**European
Thyroid Journal**

European Thyroid Association Guidelines regarding Thyroid Nodule Molecular Fine-Needle Aspiration Cytology Diagnostics

Ralf Paschke[a]   Silvia Cantara[b]   Anna Crescenzi[c]   Barbara Jarzab[d]
Thomas J. Musholt[e]   Manuel Sobrinho Simoes[f]

[a]Division of Endocrinology and Metabolism, Departments of Medicine, Pathology, Oncology and Arnie Charbonneau Cancer Institute, Calgary, AB, Canada; [b]Department of Medical, Surgical, and Neurological Sciences, Endocrinology Section, University of Siena, Siena, Italy; [c]Pathology Unit, University Hospital Campus Bio-Medico, Rome, Italy; [d]Department of Nuclear Medicine and Endocrine Oncology, Maria Sklodowska-Curie Memorial Cancer Center and Institute of Oncology, Gliwice Branch, Gliwice, Poland; [e]Endocrine Surgery Section, Department of General, Visceral, and Transplantation Surgery, University Medicine of the Johannes Gutenberg-University Mainz, Mainz, Germany; [f]Institute of Molecular Pathology and Immunology, University of Porto, Porto, Portugal

Abstract
Molecular fine-needle aspiration (FNA) cytology diagnostics has the potential to address the inherent limitation of FNA cytology which is an indeterminate (atypia of undetermined significance/follicular lesion of undetermined significance follicular neoplasm) cytology. Because of the emerging role of molecular FNA cytology diagnostics, the European Thyroid Association convened a panel of international experts to review methodological aspects, indications, results, and limitations of molecular FNA cytology diagnostics. The panel reviewed the evidence for the diagnostic value of mutation panel assessment (including at least *BRAF*, *NRAS*, *HRAS*, *KRAS*, *PAX8/PPARG*, *RET/PTC*) of targeted next generation sequencing and of a microarray gene expression classifier (GEC) test in the diagnostic assessment of an indeterminate cytology thyroid nodule. Moreover, possible surgical consequences of molecular FNA diagnostic results of thyroid nodules and the evidence that analysis of a molecular FNA diagnostic panel of somatic mutations or a microarray GEC test can alter the follow-up were reviewed. Molecular tests may help clinicians to drive patient care and the surgical decision if the analysis is performed in specialized laboratories. These molecular tests require standardization of performance characteristics and appropriate calibration as well as analytic validation before clinical interpretation.

© 2017 European Thyroid Association
Published by S. Karger AG, Basel

Introduction/Background

Fine-needle aspiration (FNA) has high sensitivity and specificity in distinguishing benign from malignant thyroid lesions [1]. However, in 2–16% of cases cytology is not diagnostic [2], i.e., the material is insufficient for diagnosis and an FNA repetition is requested. In 5–20% of cases, moreover, it is not possible to discriminate between benign and malignant nodules because of an indetermi-

Ralf Paschke, MD, PhD
Division of Endocrinology, Cumming School of Medicine
University of Calgary, HMRB, Room 382B, 3330 Hospital Dr NW
Calgary, AB T2N 4N1 (Canada)
E-Mail ralf.paschke@ucalgary.ca

Table 1. Comparison of the diagnostic categories of the classification systems

Bethesda System (USA)		Royal College of Pathologist (UK)	SIAPEC-IAP (Italy)
I	Non-diagnostic: cystic fluid only	*Thy1*: Non-diagnostic for cytological diagnosis *Thy1c*: Non-diagnostic for cytological diagnosis: cystic lesion	*TIR1*: Non-diagnostic *TIR1C*: Non-diagnostic-cystic
II	Benign	*Thy2*: Non-neoplastic *Thy2c*: Non-neoplastic cystic lesion	*TIR2*: Non-malignant
III	Atypia of undetermined significance or follicular lesion of undetermined significance (AUS/FLUS)	*Thy3a*: Neoplasm possible atypia/non-diagnostic *Thy3b*: Neoplasm possibly suggesting follicular neoplasm	*TIR3A*: Low-risk indeterminate lesion
IV	Follicular neoplasm or suspicious for follicular neoplasm (FN/SFN)		*TIR3B*: High-risk indeterminate lesion
V	Suspicious of malignancy	*Thy4*: Suspicious of malignancy	*TIR4*: Suspicious of malignancy
VI	Malignant	*Thy5*: Malignant	*TIR5*: Malignant

nate cytology, including follicular proliferation and atypia of undetermined significance (AUS) [3]. As a minority (about 20%) [4] of indeterminate thyroid nodules has been proved to be malignant at final histology, a high proportion of these patients undergo diagnostic rather than curative thyroid surgery.

Reporting of FNA cytology has improved in the past 10 years with the introduction of classification schemes in order to standardize terminology, to facilitate communication among cytopathologists, endocrinologists and surgeons, and to provide the malignancy risk for specific diagnostic categories. The Bethesda System, a 6-diagnostic-category system, is at present the most used reporting system for thyroid FNA cytology. The categories are as follows: (I) non-diagnostic/unsatisfactory; (II) benign; (III) atypia of undetermined significance/follicular lesion of undetermined significance (AUS/FLUS); (IV) follicular neoplasm/suspicious for follicular neoplasm (FN/SFN), a category that also encompasses the diagnosis of Hürthle cell neoplasm/suspicious for Hürthle cell neoplasm; (V) suspicious for malignancy, and (VI) malignant [5]. In a meta-analysis of 25,445 thyroid FNA samples reported from 8 studies using The Bethesda System for Reporting Thyroid Cytology, 9.6% of all samples were diagnosed as AUS/FLUS, 10.1% were diagnosed as FN/SFN, and 2.7% were diagnosed as suspicious for malignancy, with an average cancer risk of 15.9, 26.1, and 75.2%, respectively [6]. Excluding suspicious lesions, which have such a high malignancy risk level that a surgical strategy is called for, both the AUS/FLUS and FN/SFN, also named indeterminate cytology, have a non-negligible cancer risk. Therefore, diagnostic surgery appears to be frequently necessary even if the malignancy risk is not so high that it indicates definitive cancer with certainty.

Other classification systems for thyroid FNA used in Europe are the Italian consensus by the Italian Society for Anatomic Pathology and Cytology jointly with the Italian Division of the International Academy of Pathology (SIAPEC-IAP) [7] and the Royal College of Pathologist [8] classification. Table 1 shows the comparison of the diagnostic categories of these classification systems. The main difference resides in the Bethesda class IV/RCP Thy3f and the Italian TIR3B. The latter category includes both follicular and non-follicular patterned lesions instead of a pure follicular patterned group.

The need to improve the diagnostic performance of FNA cytology in the setting of indeterminate results has led to the search for diagnostic biomarkers, and several recently introduced molecular tests [9], such as gene expression classifier (GEC) [10], gene mutational panel [11], targeted next-generation sequencing (tNGS) [12], and the combination of mutation detection and miRNA expression [13], have been reported as promising in ruling in or out a substantial risk of cancer. Of note, the major limit to the use of molecular testing in thyroid FNA may stem from a clinicohistopathological finding: it has recently been suggested that the encapsulated, non-invasive follicular variant of papillary thyroid carcinoma (PTC) should be reclassified as a benign tumor. If this reclassification is unanimously accepted, it will lead to a dramatic reduction of the risk of malignancy for AUS/FLUS, FN/SFN, or indeterminate FNA cytologies [14] and consequently also a dramatic reduction of the risk of malignancy for associated mutations like RAS.

Paschke/Cantara/Crescenzi/Jarzab/
Musholt/Sobrinho Simoes

A specific setting among follicular lesions is represented by nodules with oncocytic (Hürthle cell) features. The clinical significance of oncocytic changes in thyroid tumors is controversial, and a more aggressive behavior of Hurthle cell carcinomas when compared with the respective usual variants is still being debated [15]. Cytomorphologic criteria alone, moreover, are not helpful in distinguishing between benign and malignant Hürthle cell neoplasms. Finally, a potential setting for molecular tests with diagnostic purpose could be the non-diagnostic material. The rate of non-diagnostic results that should not exceed 10% of cases demonstrated a percentage of malignancy estimated between 5 and 15% and represents an indirect measure of the expertise and quality of FNA practice [4]. A molecular test resolving non-diagnostic FNA cytologies has been regarded as an important goal for future research.

