Characterization of colon carcinoma growth pattern by computerized morphometry: Definition of a complexity index

L.E. FRANZÉN[1], V. HAHN-STRÖMBERG[2], H. EDVARDSSON[3] and L. BODIN[4]

[1]Department of Pathology and Cytology, Aleris Medilab, Täby and Department of Clinical Medicine, Örebro University; [2]Department of Clinical Medicine, Örebro University and Department of Pathology, University Hospital, Örebro; [3]Department of Pathology, Karlstad Regional Hospital, Karlstad; [4]Unit of Statistics and Epidemiology, University Hospital and Department of Statistics, Örebro University, Örebro, Sweden

Abstract. The invasive front of carcinomas may vary in complexity from smooth to highly complex when the front splits up into small cell clusters or even single cancer cells. The degree of complexity is usually estimated visually and semiquantitatively by a pathologist, although more objective methods based on computer-assisted image analysis are available. In this study, we compared the visual estimation of the irregularity of the tumour invasion front of colon carcinomas to different quantitative image analytical techniques and defined a complexity index for the invasive margin. Sections from 29 archived colon carcinomas were stained immunohistochemically for cytokeratin 8. Images of the tumour invasion front were read into a computer and thresholded so that the tumour tissue became black and the background white or so that the tumour front was outlined by a single pixel line. The invasive front was visually classified into four degrees of irregularity by a pathologist. The complexity of the front was then assessed using four different image analysis techniques, i.e. the estimation of fractal dimension, tumour front length, number of tumour cell clusters and lacunarity. Fractal dimension and tumour cell clusters together gave the best correlation to visual grading using a discriminant analysis. A cluster analysis and a tree diagram analysis were then performed and were found to be superior to visual estimation. The clusters represent different degrees of complexity and the result of the tree diagram analysis can be used to assign complexity indices to colon tumours. The fractal dimension separated tumours up to a certain level (1.5-1.6) of complexity. When the tumour front split up into small cell clusters, the counting of tumour cell clusters separated the cells over and above the fractal dimension. This new technique can be used to objectively and quantitatively describe the complexity of the invasive front of tumours.

Introduction

The growth of colon carcinoma into the surrounding tissue can show different patterns or different degrees of complexity regarding the tumour cell distribution at the invasive margin.

Jass et al (1) described two basic patterns of growth i.e. expansive with a smooth tumour-stroma interface and infiltrative with invasion in a diffuse manner. The detachment of cells from the main tumour mass into infiltrating isolated single cancer cells or clusters was later named as tumour budding (2). The patterns described by Jass et al (1) were revisited in 1994 and intraobserver and interobserver κ values of 0.37 and 0.41 were obtained, i.e. fair to moderate concordance (3). The poor agreement regarding the classification of the growth pattern in colorectal carcinomas between observers and for a single observer (see above) makes it necessary to look for objective methods to estimate the growth patterns.

Growth in nature appears to follow fractal geometry showing self similarity as opposed to Euclidian geometry, which describes growth in one, two or three dimensions (4). The use of fractal geometry was used in different contexts such as molecular biology and bone, vascular and tumour pathology (5). In tumour pathology, the fractal dimension was shown to be able to differentiate between the tubular, tubulovillous and villous adenomas of the colon (6). Fractal geometrical analysis was also able to differentiate severe dysplasia and cancer from benign conditions in the epithelial-connective tissue interface in the floor of the mouth (7). Lacunarity is another characteristic that may describe complexity and can vary when the fractal dimension is the same (8). Other characteristics that may reflect the complexity of a tumour border can be the length of the tumour-stromal interface and the number of tumour cells or tumour cell clusters.

