Clinical and morphological methods to improve the diagnostic accuracy of thyroid nodular goiter

Theses of a Ph D Dissertation

Tamas Solymosi

Ph D Course of Semmelweis University, Clinical Doctoral School, Program of clinical endocrinology

Supervisor:
Prof. Dr. Istvan Szabolcs

Országos Gyógyintézeti Központ
Belgyógyászati Osztály, Endokrinológia Szakprofil
INTRODUCTION

Factors determining the efficacy of the evaluation process of nodular goiter were investigated. We analyzed the diagnostic potential of markers of proliferation, including DNA cytometry, the AgNOR method, and digital morphometry. The impact of ultrasonography (US) on the diagnostic accuracy of fine needle aspiration cytology (FNAC) was also determined. Ethnographically homogenous population was studied. The majority of patient live in mildly iodine deficient region. FNAC and US were performed by the same person. The diagnostic power of FNAC was determined by reviewing 45 publications on the diagnostic role of FNAC of the thyroid.

We performed the first prospective study in the literature in order to analyze and compare the diagnostic power of US-guided FNAC in two large groups of thyroid patients, with different iodine intakes, but from an ethnographically homogenous population. A comparative study was performed in two different periods whether the US would increase the diagnostic power of FNAC or not.

We applied DNA image cytometry, AgNOR method and digital morphometry on FNAC smears in order to determine their potential to improve the diagnostic accuracy of FNAC. The first phase study was performed on three consecutive series of nodular goiters, adenomas and carcinomas of
follicular cell origin. Thereafter a second phase investigation was made with AgNOR technique on those patients operated on where the FNAC was suspicious. The differential diagnostic potential of various cytological parameters were also analyzed. Finally, we performed a prospective study in order to determine the role of a new decision strategy in two groups of patients with difficult-to-interpret FNAC report, i.e. in patients with follicular proliferation without significant atypia, and in those where the FNAC in unable to differentiate epithelial atypia caused by lymphocytic thyroiditis and/or thyroid tumor. We studied whether ultrasonographic follow-up instead of immediate surgery would be a safe and effective method in order to decrease unnecessary surgical procedures.

**OBJECTIVES**

A/ To determine the diagnostic power of thyroid FNAC on the basis of metaanalysis of 45 publications.

B/ To determine to role of iodine intake on the diagnostic accuracy of FNAC.

C/ To analyze the differential diagnostic potential of the most important parameters of thyroid cytology, i.e. nuclear size and shape, paillardization, nuclear grooves and inclusions.
D/ To investigate the role of AgNOR method and digital morphometry in differentiating nodular goiters, adenomas and carcinomas (1st phase study). To determine the role of these techniques in those cases where the FNAC report was suspicious (2nd phase study).

E/ To study the role of DNA image cytometry in discrimination of Hashimoto’s thyroiditis and MALT- and non-MALT lymphomas.

F/ To determine the impact of US on the diagnostic accuracy of FNAC. To determine the role of a new decision strategy based on ultrasonographic follow-up in two groups of patients, i.e. in those with follicular proliferation without significant atypia and in those where the differential diagnostic problem is caused by Hashimoto’s thyroiditis.

PATIENTS AND METHODS

A/ Metaanalysis of 45 studies in the literature focusing on cyto-histological comparison.

B/ To compare the efficacy of FNAC on patients living in mildly iodine deficient region (ID) and on those living in iodine sufficient area (IS). The IS and ID groups were characterized with the followings. median of urinary iodine excretion (103 and 75 µg/l, respectively), the iodine content
in the drinking water (31.0 ± 21.3 vs. 10.3 ± 6.34 µg/l, p = 0.04, Student-test). US and FNAC was performed on all nodular goiter patients. Cyto-histological comparison was made and compared in ID and IS group.

