10.1053.paor.2001.0291 available online at http://www.idealibrary.com on IDEAL

ARTICLE

Evaluation of Diagnostic Efficiency of Computerized Image Analysis Based Quantitative Nuclear Parameters in Papillary and Follicular Thyroid Tumors Using Paraffin-Embedded Tissue Sections

Nimisha GUPTA, Chitra SARKAR, Rajvir SINGH*, Asis Kumar KARAK

Department of Pathology and Biostatistics*, All India Institute of Medical Sciences, New Delhi

Computerized image analysis (IA) system has emerged in recent years as a very powerful tool for objective and reproducible quantification of histological features. It has shown considerable potential for diagnostic application in diverse histological situations. The objectives of the present study were to evaluate the discriminatory diagnostic efficiency of computerized image analysis based quantitative subvisual nuclear parameters in papillary and follicular neoplasms of thyroid. A total of 60 cases were studied. Forty-four cases belonged to training set and 16 cases belonged to a test set. A minimum of 100 nuclei was analyzed in each case using uniform 5 μm thick hematoxylin stained sections. The IA workstation comprised of an Olympus microscope, a 10 bit digital video camera, an image grabber card and a pentium 120 MHz computer. Optimas 5.2 software was utilized for data collection on 8 morphometric and 8 densitometric parameters. Multivariate stepwise discriminant statistical analysis of data was done with the help of BMDP statistical software release 7.0. Results from a training set revealed correct classification rates of 98.0%, 84.5% and 61.2% for the histological groups of hyperplastic papillae versus papillae of papillary carcinoma (group I), follicular variant of papillary carcinoma versus the broad category of follicular neoplasms consisting of both follicular adenoma and follicular carcinoma (group II) and follicular adenoma versus follicular carcinoma (group III), respectively. Results of test set revealed correct classification rates of 100%, 80% and 50% for groups I, II and III respectively. It was concluded that computerized nuclear IA parameters have potential usefulness for discriminating benign versus malignant papillary lesions of thyroid, follicular variant of papillary carcinoma versus follicular adenoma and/or follicular carcinoma but are of no value in discriminating between follicular adenoma and follicular carcinoma. (Pathology Oncology Research Vol 7, No 1, 46–55, 2001)

Keywords: thyroid tumor, image analysis, quantitative nuclear morphometry, papillary carcinoma, follicular neoplasm

Introduction

The means of establishing the correct tissue diagnosis of papillary and follicular neoplasms of thyroid is dependent on subjective recognition of an aggregate of histoarchitectural and cytologic features in hematoxylin-eosin stained tissue sections. Lack of objective markers to compensate for subjective bias, is the source of significant diagnostic discrepancy and inter-observer discordance.[18,48] The most problematic diagnostic areas appear to be the differentiation between follicular adenoma versus follicular carcinoma, follicular variant of papillary carcinoma versus follicular adenoma or follicular carcinoma and papillary structures of benign versus malignant origin.[18,35,45,48] Despite many studies[7,23,31,38-39,44,54,57-58] in search of diagnostically helpful immunohistochemical markers for primary thyroid tumors; the results have not been very encouraging so far except for medullary carcinoma.[52] The paucity of reliable objective markers is also evident at molecular level. Thus, the PTC oncogene, althongh specific for papillary carcinoma, has unreliable sensitivity[21,43,59] and ras oncogene activation is incapable of distinguishing between follicular adenoma and follicular carcinoma despite being a fair-

Received: 18 August, 2000; *accepted:* Dec 6, 2000

Correspondence: Asis Kumar KARAK, MD, PhD; Associate Professor, Department of Pathology, All India Institute of Medical Sciences, New Delhi – 110029, India, *Tel:* 91-11-6864851/Ext. 4376, *Fax:* 91-11-686 2663; E-mail: akkarak@hotmail.com / akk@medinst.ernet.in

This work has been supported partly by the research grant from the Indian Council of Medical Research, Ansari Nagar, New Delhi – 110029, India

 1219-4956/01/010046+10 $ 12.00/0

ly common and early event in follicular neoplasm[25,29,51] In recent years, computerized image analysis has emerged as a powerful tool for bias-free objective reproducible quantification of subvisual histologic features[3,24] and has already shown remarkable potential for diverse diagnostic applications.[4-6,11,13,15-17,30,37,55] In the specific context of differentiated thyroid neoplasms of follicular epithelial cell origin, although morphometric, stereologic and densitometric parameters in various combinations have been studied both in histologic sections and fine needle aspirated cytologic specimens, there is no general consensus in the results of the studies so far.[2,9,19,22,33,40-42,50,53] Moreover, image analysis based comprehensive studies in significant number of cases with multivariate discriminant analysis of data are very few,[2,53] encouraging the present investigation. The aim of the study was to evaluate the discriminatory diagnostic efficiency of computerized image analysis based quantitative subvisual nuclear parameters in papillary and follicular thyroid tumors.

Materials and methods

A total of 60 cases were studied retrospectively. Consecutively available cases from the histopathology files of All India Institute of Medical Sciences, New Delhi were entered in to the study without any selection bias. Histological material was reviewed jointly and concordant agreement was established for histological features in each case between two trained and experienced pathologists (AKK & CS). Any discordant sample was abandoned. In addition, for papillary carcinoma (both classical papillary and follicular variant), only cases with histologic or fine needle aspiration biopsy evidence of lymph node metastasis were included. This was done to dispel with confidence even the slightest possibility of misdiagnosis of florid papillary hyperplasia for conventional papillary carcinoma or follicular neoplasms (follicular adenoma/follicular carcinoma) or the follicular variant of papillary carcinoma. All cases of papillary hyperplasia had background histological changes of adenomatous goitre including macroscopic and clinical features of multinodular goitre without any follow up evidence of a neoplastic disease. For follicular carcinoma, apart from a prerequisite of unequivocal histologic evidence of vascular and/or capsular invasion, all cases showed either histological or cytological (FNA) or radioisotopic evidence of distant metastasis. For follicular adenoma, a minimum of 10 representative histological sections from the capsule of the lesion did not show any evidence suggestive of either vascular or capsular invasion and also none of the cases had any follow up evidence of distant metastasis.