We earlier used tumour characteristics such as fractal dimension and the number of tumour cell clusters to assess a relationship between the growth pattern of colon carcinoma and the expression of different adhesion proteins (ß-catenin, E-cadherin, Occludin and Claudin 2) and mutations in their
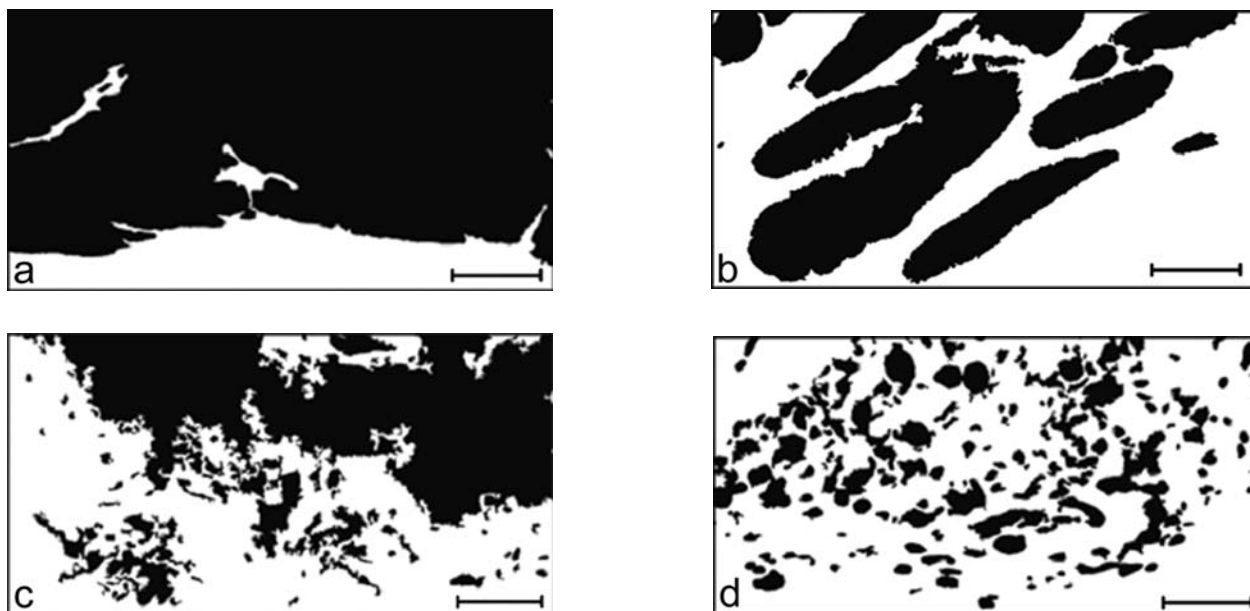
Figure 1. Images of tumours with a different growth pattern thresholded from specimens immunohistochemically stained for cytokeratin (Cam 5.2; bar represents 100 μm). (a) Tumour with a smooth infiltration margin (score 1), (b) larger tumour cell clusters are seen at the invasive front (score 2), (c) the margin has begun to break up into small cell clusters (score 3) and (d) the margin consists almost entirely of small cell clusters (score 4).

genes (9). In the present study, we compared the human visual assessment of the irregularity of the invasive front of colon carcinoma to different computer-assisted techniques (fractal dimension, number of tumour cell clusters, lacunarity and the length of the tumour-stromal interface). A complexity index, based on cluster and tree diagram analyses of the fractal dimension and the number of tumour cell clusters was constructed to estimate the grade of complexity of the tumour invasive front.

Materials and methods

Histopathological specimens. The tumours were selected from archived paraffin-embedded tissue blocks at the Department of Pathology, Örebro University Hospital from the years 2002 and 2003. Twenty-nine sections, one from each tumour, were selected from regular hematoxylin and eosin-stained slides. The selection of tumours and slides was done so that there was a representation of tumours with infiltrative and expansive patterns as they appeared under the microscope. Only carcinomas from the colon were considered since most patients with rectal carcinomas obtain local radiation to the tumour preoperatively. The study was approved by the local ethics committee in Örebro.

Five micron thick sections were cut onto silano slides for Dako TechMate Horizon (Dakopatts, Älvsjö, Sweden). After deparaffinization, the slides were treated in a microwave oven for antigen retrieval (2x5 min at 780 W in a citrate buffer at pH 6.0). The immunohistochemical processing was performed in Dako TechMate Horizon and the Dako ChemMate Reagents were used. The antibody Cam 5.2 for cytokeratin 8/18 (Becton-Dickinson, UK, dilution at 1:5) was used for the staining of the tumour cells.

Measurements. The capturing of images from the tumour-stromal interface was performed using a Leica DC200 digital camera mounted on a Leica DMRXE microscope (Leica Microsystems Wetzlar GmbH, Germany) (objective 10X). Images were digitized and stored in an uncompressed TIF-format. No compression of the images was performed. From each specimen, a median of 9 images (range 5-16; a total of 265 images) was captured, the number of images depended on the length of the tumour-stromal front. Areas with artifacts and necroses in the invasive margin were omitted. The images were digitized and thresholded so that the immunohistochemically-stained areas were black (Fig. 1). Structures measuring 20 pixels or less were omitted in order not to obtain small dots or interrupted lines, which might cause lower values of the fractal dimension (see below). The area of 20 pixels corresponds to approximately the size of a lymphocyte nucleus. Binary noise reduction or other filtering of the images was not performed in order to keep the edges of the tumour unperturbed. The images measured 1 210x950 pixels and were stored in an uncompressed TIF-format.

The visual estimation was performed on the thresholded black/white images. A scoring system from 1 to 4 was used (Fig. 1) as follows: tumour margin is regular and smooth (score 1), tumour margin is irregular with larger tumour cell clusters (score 2), tumour margin is irregular with larger and smaller clusters (score 3) and tumour margin consists almost entirely of smaller clusters (score 4). The scoring system is similar to that used by Spiro *et al* (10) in oral tongue cancer. Five tumours fulfilled the criteria of showing tumour budding, i.e. the tumour front was split up into small cell clusters with five cells or less (11).

