JOURNAL OF PATHOLOGY
J. Pathol. **185:** 366–381 (1998)

FRACTAL CHARACTERIZATION OF CHROMATIN APPEARANCE FOR DIAGNOSIS IN BREAST CYTOLOGY

ANDREW J. EINSTEIN[1,2]*, HAI-SHAN WU[2], MIGUEL SANCHEZ[3] AND JOAN GIL[2]

[1]*Department of Biomathematical Sciences, Mount Sinai School of Medicine, One Gustave L. Levy Place, Box 1023, New York, NY 10029, U.S.A.*
[2]*Department of Pathology, Mount Sinai School of Medicine, One Gustave L. Levy Place, Box 1194, New York, NY 10029, U.S.A.*
[3]*Department of Pathology, Englewood Hospital and Medical Center, 350 Engle Street, Englewood, NJ 07631, U.S.A.*

SUMMARY

This study explores the use of fractal analysis in the numerical description of chromatin appearance in breast cytology. Images of nuclei from fine-needle aspiration biopsies of the breast are characterized in terms of their Minkowski and spectral fractal dimensions, for 19 patients with benign epithelial cell lesions and 22 with invasive ductal carcinomas. Chromatin appearance in breast epithelial cell nuclear images is demonstrated to be fractal, suggesting that the three-dimensional chromatin structure in these cells also has fractal properties. A statistically significant difference between the mean spectral dimensions of the benign and malignant cases is demonstrated. The two fractal dimensions are very weakly correlated. A statistically significant difference between the benign and malignant cases in lacunarity, a fractal property characterizing the size of holes or gaps in a texture, is found over a wide range of scales. These differences are particularly pronounced at the smallest and largest scales, corresponding respectively to fine-scale texture, indicating whether chromatin is clumped or fine, and to large-scale structures like nucleoli. Logistic regression and artificial neural network classification models are developed to classify unknown cases on the basis of fractal measures of chromatin texture. Using leave-one-out cross-validation, the best logistic regression classifier correctly diagnoses 95·1 per cent of the cases. The best neural network model can correctly classify all of the cases, but it is unclear whether this is due to overtraining. Fractal dimensions and lacunarity are useful tools for the quantitative characterization of chromatin appearance, and can potentially be incorporated into image analysis devices to assure the quality and reproducibility of diagnosis by breast fine-needle aspiration biopsy.

KEY WORDS—**fractals; fractal dimension; lacunarity; logistic regression; artificial neural networks; morphometry; image analysis; breast cytology**

INTRODUCTION

Numerous studies have aimed at developing image analysis procedures for the resolution of difficult differential diagnoses in cytology and histopathology. The general approach used involves characterizing nuclei with numerical measures of factors considered subjectively by pathologists. A diagnosis is assigned on the basis of these features (also referred to as descriptors), in accordance with a prescribed classificatory approach determined and validated on the basis of representative sets of cases. Among the most useful features for cytological applications have been measures of nuclear size, pleomorphism, and chromatin appearance. This study evaluates the use of fractal descriptors of chromatin appearance in breast cytology.

In image analysis, chromatin appearance is expressed in terms of the texture of a digitized nuclear image, i.e., the spatial distribution of grey values. A nuclear image can be viewed as a surface for which the *x*- and *y*-coordinates represent position and the *z*-coordinate represents grey level. Figures 1 and 2 illustrate the texture of representative benign and malignant breast epithelial cell nuclei. While qualitatively they look different, it has proven exceedingly difficult to assess quantitatively such subjective differences in texture in a way suitable for classificatory analysis. A number of approaches have been suggested,[1] based on Markovian analysis,[2,3] run length statistics,[4] textons,[5] Fourier analysis,[6] mathematical morphology,[7] and local grey level variation.[8] The most popular approach, based on Markovian analysis, may yield over 60 highly correlated[9] features which are difficult to interpret and do not correspond to the visual impressions of pathologists.[10] Fractal geometry offers an alternative approach to chromatin texture description.

Fractal geometry provides a framework to describe mathematically objects exhibiting structure over a range of scales.[11] Such objects, which have no characteristic size but rather exhibit similar detail on many scales, are called fractals and may be described by their fractal dimensions and by measures of lacunarity. Applications of fractal geometry to pathology[12,13] have concentrated on using fractal dimensions to characterize the structural complexity or irregularity of cell[14,15] and nuclear[16] membranes, of tissue shape,[17–19] and of cell growth *in vitro*.[20,21] In this work, we use methods of fractal analysis to describe the structure of chromatin in light microscopic images of breast epithelial cell nuclei. We demonstrate differences in fractal properties between benign and malignant cases, and show how fractal dimensions and lacunarity can be used in conjunction

*Correspondence to: Andrew J. Einstein, Mount Sinai School of Medicine, Department of Biomathematical Sciences, One Gustave L. Levy Place, Box 1023, New York, NY 10029, U.S.A. E-mail: einstein@msvax.mssm.edu

CCC 0022–3417/98/040366–16 $17.50

Received 23 October 1997
Accepted 2 March 1998

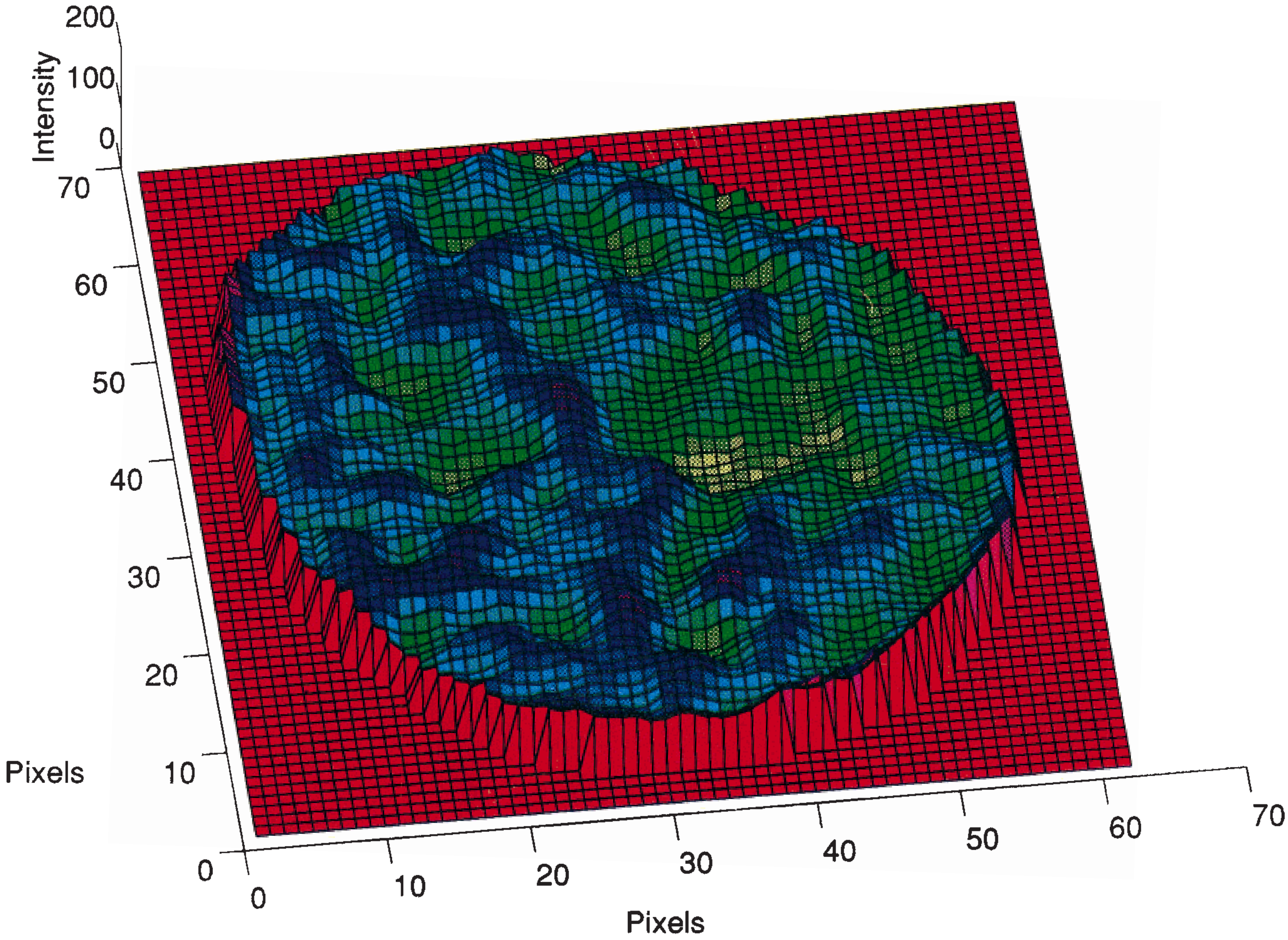

Fig. 1—Surface plot of a benign breast epithelial cell nucleus

with logistic regression and artificial neural network classifiers to diagnose breast lesions.

MATERIALS AND METHODS

Cytological and histological materials

Cytology specimens were obtained from 41 patients, of whom 22 were diagnosed with invasive ductal carcinoma and 19 as benign with no atypia. Each diagnosis was made by a cytologist and independently confirmed by a pathologist on the basis of histopathological findings. With the exception of two cases, all histopathological diagnoses were reconfirmed by a second pathologist. Cytological specimens were processed using the ultrafast Papanicolaou protocol.[22]

Image analysis

Image acquisition was performed on a self-assembled system based on a Gateway 2000 486DX2/50E microcomputer augmented with Sprynt i860 image processing boards and the Semper 6 Plus graphics program, connected to a Nikon Optiphot microscope equipped with a 100 × Nikon Plan objective through a Sony DXC-M2 camera. For each specimen, randomly selected epithelial cell nuclei were segmented, excluding overlapping and damaged nuclei as well as those with insufficient contrast. Segmentation was performed using an arc-forming method,[23] involving the selection of three points on the margin of a nucleus, the computation and display of the arc connecting these points, and the extension of this contour with additional arcs until the whole profile is outlined. We have found this approach to be the most reproducible method for segmenting cytological images. Subsequent image processing and descriptor computation was performed using software written in C and run on a Silicon Graphics Indigo^2 workstation and Silicon Graphics Power Challenge XL supercomputer. Nuclear images were normalized[8] to compensate for possible differences in staining and lighting conditions and then screened to assure that they were complete, properly segmented, and of sufficient contrast for texture analysis. A total of 2621 images were collected, with a minimum of 55 and a mean of 64 nuclei per patient.