Cytological specimens represent a useful and valuable source of cellular material for the evaluation of various molecular diagnostic tests that contribute to optimal patient care. This source suffers, however, from a large variability of sampling and processing. Cytopathology laboratories utilize numerous different methods to obtain and process cytological specimens. Therefore, the development and careful validation of molecular assays for these different cytological samples, such as direct smears, cell blocks, and liquid-based preparations, are important.

The thyroid FNA procedure should be performed under ultrasound guidance and thyroid cells are aspirated using a 23-, 25-, or 27-gauge needle for cytological evaluation. Direct cytology smears are both air dried or fixed with alcoholic solution or spray preparation. It is important to note that this preparatory technique is performed without formalin fixation. Formalin fixation leads to the crosslinking of nucleic acids and proteins, fragmentation of isolated nucleic acid material, and the possibility of sequence alterations. This point needs to be considered when other sources of samples such as cell block or cytology preparation with preserving liquid are used and specific protocols have to be applied in order to obtain high-quality nucleic acid isolation [16] Different types of specimen obtained from thyroid FNA procedure may be used for molecular analysis: (a) cells collected from FNA during the procedure [11], (b) cells retrieved from FNA-stained smears [1], (c) cells retrieved from air-dried slides [17], and (d) paraffin sections from cell blocks. All these samples have been shown to be suitable for point mutation analysis as well as for detecting chromosome rearrangements [17]. The different assays should be optimized for small size specimens and validated for the different cytology specimens. Furthermore, standardization of the performance characteristics and appropriate calibration as well as of analytic interpretation before clinical interpretation is necessary [18].

Nucleic acid recovering from direct smears offers relevant advantages: first, it permits to check the cells before nucleic acid isolation, allowing the analysis of exactly those cells that are cytologically indeterminate and the enrichment of the pathological population [16, 17]. This is not possible if a dedicated pass of FNA is washed in a tube containing nucleic acid preservative solution. Tumor enrichment mitigates the error based on the possibility that tumors may be heterogeneous with respect to a specific molecular alteration. Second, molecular tests from direct smears are performed on the already available diagnostic material without additional reaspiration of the nodule, patient discomfort, and the risk of inadequate sampling due to recent intranodular hemorrhage [16, 17].

Molecular analysis for thyroid FNA should be restricted to centers that have a low percentage of inadequate samples, an appropriate, but not too high, percentage of AUS/FLUS diagnoses, and a good track record for cytological-surgical concordance and malignancy rates. Cytological and molecular interpretation should be integrated. All molecular testing intended for clinical use must be performed only in certified clinical laboratories after appropriate validation of at least analytical sensitivity and specificity of mutation detection. Analytical sensitivity should be correlated with clinical sensitivity, so that the sensitivity is not pushed to extremely high levels, as this would lead to false-positive mutation detection.

Methods

The European Thyroid Association initiated this task force with the overall aim to review the relevant medical evidence and provide clinical recommendations on the use of molecular diagnostics in the evaluation of indeterminate cytology thyroid nodules (including the respective Bethesda categories of AUS/FLUS or FN, or SFN, or equivalent). Task force members were selected by the Task Force Chair, Ralf Paschke, based on their expertise in relevant fields, including endocrinology, molecular oncology, epidemiology, and endocrine surgery. This task force received no funding from the European Thyroid Association, industry, or any other organizations, and it is editorially independent from the European Thyroid Association. The competing interests of all task force members were collected by the Task Force Chair, prior to writing the manuscript and updated prior to submission for publication of this document (reported at the end of this manuscript). The clinical questions for these guidelines were developed by the Task Force Chair and modified by the Task Force members, such that there was final agreement by the group on the relevant topics for review.

The target users of these guidelines are endocrinologists, endocrine surgeons, and other healthcare providers caring for patients with thyroid nodules. The target population to which these guidelines apply includes adults (aged ≥18 years) with an indeterminate cytology thyroid nodule ≥1 cm in diameter on ultrasound imaging, who are otherwise potential candidates for thyroid surgery (i.e., no major contraindications to thyroid surgery), and who have no prior thyroid malignancy, no prior significant head and neck radiation exposure, and no known concurrent diagnosis of thyroid cancer (i.e., within another thyroid nodule, a cervical lymph node, or a distant metastatic focus). We did not review economic evaluations of various molecular diagnostic tests. We restricted eligible molecular diagnostic tests to those carried out on thyroid FNA biopsy (FNAB) specimens for the purpose of this review. The molecular diagnostic tests of interest for this review included: panels examining at least *BRAF*, *NRAS*, *HRAS*, *KRAS*, *PAX8/PPARG*, *RET/PTC*, with or without others, or GEC tests. Molecular tests evaluating for a single mutation were not deemed sufficiently contemporary to be relevant for this review.

Literature Review and Appraisal and Resulting Formulation of Recommendations

Each task force member was assigned one or more clinical topics and questions. There was a primary and secondary reviewer for each section; the role of the secondary reviewer was to verify content accuracy and to provide input on potential additional relevant references (if present) as well as on editing the section. Each primary reviewer was instructed to carry out an electronic citation search using one or more electronic databases (including one of either Medline or PubMed plus additional references). The Task Force Chair (R.P.) provided all members with relevant search terms to use in the execution of the electronic searches, as resources were not available for a librarian to conduct this task. The citations retrieved from electronic searches, and relevant full-text papers, were reviewed by each respective primary reviewer, with a review of the entire reference list for completeness by the second reviewer. Only English language papers were included in the review. Data were summarized in narrative form, with summary tables, as appropriate, by each respective primary reviewer, and further edited by the secondary reviewer, and thereafter by the Task Force Chair and the rest of the Task Force. A clinical validity score was evaluated according to EGAPP as good, fair, or marginal [19].

What Is the Value of Mutation Panel Assessment (Including as a Minimum *BRAF, NRAS, HRAS, KRAS, PAX8/PPARG, RET/PTC*) in the Diagnostic Assessment of an Indeterminate Cytology Thyroid Nodule?

The most frequent genetic alterations detected in PTC and follicular thyroid carcinoma (FTC) are *BRAF* and *RAS* mutations and *RET/PTC* and *PAX8/PPARG* rearrangements [20, 21]. *TERT* promoter mutations and *TP53* mutations are less frequent in well-differentiated (PTC and FTC) carcinomas than in less-differentiated carcinomas, but their detection in thyroid FNA should also be considered in the appropriate clinicocytopathological setting [22, 23].

The search for mutations of *BRAF*, *TERT* promoter, and *TP53* is recommended in cases classified in Bethesda class V as well as in the subgroups of Bethesda class III in which there are focal features suggestive of PTC. Besides helping to clarify whether or not one is dealing with a malignant tumor, the detection of *BRAF* (and, probably, also of *TERT* promoter and *TP53* mutations) may guide the type of surgery, leading to total thyroidectomy if the clinicopathological setting is appropriate [21, 24, 25].

For non-diagnostic samples (Bethesda class I) in most cases, the problem stems either from a lack of expertise in the center that performed the defective procedure or from particular characteristics of the nodule [26]. In any of the aforementioned cases, the best approach is to perform a second FNA at another institution with FNA experience. Molecular testing may help in selected cases [27].

In Bethesda class II cytological specimens (benign lesion) it is not helpful (and may even be confusing) to look for any mutations and/or rearrangements in FNA material except for rare cases when there is a strong clinical suspicion for cancer but a discordant FNA cytology result [28].

The reported sensitivities, specificities, negative (NPVs), and positive predictive values (PPVs) for studies investigating panels detecting *BRAF*, *NRAS*, *HRAS*, *KRAS*, *PAX8/PPARG*, and *RET/PTC* in AUS/FLUS, FN/SFN, or indeterminate FNA cytologies range from 18 to 100%, from 82 to 100%, from 56 to 100%, and from 19 to 100%, respectively (Table 2). Varying risks of malignancies and different rates of FTCs or PTCs for the investigated samples or knowledge of the pathologist about the molecular results are important variables that profoundly influence these statistics. Moreover, especially the risk of malignancy for *RAS*-positive FNA results varies dramatically between these studies, i.e. from 19 to 85% [11, 16, 17, 27]. Therefore, the impact of the detection of *NRAS*, *HRAS*, *KRAS*, *PAX8/PPARG*, and *RET/PTC* in AUS/FLUS, FN/SFN, or indeterminate FNA cytologies has to be determined in each specific setting.