The morphometrical calculations were performed using image analysis software. Adobe Photoshop 7.0 (Adobe Systems Inc. San Jose, CA, USA) with the Fovea Pro plug in for image analysis (Reindeer Graphics, Inc., NC, USA) was used to threshold and delineate the tumour and tumour cell clusters (Fig. 2). The number of tumour cell clusters at the infiltration front was counted after omitting the structures in
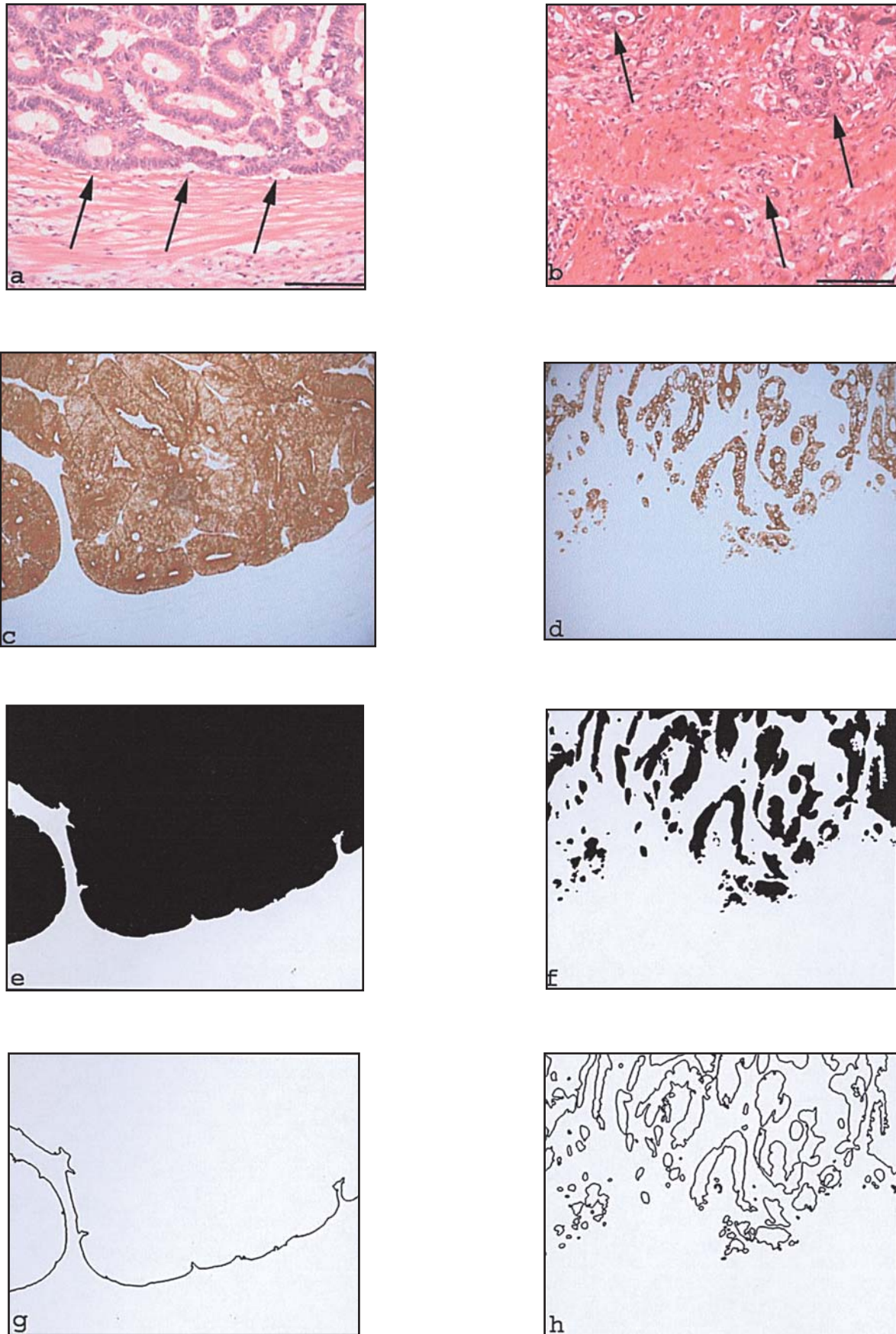
Figure 2. Two columns of images, to the left from a smooth tumour front and to the right an irregularly infiltrating tumour (bars represent 200 μm). (a and b) Shows the tumour fronts indicted by arrows. (c and d) The epithelial tumour cells were stained immunohistochemically for keratin (Cam 5.2) and stands out from the unstained background. (e and f) A thresholding procedure was performed and the images binarized showing the tumour cells in black and the background in white. (g and h) The tumour fronts are outlined after image processing.
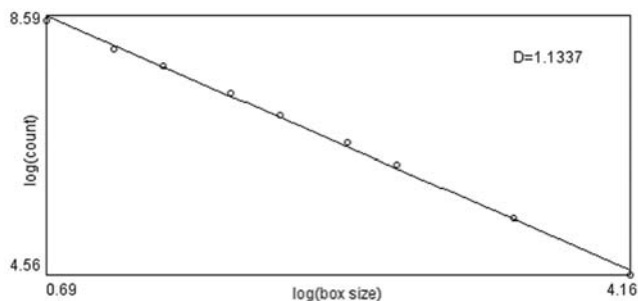
Figure 3. A typical log-log graph used to calculate the box-counting fractal dimension. There is a linear relationship between box size and the number of boxes in the interval of boxes used (2 to 64 pixels).
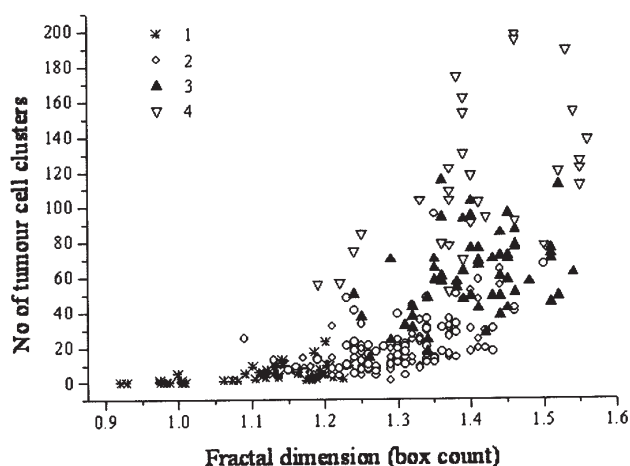

