

SYNTACTIC STRUCTURE ANALYSIS OF BRONCHUS CARCINOMAS - FIRST RESULTS

Klaus Kayser, Bernd Kiefer, Hans Ulrich Burkhardt,
Margit Shaver

Institute of Pathology, University of Heidelberg,
Im Neuenheimer Feld 220/221, D-6900 Heidelberg, FRG.

ABSTRACT

Histopathology of 100 cases of operated lung carcinoma patients was analysed by syntactic structure analysis. Material consisted of HE-stained paraffin embedded sections of 20 cases of healthy lung parenchyma, and 20 cases of each of the four major cell types of bronchus carcinomas respectively (epidermoid, adeno, small cell, large cell carcinomas). Network formed by epithelial cells, mean distance of cell centers, distribution of distances according to neighborhood sequences were measured using programmes written in BASIC. O'Callaghan's (1975) neighborhood condition was used. Reclassification of 75 cases using discriminant analysis was performed with programmes from SPSS retrospectively, whereas 25 cases were classified prior to final diagnosis. Correct reclassification could be performed in 68/75 cases (87%) retrospectively and in 23/25 prospectively. All cases with healthy lung parenchyma could be separated correctly. Analysis of network pattern revealed similar textures of epidermoid carcinomas and of small cell anaplastic carcinomas in some cases, and of adeno carcinomas compared to large cell carcinomas in some cases respectively in accordance with similar difficulties obtained by pathologists.

Keywords: Syntactic structure analysis, bronchus carcinoma, diagnostic algorithm.

INTRODUCTION

Syntactic structure analysis offers new possibilities in supporting histopathological diagnosis. In contrast to quantitative morphometric measurements as performed by several authors (Baak, Oort 1983) syntactic structure analysis deals with structural properties of an image. In order to apply this method to histopathologic diagnosis different parameters free of choice have to be predefined:

- a. order of structures to be analysed
- b. neighborhood condition to be used

In a previous study Kayser, 1985 could show reliable support of syntactic structure analysis in difficult histopathologic diagnosis pleuritis carcinomatosa - mesothelioma by use of two different orders of structures and O'Callaghan's (1975) neighborhood condition. The present study was performed to evaluate diagnostic aids in differential diagnosis of primary lung carcinomas by use of second order structures only (geometrical arrangement of epithelial cells in tumorous tissue and healthy lung parenchyma).

MATERIAL AND METHODS

100 HE-stained specimens of healthy lung and primary lung carcinomas were prescreened and the diagnostic morphological area in the tumour was marked. Histomorphologic picture of this area was imaged onto a graphic tablet connected to a TEKTRONIX computer 4051 using a projection microscope. Coordinates of center of nuclei of epithelial tumor cells and lung parenchyma cells were marked interactively. Using programmes written in BASIC sequential order of neighbors of every epithelial cell was computed by use of O'Callaghan's neighborhood condition. The network formed by all neighbors was constructed and the number of vertices N_v , number of edges N_e and the cyclomatic number N_c were calculated. Because of obtained simple and connected graphs cyclomatic number was computed:

$$N_c = N_e - N_v + 1$$

In addition, mean distances and their distribution were calculated for nearest neighbors, second nearest neighbors etc. Based upon cyclomatic number and mean distance of nearest and second nearest neighbors reclassification was performed by use of discriminant analysis taken from SPSS programmes. The data set of 75 reclassified cases (including the misclassified cases) was used for classifying additional 25 cases without knowledge of final outcome. Reference area of specimens was set 0.04 qmm.

RESULTS

The reproducibility of the graph theory approach in syntactic structure analysis was tested in previous studies and was found to be $> 95\%$ in all parameters analysed. A typical graph of an adeno carcinoma of lung is shown in Fig.1. The mean and standard error of cyclomatic number and distance of nearest neighbors, second nearest neighbors and third nearest neighbors of 75 reclassified cases is given in table 1.

