48. Marchetti I, Lessi F, Mazzanti CM, Bertacca G, Elisei R, Coscio GD, Pinchera A, Bevilacqua G. A morpho-molecular diagnosis of papillary thyroid carcinoma: BRAF V600E detection as an important tool in preoperative evaluation of fine-needle aspirates. Thyroid 2009; 19:837-842

49. Cantara S, Capezzone M, Marchisotta S, Capuano S, Busonero G, Toti P, Di Santo A, Caruso G, Carli AF, Brilli L, Montanaro A, Pacini F. Impact of proto-oncogene mutation detection in cytological specimens from thyroid nodules improves the diagnostic accuracy of cytology. The Journal of clinical endocrinology and metabolism 2010; 95:1365-1369

50. Moses W, Weng J, Sansano I, Peng M, Khanafshar E, Ljung BM, Duh QY, Clark OH, Kebebew E. Molecular testing for somatic mutations improves the accuracy of thyroid fine-needle aspiration biopsy. World journal of surgery 2010; 34:2589-2594

51. Adeniran AJ, Theoharis C, Hui P, Prasad ML, Hammers L, Carling T, Udelsman R, Chhieng DC. Reflex BRAF testing in thyroid fine-needle aspiration biopsy with equivocal and positive interpretation: a prospective study. Thyroid 2011; 21:717-723

52. Patel A, Klubo-Gwiezdzinska J, Hoperia V, Larin A, Jensen K, Bauer A, Vasko V. BRAF(V600E) mutation analysis from May-Grunwald Giemsa-stained cytological samples as an adjunct in identification of high-risk papillary thyroid carcinoma. Endocrine pathology 2011; 22:195-199

53. Canadas-Garre M, Becerra-Massare P, Lopez de la Torre-Casares M, Villar-del Moral J, Cespedes-Mas S, Vilchez-Joya R, Muros-de Fuentes T, Garcia-Calvente C, Piedrola-Maroto G, Lopez-Nevot MA, Montes-Ramirez R, Llamas-Elvira JM. Reduction of false-negative papillary thyroid carcinomas by the routine analysis of BRAF(T1799A) mutation on fine-needle aspiration biopsy specimens: a prospective study of 814 thyroid FNAB patients. Annals of surgery 2012; 255:986-992

54. Kang G, Cho EY, Shin JH, Chung JH, Kim JW, Oh YL. Role of BRAFV600E mutation analysis and second cytologic review of fine-needle aspiration for evaluating thyroid nodule. Cancer cytopathology 2012; 120:44-51

55. Mancini I, Pinzani P, Pupilli C, Petrone L, De Feo ML, Bencini L, Pazzagli M, Forti G, Orlando C. A high-resolution melting protocol for rapid and accurate differential diagnosis of thyroid nodules. The Journal of molecular diagnostics : JMD 2012; 14:501-509

56. Tonacchera M, Agretti P, Rago T, De Marco G, Niccolai F, Molinaro A, Scutari M, Candelieri A, Conforti D, Musmanno R, Di Coscio G, Basolo F, Iacconi P, Miccoli P, Pinchera A, Vitti P. Genetic markers to discriminate benign and malignant thyroid nodules with undetermined cytology in an area of borderline iodine deficiency. Journal of endocrinological investigation 2012; 35:754-759

57. Kloos RT, Reynolds JD, Walsh PS, Wilde JI, Tom EY, Pagan M, Barbacioru C, Chudova DI, Wong M, Friedman L, LiVolsi VA, Rosai J, Lanman RB, Kennedy GC. Does addition of BRAF V600E mutation testing modify sensitivity or specificity of the Afirma Gene Expression Classifier in cytologically indeterminate thyroid nodules? The Journal of clinical endocrinology and metabolism 2013; 98:E761-768

58. Agretti P, Niccolai F, Rago T, De Marco G, Molinaro A, Scutari M, Di Cosmo C, Di Coscio G, Vitale M, Maccheroni M, Vitti P, Tonacchera M. BRAF mutation analysis

in thyroid nodules with indeterminate cytology: our experience on surgical management of patients with thyroid nodules from an area of borderline iodine deficiency. Journal of endocrinological investigation 2014; 37:1009-1014

59. Beaudenon-Huibregtse S, Alexander EK, Guttler RB, Hershman JM, Babu V, Blevins TC, Moore P, Andruss B, Labourier E. Centralized molecular testing for oncogenic gene mutations complements the local cytopathologic diagnosis of thyroid nodules. Thyroid 2014; 24:1479-1487

60. Danilovic DL, Lima EU, Domingues RB, Brandao LG, Hoff AO, Marui S. Pre-operative role of BRAF in the guidance of the surgical approach and prognosis of differentiated thyroid carcinoma. European journal of endocrinology / European Federation of Endocrine Societies 2014; 170:619-625

61. Eszlinger M, Krogdahl A, Munz S, Rehfeld C, Precht Jensen EM, Ferraz C, Bosenberg E, Drieschner N, Scholz M, Hegedus L, Paschke R. Impact of molecular screening for point mutations and rearrangements in routine air-dried fine-needle aspiration samples of thyroid nodules. Thyroid 2014; 24:305-313

62. Eszlinger M, Neustadt M, Ruschenburg I, Neumann A, Franzius C, Adam S, Bacher K, Hach A, Hammoser R, Langvogt C, Molwitz T, Paschke R. [Fine-needle aspiration cytology of thyroid nodules: molecular diagnostics in a routine diagnostic setting]. Deutsche medizinische Wochenschrift (1946) 2014; 139:476-480

63. Johnson SJ, Hardy SA, Roberts C, Bourn D, Mallick U, Perros P. Pilot of BRAF mutation analysis in indeterminate, suspicious and malignant thyroid FNA cytology. Cytopathology 2014; 25:146-154

64. Liu S, Gao A, Zhang B, Zhang Z, Zhao Y, Chen P, Ji M, Hou P, Shi B. Assessment of molecular testing in fine-needle aspiration biopsy samples: an experience in a Chinese population. Experimental and molecular pathology 2014; 97:292-297

65. Poller DN, Glaysher S, Agrawal A, Caldera S, Kim D, Yiangou C. BRAF V600 co-testing in thyroid FNA cytology: short-term experience in a large cancer centre in the UK. Journal of clinical pathology 2014; 67:684-689

66. Seo JY, Kim EK, Kwak JY. Additional BRAF mutation analysis may have additional diagnostic value in thyroid nodules with "suspicious for malignant" cytology alone even when the nodules do not show suspicious US features. Endocrine 2014; 47:283-289

67. Borrelli N, Ugolini C, Giannini R, Antonelli A, Giordano M, Sensi E, Torregrossa L, Fallahi P, Miccoli P, Basolo F. Role of gene expression profiling in defining indeterminate thyroid nodules in addition to BRAF analysis. Cancer cytopathology 2015;

68. Capelli L, Marfisi C, Puccetti M, Saragoni L, De Paola F, Zaccaroni A, Chiadini E, Gagliardi L, Ferretti G, Zoli W, Ulivi P. Role of BRAF molecular analysis in the management of papillary thyroid carcinoma: analysis of cytological and histological samples. Cytopathology 2015; 26:297-302

69. Gill MS, Nayan S, Kocovski L, Cutz JC, Archibald SD, Jackson BS, Young JE, Gupta MK. Local molecular analysis of indeterminate thyroid nodules. Journal of otolaryngology - head & neck surgery = Le Journal d'oto-rhino-laryngologie et de chirurgie cervico-faciale 2015; 44:52

70. Giovanella L, Campenni A, Treglia G, Verburg FA, Trimboli P, Ceriani L, Bongiovanni M. Molecular imaging with Tc-MIBI and molecular testing for mutations in differentiating benign from malignant follicular neoplasm: a prospective comparison. European journal of nuclear medicine and molecular imaging 2015;

71. Le Mercier M, D'Haene N, De Neve N, Blanchard O, Degand C, Rorive S, Salmon I. Next-generation sequencing improves the diagnosis of thyroid FNA specimens with indeterminate cytology. Histopathology 2015; 66:215-224

72. Marino M, Monzani ML, Brigante G, Cioni K, Madeo B, Santi D, Maiorana A, Bettelli S, Moriondo V, Pignatti E, Bonacini L, Carani C, Rochira V, Simoni M. High-Resolution Melting Is a Sensitive, Cost-Effective, Time-Saving Technique for BRAF V600E Detection in Thyroid FNAB Washing Liquid: A Prospective Cohort Study. European Thyroid Journal 2015; 4:73-81

73. Park KS, Oh YL, Ki CS, Kim JW. Evaluation of the Real-Q BRAF V600E Detection Assay in Fine-Needle Aspiration Samples of Thyroid Nodules. The Journal of molecular diagnostics : JMD 2015; 17:431-437

74. Beisa A, Beisa V, Stoskus M, Ostaneviciute E, Griskevicius L, Strupas K. The value of the repeated examination of BRAF V600E mutation status in diagnostics of papillary thyroid cancer. Endokrynologia Polska 2016; 67:35-40

75. Kowalska A, Kowalik A, Palyga I, Walczyk A, Gasior-Perczak D, Kopczynski J, Lizis-Kolus K, Szyska-Skrobot D, Hurej S, Radowicz-Chil A, Chodurska R, Wypiorkiewicz E, Chlopek M, Nowak E, Niemyska K, Gozdz S. The usefulness of determining the presence of BRAF V600E mutation in fine-needle aspiration cytology in indeterminate cytological results. Endokrynologia Polska 2016; 67:41-47

76. Valderrabano P, Leon ME, Centeno BA, Otto KJ, Khazai L, McCaffrey JC, Russell JS, McIver B. Institutional prevalence of malignancy of indeterminate thyroid cytology is necessary but insufficient to accurately interpret molecular marker tests. European journal of endocrinology / European Federation of Endocrine Societies 2016;

77. Kim SW, Lee JI, Kim JW, Ki CS, Oh YL, Choi YL, Shin JH, Kim HK, Jang HW, Chung JH. BRAFV600E mutation analysis in fine-needle aspiration cytology specimens for evaluation of thyroid nodule: a large series in a BRAFV600E-prevalent population. The Journal of clinical endocrinology and metabolism 2010; 95:3693-3700

78. Koh J, Choi JR, Han KH, Kim EK, Yoon JH, Moon HJ, Kwak JY. Proper indication of BRAF(V600E) mutation testing in fine-needle aspirates of thyroid nodules. PloS one 2013; 8:e64505

79. Leslie C, Grieu-Iacopetta F, Richter A, Platten M, Murray J, Frost FA, Amanuel B, Kumarasinghe MP. BRAF p.Val600Glu (V600E) mutation detection in thyroid fine needle aspiration cell block samples: a feasibility study. Pathology 2015; 47:432-438