Figure 4. A scatter diagram of tumour cell clusters and fractal dimension (Box count) with legends for the visual gradings 1-4. These data are based on the visual grading of the tumour complexity.

the image that touched any border of the image. The tumour front was then converted to a one pixel wide line. The length of the front was calculated by counting the number of intersections between a circular grid superimposed on the image and the tumour outline. The grid consisted of 99 circles, 73 μm in diameter and 37 μm apart. The ImageJ software (Rasband, W.S., Image J, National Institute of Health, Bethesda, MD, USA, http://rsb.info.nih.gov/ij/, 1997-2003) was used to calculate the fractal dimension with the box counting method and the lacunarity. A full description of the technique can be found in Smith *et al*, 1996 (8). Briefly, a range of boxes were superimposed over the image and the number of boxes needed to cover the one pixel wide outline of the tumour was counted. A plot was then generated with the log of box size on the x-axis and the log of count on the y-axis (Fig. 3). Boxes measuring between 2 and 64 pixels (2,3,4,6,8,12,16,32 and 64) were found to give a linear relationship in the log-log plot (Fig. 3).

Statistics. The Pearson correlation coefficient was used to find correlations between different image analyses parameters obtained from the computer. A discriminant analysis showed the connection between the visual grading and the image analysis data. We performed a cluster analysis

of the 265 images with Ward's algorithm in order to find homogeneous subsets of images with respect to the objective measurements obtained from the image analysis. Cluster solutions from the Ward method were refined with the K-means algorithm using the cluster centers from the Ward method as starting points. A tree diagram analysis or recursive partitioning provided a nonlinear and non-additive approach to classify and predict classifications of images, based on visual gradings and on cluster solutions.

Computations were performed with the statistical packages SPSS and SPLUS. The principles for cluster analysis and tree diagram analysis from Everitt *et al* (12) and Zhang and Singer (13) were followed.

Results

The sections used for analysis were selected from specimens stained with hematoxylin and eosin (Fig. 2a and b). After immunohistochemical staining for keratin, the tumour epithelial cells became dark brown and stood out against the transparent background (Fig. 2c and d). Thresholding and binarization of the image made the epithelium black and the non-epithelial background white (Fig. 2e and f). These images were used for the calculations of the number of tumour cell clusters. Then an outlining of the tumour cells and cell clusters was done to assess fractal dimension, lacunarity and tumour front length (Fig. 2g and h).

The visual and semiquantitative grading of the complexity of the tumour infiltrative front was performed in thresholded images and scored into grades 1 through 4 (Fig. 1). The number of images assigned in the different score groups was 50 (grade 1), 125 (grade 2), 59 (grade 3) and 31 (grade 4). Five of the tumours had the same score in all the images, whereas the score in the others differed between the images.

Four characteristics were assessed with image analysis, fractal dimension, tumour front length, tumour cell clusters and lacunarity. The correlations between these four variables are shown in Table I. The correlations between the first three variables are fairly strong whereas lacunarity had small and negative correlations with the other variables. None of the 265 images can be classified as an isolated outlier with extreme image analysis data. Hence, the 265 images are used in the preceding analysis and in particular are included in the cluster analyses.

The visual grading with scores 1 to 4 was correlated to the image analysis data through a discriminant analysis. An analysis using all four morphometrical variables reproduced the grading correctly for 78.1% of the cases when an á priori probability for each grading was proportional to the number of cases in each grading (i.e. 18.9, 47.2, 22.3 and 11.7% respectively) and 80.4% correct when an equal distribution was assumed (i.e. 25% for each grading). We tried to perform the discriminant analysis with single or pairs of image analysis variables and the combination of tumour cell clusters and fractal dimension scored almost as good as the complete set of four variables, up to 79.1% correct. Less successful results were obtained for lacunarity, either alone or in any combination with other variables.

Fig. 4 shows the scatter diagram between tumour cell clusters and fractal dimension where the visual gradings have

Table I. Correlations (Pearson's) between image analysis variables.