Fractal dimension calculations

Several related mathematical formulations of fractal dimension have been suggested, such as Hausdorf–

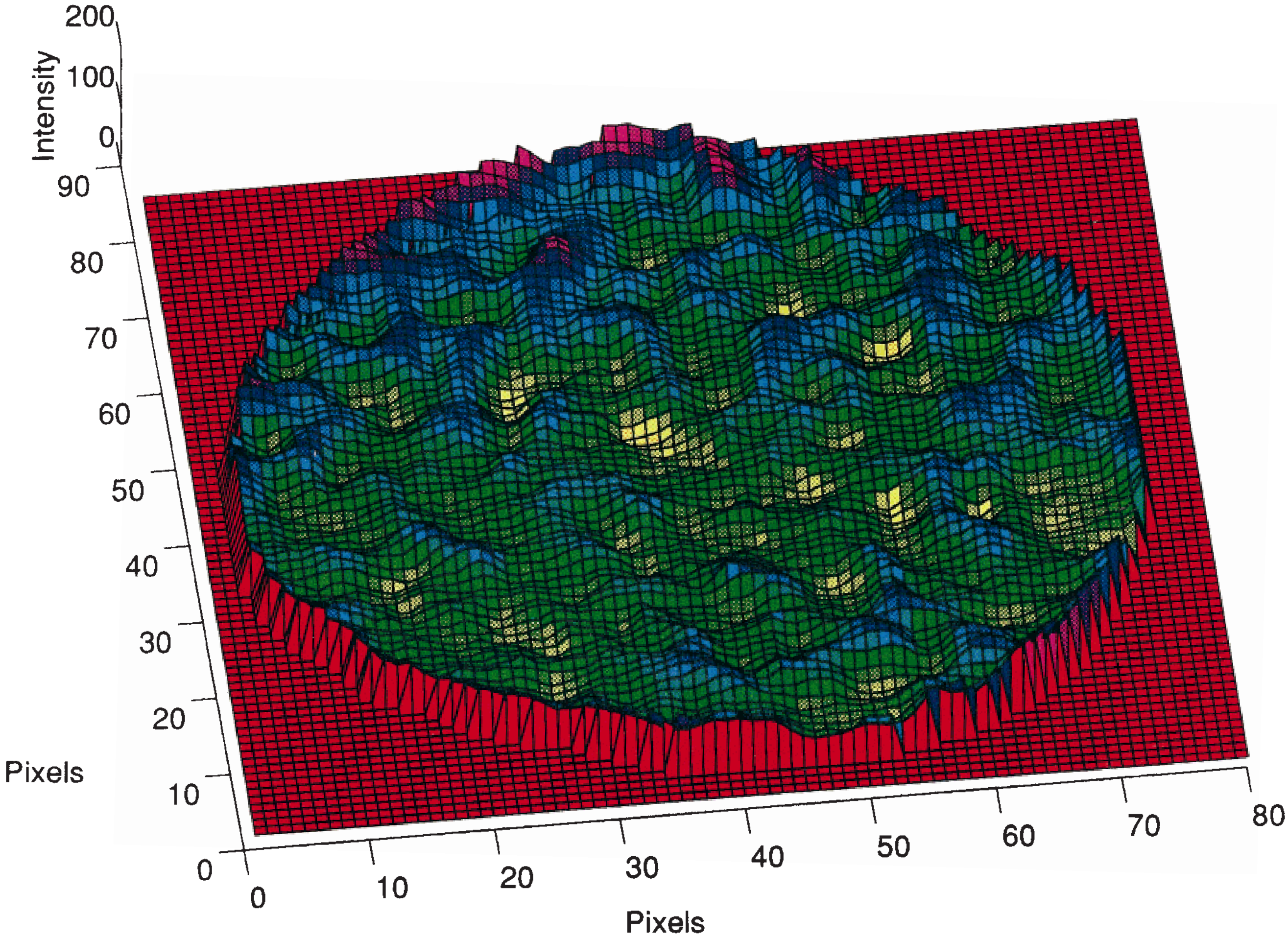

Fig. 2—Surface plot of a malignant breast epithelial cell nucleus

Besicovitch dimension, Minkowski–Bouligand dimension, Kolmogorov box-counting dimension, spectral dimension, and Korčak dimension.[24] Each method for determining fractal dimension may characterize a particular aspect of the fractal nature of a nuclear image; not all fractal dimensions are applicable in every context. In particular, it is important to make a distinction between self-similar and self-affine fractals, since the methods used to calculate dimensions differ between these two classes of fractals. Mandelbrot[25] illustrates this difference by comparing measurement on the earth's surface, where the choice of north–south and east–west as the coordinate axes is in a sense arbitrary and distances are meaningful, to measurement on a graph of volume versus pressure, where the choice of coordinate axes is canonical but distances $\sqrt{(\Delta V)^2+(\Delta p)^2}$ are meaningless and therefore cannot be used in determining the fractal dimension. In the first case, linear log–log scaling would demonstrate that an object is self-similar, while in the second case, linear log–log scaling would show that it is self-affine. Since for our surface plots the x- and y-coordinates represent position and the z-coordinate represents grey level, fractal characterization of the images should use methods appropriate for self-affine objects. In this study, we characterize breast epithelial cell nuclei using the Minkowski–Bouligand fractal dimension (D_{MB}) and the spectral fractal dimension (D_S).

Minkowski–Bouligand dimensions—Minkowski–Bouligand dimensions were determined using a modification of the variation method of Dubuc *et al.*[26,27] This method is based on the concept of ε-variation. In the digitized grey scale image of a nucleus, an intensity value between 0 and 255 is associated with each pixel. The ε-oscillation for a given pixel is defined as the difference between the maximum and minimum intensity values of all pixels within a distance ε of the given pixel, i.e., all pixels in a disc of radius ε, centred at the given pixel. In terms of the surface plot, the distance ε is measured in the x–y plane, thus avoiding the meaningless distances discussed in the previous paragraph. The ε-variation for the whole nucleus, denoted as $V_f(\varepsilon)$, is the sum of the ε-oscillations for all pixels in the image. The Minkowski dimension is estimated by 3 minus the slope of the least-squares line fitting the plot of log ε versus log $V_f(\varepsilon)$ over a suitable range of ε's. In this study, $V_f(\varepsilon)$ was computed for each nucleus at 145 values of ε, ranging between 1 and 20 pixels (0·135 and 2·69 μm), and also over a shorter range of resolutions, considering 21

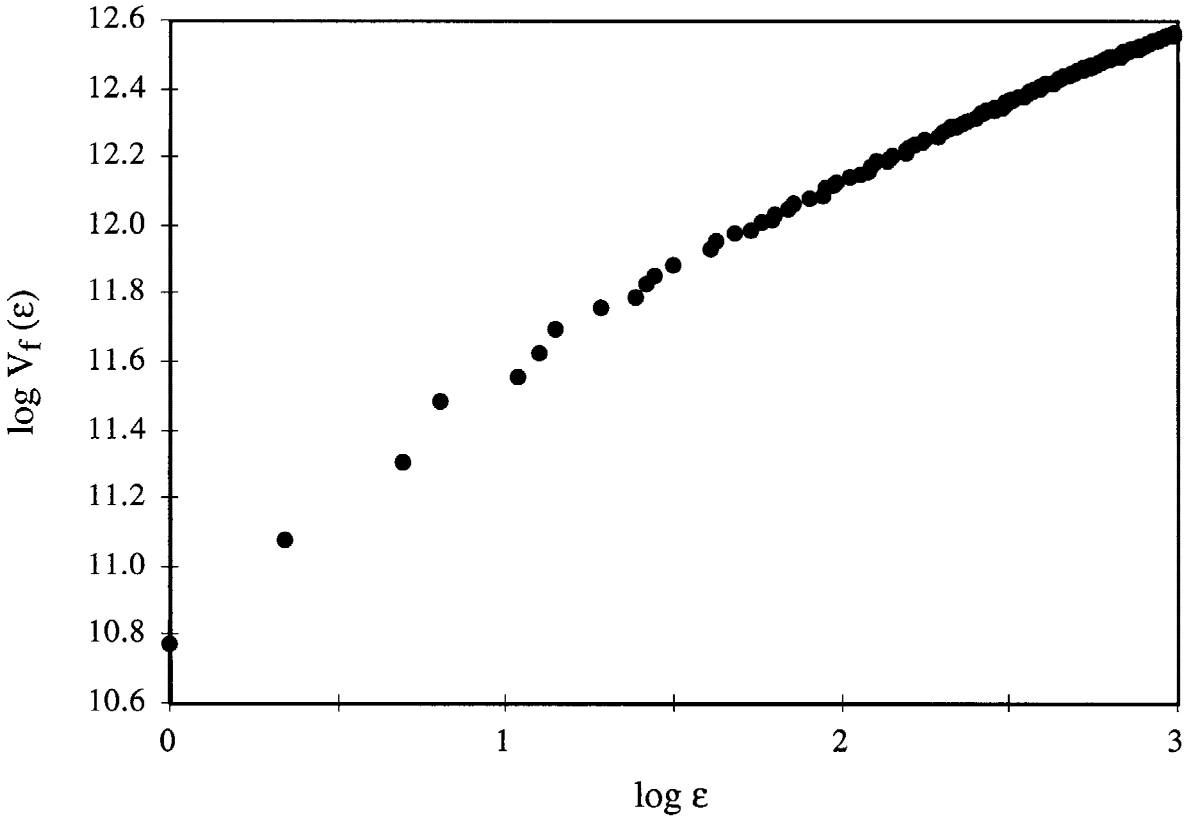

Fig. 3—Variation method plot used to determine the Minkowski dimension for a typical malignant breast epithelial cell nucleus (D_{MB}=2·49, r=0·991)

values of ε between 1 and 7 pixels (0·135 and 0·942 μm). This method is illustrated in Fig. 3.

Spectral dimensions—Spectral dimensions of the nuclei were determined using the iterative spectral dimension method. Spectral dimensions are based on Fourier analysis. The power spectrum $|Z(u,v)|^2$ of an image $z(x,y)$ is the square of the magnitude of its two-dimensional discrete Fourier tranform $Z(u,v)$. Fractal images are typified by a power spectrum in which there is a $1/f^\beta$ dependence on frequency,[28] i.e., $|Z|^2 = Cf^\beta$. In particular, Voss[29] has shown that for statistically self-affine fractal Brownian motion, the spectral exponent β is related to a fractal similarity dimension D_s by the equation $D_s = (7\text{-}\beta)/2$. However, since power spectra are determined from an image's two-dimensional Fourier transform, conventional methods for determining the spectral dimension of surface[28] require a rectangular image, which poses a problem since nuclei have irregular shapes. The requirement of rectangularity can be overcome using an iterative method to determine the spectral dimension.[30] This approach embeds a nuclear image in a rectangular image and iteratively fills in the remainder of the image so that it shares similar properties to the nucleus in the Fourier domain.