80. Nikiforov YE, Carty SE, Chiosea SI, Coyne C, Duvvuri U, Ferris RL, Gooding WE, Hodak SP, LeBeau SO, Ohori NP, Seethala RR, Tublin ME, Yip L, Nikiforova MN. Highly accurate diagnosis of cancer in thyroid nodules with follicular neoplasm/suspicious for a follicular neoplasm cytology by ThyroSeq v2 next-generation sequencing assay. Cancer 2014; 120:3627-3634

81. Jeong SH, Hong HS, Lee EH, Cha JG, Park JS, Kwak JJ. Outcome of thyroid nodules characterized as atypia of undetermined significance or follicular lesion of undetermined significance and correlation with Ultrasound features and BRAF(V600E) mutation analysis. AJR American journal of roentgenology 2013; 201:W854-860

82. Hyeon J, Ahn S, Shin JH, Oh YL. The prediction of malignant risk in the category "atypia of undetermined significance/follicular lesion of undetermined significance" of the Bethesda System for Reporting Thyroid Cytopathology using subcategorization and BRAF mutation results. Cancer cytopathology 2014; 122:368-376

83. Park HJ, Moon JH, Yom CK, Kim KH, Choi JY, Choi SI, Ahn SH, Jeong WJ, Lee WW, Park SY. Thyroid "atypia of undetermined significance" with nuclear atypia has high rates of malignancy and BRAF mutation. Cancer cytopathology 2014; 122:512-520

84. Lee ST, Kim SW, Ki CS, Jang JH, Shin JH, Oh YL, Kim JW, Chung JH. Clinical implication of highly sensitive detection of the BRAF V600E mutation in fine-needle aspirations of thyroid nodules: a comparative analysis of three molecular assays in 4585 consecutive cases in a BRAF V600E mutation-prevalent area. The Journal of clinical endocrinology and metabolism 2012; 97:2299-2306

85. Ohori NP, Singhal R, Nikiforova MN, Yip L, Schoedel KE, Coyne C, McCoy KL, LeBeau SO, Hodak SP, Carty SE, Nikiforov YE. BRAF mutation detection in indeterminate thyroid cytology specimens: underlying cytologic, molecular, and pathologic characteristics of papillary thyroid carcinoma. Cancer cytopathology 2013; 121:197-205

86. Afkhami M, Karunamurthy A, Chiosea S, Nikiforova MN, Seethala R, Nikiforov YE, Coyne C. Histopathologic and Clinical Characterization of Thyroid Tumors Carrying the BRAF(K601E) Mutation. Thyroid 2016; 26:242-247

87. Trimboli P, Treglia G, Condorelli E, Romanelli F, Crescenzi A, Bongiovanni M, Giovanella L. BRAF-mutated carcinomas among thyroid nodules with prior indeterminate FNA report: a systematic review and meta-analysis. Clinical endocrinology 2016; 84:315-320

88. Lee WS, Palmer BJ, Garcia A, Chong VE, Liu TH. BRAF mutation in papillary thyroid cancer: A cost-utility analysis of preoperative testing. Surgery 2014; 156:1569-1577; discussion 1577-1568

89. An JH, Song KH, Kim SK, Park KS, Yoo YB, Yang JH, Hwang TS, Kim DL. RAS mutations in indeterminate thyroid nodules are predictive of the follicular variant of papillary thyroid carcinoma. Clinical endocrinology 2015; 82:760-766

90. Yoon JH, Kwon HJ, Lee HS, Kim EK, Moon HJ, Kwak JY. RAS Mutations in AUS/FLUS Cytology: Does it Have an Additional Role in BRAFV600E Mutation-Negative Nodules? Medicine 2015; 94:e1084

91. Marshall CJ, Hall A, Weiss RA. A transforming gene present in human sarcoma cell lines. Nature 1982; 299:171-173

92. Shimizu K, Goldfarb M, Perucho M, Wigler M. Isolation and preliminary characterization of the transforming gene of a human neuroblastoma cell line. Proceedings of the National Academy of Sciences of the United States of America 1983; 80:383-387

93. Radkay LA, Chiosea SI, Seethala RR, Hodak SP, LeBeau SO, Yip L, McCoy KL, Carty SE, Schoedel KE, Nikiforova MN, Nikiforov YE, Ohori NP. Thyroid nodules with KRAS mutations are different from nodules with NRAS and HRAS mutations with regard to cytopathologic and histopathologic outcome characteristics. Cancer cytopathology 2014; 122:873-882

94. Kunavisarut T. Diagnostic biomarkers of differentiated thyroid cancer. Endocrine 2013; 44:616-622

95. Mitsutake N, Miyagishi M, Mitsutake S, Akeno N, Mesa C, Jr., Knauf JA, Zhang L, Taira K, Fagin JA. BRAF mediates RET/PTC-induced mitogen-activated protein kinase activation in thyroid cells: functional support for requirement of the RET/PTC-RAS-BRAF pathway in papillary thyroid carcinogenesis. Endocrinology 2006; 147:1014-1019

96. Guerra A, Sapio MR, Marotta V, Campanile E, Moretti MI, Deandrea M, Motta M, Limone PP, Fenzi G, Rossi G, Vitale M. Prevalence of RET/PTC rearrangement in benign and malignant thyroid nodules and its clinical application. Endocrine journal 2011; 58:31-38

97. Elisei R, Romei C, Vorontsova T, Cosci B, Veremeychik V, Kuchinskaya E, Basolo F, Demidchik EP, Miccoli P, Pinchera A, Pacini F. RET/PTC rearrangements in thyroid nodules: studies in irradiated and not irradiated, malignant and benign thyroid lesions in children and adults. The Journal of clinical endocrinology and metabolism 2001; 86:3211-3216

98. Su X, Li Z, He C, Chen W, Fu X, Yang A. Radiation exposure, young age, and female gender are associated with high prevalence of RET/PTC1 and RET/PTC3 in papillary thyroid cancer: a meta-analysis. Oncotarget 2016; 7:16716-16730

99. de Groot JW, Links TP, Plukker JT, Lips CJ, Hofstra RM. RET as a diagnostic and therapeutic target in sporadic and hereditary endocrine tumors. Endocrine reviews 2006; 27:535-560

100. Kroll TG, Sarraf P, Pecciarini L, Chen CJ, Mueller E, Spiegelman BM, Fletcher JA. PAX8-PPARgamma1 fusion oncogene in human thyroid carcinoma [corrected]. Science (New York, NY) 2000; 289:1357-1360

101. Zhang Y, Yu J, Lee C, Xu B, Sartor MA, Koenig RJ. Genomic binding and regulation of gene expression by the thyroid carcinoma-associated PAX8-PPARG fusion protein. Oncotarget 2015; 6:40418-40432

102. Pasca di Magliano M, Di Lauro R, Zannini M. Pax8 has a key role in thyroid cell differentiation. Proceedings of the National Academy of Sciences of the United States of America 2000; 97:13144-13149

103. Raman P, Koenig RJ. Pax-8-PPAR-gamma fusion protein in thyroid carcinoma. Nature reviews Endocrinology 2014; 10:616-623

104. Castro P, Rebocho AP, Soares RJ, Magalhaes J, Roque L, Trovisco V, Vieira de Castro I, Cardoso-de-Oliveira M, Fonseca E, Soares P, Sobrinho-Simoes M. PAX8-PPARgamma rearrangement is frequently detected in the follicular variant of papillary thyroid carcinoma. The Journal of clinical endocrinology and metabolism 2006; 91:213-220

105. Nikiforova MN, Lynch RA, Biddinger PW, Alexander EK, Dorn GW, 2nd, Tallini G, Kroll TG, Nikiforov YE. RAS point mutations and PAX8-PPAR gamma rearrangement in thyroid tumors: evidence for distinct molecular pathways in thyroid follicular carcinoma. The Journal of clinical endocrinology and metabolism 2003; 88:2318-2326

106. Jeong SH, Hong HS, Kwak JJ, Lee EH. Analysis of RAS mutation and PAX8/PPARgamma rearrangements in follicular-derived thyroid neoplasms in a Korean population: frequency and ultrasound findings. Journal of endocrinological investigation 2015; 38:849-857

107. Armstrong MJ, Yang H, Yip L, Ohori NP, McCoy KL, Stang MT, Hodak SP, Nikiforova MN, Carty SE, Nikiforov YE. PAX8/PPARgamma rearrangement in thyroid nodules predicts follicular-pattern carcinomas, in particular the encapsulated follicular variant of papillary carcinoma. Thyroid 2014; 24:1369-1374

108. French CA, Alexander EK, Cibas ES, Nose V, Laguette J, Faquin W, Garber J, Moore F, Jr., Fletcher JA, Larsen PR, Kroll TG. Genetic and biological subgroups of low-stage follicular thyroid cancer. The American journal of pathology 2003; 162:1053-1060

109. French CA, Fletcher JA, Cibas ES, Caulfield C, Allard P, Kroll TG. Molecular detection of PPAR gamma rearrangements and thyroid carcinoma in preoperative fine-needle aspiration biopsies. Endocrine pathology 2008; 19:166-174

110. Collins K, Mitchell JR. Telomerase in the human organism. Oncogene 2002; 21:564-579

111. Melo M, da Rocha AG, Vinagre J, Batista R, Peixoto J, Tavares C, Celestino R, Almeida A, Salgado C, Eloy C, Castro P, Prazeres H, Lima J, Amaro T, Lobo C, Martins MJ, Moura M, Cavaco B, Leite V, Cameselle-Teijeiro JM, Carrilho F, Carvalheiro M, Maximo V, Sobrinho-Simoes M, Soares P. TERT promoter mutations are a major indicator of poor outcome in differentiated thyroid carcinomas. The Journal of clinical endocrinology and metabolism 2014; 99:E754-765

112. Umbricht CB, Conrad GT, Clark DP, Westra WH, Smith DC, Zahurak M, Saji M, Smallridge RC, Goodman S, Zeiger MA. Human telomerase reverse transcriptase gene expression and the surgical management of suspicious thyroid tumors. Clinical cancer research : an official journal of the American Association for Cancer Research 2004; 10:5762-5768

113. Liou MJ, Chan EC, Lin JD, Liu FH, Chao TC. Human telomerase reverse transcriptase (hTERT) gene expression in FNA samples from thyroid neoplasms. Cancer letters 2003; 191:223-227

114. Greco A, Miranda C, Pierotti MA. Rearrangements of NTRK1 gene in papillary thyroid carcinoma. Molecular and cellular endocrinology 2010; 321:44-49

115. Berlingieri MT, Pierantoni GM, Giancotti V, Santoro M, Fusco A. Thyroid cell transformation requires the expression of the HMGA1 proteins. Oncogene 2002; 21:2971-2980

116. Lappinga PJ, Kip NS, Jin L, Lloyd RV, Henry MR, Zhang J, Nassar A. HMGA2 gene expression analysis performed on cytologic smears to distinguish benign from malignant thyroid nodules. Cancer cytopathology 2010; 118:287-297

