Journal of Cellular Biochemistry Supplements 28/29:21–38 (1997)

Quantitation of Preinvasive Neoplastic Progression in Animal Models of Chemical Carcinogenesis

James W. Bacus,[1]* James V. Bacus,[1] Gary D. Stoner,[2] Richard C. Moon,[3] Gary J. Kelloff,[4] and Charles W. Boone[4]

[1]Bacus Laboratories, Inc., Elmhurst, Illinois
[2]Ohio State University Comprehensive Cancer Center, Arthur G. James Cancer Hospital, Columbus, Ohio
[3]University of Illinois, Specialized Cancer Center, Chicago, Illinois
[4]Chemoprevention Branch, Division of Cancer Prevention and Control, National Cancer Institute, National Institutes of Health, Bethesda, Maryland

Abstract An assay method that precisely quantitates the cellular and tissue changes associated with early, preinvasive neoplasia is much needed as a surrogate endpoint biomarker (SEB) in clinical trials to predict the potential efficacy of chemopreventive agents in bringing about cancer incidence reduction. Quantification of histological changes at the tissue level are potentially powerful SEB's since these visually apparent changes are common in all neoplastic development, regardless of tissue type or neoplastic cause. Currently, subjective inspection of the histological appearance of sectioned and stained material, or ''grading,'' by experienced pathologists is used to evaluate neoplastic progression. This has well-known limitations of reproducibility, accuracy, and resolution of grading scale. Since neoplastic changes are visually apparent and morphologic in nature, quantification by image analysis is a measurement modality of choice.

Image analysis was implemented through the use of high-resolution ''tiled'' images of complete tissue sections. A histological grading system, or ''scale,'' was developed that could be expressed in terms of normal deviate units of multiple and different morphometric descriptors. Neoplastic growth was characterized quantitatively with multiple measurements on each tissue image tile, which were combined into a single number for each tile, i.e., a histologic grade per tile, and parameters from the distributions of these measurements were used to represent the histologic grade for the entire region considered. This concept provided a uniform final scale in similar units of measurement, regardless of which tissues were graded. Also, the grading scale automatically adjusted measurement variance for different tissues by using normal tissue for each different type to obtain the normalization to standard deviation (z) units. This further defined a uniform final scale and maintained standard references.

Using this method, results from two well-known animal models of carcinogenesis, squamous cell carcinoma of SENCAR mouse skin induced by benzo(*a*)pyrene (B[*a*]P), and squamous cell carcinoma of the rat esophagus induced by *N*-nitrosomethylbenzylamine (NMBA), were compared to each other. Image analysis was performed on skin tissue sections from a total of 64 SENCAR mice, and esophagus tissue sections from 96 Fischer-344 rats. In both cases, a quantitative expression of the preinvasive neoplastic response to the carcinogen as a function of time of exposure was expressed along a continuous grading scale in standard deviation units (z) . In the SENCAR mouse skin animal model, similar cohorts of 4 mice at 20 weeks showed significant modulation of B[*a*]P-induced neoplasia by treatment with the antiproliferative agent difluoromethylornithine, $P < .05$. In the rat esophagus animal model, similar cohorts of 6 rats at 10 and 15 weeks showed significant modulation of NMBA-induced neoplasia by treatment with the antimutagen phenethyl isothiocyanate, $P < .05$. J. Cell. Biochem. Suppls. 28/29:21–38. © 1998 Wiley-Liss, Inc.

Key words: preinvasive neoplasia; image analysis; chemical carcinogenesis; rat; mouse

Cancer chemoprevention attempts to prevent clinical cancer by the administration of synthetic chemical or dietary constituents prior to or during the early phases of preinvasive (intraepithelial) neoplasia, in an attempt to block or delay progression to invasion [1] (Fig. 1). An assay method to precisely quantitate the cellular and tissue changes associated with early, preinvasive neoplasia is much needed as a surrogate endpoint biomarker in clinical trials to predict the potential efficacy of chemopreventive agents in bringing about cancer incidence

*Correspondence to: James W. Bacus, PhD, Bacus Laboratories, Inc., 910 Riverside Drive, Unit 8A, Elmhurst, IL 60126.
Received 3 July 1997; Accepted 1 October 1997

reduction [2]. This paper presents such an assay method, using computer-assisted image analysis,which is generally applicable to tissues of any organ system, including breast, prostate, colon, aerodigestive tract, etc., in humans and in animal models of carcinogenesis.