Lacunarity analysis

Fractal dimensions quantify textural complexity and irregularity and therefore can be used to discriminate between diagnostic categories for which these properties differ. However, they do not provide a unique description of an entire textured surface, for as is illustrated in Fig. 4, two fractals can have strikingly different appearances yet still possess the same fractal dimensions. Mandelbrot[34] introduced the notion of lacunarity to describe one particular aspect of the texture of a fractal: the largeness of its gaps or holes. As such, lacunarity would appear to describe features such as voids, chromatin clearings and nucleoli, which pathologists regard as important in the diagnosis of malignancy.

Several expressions for lacunarity have been suggested in the physics literature. The approach that we follow here is based on the gliding box method of Allain and Cloitre.[35] The gliding box method is designed for binary images, i.e., images in which all pixels take the value 0 or 1, and while it can be modified for grey scale data, this results in a loss of resolution.[36] Nuclear images can be binarized by thresholding: all pixels with grey values greater than a specified level take the value 1 (white), while the remaining pixels are assigned to 0 (black).

In applying the gliding box method to irregularly shaped images such as nuclei, complications may arise in determining the lacunarity; for this reason, we determine the *weighted* lacunarity.[37] The precise nature of these complications is beyond the scope of this paper, and the interested reader is referred elsewhere for an explanation, as well as for the mathematical derivation of and motivation for the weighted lacunarity formula.[37]

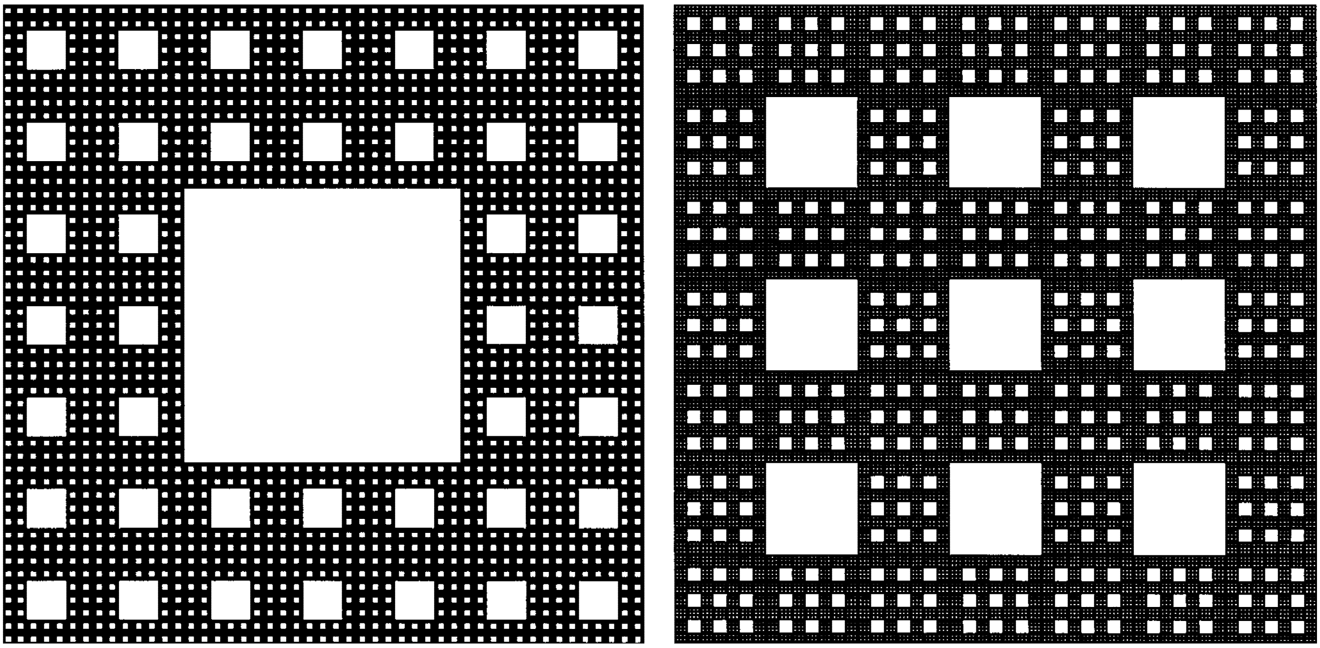

Fig. 4—Two mathematical fractals, called Sierpinski carpets, with identical fractal dimension but different appearance. A Sierpinski carpet[31] is constructed[32,33] by beginning with a square and subdividing it into b^2 subsquares, out of which l^2 subsquares are cut. This process can be iteratively repeated on each of the remaining subsquares, at each stage removing the l^2 subsquares from the same positions. The Figure shows three stages. It can be mathematically shown that the fractal (Hausdorf) dimension of a Sierpinski carpet is given by the formula[32] $D=\log(b^2-l^2)/\log(b)$. Since $b=7$ and $l=3$ for both carpets, they have a common fractal dimension of log 40/log 7=1·896. They differ in terms of which subsquares are removed. While fractal dimension is insufficient to discriminate between these carpets, the holes are larger in the carpet on the left, while the carpet on the right is more homogeneous. This difference can be quantified in terms of lacunarity

The gliding box method characterizes texture using an $s \times s$ pixel 'gliding box'. The box is initially placed at the upper left corner of the image. The number of white pixels in the image contained in the gliding box is counted. This value is denoted n_1 and can take any value from 0 to s^2. The gliding box is moved one pixel to the right, and the number of white pixels at this new location, denoted n_2, is counted. In a similar manner, the box glides over the entire image, moving to all N positions at which it covers at least one pixel of the image, at each location recording the number of white pixels, n_i, in the image. Additionally, at each position a weighting factor w_i, equal to the number of pixels in the gliding box that are part of the image, is recorded. Weighted lacunarity is defined as

$$\Lambda'(s)=\sum_{i=1}^{N} w_i \sum_{i=1}^{N} (n_i^2/w_i) \Bigg/ \left(\sum_{i=1}^{N} n_i\right)^2 .$$

Weighted lacunarity curves for the mathematical fractals in Fig. 4 are shown in Fig. 5. While these fractals have the same dimensions, they can be differentiated on the basis of the lacunarity curves. Thus, lacunarity complements fractal dimensions in characterizing texture. In fact, lacunarity analysis is a useful method not just for mathematical fractals, but for real-world data as well. Figure 6 illustrates a pair of nuclei with the same spectral fractal dimensions (2·36) and Minkowski dimensions (2·82) over the range 0·135–0·942 μm, together with their normalized weighted lacunarity curves. The nucleus on the left has a prominent nucleolus while the nucleus on the right does not, so we would expect the nucleus on the left to be more lacunar. This is confirmed by the lacunarity curves.

Statistical analysis

Statistical analysis was performed on an Apple Power Macintosh 8100/110 using JMP and Microsoft EXCEL.

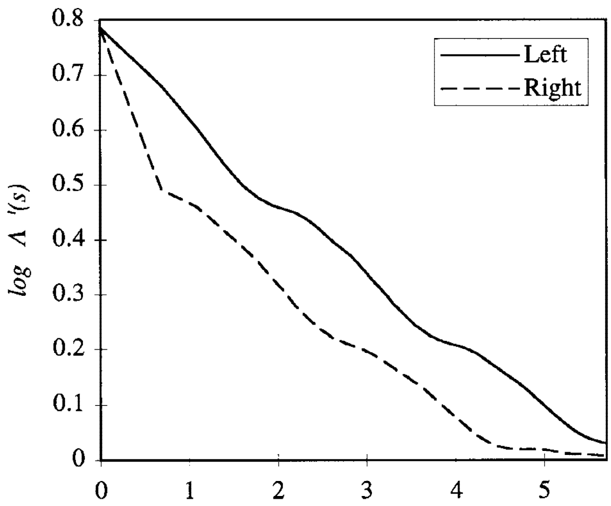

Fig. 5—Weighted lacunarity curves of the Sierpinski carpets in Fig. 4. The carpet on the left has larger holes and is therefore more lacunar

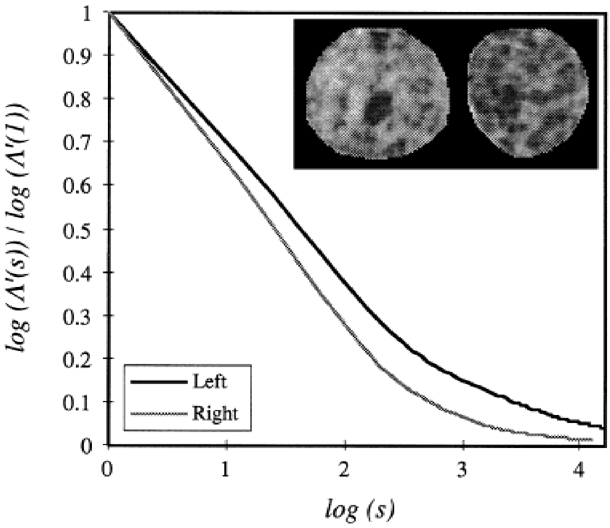

Fig. 6—Two nuclei with equal fractal dimensions and their weighted lacunarity curves. The nucleus on the left is more lacunar

Pearson correlation coefficients were determined for plots of log ε versus log $V_f(\varepsilon)$ to assess the appropriateness of the fractal model. Mean areas and Minkowski dimensions were determined for each patient. Mean spectral dimensions were determined only considering nuclei with fractal dimensions D_s between 2 and 3 and multiplicative prefactors C less than 2000. The 239 nuclei not meeting these criteria were considered to have converged incorrectly to nonsensical values and were therefore excluded from determinations of mean spectral dimensions. In addition, the area of each nucleus was determined. Fractal dimensions were compared between benign and malignant cases using two-way analysis of variance (ANOVA), treating diagnosis as a fixed effect and patient as a nested random effect. Correlation coefficients between Minkowski and spectral dimensions were determined. Mean areas and lacunarities at each box size and threshold were determined for each patient. Patient means of lacunarity were compared between benign and malignant cases at each box size and threshold using Student's t-test.