	Fractal dimension	Tumour front length	Lacunarity
Tumour cell clusters	+0.66	+0.69	-0.18
Fractal dimension	-	+0.92	-0.27
Tumour front length	-	-	-0.22

Table II. Cluster analysis.

	Cluster				
Measure	1	2	3	4	5
N	31	69	78	65	22
Tumour cell clusters	4.97	13.51	22.12	65.54	134.68
Fractal dimension	1.07	1.22	1.34	1.42	1.44

Mean values for two measures used in a cluster analysis of 265 images from 29 patients. The cluster analysis was performed with K-means clustering after initial seed values were found with Ward's method.
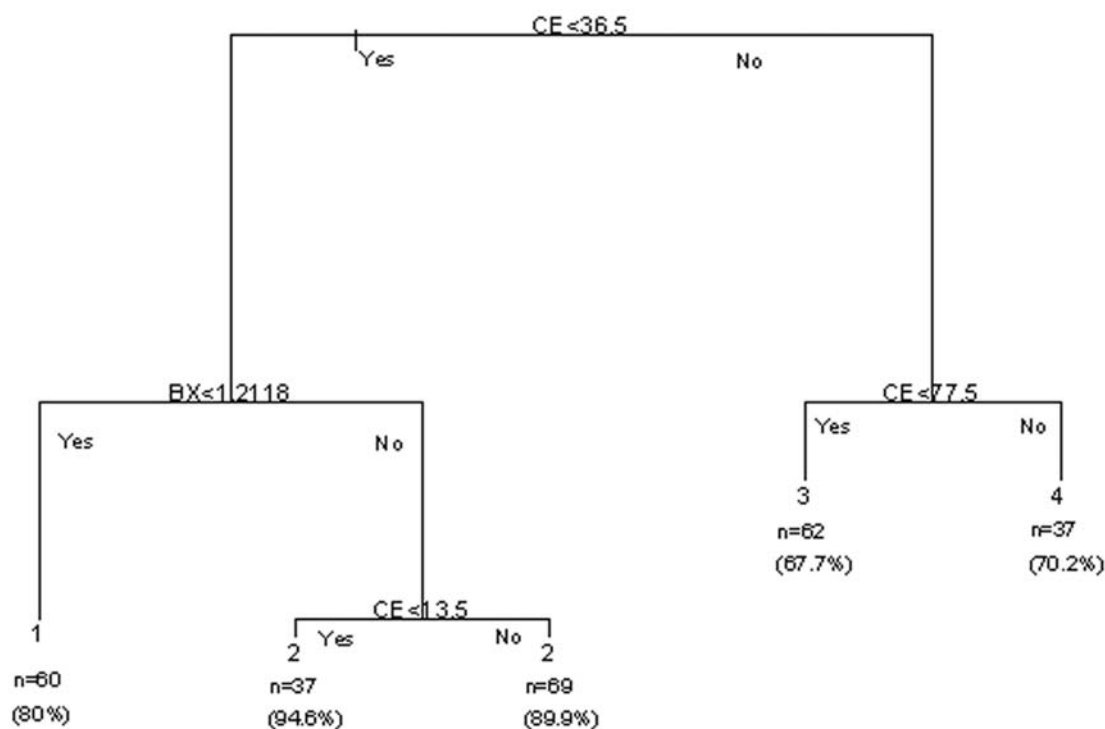

Figure 5. A tree diagram based on the visual grading of tumour complexity showing a partitioning rule for obtaining the four visual grades, 1-4, based on the number of tumour cell clusters (CE) and fractal dimension (BX). Classification is 80.3% correct overall and the correct percentage for each terminal node is shown in brackets.

different legends. The figure indicates a certain degree of clustering, although the four gradings are not clearly separated.

To be able to quickly determine a classification into the four visual gradings, we performed a tree diagram analysis, (Fig. 5). The analysis was conducted so that decision nodes (terminal points) of at least 35 images could be produced. Only four decision points were necessary to obtain a correct classification as high as 80.3%. Allowing for a smaller number of images in the terminal points, down to 20, in this case gave the same classification percentage, though with the disadvantage of a more complicated decision algorithm.

