Characterization of thyroid ‘follicular neoplasms’ in fine-needle aspiration cytological specimens using a panel of immunohistochemical markers: a proposal for clinical application

E Saggiorato, R De Pompa¹, M Volante², S Cappia¹, F Arecco, A P Dei Tos³, F Orlandi and M Papotti⁴

Section of Endocrinology, Department of Clinical and Biological Sciences, University of Turin, Regione Gonzole 10, 10043 Orbassano (TO), Italy
¹Division of Pathology, Department of Clinical and Biological Sciences, University of Turin, Turin, Italy
²Department of Biomedical Sciences and Oncology, University of Turin, Turin, Italy
³Department of Pathology, Regional Hospital, Treviso, Italy

(Requests for offprints should be addressed to E Saggiorato; Email: enrico.saggiorato@unito.it)

Abstract

The distinction of benign from malignant follicular thyroid neoplasms remains a difficult task in diagnostic fine-needle aspiration cytology, and some discrepant results have been reported for the individual immunocytochemical markers of malignancy proposed so far. The aim of this study was to test if the combined use of a panel of markers could improve the diagnostic accuracy in the preoperative cytological evaluation of ‘follicular neoplasms’ in an attempt to reduce the number of thyroidectomies performed for benign lesions. The immunocytochemical expression of galectin-3, HBME-1, thyroperoxidase, cytokeratin-19 and keratan-sulfate was retrospectively analyzed in 125 consecutive fine-needle aspiration samples (cell blocks) of indeterminate diagnoses of ‘follicular thyroid neoplasm’, and compared with their corresponding surgical specimens, including 33 follicular carcinomas, 42 papillary carcinomas and 50 follicular adenomas. Statistical analysis on each marker confirmed that galectin-3 and HBME-1 were the most sensitive (92% and 80% respectively) and specific (94% and 96% respectively) molecules. The use of these two markers sequentially in non-oncocytic lesions (testing HBME-1 as a second marker whenever galectin-3 proved negative) increased the sensitivity and specificity up to 97% and 95% respectively. In oncocytic lesions, HBME-1 proved to be less sensitive, and the sequential combination of galectin-3 and cytokeratin-19 reached 100% of both specificity and sensitivity. Our data showed that, as compared with the use of single markers, the sequential combination of two markers represents the most accurate immunohistochemical panel in managing patients with a fine-needle aspiration biopsy diagnosis of ‘follicular neoplasms’, especially in otherwise controversial categories such as oncocytic tumours. The combination of three or more markers did not substantially improve the diagnostic accuracy of the test.

Introduction

Fine-needle aspiration biopsy (FNAB) is a well-established diagnostic technique for the preoperative evaluation of thyroid nodules, allowing a significant reduction in the number of surgical operations (Mazzaferri 1993, Hegedus et al. 2003). However, an important limitation of FNAB is the lack of sensitivity in the evaluation of ‘follicular neoplasms’, due to its inability to differentiate benign (follicular adenomas)
from malignant (follicular carcinomas) lesions (Baloch et al. 2002, Castro & Gharib 2003). Moreover, the cytological diagnosis of a follicular variant of papillary carcinoma represents an additional problem, since the diagnostic cytological criteria of papillary carcinoma (cytoplasmic invaginations into the nucleus or abundant nuclear grooves) may be lacking (in such cases the nuclei show coarse chromatin similar to those of follicular carcinoma) or not completely unequivocal (features suggestive but not diagnostic of papillary carcinomas), as also reported in histological practice (Lloyd et al. 2004); thus the tumour is mainly inclined to be cytologically interpreted as a ‘follicular neoplasm’ (Kini 1987).

To date, no effective methods have been established to identify preoperatively malignant follicular tumours of the thyroid.

Although several additional procedures, such as advanced image systems (Baloch et al. 2000, Harms et al. 2002, Papini et al. 2002) and modified techniques for thyroid biopsy (Tangpricha et al. 2001), have been developed to increase the FNAB diagnostic accuracy of thyroid lesions, nowadays none of them has led to any significant improvement in solving the ‘follicular neoplasm’ dilemma (Yang et al. 2003). Thus, the identification of new diagnostic approaches is of paramount importance in order to provide new criteria for an early preoperative diagnosis, and reduce the diagnostic category of ‘indeterminate’ FNABs.

To this purpose, recent advances in molecular diagnostics, such as immunocytochemistry, enzyme activity assays and RT-PCR, allowed a further analysis of FNAB material in an attempt to differentiate benign from malignant thyroid nodules (Haugen et al. 2002). Several molecules involved in carcinogenic processes have been proposed as markers of thyroid malignancy (oncogene products, altered enzymes, integrins, cadherins, lectins, etc.) (Ringel 2000, Serini et al. 1999, Nikiforova et al. 2003, Cohen et al. 2004). Among them, telomerase (Saji et al. 1999, Liou et al. 2003), high mobility group proteins I(Y) (HMGI(Y)) (Chiappetta et al. 1998), the cell surface mesothelial antigen HBME-1 (Sack et al. 1997, van Hoeven et al. 1998, Prasad et al. 2004), thyroperoxidase (TPO) (Henry et al. 1994), cytokeratin-19 (CK19) (Khurana et al. 2001, Weber et al. 2004), thyroglobulin (Tg) (Christensen et al. 1994), thyroperoxidase (TPO) (Henry et al. 1994), cytokeratin-19 (CK19) (Khurana et al. 2003) and galectin-3 (GAL3) (Bartolazzi et al. 2001, Lloyd 2001) seem to be the most promising molecules in significantly increasing FNAB accuracy. The molecular detection of specific alterations, mainly restricted to papillary carcinoma, such as BRAF mutations and RET/PTC translocations, have also been proposed in FNAB samples (Cheung et al. 2001a, Cohen et al. 2004, Salvatore et al. 2004) but, although they may represent promising future methods, they are not easily accessible to most routine laboratories.