C/ The occurrence and type of papillarization was determined on 11510 smears. The occurrence of nuclear grooves and inclusions was detected on cytological smears of 861 patients operated on. Papillarization with honeycomb pattern in colloid background was defined as benign when the occurrence of nuclear grooves and inclusions as below 1:1000, while irregular papillarization with occurrence of inclusions and grooves above 5:1000 was defined as malignant. The remaining cases were named indeterminate.

The size (µ²) and shape (the proportion of the great and small diameter) of the nucleus was determined on 27 carcinomas, 26 microfollicular adenomas, 10 Hürthle-cell adenomas, 6 Hashimoto’s thyroiditis, 27 nodular goiters (6 with, 21 without previous thyrostatic treatment) operated on consecutively.

D/ In the first phase study the AgNOR method and digital morphometry was applied to intraoperative imprint smears of three series of patients (nodular goiters, adenomas and carcinomas of follicular origin) operated on consecutively. In the 2nd phase study these methods were applied on FNAC
smears of 52 consecutive patients later underwent on surgery because of suspicious FNAC report.

D/ A comparative DNA image cytometry study was performed in 25 Hashimoto’s thyroiditis, in 5 histologically verified thyroid MALT-lymphomas and in 8 B-cell type non-Hodgkin lymphoma bone marrow smears.

E/ In the first study we compared the results of two periods. In Period 1 (between 01.01.1993 and 12.31.1994) we did not performed US and patients with cellular follicular lesions were categorized as having suspicious FNAC finding and were regularly sent for surgery. In Period 2 (between 01.01.1997 and 12.31.1998) all of the patients underwent thyroid US. US-guided FNAC was performed on non-palpable or difficult to palpate nodules and on those with mixed echopattern. More than 2/3 of patients with cellular follicular lesions was categorized as benign follicular proliferation. Another study was performed between 01.01.1997 and 12.31.2000. on 84 patients where the FNAC was not able to disclose the possibility of thyroid tumor beside lymphocytic thyroiditis, but based on previous experiences the chance of malignancy was held less than 20%. We proposed ultrasonographic follow-up for these patients, and surgery was advised only for those, whose
nodule increased in size more than 33% on any follow-up occasion.

RESULTS

A/ Surgical exploration was performed in every fifth patient who underwent. The most important statistical parameters of FNAC are as follows: sensitivity 91.6%, specificity 73.2%, diagnostic accuracy 78.5%, positive predictive value 58.1%. The introduction of FNAC in the evaluation of thyroid nodules caused a clear-cut disadvantage for a 3.2% of the patients investigated, and a clear-cut advantage for 4.4% of the patients.

B/ Elevated rates of multinodularity (59.6% vs. 49.6%, P<0.001), nonpalpable nodules (47.6% vs. 37.3%, P<0.001) and nondiagnostic US-FNACs (8.8% vs. 5.1%, P=0.008) and a lower malignancy rate (1.2% vs. 2.3%, P=0.006) were found in the ID area. Cytohistological comparison (ID N=416, IS N=97) revealed that the sensitivity, specificity, and diagnostic accuracy of US-FNAC were similar in the two areas (95.5% vs. 92.3%, 78.3% vs. 71.1%, 82.4% vs. 80.6%, IS vs. ID area, respectively), while a lower malignancy rate and a higher ratio of benign to malignant tumors were observed in the ID than in the IS area (10.1% vs. 22.6%, P<0.001, and 4.38 vs. 1.50, P<0.001, respectively). This
resulted in a lower positive predictive value of US-FNAC in the ID (36/106) than in the IS area (21/36, P=0.001), because the rate of false US-FNAC was higher in benign (67/184) than in malignant tumors (4/61, P<0.001).

The ratio of papillary to follicular cancer was significantly higher in ID compared with IS region (7:28 vs. 0:19, P<0.001).