In general, the cases included in the present study by means of above-mentioned procedure were considered to be best possible "gold standard" representatives of various diagnostic entities not only form the point of inter-observer concordance, but also the true biologic behavior of several neoplastic lesions. All subsequent data acquisition and analysis were done with reference to this diagnostic standard.

Table 1. **Lesion-wise distribution of 60 cases between the training and test set for image analysis**

Histologic Diagnosis (number of cases)	*Training Set (n = 44)*	*Test Set (n = 16)*
Papillary carcinoma (27)		
Classical	10	4
Follicular variant	9	4
Follicular neoplasms (26)		
Carcinoma	10	3
Adenoma	10	3
Papillary hyperplasia (7)	5	2

All cases were coded by` a technician to make provision for a "histologically blind" study situation for data collection without the actual knowledge of already assigned "gold standard" diagnosis for the cases. The cases were divided into training and test sets. Forty-four cases constituted the training set in order to find out the possible usefulness of image analytic parameters for diagnostic discrimination amongst histologic entities. The sixteen remaining cases were allocated to test set in order to check the validity of the discriminatory function of the image analytic outcome of the training set. The lesion-wise break up of the training and test set of cases of image analysis is shown in *Table 1*.

Staining protocol

One representative block of formaldehyde fixed paraffin-embedded tissue was selected from each case. Uniform 5 μm thick sections were cut for all the cases by the same technician using the same microtome. On completion of tissue section preparation for all 60 cases, single batch staining was performed to ensure uniform staining condition for each case. Sections were deparaffinized, hydrated and stained with Mayer's hematoxylin for 15 minutes at room temperature, washed in running tap water for 30 minutes, dehydrated and mounted with coverslips using DPX as mounting medium. No cytoplasmic counterstain was used.[36]

Hardware

The image analysis workstation included a trinocular microscope (Olympus BX 50, Olympus Corporation, Japan), a 10 bit digital video camera (Xilix Microimager, Xilix Technologies Corporation, Richmond, Canada), a F-64 Oculus image grabber card (Coreco Inc, Stlament, Ouebec, Canada), a 120 MHz pentium computer (Celebris XL, Digital Corporation, USA) equipped with 4 MB RAM, 1.5

GB hard disk, one 3.5" floppy diskette drive, a mouse, key board, a 17" high resolution color monitor and Hewlett Packard 4L Laserjet Printer and an UPS (Uninterrupted Power Supply) source.

Software

OPTIMAS (Optimas Corporation, USA) version 5.2 (October 1995) running in Microsoft Windows 95 graphics environment.

System calibration for data collection

Morphometric spatial measurements were calibrated in terms of micrometer with the help of a Nikon micrometer slide. Densitometric values were based on 0-65355 gray scale obtained by converting the images to 16-bit image format. Constancy of illumination was maintained to ensure reproducible and comparable densitometric values during different measurement sessions. The linearity of the system was checked using different filters with known optical densities. Glare reduction was achieved with Köhler illumination and Abbe condenser setting of numeric aperture to 0.8. Background decalibration with bright and dark field images was done to correct for the effect of illumination inhomogeneities.

Data collection

A total of 16 nuclear image analytic descriptors comprising of 8 morphometric (Area, Area Equivalent Diameter, Perimeter, Perimeter Equivalent Diameter, Major Axis Length, Breadth, Circularity, Rectangularity) and 8 densitometric (Gray Value, Integrated Gray Value, Integrated Log Inverse Gray Value, Gray Value Surface Area Density, Log Inverse Gray Value Surface Area Density, Gray Value Roughness, Log Inverse Gray Value Roughness, Fractal Texture) parameters were included in the thyroid measurement set of data collection program. Image acquisition was done using 40 X (dry) objective lens with numerical aperture of 0.65 and a focal length of 3 μm, using a 550 ± 10 nm filter for contrast enhancement. A total of at least 100 nuclei was analyzed in each case with sampling of 20-30 nuclei from five different areas of tissue section i.e. center and four corners of the section. Precautions were taken to include only intact whole nuclei representing the actual lesion, avoiding the nuclei of stromal cells. Overlapped and fragmented nuclei were disregarded. A careful accurate manual tracing of each nuclear boundary was done with the help of drawing tools, followed by data extraction. Finally, the data along with the measurement headings and case identification label were exported to linked Microsoft Excel spreadsheets for storage and subsequent analysis.

Statistical analysis of data

Descriptive statistics and multivariate discriminant analysis have been applied, wherever necessary. Stepwise discriminant analysis program belonged to BMDP Statistical Software release 7.0 (BMDP Statistical Software Inc., Los Angeles, CA).

Using the training set of cases (*Table 1*), an attempt was made to find out the best diagnostic class discriminatory power of multiple variables in combined strength. Composition of the variables in a particular set most suited to provide the best discrimination in a given situation of diagnostic demand, is arrived at by means of stepwise discriminatory analysis of each of 16 variables included in the present study.

Statistical analysis of training set of 44 cases involved pooling of nuclear image analytic data for a particular histological class. Thus, a total of 1012 nuclei (a minimum sampling of at least 100 nuclei x 10 cases) represented the histological class of classical papillary carcinoma of thyroid. Similarly, 904 nuclei of follicular variant of papillary carcinoma (9 cases), 1015 nuclei of follicular adenoma (10 cases), 1013 nuclei of follicular carcinoma (10 cases) and 517 nuclei of papillary hyperplasia (5 cases) represented the respective histologic classes.

Finally, the statistical attributes of discriminant classification functions derived from the training set of 44 cases were subjected to a validity test in terms of classification error rate when applied to 16 "unknown" cases belonging to the test set (*Table 1*).