GAL3 polypeptide is a member of the oligosaccharide-selective binding protein family known as lectins (Barondes et al. 1994). GAL3 plays important roles in cell–cell and cell–matrix interactions, extracellular matrix organization, mRNA splicing, cell growth and apoptosis, neoplastic transformation and metastasization (Liu et al. 2002, Yang & Liu 2003). As reported by several studies, this lectin is expressed in thyroid carcinoma but not in normal thyrocytes and in benign lesions such as follicular adenoma (Xu et al. 1995, Fernandez et al. 1997). In particular, we previously reported that cytoplasmic GAL3, selectively expressed in malignant thyroid cells, is easily detectable also on FNAB cytological samples (Orlandi et al. 1998, Bartolazzi et al. 2001).

The HBME-1 antigen, which is recognized by the monoclonal antibody HBME-1 (Sheibani et al. 1992), was shown to be a useful marker of follicular-derived malignant tumours of the thyroid in both surgical and FNAB cytological samples (Miettinen & Karkkainen 1996, Sack et al. 1997, van Hoeven et al. 1998, Cheung et al. 2001b).

TPO is an enzyme present in all non-malignant thyroid follicular cells. It is immunologically altered in cancer cells and one of the monoclonal antibodies raised against intact TPO (MoAB47) fails to recognize the TPO forms of malignant thyroid tumours (De Micco et al. 1991). Some studies have shown that TPO immunostaining highly improves the FNAB diagnostic accuracy of thyroid follicular lesions (Henry et al. 1994, Christensen et al. 2000, Raffaelli et al. 2001, Weber et al. 2004).

CK19 is a low molecular weight cytokeratin widely present in simple epithelia and in basal cell layers of stratified epithelium (Rapheal et al. 1994, Porter & Lane 2003). In both FNAB cytological and surgical samples of thyroid tumours, CK19 has been found to be strongly and diffusely expressed in papillary carcinoma, whereas it is heterogeneously expressed in follicular carcinomas and absent or focally expressed in follicular adenomas (Schröder et al. 1996, Fonseca et al. 1997, Miettinen et al. 1997, Beesley & McLaren 2002, Khurana et al. 2003).

Keratan-sulphate (KS) is a glycoprotein complex having an unusual relative ratio of highly sulphated disaccharides and carrying a novel non-characterized oligosaccharide moiety at the non-reducing terminus, which is recognized by the monoclonal antibody 373E1 (Magro et al. 2003). This aberrant glycoprotein complex has been found almost exclusively in thyroid papillary carcinomas. Moreover, it has been reported
that papillary carcinoma-specific KS-bearing macromolecules are unique glycoforms of thyroglobulin and transferrin (Magro et al. 2003). The peculiar KS epitope has been proposed as a novel marker for the differential diagnosis of benign and malignant thyroid tumours (Magro et al. 2003).

However, based on the data in the literature, it is clear that no ‘magic marker’ that distinguishes follicular adenomas from carcinomas has yet been found (Baloch & LiVolsi 2002). In fact, several reports on this topic have provided discrepant results. GAL3 was found in the hands of ourselves and others (Gasbarri et al. 1999, Inohara et al. 1999, Saggiorato et al. 2001, 2004, Papotti et al. 2002, Volante et al. 2004b, Weber et al. 2004) to be a highly sensitive marker, but this was not confirmed by other authors (Nascimento et al. 2001, Martins et al. 2002, Niedziela et al. 2002, Takano et al. 2003, Mehrotra et al. 2004). The same holds true for HBME-1 and CK19 (Sahoo et al. 2001, Mai et al. 2002). Many of these discrepancies stem from the apparently false-positive staining of some of these markers in normal thyroid or adenomas (Bartolazzi et al. 2003, Saggiorato et al. 2004).

In the attempt to solve the aforementioned issues, we designed this study on a relatively large number of histologically proven thyroid follicular lesions, and tested a panel of immunohistochemical markers (GAL3, HBME-1, TPO, CK19 and KS) on both FNAB cell blocks and all the corresponding surgical specimens. The goal was not only to validate the role of each individual molecule, but also to establish if any combination of these selected markers could be helpful in identifying malignant tumours with higher accuracy in the indeterminate subset of FNABs, thus reducing the number of thyroidectomies for benign follicular lesions. In this respect, we showed that the combination of GAL3 and HBME-1 for conventional follicular lesions, or GAL3 and CK19 in the case of oncocytic lesions, provides the highest sensitivity in FNAB samples as compared with a much larger (and expensive) panel of markers. Furthermore, we have developed a clinico-pathological algorithm for the management of patients with an indeterminate FNAB, based on the results of these two marker combination analyses.

Materials and methods

Cytological and histological samples

From the pathology files of San Luigi Hospital (Orbassano, Turin, Italy) and San Giovanni Battista Hospital (Turin, Italy), in the years 2002–2004, 125 consecutive cases of FNAB of the thyroid were collected, in which a diagnosis of ‘follicular neoplasm’ was made, and the subsequent surgical specimens were filed in the same hospitals and were available for review.