Two well-known animal models of carcinogenesis were chosen for study: squamous cell carcinoma of SENCAR mouse skin induced by painting with benzo(*a*)pyrene (B[*a*]P) [3], and squamous cell carcinoma of the rat esophagus induced by subcutaneous administration of *N*-nitrosomethylbenzylamine (NMBA) [4]. To evaluate the ability of imaging methods to measure carcinogen-induced nuclear and histological changes inhibited or reduced by a chemopreventive agent, the antiproliferative difluoromethylornithine (DFMO) was tested in the SENCAR mouse skin model [5], and the antimutagen phenethyl isothiocyanate (PEITC) was tested in the rat esophagus model [6]. Both chemopreventive agents were administered orally. A main goal of these studies was to develop surrogate endpoint biomarker assays for animal model studies and for human clinical chemoprevention trials, which are not only more quantitative and precise, thereby requiring fewer subjects to reach statistical significance, but also more sensitive, analyzing earlier lesions in the intraepithelial neoplastic process.

METHODS

Experimental Design and Specimen Preparation

For the rat esophagus study, 96 male Fischer-344 rats were acclimated on AIN-76A diet from Teklad, Madison, WI for a week, then randomly assigned to four groups of 24 animals, each group being further divided so that 6 animals were sacrificed at each of 10, 15, 20, and 25 weeks. Two treatment groups were used. One was administered vehicle (20% dimethylsulfoxide in water) containing 0.5 mg/kg of carcinogen (NMBA) subcutaneously (s.c.) 3 times a week for 5 weeks, and the other was given the same treatment along with a dietary supplement of 3 PEITC/g diet. The PEITC was administered in the diet 2 weeks before, during, and for 17 weeks after NMBA treatment. Of two control groups, one received vehicle solution without carcinogen s.c. 3 times a week for 5 weeks with no chemopreventive (PEITC), and the other was given the chemopreventive (3 µmol PEITC/g diet) for a period of 24 weeks without either carcinogen or vehicle treatment (chemopreventive control). At sacrifice, the esophagus of each rat was sectioned lengthwise, opened up to expose the epithelium, pinned on a small index card, and fixed in 10% neutral buffered formalin for 24 hours. It was then placed in 70% ethanol for 48 hours. After fixation, the flattened esophagus was cut in

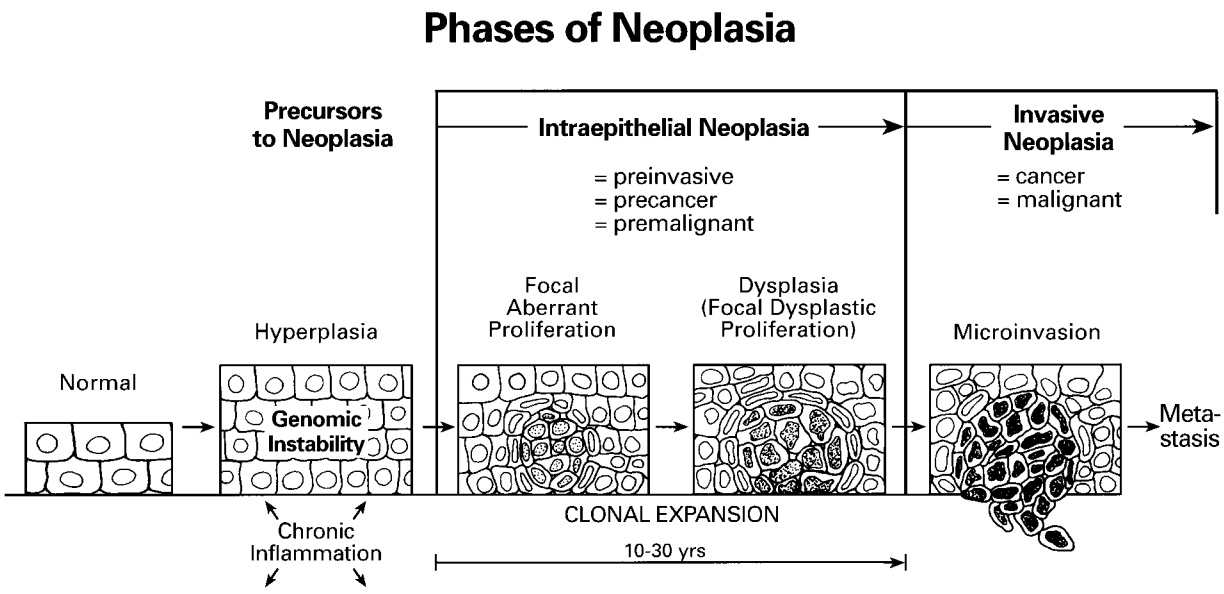


Fig. 1. Overview of intraepithelial neoplastic progression. The average time between onset of a visible neoplastic lesion within the epithelium and invasive neoplasia is more than 10 years.