Classification approaches: logistic regression and neural networks

A variety of classificatory algorithms, such as logistic regression, discriminant analysis, and artificial neural networks, can be used to diagnose cases on the basis of morphometric descriptors.[38] Using a training set of cases with known features and diagnoses, a classification algorithm can be trained or optimized for the particular diagnostic setting of interest. This process involves determining an optimal set of features (e.g., mean spectral dimension and mean nuclear area) and parameters (e.g., regression coefficients in logistic regression or neuron weights and number of hidden neurons in an artificial neural network) in the classificatory model. Subsequently, unknown cases can be diagnosed using the trained classifier. A useful approach to evaluate the efficacy of a classificatory algorithm for a data set is to perform leave-one-out cross-validation.[39] This is often referred to as jack-knife analysis,[40] and here we use the terms interchangeably. Since here we consider 41 cases, a single run of a jack-knife analysis involves training a classifier on 40 of the cases and testing the accuracy of its diagnosis on the remaining 'masked case', repeating this process 41 times so that each case serves as the masked case.

We performed jack-knife analyses using logistic regression models[41] with a variety of fractal features as well as measures of nuclear area. The jack-knife analysis was repeated for each combination of features considered, except in a few instances where it was only necessary to consider a few patients to verify that a combination of features resulted in poor classificatory performance. In the terminology of logistic regression, the 'covariates' were the features while the 'independent variable' was the diagnosis. The fractal textural features included means and standard deviations of spectral and Minkowski–Bouligand fractal dimensions, of weighted lacunarity at a variety of box sizes, and of the multiplicative prefactor C, which can be viewed as a measure of lacunarity in the frequency domain. The weighted lacunarity measures considered included weighted lacunarity $\Lambda'(s)$, log weighted lacunarity log $[\Lambda'(s)]$ [also referred to as log $\Lambda'(s)$], normalized weighted lacunarity $\Lambda'(s)/\Lambda'(1)$ [also referred to as normalized $\Lambda'(s)$], and normalized log weighted lacunarity $\log[\Lambda'(s)]/\log[\Lambda'(1)]$ [also referred to as normalized log $\Lambda'(s)$]. The weighted lacunarity measures considered were further restricted on the basis of the results of the statistical comparisons of these features between benign and malignant cases. The measures of nuclear area considered were the mean, minimum, and standard deviation per patient. Logistic regression was performed using JMP (SAS Institute, Cary, NC, U.S.A.) on an Apple Power Macintosh 8100/110. The various classificatory models were evaluated in terms of accuracy, sensitivity, specificity, predictive values, and kappa (κ) statistics. In some models, the classifier was unable to predict the diagnosis for certain patients, resulting in three categories of 'diagnoses': benign (negative test), malignant (positive test), and no prediction (uncertain test); the determination of performance measures in such models follows the approach of Garcia-Romero *et al.*[42]

Jack-knife classification was also performed with back-propagation artificial neural networks.[43] General overviews of neural networks and their applications to pathology can be found in several recent reports.[40,44–46] Several neural network models were compared, differing from each other on the basis of architecture; the selection of input neurons, i.e., area and fractal texture features; and the coding of diagnoses in output neurons. Neural networks were trained using the Levenberg–Marquardt learning rule,[47–49] which is markedly faster than standard conjugate gradient methods, thus enabling more neural network models to be considered. The models were implemented in the MATLAB Neural Networks Toolbox (The MathWorks, Inc., Natick, MA, U.S.A.) on several Silicon Graphics computers, including a Power Challenger XL supercomputer.

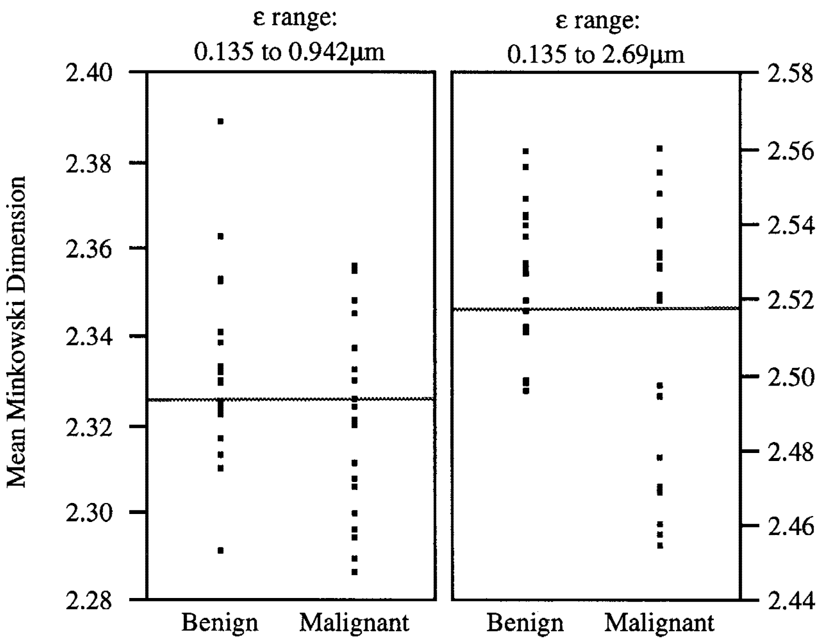

Fig. 7—Mean Minkowski dimensions for the 41 cases. Horizontal lines represent mean values

RESULTS

Fractality of chromatin texture

The log–log plots of $V_f(\varepsilon)$ vs. ε were linear, with correlation coefficients r greater than 0·95 in 98 per cent of the nuclei (median r=0·981, minimum r=0·914), as illustrated in Fig. 3, and fractal Minkowski dimensions of the nuclei were strictly greater than their topological dimensions. Thus, we conclude that the fractal model is an appropriate one; chromatin appearance in breast cytology specimens is fractal over the range of resolutions considered. As such, it is appropriate to use fractal dimensions as descriptors of chromatin appearance in image analysis. Over the smaller range of resolutions, the fit was even better, with a median r of 0·993 and a minimum of 0·985.

Fractal dimensions

Minkowski dimensions for the 41 cases are shown in Fig. 7. The patient means of Minkowski dimension averaged 2·527 for the benign cases and 2·510 for the malignancies. This difference was marginally significant (p=0·067). Over the smaller range, average patient means of Minkowski dimension were 2·333 for the benign cases and 2·319 for the malignant cases, a statistically significant difference (p=0·044). Thus, Minkowski dimensions were smaller over this range. The dependence of Minkowski dimension on the range of resolutions considered may suggest that fractal dimension changes depend on the range of resolutions, or may simply be a consequence of edge effects.

Spectral dimensions for the 41 patients are shown in Fig. 8. The average mean spectral dimension was 2·853 for the benign cases and 2·796 for the malignancies, a difference which was determined to be very highly significant (p=0·000011). Spectral and Minkowski dimensions for individual nuclei were weakly correlated (r= − 0·166 for the 2382 non-excluded nuclei), as were patient means of spectral and Minkowski dimensions (r= − 0·134). These correlations are illustrated in Fig. 9.

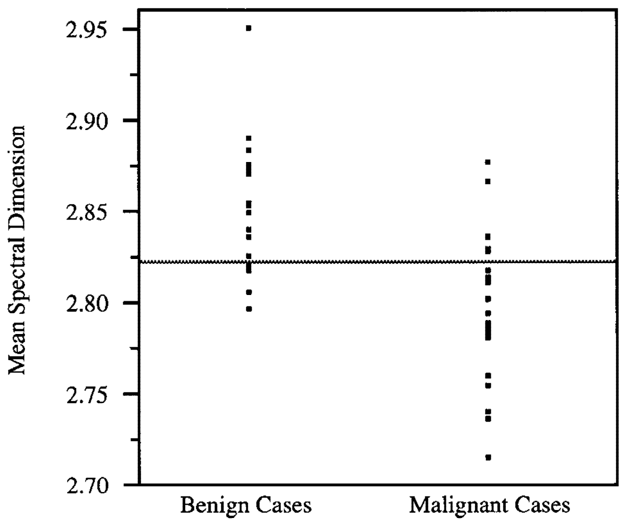

Fig. 8—Mean spectral dimensions for the 41 cases. The horizontal line represents the mean value

Lacunarity

Figure 10 illustrates weighted lacunarities of the 41 cases thresholded at the third quartile of the intensity histogram, as a function of the box size. At each scale, all of the most lacunar cases are seen to be malignancies, and patient means of weighted lacunarities are greater in

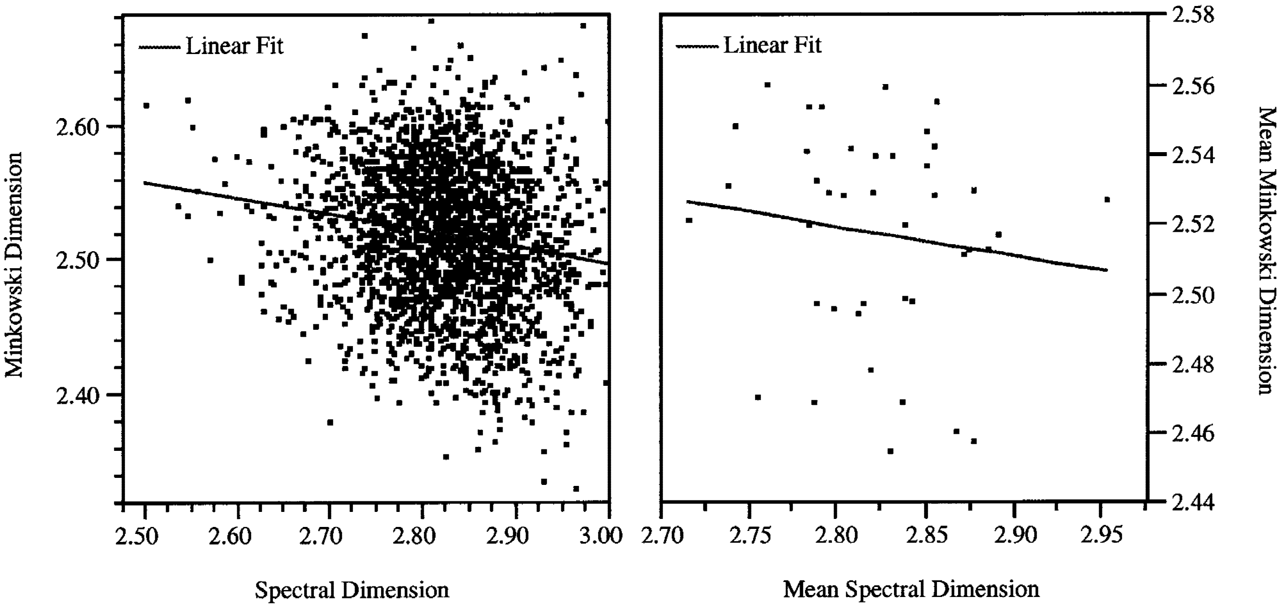

Fig. 9—Correlations between Minkowski and spectral dimensions

the malignant cases then in the benign ones. Figure 11 demonstrates the significance of these differences. For the images thresholded at the third quartile, the differences in patient means are significant at almost all box sizes and the maximum p value of 0·0505 at s=6 pixels (0·807 μm) is barely above the cut-off for significance. The difference is especially significant at the smallest and largest scales. These correspond respectively to fine-scale texture, such as whether chromatin is clumped or fine, and to large-scale structures like nucleoli. While increased nucleolar number or size is not pathognomonic for malignancy, nucleolar alterations provide diagnostically important information for Pap-stained breast cytology specimens:[50] eosinophilic macronucleoli[51] as well as excessive variability in size, shape, and numbers of nucleoli[52] tend to be associated with malignancy. Differences in weighted lacunarity between benign and malignant cases are not as pronounced when the images are segmented at the first and second quartiles of their intensity histograms. While the plot of p as a function of s follows the same shape at the three thresholds, p values are higher at the second quartile than at the third quartile, and even higher at the first quartile.