The detection of this panel of mutations has been performed in fresh FNA material, sometimes obtained by extra FNA passes and in air-dried, stained smears or liquid cytology material. Detecting these mutations in material obtained from air-dried, stained smears or liquid cytology material has obvious advantages such as no need for obtaining extra FNA material, storing extra FNA material, and most importantly performing molecular diagnostics on exactly the same sample as the cytologist used for his morphologic diagnosis [16, 17].

Table 2. 7-mutation panel reports (modified from Nishino, 2016 [56])

First author, year	Cytological category	Cases, n	Prevalence of malignancy based on cytology, %[d]	Test performance, %[d]			
				Sens.	Spec.	NPV	PPV
Nikiforov, 2009 [57][1]	AUS/FLUS	21	14	100	100	100	100
	FN/SFN	23	52	75	100	79	100
Cantara, 2010 [40][a, 1]	Indeterminate	41	17	86	97	97	86
Nikiforov, 2011 [11][1]	AUS/FLUS	247	14	63	99	94	88
	FN/SFN	214	27	57	97	86	87
Beaudenon-Huibregtse, 2014 [41][2]	AUS/FLUS	22	50	36	82	56	67
	FN/SFN	19	32	67	92	86	80
Eszlinger, 2014 [17][b, 3]	Indeterminate	141	16	18	86	85	19
Eszlinger, 2015 [16][3]	Thy3 (FN/SFN)	163	28	49	92	82	71
Labourier, 2015 [13][c, 2]	AUS/FLUS and FN/SFN	109	32	69	86	85	71

Sens., sensitivity; Spec., specificity. [1] Prospective single-center study. [2] Prospective multiple-center study. [3] Retrospective single-center study. [a] The study by Cantara et al. also tested for gene fusions involving TRK. The authors used a 4-category cytology reporting systgem ("inadequate," "benign," "indeterminate," and "suspicious for thyroid cancer"). [b] The malignant cases in the data set in the 2014 study by Eszlinger et al. were enriched for follicular thyroid carcinomas relative to papillary carcinomas. The authors indicated that cytologically diagnostic specimens were reported in 3 categories ("benign," "indeterminate," and "malignant"). [c] The study by Labourier et al. compared performance of the 7-gene panel with and without a microRNA expression classifier. Only the data for the 7-gene panel were extracted here. Some of the data may overlap with those of Beaudenon-Huibregtse et al. [d] Prevalence of malignancy, test results, and test performance calculations based on resected nodules only. Only data for nodules with indeterminate cytology (AUS/FLUS, FN/SFN, or comparable categories), satisfactory molecular testing results using a 7-gene panel (*BRAF*, *HRAS*, *KRAS*, *NRAS*, *RET/PTC1*, *RET/PTC3*, and *PAX8/PPARG*), and follow-up reference histology were extracted from the above studies for this table. Cases with non-diagnostic molecular testing results or those lacking follow-up surgical resection were excluded.

Recommendations

Cytopathology practice results, patient characteristics, and prevalence of malignancy within the population being tested all have an impact on the NPVs and PPVs for molecular testing. For *cytologically indeterminate nodules* consider, if available, the detection of *BRAF* and *RET/PTC*, and, possibly, *PAX8/PPARG* and *RAS* mutations. The significance of detecting *RAS* mutations needs to be clarified.

The search for molecular markers in Bethesda class II cases is not recommended. (*This recommendation is not shared by one author.*)

The relevance for selected mutations such as *BRAF*, *TERT* promoter, and *TP53* in Bethesda class V cases in which there is suspicion for PTC needs to be further investigated.

Should a Microarray GEC Test Be Routinely Used on an Indeterminate (AUS/FLUS or FN/SFN) Cytology Thyroid Nodule to Exclude a Thyroid Malignancy in Lieu of Diagnostic Thyroid Surgery or Continued Close Follow-Up?

The evaluation of mRNA expression in indeterminate cytology thyroid nodules using microarray technology is referred to as GEC testing. The core data of the studies are summarized in Table 3. The largest study describing analytic validity (diagnostic performance) of the GEC test was reported by Alexander et al. [10], who conducted a prospective, multicenter study in the USA; this study was funded by industry (Veracyte). The authors performed central, blinded, gold-standard, histologic confirmation of all specimens included in the final analysis [10]. In this study, 577 cytologically indeterminate aspirates were obtained, 413 of which had histopathological confirmation (which was required for in-

Table 3. Evidence of molecular FNA diagnostics can reduce diagnostic lobectomies in individuals with indeterminate thyroid cytology who do not have thyroid cancer and clinical utility of GEC

First author, year	Indeterminate nodules, total N	GEC result		Nodules operated, total N	Benign nodules (correctly classified by GEC), n (n)[1]	Malignant nodules (correctly classified by GEC), n (n)[1]	Sens., % (CI)	Spec., % (CI)	Clinical impact, n/N (%)		Data on follow-up of benign tumors	EGAPP clinical validity
		B	S						avoid surgery	recommend surgery		
Alexander, 2012 [10]	265	100	165	265	180 (93)	85 (78)	92 (84–97)	52 (44–59)	No recommendation	No recommendation	No	Fair
McIver, 2014 [29]	60	16	44	36	30 (3)	6 (5)	83	10	12/16 (75)[3]	40/44[4]	11 months in 16 patients	Fair
Alexander, 2014 [38]	339[2]	174	148	132	78 (10)	54 (53)	98	13	170/174 (98)[5]	141/148 (95)	8 months in 41% of 174 patients	Fair
Duick, 2012 [39]	395	368	27	28	NG	NG	NG	NG	340/368 (92)	–	No	Marginal
Lastra, 2014 [31]	132	70	62	50	2	22	NG	NG	68/70 (97)	62/132	Only 2 patients	Fair
Harrel, 2014 [30]	58	20	36	35	3 of 5	21 of 30	NG	NG	19/58 (32)	36/58	Not given	Marginal
Nikiforov, 2014 [12]	143			143	96 (97/101)	83 (35/42)	90	93	Not given	Not given	Not given	Fair
Marti, 2015 [32]	94[A]	24	70	44	2 (2)	24 (42)	91	51	Not given	Not given	Not given	Fair
	71[B]	37	34	26	5 (5)	3 (21)						

EGAPP clinical validity score: good, fair, or marginal. B, benign; S, suspicious; Sens., sensitivity; Spec., specificity; NG, not given. [1] With the exception of Nikiforov [12] where the values represent % (n/n). [2] For 17 tumors the Afirma test was non-diagnostic. [3] 4 patients underwent surgery (2 immediately, 2 in the follow-up period), 1 patient with 3.2-cm FTC. [4] 27/32 patients with a suspicious GEC result had benign histopathological results following surgery. [5] Surgery was recommended in 4/174 patients but performed in 11/174 patients with benign GEC classification for personal reasons or compressive symptoms; follow-up only in 41% of the 174 patients. [A] Results of the Memorial Sloan Kettering Cancer Center (MSK; n = 94). [B] Results of the Mount Sinai Beth Israel Hospital (MSBI; n = 71).

clusion in the study), and of these 389 were allocated to GEC testing (excluding those used in training or analytic verification) [10]. The cytological diagnoses included in the indeterminate category for this study included AUS/FLUS, FN/SFN, Hürthle cell neoplasm, or suspicious for malignancy, with the AUS/FLUS category comprising almost half of the specimens in the final analysis (49%, 129/265). Of the 389 samples allocated to GEC testing, 379 were processed for a GEC result, 328 had a valid GEC result, and 312 had a valid GEC result with gold-standard histopathological confirmation [10]. However, ultimately 265 were included in the final analysis [10]. Of these 265 non-consecutively obtained indeterminate cytology thyroid nodules, 85 (32%) were confirmed to be histopathologically malignant. In the 265 indeterminate cytology nodules, the sensitivity of the GEC test was 92% (95% confidence interval [CI], 84–97%, 78/85) and the specificity was 52% (95% CI, 44–59%, 93/180) [10]. It is important to note that all the patients included in the final analysis underwent thyroid surgery (as a pre-established inclusion criterion); therefore, long-term outcomes of patients avoiding thyroid surgery on the basis of a negative GEC test cannot be extrapolated from this study.