At operation, all patients had a histological diagnosis of follicular adenoma or well-differentiated thyroid carcinoma. The preoperative FNAB diagnosis was:

1. ‘follicular neoplasm’ (not otherwise specified) in 73/125 samples, histologically classified as follicular adenomas (43 cases), follicular carcinomas (19 cases) and follicular variant of papillary carcinomas (11 cases);
2. ‘oncocytic cell neoplasm’ in 24/125 specimens, histologically proven as oncocytic adenomas (seven cases), oncocytic carcinomas (14 cases) and oncocytic (follicular) variant of papillary carcinomas (three cases);
3. ‘follicular/papillary neoplasm’ (suspected papillary carcinoma) in 28/125 samples, histologically confirmed as pure follicular variants (14 cases), or having minor classical (12 cases) or solid/trabecular (two cases) components. The whole series included 22 males and 103 females, with a median age of 46 years (range 11–89 years). Written informed consent was obtained from each patient, and the study was approved by the San Luigi Hospital review board.

FNABs

Preoperative FNABs were performed with a 22 gauge needle attached to a 30 ml plastic syringe. The aspirated fluid was in part expelled and smeared onto charged slides, fixed and stained with a rapid haematoxylin and eosin (H&E) method for the evaluation of adequacy. The remaining material was used to prepare alcohol-fixed cell blocks (43). Sections of cell block serial to those employed for H&E routine evaluation were used for immunohistochemistry.

Surgical specimens

The corresponding surgical samples of all patients were formalin fixed and paraffin embedded for routine histopathological diagnosis and immunohistochemical staining.

Histopathological classification

For all the 125 FNAB cases included in the study, histopathological review and classification of the corresponding surgical samples was performed, according to WHO thyroid tumour classification (De Lellis et al. 2004). From three to ten H&E-stained slides as well as a paraffin block representative of the lesion were available for review and for the immunocytochemical analyses.
procedure. A total of 50 follicular adenomas, 33 follicular carcinomas and 42 papillary carcinomas with a predominant degree of follicular pattern of growth were examined. Follicular adenomas consisted of 32 microfollicular, 11 trabecular/solid and seven oncocytic cell types whereas, in the follicular carcinoma group, 14 had oncocytic features and 19 were conventional follicular carcinomas (with a minor trabecular or insular component in six of them). Among papillary carcinomas, we observed 25 follicular variants, three oncocytic follicular variants and 14 cases of follicular variant with minor classical (12 cases) and solid/trabecular (two cases) features. Serial sections of a representative block were obtained from all cases for immunohistochemistry.

**Antibodies**

The primary monoclonal antibodies used in this study (with their working dilutions) were as follows: purified 9C4 to GAL3 (1:400; Novocastra Laboratories Ltd, Newcastle, Tyne and Wear, UK), HBME-1 to HBME-1 antigen and MoAb47 to TPO (1:500 and 1:200 respectively; DakoCytomation, Glostrup, Denmark), b170 to CK19 (1:150; Novocastra Laboratories Ltd), 9C4 to GAL3 (1:400; Novocastra Laboratories Ltd, Newcastle, Tyne and Wear, UK), HBME-1 to HBME-1 antigen and MoAb47 to TPO (1:500 and 1:200 respectively; DakoCytomation, Glostrup, Denmark), and 373E1 to KS (1:100; LabVision Corporation, Fremont, CA, USA).

**Immunoperoxidase technique**

Both FNAB cytological cell blocks and surgical specimens were cut at 4 μm thickness, overlaid on poly-l-lysine-coated slides, dewaxed in xylene, rehydrated in decreasing ethanol concentrations and incubated for 10 min in phosphate-buffered saline (PBS) (pH 7.4). A heat-induced antigen retrieval procedure was carried out for all the antigens by placing slides in 0.01 M sodium citrate buffer (pH 6.0) (for GAL3, HBME-1 and KS) and in 0.01 M EDTA buffer (pH 8.0) (for TPO and CK19) in a microwave oven set at high power for three consecutive cycles of 5 min each. These slides were left to cool for 20 min at room temperature and rinsed in PBS. Endogenous peroxidase activity was quenched with methanol-hydrogen peroxide (3%) for 15 min at room temperature. Slides were then washed twice in PBS for 5 min. Cell blocks and tissue sections were then incubated in the blocking solution (ChemMate Buffer Kit; DakoCytomation), and subsequently with the primary antibodies (see above) in a humidified chamber at room temperature for 45 min (HBME-1 and CK19), or at 4 °C overnight (GAL3, TPO and KS). After a prolonged wash in PBS, the sections were incubated with an enzyme-labelled polymer pre-conjugated with anti-mouse secondary antibodies (a biotin-free detection system) (Envision System; DakoCytomation) at room temperature for 30 min. The sections were finally washed three times in PBS, and incubated with 3'-3'-diaminobenzidine-tetrahydrochloride for 10 min. Slides were subsequently rinsed in tap water, counterstained with Mayer’s haemalum solution, mounted in Entellan (Merck, Darmstadt, Germany) and examined with a DM-RBE Leica photomicroscope (Leica Microsystems AG, Wetzlar, Germany). Positive controls were histiocytes for GAL3, a case of mesothelioma for HBME-1, normal thyroid follicles for TPO, skin for CK19 and a lymph node metastasis of a papillary thyroid carcinoma for KS. Negative controls were obtained by omitting the primary antibody.