C/ Determination of nuclear size and nuclear shape has limited value in the cytological differential diagnostics, because there is a great overlap between benign and malignant cases. Nuclear atypia characteristic for papillary cancer is a powerful tool in the diagnosis. If the occurrence of nuclear grooves and/or that of nuclear inclusions exceeds 5/1000, the probability of papillary cancer is greater than 95%. On the other hand in oxyphilic lesions even these nuclear signs have limited value, because more than 20% of Hashimoto’s thyroiditis and oxyphilic adenoma are characterized with an occurrence above the limit.

Papillarization is not a sign of malignancy, but papillarization combined with nuclear atypia are useful signs.

D/ First phase study. The mean AgNOR count per cell, the mean AgNOR area per cell and the mean standard deviation of AgNOR area per cell (mSDAa) were significantly higher in malignant than in benign lesions, and as regards the latter
two parameters, there were only two and one cases, respectively, of overlap. A cut off mSDAa of 1.9 µm² correctly categorized all but one of the atypical adenoma cases. A strong correlation was observed between the mean AgNOR and mean nuclear area (r = 0.88). For the mean nuclear area a greater overlap was found between benign and malignant groups.

In the second phase study the above mentioned AgNOR parameters except for mean AgNOR count per cell was also significantly higher in malignant than in benign cases, but a much greater overlap was found between the malignant and benign cases. As regards the mSDAa more than 30% of benign cases exceeded the cut off value of 1.9 µm². The most important practical consequence of these studies is that we did not found any carcinomas of follicular cell origin with a mean AgNOR area per cell below 5 µm². Nevertheless, in more than 50% of lymphocytic thyroiditis cases (either Hashimoto’s thyroiditis or focal form) the mean AgNOR area per cell was above 5 µm².

E/ DNA image cytometry is unable to discriminate between Hashimoto’s thyroiditis and MALT-lymphoma. On the other hand MALT- and non-MALT B-cell lymphomas could be distinguished without overlap based on 5cER and 2cDI.
The number of FNAC, the positive predictive value of FNAC, the number of malignancies, the sensitivity, the specificity and the diagnostic accuracy was higher in Period 2 compared with Period 1; 3446 vs. 1448, 45% vs. 16%, 39 vs. 29, 92% vs. 76%, 65% vs. 87% (\(\chi^2\) = 30.6, P< 0.001), and 66% vs. 87% (\(\chi^2\) = 44.7, P< 0.001), respectively. The number of unnecessary operation decreased substantially, while the risk of malignancy was only 1/59 in this group of patients. There was a greater increase in the number of FNACs (2.38x) than in the number of malignancies diagnosed (1.95x) in Period 2.

The new decision strategy based on US follow-up instead of immediate surgery seems a safe and effective method. 29 of 84 patients entered the study went on surgery. In two of these cases surgery disclosed papillary microcancer while in a third case papillary cancer with a maximal diameter of 13 mm. The histopathological diagnosis in the remaining 26 cases was benign. In more than 50% of the 84 cases follow-up examination disclosed that the suspicious lesion was in fact not a nodule but a focal lesion.

CONCLUSIONS, NEW SCIENTIFIC CONSIDERATIONS
1. The drawbacks of the introduction of FNAC in the diagnostic algorithm are not negligible. The introduction of FNAC in the evaluation of thyroid nodules caused a clear-cut disadvantage for a relevant number (ranging from 3.2 to 6.7%) of the patients investigated, and a clear-cut advantage for 4.4% of the patients. Many patients who avoid surgery are suspected of receiving an advantage from FNAC, while some of them harbor a malignant tumor. Since we do not know who they are, the entire population must be regularly followed up. The statistical analysis of cytohistological comparisons published in the literature is not equivocal. In more than 40% of publications the authors analyzed their results by misusing statistical methods. This resulted in all of these cases in better than the true statistical parameters.

2. The well-known cytological signs (size and shape of the nucleus, papillarization) are not suitable for discrimination between benign and malignant thyroid lesions. Counting of nuclear grooves and inclusions characteristic for papillary cancer is a good tool in disclosing this cancer type. On the other hand in the case of an oxyphilic lesion, even these signs have limited value.