The diagnostic performance of the GEC test was subsequently reported in 4 respective studies from the USA [29–32]. In each of these studies, treating physicians were aware of the GEC result at the time of surgical decision-making and not all patients who underwent a GEC test had gold-standard histopathological confirmation of the diagnosis; furthermore, there was no reported blinding of the histopathologists who ascertained the gold-standard diagnosis (when performed) [29–32]. McIver et al. [29] reported in a prospective register study that patients at the Mayo Clinic (Rochester and Jacksonville sites) were offered the GEC test if the FNA was performed on site and the cytology was reported as SFN, suspicious for Hürthle cell neoplasm, or AUS/FLUS; approximately 83% (30/36) of the specimens included in the final analysis were SFN or suspicious for Hürthle cell neoplasm. In this study, criteria for offering GEC testing to patients with indeterminate cytology thyroid nodules included willingness of the patient to defer surgery, lack of compressive symptoms, and absence of higher-risk features (including head and neck radiation history, family history of thyroid cancer, prior history of thyroid cancer, or worrisome imaging characteristics) [29]. A GEC "suspicious" result prompted a recommendation for surgery in this study, whereas a GEC "benign" result prompted a recommendation for repeat ultrasound in 3–6 months and thereafter yearly for 5 years [29]. McIver et al. [29] reported that of 105 indeterminate cytology FNAs obtained during the study, 72 were subjected to GEC testing, 60 had GEC results, 12 had insufficient samples for

a GEC result, and of those 36 underwent surgery with histopathological confirmation of the diagnosis. Of the 36 samples with GEC results and histopathological confirmation of the diagnosis, the prevalence of malignancy was 17% (6/36). The authors of this study estimated that the sensitivity of the GEC test was 83% (5/6) and the specificity was 10% (3/30) in histopathologically confirmed specimens. The median follow-up for GEC "benign" nodules in this study was only 9.5 months at the time of publication [29].

Harrell and Bimston [30] also reported the results of a single-center retrospective study from Florida, in which all histopathological specimens were read centrally by an experienced pathologist, who was not blinded to the GEC result. These authors reported on 58 indeterminate cytology specimens, including the categories of either AUS/FLUS or FN (the latter being follicular or Hürthle cell predominant). The prevalence of malignancy in those with histopathological confirmation in this study was 51% (18/35). In this study, the GEC test sensitivity was 94% and specificity was 24%. Harrell et al. [30] indicated that the patients with a GEC "benign" result were advised to return for a yearly follow-up, including neck imaging, but no follow-up data were reported.

In another single-center retrospective, unblinded study, Lastra et al. [31] reported on the University of Pennsylvania experience, in which 132 cases of indeterminate cytology nodules were subject to GEC testing, including the categories of AUS/FLUS, FN, or FN with oncocytic features, with more than half of the cases being in the AUS/FLUS category (52%, 68/132). The prevalence of malignancy in the 50 cases in this study that had histopathological confirmation of diagnosis was 44% (22/50). In calculations using data from this study [31], the GEC test sensitivity was 100% (22/22) and specificity was 7% (2/28). No follow-up data were reported on patients who did not undergo surgery in this study.

Most recently, Marti et al. [32] reported a retrospective analysis of all indeterminate thyroid nodules evaluated with the GEC at 2 New York-based institutions, including Memorial Sloan Kettering Cancer Center (94 nodules) and Mount Sinai Beth Israel Hospital (71 nodules). In this study, the GEC test was reported as "benign" in 38.3% of cases at Memorial Sloan Kettering Cancer Center and 52.1% of cases at Mount Sinai Beth Israel Hospital. The authors estimated that the PPVs of the GEC "suspicious" results were 57.1 and 14.3% at respective institutions, whereas the NPVs of a GEC "benign" result were estimated to be 86–92% at Memorial Sloan Kettering Cancer Center and 95–98% at Mount Sinai Beth Israel Hospital.

The authors of this study stressed the importance of knowing the prevalence of malignancy in indeterminate thyroid nodules at respective institutions, as such information was considered critical in interpreting the GEC results.

In summary, based on the above studies that reported on sensitivity and specificity, GEC test sensitivity was reported to range from 83 to 100% and specificity from 7 to 52%, whereas the prevalence of malignancy in histopathologically confirmed study populations ranged from 17 to 51% [10, 29–31].

Long-term outcomes of patients undergoing GEC testing, including those who do not undergo thyroid surgery based on a negative GEC result, are important to be considered for clinical recommendations. Unfortunately, currently, such long-term outcome data are lacking, with available studies reporting median follow-up periods of less than 1 year in duration [10, 29].

Recommendation
In an individual with an indeterminate cytology thyroid nodule, a GEC test cannot be recommended as clinical routine at this time to exclude malignancy in lieu of diagnostic thyroid surgery or close surveillance. (*This recommendation is not shared by one author.*)

This recommendation recognizes the lack of long-term outcome data on patients with indeterminate cytology thyroid nodules who do not undergo thyroid surgery on the basis of a negative GEC test and variable results for sensitivity, specificity, NPV, and PPV in independent studies.

What Is the Value of tNGS in the Diagnostic Assessment of an Indeterminate Cytology Thyroid Nodule?

tNGS sequencing is a promising method to examine multiple genes simultaneously with high sensitivity potentially achieving not only high PPVs but also high NPVs [33] and with low input of starting material. Preliminary results for thyroid FNAs have thus far only been published by 2 groups [12, 33, 34], and clinical validation results are not yet available. Data for lung and other tumors show that next generation sequencing is robust like Sanger sequencing in routine diagnostics and, in addition, is able to reveal mutations at low percentage and to screen the mutational status of different critical samples offering innovative diagnostic opportunities [35–37]. Furthermore, methodological problems like result interpretation (e.g., for unknown mutations), definition of

cutoffs for mutation calling, and bioinformatics analysis need to be solved, and common standard operation procedures (SOPs) need to be defined.

tNGS analyzing larger mutation panels was used for the molecular diagnostics of limited numbers of FN/SFN, AUS/FLUS, and indeterminate FNA samples reporting sensitivities from 71 to 91%, specificities from 89 to 93%, NPVs from 92 to 97%, and PPVs from 63 to 83%. If reproduced in larger sample sets this tNGS approach has the potential to become a simultaneous rule-in and rule-out test if NPVs >95% can be confirmed.

Recommendation

The substitution of classic mutation detection methods for thyroid FNAs with the tNGS approach is most promising, although it has not been confirmed in other laboratories. Larger tNGS mutation panels have the potential to become a simultaneous rule-in and rule-out test if NPVs >95% can be confirmed.

Possible Surgical Consequences of Molecular FNA Diagnostic Results of Thyroid Nodules

The preoperative knowledge of a verified thyroid carcinoma will have an impact on the surgical strategy and therefore will considerably increase the success of the oncologic treatment.

The clinical setting may lead to two different basic diagnostic approaches in the decision-making for thyroid surgery due to suspicious nodes:
1. Rule out malignancy in order to avoid surgery
2. Verify malignancy preoperatively and determine the type of thyroid cancer to optimize surgical strategy and treatment success, while avoiding secondary interventions

In order to comply with the rule-in and rule-out strategies the necessary NPV for rule-out tests should be >95% according to the National Comprehensive Cancer Network guidelines [38], while the ideal PPV should be >95% for rule-in strategies leading to more radical resections (total thyroidectomy). However, apart from the indication for primary radical surgery a lower risk of malignancy prompts the recommendation for diagnostic resection. Current guidelines recommend total or near total thyroidectomy or lobectomy in case of FNA results suspicious for malignancy (Bethesda class V), which is associated with a 60–75% risk of malignancy. Similarly lobectomy is recommended in case of Bethesda class IV with a risk of malignancy ranging from 15 to 30% [5].

What Is the Evidence that Molecular FNA Diagnostics Can Reduce Diagnostic Lobectomies in Individuals with Indeterminate Thyroid Cytology Who Do Not Have Thyroid Cancer?

In a recent meta-analysis [6], 22% of 25,448 FNAB specimens were classified in one of the indeterminate categories ("follicular lesion of undetermined significance (FLUS)," "follicular neoplasm (FN)," or "suspicious for malignancy") according to the Bethesda Cytologic Classification System. 6,362 (25%) underwent surgery confirming malignancy among indeterminate FNAB specimens in 31%. The cytology categories AUS/FLUS, FN, and suspicious for malignancy corresponded to ascending malignancy rates of 16, 26, and 75%, respectively. If additional diagnostics could rule out malignancy this would potentially prevent 84, 74, and 25% of the interventions assuming that there is no other reason (e.g., compression symptoms, growing nodules, thyrotoxicosis) for surgery.