Results

The analysis was purposedly oriented towards actually encountered diagnostic dilemmas in histopathology, namely papillary hyperplasia (PH) versus classical papillary carcinoma (CLPCT); follicular adenoma (FA) versus follicular carcinoma (FC); follicular variant of papillary carcinoma (FVPCT) versus combined follicular adenoma and follicular carcinoma (FN: Follicular Neoplasm group). In addition, the status in following situations was also examined: combined classical and follicular variant of papillary carcinoma (allPCT) versus FN group and allPCT group versus FN group versus PH.

A. Classification efficiency of multivariate functions

This is presented in *Table 2*. The highest degree (approximately 98%) of sensitivity and specificity was achieved for PH versus CLPCT. This was followed by approximately 85% sensitivity and specificity for FVPCT versus FN. The values did not change significantly in a situation of FVPCT and CLPCT cases together (allPCT) versus FN. The lowest degree of sensitivity and specificity

Table 2. **Efficiency of multivariate classification function of IA parameters**

Group	*Histologic diagnosis*	*% of correct classification*
PH	PH	97.6
vs	CLPCT	98.2
CLPCT	COMBINED	**97.9**
FA	FA	71.5
vs	FC	50.9
FC	COMBINED	**61.2**
FVPCT	FVPCT	86.7
vs	FN	83.5
FN	COMBINED	**85.1**
AllPCT	AllPCT	89.0
vs	FN	82.3
FN	COMBINED	**85.7**
AllPCT	AllPCT	86.3
vs	FN	82.0
FN	PH	77.3
vs	COMBINED	**81.9**
PH		

(50.9% and 71.5% for FC; 71.5% and 50.9% for FA) was observed for FA versus FC.

Finally, the results of discriminant analysis encompassing three lesions together (allPCT versus FN versus PH) showed a marked loss of sensitivity (20%) but not specificity for diagnosis of PH. Sensitivity and specificity for other two lesions of allPCT and FN were only marginally affected.

B. Composition of multivariate sets responsible for efficiency of classification functions

Step-wise multivariate discriminant analysis of sixteen IA parameters automatically eliminated the parameters without any statistically significant attribute of discriminatory classification function. At the end of the run, the analytic procedure generated a set of variables arranged in order of merit in terms of significant F values. Thus, each of the parameters within the set had significant discriminatory strength whereas the combined discriminatory strength of the entire set was maximum possible in a given situation. The relative composition of multivariate sets, however, varied depending upon the variety of situational demand (*Table 3*).

C. Distribution of canonical variable(s) of multivariate sets

Histograms (not shown) of canonical variable demonstrated adequate separation of mean coordinate values for CLPCT versus PH, FVPCT versus FN and allPCT versus FN. However, mean coordinates for FA versus FC were too close to each other, obliterating the possibility of any useful discriminatory function between follicular adenoma and follicular carcinoma by means of IA parameters.

Mean coordinates for canonical variables 1 and 2 for allPCT versus FN versus PH showed significant distance from each other. Plotting of canonical variables 1 and 2 for each individual case in the group, however, brought out a central zone of overlap.

D. Comparative evaluation of efficiency for morphometric versus densitometric versus combined morphometric-densitometric parameters

Results are presented in *Table 4.* It is observed that, in general, results are superior with combined parameters compared to use of either morphometric or densitometric subset alone. However, efficiency of densitometric subset almost equalled to that of combined parameters in FVPCT versus FN. Similarly, there was only marginal inferiority of densitometric variables to combined parameters in allPCT versus FN.

Test Set

The test set comprised of a total of 16 cases including 4 cases of CLPCT, 4 cases of FVPCT, 3 cases of FC, 3 cases of FA and 2 cases of PH. The cases were first allocated to form three test groups with differing demands of diagnostic discrimination. Each case within a particular test group was then subjected to calculation of discriminant value using the mathematical attributes of classification function of the appropriate multivariate subset generated by the training set.

Table 5 shows the composition of three test groups and the application of classification function of appropriate multivariate subset of training set. Mathematical attributes of the multivariate subsets are also presented.

The discriminant value for classifying a test case was calculated as follows:

Discriminant value for possibility A = [(mean of variable 1 of test case in reference to possibility A x mathematical attribute of variable 1 of classification function in reference to possibility A) + (mean of variable N of test case in reference to possibility A x mathematical attribute of variable N of classification function in reference to possibility A)] – CONSTANT value for possibility A.

Discriminant value for possibility B was calculated in a similar way as above by substituting all the values appropriately in terms of possibility B.

A test case with higher discriminant value for a particular histologic class was assigned diagnostic commitment to that class.

Table 3. **Statistical moments of multivariate sets derived from step-wise discriminant analysis of a total of 16 1A parameters (8 morphometric, 8 densitometric)**

A. PH versus CLPCT

Parameters	*Mean ± SD*		*F statistic*	*P value*
	PH	*CLPCT*		
1. peried	6.93 ± 1.35	9.20 ± 0.96	1252.64	< 0.0001
2. ar	30.95 ± 13.23	46.53 ± 10.11	1015.36	< 0.0001
3. ligvsad	3.25 ± 2.05	11.53 ± 9.62	917.15	< 0.0001
4. bre	5.72 ± 1.19	7.18 ± 0.95	801.58	< 0.0001
5. gv	340.17 ± 76.43	326.60 ± 86.51	672.90	< 0.0001
6. gvr	9465.59 ± 6099.22	17623.87 ± 11653.87	594.87	< 0.0001
7. ligvr	0.18 ± 0.28	0.27 ± 0.32	544.27	< 0.0001
8. maxle	7.13 ± 1.56	8.95 ± 1.06	479.33	< 0.0001
9. ared	6.15 ± 1.24	7.65 ± 0.82	428.96	< 0.0001
10. rec	0.73 ± 0.02	0.71 ± 0.03	407.49	< 0.0001
11. gvsad	194.01 ± 329.07	283.07 ± 366.27	371.82	< 0.0001