To base the classification solely on the image analysis data, we started with a cluster analysis using Ward's algorithm supplemented with a refinement and re-allocation of cluster
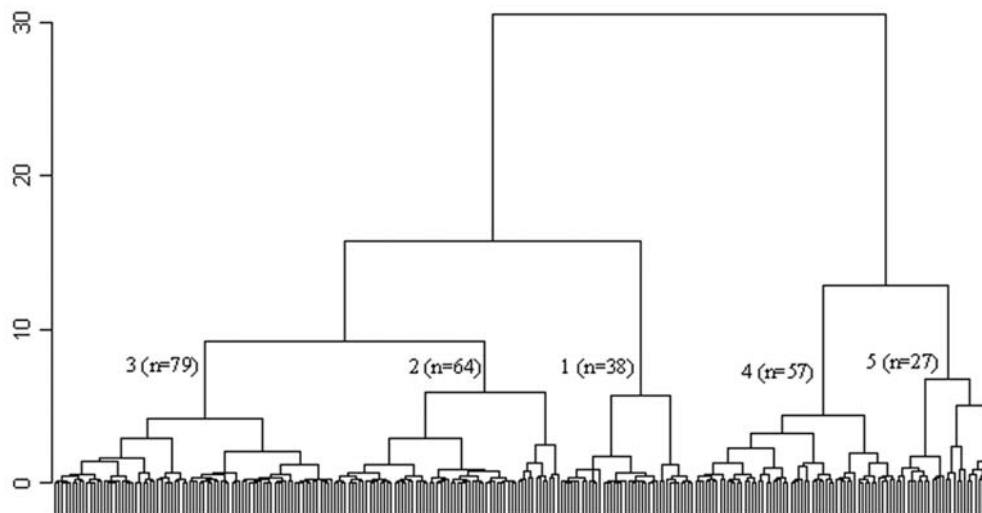
Figure 6. A dendrogram of hierarchical cluster analysis with Ward's method using observations on tumour cell clusters and fractal dimension for 265 images. Five selected clusters and their sizes are shown. The vertical axis represents the distances in aggregation of hierarchical clusters.
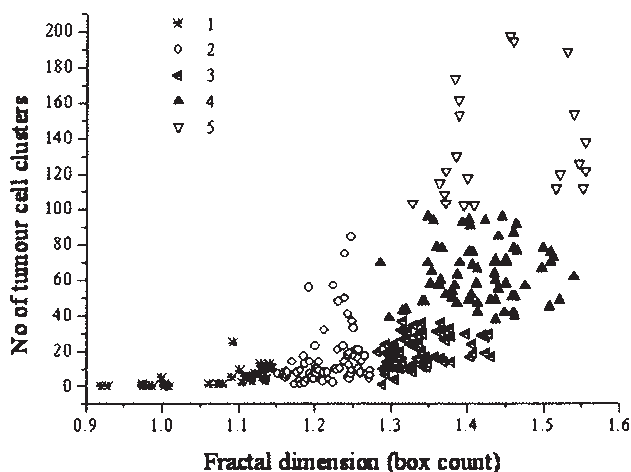

Figure 7. A scatter diagram of tumour cell clusters and fractal dimension with legends for the cluster solution with five clusters based on two variables i.e. tumour cell clusters and fractal dimension. These data are based on the cluster analysis presented in TII in the dendrogram in Fig. 6.

members through the K-means algorithm. An analysis using the four variables as well as an analysis based only on tumour cell clusters and fractal dimension produced almost similar results with respect to cluster structure and allocations of images to the clusters. Interpretations of dendrograms and an application of the Calinski and Harabasz criterion (14) indicated preferences for the four- or fivecluster solutions. The dendrogram for Ward's method applied with tumour cell clusters and fractal dimension is shown in Fig. 6.

With the cluster solution from Ward's method as initial seed the K-means algorithm produced the final five-cluster solution. The mean values and the number of images in each cluster are shown in Table II. The number of images in each cluster is similar to but not exactly the same as with Ward's method.

The cluster separation for the five clusters with respect to tumour cell clusters and fractal dimension is shown in Fig. 7.

The cluster structure is now much more distinct than that based on the visual grading (Fig. 4).

The application of the cluster results requires decision rules on where to allocate the images. The tree diagram analysis, with minimum terminal node size set to 20, resulted in rules that gave a correct classification into the five clusters of 97.0% with 8 terminal nodes as shown in Fig. 8.

Discussion

A morphological grading of certain characteristics of cells and tissues is traditionally based on semiquantitative estimations by the viewer. Such estimations do not show a good reproducibility and low κ-values for either inter- or intraindividual estimations are often found. For instance, in the colorectum, the invasive patterns of colorectal carcinoma show only fair interindividual (κ 0.37) to moderate intraindividual (κ 0.41) concordance (3). The same is true for the grading of dysplasia in advanced colorectal adenomas (κ = 0.20 and 0.42 for intra- and interindividual grading, respectively) (15). New and more sophisticated image analysis software for computer-based analysis has made it possible to quantitatively analyze complex biological structures in a standardized way. The present methodological study was performed in order to compare different image analysis techniques (fractal dimension, tumour cell clusters, lacunarity and tumour front length) to assess the complexity of the tumour-stroma interface of colon carcinomas in comparison to the visual, semiquantitative estimation performed by a pathologist and to define a complexity index based on cluster and tree diagram analyses.