Logistic regression classification

The results of the jack-knife analyses of logistic regression classification performance are summarized in Tables I–III.

As lacunarity is not described by a single number as is, for example, spectral dimension, but rather by a function, it was necessary to restrict the lacunarity measures considered in the logistic regression models to a few well-chosen ones. Since Fig. 11 reveals that differences in

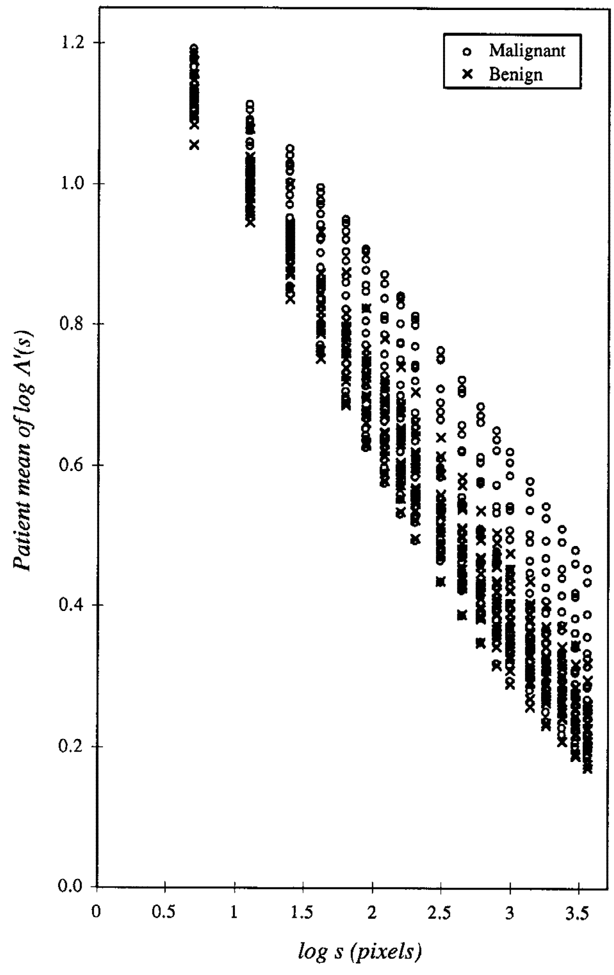

Fig. 10—Weighted lacunarities of the 41 cases as a function of the box size

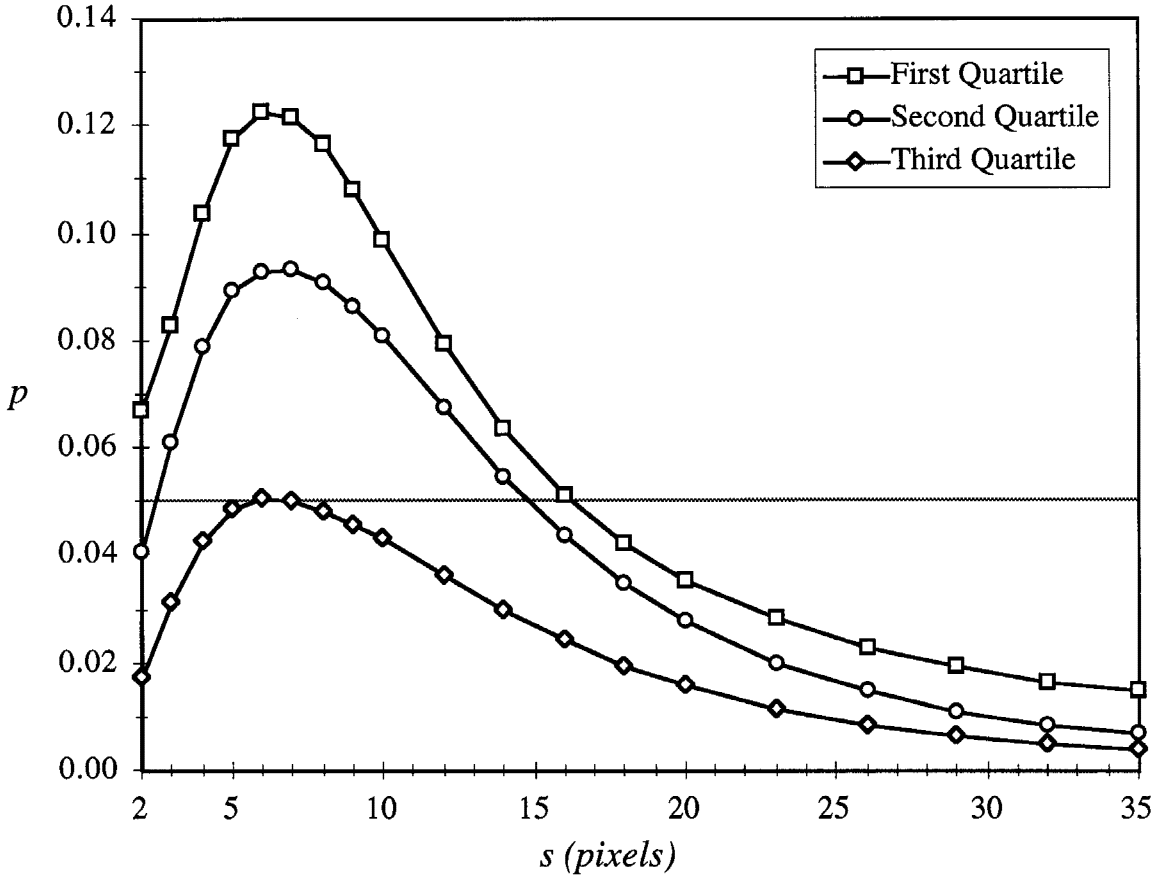

Fig. 11—p as a function of the box size s for nuclear images thresholded at the first, second, and third quartiles of the intensity histogram. p is the significance level of the t-test comparing the mean lacunarity per case of malignant and benign cases

Table I—Two-feature and three-feature logistic regression models

Features included			SE	SP	PPV	NPV	ACC	κ	Misclassified cases
			Performance measures						
$D_s(m)$	$\Lambda'(35)(m)$	C(sd)	0·82	0·79	0·82	0·79	0·80	0·61	2, 10, 17, 18, 21, 23, 36, 39
$D_s(m)$	Log $\Lambda'(35)(m)$	C(sd)	0·82	0·79	0·82	0·79	0·80	0·61	2, 10, 17, 18, 21, 23, 36, 39
$D_s(m)$	Normalized $\Lambda'(35)(m)$	C(sd)	0·82	0·79	0·82	0·79	0·80	0·61	2, 10, 17, 18, 21, 23, 36, 39
$D_s(m)$	Normalized log $\Lambda'(35)(m)$	C(sd)	0·95	0·84	0·88	0·94	0·90	0·80	10, 17, 18, 39
$D_s(m)$	Normalized log $\Lambda'(35)(m)$		0·86	0·79	0·83	0·83	0·83	0·66	2, 10, 17, 18, 23, 36, 39
$D_s(m)$		C(sd)	0·77	0·74	0·77	0·74	0·76	0·51	8, 10, 17, 18, 19, 32. 36, 37, 39, 41
	Normalized log $\Lambda'(35)(m)$	C(sd)	0·68	0·79	0·79	0·68	0·73	0·47	1, 2, 10, 16, 20, 23, 27, 33, 38, 39, 40

SE=sensitivity; SP=specificity; PPV=positive predictive value; NPV=negative predictive value; ACC=accuracy; D_s=spectral fractal dimension; (m)=mean; $\Lambda'(s)$=weighted lacunarity with a box size of s; C=multiplicative prefactor; (sd)=standard deviation.
Cases 1–19 are benign; cases 20–41 are malignant.

lacunarity between benign and malignant cases are more pronounced when images are thresholded at the third quartile of the intensity histogram than at the first or second quartile, the only lacunarity measures considered were those associated with third quartile thresholding. Further, since these differences are most pronounced at the smallest and largest scales considered (2 and 35 pixels), the only lacunarity measures incorporated into the classificatory models were for box sizes of 2 and 35 pixels.

Even having limited the number of lacunarity measures considered, it is still unfeasible to perform the jack-knife analysis using all possible combinations of features to determine the group of features for which performance is best. Instead, it is necessary to adopt a strategy to achieve good performance while limiting the number of regressions performed. We began with a model including as covariates, members of the three classes of textural descriptors for which the difference between benign and malignant cases was most

Table II—Four-feature logistic regression models

Features included				Performance measures						Misclassified cases	Non-convergent cases
				SE	SP	PPV	NPV	ACC	κ		
$D_s(m)$	Normalized $\Lambda'(35)(m)$	C(sd)	$\Lambda'(2)$(sd)	0·86	0·84	0·86	0·89	0·85	0·71	10, 12, 17, 27, 39	22
$D_s(m)$	Normalized $\Lambda'(35)(m)$	C(sd)	Log $\Lambda'(2)$(sd)	0·86	0·84	0·86	0·84	0·85	0·71	10, 12, 17, 22, 27, 39	
$D_s(m)$	Normalized $\Lambda'(35)(m)$	C(sd)	Normalized $\Lambda'(2)$(sd)	0·91	0·89	0·95	0·89	0·90	0·81	10, 23, 41	17
$D_s(m)$	Normalized $\Lambda'(35)(m)$	C(sd)	Normalized log $\Lambda'(2)$(sd)	0·86	0·84	0·90	1·00	0·85	0·73	10, 12	17, 23, 27, 41
$D_s(m)$	Normalized $\Lambda'(35)(m)$	C(sd)	Area (m)	0·86	0·89	0·95	0·85	0·88	0·76	17, 21, 26, 39	2
$D_s(m)$	Normalized log $\Lambda'(35)(m)$	C(sd)	$\Lambda'(2)$(sd)	0·82	0·84	0·86	0·80	0·83	0·66	10, 12, 17, 22, 27, 39, 41	
$D_s(m)$	Normalized log $\Lambda'(35)(m)$	C(sd)	Log $\Lambda'(2)$(sd)	0·82	0·79	0·82	0·79	0·80	0·61	10, 12, 17, 18, 22, 27, 39, 41	
$D_s(m)$	Normalized log $\Lambda'(35)(m)$	C(sd)	Normalized $\Lambda'(2)$(sd)	0·77	0·89	0·94	0·77	0·83	0·67	10, 20, 23, 28, 39, 41	17
$D_s(m)$	Normalized log $\Lambda'(35)(m)$	C(sd)	Normalized log $\Lambda'(2)$(sd)	0·77	0·84	0·85	0·76	0·80	0·61	10, 12, 17, 23, 27, 28, 39, 41	
$D_s(m)$	Normalized log $\Lambda'(35)(m)$	C(sd)	Area (m)	0·82	0·84	0·90	0·80	0·83	0·67	10, 17, 21, 26, 29, 39	2

SE=sensitivity; SP=specificity; PPV=positive predictive value; NPV=negative predictive value; ACC=accuracy; D_s=spectral fractal dimension; (m)=mean; $\Lambda'(s)$=weighted lacunarity with a box size of s; C=multiplicative prefactor; (sd)=standard deviation.