B. FA versus FC

	FA	*FC*		
1. ar	37.42 ± 7.14	39.49 ± 12.23	81.74	< 0.0001
2. ared	6.87 ± 0.64	7.01 ± 1.03	60.22	< 0.0001
3. igv	12642.09 ± 4428.55	13003.08 ± 6000.78	48.71	< 0.0001
4. gvsad	275.77 ± 381.11	262.76 ± 364.77	38.80	< 0.0001
5. ligvsad	32.74 ± 14.06	34.48 ± 14.09	32.68	< 0.0001

C. FVPCT versus FN

	FVPCT	*FN*		
1. ligvsad	12.10 ± 9.91	33.61 ± 14.10	1721.96	< 0.0001
2. igv	21266.54 ± 8319.73	12822.40 ± 5274.61	1199.51	< 0.0001
3. fractex	4.48 ± 1.78	4.46 ± 1.82	830.62	< 0.0001
4. bre	7.42 ± 1.07	6.59 ± 0.91	521.45	< 0.0001
5. rec	0.72 ± 0.03	0.72 ± 0.02	437.91	< 0.0001
6. ar	50.62 ± 11.55	38.45 ± 10.06	386.02	< 0.0001
7. iligv	25.73 ± 9.82	23.36 ± 8.67	340.60	< 0.0001
8. ared	7.97 ± 0.90	6.94 ± 0.86	303.66	< 0.0001

D. allPCT versus FN

	allPCT	*FN*		
1. ligvsad	11.80 ± 9.76	33.61 ± 14.10	3151.50	< 0.0001
2. peried	9.19 ± 0.97	7.96 ± 0.99	1979.84	< 0.0001
3. ar	48.46 ± 11.00	38.45 ± 11.06	1418.07	< 0.0001
4. gvr	19060.48 ± 11962.90	14839.14 ± 8626.36	1076.41	< 0.0001
5. igv	18250.15 ± 7958.30	12822.40 ± 5274.61	865.26	< 0.0001
6. gv	366.15 ± 97.85	326.80 ± 71.24	724.92	< 0.0001
7. ared	7.80 ± 0.88	6.94 ± 0.86	627.37	< 0.0001
8. cir	17.52 ± 1.63	16.54 ± 0.95	560.63	< 0.0001
9. gvsad	278.94 ± 365.93	269.28 ± 373.01	499.55	< 0.0001
10. iligv	27.42 ± 9.78	23.36 ± 8.67	454.23	< 0.0001
11. fractex	4.58 ± 1.75	4.46 ± 1.82	415.25	< 0.0001
12. rec	0.72 ± 0.03	0.72 ± 0.02	381.39	< 0.0001
13. bre	7.29 ± 1.01	6.59 ± 0.91	352.73	< 0.0001

E. allPCT versus FN versus PH

Parameters	*Mean ± SD* *allPCT*	*FN*	*PH*	*F statistic*	*P value*
1. ligvsad	11.80 ± 9.76	33.61 ± 14.10	3.25 ± 2.05	2279.62	< 0.0001
2. peried	9.19 ± 0.97	7.96 ± 0.99	6.93 ± 1.35	1622.52	< 0.0001
3. ar	48.46 ± 11.00	38.45 ± 10.06	30.95 ± 13.23	1232.78	< 0.0001
4. ared	7.80 ± 0.88	6.94 ± 0.86	6.15 ± 1.24	1014.95	< 0.0001
5. gvr	19060.48 ± 11962.90	14839.14 ± 8626.36	9465.59 ± 6099.22	837.34	< 0.0001
6. maxle	9.12 ± 1.11	7.93 ± 1.12	7.13 ± 1.56	714.85	< 0.0001
7. cir	17.52 ± 1.63	16.54 ± 0.95	16.03 ± 0.99	622.29	< 0.0001
8. ligvr	0.27 ± 0.33	0.25 ± 0.32	0.18 ± 0.28	547 91	< 0.0001
9. gv	366.15 ± 97.85	326.80 ± 71.24	340.17 ± 76.43	496.46	< 0.0001
10. igv	18250.15 ± 7958.30	12822.40 ± 5274.61	11134.77 ± 7068.21	455.99	< 0.0001
11. rec	0.72 ± 0.03	0.72 ± 0.02	0.73 ± 0.02	416.33	< 0.0001
12. iligv	27.42 ± 9.78	23.36 ± 8.67	16.76 ± 7.24	383.23	< 0.0001
13. gvsad	278.94 ± 365.93	269.28 ± 373.01	194.01 ± 329.07	355.08	< 0.0001

Abbreviations: *ar*, area; *ared*, area equivalent diameter; *bre*, breadth; *cir*, circularity; *fractex*, fractal texture; *gv*, mean gray value; *gvr*, gray value roughness; *gvsad*, gray value surface area density; *igv*, integrated gray value; *iligv*, integrated log inverse gray value; *ligvr*, log inverse gray value roughness; *ligvsad*, log inverse gray value surface area density; *maxle*, major axis, length; *peri*, perimeter; *peried*, perimeter equivalent diameter; *rec*, rectangularity.

6/6 (100%) cases of test group-I (PH vs CLPCT) were correctly classified. 8/10 (80%) cases of test group-II (FVPCT vs FN) were correctly classified. One case of each belonging to FVPCT and FN group were wrongly classified. 3/6 (50%) cases of test group-III (FA vs FC) were correctly classified. All three cases of follicular adenoma were wrongly classified as follicular carcinoma.

Discussion

The important findings of this study were absolute discrimination of papillary structures of benign versus malignant origin and significant discrimination between follicular variant of papillary carcinoma and the broad group of follicular neoplasms. However, image analysis was found to be not useful in distinguishing between follicular adenoma and follicular carcinoma.