**Scoring of staining and statistical evaluation**

GAL3, HBME-1, TPO, CK19 and KS immunostainings were evaluated blindly by three independent observers (ES, MP and RDP) without knowledge of the previously established histological diagnosis, using the following semiquantitative scale: −no reactivity, +/− focal reactivity, <10% of the neoplastic cells, + moderate reactivity, 10–50% of the neoplastic cells and ++ diffuse positivity, >50% of the neoplastic cells.

Immunostainings for GAL3 (cytoplasmic expression only), HBME-1, CK19 and KS were considered positive in the presence of >10% immunoreactive neoplastic cells, whereas for TPO a positive case had >50% neoplastic cells stained.

Sensitivity, specificity, positive/negative predictive values and diagnostic accuracy of each marker were calculated as previously described (Orlandi et al. 1998), and their sequential and simultaneous combinations were compared. The overall sensitivity and specificity of sequential joint tests were computed as: (true positive test A + true positive test B)/(true positive test A + false negative test A) and (true negative test B)/(true negative test B + false positive test B) (Di Orio 1994) respectively. The histological diagnosis on the surgical specimens was regarded as the gold standard.

The aforementioned statistical parameters were computed in the adenoma vs carcinoma groups and, separately, in both oncocytic and non-oncocytic conventional neoplasms.

The differences of each marker expression between cytological and histological counterparts, adenomas and carcinomas, and papillary and follicular carcinomas were analyzed by means of the χ² test, where a value of $P \leq 0.05$ was considered as statistically significant. Differences of marker expression between
oncocytic cell carcinomas and non-oncocytic carcinomas were also analyzed using two-tail Fisher exact test, where a value of $P \leq 0.01$ was considered as significant.

Results

Thyroid adenomas

Immunohistochemical analysis of FNAB cytological specimens of adenomas with anti-GAL3 antibody showed absent or scant (<10%) cytoplasmic expression in the neoplastic cells of 47 out of 50 cases, whereas nuclear staining was observed in focal areas or in single cells of several adenoma specimens. Three cases (6%) were diffusely positive in more than 80% of neoplastic cells. Histologically, these cases were proven adenomas associated with marked cytological atypia and incomplete (non-full thickness) capsular penetration. Such surgical specimens had the same pattern of GAL3 staining. Normal follicular cells surrounding each tumour nodule were GAL3 negative. Macrophages, endothelial cells and polymorphonuclear inflammatory cells expressed GAL3, as expected.

HBME-1 was negative or had cell-membrane and apical immunoreactivity in less than 10% of neoplastic cells in 48 out of 50 cytological cases. Two samples (4%) had cell membrane and apical staining of more than 80% of tumour cells. The same pattern of cell block staining was observed in the corresponding surgical specimens, which were confirmed to be adenomas, according to currently accepted diagnostic criteria. Normal follicular thyroid cells were negative for HBME-1, as well as fibroblasts, endothelial cells, histiocytes, lymphocytes and polymorphonuclear inflammatory cells.

TPO immunoreactivity in more than 50% of neoplastic cells was observed in 43 out of 50 cytological cases (Fig. 1A). Seven adenomas (14%) had cytoplasmic expression in less than 40% of neoplastic cells. The same pattern of positive-staining cell block cases was observed in the corresponding surgical specimens (Fig. 1B). Peri-tumoral thyroid gland was strongly positive. Fibroblasts, endothelial cells, histiocytes, lymphocytes and polymorphonuclear inflammatory cells were unreactive.

CK19 showed absent or focal cytoplasmic expression (<10% of neoplastic cells) in 45 out of 50 cases. Five cases (10%) showed positive staining in the cytoplasm of more than 70% of neoplastic cells. The same pattern of cell block staining was observed in the corresponding surgical specimens. The normal gland surrounding each tumour specimen was either negative or only focally positive. None of the non-follicular cells had any immunoreactivity.

All but one FNAB cytological cases were negative or focally stained (<10%) for KS. The remaining case (2%) was positive in more than 70% of the cells with both cell membrane and cytoplasmic patterns. The same pattern of cell block staining was observed in the corresponding surgical specimens. Normal thyroid surrounding tumour nodules was negative or contained occasionally positive cells. None of the non-follicular cells had any immunoreactivity. The prevalences of individual marker expression in the benign tumours are summarized in Table 1.

Thyroid follicular carcinomas

Twenty-eight out of thirty-three (84.8%) FNAB sections of follicular carcinoma had a cytoplasmic GAL3 expression in more than 10% malignant cells, whereas five samples were either negative or focally (<10%) positive. All but one (a follicular carcinoma with an insular component) corresponding surgical specimens had the same pattern of staining.

HBME-1 had cell-membrane and apical staining of more than 10% neoplastic cells in 21 out of 33 cases (63.6%). Ten out of the twelve negative cell blocks were oncocytic tumours, corresponding to 71% of all oncocytic cell carcinomas of this series. All the corresponding surgical cases showed the same pattern of staining.

Twenty-three out of thirty-three (69.7%) cytological specimens of follicular carcinomas were either TPO negative or had cytoplasmic TPO immunoreactivity in less than 50% of neoplastic cells (Fig. 1C), whilst the remaining ten samples were diffusely positive (>60% of cells). All the corresponding surgical specimens showed the same pattern of staining (Fig. 1D).

CK19 was positive in 21 out of 33 (63.6%) FNABs of follicular carcinomas. All but two corresponding surgical cases showed the same staining pattern. Two CK19-positive surgical specimens where either negative or only focally (<8%) positive in their corresponding FNAB material.