117. Bartolazzi A, Gasbarri A, Papotti M, Bussolati G, Lucante T, Khan A, Inohara H, Marandino F, Orlandi F, Nardi F, Vecchione A, Tecce R, Larsson O, Thyroid Cancer Study G. Application of an immunodiagnostic method for improving preoperative diagnosis of nodular thyroid lesions. Lancet 2001; 357:1644-1650

118. Maruta J, Hashimoto H, Yamashita H, Yamashita H, Noguchi S. Immunostaining of galectin-3 and CD44v6 using fine-needle aspiration for distinguishing follicular carcinoma from adenoma. Diagnostic cytopathology 2004; 31:392-396

119. Matesa N, Samija I, Kusic Z. Accuracy of fine needle aspiration biopsy with and without the use of tumor markers in cytologically indeterminate thyroid lesions. Collegium antropologicum 2010; 34:53-57

120. Yip L, Farris C, Kabaker AS, Hodak SP, Nikiforova MN, McCoy KL, Stang MT, Smith KJ, Nikiforov YE, Carty SE. Cost impact of molecular testing for indeterminate thyroid nodule fine-needle aspiration biopsies. The Journal of clinical endocrinology and metabolism 2012; 97:1905-1912

121. Lee L, How J, Tabah RJ, Mitmaker EJ. Cost-effectiveness of molecular testing for thyroid nodules with atypia of undetermined significance cytology. The Journal of clinical endocrinology and metabolism 2014; 99:2674-2682

122. Nikiforova MN, Wald AI, Roy S, Durso MB, Nikiforov YE. Targeted next-generation sequencing panel (ThyroSeq) for detection of mutations in thyroid cancer. The Journal of clinical endocrinology and metabolism 2013; 98:E1852-1860

123. Nishino M. Molecular cytopathology for thyroid nodules: A review of methodology and test performance. Cancer cytopathology 2016; 124:14-27

124. ThermoFisher Scientific. Ion AmpliSeq™ Cancer Hotspot Panel v2. 2107; https://www.thermofisher.com/order/catalog/product/4475346. Accessed 25-03-2017.

125. Chudova D, Wilde JI, Wang ET, Wang H, Rabbee N, Egidio CM, Reynolds J, Tom E, Pagan M, Rigl CT, Friedman L, Wang CC, Lanman RB, Zeiger M, Kebebew E, Rosai J, Fellegara G, LiVolsi VA, Kennedy GC. Molecular classification of thyroid nodules using high-dimensionality genomic data. The Journal of clinical endocrinology and metabolism 2010; 95:5296-5304

126. Alexander EK, Kennedy GC, Baloch ZW, Cibas ES, Chudova D, Diggans J, Friedman L, Kloos RT, LiVolsi VA, Mandel SJ, Raab SS, Rosai J, Steward DL, Walsh PS, Wilde JI, Zeiger MA, Lanman RB, Haugen BR. Preoperative diagnosis of benign thyroid nodules with indeterminate cytology. The New England journal of medicine 2012; 367:705-715

127. Alexander EK, Schorr M, Klopper J, Kim C, Sipos J, Nabhan F, Parker C, Steward DL, Mandel SJ, Haugen BR. Multicenter clinical experience with the Afirma gene expression classifier. The Journal of clinical endocrinology and metabolism 2014; 99:119-125

128. Duick DS, Klopper JP, Diggans JC, Friedman L, Kennedy GC, Lanman RB, McIver B. The impact of benign gene expression classifier test results on the endocrinologist-patient decision to operate on patients with thyroid nodules with indeterminate fine-needle aspiration cytopathology. Thyroid 2012; 22:996-1001

129. Witt RL. Outcome of thyroid gene expression classifier testing in clinical practice. Laryngoscope 2016; 126:524-527

130. McIver B, Castro MR, Morris JC, Bernet V, Smallridge R, Henry M, Kosok L, Reddi H. An independent study of a gene expression classifier (Afirma) in the evaluation of

cytologically indeterminate thyroid nodules. The Journal of clinical endocrinology and metabolism 2014; 99:4069-4077

131. Yang SE, Sullivan PS, Zhang J, Govind R, Levin MR, Rao JY, Moatamed NA. Has Afirma gene expression classifier testing refined the indeterminate thyroid category in cytology? Cancer cytopathology 2016; 124:100-109

132. Brauner E, Holmes BJ, Krane JF, Nishino M, Zurakowski D, Hennessey JV, Faquin WC, Parangi S. Performance of the Afirma Gene Expression Classifier in Hurthle Cell Thyroid Nodules Differs from Other Indeterminate Thyroid Nodules. Thyroid 2015; 25:789-796

133. Celik B, Whetsell CR, Nassar A. Afirma GEC and thyroid lesions: An institutional experience. Diagnostic cytopathology 2015; 43:966-970

134. Marti JL, Avadhani V, Donatelli LA, Niyogi S, Wang B, Wong RJ, Shaha AR, Ghossein RA, Lin O, Morris LG, Ho AS. Wide Inter-institutional Variation in Performance of a Molecular Classifier for Indeterminate Thyroid Nodules. Annals of surgical oncology 2015;

135. Noureldine SI, Olson MT, Agrawal N, Prescott JD, Zeiger MA, Tufano RP. Effect of Gene Expression Classifier Molecular Testing on the Surgical Decision-Making Process for Patients With Thyroid Nodules. JAMA otolaryngology-- head & neck surgery 2015; 141:1082-1088

136. Aragon Han P, Olson MT, Fazeli R, Prescott JD, Pai SI, Schneider EB, Tufano RP, Zeiger MA. The impact of molecular testing on the surgical management of patients with thyroid nodules. Annals of surgical oncology 2014; 21:1862-1869

137. Wu JX, Lam R, Levin M, Rao J, Sullivan PS, Yeh MW. Effect of malignancy rates on cost-effectiveness of routine gene expression classifier testing for indeterminate thyroid nodules. Surgery 2016; 159:118-129

138. Li H, Robinson KA, Anton B, Saldanha IJ, Ladenson PW. Cost-effectiveness of a novel molecular test for cytologically indeterminate thyroid nodules. The Journal of clinical endocrinology and metabolism 2011; 96:E1719-1726

139. Heinzel A, Muller D, Behrendt FF, Giovanella L, Mottaghy FM, Verburg FA. Thyroid nodules with indeterminate cytology: molecular imaging with (9)(9)mTc-methoxyisobutylisonitrile (MIBI) is more cost-effective than the Afirma gene expression classifier. European journal of nuclear medicine and molecular imaging 2014; 41:1497-1500

140. Labourier E. Utility and cost-effectiveness of molecular testing in thyroid nodules with indeterminate cytology. Clinical endocrinology 2016; 85:624-631

141. Macias AA, Eappen S, Malikin I, Goldfarb J, Kujawa S, Konowitz PM, Kamani D, Randolph GW. Successful intraoperative electrophysiologic monitoring of the recurrent laryngeal nerve, a multidisciplinary approach: The Massachusetts Eye and Ear Infirmary monitoring collaborative protocol with experience in over 3000 cases. Head & neck 2016; 38:1487-1494

142. Harrell RM, Bimston DN. Surgical utility of Afirma: effects of high cancer prevalence and oncocytic cell types in patients with indeterminate thyroid cytology. Endocrine practice : official journal of the American College of Endocrinology and the American Association of Clinical Endocrinologists 2014; 20:364-369

143. Santhanam P, Khthir R, Gress T, Elkadry A, Olajide O, Yaqub A, Driscoll H. Gene expression classifier for the diagnosis of indeterminate thyroid nodules: a meta-analysis. Medical oncology (Northwood, London, England) 2016; 33:14

144. Lastra RR, Pramick MR, Crammer CJ, LiVolsi VA, Baloch ZW. Implications of a suspicious afirma test result in thyroid fine-needle aspiration cytology: an institutional experience. Cancer cytopathology 2014; 122:737-744

145. He H, Jazdzewski K, Li W, Liyanarachchi S, Nagy R, Volinia S, Calin GA, Liu CG, Franssila K, Suster S, Kloos RT, Croce CM, de la Chapelle A. The role of microRNA genes in papillary thyroid carcinoma. Proceedings of the National Academy of Sciences of the United States of America 2005; 102:19075-19080

146. Pallante P, Visone R, Ferracin M, Ferraro A, Berlingieri MT, Troncone G, Chiappetta G, Liu CG, Santoro M, Negrini M, Croce CM, Fusco A. MicroRNA deregulation in human thyroid papillary carcinomas. Endocrine-related cancer 2006; 13:497-508

147. Nikiforova MN, Nikiforov YE. Molecular diagnostics and predictors in thyroid cancer. Thyroid 2009; 19:1351-1361

148. Nikiforova MN, Tseng GC, Steward D, Diorio D, Nikiforov YE. MicroRNA expression profiling of thyroid tumors: biological significance and diagnostic utility. The Journal of clinical endocrinology and metabolism 2008; 93:1600-1608

149. Agretti P, Ferrarini E, Rago T, Candelieri A, De Marco G, Dimida A, Niccolai F, Molinaro A, Di Coscio G, Pinchera A, Vitti P, Tonacchera M. MicroRNA expression profile helps to distinguish benign nodules from papillary thyroid carcinomas starting from cells of fine-needle aspiration. European journal of endocrinology / European Federation of Endocrine Societies 2012; 167:393-400

150. Aragon Han P, Weng CH, Khawaja HT, Nagarajan N, Schneider EB, Umbricht CB, Witwer KW, Zeiger MA. MicroRNA Expression and Association with Clinicopathologic Features in Papillary Thyroid Cancer: A Systematic Review. Thyroid 2015; 25:1322-1329

151. Li X, Abdel-Mageed AB, Mondal D, Kandil E. MicroRNA expression profiles in differentiated thyroid cancer, a review. International journal of clinical and experimental medicine 2013; 6:74-80

152. Dettmer M, Perren A, Moch H, Komminoth P, Nikiforov YE, Nikiforova MN. Comprehensive MicroRNA expression profiling identifies novel markers in follicular variant of papillary thyroid carcinoma. Thyroid 2013; 23:1383-1389

153. Dettmer M, Vogetseder A, Durso MB, Moch H, Komminoth P, Perren A, Nikiforov YE, Nikiforova MN. MicroRNA expression array identifies novel diagnostic markers for conventional and oncocytic follicular thyroid carcinomas. The Journal of clinical endocrinology and metabolism 2013; 98:E1-7

154. Shen R, Liyanarachchi S, Li W, Wakely PE, Jr., Saji M, Huang J, Nagy R, Farrell T, Ringel MD, de la Chapelle A, Kloos RT, He H. MicroRNA signature in thyroid fine needle aspiration cytology applied to "atypia of undetermined significance" cases. Thyroid 2012; 22:9-16