The distinction between papillary structures of benign and malignant origin is a documented diagnostic problem.[15,34] Our study revealed a correct classification rate of 97.9% in the training set and 100% in the test set for papillae of hyperplastic origin versus papillae of papillary carcinoma. Not many image analysis studies have addressed this particular issue of diagnostic confusion. However, Holschbach et al[28] tested the potential application of three-dimensional reconstruction of tissue structures from serial tissue sections and reported better discrimination efficiency in above context. But the authors themselves admitted that three-dimensional reconstruction of biological tissue is time consuming and unsuitable for use in routine practice. Several other studies have attempted to distinguish benign versus malignant follicular neoplasms[12,20,47] but none of the authors have separately studied the benign versus malignant papillary lesions of thyroid.

Before the description of follicular variant of papillary carcinoma by Chem and Rosai[10] in 1977, older studies erroneously included many cases of follicular variant of papillary carcinoma as follicular carcinoma.[1] Further, the diagnostic problems between benign lesions of partly encapsulated hyperplastic nodules and pseudoinvasion after fine needle aspiration versus malignancies (namely the follicular variant of papillary carcinoma, follicular carcinoma and follicular variant of medullary carcinoma) is well recognized.[35] The training set of the present study revealed an overall correct classification rate of 85.1% for the histologic group of follicular variant of papillary carcinoma versus the broad histologic category of follicular neoplasms comprising of both follicular adenoma and follicular carcinoma. Comparable image analysis studies on this specific aspect are lacking in the literature. Artacho-Perula and colleagues,[2] in their recent report have shown good discrimination efficiency of papillary carcinoma in general versus follicular adenoma and carcinoma, but have not attempted to analyze their data in the specific diagnostic context of follicular variant of papillary carcinoma vis-à-vis follicular neoplasms. Analysis of our data on all papillary carcinomas irrespective of morphologic variants versus follicular neoplasms showed a correct classification rate of 89.0% for papillary carcinoma. This is similar to Artacho-Perula et al's[2] reported 85.0% classification efficiency for papillary carcinoma.

Table 4. **Comparative efficiency of classification functions of morphometric or densitometric subsets alone, vis-à-vis combined morphometric – densitometric parameters**

Groups	*% of correct classification using*		
	Morphometric Parameters (n = 8)	*Densitometric Parameters (n = 8)*	*All Parameters (n = 16)*
PH vs CLPCT	94.5	89.0	98.0
FA vs FC	58.5	55.5	61.2
FVPCT Vs FN	75.3	84.4	84.5
allPCT vs FN	78.5	83.7	85.6
allPCT vs FN vs PH	67.0	80.3	83.4

Accurate distinction of follicular adenoma from follicular carcinoma perhaps poses maximum difficulty in thyroid pathology, since the only available means are histologic demonstration of unequivocal vascular or capsular invasion. In an analysis of observer variations for histologic diagnosis of thyroid cancer, follicular carcinoma was the most common diverging diagnosis where the final diagnosis was a benign lesion.[18,48] Also, lack of cytologic parameters of distinction between follicular adenoma and follicular carcinoma is the main reason for failure of minimally invasive investigative procedure of fine needle aspiration (FNAC) to offer further diagnostic direction for follicular neoplasm group of lesions. The present study investigated the role of image analysis based sixteen nuclear parameters for distinguishing between follicular adenoma and follicular carcinoma. Results however showed that morphometric and densitometric parameters were inefficient for discriminating the two lesions. Correct classification rates of only 61.2% and 50.0% for the training set and test set respectively were observed. These results are in complete concordance with the recent study by Artacho-Perula et al[2] who reported poor discrimination efficiency between follicular adenoma and follicular carcinoma. Similarly, Schurmann and colleagues[49] in their flow cytometric study of Feulgen stained cell suspension observed that morphometric, sterologic and densitometric nuclear parameters are incapable of providing distinction between follicular adenoma and follicular carcinoma. However reports of few other studies showed contrary results claiming significant to total discrimination between follicular adenoma and carcinoma.[22,32,41,53] A detailed analysis of the study designs and mode of statistical analysis of results provided explanation for the apparently discordant observations of several investigations. Thus, although Kriete et al's[32] observations were based on statistically significant differences of measured value of parameters, they neither used multivariate discriminant analysis nor evaluated the discriminatory power of parameters in a test set of cases. Presence of some degree of differences in parameter values alone, even if statistically significant, may not be endowed with an efficient discriminatory power in a test situation. This is evident in the study by Salmon and colleagues[47] wherein, despite nuclear chromatin descriptors having significant differences between follicular adenoma and follicular carcinoma, the authors concluded that such differences are of limited value when applied to the diagnosis of individual cases. Results of studies by De Santis et al[14] and Galera-Davidson et al[22] indicated that morphometric and densitometric nuclear features might be helpful in differentiating between follicular adenoma and carcinoma. However, the number of cases in both of the studies were too small for any valid conclusion. De Santis et al[14] studied only 3 cases of each histologic category while Galera-Davidson et al's[22] observations were based on 6 cases of follicular adenoma and follicular carcinoma each. Results of a recent report by Tsybrovskyy and colleagues[53] however defy explanation in the light of our experience from the present study. The authors studied sixteen planimetric and densitometric nuclear features in 19 follicular adenomas, 12 minimally invasive and 3 widely invasive follicular carcinomas by means of a semi-automatic image analysis system using Feulgen stained paraffin sections. They reported 100% sensitivity and 94.7% specificity of the classification rule derived from multivariate discriminant analysis of their data.

For the purpose of close comparison with the present study in terms of case material, image analysis methodology and mode of statistical analysis, recently published work by Artacho-Perula et al[2] is most suitable. The authors analyzed quantitative nuclear parameters by morphometric and sterological methods in 55 cases comprising of papillary carcinoma, follicular adenoma and carcinoma, followed by stepwise discriminant analysis of data. Their results are highly comparable to those of the present study. The overall accuracy rate of discrimination was 75% with an efficiency of 85% for papillary carcinoma, compared to our corresponding figures of 85.7% and 89.0% respectively. The worst discrimination rate of 69.0% was seen between follicular adenoma and follicular carcinomas. This is in agreement with the corresponding value of 61.2% observed in the present study.