Only 21.2% (seven out of 33) of FNAB cytological samples of follicular cancers had more than 10% cells positive for KS (at both cell-membrane and cytoplasmic level) (Fig. 1E), whereas in the surgical specimen group one-third (10/33) of cases were positive (Fig. 1F). The prevalences of individual marker expression in follicular carcinomas are summarized in Table 1.
Figure 1 Immunoperoxidase staining of immunohistochemical markers in thyroid tumors. TPO is expressed in a follicular adenoma in both (A) the FNAB cytological cell block and (B) the corresponding surgical specimen. Conversely, in a follicular carcinoma, TPO is negative in both (C) the cytological cell block and (D) the corresponding surgical specimen. KS is strongly expressed in these cases of (E and F) follicular carcinoma and (G and H) follicular variant of papillary carcinoma in both (E and G) cytological cell blocks and (F and H) the corresponding surgical specimens. Non-oncocytic carcinomas generally diffusely express (I) GAL3 and (M) HBME-1 and (O) only focally CK19. Conversely, oncocytic follicular carcinomas are strongly positive for (L) GAL3 and (P) CK19, rather than (N) HBME-1. (Immunoperoxidase, ×160).
Thyroid papillary carcinomas

GAL3 had diffuse (over two-thirds of cancer cells) cytoplasmic expression in 41 out of 42 (97.6%) FNAB cytological cases. All the corresponding surgical specimens were moderately to diffusely stained for GAL3.

Thirty-nine out of forty-two (92.9%) FNAB cytological samples had diffuse HBME-1 immunoreactivity (>50% of malignant cells), whilst three cases were either negative or only focally (<10%) positive. One out of the three oncocytic papillary carcinomas was negative. All corresponding surgical specimens showed the same staining pattern.

With regard to TPO, 37 out of 42 (76.2%) FNAB cytological specimens were either negative or stained <50% of neoplastic cells. All the corresponding surgical samples expressed TPO with the same pattern of staining.

CK19 was positive in 36 out of 42 (85.7%) cases. Six specimens were either negative or had focal (<10%) immunoreactivity. All but one corresponding surgical cases had the same pattern of staining.

KS was diffusely expressed in 29 out of 42 (69.0%) FNAB cytological specimens (Fig. 1G), whereas the other 13 were either negative or showed only occasional positive cells. All but three corresponding surgical samples showed the same pattern of staining (Fig. 1H). The prevalences of individual marker expression in papillary carcinomas are summarized in Table 1.

Statistical analysis

Single marker

No significant differences were observed between the expression of each marker in FNAB cytological and surgical specimens. All five markers were significantly more expressed (GAL3, HBME-1, CK19, KS) or reduced (TPO) in carcinomas than adenomas \( (P<0.001) \). Moreover, GAL3, HBME-1, CK19 and KS expression, as well as TPO immunoreactivity reduction were significantly more associated with papillary carcinomas than with follicular carcinomas \( (P<0.05) \). In the papillary carcinoma group, no differences were observed for all markers tested among pure follicular variants and cases showing minor classical or solid/trabecular growth patterns.

Although all markers were highly specific \( (>90\%) \) in both cytological and surgical samples, they noticeably differed in sensitivity, which ranged from 48% for KS to 92% for GAL3 in cytological specimens (Table 2), and from 56% for KS to 94% for GAL3 in the corresponding surgical specimens.

GAL3 and HBME-1 provided a higher accuracy than other markers in both FNAB cytological and surgical materials. In particular, GAL3 was less specific but more sensitive than HBME-1 (Table 2). A progressive decrease of accuracy values was found for TPO, CK19 and KS (Table 2). No significant differences were observed between the expression of GAL3 (Fig. II and L), TPO, CK19 and KS in oncocytic cell tumours and non-oncocytic ones on both cytological and surgical specimens. Conversely, HBME-1 expression was significantly more associated with non-oncocytic carcinomas than oncocytic ones \( (P<0.001) \) in both FNAB cytological (Fig. 1M and N) and surgical samples. HBME-1 showed a lower diagnostic accuracy in oncocytic lesions (50%) than CK19 (Fig. 1O and P) and TPO (Table 3).

Table 1 Prevalence of positivity of immunohistochemical markers in the benign and malignant thyroid tumour groups.

<table>
<thead>
<tr>
<th></th>
<th>Cytological samples</th>
<th>Surgical samples</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>GAL3</td>
<td>HBME-1</td>
</tr>
<tr>
<td>All tumours</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Follicular adenoma</td>
<td>3/50</td>
<td>2/50</td>
</tr>
<tr>
<td>Follicular carcinomas</td>
<td>28/33</td>
<td>21/33</td>
</tr>
<tr>
<td>Papillary carcinomas*</td>
<td>41/42</td>
<td>39/42</td>
</tr>
<tr>
<td>Oncocytic subset of tumours</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Follicular adenoma</td>
<td>0/7</td>
<td>1/7</td>
</tr>
<tr>
<td>Follicular carcinomas</td>
<td>13/14</td>
<td>4/14</td>
</tr>
<tr>
<td>Papillary carcinomas*</td>
<td>2/3</td>
<td>2/3</td>
</tr>
<tr>
<td>Non-oncocytic subset of tumours</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Follicular adenoma</td>
<td>3/43</td>
<td>1/43</td>
</tr>
<tr>
<td>Follicular carcinomas</td>
<td>15/19</td>
<td>17/19</td>
</tr>
</tbody>
</table>

*Histologically proven as pure follicular variants of papillary carcinomas or papillary carcinomas with a predominant follicular pattern of growth.
### Table 2 Discrimination between benign and malignant thyroid lesions by single marker or by simultaneous/sequential combinations of markers in cytological specimens.