3. Ours was the first publication in the literature focusing on the AgNOR method applied to thyroid cytological smears. Many parameters determined with this method are able to
differentiate between benign and malignant thyroid lesions. AgNOR method is superior to digital morphometry in the differential diagnostic. The most important parameters are mean AgNOR area and mean SD of AgNOR area per cell. There is no publication in the literature focusing on suspicious lesions. From a practical point-of-view this is more important, than to analyze a new tool in consecutively operated patients. We performed such an investigation termed 2nd phase study with AgNOR method. Similarly to 1st phase study malignant lesions were characterized with significantly greater proliferative activity, but the overlap between benign and malignant lesions was greater. A cut off value of 1.9 µm$^2$ correctly categorized all but one carcinoma, but more than 30% of benign cases fell in the malignant range. The mean AgNOR area was greater than 5 µm$^2$ in all cancer of follicular origin, and was less than 5 µm$^2$ in 905 of adenomas. On the other hand in more than 50% of lymphocytic thyroiditis cases exceeded this cut-off value. So AgNOR method seems a valuable new tool in distinguishing benign and malignant lesions even in the special group of patients, whose FNAC report is suspicious. If the mean AgNOR area of a suspicious lesion is below 5 µm$^2$, the possibility of malignancy is very low.
4. The introduction of US permits a higher diagnostic accuracy of the evaluation process of nodular goiter. This is achieved not only by lessening false negative reports but more importantly by decreasing false positive results. US permits to correctly diagnose difficult-to-palpate lesion, and is able to diagnose not only nonpalpable nodules but to prevent overdiagnose a surface irregularity not infrequently seen in nodular goiters and frequently seen in Hashimoto’s thyroiditis. Moreover US gives objective and reproducible data on the volume of a nodule, which is the prerequisite of the correct follow-up investigation. These are the basis for a new decision strategy in the case of patients with cellular follicular lesion, where an immediate surgery is in more than 95% of cases unnecessary.

5. In part of those cytological pictures which are characteristic of a follicular tumor, i.e. in those where the lesion is smaller than 2.5 cm, there is a normal diploid histogram we can offer instead of immediate surgery US follow-up. We first published our results in the literature in this field. This approach seems safe (the rate of malignancy is below 2%, and even in those cases there is no drawback of delay) and effective (more than 50% of patients are avoid of surgery after 5-year follow-up). A delay of a few years can be accepted in the diagnosis and therapy of the small proportion...
(less than 2%) of patients where the disease is a low-grade follicular carcinoma.

6. There has not been published any paper in the literature as yet focusing on the differential diagnostic problem caused by epithelial atypia in lymphocytic thyroiditis. Our new approach is not to operate these patients but to perform a close follow up in order to decrease unnecessary operations. Our results demonstrate that this is a safe method. This approach has particular interest in those cases where the second investigation discloses that the patient had no nodule but only a focal hypoechogenic lesion at first investigation.

7. Ours was the first prospective study to analyze and compare the diagnostic power of US-guided FNAC performed by one investigator in two large groups of thyroid patients, with different iodine intakes, but from an ethnographically homogenous population.

We have been demonstrated a significantly higher positive predictive value of US-FNAC in the IS than in the ID area because in the latter area the ratio of benign to malignant tumors is significantly higher than in the former, and the diagnostic accuracy is significantly worse in benign than in malignant thyroid tumors. Our results confirm the fact that in ID regions the ratio of follicular to papillary cancer is significantly higher than in IS area. These results may explain
the difference in the evaluation protocols of nodular goiter of the American and European approach.
PUBLICATIONS


6. Irányelvek a pajzsmirigyrák kivizsgálása, kezelése és követése során. Szerk Ésik O. Orvosi Hetilap 2001 (142), Suppl. 1, 2515-2551


ABSTRACTS AND ORAL PRESENTATIONS


BOOK CHAPTERS