Cases 1–19 are benign; cases 20–41 are malignant.

For purposes of performance measures, benign non-convergent cases are counted as false positives and malignant non-convergent cases are counted as false negatives, following the approach of Garcia-Romero *et al.*[42]

Table III—Five-feature logistic regression models

Features included					Performance measures						Misclassified cases
					SE	SP	PPV	NPV	ACC	κ	
$D_s(m)$	Normalized $\Lambda'(35)(m)$	C(sd)	$\Lambda'(2)$(sd)	Area (m)	0·82	0·95	0·95	0·82	0·88	0·76	12, 21, 26, 27, 39
$D_s(m)$	Normalized $\Lambda'(35)(m)$	C(sd)	Log $\Lambda'(2)$(sd)	Area (m)	0·82	0·95	0·95	0·82	0·88	0·76	12, 21 27, 39, 40
$D_s(m)$	Normalized $\Lambda'(35)(m)$	C(sd)	Normalized $\Lambda'(2)$(sd)	Area (m)	0·82	1·00	1·00	0·83	0·90	0·81	20, 21, 23, 39
$D_s(m)$	Normalized $\Lambda'(35)(m)$	C(sd)	Normalized log $\Lambda'(2)$(sd)	Area (m)	0·91	1·00	1·00	0·90	0·95	0·90	21, 39
$D_s(m)$	Normalized log $\Lambda'(35)(m)$	C(sd)	$\Lambda'(2)$(sd)	Area (m)	0·77	0·95	0·94	0·78	0·85	0·71	12, 21, 26, 27, 39, 40
$D_s(m)$	Normalized log $\Lambda'(35)(m)$	C(sd)	Log $\Lambda'(2)$(sd)	Area (m)	0·00	0·00	0·00	0·00	0·00	− 0·04	*
$D_s(m)$	Normalized log $\Lambda'(35)(m)$	C(sd)	Normalized $\Lambda'(2)$(sd)	Area (m)	0·86	0·89	0·90	0·85	0·88	0·76	10, 12, 20, 21, 39
$D_s(m)$	Normalized log $\Lambda'(35)(m)$	C(sd)	Normalized log $\Lambda'(2)$(sd)	Area (m)	0·86	1·00	1·00	0·86	0·93	0·85	20, 21, 39

SE=sensitivity; SP=specificity; PPV=positive predictive value; NPV=negative predictive value; ACC=accuracy; D_s=spectral fractal dimension; (m)=mean; Λ'(s)=weighted lacunarity with a box size of s; C=multiplicative prefactor; (sd)=standard deviation.

Cases 1–19 are benign; cases 20–41 are malignant.

*Cases 12, 39, and 40 were misclassified, while no convergence was attained in the remainder of cases.

Table IV—Numbers of cases incorrectly diagnosed in jack-knife analyses using various artificial neural network models with the features mean D_s, mean normalized $\Lambda'(35)$, standard deviation of C, standard deviation of normalized log $\Lambda'(2)$, and mean area

No. of hidden neurons		2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
Output coding	0 to 1	7	10	8	6	7	10	12	8	11	11	11	12	8	8	10	8	9	7	11
	− 1 to 1	6	6	5	7	7	8	3	6	7	6	5	6	7	7	9	6	5	7	6
	− 1 to 1	6	4	7	5	8	6	2	6	6	8	6	4	5	6	6	8	7	5	6

statistically significant: mean spectral dimension, standard deviation of the multiplicative prefactor C, and mean lacunarities at a box size of 35 pixels. As is shown in Table I, the best performance among the three-feature models was attained by a model including mean spectral dimension, standard deviation of C, and mean normalized log $\Lambda'(35)$. For this model, there were three false-positive diagnoses and one false negative, so 37 of the 41 cases were correctly diagnosed. Classificatory performance was poorer in models including just two of these textural features.

Our next attempt to improve classification performance, summarized in Table II, was to add a fourth feature to the models, also chosen from those demonstrating a significant difference between benign and malignant cases. The added features included the standard deviations of the (four) weighted lacunarity functions at a box size of 2 pixels, and mean nuclear area. In one model there were only two false positives; however, in this model there were also four cases for which maximum likelihood estimates of parameters in the model failed to converge, making it impossible to classify the cases as benign or malignant. Interestingly, while the three-feature model including mean normalized log $\Lambda'(35)$ performed better than the three-feature model including mean normalized $\Lambda'(35)$, the opposite was true for the four-feature models, where each model including mean normalized $\Lambda'(35)$ outperformed its counterpart with mean normalized log $\Lambda'(35)$. This finding suggests that a stepwise approach[41] to feature selection is likely to miss optimal sets of features in such classifactory scenarios. Thus, it is necessary to use a more flexible approach, as we have done.

Continuing further, we considered five-feature models, adding both one of the four $\Lambda'(2)$ measures and mean area. They are summarized in Table III. The best performance was attained by a model including mean normalized $\Lambda'(35)$ and the standard deviation of normalized log $\Lambda'(2)$. In this 'best' model, all cases were classified correctly except for two false negatives. Additionally, other variations of five-feature models were considered, including features such as mean $\Lambda'(35)$, mean log $\Lambda'(35)$, Minkowski dimension, and minimum nuclear area, but none performed as well.

Finally, we considered some six-feature models, adding features such as Minkowski dimension and minimum nuclear area to the best five-feature model, but none improved on the performance of the best five-feature model. Thus, in this study, the highest accuracy attained using logistic regression to diagnose the breast cytology specimens was 95·1 per cent.

Artificial neural network classification

While the logistic regression models considered differed in terms of the features used as covariates, artificial neural network models can differ in terms of a host of parameters.[53] These include not only the features included as inputs, but also the number and arrangement of hidden layers, the coding of diagnoses, the transfer function(s) used, training and testing tolerances, the order in which training facts are presented, and the initial values chosen (typically at random) for the neuron weights. We began by considering neural networks with the five features in the 'best' logistic regression model assigned to input neurons. Output diagnoses were coded both on a 0 to 1 scale, where 0 corresponded to a benign and 1 to a malignant diagnosis, and also on a − 1 to 1 scale, where − 1 corresponded to a benign and 1 to a malignant diagnosis. The number of hidden neurons was varied systematically from 2 to 20. Initially, a variety of transfer functions were considered, including linear, log-sigmoid, and tan-sigmoid (hyperbolic tangent) transfer functions. However, in these initial runs, convergence of the neural network weights was reliably attained only when the transfer functions in both layers were tan-sigmoid, so subsequent runs were restricted to networks with this transfer function.

Jack-knife analyses for neural networks with these five features are summarized in Table IV. Every entry in the table represents the number of incorrect diagnoses for a single jack-knife analysis with the specified output coding and number of hidden neurons. It is evident from the Table that better performance was attained using the − 1 to 1 scale. Another feature of neural network training illustrated by the Table is that repeating a jack-knife analysis need not yield identical performance. Due to the different initial values chosen at random for the neuron weights, repeat analyses resulted in differences of up to three misclassifications. Nevertheless, the two jack-knife analyses with the highest classificatory accuracies were the two runs with eight hidden neurons. These resulted in two and three of the 41 cases misclassified, a performance comparable to that in the best logistic regression classification model.

Since the best features for neural network classification need not be the same as those for logistic regression

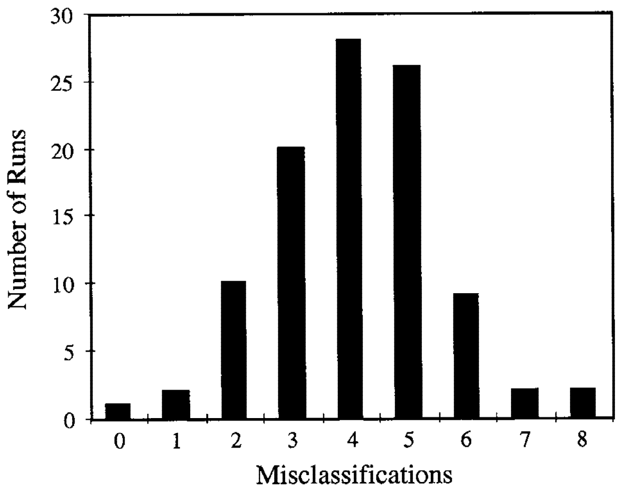

Fig. 12—Histogram of the number of misclassifications out of 41 cases for a jack-knifed neural network model with eight hidden neurons, tan-sigmoid transfer functions, and input neurons corresponding to mean spectral dimension, mean area, mean normalized $\Lambda'(35)$, and standard deviation of normalized $\log \Lambda'(2)$. One hundred observations; mean=4·08 misclassifications; standard deviation=1·44 misclassifications

classification, we next considered neural network models with a variety of features. Two hundred combinations of features, denoted 'feature sets', were considered. Each of these feature sets included mean spectral dimension and mean nuclear area, and from zero to five additional features:

(1) mean Minkowski dimension (yes/no): two options
(2) standard deviation (sd) of spectral dimension (yes/no): two options
(3) mean multiplicative prefactor C (yes/no): two options
(4) mean of (log/not log) (normalized/unnormalized) $\Lambda'(35)$ (yes/no): five options
(5) sd of (log/not log) (normalized/unnormalized) $\Lambda'(2)$ (yes/no): five options

Jack-knife analyses were performed using neural networks with eight hidden neurons, tan-sigmoid transfer functions, and each of the 200 feature sets. For two of the feature sets, only one of the 41 cases was misclassified, while for 11 feature sets there were two misclassifications, as in the 'best' feature set for logistic regression.