155. Weber F, Teresi RE, Broelsch CE, Frilling A, Eng C. A limited set of human MicroRNA is deregulated in follicular thyroid carcinoma. The Journal of clinical endocrinology and metabolism 2006; 91:3584-3591

156. Stokowy T, Wojtas B, Fujarewicz K, Jarzab B, Eszlinger M, Paschke R. miRNAs with the potential to distinguish follicular thyroid carcinomas from benign follicular thyroid tumors: results of a meta-analysis. Hormone and metabolic research = Hormon- und Stoffwechselforschung = Hormones et metabolisme 2014; 46:171-180

157. Zhang Y, Zhong Q, Chen X, Fang J, Huang Z. Diagnostic value of microRNAs in discriminating malignant thyroid nodules from benign ones on fine-needle aspiration samples. Tumour biology : the journal of the International Society for Oncodevelopmental Biology and Medicine 2014; 35:9343-9353

158. Wei WJ, Shen CT, Song HJ, Qiu ZL, Luo QY. MicroRNAs as a potential tool in the differential diagnosis of thyroid cancer: a systematic review and meta-analysis. Clinical endocrinology 2016; 84:127-133

159. Keutgen XM, Filicori F, Crowley MJ, Wang Y, Scognamiglio T, Hoda R, Buitrago D, Cooper D, Zeiger MA, Zarnegar R, Elemento O, Fahey TJ, 3rd. A panel of four miRNAs accurately differentiates malignant from benign indeterminate thyroid lesions on fine needle aspiration. Clinical cancer research : an official journal of the American Association for Cancer Research 2012; 18:2032-2038

160. Vriens MR, Weng J, Suh I, Huynh N, Guerrero MA, Shen WT, Duh QY, Clark OH, Kebebew E. MicroRNA expression profiling is a potential diagnostic tool for thyroid cancer. Cancer 2012; 118:3426-3432

161. Rossi ED, Bizzarro T, Martini M, Capodimonti S, Sarti D, Cenci T, Bilotta M, Fadda G, Larocca LM. The evaluation of miRNAs on thyroid FNAC: the promising role of miR-375 in follicular neoplasms. Endocrine 2016;

162. Igci YZ, Ozkaya M, Korkmaz H, Bozgeyik E, Bayraktar R, Ulasli M, Erkilic S, Eraydin A, Oztuzcu S. Expression Levels of miR-30a-5p in Papillary Thyroid Carcinoma: A Comparison Between Serum and Fine Needle Aspiration Biopsy Samples. Genetic testing and molecular biomarkers 2015;

163. de Matos LL, Del Giglio AB, Matsubayashi CO, de Lima Farah M, Del Giglio A, da Silva Pinhal MA. Expression of CK-19, galectin-3 and HBME-1 in the differentiation of thyroid lesions: systematic review and diagnostic meta-analysis. Diagnostic pathology 2012; 7:97

164. Bartolazzi A, Orlandi F, Saggiorato E, Volante M, Arecco F, Rossetto R, Palestini N, Ghigo E, Papotti M, Bussolati G, Martegani MP, Pantellini F, Carpi A, Giovagnoli MR, Monti S, Toscano V, Sciacchitano S, Pennelli GM, Mian C, Pelizzo MR, Rugge M, Troncone G, Palombini L, Chiappetta G, Botti G, Vecchione A, Bellocco R. Galectin-3-expression analysis in the surgical selection of follicular thyroid nodules

with indeterminate fine-needle aspiration cytology: a prospective multicentre study. The Lancet Oncology 2008; 9:543-549

165. Zhang L, Krausz T, DeMay RM. A Pilot Study of Galectin-3, HBME-1, and p27 Triple Immunostaining Pattern for Diagnosis of Indeterminate Thyroid Nodules in Cytology With Correlation to Histology. Applied immunohistochemistry & molecular morphology : AIMM / official publication of the Society for Applied Immunohistochemistry 2015; 23:481-490

166. Collet JF, Hurbain I, Prengel C, Utzmann O, Scetbon F, Bernaudin JF, Fajac A. Galectin-3 immunodetection in follicular thyroid neoplasms: a prospective study on fine-needle aspiration samples. British journal of cancer 2005; 93:1175-1181

167. Coli A, Bigotti G, Parente P, Federico F, Castri F, Massi G. Atypical thyroid nodules express both HBME-1 and Galectin-3, two phenotypic markers of papillary thyroid carcinoma. Journal of experimental & clinical cancer research : CR 2007; 26:221-227

168. Bartolazzi A, Bellotti C, Sciacchitano S. Methodology and technical requirements of the galectin-3 test for the preoperative characterization of thyroid nodules. Applied immunohistochemistry & molecular morphology : AIMM / official publication of the Society for Applied Immunohistochemistry 2012; 20:2-7

169. Rossi ED, Raffaelli M, Mule A, Miraglia A, Lombardi CP, Vecchio FM, Fadda G. Simultaneous immunohistochemical expression of HBME-1 and galectin-3 differentiates papillary carcinomas from hyperfunctioning lesions of the thyroid. Histopathology 2006; 48:795-800

170. Pennelli G, Mian C, Pelizzo MR, Naccamulli D, Piotto A, Girelli ME, Mescoli C, Rugge M. Galectin-3 cytotest in thyroid follicular neoplasia: a prospective, monoinstitutional study. Acta cytologica 2009; 53:533-539

171. Ito Y, Yoshida H, Tomoda C, Miya A, Kobayashi K, Matsuzuka F, Yasuoka H, Kakudo K, Inohara H, Kuma K, Miyauchi A. Galectin-3 expression in follicular tumours: an immunohistochemical study of its use as a marker of follicular carcinoma. Pathology 2005; 37:296-298

172. Oestreicher-Kedem Y, Halpern M, Roizman P, Hardy B, Sulkes J, Feinmesser R, Stern Y. Diagnostic value of galectin-3 as a marker for malignancy in follicular patterned thyroid lesions. Head & neck 2004; 26:960-966

173. Torregrossa L, Faviana P, Camacci T, Materazzi G, Berti P, Minuto M, Elisei R, Vitti P, Miccoli P, Basolo F. Galectin-3 is highly expressed in nonencapsulated papillary thyroid carcinoma but weakly expressed in encapsulated type; comparison with Hector Battifora mesothelial cell 1 immunoreactivity. Human pathology 2007; 38:1482-1488

174. Saggiorato E, De Pompa R, Volante M, Cappia S, Arecco F, Dei Tos AP, Orlandi F, Papotti M. Characterization of thyroid 'follicular neoplasms' in fine-needle aspiration cytological specimens using a panel of immunohistochemical markers: a proposal for clinical application. Endocrine-related cancer 2005; 12:305-317

175. Aiad HA, Kandil MA, Asaad NY, El-Kased AM, El-Goday SF. Galectin-3 immunostaining in cytological and histopathological diagnosis of thyroid lesions. Journal of the Egyptian National Cancer Institute 2008; 20:36-46

176. Asioli S, Maletta F, Pacchioni D, Lupo R, Bussolati G. Cytological detection of papillary thyroid carcinomas by nuclear membrane decoration with emerin staining. Virchows Archiv : an international journal of pathology 2010; 457:43-51

177. Raggio E, Camandona M, Solerio D, Martino P, Franchello A, Orlandi F, Gasparri G. The diagnostic accuracy of the immunocytochemical markers in the pre-operative evaluation of follicular thyroid lesions. Journal of endocrinological investigation 2010; 33:378-381

178. Cheung CC, Ezzat S, Freeman JL, Rosen IB, Asa SL. Immunohistochemical diagnosis of papillary thyroid carcinoma. Modern pathology : an official journal of the United States and Canadian Academy of Pathology, Inc 2001; 14:338-342

179. Fadda G, Rossi ED, Raffaelli M, Pontecorvi A, Sioletic S, Morassi F, Lombardi CP, Zannoni GF, Rindi G. Follicular thyroid neoplasms can be classified as low- and high-risk according to HBME-1 and Galectin-3 expression on liquid-based fine-needle cytology. European journal of endocrinology / European Federation of Endocrine Societies 2011; 165:447-453

180. Lacoste-Collin L, d'Aure D, Berard E, Rouquette I, Delisle MB, Courtade-Saidi M. Improvement of the cytological diagnostic accuracy of follicular thyroid lesions by the use of the Ki-67 proliferative index in addition to cytokeratin-19 and HBME-1 immunomarkers: a study of 61 cases of liquid-based FNA cytology with histological controls. Cytopathology 2014; 25:160-169

181. Rossi ED, Raffaelli M, Minimo C, Mule A, Lombardi CP, Vecchio FM, Fadda G. Immunocytochemical evaluation of thyroid neoplasms on thin-layer smears from fine-needle aspiration biopsies. Cancer 2005; 105:87-95

182. Khurana KK, Truong LD, LiVolsi VA, Baloch ZW. Cytokeratin 19 immunolocalization in cell block preparation of thyroid aspirates. An adjunct to fine-needle aspiration diagnosis of papillary thyroid carcinoma. Archives of pathology & laboratory medicine 2003; 127:579-583

183. Baloch ZW, Abraham S, Roberts S, LiVolsi VA. Differential expression of cytokeratins in follicular variant of papillary carcinoma: an immunohistochemical study and its diagnostic utility. Human pathology 1999; 30:1166-1171

184. Kjellman P, Wallin G, Hoog A, Auer G, Larsson C, Zedenius J. MIB-1 index in thyroid tumors: a predictor of the clinical course in papillary thyroid carcinoma. Thyroid 2003; 13:371-380

185. Pisani T, Pantellini F, Centanni M, Vecchione A, Giovagnoli MR. Immunocytochemical expression of Ki67 and laminin in Hurthle cell adenomas and carcinomas. Anticancer research 2003; 23:3323-3326

186. Addati T, Achille G, Centrone M, Petroni S, Popescu O, Russo S, Grammatica L, Simone G. TROP-2 expression in papillary thyroid cancer: a preliminary cyto-histological study. Cytopathology 2015;

187. Coban I, Cakir A, Unal TD, Bassullu N, Karpuz V, Dogusoy GB, Alper M. Emerin expression in well differentiated epithelial lesions of thyroid: implications in papillary thyroid carcinoma diagnosis and predicting malignant behavior. Pathology oncology research : POR 2015; 21:357-366

188. Chandan VS, Faquin WC, Wilbur DC, Khurana KK. The role of immunolocalization of CD57 and GLUT-1 in cell blocks in fine-needle aspiration diagnosis of papillary thyroid carcinoma. Cancer 2006; 108:331-336

189. Troncone G, Volante M, Iaccarino A, Zeppa P, Cozzolino I, Malapelle U, Palmieri EA, Conzo G, Papotti M, Palombini L. Cyclin D1 and D3 overexpression predicts malignant behavior in thyroid fine-needle aspirates suspicious for Hurthle cell neoplasms. Cancer 2009; 117:522-529