Table 1 Parameters of texture analysis according to cell type (mean and confidence limits, $p > 0.95$)

Cell type	cyclomatic number	distance of nearest neighbor (μ)	distance of 2nd nearest neighbor (μ)	distance of 3rd nearest neighbor (μ)
Healthy	25 \pm 2.5	13.0 \pm 0.8	19.7 \pm 1.2	26.4 \pm 1.8
Epidermoid	265 \pm 25	8.4 \pm 0.4	10.7 \pm 0.6	12.7 \pm 0.7
Adeno	114 \pm 15	9.7 \pm 0.6	13.1 \pm 0.7	16.5 \pm 0.8
Small cell	281 \pm 13	7.1 \pm 0.4	9.2 \pm 0.6	11.0 \pm 0.6
Large cell	137 \pm 16	11.3 \pm 0.5	15.0 \pm 0.7	18.1 \pm 0.9

2 groups of obtained parameters can be easily separated, i.e. healthy lung parenchyma from carcinomatous tissue. Linear regression of neighboring cells and cyclomatic number is shown in Fig.2. Curve of normal tissue is completely different compared to curves of carcinomatous tissue.

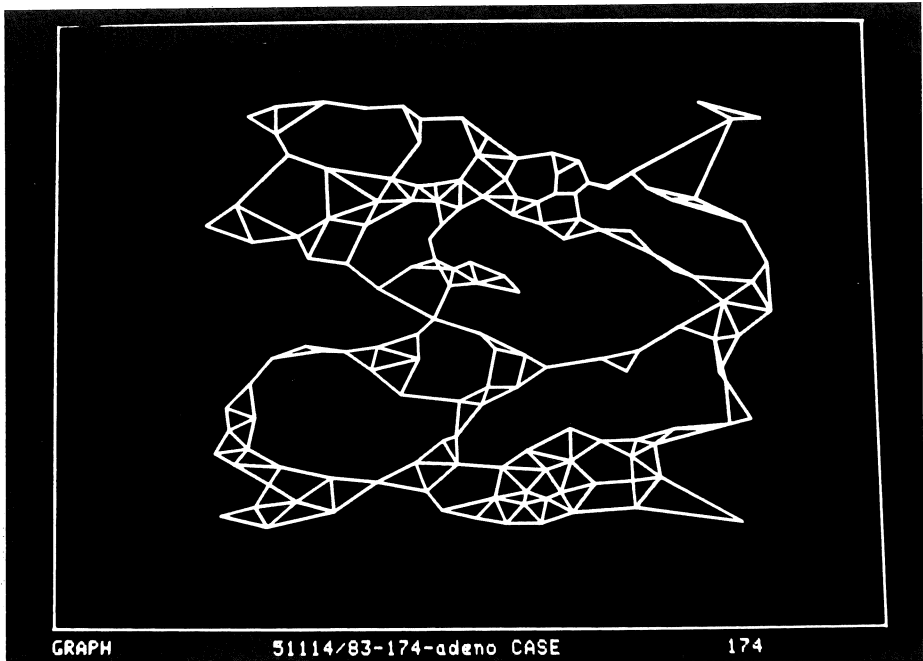


Fig. 1. Typical graph of adeno carcinoma of lung. Vertices correspond to centers of gravity of cells, edges (lines) correspond to neighboring cells.

Obtained data of reclassification of retrospectively classified cases (learn set) and of prospectively classified cases using discriminant analysis are given in table 2. Correct separation of healthy tissue from cancer cells could be obtained in all cases.

Table 2 Syntactic structure analysis

Actual cases	Learn set (N=75) Predicted cases				
	Healthy	Epidermoid	Adeno	Small cell	Large cell
Healthy	15	0	0	0	0
Epidermoid	0	12	0	1	2
Adeno	0	0	13	0	2
Small cell	0	2	0	13	0
Large cell	0	0	0	0	15

Actual cases	Prospective (N=25) Predicted cases				
	Healthy	Epidermoid	Adeno	Small cell	Large cell
Healthy	5	0	0	0	0
Epidermoid	0	4	0	1	0
Adeno	0	0	4	0	1
Small cell	0	0	0	5	0
Large cell	0	0	0	0	5