GEC (based on Affymetrix Human Exon 1.0 GeneChip) [10] aims to rule out malignancy in FNABs that were classified as either "follicular lesion of undetermined significance (FLUS)" or "follicular neoplasm (FN)" according to the Bethesda Cytologic Classification System. Five nonrandomized studies (4 prospective, 1 retrospective) reported on patients with intermediate thyroid cytology but benign classification based on the GEC. Based on the results of the GEC, thyroid surgery was prevented in 32–98% of patients. However, only 2 studies provided limited follow-up information on a subgroup of patients that did not undergo surgery (11 and 8 months). Moreover, the calculation of the NPV of the test was based on the assumption that all thyroid nodules classified as benign and that were not operated on did not include false negative results. Despite promising initial results generated by industry-funded studies in North America, the value of the GEC still has to be demonstrated in the routine clinical setting [10, 29, 39]. Marti et al. [32] evaluated the test performance in 2 different clinical centers. Patient populations of both centers demonstrated significantly different risks for thyroid cancer, which had a profound impact on the test performance as well as NPV and PPV. The authors demonstrated that PPV and NPV depend heavily on the prevalence of malignant disease in the population under study. In the studied centers GEC results had little or no impact on the decision for or against surgery. The GEC would be expected to provide the most useful information in a practice setting with a prevalence of malignancy in intermediate thyroid nodules with a risk of malignancy of 15–21%. However,

in most practice settings the associated risk of malignancy differs from this range.

In summary, the percentage of individuals undergoing thyroid surgery after GEC testing has been reported to range from 38 to 60% with the majority of those preceded by a "suspicious" GEC result [10, 29–31]. In the only study reporting the rate of surgery in a concurrently followed cohort of individuals with indeterminate cytology who did not undergo GEC testing (likely in part due to not meeting pre-established institutional criteria for offering the GEC test), McIver et al. [29] reported that the rate of thyroid surgery was relatively comparable at 55%. However, Duick et al. [28] reported a historical thyroid surgery rate of 74% in patients with indeterminate cytology nodules. Some of the most frequently reported reasons for individuals with indeterminate cytology thyroid nodules and a GEC benign result to undergo surgery included the presence of compressive symptoms, higher risk of malignancy based on history or imaging characteristics, and personal preference. This leads to the question why the GEC was performed in these patients with other indications for surgery. Molecular diagnostics that rule out malignant disease may reduce the number of thyroid resections in the short-term follow-up if the risk of malignancy associated with indeterminate thyroid cytology ranges between 15 and 21% and the test results in an NPV are >95% in a defined patient population.

Mutation panels that screen for the most frequent genetic alterations in differentiated thyroid carcinoma including *BRAF*, *NRAS*, *KRAS*, *RET/PTC*, *NTRK*, and *PAX8/PPARG* are predominantly intended to identify thyroid cancer and are not able to rule out malignancy with an NPV >95%. Therefore, surgery has to be recommended despite the exclusion of these mutations [11, 16, 17, 28, 40]. A tNGS analysis of the above-mentioned genes plus additional genes (*AKT*, *TP53*, *GNAS*, *PTEN*, *TERT*, *CTNNB1*, *PIK3CA*, and *TSHR*) [12] included 143 FNA samples with indeterminate cytology. Among all FN/SFN nodules, the extended preoperative mutation detection was performed with 90% sensitivity (95% CI, 80–99%), 93% specificity (95% CI, 88–98%), a PPV of 83% (95% CI, 72–95%), an NPV of 96% (95% CI, 92–100%), and 92% accuracy (95% CI, 88–97%). While these results almost meet the criteria of an ideal test, 91 FNA samples in the study were analyzed retrospectively and the pathologist was aware of the histologic diagnosis. Therefore, the impact on the decision to recommend surgery and the surgical strategy cannot be estimated.

Recommendation

Due to a lack of long-term follow-up data, the unknown risk of necessary surgery in the further course of the disease and missing information on the patient's symptoms that may lead to surgery the overall evidence is very weak.

What Is the Evidence that Molecular FNA Diagnostics Can Increase the Number of Primary Total Thyroidectomies and Reduce the Number of 2-Stage Surgeries?

The reduction of the number of 2-stage surgeries is dependent on the preoperative diagnosis of thyroid cancer especially in cases with non-malignant or with indeterminate FNA cytology. Molecular markers that are almost exclusively associated with thyroid cancer qualify for the "diagnosis" of malignant disease in the absence of histologic proof. However, studies that report on the clinical impact do not exist.

Several studies investigated the use of mutation screening for predominant genetic alteration in PTC and FTC. Various genes were included, e.g. *BRAF*, *RET/PTC*, *TRK*, *PAX8/PPARG*, and *RAS*. Cantara et al. [40] reported the results of the analysis of 235 FNAs in 174 patients. All non-malignant FNAs (53 non-diagnostic, 41 indeterminate, 87 benign, and 54 suspicious for malignancy) were included. Indeterminate FNA samples revealed mutations in 7/41 cases [40]. Excluding the detection of *RAS* mutations in follicular adenomas the molecular genetic information would have a potential impact in 61 patients for all FNAs but only in 4/41 patients with indeterminate FNA. Moreover, the indication for surgery was determined prior to the inclusion into the study and not dependent on the molecular genetic analysis. However, the authors recommended primary thyroidectomy in case of mutation-positive FNA including *RAS* mutations. It is not reported how many patients would have undergone primary thyroidectomy for bilateral disease irrespective of the molecular genetic result or in how many cases the diagnosis could be established by frozen section.

The study of Beaudenon-Huibregtse et al. [41] subjected patients to molecular genetic analysis of *BRAF*, *KRAS*, *HRAS*, *NRAS* mutations, and *PAX8/PPARG* and *RET/PTC* rearrangements irrespective of the cytology result. They reported the identification of oncogenic mutations in 25/53 with indeterminate cytology, 4/26 with a benign cytology result, and 2/2 with non-diagnostic cy-

tology. Potentially in 31 of 618 patients (546 with benign cytologies; 5%) 2-stage surgery or suboptimal surgery could have been avoided. However, the authors do not provide information on intraoperative findings or on how the surgical strategy was changed in the patients with mutation-positive nodules.

Similarly Moses et al. [42] and Eszlinger et al. [17] demonstrated that in about 11% of all FNAs that were classified as non-malignant (Bethesda classes I–V) the identification of oncogenic mutations could potentially alter the surgical strategy. Again in both studies, no information is provided on how the surgical strategy was changed in patients with mutation-positive nodules.

Nikiforov et al. [11] reported their results including 967 FNAs of 762 patients with AUS/FLUS, FN/SFN, and also 52 suspicious FNAs. Oncogenic mutations were present in 89% of FNAs. The additional mutational analysis of *BRAF* V600E, *NRAS* codon 61, *HRAS* codon 61, and *KRAS* codons 12/13 point mutations and *RET/PTC1*, *RET/PTC3*, and *PAX8/PPARG* rearrangements in the different cytology categories enabled the identification of thyroid malignancies and potentially a preoperative stratification of the surgical therapy between lobectomy and thyroidectomy in 38 of 513 FNAs including Bethesda classes IV and V. Again, information on the exact impact on the surgical strategy was not provided but primary thyroidectomy in case of oncogenic mutations is recommended.

The same group reported on the extended tNGS mutation panel including the additional genes *AKT*, *TP53*, *GNAS*, *PTEN*, *TERT*, *CTNNB1*, *PIK3CA*, and *TSHR* [12]. 143 FNA samples with indeterminate cytology were analyzed. The extended mutation panel achieved a PPV of 83% (95% CI, 72–95%) and an NPV of 96% (95% CI, 92–100%). 91 FNA samples in the study were analyzed retrospectively, and the pathologist was aware of the histologic diagnosis. Therefore, the impact on the decision to recommend surgery and the surgical strategy is unclear and has to be evaluated in a larger prospective trial.