190. Mills LJ, Poller DN, Yiangou C. Galectin-3 is not useful in thyroid FNA. Cytopathology 2005; 16:132-138

191. Brito JP, Gionfriddo MR, Al Nofal A, Boehmer KR, Leppin AL, Reading C, Callstrom M, Elraiyah TA, Prokop LJ, Stan MN, Murad MH, Morris JC, Montori VM. The accuracy of thyroid nodule ultrasound to predict thyroid cancer: systematic

review and meta-analysis. The Journal of clinical endocrinology and metabolism 2014; 99:1253-1263

192. Remonti LR, Kramer CK, Leitao CB, Pinto LC, Gross JL. Thyroid ultrasound features and risk of carcinoma: a systematic review and meta-analysis of observational studies. Thyroid 2015; 25:538-550

193. Jeh SK, Jung SL, Kim BS, Lee YS. Evaluating the degree of conformity of papillary carcinoma and follicular carcinoma to the reported ultrasonographic findings of malignant thyroid tumor. Korean journal of radiology 2007; 8:192-197

194. Lee SH, Baek JS, Lee JY, Lim JA, Cho SY, Lee TH, Ku YH, Kim HI, Kim MJ. Predictive factors of malignancy in thyroid nodules with a cytological diagnosis of follicular neoplasm. Endocrine pathology 2013; 24:177-183

195. Kim DS, Kim JH, Na DG, Park SH, Kim E, Chang KH, Sohn CH, Choi YH. Sonographic features of follicular variant papillary thyroid carcinomas in comparison with conventional papillary thyroid carcinomas. Journal of ultrasound in medicine : official journal of the American Institute of Ultrasound in Medicine 2009; 28:1685-1692

196. Ozdemir D, Ersoy R, Cuhaci N, Arpaci D, Ersoy EP, Korukluoglu B, Guler G, Cakir B. Classical and follicular variant papillary thyroid carcinoma: comparison of clinical, ultrasonographical, cytological, and histopathological features in 444 patients. Endocrine pathology 2011; 22:58-65

197. Brophy C, Stewart J, O'Donovan N, McCarthy J, Murphy M, Sheahan P. Impact of Microcalcifications on Risk of Malignancy in Thyroid Nodules with Indeterminate or Benign Cytology. Otolaryngology--head and neck surgery : official journal of American Academy of Otolaryngology-Head and Neck Surgery 2016; 154:46-51

198. Gweon HM, Son EJ, Youk JH, Kim JA. Thyroid nodules with Bethesda system III cytology: can ultrasonography guide the next step? Annals of surgical oncology 2013; 20:3083-3088

199. Lee KH, Shin JH, Oh YL, Hahn SY. Atypia of undetermined significance in thyroid fine-needle aspiration cytology: prediction of malignancy by US and comparison of methods for further management. Annals of surgical oncology 2014; 21:2326-2331

200. Pompili G, Tresoldi S, Primolevo A, De Pasquale L, Di Leo G, Cornalba G. Management of thyroid follicular proliferation: an ultrasound-based malignancy score to opt for surgical or conservative treatment. Ultrasound in medicine & biology 2013; 39:1350-1355

201. Yoo WS, Choi HS, Cho SW, Moon JH, Kim KW, Park HJ, Park SY, Choi SI, Choi SH, Lim S, Yi KH, Park do J, Jang HC, Park YJ. The role of ultrasound findings in the management of thyroid nodules with atypia or follicular lesions of undetermined significance. Clinical endocrinology 2014; 80:735-742

202. Chng CL, Kurzawinski TR, Beale T. Value of sonographic features in predicting malignancy in thyroid nodules diagnosed as follicular neoplasm on cytology. Clinical endocrinology 2014;

203. Tutuncu Y, Berker D, Isik S, Akbaba G, Ozuguz U, Kucukler FK, Gocmen E, Yalcin Y, Aydin Y, Guler S. The frequency of malignancy and the relationship between malignancy and ultrasonographic features of thyroid nodules with indeterminate cytology. Endocrine 2014; 45:37-45

204. Yoon JH, Kwak JY, Kim EK, Moon HJ, Kim MJ, Kim JY, Koo HR, Kim MH. How to approach thyroid nodules with indeterminate cytology. Annals of surgical oncology 2010; 17:2147-2155

205. Khoncarly SM, Tamarkin SW, McHenry CR. Can ultrasound be used to predict malignancy in patients with a thyroid nodule and an indeterminate fine-needle aspiration biopsy? Surgery 2014; 156:967-970

206. Yoon JH, Lee HS, Kim EK, Moon HJ, Kwak JY. A nomogram for predicting malignancy in thyroid nodules diagnosed as atypia of undetermined significance/follicular lesions of undetermined significance on fine needle aspiration. Surgery 2014; 155:1006-1013

207. Dutta S, Thaha MA, Smith DM. Do sonographic and cytological features predict malignancy in cytologically indeterminate thyroid nodules? Annals of the Royal College of Surgeons of England 2011; 93:361-364

208. Parikh PP, Allan BJ, Lew JI. Surgeon-performed ultrasound predictors of malignancy in patients with Hurthle cell neoplasms of the thyroid. The Journal of surgical research 2013; 184:247-252

209. Batawil N, Alkordy T. Ultrasonographic features associated with malignancy in cytologically indeterminate thyroid nodules. European journal of surgical oncology : the journal of the European Society of Surgical Oncology and the British Association of Surgical Oncology 2014; 40:182-186

210. Garino F, Deandrea M, Motta M, Mormile A, Ragazzoni F, Palestini N, Freddi M, Gasparri G, Sgotto E, Pacchioni D, Limone PP. Diagnostic performance of elastography in cytologically indeterminate thyroid nodules. Endocrine 2015; 49:175-183

211. Macias CA, Arumugam D, Arlow RL, Eng OS, Lu SE, Javidian P, Davidov T, Trooskin SZ. A risk model to determine surgical treatment in patients with thyroid nodules with indeterminate cytology. Annals of surgical oncology 2015; 22:1527-1532

212. Turanli S, Pirhan Y, Ozcelik CK, Cetin A. Predictors of malignancy in patients with a thyroid nodule that contains Hurthle cells. Otolaryngology--head and neck surgery : official journal of American Academy of Otolaryngology-Head and Neck Surgery 2011; 144:514-517

213. Lippolis PV, Tognini S, Materazzi G, Polini A, Mancini R, Ambrosini CE, Dardano A, Basolo F, Seccia M, Miccoli P, Monzani F. Is elastography actually useful in the presurgical selection of thyroid nodules with indeterminate cytology? The Journal of clinical endocrinology and metabolism 2011; 96:E1826-1830

214. Kayilioglu SI, Dinc T, Sozen I, Senol K, Katar K, Karabeyoglu M, Tez M, Coskun F. Thyroid nodules with atypia or follicular lesions of undetermined significance (AUS/FLUS): analysis of variables associated with outcome. Asian Pacific journal of cancer prevention : APJCP 2014; 15:10307-10311

215. Cakir B, Ersoy R, Cuhaci FN, Aydin C, Polat B, Kilic M, Yazgan A. Elastosonographic strain index in thyroid nodules with atypia of undetermined significance. Journal of endocrinological investigation 2014; 37:127-133

216. Calo PG, Medas F, Santa Cruz R, Podda F, Erdas E, Pisano G, Nicolosi A. Follicular nodules (Thy3) of the thyroid: is total thyroidectomy the best option? BMC surgery 2014; 14:12

217. Cantisani V, Maceroni P, D'Andrea V, Patrizi G, Di Segni M, De Vito C, Grazhdani H, Isidori AM, Giannetta E, Redler A, Frattaroli F, Giacomelli L, Di Rocco G, Catalano C, D'Ambrosio F. Strain ratio ultrasound elastography increases the accuracy of colour-Doppler ultrasound in the evaluation of Thy-3 nodules. A bi-centre university experience. European radiology 2015;

218. Rago T, Di Coscio G, Basolo F, Scutari M, Elisei R, Berti P, Miccoli P, Romani R, Faviana P, Pinchera A, Vitti P. Combined clinical, thyroid ultrasound and cytological

features help to predict thyroid malignancy in follicular and Hupsilonrthle cell thyroid lesions: results from a series of 505 consecutive patients. Clinical endocrinology 2007; 66:13-20

219. Cantisani V, Ulisse S, Guaitoli E, De Vito C, Caruso R, Mocini R, D'Andrea V, Ascoli V, Antonaci A, Catalano C, Nardi F, Redler A, Ricci P, De Antoni E, Sorrenti S. Q-elastography in the presurgical diagnosis of thyroid nodules with indeterminate cytology. PloS one 2012; 7:e50725

220. Matthey-Gie ML, Walsh SM, O'Neill AC, Lowery A, Evoy D, Gibbons D, Prichard RS, Skehan S, McDermott EW. Ultrasound predictors of malignancy in indeterminate thyroid nodules. Irish journal of medical science 2014; 183:633-637

221. Mohey N, Hassan TA, Abdel-Baki S. Role of combined grey scale US and US tissue elastography in differentiating solid thyroid nodules. Egyptian Journal of Radiology and Nuclear Medicine 2013; 44:505-512

222. Rago T, Scutari M, Santini F, Loiacono V, Piaggi P, Di Coscio G, Basolo F, Berti P, Pinchera A, Vitti P. Real-time elastosonography: useful tool for refining the presurgical diagnosis in thyroid nodules with indeterminate or nondiagnostic cytology. The Journal of clinical endocrinology and metabolism 2010; 95:5274-5280

223. Lee KH, Shin JH, Ko ES, Hahn SY, Kim JS, Kim JH, Oh YL. Predictive factors of malignancy in patients with cytologically suspicious for Hurthle cell neoplasm of thyroid nodules. International journal of surgery (London, England) 2013; 11:898-902

224. Dogan L, Gulcelik MA, Karaman N, Baskan E, Kahraman YS, Aksel B, Ozaslan C. The contribution of clinical and radiological features to the diagnosis in AUS/FLUS and FN/SFN thyroid nodules. Acta Medica Mediterranea 2016; 32:151-155

225. Moon HG, Jung EJ, Park ST, Ha WS, Choi SK, Hong SC, Lee YJ, Joo YT, Jeong CY, Choi DS, Ryoo JW. Role of ultrasonography in predicting malignancy in patients with thyroid nodules. World journal of surgery 2007; 31:1410-1416

226. Alexander EK, Marqusee E, Orcutt J, Benson CB, Frates MC, Doubilet PM, Cibas ES, Atri A. Thyroid nodule shape and prediction of malignancy. Thyroid 2004; 14:953-958