Since Table IV showed that there is some inter-run variability in neural network accuracy, owing to the dependence of the neuron weights converged upon on their randomly chosen initial values, we next repeated the jack-knife analyses for the 13 feature sets with one or two misclassifications. For each feature set, the jack-knife analysis was repeated 100 times. As observed previously, the number of misclassifications varied from run to run. Figure 12 is a histogram of the number of misclassifications for 100 jack-knife analyses using a neural network with mean spectral dimension, mean area, mean normalized $\Lambda'(35)$, and sd of normalized $\log \Lambda'(2)$ as features. While in the original run using this feature set there were two misclassifications, over the 100 runs there were four misclassifications on average. However, for one run there were no misclassifications, and for two runs there was only a single misclassified case. The performance was typical for the 13 feature sets repeated.

DISCUSSION

Although biological structures have certain characteristic size scales, some aspects of morphology are better described in terms of fractal geometry than in terms of Euclidean geometry. Our study shows that chromatin appearance in breast epithelial cell nuclei has fractal properties and can be described by fractal dimensions and lacunarity measurements, which differ between patients with benign and malignant breast lesions. These fractal descriptors can be used in the diagnosis of cytological specimens. That nuclear appearance is fractal is highly suggestive that three-dimensional chromatin structure is also fractal, as has been hypothesized by Orlando and Paro.[54] Supportive of this, Pentland[55] has mathematically proven that under certain assumptions, a three-dimensional surface is fractal if and only if its image intensity surface is fractal.

While for some types of mathematical objects, different fractal dimensions are necessarily equivalent, the two approaches to characterize fractal dimension that we considered here were remarkably different. Spectral dimensions were considerably greater then Minkowski dimensions and there was only a weak correlation between these two indices of fractal structure. Thus, the two approaches seem to measure different aspects of the fractal nature of nuclear chromatin, underscoring the fact that no single parameter can completely describe the fractal nature of biological structure. This observation is consistent with the results obtained in geology by Cox and Wang,[56] who found that fractal dimension may vary systematically, depending on the measurement method. As Mandlebrot[25] has recently commented, while initially a pre-eminent position was given to Hausdorf–Besicovitch dimension, it is now clear that fractal dimension is a multifaceted concept. The different computational methods need not even theoretically yield the same value for fractal dimension, so a surface may have several distinct fractal dimensions.[25]

As shown above, lacunarity also provides useful diagnostic information, complementing fractal dimensions in characterizing fractal properties of chromatin texture. Optimal feature sets for both logistic regression and neural network classifiers included lacunarity features, suggesting that fractal dimensions are insufficient to characterize entirely the fractality of nuclear texture. As mentioned above, the difference in weighted lacunarity between benign and malignant cases was especially significant on scales corresponding to fine-scale texture, such as chromatin coarseness, and to large-scale structures like nucleoli. The ubiquity of these qualitative features in pathological diagnosis suggests that lacunarity, as a numerical measure of them, should have a broad range of applications in quantitative diagnostic pathology.

The logistic regression model with the optimal feature set resulted in classificatory accuracy exceeding 95 per cent. This model shows particular promise and would be a good candidate for clinical trials to evaluate the diagnostic potential of fractal measures of chromatin appearance. The presence of jack-knife runs with zero and one misclassifications indicates that an optimized neural network can improve on logistic regression classification. However, this improved performance requires further validation, since it could be attributable to overtraining.[38,57] Neural network overtraining is most commonly manifested in terms of a disparity in classifier performance between training and test sets. Two competing factors are at work in neural network training: (1) the network learns general characteristics of the classificatory task, and (2) the network learns particular characteristics of the members of the training set. As the criterion for ending neural network training becomes more stringent and neural network training is allowed to proceed longer, the balance generally shifts towards the latter factor.[57–59] Consequently, the performance on new cases, such as the masked case in our jack-knife analysis, for which the particular characteristics of the training set are not applicable, can be markedly reduced.

This 'classical' form of overtraining is not present in our study. The optimized neural network achieved a 100 per cent classification rate of test cases on which it had not been trained. Nevertheless, a more subtle form of overtraining may come into play. By considering many neural network models, the networks at each step in a single jack-knife analysis may still focus on particular characteristics of the masked case. Although a network cannot learn particulars of the masked case through its training process, it may effectively 'learn' these particulars through a process of selection. By considering for each case 100 nearly identical versions of the same network, differing only in the randomly selected initial weights, some of these may be predisposed to converging to neuron weights resulting in a correct classification for the masked case. A jack-knife analysis in which such neural networks are trained for all of the most diagnostically challenging cases may exhibit perfect classificatory performance, although some of the networks trained in this jack-knife run may not correctly diagnose further new cases. The scope of this potential problem remains unclear and the issue of overtraining in repeated jack-knife analyses requires further study. While literature exists on neural network generalization, these studies typically address the classical form of overtraining. For practical purposes, the generalizability of the 'optimal' neural networks trained in the best jack-knife analyses could be evaluated through prospective testing on new cases.

Thus, this study shows how fractal texture features can be used in the cytological diagnosis of breast cancer to attain a diagnostic accuracy of 95·1 per cent using logistic regression, and potentially approaching 100 per cent using neural networks. Other investigators have also developed image analysis-based systems for the cytological diagnosis of breast epithelial cell lesions. King *et al.*,[60] using texton[5]-based measures of chromatin texture[61] and stepwise discriminant analysis for classification, attained an accuracy of 92·3 per cent but a sensitivity of only 78·6 per cent. Hutchinson *et al.*[62] used run-length[4] texture features as well as features from a low-resolution contextual analysis, combined with stepwise discriminant analysis for classification. Excluding fibroadenoma cases, this method correctly classified 93 per cent of cases. Wolberg *et al.*[63,64] quantified texture using a measure of grey scale variation, and classified cases using MSM-Tree, a decision tree method. Ten-fold cross-validation, which is comparable to the jack-knife analysis but excludes one-tenth of the cases from the training set at a time, rather than a single case at a time, estimated accuracy at 97·2 per cent; a subsequent prospective study correctly classified all the benign and malignant cases. However, not all the benign masses in this study were histologically confirmed.

Recently there has been considerable interest within the cytology community in image analysis devices for quality assurance, particularly in the area of cervical cytology, where the United States Food and Drug Administration has approved two instruments for rescreening.[65,66] A review[67] of 83 published series on fine-needle aspiration (FNA) biopsy of the breast reported false-negative rates ranging from 0 to 35 per cent, indicating that breast FNA diagnostic performance varies widely between institutions. Thus, ancillary image analysis techniques could serve an important role for quality assurance in this setting as well. Although diagnostic accuracy has been quite high in pilot studies using image analysis for breast FNA diagnosis, it is commonly observed that test efficacy in practice is less than that observed in pilot studies; this phenomenon may be attributable to a number of reasons, such as the broader spectrum of disease observed in practice and the differing conditions under which a test may be administered.[68–70] Ultimately, the most effective approach may well be some hybrid combining the best features of the various methods developed to date. In this study, we limited the number of features considered, since our purpose was to test the usefulness of fractal texture descriptors. Nevertheless, the inclusion of other descriptors may contribute to higher diagnostic performance in practice. Observations in past studies suggest that, in many cases, features based on the size distribution of nuclear areas are extremely useful in cytological diagnosis.[71] Other useful features may include contextual descriptors, such as those considered by Hutchinson *et al.*,[62] and other classes of texture descriptors. Similarly, the incorporation of fractal texture features could improve the performance of the classificatory models in other studies. A further area in which more sophisticated classificatory models may be needed is in the diagnosis of borderline epithelial lesions of the breast, where even the histopathological diagnosis is fraught with high inter-observer variability.[72] Ongoing efforts are aimed at developing image analysis techniques to assist cytologists in making this subtle distinction.

ACKNOWLEDGEMENTS

We thank Drs Craig Benham, Joseph Malinsky, Sylvan Wallenstein, and Carol Bodian of the Department of Biomathematical Sciences, Mount Sinai School of Medicine, for their helpful advice and discussion. We also thank our pathology colleagues: Dr Ira Bleiweiss for histopathological diagnoses; Dr David Burstein for cytological diagnoses; Dr Grace Yang (NYU) and Dr Syed Hoda (Cornell) for providing materials; and Joseph Silberfarb and Cynthia Sheppard for technical assistance. This material was presented in part at the Second International Symposium on Fractals in Biology and Medicine, Ascona, Switzerland, 6–9 March 1996. Andrew Einstein was supported by a traineeship on NIH MSTP Training Grant GM07280 and by a fellowship from the Louis and Rachel Rudin Foundation.