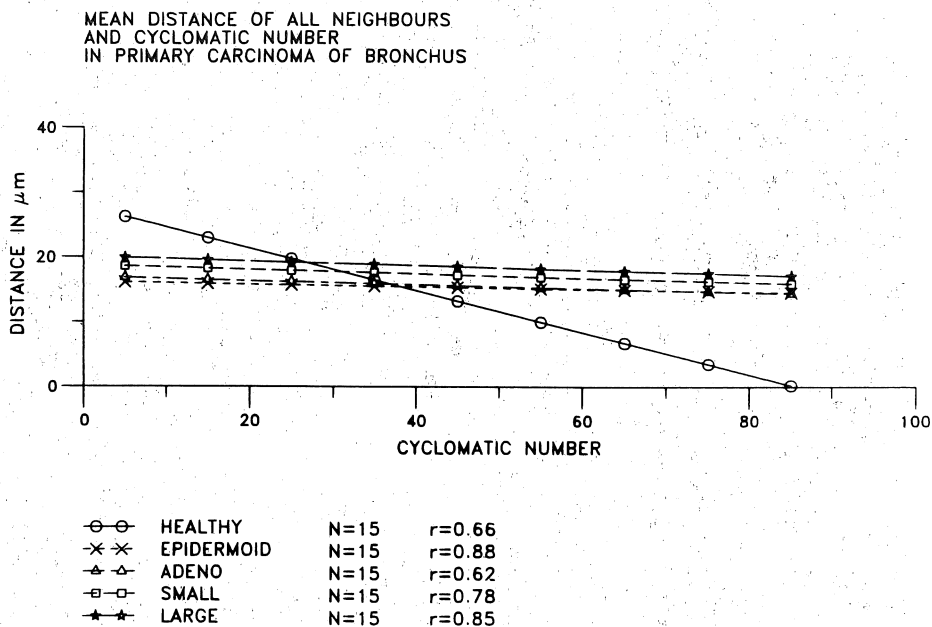


Fig. 2. Linear regression curves of mean distance of neighboring cells and cyclomatic number of healthy and carcinomatous lung tissue.

DISCUSSION

In routine histopathology major difficulties exist for distinguishing epidermoid carcinomas from small cell anaplastic carcinomas and for separating undifferentiated adeno carcinomas from large cell undifferentiated carcinomas of lung respectively. Special stainings do not help in many cases and specific monoclonal antibodies have until now not been successful in distinguishing the subtypes. Histopathological diagnosis is based upon tissue textures in general and only partly upon morphometric properties of cells or of other structures. Similar to previous approaches analysing colon mucosa and pleural tissue (Kayser 1985) syntactic structure analysis is useful in separating healthy lung parenchyma from primary carcinomas of lung. Due to limited computer power this approach of analysing textures of primary lung carcinomas was limited to simple descriptive parameters as cyclomatic number and average distances of neighboring cells. Computation of linear regression of average distance of nearest neighbors and cyclomatic number showed high correlation coefficients and different regression curves in healthy lung parenchyma compared to carcinomatous tissue. Naturally this effect can be noted in correct separation of healthy lung tissue from carcinomatous tissue by discriminant analysis in both, learn set and prospective data set. In this context we will not discuss the problems of carcinomas containing squamous- and adeno carcinoma cells, or squamous carcinoma cells and small carcinoma cells. Similar to difficulties obtained in routine histopathology

separation of epidermoid carcinoma from small cell anaplastic carcinoma as well as undifferentiated adeno carcinoma from large cell carcinoma could not be performed in all cases. The obtained misclassifications in epidermoid carcinoma - small cell carcinoma, and adeno-large cell carcinoma, can probably be improved when, in addition, the syntactic structure parameters, the nuclear and cytoplasmic sizes and nuclear cytoplasmic ratios of the tumour cells are included in the analysis. From theoretical point of view syntactic analysis of properties of textures is closely related to algorithm performed by pathologists, and similar results as well as similar difficulties in separating certain diagnoses should be obtained. Because of very limited area analysed compared to complete tumor volume major difficulties in sampling exist. Selecting certain areas considered to be characteristic for corresponding diagnosis are performed in routine pathology as well as in this study. In order to transform syntactic structure analysis into automated pattern analysis algorithms dealing with both problems have to be solved in advance. At the moment specificity and sensitivity of graph theory approach in histomorphologic diagnosis using interactive measurements is comparable to results obtained by pathologists.

REFERENCES

- Baak JPA, Oort J. A Manual of Morphometry in Diagnostic Pathology. Springer Verlag Berlin, Heidelberg, New York 1983.
- Kayser K. Analysis of adenomatous structures in histopathology. J Quant Cytology 1985 (in press).
- O'Callaghan JF. An alternative definition for neighborhood of a point. IEEE Trans Comput 24 (1975) 1121-1125.