Table 5. **Compsotion of test groups versus application of multivariate classification function of training set**

Test group	*Histologic entities*	*Classification function of training set group*
I	CLPCT : 4	(a) PH versus CLPCT
	PH : 2	
II	FVPCT : 4	(b) FVPCT versus FN
	FN : 6	
III	FA : 3	(c) FA versus FC
	FC : 3	

(A) *PH versus CLPCT*

Variable	*PH*	*CLPCT*
ar	– 16.40011	– 17.33611
ared	– 2664.99023	– 2637.23926
peried	– 63.63652	– 54.64731
maxle	1211.82349	1200.37683
bre	1614.13037	1602.11292
rec	15662.03520	15532.86330
gv	0.06879	0.04620
gvr	– 0.00090	– 0.00068
gvsad	– 0.05838	– 0.05204
ligvr	76.00646	64.02271
ligvsad	– 0.38807	– 0.17286
Constant	– 6008.14111	– 5968.4113

(B) *FVPCT versus FN*

Variable	*FVPCT*	*FN*
ar	– 48.09598	– 47.67336
ared	548.92755	543.84650
bre	29.39716	30.50536
rec	1035.07434	1048.83655
igv	0.00168	0.00137
iligv	– 0.23407	– 0.27980
ligvsad	0.50384	0.61750
fractex	2.58729	2.41610
Constant	– 1481.63794	– 1475.71924

(C) *FA versus FC*

Variable	*FA*	*FC*
ar	– 53.96203	– 53.40494
ared	639.07996	633.36603
igv	– 0.00002	– 0.00012
gvsad	– 0.00277	– 0.00314
ligvsad	0.49531	0.50539
Constant	– 1194.81604	– 1175.50586

In general, we found superior results with combined morphometric and densitometric parameters compared to the use of either set of parameter alone. However, in the histologic groups of papillary carcinoma (follicular variant or combined classical and follicular variant) versus follicular neoplasms, densitometric parameters alone yielded results almost equivalent to combined morphometric and densitometric parameters. The independent diagnostic contribution of morphometric and densitometric parameters was demonstrated by factor and cluster analysis by Tsybrovskyy et al[53] in their image analysis study on follicular adenoma and follicular carcinoma of thyroid. This explains the superior classification efficiency of combined parameters noted in the present study. Failure of morphometric parameters to cause an additive improvement of densitometric results for correct classification of papillary carcinoma, as observed in this study, underscores the crucial diagnostic importance of densitometry related optical quality of haematoxylin stained nuclei of papillary carcinoma. This particular quality represents the so-called ''optically clear'' or ''ground glass'' nuclei in conventional histology.

Densitometric properties of cell nuclei are mostly studied for estimation of nuclear DNA content or ploidy using Feulgen stain because of its stoichiometric binding property. Although it has been claimed that hematoxylin and Feulgen staining of cell nuclei are correlated[56], subsequent studies have proved lack of correlation between the optical density of hematoxylin stained cell nuclei and nuclear DNA content.[17,26] Hematoxylin is, thus, unsuitable for DNA ploidy analysis. Nevertheless, the densitometric properties of hematoxylin stained nuclei could be exploited by image analysis for diagnostic discrimination in appropriate situations. Erler and colleagues[17] used hematoxylin and eosin stained histologic sections to evaluate twenty-two nuclear morphometric and densitometric parameters for discriminating benign and malignant hepatocytes. The authors reported positive and negative predictive values of 90.5% and 84.6% respectively for five densitometric parameters of hematoxylin stained nuclei. In the present study, evaluation of densitometric properties of hematoxylin stained nuclei was thought to be particularly pertinent in view of the crucial diagnostic importance accorded to ''optically clear'' or ''ground glass'' nuclei in histologic assessment of thyroid tumors.[8,27] Other nuclear features like grooving and pseudonuclei, although helpful and supportive, are in general considered to be less specific for diagnosis of papillary carcinoma.[1] However, the ground glass nuclear morphology observed in hematoxylin and eosin stained histologic sections, though accorded the highest diagnostic specificity for papillary carcinoma is alleged to be an artefact related to paraffin embedding independent of the type of fixative used and this feature is thus not seen in frozen sections or smears.[46] The phenom-

enon may actually reflect some intrinsic physiochemical alteration of the chromatin structure or associated nuclear proteins.[46] In the present study of nuclear parameters, cytoplasmic counterstain of eosin was omitted in order to minimize noise during image processing.

In order to obtain consistent and reproducible image analysis results, strict adherence to uniform standardized staining procedure and uniform thickness of tissue sections are of utmost importance. In this study, histologic sections were cut using the same microtome with fixed settings by the same technician and the entire study material was stained in single batch to avoid staining variability. Since hematoxylin binds with nuclear DNA non-stoichiometrically, certain degree of staining variability is anticipated. Self-evidently, this variability could be a source of error in terms of intra and inter-observer reproducibility of densitometric parameters. Morphometric measurements however, remain unaffected. Thus, although the results of the present study bring out the potentially useful discriminantory power of subvisual densitometric nuclear parameters, a standardized reproducible staining technique must be ensured before the technique could be useful in practice. Analogous to image analysis of Feulgen stained material for estimation of DNA ploidy, use of an internal or external tissue control along with appropriate correction factor for variability of staining density would presumably be successful for hematoxylin staining also. However, this was not tested in the present study. Mayer's hematoxylin instead of others was used since it is a progressive stain without the need for differentiation. Further, this stain is reported to produce most consistent staining results.[36]