<table>
<thead>
<tr>
<th></th>
<th>SN (%)</th>
<th>SP (%)</th>
<th>PV+ (%)</th>
<th>PV− (%)</th>
<th>AC (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Single marker</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GAL3</td>
<td>92.00</td>
<td>94.00</td>
<td>95.83</td>
<td>88.68</td>
<td>92.80</td>
</tr>
<tr>
<td>HBME-1</td>
<td>80.00</td>
<td>96.00</td>
<td>96.77</td>
<td>76.19</td>
<td>86.40</td>
</tr>
<tr>
<td>TPO</td>
<td>80.00</td>
<td>86.00</td>
<td>89.55</td>
<td>74.14</td>
<td>82.40</td>
</tr>
<tr>
<td>CK19</td>
<td>76.00</td>
<td>90.00</td>
<td>91.94</td>
<td>71.43</td>
<td>81.60</td>
</tr>
<tr>
<td>KS</td>
<td>48.00</td>
<td>98.00</td>
<td>97.30</td>
<td>55.68</td>
<td>68.00</td>
</tr>
<tr>
<td><strong>Two simultaneous markers</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GAL3 + HBME-1</td>
<td>97.33</td>
<td>90.00</td>
<td>93.59</td>
<td>95.74</td>
<td>94.40</td>
</tr>
<tr>
<td>GAL3 + CK19</td>
<td>98.67</td>
<td>84.00</td>
<td>90.24</td>
<td>97.67</td>
<td>92.80</td>
</tr>
<tr>
<td>HBME-1 + CK19</td>
<td>92.00</td>
<td>88.00</td>
<td>92.00</td>
<td>88.00</td>
<td>90.40</td>
</tr>
<tr>
<td>GAL3 + TPO</td>
<td>96.00</td>
<td>80.00</td>
<td>87.80</td>
<td>93.02</td>
<td>89.60</td>
</tr>
<tr>
<td>HBME-1 + TPO</td>
<td>90.67</td>
<td>82.00</td>
<td>88.31</td>
<td>85.42</td>
<td>87.20</td>
</tr>
<tr>
<td><strong>Three simultaneous markers</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GAL3 + HBME-1 + CK19</td>
<td>100.00</td>
<td>82.00</td>
<td>78.57</td>
<td>97.67</td>
<td>92.60</td>
</tr>
<tr>
<td>GAL3 + HBME-1 + TPO</td>
<td>98.67</td>
<td>76.00</td>
<td>86.05</td>
<td>94.40</td>
<td>89.60</td>
</tr>
<tr>
<td>GAL3 + TPO + CK19</td>
<td>98.67</td>
<td>70.00</td>
<td>83.15</td>
<td>97.22</td>
<td>87.20</td>
</tr>
<tr>
<td><strong>Two sequential markers</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GAL3 + HBME-1*</td>
<td>97.33</td>
<td>95.74</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>GAL3 + CK19*</td>
<td>98.67</td>
<td>89.36</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>GAL3 + TPO*</td>
<td>96.00</td>
<td>85.11</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
</tbody>
</table>

SN, sensitivity; SP, specificity; PV+, positive predictive value; PV−, negative predictive value; AC, accuracy; – not computed.

*Marker of second instance. The highest results are shaded.

### Table 3 Discrimination between oncocytic and non-oncocytic thyroid FNAB indeterminate lesions by single marker or by simultaneous/sequential combinations of markers.

<table>
<thead>
<tr>
<th></th>
<th>SN (%)</th>
<th>SP (%)</th>
<th>PV+ (%)</th>
<th>PV− (%)</th>
<th>AC (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Single marker</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GAL3</td>
<td>88.00</td>
<td>100.00</td>
<td>100.00</td>
<td>77.78</td>
<td>91.67</td>
</tr>
<tr>
<td>HBME-1</td>
<td>64.71</td>
<td>100.00</td>
<td>100.00</td>
<td>53.85</td>
<td>75.00</td>
</tr>
<tr>
<td>TPO</td>
<td>64.71</td>
<td>85.71</td>
<td>91.67</td>
<td>50.00</td>
<td>70.83</td>
</tr>
<tr>
<td>CK19</td>
<td>35.29</td>
<td>85.71</td>
<td>85.71</td>
<td>35.29</td>
<td>50.00</td>
</tr>
<tr>
<td>KS</td>
<td>23.53</td>
<td>95.71</td>
<td>92.00</td>
<td>35.29</td>
<td>50.00</td>
</tr>
<tr>
<td><strong>Two simultaneous markers</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GAL3 + HBME-1</td>
<td>97.33</td>
<td>90.00</td>
<td>93.59</td>
<td>95.74</td>
<td>94.40</td>
</tr>
<tr>
<td>GAL3 + CK19</td>
<td>98.67</td>
<td>84.00</td>
<td>90.24</td>
<td>97.67</td>
<td>92.80</td>
</tr>
<tr>
<td>HBME-1 + CK19</td>
<td>92.00</td>
<td>88.00</td>
<td>92.00</td>
<td>88.00</td>
<td>90.40</td>
</tr>
<tr>
<td>GAL3 + TPO</td>
<td>96.00</td>
<td>80.00</td>
<td>87.80</td>
<td>93.02</td>
<td>89.60</td>
</tr>
<tr>
<td>HBME-1 + TPO</td>
<td>90.67</td>
<td>82.00</td>
<td>88.31</td>
<td>85.42</td>
<td>87.20</td>
</tr>
<tr>
<td><strong>Two sequential markers</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GAL3 + HBME-1*</td>
<td>97.33</td>
<td>95.74</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>GAL3 + CK19*</td>
<td>98.67</td>
<td>89.36</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>GAL3 + TPO*</td>
<td>96.00</td>
<td>85.11</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
</tbody>
</table>

SN, sensitivity; SP, specificity; PV+, positive predictive value; PV−, negative predictive value; AC, accuracy; – not computed.