227. Chin C, Franklin J, Sowerby L, Fung K, Yoo J. Stratification of intermediate-risk fine-needle aspiration biopsies. Journal of otolaryngology - head & neck surgery = Le Journal d'oto-rhino-laryngologie et de chirurgie cervico-faciale 2010; 39:393-396

228. Kim DW, Lee EJ, Jung SJ, Ryu JH, Kim YM. Role of sonographic diagnosis in managing Bethesda class III nodules. AJNR American journal of neuroradiology 2011; 32:2136-2141

229. Lee MJ, Hong SW, Chung WY, Kwak JY, Kim MJ, Kim EK. Cytological results of ultrasound-guided fine-needle aspiration cytology for thyroid nodules: emphasis on correlation with sonographic findings. Yonsei medical journal 2011; 52:838-844

230. Maia FF, Matos PS, Pavin EJ, Vassallo J, Zantut-Wittmann DE. Value of ultrasound and cytological classification system to predict the malignancy of thyroid nodules with indeterminate cytology. Endocrine pathology 2011; 22:66-73

231. Norlen O, Popadich A, Kruijff S, Gill AJ, Sarkis LM, Delbridge L, Sywak M, Sidhu S. Bethesda III thyroid nodules: the role of ultrasound in clinical decision making. Annals of surgical oncology 2014; 21:3528-3533

232. Rosario PW, Salles DS, Bessa B, Purisch S. Contribution of scintigraphy and ultrasonography to the prediction of malignancy in thyroid nodules with indeterminate cytology. Arquivos brasileiros de endocrinologia e metabologia 2010; 54:56-59

233. Rosario PW. Thyroid nodules with atypia or follicular lesions of undetermined significance (Bethesda Category III): importance of ultrasonography and cytological subcategory. Thyroid 2014; 24:1115-1120

234. Gulcelik NE, Gulcelik MA, Kuru B. Risk of malignancy in patients with follicular neoplasm: predictive value of clinical and ultrasonographic features. Archives of otolaryngology--head & neck surgery 2008; 134:1312-1315

235. Kamaya A, Lewis GH, Liu Y, Akatsu H, Kong C, Desser TS. Atypia of undetermined significance and follicular lesions of undetermined significance: sonographic assessment for prediction of the final diagnosis. Journal of ultrasound in medicine : official journal of the American Institute of Ultrasound in Medicine 2015; 34:767-774

236. Kwak JY, Han KH, Yoon JH, Moon HJ, Son EJ, Park SH, Jung HK, Choi JS, Kim BM, Kim EK. Thyroid imaging reporting and data system for US features of nodules: a step in establishing better stratification of cancer risk. Radiology 2011; 260:892-899

237. Horvath E, Majlis S, Rossi R, Franco C, Niedmann JP, Castro A, Dominguez M. An ultrasonogram reporting system for thyroid nodules stratifying cancer risk for clinical management. The Journal of clinical endocrinology and metabolism 2009; 94:1748-1751

238. Maia FF, Matos PS, Pavin EJ, Zantut-Wittmann DE. Thyroid imaging reporting and data system score combined with Bethesda system for malignancy risk stratification in thyroid nodules with indeterminate results on cytology. Clinical endocrinology 2015; 82:439-444

239. Park VY, Kim EK, Kwak JY, Yoon JH, Moon HJ. Malignancy risk and characteristics of thyroid nodules with two consecutive results of atypia of undetermined significance or follicular lesion of undetermined significance on cytology. European radiology 2015; 25:2601-2607

240. Russ G, Royer B, Bigorgne C, Rouxel A, Bienvenu-Perrard M, Leenhardt L. Prospective evaluation of thyroid imaging reporting and data system on 4550 nodules with and without elastography. European journal of endocrinology / European Federation of Endocrine Societies 2013; 168:649-655

241. Yoon JH, Kwon HJ, Kim EK, Moon HJ, Kwak JY. Subcategorization of atypia of undetermined significance/follicular lesion of undetermined significance (AUS/FLUS): a study applying Thyroid Imaging Reporting and Data System (TIRADS). Clinical endocrinology 2016;

242. Raber W, Kaserer K, Niederle B, Vierhapper H. Risk factors for malignancy of thyroid nodules initially identified as follicular neoplasia by fine-needle aspiration: results of a prospective study of one hundred twenty patients. Thyroid 2000; 10:709-712

243. Lyshchik A, Higashi T, Asato R, Tanaka S, Ito J, Mai JJ, Pellot-Barakat C, Insana MF, Brill AB, Saga T, Hiraoka M, Togashi K. Thyroid gland tumor diagnosis at US elastography. Radiology 2005; 237:202-211

244. Itoh A, Ueno E, Tohno E, Kamma H, Takahashi H, Shiina T, Yamakawa M, Matsumura T. Breast disease: clinical application of US elastography for diagnosis. Radiology 2006; 239:341-350

245. Rago T, Santini F, Scutari M, Pinchera A, Vitti P. Elastography: new developments in ultrasound for predicting malignancy in thyroid nodules. The Journal of clinical endocrinology and metabolism 2007; 92:2917-2922

246. Asteria C, Giovanardi A, Pizzocaro A, Cozzaglio L, Morabito A, Somalvico F, Zoppo A. US-elastography in the differential diagnosis of benign and malignant thyroid nodules. Thyroid 2008; 18:523-531

247. Veer V, Puttagunta S. The role of elastography in evaluating thyroid nodules: a literature review and meta-analysis. European archives of oto-rhino-laryngology : official journal of the European Federation of Oto-Rhino-Laryngological Societies (EUFOS) : affiliated with the German Society for Oto-Rhino-Laryngology - Head and Neck Surgery 2015; 272:1845-1855

248. Nell S, Kist JW, Debray TP, de Keizer B, van Oostenbrugge TJ, Borel Rinkes IH, Valk GD, Vriens MR. Qualitative elastography can replace thyroid nodule fine-needle aspiration in patients with soft thyroid nodules. A systematic review and meta-analysis. European journal of radiology 2015; 84:652-661

249. Abdelrahman SF, Ali FH, El-Sayed Khalil M, El Masry MR. Ultrasound elastography in the diagnostic evaluation of indeterminate thyroid nodule s. Egyptian Journal of Radiology and Nuclear Medicine 2015; 46:639-648

250. Trimboli P, Treglia G, Sadeghi R, Romanelli F, Giovanella L. Reliability of real-time elastography to diagnose thyroid nodules previously read at FNAC as indeterminate: a meta-analysis. Endocrine 2014;

251. Park SH, Kim SJ, Kim EK, Kim MJ, Son EJ, Kwak JY. Interobserver agreement in assessing the sonographic and elastographic features of malignant thyroid nodules. AJR American journal of roentgenology 2009; 193:W416-423

252. Sebag F, Vaillant-Lombard J, Berbis J, Griset V, Henry JF, Petit P, Oliver C. Shear wave elastography: a new ultrasound imaging mode for the differential diagnosis of benign and malignant thyroid nodules. The Journal of clinical endocrinology and metabolism 2010; 95:5281-5288

253. Kim DW. Computed tomography features of papillary thyroid carcinomas. Journal of computer assisted tomography 2014; 38:936-940

254. Lee C, Chalmers B, Treister D, Adhya S, Godwin B, Ji L, Groshen S, Grant E. Thyroid lesions visualized on CT: sonographic and pathologic correlation. Academic radiology 2015; 22:203-209

255. Treglia G, Caldarella C, Saggiorato E, Ceriani L, Orlandi F, Salvatori M, Giovanella L. Diagnostic performance of (99m)Tc-MIBI scan in predicting the malignancy of thyroid nodules: a meta-analysis. Endocrine 2013; 44:70-78

256. Gharib H, Papini E. Thyroid nodules: clinical importance, assessment, and treatment. Endocrinology and metabolism clinics of North America 2007; 36:707-735, vi

257. Lumachi F, Varotto L, Borsato S, Tregnaghi A, Zucchetta P, Marzola MC, Cecchin D, Bui F. Usefulness of 99mTc-pertechnetate scintigraphy and fine-needle aspiration cytology in patients with solitary thyroid nodules and thyroid cancer. Anticancer research 2004; 24:2531-2534

258. Saggiorato E, Angusti T, Rosas R, Martinese M, Finessi M, Arecco F, Trevisiol E, Bergero N, Puligheddu B, Volante M, Podio V, Papotti M, Orlandi F. 99mTc-MIBI Imaging in the presurgical characterization of thyroid follicular neoplasms: relationship to multidrug resistance protein expression. Journal of nuclear medicine : official publication, Society of Nuclear Medicine 2009; 50:1785-1793

259. Vattimo A, Bertelli P, Cintorino M, Burroni L, Volterrani D, Vella A. Identification of Hurthle cell tumor by single-injection, double-phase scintigraphy with technetium-99m-sestamibi. Journal of nuclear medicine : official publication, Society of Nuclear Medicine 1995; 36:778-782

260. Boi F, Lai ML, Deias C, Piga M, Serra A, Uccheddu A, Faa G, Mariotti S. The usefulness of 99mTc-SestaMIBI scan in the diagnostic evaluation of thyroid nodules with oncocytic cytology. European journal of endocrinology / European Federation of Endocrine Societies 2003; 149:493-498

261. Hindie E, Ugur O, Fuster D, O'Doherty M, Grassetto G, Urena P, Kettle A, Gulec SA, Pons F, Rubello D, Parathyroid Task Group of the E. 2009 EANM parathyroid guidelines. European journal of nuclear medicine and molecular imaging 2009; 36:1201-1216

262. Wale A, Miles KA, Young B, Zammit C, Williams A, Quin J, Dizdarevic S. Combined (99m)Tc-methoxyisobutylisonitrile scintigraphy and fine-needle aspiration cytology offers an accurate and potentially cost-effective investigative strategy for the assessment of solitary or dominant thyroid nodules. European journal of nuclear medicine and molecular imaging 2014; 41:105-115

263. Madorin CA, Owen R, Coakley B, Lowe H, Nam KH, Weber K, Kushnir L, Rios J, Genden E, Pawha PS, Inabnet WB, 3rd. Comparison of radiation exposure and cost between dynamic computed tomography and sestamibi scintigraphy for preoperative localization of parathyroid lesions. JAMA Surg 2013; 148:500-503

264. Heller M, Zanocco K, Zydowicz S, Elaraj D, Nayar R, Sturgeon C. Cost-effectiveness analysis of repeat fine-needle aspiration for thyroid biopsies read as atypia of undetermined significance. Surgery 2012; 152:423-430

265. Upadhyay M, Samal J, Kandpal M, Singh OV, Vivekanandan P. The Warburg effect: insights from the past decade. Pharmacol Ther 2013; 137:318-330

266. Mitchell JC, Grant F, Evenson AR, Parker JA, Hasselgren PO, Parangi S. Preoperative evaluation of thyroid nodules with 18FDG-PET/CT. Surgery 2005; 138:1166-1174; discussion 1174-1165