Growth in nature shows self-similarity under scale changes and thus follows fractal geometry. Natural objects, in contrast to mathematical fractals, are not the result of a construction by iteration. In nature, the fractal dimension shows self-similarity only over a limited scale range and if a linear segment is present on the log-log graph, the gradient of this will accurately reflect that dimension (4). The box sizes used in this study resulted in linear log-log plots (Fig. 3). Binary noise and obliquely cut small parts of tumour cell cytoplasm that stain
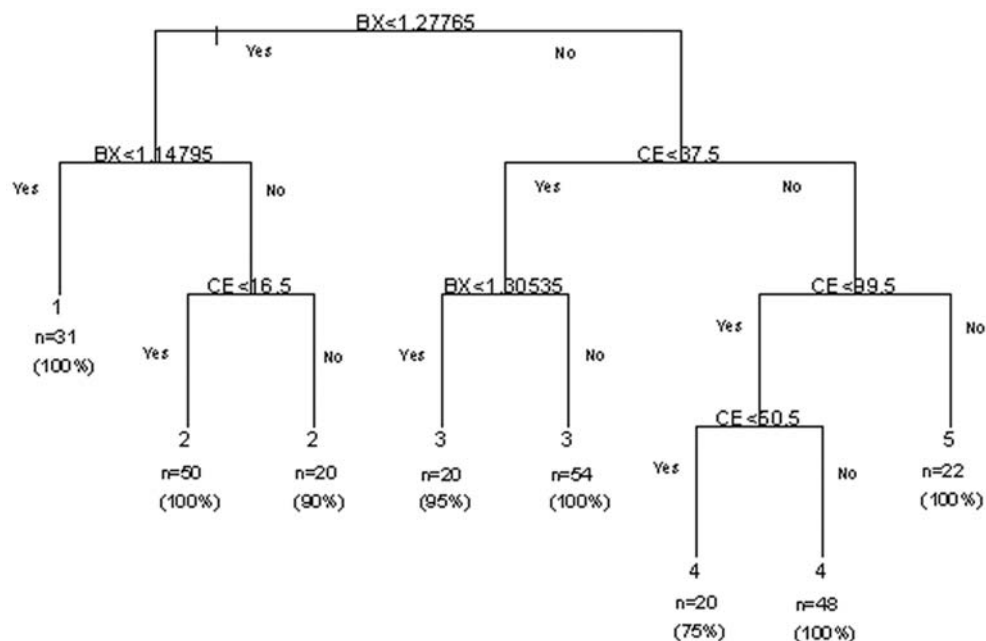
Figure 8. A tree diagram showing a partitioning rule for obtaining five clusters based on the number of tumour cell clusters (CE) and fractal dimension (BX). Classification is 97.0% correct overall and the correct percentage for each terminal node is shown in brackets.

with the cytokeratin marker cause dots and lose ends in the image and interfere with the fractal dimension calculation. Prior to analysis, we therefore omitted all structures measuring 20 microns or less. In this way, we avoided dots and loose ends in the image. Twenty microns corresponds to the size of a lymphocyte nucleus with the magnification used and is much smaller than tumour cells and clusters. We therefore do not think that tumour cells or clusters were disregarded.

The application of fractal geometry enabled measurements of the fractal dimension to describe the 'complexity' of tissue structures (16). However, there are different mathematical approaches to the calculation of the fractal dimension (4), which can render different results for the same structure (8). In the present study, we chose the box counting technique, which is widely used for biological material (4,5) and has been described as a useful method of quantifying the nature of tumour borders in colorectal carcinomas which are often subjectively-divided into 'pushing' or 'infiltrative' types (17). The fractal dimension ranged up to ~1.55 in this study, which is comparable to what was found in squamous cell carcinomas, 1.6 (7) and slightly higher than in basal cell carcinomas of the infiltrative type, up to 1.468 (18).

Fractals do not uniquely describe irregular binary features of biological objects and objects that look very different may have very similar fractal dimensions (8). One measure that can distinguish such objects is the characteristic 'lacunarity', which is supposed to measure the 'differences in texture'. In a general sense, lacunarity is a measure of the non-uniformity (heterogeneity) of structure or the degree of structural variance within an object (8) and not the 'self-similarity' that is typical for a structure with fractal properties. The lacunarity correlated poorly to the visual scoring and to the other morphometrical characteristics. This may indicate that irregularities of the tumour invasion front have fractal characteristics with high 'self-similarity' rather than a high structural variance.

The length of the tumour-stromal interface can also serve as an indirect measure of the complexity of the tumour invasion front. This characteristic correlates highly with the fractal dimension but does not have the inherent self similarity. In the cluster analysis, the tumour front length did not add much information when the four parameters were correlated to the visual grading and was therefore not used.

The tendency of the tumour margin to split up into cell clusters or single cells reflects the cohesiveness of the cells indicating a lower degree of differentiation. In the visual scoring system we use, this property of the tumour rendered it a higher score. In the image analysis, it appears (Fig. 7) that the fractal dimension separates tumours up to values of 1.4 to 1.6 and that tumours within this range are further separated by the number of tumour cell clusters. Thus, this variable appears to give additional information over and above the fractal dimension alone.

Altogether four variables of the tumour invasion front were analyzed and the solution based on the four variables did not differ significantly from the two variable solution. The two variable solution with fractal dimension and tumour cell clusters was chosen since it is a much simpler solution without any notable loss of precision. The visual grading correlated well with the various morphometrical characteristics, though there was a substantial overlap between the clusters with legends for the visual gradings (Fig. 4) compared to the better delineated clusters obtained mathematically with legends for the cluster solution (Fig. 7). This indicates a better and more distinct separation between clusters of tumours with different invasive patterns and therefore a more accurate classification. However, there was not a great difference between the 4- and 5-cluster solutions, though the 5-cluster solution is more robust to changes in data and method.