*Marker of second instance. The highest results are shaded.
Simultaneous and sequential associations of markers

The highest diagnostic accuracy of the two-marker simultaneous combinations in discriminating carcinomas from adenomas on FNAB cytological samples was obtained by GAL3 plus HBME-1 (94.40%) and GAL3 plus CK19 (92.80%). The simultaneous association of GAL3 with HBME-1 showed a higher specificity (90.00%) and a lower sensitivity (97.33%) than that of GAL3 with CK19 (Table 2).

The best three-marker simultaneous combination was GAL3, HBME-1 and CK19, which had a sensitivity of 100%, but a relatively low specificity (82.00%) (Table 2).

The sequential associations of HBME-1 or CK19 to GAL3 showed a higher combined sensitivity (97.33% and 98.67% respectively) and specificity (95.74% and 89.36% respectively) than the other sequential associations evaluated (Table 2). GAL3/HBME-1 and GAL3/CK19 sequential associations were superior to the corresponding simultaneous combinations in terms of specificity (95.74% and 89.36% vs 90.00% and 84.00% respectively), and identical in terms of sensitivity (Table 2).

On FNAB cytological specimens of oncocytic cell neoplasms, the simultaneous combination of GAL3 with CK19 showed 100% of both sensitivity and specificity, strictly followed by that of GAL3 with TPO. The same values of sensitivity and specificity were obtained by a sequential coupling of the two markers (Table 3).

In non-oncocytic cytological FNAB samples, the highest values of sensitivity, specificity and accuracy of the two-marker simultaneous combinations were obtained by GAL3 with HBME-1 (98.28%, 90.70% and 95.05% respectively) (Table 3). In non-oncocytic tumours, the GAL3/HBME-1 sequential association was superior to its simultaneous combination in terms of specificity (97.50% vs 90.70%) and identical in terms of sensitivity (Table 3).

Discussion

In this study, the expression of selected markers (GAL3, HBME-1, TPO, CK19 and KS) was assessed in a relatively large series of FNAB cytological and surgical thyroid samples using commercial species-specific monoclonal antibodies and a biotin-free detection system, in order to avoid interfering technical factors (e.g. aspecific antigen cross-reactivity of non-species-specific antibodies and endogenous biotin-like activity) (Herrmann et al. 2002, Bartolazzi et al. 2003, Saggiorato et al. 2004) which may be responsible for most of the reported discrepancies, particularly in the cases of GAL3 and HBME-1 (Nascimento et al. 2001, Mai et al. 2002, Mehrotra et al. 2004).

A first matter was the selection of markers. According to the reports in the literature, GAL3, HBME-1 and CK19 proved the most reliable ones, due to their reproducibility and also to the availability of commercial antibodies. All three selected markers (GAL3, HBME-1 and CK19) were claimed to be useful for both follicular and papillary carcinomas, although the latter are generally more strongly reactive for such molecules. To expand the search for an optimal panel of markers, we also introduced two novel antibodies, claimed to be able to distinguish benign from malignant differentiated follicular tumours: anti-TPO, which was actually tested more than 10 years ago (De Micco et al. 1991), but became commercially available only recently, and anti-KS, which is reported mainly expressed in papillary carcinomas.

A second important point was to set cut-off values for each marker, being their expression not subjected to the all-or-none law. We have thus defined the most appropriate cut-off in an attempt to reduce false negative results (that means to ensure thyroidectomy to all patients with carcinomas) as much as possible, to the detriment of a slight increase of false positive diagnoses. According to the aforementioned criteria, the best compromise between sensitivity and specificity was reached setting at 10% the cut-off value for GAL3, HBME-1, CK19 and KS expression, and at 50% that for TPO. In particular, the cut-off values (HBME-1, CK19 and TPO) here assigned were slightly more restrictive than those reported by some other authors (van Hoeven et al. 1998, Christensen et al. 2000, Raffaelli et al. 2001, Khurana et al. 2003). We suggest that these discrepancies may depend on the use of a biotin-free detection system, which is potentially able to reduce false stainings.

Based on the selected cut-offs for each marker, it was clear from our series that GAL3 was superior to other markers in identifying the majority of malignant tumours.

HBME-1 and CK19 were here confirmed to be valuable markers, while the novel KS was acceptable for papillary carcinoma cases only. Intact TPO was under-expressed in 80% of malignant cases, but its association with GAL3 did not considerably improve the diagnostic accuracy.

The results of our statistical analysis have indicated that the use of three or more markers at a time does not increase the FNAB diagnostic accuracy up to 100%, and is thus useless, especially at a time of money restrictions and scant resources. In practical terms,
we observed that just two markers (one being GAL3 and the other HBME-1 or CK19) are able to provide a good diagnostic accuracy (>90%) with a sensitivity as high as 97% and a specificity of 90%.