267. Vriens D, de Wilt JH, van der Wilt GJ, Netea-Maier RT, Oyen WJ, de Geus-Oei LF. The role of [18F]-2-fluoro-2-deoxy-d-glucose-positron emission tomography in thyroid nodules with indeterminate fine-needle aspiration biopsy: systematic review and meta-analysis of the literature. Cancer 2011; 117:4582-4594

268. Kresnik E, Gallowitsch HJ, Mikosch P, Stettner H, Igerc I, Gomez I, Kumnig G, Lind P. Fluorine-18-fluorodeoxyglucose positron emission tomography in the preoperative assessment of thyroid nodules in an endemic goiter area. Surgery 2003; 133:294-299

269. de Geus-Oei LF, Pieters GF, Bonenkamp JJ, Mudde AH, Bleeker-Rovers CP, Corstens FH, Oyen WJ. 18F-FDG PET reduces unnecessary hemithyroidectomies for thyroid nodules with inconclusive cytologic results. Journal of nuclear medicine : official publication, Society of Nuclear Medicine 2006; 47:770-775

270. Sebastianes FM, Cerci JJ, Zanoni PH, Soares J, Jr., Chibana LK, Tomimori EK, de Camargo RY, Izaki M, Giorgi MC, Eluf-Neto J, Meneghetti JC, Pereira MA. Role of 18F-fluorodeoxyglucose positron emission tomography in preoperative assessment of cytologically indeterminate thyroid nodules. The Journal of clinical endocrinology and metabolism 2007; 92:4485-4488

271. Traugott AL, Dehdashti F, Trinkaus K, Cohen M, Fialkowski E, Quayle F, Hussain H, Davila R, Ylagan L, Moley JF. Exclusion of malignancy in thyroid nodules with indeterminate fine-needle aspiration cytology after negative 18F-fluorodeoxyglucose positron emission tomography: interim analysis. World journal of surgery 2010; 34:1247-1253

272. Hales NW, Krempl GA, Medina JE. Is there a role for fluorodeoxyglucose positron emission tomography/computed tomography in cytologically indeterminate thyroid nodules? American journal of otolaryngology 2008; 29:113-118

273. Deandreis D, Al Ghuzlan A, Auperin A, Vielh P, Caillou B, Chami L, Lumbroso J, Travagli JP, Hartl D, Baudin E, Schlumberger M, Leboulleux S. Is (18)F-fluorodeoxyglucose-PET/CT useful for the presurgical characterization of thyroid nodules with indeterminate fine needle aspiration cytology? Thyroid 2012; 22:165-172

274. Smith RB, Robinson RA, Hoffman HT, Graham MM. Preoperative FDG-PET imaging to assess the malignant potential of follicular neoplasms of the thyroid. Otolaryngology--head and neck surgery : official journal of American Academy of Otolaryngology-Head and Neck Surgery 2008; 138:101-106

275. Valderrabano P, Montilla-Soler J, Mifsud M, Leon M, Centeno B, Khazai L, Padhya T, McCaffrey T, Russell J, McIver B, Otto K. Hypermetabolism on (18)F-Fluorodeoxyglucose Positron Emission Tomography Scan Does Not Influence the Interpretation of Thyroid Cytopathology, and Nodules with a SUVmax <2.5 Are Not at Increased Risk for Malignancy. Thyroid 2016; 26:1300-1307

276. Kim JM, Ryu JS, Kim TY, Kim WB, Kwon GY, Gong G, Moon DH, Kim SC, Hong SJ, Shong YK. 18F-fluorodeoxyglucose positron emission tomography does not predict malignancy in thyroid nodules cytologically diagnosed as follicular neoplasm. The Journal of clinical endocrinology and metabolism 2007; 92:1630-1634

277. Boellaard R, Delgado-Bolton R, Oyen WJ, Giammarile F, Tatsch K, Eschner W, Verzijlbergen FJ, Barrington SF, Pike LC, Weber WA, Stroobants S, Delbeke D, Donohoe KJ, Holbrook S, Graham MM, Testanera G, Hoekstra OS, Zijlstra J, Visser E, Hoekstra CJ, Pruim J, Willemsen A, Arends B, Kotzerke J, Bockisch A, Beyer T, Chiti A, Krause BJ. FDG PET/CT: EANM procedure guidelines for tumour imaging: version 2.0. European journal of nuclear medicine and molecular imaging 2015; 42:328-354

278. Vriens D, Visser EP, de Geus-Oei LF, Oyen WJ. Methodological considerations in quantification of oncological FDG PET studies. European journal of nuclear medicine and molecular imaging 2010; 37:1408-1425

279. Pathak KA, Klonisch T, Nason RW, Leslie WD. FDG-PET characteristics of Hurthle cell and follicular adenomas. Annals of nuclear medicine 2016; 30:506-509

280. Boellaard R, O'Doherty MJ, Weber WA, Mottaghy FM, Lonsdale MN, Stroobants SG, Oyen WJ, Kotzerke J, Hoekstra OS, Pruim J, Marsden PK, Tatsch K, Hoekstra CJ, Visser EP, Arends B, Verzijlbergen FJ, Zijlstra JM, Comans EF, Lammertsma AA, Paans AM, Willemsen AT, Beyer T, Bockisch A, Schaefer-Prokop C, Delbeke D, Baum RP, Chiti A, Krause BJ. FDG PET and PET/CT: EANM procedure guidelines for tumour PET imaging: version 1.0. European journal of nuclear medicine and molecular imaging 2010; 37:181-200

281. Treglia G, Taralli S, Salsano M, Muoio B, Sadeghi R, Giovanella L. Prevalence and malignancy risk of focal colorectal incidental uptake detected by (18)F-FDG-PET or PET/CT: a meta-analysis. Radiology and oncology 2014; 48:99-104

282. Nayan S, Ramakrishna J, Gupta MK. The Proportion of Malignancy in Incidental Thyroid Lesions on 18-FDG PET Study: A Systematic Review and Meta-analysis. Otolaryngology--head and neck surgery : official journal of American Academy of Otolaryngology-Head and Neck Surgery 2014; 151:190-200

283. Nakahira M, Saito N, Murata S, Sugasawa M, Shimamura Y, Morita K, Takajyo F, Omura G, Matsumura S. Quantitative diffusion-weighted magnetic resonance imaging as a powerful adjunct to fine needle aspiration cytology for assessment of thyroid nodules. American journal of otolaryngology 2012; 33:408-416

284. Brown AM, Nagala S, McLean MA, Lu Y, Scoffings D, Apte A, Gonen M, Stambuk HE, Shaha AR, Tuttle RM, Deasy JO, Priest AN, Jani P, Shukla-Dave A, Griffiths J. Multi-institutional validation of a novel textural analysis tool for preoperative stratification of suspected thyroid tumors on diffusion-weighted MRI. Magnetic resonance in medicine 2015;

285. Razek AA, Sadek AG, Kombar OR, Elmahdy TE, Nada N. Role of apparent diffusion coefficient values in differentiation between malignant and benign solitary thyroid nodules. AJNR American journal of neuroradiology 2008; 29:563-568

286. Chen X, Li WL, Zhang YL, Wu Q, Guo YM, Bai ZL. Meta-analysis of quantitative diffusion-weighted MR imaging in the differential diagnosis of breast lesions. BMC cancer 2010; 10:693

287. Wang J, Takashima S, Takayama F, Kawakami S, Saito A, Matsushita T, Momose M, Ishiyama T. Head and neck lesions: characterization with diffusion-weighted echo-planar MR imaging. Radiology 2001; 220:621-630

288. Bozgeyik Z, Coskun S, Dagli AF, Ozkan Y, Sahpaz F, Ogur E. Diffusion-weighted MR imaging of thyroid nodules. Neuroradiology 2009; 51:193-198

289. Dilli A, Ayaz UY, Cakir E, Cakal E, Gultekin SS, Hekimoglu B. The efficacy of apparent diffusion coefficient value calculation in differentiation between malignant and benign thyroid nodules. Clinical imaging 2012; 36:316-322

290. Erdem G, Erdem T, Muammer H, Mutlu DY, Firat AK, Sahin I, Alkan A. Diffusion-weighted images differentiate benign from malignant thyroid nodules. Journal of magnetic resonance imaging : JMRI 2010; 31:94-100

291. Mutlu H, Sivrioglu AK, Sonmez G, Velioglu M, Sildiroglu HO, Basekim CC, Kizilkaya E. Role of apparent diffusion coefficient values and diffusion-weighted magnetic resonance imaging in differentiation between benign and malignant thyroid nodules. Clinical imaging 2012; 36:1-7

292. Schueller-Weidekamm C, Kaserer K, Schueller G, Scheuba C, Ringl H, Weber M, Czerny C, Herneth AM. Can quantitative diffusion-weighted MR imaging differentiate benign and malignant cold thyroid nodules? Initial results in 25 patients. AJNR American journal of neuroradiology 2009; 30:417-422

293. Chen L, Xu J, Bao J, Huang X, Hu X, Xia Y, Wang J. Diffusion-weighted MRI in differentiating malignant from benign thyroid nodules: a meta-analysis. BMJ Open 2016; 6:e008413

294. Friedrich-Rust M, Meyer G, Dauth N, Berner C, Bogdanou D, Herrmann E, Zeuzem S, Bojunga J. Interobserver agreement of Thyroid Imaging Reporting and Data System (TIRADS) and strain elastography for the assessment of thyroid nodules. PloS one 2013; 8:e77927

295. Chou CK, Chen RF, Chou FF, Chang HW, Chen YJ, Lee YF, Yang KD, Cheng JT, Huang CC, Liu RT. miR-146b is highly expressed in adult papillary thyroid carcinomas with high risk features including extrathyroidal invasion and the BRAF(V600E) mutation. Thyroid 2010; 20:489-494

296. National Cancer Institute. The website of the National Cancer Institute. https://www.cancer.gov, 2017.

297. Cancer Genome Atlas Research N. Integrated genomic characterization of papillary thyroid carcinoma. Cell 2014; 159:676-690

298. Hyman DM, Puzanov I, Subbiah V, Faris JE, Chau I, Blay JY, Wolf J, Raje NS, Diamond EL, Hollebecque A, Gervais R, Elez-Fernandez ME, Italiano A, Hofheinz RD, Hidalgo M, Chan E, Schuler M, Lasserre SF, Makrutzki M, Sirzen F, Veronese ML, Tabernero J, Baselga J. Vemurafenib in Multiple Nonmelanoma Cancers with BRAF V600 Mutations. The New England journal of medicine 2015; 373:726-736