In conclusion, a computer-based morphometrical technique was found to be superior to visual grading of complexity of the invasive front of colon carcinomas. The fractal dimension together with the number of tumour cell clusters were the best characteristics to objectively describe the complexity of the invasion front. A cluster analysis followed by a tree diagram analysis resulted in rules that gave correct classification in 97% of the images. The five clusters can be used to discriminate tumours regarding their irregularity as a complexity index. We recently used the complexity index to study the expression of different cellular adhesion proteins (ß-catenin, E-cadherin, Occludin and Claudin 2) in different growth patterns (9). No correlation was found between the growth pattern of colon carcinomas and the expression of different adhesion proteins or to mutations in their genes indicating that the complexity of colon carcinoma invasion is not dependent on genetic derangements in genes of the adhesion proteins themselves or their distribution. Rather, aberrations in the function of other proteins related to the adhesive proteins could be responsible. The fractal dimension was earlier used as a single variable to estimate the invasive pattern of tumours (7,17,18). In contrast, the complexity index now presented takes into account the irregularity of the tumour border (fractal dimension) as well as the splitting up of tumour cell clusters (tumour cell clusters). It can be used in scientific studies when detailed knowledge of the invasive pattern is of importance.

Acknowledgements

References

1. Jass JR, Atkin WS, Cuzik J, Bussey HJ, Morson BC, Northover JM and Todd IP: The grading of rectal cancer: historical perspectives and a multivariate analysis of 447 cases. Histopathology 10: 437-459, 1986.
2. Hase K, Shatney C, Johnson D, Trollope M and Vierra M: Prognostic value of tumor 'budding' in patients with colorectal cancer. Dis Colon Rectum 36: 627-635, 1993.
3. Deans GT, Heatley M, Anderson N, Patterson CC, Rowlands BJ, Parks TG and Spence RA: Jass' classification revisited. J Am Coll Surg 179: 11-17, 1994.
4. Cross SS: The application of fractal geometric analysis to microscopic images. Micron 1: 101-113, 1994.
5. Cross SS: Fractals in pathology: J Pathol 182: 1-8, 1997.
6. Cross SS, Bury JP, Silcocks PB, Stephenson TJ and Cotton DWK: Fractal geometric analysis of colorectal polyps. J Pathol 172: 317-323, 1994.
7. Landini G and Rippin JW: Fractal dimensions of the epithelial-connective tissue interfaces in premalignant and malignant epithelial lesions of the floor of the mouth. Analyt Quant Cytol Histol 15: 144-149, 1993.
8. Smith TG, Lange GD and Marks WB: Fractal methods and results in cellular morphology - dimensions, lacunarity and multifractals. J Neurosci Methods 69: 123-136, 1996.
9. Hahn-Stromberg V, Edvardsson H, Bodin L and Franzen L: Disturbed expression of E-cadherin, beta-catenin and tight junction proteins in colon carcinoma is unrelated to growth pattern and genetic polymorphisms. APMIS 116: 253-262, 2008.
10. Spiro RH, Guillamondegui O Jr , Paulino AF and Huvos AG: Pattern of invasion and margin assessment in patients with oral tongue cancer. Head and Neck 21: 408-413, 1999.
11. Ueno H, Murphy J, Jass JR, Mochizuki H and Talbot IC: Tumour 'budding' as an index to estimate the potential of aggressiveness in rectal cancer. Histopathology 40: 127-132, 2002.
12. Everitt BS, Landau S and Leese M: Cluster Analysis. 4th edition, Arnold, London, 2001.
13. Zhang H and Singer B: Recursive Partitioning in the Health Sciences (Statistics for Biology and Health). Springer, New York, 1999.
14. Calinski RB and Harabasz J: A dendrite method for cluster analysis. Comm Stat 3: 1-27, 1974.
15. Terry MB, Neugut AI, Bostick RM, Potter JD, Haile RW and Fenoglio-Preisler CM: Reliability in the classification of advanced colorectal carcinoma. Cancer Epidemiol Biomarkers Prev 11: 660-663, 2002.
16. Losa GA and Nonnenmacher TF: Self-similarity and fractal irregularity in pathologic tissues. Mod Pathol 9: 174-182, 1996.
17. Cross SS, Scholfield JH, Kennedy A and Cotton DWK: Measuring the fractal dimension of tumour borders. J Pathol 168: 117A, 1992.
18. Miracco C, Bianciardi G, Perrone A, Bruni A, Lazzi S and Luzi P: Fractal dimension of epithelial-connective tissue interface in basal cell carcinoma of the skin. In: Fractals in Biology and Medicine. Losa GA, Merlini D, Nonnenmacher TF, Weibel ER (eds). Vol 2, Birkhäuser Verlag, Basel, pp284-293, 1998.