References

1. *Ain KB:* Papillary thyroid carcinoma. Etiology, assessment and therapy. Endocrinol Metab Clin North Amer 24:711-760, 1995.
2. *Artacho-Perula E, Roldan-Villalobos R, Blanco-Garcia F, et al:* Objective differential classification of thyroid lesions by nuclear quantitative assessment. Histol. Hisopathol 12:425-431, 1997.
3. *Auer G, Askensten U, Ahrens O:* Cytophometry. Hum Pathol 20:518-527, 1989.
4. *Baak JP, Blanco AA, Kurver PH, et al:* Quantitation of borderline and malignant mucinous ovarian tumors. Histopathology 5:353-360, 1981.
5. *Baak JP, Kurver PH, Diegenbach PC, et al:* Discrimination of hyperplasia and carcinoma of endometrium by quantitative microscopy. A feasibility study. Histopathology 5: 61-68, 1981.
6. *Baak JP, Kurver PH, Overdiep SH, et al:* Quantitative, microscopical, computer-aided diagnosis of endometrial hyperplasia or carcinoma in individual patients. Histopathology 5:689-695, 1981.
7. *Cabaret V, Vilain MO, Delobelle-Deroide A, et al:* Immunohistochemical demonstration of ceruloplasmin and lactoferrin in a series of 59 thyroid tumors. Ann Pathol 12:347-352, 1992.
8. *Carcangiu ML, Zampi G, Pupi A, et al:* Papillary carcinoma of the thyroid: A clinicopathologic study of 241 cases treated at the University of Florence, Italy. Cancer 55:805-828, 1985.
9. *Cavallari V, Maiorana A, La Rosa GL, et al:* Morphometric studies on fine-needle aspirates from follicular proliferative lesions of the thyroid. Pathologica 81:441-446, 1989.
10. *Chem KT, Rosai J:* Follicular variant of thyroid papillary carcinoma. A clinicopathologic study of six cases. Am J Surg Pathol 1:123-130, 1977.
11. *Colgan TJ, Norris HJ, Foster W, et al:* Predicting the outcome of endometrial hyperplasia by quantitative analysis of nuclear features using linear discriminant function. Int J Gynecol Pathol 1:347-352, 1983.
12. *Collin F, Salmon I, Rahier I, et al.* Quantitative nuclear cell image analysis of thyroid tumors from archival material. Hum Pathol 22:191-196, 1991.
13. *Dawson AE, Austin RE Jr, Weinberg DS:* Nuclear grading of breast carcinoma by image analysis: Classification by multivariate and neural network analysis. Am J Clin Pathol 95:529-37, 1991.
14. *De Santis M, Sciarretta F, Sudano L, et al*: Morphometric evaluation of histologic sections of thyroid gland in benign and malignant follicular lesions. Diagn Cytopathol 3:60-67, 1987.
15. *Diamond DA, Berry SJ, Umbricht C, et al:* Computerized image analysis of nuclear shape as a prognostic factor for prostatic cancer. Prostate 3:321-332, 1982.
16. *Epstein JI, Berry SJ, Eggleston JC:* Nuclear roundness factor : A predictor of progression in untreated stage A2 prostate cancer. Cancer 54:1666-1671, 1984.
17. *Erler BS, Truong HM, Kim SS, et al:*A study of hepatocellular carcinoma using morphometric and densitometric image analysis. Am J Clin Pathol 100:151-157, 1993.
18. *Fassina AS, Montesco MC, Ninfo V, et al:* Histological evaluation of thyroid carcinomas : reproducibility of the ``WHO'' classification. Tumori 79:314-320, 1993.
19. *Ferrer-Roca O, Ballester Guardia E, Martin J:* Nuclear chromatin texture to differentiate follicular and papillary carcinoma of the thyroid. Pathol Res Pract 185:561-566, 1989.
20. *Ferrer-Roca O, Ballester-Guardia E, Martin-Rodriguez JA:* Morphometric, densitometric and flow cytometric criteria for the automated classification of thyroid lesions. Analyt Quant Cytol Histol 12:48-55, 1990.
21. *Fugazzola L, Pierotti MA, Vigano E, et al:* Molecular and biochemical analysis of RET/PTC4, a novel oncogenic rearrangement between Ret and Ele 1 genes in post-Chernobyl papillary thyroid cancer. Oncogene 13:1093-1097, 1996.
22. *Galera-Davidson H, Bibbo M, Bartels PH, et al:* Differential diagnosis between follicular adenoma and follicular carcinoma of the thyroid by marker features. Anal Quant Cytol Histol 8:195-200, 1986.
23. *Ghali VS, Jimenez EJS, Garch RL*: Distribution of Leu-7 antigen (HNK-1) in thyroid tumors: Its usefulness as a diagnostic marker for follicular and papillary carcinomas. Hum Pathol 23:21-25, 1992.
24. *Gil J:* Image analysis in anatomic pathology : what are the issues? Hum Pathol 20:203-204, 1989.
25. *Grebe SK, Hay ID:* Follicular thyroid cancer. Endocrinol Metab Clin North Amer 24:761-801, 1995.
26. *Gurley AM, Hidvegi DF, Bacus JW, et al:* Comparison of the Papanicolaou and Feulgen staining methods for DNA quantification by image analysis. Cytometry 11:468-474, 1990.
27. *Hedinger CE, Williams ED, Sobin LH:* Histologic Typing of Thyroid Tumors. Springer-Verlag, Berlin, 1988.
28. *Holschbach A, Kriete A, Schaffer R:* Differential diagnosis of papillary carcinomas of the thyroid using image analysis and three dimensional reconstruction from serial sections. Verh Dtsch Ges Pathol 74:270-274, 1990.