The last point was how to plan the immunohistochemical panel analysis, keeping in mind the cost-benefit problems and the turn-around time of pathology reports. Hence, we wanted to check whether there was any benefit from the simultaneous (two or more markers simultaneously tested in 1 day) or sequential (one marker first, followed by a second marker in the case of a first marker-negative result only) analysis of the selected molecules.

We have here shown that sequential analysis of GAL3 and HBME-1 or GAL3 and CK19 gave the same sensitivity (approximately 97%) as their simultaneous combinations, and a higher specificity (approximately 90%) than the simultaneous one, reducing the costs due the immunohistochemical procedures. Finally, we compared the accuracy of the different marker combinations in identifying malignant tumours of the oncocytic type vs the conventional type. Interestingly, we observed a marked difference in the ability to recognize oncocytic carcinomas when GAL3 was sequentially associated with CK19 (reaching 100% of both sensitivity and specificity), rather than HBME-1. Conversely, the latter proved more valuable in identifying follicular (and papillary) carcinomas of non-oncocytic types (98.28 of sensitivity and 97.50 of specificity). This finding is in agreement with the data in the literature on the reduced expression of HBME-1 by oxyphilic thyroid tumours (Mai et al. 2002, Volante et al. 2004a), and indicates that the selection of immunohistochemical markers for FNABs should be driven by the morphological features of cytological material (e.g. HBME-1 should be avoided in the presence of mitochondrion-rich oncocytic cells, as its sensitivity for such malignant tumours is as low as 35%).

Taken together, our results have provided useful suggestions for the diagnostic management of patients bearing thyroid nodules with an indeterminate FNAB result. In our opinion, patients with FNAB cytology of ‘follicular neoplasm’ and a positive GAL3 cytological test should be referred directly to the surgeon, without any other immunohistochemical evaluation and regardless of the clinical and ultrasound features. In the case of a negative GAL3 test, we propose the evaluation by a second marker represented by HBME-1 for the non-oncocytic follicular neoplasms or CK19 for the oncocytic cell tumours. The negative result of the second marker (HBME-1 or CK19) should encourage the clinician to use a follow-up programme, including a further FNAB 6–12 months later (possibly with immunohistochemical re-evaluation), and referring the patient to the surgeon when an increase in the nodule size or a positive result in a sequential test is observed. At worst, a diagnosis of thyroid malignancy will be just slightly postponed. Conversely, in the case of a positive result with the second marker the patient should be referred to the surgeon for thyroidectomy (Fig. 2). Needless to say patients with an indeterminate

Figure 2 Proposal of practical management of an indeterminate thyroid FNAB result ('follicular neoplasm'), based on two markers sequentially combined (GAL3 + CK19 or GAL3 + HBME-1).
FNAB result, but with clinical signs strongly suggestive of malignancy (e.g. hard and fixed thyroid nodule, lymphadenopathies, ultrasonographic features suggestive of malignancy) should be treated accordingly and referred to the surgeon, irrespective of the FNAB cytology results.

Hence, we propose a diagnostic algorithm, of low cost and easy interpretation, able to optimize the management of patients with thyroid nodular disease, reducing the number of unnecessary thyroidectomies for benign neoplasms.

Acknowledgements

We wish to thank Dr Carlotta Sacerdote (Unit of Cancer Epidemiology, University of Turin) for her invaluable help in the statistical analysis of these data, Dr Nicoletta Bergero for technical assistance and Dr Armando Bartolazzi (Department of Pathology, Saint Andrea Hospital, Rome) for helpful discussion.

Funding

This research was supported by grants from the Regione Piemonte (D.D. n.63 of May 27 2002, D.G.R. n.69-8040 of December 16 2002, D.D. n.104 of July 21 2003 and D.D. n.173 of October 30 2003), Turin, and from the Italian Ministry of University and Research, Rome, and from the Fondazione Inter-nazionale di Ricerca in Medicina Sperimentale, Turin, Italy. R D P is a recipient of a research fellowship from the Fondazione Speciale Oncologia, Turin, Italy. The authors declare that there is no conflict of interest that would prejudice the impartiality of this scientific work.

References


Baloch ZW, Fleisher S, LiVolsi VA & Gupta PK 2002 Diagnosis of ‘follicular neoplasm’: a gray zone in thyroid fine-needle aspiration cytology. Diagnostic Cytopathology 26 41–44.


Herrmann ME, LiVolsi VA, Pasha TL, Roberts SA, Wojcik EM & Baloch ZW 2002 Immunohistochemical expression of galectin-3 in benign and malignant thyroid lesions. Archives of Pathology and Laboratory Medicine 126 710–713.
Lloyd RV 2001 Distinguishing benign from malignant thyroid lesions: galectin 3 as the latest candidate. Endocrine Pathology 12 255–257.
Mai KT, Bokhary R, Yazdi HM, Thomas J & Commons AS 2002 Reduced HBME-1 immunoreactivity of papillary thyroid carcinoma and papillary thyroid carcinoma-related neoplastic lesions with Hurthle cell and/or apocrine-like changes. Histopathology 40 133–142.


Tangpricha V, Chen BJ, Swan NC, Sweeney AT, de las Morenas A & Safer JD 2001 Twenty-one-gauge needles provide more cellular samples than twenty-five-gauge needles in fine-needle aspiration biopsy of the thyroid but may not provide increased diagnostic accuracy. Thyroid 11 973–976.