All reviewed studies are based on mutational panels including at least 5 different genes that are associated with thyroid cancer. However, the risk of malignant disease varies for each mutation. While *BRAF*, *RET/PTC*, and *PAX8/PPARG* mutations predict thyroid malignancy with more than 95% probability, the risk of malignancy associated with *RAS* mutations varies between 19 and 87% [11, 16, 17, 43].

In summary, predominantly retrospective, non-randomized data indicate that potentially in 11–47% of patients (Table 2) with indeterminate FNAs, which also include suspicious cytologies in some studies, the surgical strategy may be changed from lobectomy or partial resection to primary thyroidectomy due to additional information on molecular genetic alteration. However, the potential impact is significantly influenced by the following:

– The overall risk of malignancy in the indeterminate FNA categories
– The risk of malignancy associated with a specific genetic alteration: total thyroidectomy may not be recommended for *RAS* mutations due the lower PPV associated with these mutations and possible overtreatment in a number of patients. Therefore, not the global PPV of the mutation panel but the PPV of a certain mutation has to be taken into account when defining the surgical strategy
– High interobserver variation for the histologic diagnosis of adenoma or FTC and the possible reclassification of adenomas to microinvasive FTCs after the detection of *RAS* mutations [16, 44]
– The frequency of total thyroidectomy for benign disease
– The recommendations to perform total thyroidectomy for a specific tumor stage (less than total thyroidectomy [lobectomy] is proposed/recommended for selected thyroid cancers 1–4 cm in size)

Unpublished results of the PETS II study including 22,008 patients revealed that 78% of the patients underwent surgery without preoperative FNA leading to a carcinoma rate of 12% [45]. Similar results were reported by Wienhold et al. [46].

Detailed epidemiologic studies about the frequency of completion thyroidectomy do not exist but some studies report a frequency of completion thyroidectomy in about 3–16% of thyroid carcinomas [47, 48].

Recommendation

Molecular diagnostics identifying mutations that predict thyroid cancer with >95% probability can reduce the number of completion thyroidectomies. Other mutations, e.g. *NRAS*, *HRAS*, or *KRAS* mutations, that are associated with less risk of malignancy (60–87%) are qualified to indicate surgery (e.g., lobectomy) but there is insufficient evidence to recommend more radical resection.

Paschke/Cantara/Crescenzi/Jarzab/
Musholt/Sobrinho Simoes

What Is the Evidence that a Strategy of Decision-Making on the Extent of Thyroidectomy according to Molecular FNA Diagnostics Reduces Complication Rates (Permanent Hypoparathyroidism, Recurrent Nerve Paresis, and Complications Caused by Secondary Anesthesia) in Individuals with Thyroid Nodules?

Secondary surgical interventions are associated with the risk of permanent hypoparathyroidism, recurrent nerve paresis, and risks associated with surgical procedures and anesthesia in general. The frequency of complications following completion thyroidectomy is dependent on the timing of the secondary intervention [47] as well as the type of the primary resection (bilateral resection, lobectomy).

Assuming that the FNA molecular diagnostic profile only leads to primary total thyroidectomy in patients with malignant disease and overtreatment can be avoided, the long-term health outcome would be improved in a subgroup of patients that otherwise might experience permanent complications (1–4% permanent recurrent nerve paresis, about 3% permanent hypoparathyroidism).

What Is the Evidence that a Strategy of Decision-Making on the Extent of Thyroidectomy according to Molecular FNA Diagnostic Profile Reduces the Local Recurrence Rate of Thyroid Carcinomas or Improves Long-Term Survival?

Unpublished results of the PETS II study including 22,008 patients with 2,577 thyroid carcinomas (12%) revealed that completion thyroidectomy within 6 months was necessary in 11% of carcinoma patients [45]. 16% of patients with thyroid carcinoma were operated on for recurrent disease. Predominantly, these recurrences are located within the thyroid bed or lymph nodes of the central compartment. Prophylactic central lymph node dissection is usually not performed following total thyroidectomy and postoperative diagnosis of thyroid cancer despite the fact that it might be indicated if the diagnosis was made prior to surgery.

Therefore, the following clinical situation may contribute to a possible improved health outcome:

Total thyroidectomy instead of hemithyroidectomy or other subtotal resection

Total thyroidectomy and central lymph node dissection instead of total thyroidectomy without central lymph node dissection

Total thyroidectomy for differentiated thyroid carcinoma has been recommended in many countries in recent guidelines. The identification of specific mutations such as *BRAF* V600E and *TERT* in thyroid nodules >1 cm, which are associated with malignant tumors in almost 100% of cases, justify a total thyroidectomy and possibly prophylactic central lymph node dissection. However, as outlined in recent revisions of the guidelines of the British Thyroid Association [49] and the American Thyroid Association [50] a more conservative resection is advocated for low-risk thyroid carcinomas.

The benefit of prophylactic central node dissection for PTC is continuously debated, and prospective randomized controlled trials comparing total thyroidectomy with and without prophylactic lymph node dissection are unfeasible. Metareviews of retrospective data imply a reduction of local recurrences by prophylactic central lymph node dissection of about 5% [51, 52].

In summary, there is minor evidence that long-term health outcome may be improved by FNA molecular diagnostics.

Recommendation
Molecular diagnostics may improve long-term health outcome of surgery for thyroid cancer by reduction of completion thyroidectomies and local recurrences.

What Is the Evidence that Molecular FNA Diagnostics with a GEC Can Alter the Follow-Up for Thyroid Nodules?

GEC assigns a thyroid FNA to 1 of 2 categories, "benign" or "suspicious," on the basis of an expression pattern of mRNA extracted from 1 or 2 dedicated FNA needle passes [10, 39]. Thus, it is designed as a rule-out test that recognizes benign tumors. Several subsequent papers [29, 30, 32] evaluated the clinical validity of GEC but their results were referred to the histopathological outcome of the surgery and not to the follow-up of non-operated patients who did not undergo surgery because their GEC result was benign. A short overview of their data is given in Table 3.

The National Comprehensive Cancer Network (NCCN) thyroid carcinoma guidelines, version 2.2013, recommends molecular testing of indeterminate thyroid lesions with the following limitation: "If molecular testing predicts a risk of malignancy comparable to the risk of malignancy seen with a benign FNA cytology (approximately 5% or less), consider observation" [38].

The careful review of all studies did not reveal analyses devoted to the follow-up of nodules classified "benign" by GEC. The question is even how long should the follow-up be. According to Nou et al. [53], who followed up 2,010 benign nodules for an average of 8.5 years (range 0.25–18), repeated thyroid nodule evaluation should be done 2–4 years after the initial benign FNA. 18 false-negative thyroid malignancies were diagnosed at a mean 4.5 years (range 0.3–10) after the initial benign biopsy. Noteworthy, no distant metastases were observed and all patients were alive at an average of 11 years following the FNA. According to Lee et al. [54] a long-term follow-up of thyroid nodules, initially benign by FNA, does not improve the cancer detection rate. A group of 738 patients with FNA-confirmed benign thyroid nodules were observed at the MD Anderson Center. 92 of them were operated on after initial ultrasonography, whereas the 646 ones remaining were subjected to further follow-up. Among them, 366 patients had at least 1 ultrasonography. 226 of them were monitored for a period shorter than 3 years (median 13 months, range 1–35), while 140 subjects had a long-term follow-up (>3 years, median 57 months, range 36–154). The number of ultrasonographies and repeated FNAs was significantly higher in the long-term group; however, surgical resection of the nodule as well as malignancy rate at final histopathological examination was similar between the 2 groups. Thus, stopping further follow-up after 3 years may be considered [54]. However, a longer follow-up of 8.1 years was proposed by Liel et al. [55]. Based on the data of Nou et al. [53] a 2-year follow-up should be sufficient.

In the multicenter clinical study, reported by Alexander et al. [10], of the 174 GEC benign tumors ultimately 11 patients were operated on, mostly because of personal preference. Histopathological examination revealed a benign thyroid tumor in 10 of them and PTC in 1 patient [10]. This case raises some concerns on the validity of a GEC benign outcome for follow-up of non-operated patients.

The above-mentioned studies by McIver et al. [29], Harrel and Bimston [30], and Marti et al. [32], devoted to GEC clinical utility, did not present follow-up data.