REFERENCES

1. Doudkine A, Macaulay C, Poulin N, Palcic B. Nuclear texture measurements in image cytometry. *Pathologica* 1995; **87:** 286–299.
2. Haralick RM, Shanumgam K, Dinstein I. Textual features for image classification. *IEEE Trans Syst Man Cybern* 1973; **SMC-3:** 610–621.
3. Pressman NJ. Markovian analysis of cervical cell images. *J Histochem Cytochem* 1976; **24:** 138–144.
4. Galloway MM. Texture analysis using gray level run lengths. *Comput Graph Image Process* 1975; **4:** 172–179.
5. Julesz B. Textons, the elements of texture perception, and their interactions. *Nature* 1981; **290:** 91–97.
6. Brigham E. Fast Fourier Transform and Its Applications. Englewood Cliffs, NJ: Prentice Hall, 1988.
7. Serra J. Image Analysis and Mathematical Morphology. London: Academic Press, 1982.
8. Einstein AJ, Barba J, Unger PD, Gil J. Nuclear diffuseness as a measure of texture: definition and application to the computer-assisted diagnosis of parathyroid adenoma and carcinoma. *J Microsc* 1994; **176:** 158–166.
9. Diegenbach PC, Baak JPA. Quantitative nuclear image analysis: differentiation between normal, hyperplastic, and malignant appearing uterine glands in a paraffin section. III. The use of texture features for differentiation. *Eur J Obstet Gynecol Reprod Biol* 1978; **8:** 109–116.
10. Beil M, Irinopoulou T, Vassy J, Rigaut JP. Chromatin texture analysis in three-dimensional images from confocal scanning laser microscopy. *Anal Quant Cytol Histol* 1995; **17:** 323–331.
11. Mandelbrot BB. The Fractal Geometry of Nature. New York: W. H. Freeman, 1983.
12. Cross SS. Fractals in pathology. *J Pathol* 1997; **182:** 1–8.
13. Losa GA, Nonnenmacher TF. Self-similarity and fractal irregularity in pathologic tissues. *Mod Pathol* 1996; **9:** 174–182.
14. Losa GA, Baumann G, Nonnenmacher TF. Fractal dimension of pericellular membranes in human lymphocytes and lymphoblastic leukemia cells. *Pathol Res Pract* 1992; **188:** 680–686.
15. Smith TG Jr, Marks WB, Lange GD, Sheriff WH Jr, Neale EA. A fractal analysis of cell images. *J Neurosci Methods* 1989; **27:** 173–180.
16. Landini G, Rippin JW. An 'asymptotic fractal' approach to the morphology of malignant cell nuclei. *Fractals* 1993; **1:** 326–335.
17. Cross SS, Bury JP, Silcocks PB, Stephenson TJ, Cotton DW. Fractal geometric analysis of colorectal polyps. *J Pathol* 1994; **172:** 317–323.
18. Landini G, Rippin JW. How important is tumour shape? Quantification of the epithelial–connective tissue interface in oral lesions using local connected fractal dimension analysis. *J Pathol* 1996; **179:** 210–217.
19. Fazzalari NL, Parkinson IH. Fractal dimension and architecture of trabecular bone. *J Pathol* 1996; **178:** 100–105.
20. Vilela MJ, Martins ML, Boschetti SR. Fractal patterns for cells in culture. *J Pathol* 1995; **177:** 103–107.
21. Matsuyama T, Matsushita M. Fractal morphogenesis by a bacterial cell population. *Crit Rev Microbiol* 1993; **19:** 117–135.
22. Yang GC, Alvarez II. Ultrafast Papanicolaou stain. An alternative preparation for fine needle aspiration cytology. *Acta Cytol* 1995; **39:** 55–60.
23. Einstein AJ, Gil J, Wallenstein S, *et al.* Reproducibility and accuracy of interactive segmentation procedures for image analysis in cytology. *J Microsc* 1997; **188:** 136–148.
24. Russ JC. Fractal Surfaces. New York: Plenum Press, 1994.
25. Mandelbrot BB. A fractal's lacunarity, and how it can be tuned and measured. In: Nonnenmacher TF, Losa GA, Weibel ER, eds. Fractals in Biology and Medicine. Basel: Birkhäuser Verlag, 1993; 8–21.
26. Dubuc B, Roques-Carmes C, Tricot C, Zucker SW. The variation method: a technique to estimate the fractal dimension of surfaces. *SPIE Visual Commun Image Processing II* 1987; **845:** 241–248.
27. Dubuc B, Zucker SW, Tricot C, Quiniou JF, Wehbi D. Evaluating the fractal dimension of surfaces. *Proc R Soc London Ser A* 1989; **425:** 113–127.
28. Anguiano E, Pancorbo M, Aguilar M. Fractal characterization by frequency analysis. I. Surfaces. *J Microsc* 1993; **172:** 223–232.
29. Voss RF. Random fractal forgeries. In: Earnshaw RA, ed. Fundamental Algorithms for Computer Graphics. Berlin: Springer-Verlag, 1985; 805–835.
30. Wu H-S, Gil J, Einstein AJ. Fractal characterization of irregularly shaped images by frequency analysis. *Fractals* 1997; **5:** 665–672.
31. Sierpinski W. Sur une courbe cantorienne qui contient une image biunivoque et continue de toute courbe donneé. *C R Acad Sci* 1916; **162:** 629–632.
32. Gefen Y, Aharony A, Mandelbrot BB. Phase transitions on fractals: III. Infinitely ramified lattices. *J Phys A: Math Gen* 1984; **17:** 1277–1289.
33. Lin B, Yang ZR. A suggested lacunarity expression for Sierpinski carpets. *J Phys A: Math Gen* 1986; **19:** L49–L52.
34. Mandelbrot B. Corrélations et texture dans un nouveau modèle d'Univers hiérarchisé, basé sur les ensembles trémas. *C R Acad Sci Ser A* 1979; **288:** 81–83.
35. Allain C, Cloitre M. Characterizing the lacunarity of random and deterministic fractal sets. *Phys Rev A* 19991; **44:** 3552–3558.
36. Plotnick RE, Gardner RH, Hargrove WW, Prestegaard K, Perlmutter M. Lacunarity analysis: a general technique for the analysis of spatial patterns. *Phys Rev E* 1996; **53:** 5461–5468.
37. Einstein AJ, Wu H-S, Gil J. Self-affinity and lacunarity of chromatin texture in benign and malignant breast epithelial cell nuclei. *Phys Rev Lett* 1998; **80:** 397–400.
38. Einstein AJ, Gil J. Classification procedures for diagnosis based on multiple morphometric parameters. *Acta Stereol* 1996; **15:** 15–24.
39. Efron B, Tibshirani RJ. An Introduction to the Bootstrap. New York: Chapman & Hall, 1993.
40. Kattan MW, Beck JR. Artificial neural networks for medical classification decisions. *Arch Pathol Lab Med* 1995; **119:** 672–677.
41. Hosmer DW, Lemeshow S. Applied Logistic Regression. New York: John Wiley, 1989.
42. Garcia-Romero H, Garcia-Barrios C, Ramos-Gutierrez F. Effects of uncertain results on sensitivity and specificity of diagnostic tests. *Lancet* 1996; **348:** 1745–1746.
43. Rumelhart DE, Hinton GE, Williams RJ. Learning representations by back-propagating errors. *Nature* 1986; **323:** 533–536.
44. Dytch HE, Wied GL. Artificial neural networks and their use in quantitative pathology. *Anal Quant Cytol Histol* 1990; **12:** 379–393.
45. Deligdisch L, Einstein AJ, Guera D, Gil J. Ovarian dysplasia in epithelial inclusion cysts. A morphometric approach using neural networks. *Cancer* 1995; **76:** 1027–1034.
46. Cross SS, Harrison RF, Kennedy RL. Introduction to neural networks. *Lancet* 1995; **346:** 1075–1079.
47. Levenberg K. A method for the solution of certain non-linear problems in least squares. *Q Appl Math* 1944; **2:** 164–168.
48. Marquardt DW. An algorithm for least-squares estimation of nonlinear parameters. *J Soc Indust Appl Math* 1963; **11:** 431–441.
49. Hagan MT, Manhaj MB. Training feedforward networks with the Marquardt algorithm. *IEEE Trans Neural Net* 1994; **5:** 989–993.
50. Oertel YC. Fine Needle Aspiration of the Breast. Boston: Butterworth, 1987.
51. Kline TS, Kline IK. Guides to Clinical Aspiration Biopsy: Breast. New York: Igaku-Shoin, 1989.
52. Frost JK. The Cell in Health and Disease: An Evaluation of Cellular Morphologic Expression of Biologic Behavior. Basel: Karger, 1969.
53. Truong H, Morimoto R, Walts AE, Erler B, Marchevsky A. Neural networks as an aid in the diagnosis of lymphocyte-rich effusions. *Anal Quant Cytol Histol* 1995; **17:** 48–54.
54. Orlando V, Paro R. Chromatin multiprotein complexes involved in the maintenance of transcription patterns. *Curr Opin Genet Dev* 1995; **5:** 174–179.
55. Pentland AP. Fractal-based description of natural scenes. *IEEE Trans Pattern Anal Machine Intell* 1984; **PAMI-6:** 661–674.
56. Cox BL, Wang JSY. Fractal surfaces: measurement and applications in the earth sciences. *Fractals* 1993; **1:** 87–115.
57. Astion ML, Wilding P. The application of backpropagation neural networks to problems in pathology and laboratory medicine. *Arch Pathol Lab Med* 1992; **116:** 995–1001.
58. Amari S, Murata N, Müller K-R, Finke M, Yang H. Statistical theory of overtraining—is cross-validation asymptotically effective? In: Touretzky DS, Mozer MC, Masselmo ME, eds. Advances in Neural Information Processing Systems. Vol. 8. Cambridge, MA: MIT Press, 1996; 176–182.
59. Chauvin Y. Generalization dynamics in LMS trained linear networks. In: Lippmann RE, Moody JE, Touretzky DS, eds. Advances in Neural Information Processing Systems. Vol. 3. San Mateo, CA: Morgan Kauffman, 1991; 890–896.
60. King EB, Chew KL, Duarte L, *et al.* Image cytometric classification of premalignant breast disease in fine needle aspirates. *Cancer* 1988; **62:** 114–124.

61. Young IT, Verbeek PW, Mayall BH. Characterization of chromatin distribution in cell nuclei. *Cytometry* 1986; **7:** 467–474.
62. Hutchinson ML, Isenstein LM, Zahniser DJ. High-resolution and contextual analysis for the diagnosis of fine needle aspirates of breast. *Anal Quant Cytol Histol* 1991; **13:** 351–355.
63. Wolberg WH, Street WN, Mangasarian OL. Image analysis and machine learning applied to breast cancer diagnosis and prognosis. *Anal Quant Cytol Histol* 1995; **17:** 77–87.
64. Wolberg WH, Street WN, Mangasarian OL. Machine learning techniques to diagnose breast cancer from image-processed nuclear features of fine needle aspirates. *Cancer Lett* 1994; **77:** 163–171.
65. Wilbur DC, Bonfiglio TA, Rutkowski MA, *et al*. Sensitivity of the AutoPap 300 QC System for cervical cytologic abnormalities. Biopsy data confirmation. *Acta Cytol* 1996; **40:** 127–132.
66. Koss LG, Lin E, Schreiber K, Elgert P, Mango L. Evaluation of the PAPNET cytologic screening system for quality control of cervical smears. *Am J Clin Pathol* 1994; **101:** 220–229.
67. Layfield LJ, Glasgow BJ, Cramer H. Fine-needle aspiration in the management of breast masses. *Pathol Annu* 1989; **24:** 23–62.
68. Ransohoff DF, Feinstein AR. Problems of spectrum and bias in evaluating the efficacy of diagnostic tests. *N Engl J Med* 1978; **299:** 926–930.
69. Begg CB. Biases in the assessment of diagnostic tests. *Stat Med* 1987; **6:** 411–423.
70. Einstein AJ, Bodian CA, Gil J. The relationships among performance measures in the selection of diagnostic tests. *Arch Pathol Lab Med* 1997; **121:** 110–117.
71. Unger PD, Watson CW, Liu Z, Gil J. Morphometric analysis of neoplastic renal aspirates and benign renal tissue. *Anal Quant Cytol Histol* 1993; **15:** 61–66.
72. Rosai J. Borderline epithelial lesions of the breast. *Am J Surg Pathol* 1991; **15:** 209–221.