29.[2]*Jenkins RB, Hay ID, Herath JF, et al:* Frequent occurrence of cytogenetic abnormalities in sporadic non-medullary thyroid carcinoma. Cancer 66:1213-1220, 1990.
30.[2]*Koss LG, Bartels PH, Bibbo M, et al:* Computer discrimination between benign and malignant urothelial cells. Acta Cytol 19:378-391, 1975.
31.[2]*Kotani T, Asada Y, Aratake Y, et al.* Diagnostic usefulness of dipeptidyl aminopeptidase IV monoclonal antibody in paraffin embedded thyroid follicular tumors. J Pathol 168:41-45, 1992.
32.[2]*Kriete A, Romen W, Schäffer R, et al:* Computer analysis of chromatin arrangement and nuclear texture in follicular thyroid tumors. Histochemistry 78:227-230, 1983.
33.[2]*La Rosa GL, Cavallari V, Giuffrida D, et al:* The morphometric analysis of cell nuclei from fine needle aspirates of thyroid follicular lesions does not improve the diagnostic accuracy of traditional cytologic examination. J Endocrinol Invest 13:701-707, 1990.
34.[2]*Livolsi VA:* Surgical Pathology of the thyroid. W B Saunders, Philadelphia, 1990.
35.[2]*Livolsi VA, Asa SL:* The demise of follicular carcinoma of the thyroid gland. Thyroid 4:233-236, 1994.
36.[2]*Luna LG:* Manual of Histologic Staining Methods of the Armed Forces Institute of Pathology. McGraw Hill Publications, New York, 1968.
37.[2]*Marchevsky AM, Hauptman E, Gil J, et al:* Computerized interactive morphometry as an aid in the diagnosis of pleural effusions. Acta Cytol. 31:131-136, 1987.
38.[2]*McLaren KM, Cossar DW:* The immunohistochemical localization of S-100 in the diagnosis of papillary carcinoma of the thyroid. Hum Pathol 27:633-636, 1996.
39.[2]*Miettinen M, Franssila K, Lehto VP, et al:* Expression of intermediate filament proteins in thyroid gland and thyroid tumors. Lab Invest 50:262-270, 1984.
40.[2]*Montironi R, Alberti R, Sisti S, et al.* Discrimination between follicular adenoma and follicular carcinoma of the thyroid: preoperative validity of cytometry on aspiration smears. Appl Pathol 7:367-374, 1989.
41.[2]*Nafe R, Fritsch RS, Soudah B, et al:* Histomorphometry in paraffin sections of thyroid tumors. Pathol Res Pract 188:1042-1048, 1992.
42.[2]*Nagashima T, Suzuki M, Oshida M, et al:* Morphometry in cytologic evaluation of thyroid follicular lesions. Cancer 84:115-118, 1998.
43.[2]*Pasini B, Ceccherini I, Romeo G:* Ret mutations in human disease. Trends Genet 12:138-144, 1996.
44.[2]*Raphael SJ, McKeown-Eyssen G, Asa SL:* High molecular-weight cytokeratin and cytokeratin-19 in diagnosis of thyroid tumors. Mod Pathol 7:295-300, 1994.
45.[2]*Rosai J, Carcangiu ML:* Pitfalls in the diagnosis of thyroid neoplasms. Pathol Res Pract 182:169-179, 1987.
46.[2]*Rosai J, Carcangiu ML, Delellis RA:* Papillary carcinoma. In: Tumors of the thyroid gland. (Eds: Rosai J and Sobin LH), Armed Forces Institute of Pathology, 1992, pp. 65-121.
47.[2]*Salmon I, Kiss R, Franc B, et al:* Comparison of morphonuclear features in normal, benign and neoplastic thyroid tissue by digital cell image analysis. Analyt Quant. Cytol Histol. 14:47-54, 1992.
48.[2]*Saxen E, Franssilla K, Bjarnason O, et al:* Observer variation in histologic classification of thyroid cancer. Acta Path Microbiol Scand (A) 86:483-486, 1978.
49.[2]*Schüurmann G, Mattfeldt T, Feichter G, et al:* Differential diagnosis of follicular neoplasm of the thyroid. Langenbecks Arch Chir 375:95-101, 1990.
50.[2]*Schürmann G, Mallfeldt T, Feichter G, et al:* Stereology, flow cytometry and immunohistochemistry of follicular neoplasms of the thyroid gland. Hum Pathol 22:179-184, 1991.
51.[2]*Sozzi G, Miozzo M, Cariani TC, et al:* A t(2:3)(q12-13; p24-25) in follicular thyroid adenomas. Cancer Genet Cytogenet 64:38-41, 1992.
52.[2]*Stanta G, Carcangiu ML, Rosai J:* The biochemical and immunohistochemical profile of thyroid neoplasia. Pathol Annual 23 (Part I): 129-157, 1988.
53.[2]*Tsybrovskyy O, Vassilenko I, Mannweiler S, et al:* Multivariate karyometric approach in differential diagnosis of follicular thyroid neoplasms. A study of 31 cases. Virchows Arch 433:135-143, 1998.
54.[2]*Tuccari G, Barresi G:* Tissue polypeptide antigen in thyroid tumors of follicular cell origin: an immunohistochemical re-evaluation for diagnostic purposes. Histopathology 16:377-381, 1990.
55.[2]*Walts AE, Marchevsky AM:* Computerized interactive morphometry: An expert system for diagnosis of lymphoid-rich effusions. Am J Clin Pathol 92:765-772, 1989.
56.[2]*Wied GL, Bibbo M, Bartels PH:* Computer analysis of microscopic images: Application in cytopathology. Pathol Ann 16:367-409, 1981.
57. *Xu XC, el-Naggar AK, Lotan R :* Differential expression of galectin-1 and galectin-3 in thyroid tumors. Am J Pathol 147: 815-822, 1995
58.[2]*Yamshita H, Noguchi S, Murakami N, et al.* Immunohistological differentiation of benign thyroid follicular cell tumors from malignant ones: usefulness of anti-peroxidase and JT-95 antibodies. Acta Pathol Jpn 43:670-673, 1993.
59.[2]*Zou M, Shi Y, Farid NR:* Low rate of Ret proto-oncogene activation (PTC/ret TPC) in papillary carcinomas from Saudi Arabia. Cancer 73:176-180, 1994.