To date there have been no studies validating the use of GEC in European patients. It is noteworthy that the frequency of nodular goiter in Europe, especially in the central part, is much higher than that in the USA, an iodine-sufficient country. Thus, if the frequency of nodular goiter is higher, the relative prevalence of malignant tumors will be lower than that in the USA. It should be emphasized that the NPV depends on the prevalence and therefore an independent validation for moderately iodine-deficient areas is necessary.

Recommendation

1. These data with reference to GEC are unbiased with regard to centers in which thyroid cancer prevalence in indeterminate lesions (Bethesda class III) is about 30%.
2. However, clinical utility of GEC has not yet been evaluated in centers with thyroid cancer prevalence lower than 30%.
3. We cannot recommend the use of GEC in clinical practice, because there is a lack of follow-up data for non-operated patients and there are no studies validating the use of GEC in European patients with a higher frequency of nodular goiter.

What Is the Evidence that Molecular FNA Diagnostics by Analyzing a Panel of Somatic Mutations Can Alter the Follow-Up?

It is commonly accepted that positive rule-in tests constitute an indication for surgery due to a substantial risk of malignancy [9]. However, the important question arises of what to do when the FNA result is indeterminate and the panel of somatic mutations is positive only with respect to *RAS* mutation? Total thyroidectomy has been proposed for such cases due to a high risk of cancer [11, 43]. This is not supported by results of others who demonstrated that the presence of *RAS* mutation was associated with thyroid cancer in 19, 38, and 57% of cases while *BRAF* mutations were associated with cancer in 100% cancer samples [16, 27]. Even those who propose total thyroidectomy found that *RAS* mutations in benign thyroid adenomas do not require extensive surgery as for thyroid cancer [11]. No data regarding the follow-up of patients with these nodules have been published.

The analysis of a panel of somatic mutations together with FNA in follicular lesions increased the diagnostic sensitivity from 60 to 80% in comparison with FNA alone. The risk of malignancy increased to 71% for mutation-positive FN and decreased to 18% for mutation-negative follicular lesions when compared to a 28% risk of malignancy evaluated by FNA only [16]. Unfortunately, also in this study no data with reference to follow-up were given.

Recommendation

A change of the follow-up of patients with thyroid nodules based on the analysis of a panel of somatic mutations *is only supported by unbiased data on malignancy of thyroid nodules. Panel-negative results should not be considered as evidence for a benign tumor.*

There are no follow-up data to decide whether RAS-positive nodules should be interpreted as malignant or not.

Conclusion

Recently, several molecular tests have been introduced into the clinical routine, such as GEC, seven-gene mutational panel, and tNGS. GEC recognizes benign lesions on the basis of an expression pattern of mRNA extracted from 1 or 2 dedicated FNA needle passes. GEC NPVs and PPVs are controversial in different studies and even the diagnostic performance needs to be clarified by the evaluation of the follow-up of patients who avoided surgery as a consequence of a GEC benign result. We recommend that in an individual with an indeterminate cytology a GEC test should not be routinely used to exclude malignancy.

Mutation panels intended to identify malignancies must include at least *BRAF*, *RAS* point mutations and *RET/PTC*, *NTRK*, *PAX8/PPARG* rearrangements. Several "homemade" methods comprising PCR with final Sanger sequencing and some commercial methods are available to screen for these alterations with the limitation that they cannot rule out malignancy with an NPV >95%. Therefore, surgery has to be recommended despite the exclusion of these mutations. Moreover, the role of *RAS* mutations is still unclear. While *BRAF*, *RET/PTC*, and *PAX8/PPARG* mutations predict thyroid malignancy with more than 95% probability, the risk of malignancy associated with *RAS* mutations is 57–87%. Thus, total thyroidectomy may not be recommended for *RAS* mutations due the lower PPV associated but lobectomy is preferred.

The use of tNGS with larger panels of genes offers an alternative method which is possibly able to achieve an NPV of 95% or more and a high PPV and sensitivity. However, issues such as result interpretation, cutoffs for mutation calling, bioinformatics analysis, and common SOPs need to be addressed. Therefore, the use of tNGS should currently be limited to research settings.

New molecules are emerging in thyroid cancer. In particular, investigations of the role of miRNAs in discriminating benign from malignant lesions in FNAC reported very preliminary results which need to be corroborated in larger studies with final histology as gold standard and adequate follow-up. Moreover, issues such as result interpretation, cutoffs, bioinformatics analysis, and common SOPs need to be addressed.

In conclusion, molecular tests may help clinicians to drive patient care and the surgical decision if the analysis is performed in specialized laboratories. These molecular tests require standardization of performance characteristics and appropriate calibration as well as analytic interpretation before clinical interpretation.

Disclosure Statement

The authors have no conflicts of interest to declare.

References

1 Gharib H, Goellner JR: Fine-needle aspiration biopsy of the thyroid: an appraisal. Ann Intern Med 1993;118:282–289.

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

3 Baloch ZW, Fleisher S, Livolsi VA, Gupta PK: Diagnosis of "follicular neoplasm": a gray zone in thyroid fine-needle aspiration cytology. Diagn Cytopathol 2001;26:41–44.

4 Gharib H, Papini E, Paschke R, Duick DS, Valcavi R, Hegedüs L, Vitti P; AACE/AME/ETA Task Force on Thyroid Nodules: American Association of Clinical Endocrinologists, Associazione Medici Endocrinologi and European Thyroid Association Medical Guidelines for Clinical Practice for the Diagnosis and Management of Thyroid Nodules. Endocr Pract 2010;16(suppl 1):1–43.

5 Baloch ZW, LiVolsi VA, Asa SL, Rosai J, Merino MJ, Randolph G, Vielh P, DeMay RM, Sidawy MK, Frable WJ: 2008 Diagnostic terminology and morphologic criteria for cytologic diagnosis of thyroid lesions: a synopsis of the National Cancer Institute Thyroid Fine-Needle Aspiration State of the Science Conference. Diagn Cytopathol 2008;36:425–437.

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

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. J Endocrinol Invest 2014;37:593–599.

8 Lobo C, McQueen A, Beale T, Kocjan G: The UK Royal College of Pathologists thyroid fine-needle aspiration diagnostic classification is a robust tool for the clinical management of abnormal thyroid nodules. Acta Cytol 2011;55:499–506.

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

10 Alexander EK, Kennedy GC, Baloch ZW, et al: Preoperative diagnosis of benign thyroid nodules with indeterminate cytology. N Engl J Med 2012;367:705–715.

11 Nikiforov YE, Ohori NP, Hodak SP, et al: Impact of mutational testing on the diagnosis and management of patients with cytologically indeterminate thyroid nodules: a prospective analysis of 1056 FNA samples. J Clin Endocrinol Metab 2011;96:3390–3397.

12 Nikiforov YE, Carty SE, Chiosea SI, et al: Highly accurate diagnosis of cancer in thyroid nodules with follicular neoplasm/suspicious for a follicular neoplasm cytology by Thyro-Seq v2 next-generation sequencing assay. Cancer 2014;120:3627–3634.

13 Labourier E, Shifrin A, Busseniers AE, et al: Molecular testing for miRNA, mRNA and DNA on fine needle aspiration improves the preoperative diagnosis of thyroid nodules with indeterminate cytology. J Clin Endocrinol Metab 2015;100:2743–2750.

14 Strickland KC, Howitt BE, Marqusee E, Alexander EK, Cibas ES, Krane JF, Barletta JA: The impact of noninvasive follicular variant of papillary thyroid carcinoma on rates of malignancy for fine-needle aspiration diagnostic categories. Thyroid 2015;25:987–992.

15 Auger M: Hürthle cells in fine-needle aspirates of the thyroid: a review of their diagnostic criteria and significance. Cancer Cytopathol 2014;122:241–249.

16 Eszlinger M, Piana S, Moll A, et al: Molecular testing of thyroid fine-needle aspirations improves pre-surgical diagnosis and supports the histologic identification of minimally invasive follicular thyroid carcinomas. Thyroid 2015;25:401–409.

17 Eszlinger M, Krogdahl A, Munz S, et al: 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.

18 Jennings L, Van Deerlin VM, Gulley ML; College of American Pathologists Molecular Pathology Resource Committee: Recommended principles and practices for validating clinical molecular pathology tests. Arch Pathol Lab Med 2009;133:743–755.

19 Teutsch SM, Bradley LA, Palomaki GE, Haddow JE, Piper M, Calonge N, Dotson WD, Douglas MP, Berg AO; EGAPP Working Group: The Evaluation of Genomic Applications in Practice and Prevention (EGAPP) Initiative: methods of the EGAPP Working Group. Genet Med 2009;11:3–14.