299. Chapman PB, Hauschild A, Robert C, Haanen JB, Ascierto P, Larkin J, Dummer R, Garbe C, Testori A, Maio M, Hogg D, Lorigan P, Lebbe C, Jouary T, Schadendorf D, Ribas A, O'Day SJ, Sosman JA, Kirkwood JM, Eggermont AM, Dreno B, Nolop K, Li J, Nelson B, Hou J, Lee RJ, Flaherty KT, McArthur GA, Group B-S. Improved

survival with vemurafenib in melanoma with BRAF V600E mutation. The New England journal of medicine 2011; 364:2507-2516

300. Pagan M, Kloos RT, Lin CF, Travers KJ, Matsuzaki H, Tom EY, Kim SY, Wong MG, Stewart AC, Huang J, Walsh PS, Monroe RJ, Kennedy GC. The diagnostic application of RNA sequencing in patients with thyroid cancer: an analysis of 851 variants and 133 fusions in 524 genes. BMC bioinformatics 2016; 17 Suppl 1:6

301. Ciregia F, Giusti L, Molinaro A, Niccolai F, Agretti P, Rago T, Di Coscio G, Vitti P, Basolo F, Iacconi P, Tonacchera M, Lucacchini A. Presence in the pre-surgical fine-needle aspiration of potential thyroid biomarkers previously identified in the post-surgical one. PloS one 2013; 8:e72911

302. Pagni F, Mainini V, Garancini M, Bono F, Vanzati A, Giardini V, Scardilli M, Goffredo P, Smith AJ, Galli M, De Sio G, Magni F. Proteomics for the diagnosis of thyroid lesions: preliminary report. Cytopathology 2014;

303. Bossuyt PM, Reitsma JB, Linnet K, Moons KG. Beyond diagnostic accuracy: the clinical utility of diagnostic tests. Clinical chemistry 2012; 58:1636-1643

304. Nikiforov YE, Seethala RR, Tallini G, Baloch ZW, Basolo F, Thompson LD, Barletta JA, Wenig BM, Al Ghuzlan A, Kakudo K, Giordano TJ, Alves VA, Khanafshar E, Asa SL, El-Naggar AK, Gooding WE, Hodak SP, Lloyd RV, Maytal G, Mete O, Nikiforova MN, Nose V, Papotti M, Poller DN, Sadow PM, Tischler AS, Tuttle RM, Wall KB, LiVolsi VA, Randolph GW, Ghossein RA. Nomenclature Revision for Encapsulated Follicular Variant of Papillary Thyroid Carcinoma: A Paradigm Shift to Reduce Overtreatment of Indolent Tumors. JAMA Oncol 2016; 2:1023-1029

305. Cerutti JM, Oler G, Delcelo R, Gerardt R, Michaluart P, Jr., de Souza SJ, Galante PA, Huang P, Riggins GJ. PVALB, a new Hurthle adenoma diagnostic marker identified

through gene expression. The Journal of clinical endocrinology and metabolism 2011; 96:E151-160

306. Evangelisti C, de Biase D, Kurelac I, Ceccarelli C, Prokisch H, Meitinger T, Caria P, Vanni R, Romeo G, Tallini G, Gasparre G, Bonora E. A mutation screening of oncogenes, tumor suppressor gene TP53 and nuclear encoded mitochondrial complex I genes in oncocytic thyroid tumors. BMC cancer 2015; 15:157

307. Finley DJ, Zhu B, Fahey TJ, 3rd. Molecular analysis of Hurthle cell neoplasms by gene profiling. Surgery 2004; 136:1160-1168

308. Ganly I, Ricarte Filho J, Eng S, Ghossein R, Morris LG, Liang Y, Socci N, Kannan K, Mo Q, Fagin JA, Chan TA. Genomic dissection of Hurthle cell carcinoma reveals a unique class of thyroid malignancy. The Journal of clinical endocrinology and metabolism 2013; 98:E962-972

309. Corver WE, Ruano D, Weijers K, den Hartog WC, van Nieuwenhuizen MP, de Miranda N, van Eijk R, Middeldorp A, Jordanova ES, Oosting J, Kapiteijn E, Hovens G, Smit J, van Wezel T, Morreau H. Genome haploidisation with chromosome 7 retention in oncocytic follicular thyroid carcinoma. PloS one 2012; 7:e38287

Table 1. Overview of classification systems for thyroid cytology

Bethesda System for the Reporting of Thyroid Cytology (4)		**British Thyroid Association (BTA)(6)**		**SIAPEC-IAP (Italy)(7)**			
Category	*Description*	*Category*	*Description*	*Category*	*Description*	**Malignancy rate (4)**	**Proposed management (2015 ATA guidelines (23))**
I	Nondiagnostic / unsatisfactory	Thy1 Thy1c	Nondiagnostic Nondiagnostic Cystic lesion	TIR1 TIR1c	Nondiagnostic Nondiagnostic-cystic	1%-4%	Repeat FNAC with US guidance
II	Benign	Thy2 Thy2c	Nonneoplastic Nonneoplastic Cystic lesion	TIR2	Nonmalignant / benign	0%-3%	No clinical follow-up or treatment required
III	Atypia of undetermined significance / follicular lesion of undetermined significance (AUS/FLUS)	Thy3a	Atypical features present	TIR3a	Low-risk indeterminate lesion	~5%-15%	Repeat FNAC. If second Bethesda III result, consider additional tests and/or diagnostic hemithyroidectomy
IV	Follicular neoplasm / suspicious of a follicular neoplasm, including Hürthle cell (oncocytic) type	Thy3f	Suspicious of follicular neoplasm	TIR3b	High-risk indeterminate lesion	15%-30%	Consider additional tests and/or diagnostic hemithyroidectomy
V	Suspicious of malignancy	Thy4	Suspicious of malignancy	TIR4	Suspicious of malignancy	60%-75%	Thyroid surgery recommended. Consider preoperative additional (molecular) testing to determine extent of surgery
VI	Malignant	Thy5	Malignant	TIR5	Malignant	97%-99%	Thyroid surgery recommended

Table 2. Overview of test performance and utility of main additional diagnostics in indeterminate thyroid nodules

			Sensitivity	Specificity	Main advantages	Main limitations	Cost-effectiveness
Molecular Biomarkers							
	Gene Mutation Analysis and Gene Expression						
		BRAF	0%-83% (34,59,61,65)	99%-100% (21,29-31,35-38,41-76)	Perfect specificity at low cost	Strong geographical variation in occurrence, clinical utility likely limited to gene mutation panels in countries other than South Korea	Presumed, though unpublished; €7.50 to $123 per test (53,63,72,88)
		RAS	0%-77% (39,60)	75%-100% (39,60,90)	High prevalence, frequently detected	Often found in follicular adenomas (false-positive); clinical utility limited to gene mutation panels	Unpublished
		RET/PTC	0%-29% (31,49)	73%-100% (29,31)	Specific for PTC	Low prevalence; clinical utility limited to gene mutation panels	Unpublished
		PAX8/PPARγ	0%-29% (55,59,109)	96%-100% (37,59,70)	No significant advantages	Low prevalence; utility limited to gene mutation panels	Unpublished
		7-gene mutation panel	18%-69% (22,61)	86%-99% (37,61)	Comparatively inexpensive mutation panel	Specificity often insufficient for surgical decision-making	USA: likely (120); Europe: unlikely (14); $425 to $1,700 per test (120,121)
		NGS	71%-91% (21,71)	89%-93% (71,80)	Highly accurate; rapidly advancing technology	Limited availability outside the USA; limited clinical validation studies	Unpublished; €230 to $3,200 per test (123,124)
		Afirma® GEC	83%-100% (130,131,134, 144)	10%-52% (126,130)	High rule-out capacity	Limited availability outside the USA; limited high-quality clinical validation studies	Unlikely (14,121,137-140); $3,500 ($1,750 to $7,000) per test (121,131,137)
	MicroRNA		57%-100% (22,159,161)	58%-100% (149,161)	Stable expression irrespective of preservation medium (148,161)	Limited clinical validation, research ongoing	Unpublished
	Immunocytochemistry						
		Galectin-3	0%-92% (44-174)	68%-100% (44,175-177)	Global availability; inexpensive.	Limited current application in cytology; no methodological consensus; limited validation studies for combinations of immunostains.	Unpublished. Up to €20 per test.
		HBME-1	61%-100% (174,176,180, 181)	75%-96% (174,176,180,181)			
		CK-19	76%-88% (174,180,182)	80%-100% (174,180,182)			

Table 2 (continued). Overview of test performance and utility of main additional diagnostics in indeterminate thyroid nodules

Conventional imaging						
	Ultrasound	Dependent on (combination of) feature(s)	Dependent on (combination of) feature(s)	Global availability, low cost	Operator dependency; limited prospective clinical validation; diagnostic accuracy of individual US features insufficient for surgical decision-making	Presumed, though unpublished.
	Elastosonography	47% to 97% (19,222)	6% to 100% (213,221,249)	Global availability, low cost, easily performed during standard US work-up	Operator dependency; limited clinical utility studies; alternative elasticity cut-off possibly more useful	Presumed, though unpublished.
	Computed Tomography	Unavailable	Unavailable	Unavailable	Not investigated in indeterminate thyroid nodules	n.a.
Functional and Molecular Imaging						
	^{99m}Tc-MIBI scintigraphy	56%-79% (19,258)	52%-96% (19,70)	More widely available and lower cost than PET	Limited test performance; limited clinical validation studies; exposure to limited dose of ionizing radiation	Unclear. USA $669-$1,156, Europe: €119-€500 per scan (139,262,263)
	FDG-PET	77%-100% (266,268-270,273)	33%-64% (269,272)	High rule-out capacity; increasing global availability	Exposure to limited dose of ionizing radiation.	USA: unpublished. Europe: likely (14).
	DW-MRI	Unpublished	Unpublished	No ionizing radiation	Limited evidence; no methodological consensus; research ongoing	unpublished

BRAF: BRAF point mutation analysis. GEC: Gene Expression Classifier. n.a.: not applicable. NGS: Next Generation Sequencing. RAS: *RAS* point mutation analysis.

Essential points

- Indeterminate thyroid cytology (Bethesda category III and IV) corresponds to follicular-patterned benign and malignant lesions, which are difficult to differentiate on cytology alone.
- Approximately 25% of indeterminate thyroid nodules harbor malignancy.
- The value of additional diagnostics is best defined by endpoints such as desired minimal rates of accurately prevented unbeneficial surgeries (rule-out capacity) or accurately diagnosed carcinomas (rule-in capacity).
- None of the diagnostic techniques currently available has near-perfect sensitivity, near-perfect specificity and cost-effectiveness.
- A multimodal stepwise approach using a sensitive rule-out and specific rule-in test might offer the most conclusive diagnosis for indeterminate thyroid nodules.
- The decision favoring or opposing a certain diagnostic technique strongly depends on population-dependent variations in cytology (e.g. Hürthle cell cytology), tumor genetics and prevalence of malignancy, and on the costs and feasibility of the desired diagnostic in the local patient population.