20 Sobrinho-Simões M, Máximo V, Rocha AS, Trovisco V, Castro P, Preto A, Lima J, Soares P: Intragenic mutations in thyroid cancer. Endocrinol Metab Clin North Am 2008;37:333–362.

21 Xing M: Molecular pathogenesis and mechanisms of thyroid cancer. Nat Rev Cancer 2013;13:184–199.

22 Nikiforova MN, Kimura ET, Gandhi M, et al: BRAF mutations in thyroid tumors are restricted to papillary carcinomas and anaplastic or poorly differentiated carcinomas arising from papillary carcinomas. J Clin Endocrinol Metab 2003;88:5399–5404.

23 Vinagre J, Almeida A, Pópulo H, Batista R, Lyra J, Pinto V, Coelho R, Celestino R, Prazeres H, Lima L, Melo M, da Rocha AG, Preto A, Castro P, Castro L, Pardal F, Lopes JM, Santos LL, Reis RM, Cameselle-Teijeiro J, Sobrinho-Simões M, Lima J, Máximo V, Soares P: Frequency of TERT promoter mutations in human cancers. Nat Commun 2013;4:2185.

24 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, Máximo V, Sobrinho-Simões M, Soares P: TERT promoter mutations are a major indicator of poor outcome in differentiated thyroid carcinomas. J Clin Endocrinol Metab 2014;99:E754–E765.

25 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. Clin Endocrinol 2016;84:315–320.

26 Renshaw AA: Significance of repeatedly non-diagnostic fine-needle aspirations. Am J Clin Pathol 2011;135:750–752.

27 Krane JF, Cibas ES, Alexander EK, Paschke R, Eszlinger M: Molecular analysis of residual ThinPrep material from thyroid FNAs increases diagnostic sensitivity. Cancer Cytopathol 2015;123:356–361.

28 Duick DS, Klopper JP, Diggans JC, et al: 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.

29 McIver B, Castro MR, Morris JC, et al: An independent study of a gene expression classifier (Afirma) in the evaluation of cytologically indeterminate thyroid nodules. J Clin Endocrinol Metab 2014;99:4069–4077.

30 Harrell RM, Bimston DN: Surgical utility of Afirma: effects of high cancer prevalence and oncocytic cell types in patients with indeterminate thyroid cytology. Endocr Pract 2014;20:364–369.

31 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 Cytopathol 2014;122:737–744.

32 Marti JL, Avadhani V, Donatelli LA, et al: Wide inter-institutional variation in performance of a molecular classifier for indeterminate thyroid nodules. Ann Surg Oncol 2015;22:3996.

33 Nikiforova MN, Wald AI, Roy S, Durso MB, Nikiforov YE: Targeted next-generation sequencing panel (ThyroSeq) for detection of mutations in thyroid cancer. J Clin Endocrinol Metab 2013;98:E1852–E1860.

34 Le Mercier M, D'Haene N, De Nève 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.

35 Coco S, Truini A, Vanni I, Dal Bello MG, Alama A, Rijavec E, Genova C, Barletta G, Sini C, Burrafato G, Biello F, Boccardo F, Grossi F: Next generation sequencing in non-small cell lung cancer: new avenues toward the personalized medicine. Curr Drug Targets 2015;16:47–59.

36 Malapelle U, Vigliar E, Sgariglia R, Bellevicine C, Colarossi L, Vitale D, Pallante P, Troncone G: Ion torrent next-generation sequencing for routine identification of clinically relevant mutations in colorectal cancer patients. J Clin Pathol 2015;68:64–68.

37 Chevrier S, Arnould L, Ghiringhelli F, Coudert B, Fumoleau P, Boidot R: Next-generation sequencing analysis of lung and colon carcinomas reveals a variety of genetic alterations. Int J Oncol 2014;45:1167–1174.

38 Tuttle RM, Haddad RI, Ball DW, et al: Thyroid carcinoma, version 2.2014. J Natl Compr Canc Netw 2014;12:1671–1680; quiz 1680.

39 Alexander EK, Schorr M, Klopper J, et al: Multicenter clinical experience with the Afirma gene expression classifier. J Clin Endocrinol Metab 2014;99:119–125.

40 Cantara S, Capezzone M, Marchisotta S, et al: Impact of proto-oncogene mutation detection in cytological specimens from thyroid nodules improves the diagnostic accuracy of cytology. J Clin Endocrinol Metab 2010;95:1365–1369.

41 Beaudenon-Huibregtse S, Alexander EK, Guttler RB, et al: Centralized molecular testing for oncogenic gene mutations complements the local cytopathologic diagnosis of thyroid nodules. Thyroid 2014;24:1479–1487.

42 Moses W, Weng J, Sansano I, et al: Molecular testing for somatic mutations improves the accuracy of thyroid fine-needle aspiration biopsy. World J Surg 2010;34:2589–2594.

43 Gupta N, Dasyam AK, Carty SE, et al: RAS mutations in thyroid FNA specimens are highly predictive of predominantly low-risk follicular-pattern cancers. J Clin Endocrinol Metab 2013;98:E914–E922.

44 Franc B, de la Salmonière P, Lange F, et al: Interobserver and intraobserver reproducibility in the histopathology of follicular thyroid carcinoma. Hum Pathol 2003;34:1092–1100.

45 Dralle H, Stang A, Sekulla C, et al: Surgery for benign goiter in Germany: fewer operations, changed resectional strategy, fewer complications (in German). Chirurg 2014;85:236–245.

46 Wienhold R, Scholz M, Adler JR, et al: The management of thyroid nodules: a retrospective analysis of health insurance data. Dtsch Arztebl Int 2013;110:827–834.

47 Glockzin G, Hornung M, Kienle K, et al: Completion thyroidectomy: effect of timing on clinical complications and oncologic outcome in patients with differentiated thyroid cancer. World J Surg 2012;36:1168–1173.

48 Barczynski M, Konturek A, Stopa M, et al: Total thyroidectomy for benign thyroid disease: is it really worthwhile? Ann Surg 2011;254:724–729; discussion 729–730.

49 Perros P, Colley S, Boelaert K, Evans C, Evans RM, Gerrard GE, Gilbert JA, Harrison B, Johnson SJ, Giles TE, Moss L, Lewington V, Newbold KL, Taylor J, Thakker RV, Watkinson J, Williams GR; British Thyroid Association: Guidelines for the management of thyroid cancer. Clin Endocrinol 2014;81(supp 1):1–122 .

50 Haugen BR, Alexander EK, Bible KC, Doherty GM, Mandel SJ, Nikiforov YE, Pacini F, Randolph GW, Sawka AM, Schlumberger M, Schuff K, 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. Thyroid 2016;26: 1–133.

51 Sitges-Serra A, Lorente L, Mateu G, Sancho JJ: Therapy of Endocrine Disease: central neck dissection: a step forward in the treatment of papillary thyroid cancer. Eur J Endocrinol 2015;173:R199–R206.

52 Wang TS, Cheung K, Farrokhyar F, Roman SA, Sosa JA: A meta-analysis of the effect of prophylactic central compartment neck dissection on local regional recurrence rates in patients with papillary thyroid cancer. Ann Surg Oncol 2013;20:3477–3483.

53 Nou E, Kwong N, Alexander LK, Cibas ES, Marqusee E, Alexander EK: Determination of the optimal time interval for repeat evaluation after a benign thyroid nodule aspiration. J Clin Endocrinol Metab 2014;99:510–516.

54 Lee S, Skelton TS, Zheng F, et al: The biopsy-proven benign thyroid nodule: is long-term follow-up necessary? J Am Coll Surg 2013;217: 81–88; discussion 88–89.

55 Liel Y, Ariad S, Barchana M: Long-term follow-up of patients with initially benign thyroid fine-needle aspirations. Thyroid 2001;11: 775–778.

56 Nishino M: Molecular cytopathology for thyroid nodules: a review of methodology and test performance. Cancer Cytopathol 2016;12: 14–27.

57 Nikiforov YE, Steward DL, Robinson-Smith TM, et al: Molecular testing for mutations in improving the fine-needle aspiration diagnosis of thyroid nodules. J Clin Endocrinol Metab 2009;94:2092–2098.