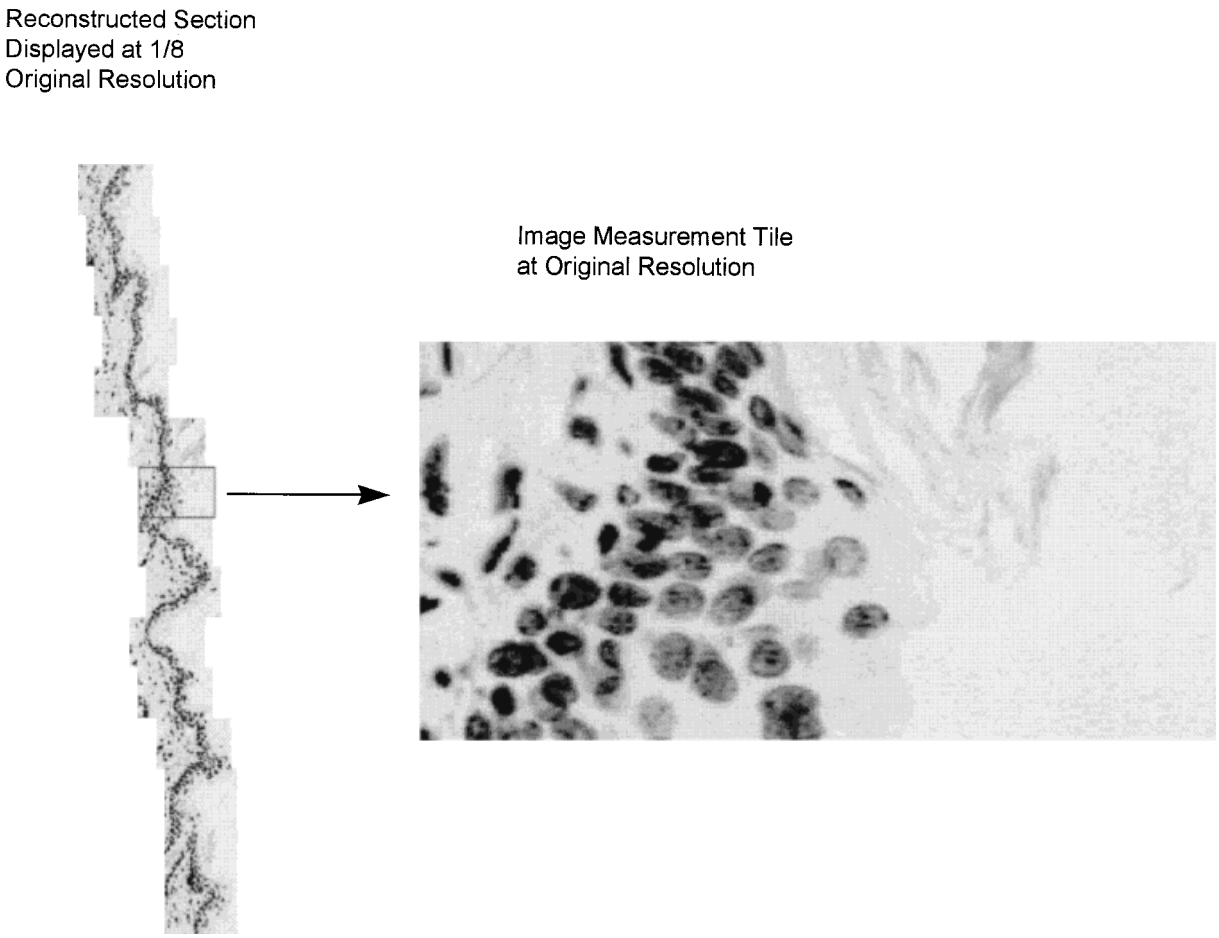


Fig. 2. Image "tiling" data acquisition and reconstruction. Software reconstruction enables observation of the entire tissue section while retaining the ability to display the "tiles" at the original digital resolution.

TABLE I. Measurements Used to Determine the Preinvasive Neoplastic Grade for Mouse Skin

x1	Summed optical density
x2	Markovian difference moment texture
x3	Markovian sum average texture
x4	Standard deviation of optical density texture
x5	Optical density peaks of .06 OD between adjacent image pixels (granule texture counts)
x6	Optical density differences of .10 OD between image pixels separated by 10 μm (counts of a specified "run length" texture along a scan line in the image)

TABLE II. Measurements Used to Determine the Preinvasive Neoplastic Grade for Rat Esophagus

x1	Area
x2	Summed optical density
x3	Optical density valleys of .01 OD between adjacent image pixels (small texture "hole" counts)
x4	Optical density differences of .05 between image pixels separated by 14.38 μm (counts of a specified "run length" texture along a scan line in the image)

thirds and embedded in paraffin blocks, oriented so that histological sections would be cut through the esophageal wall in a saggital plane. A specimen of liver from each rat was embedded in paraffin along with the esophagus. Sections were cut at 5 μm, and Feulgen stained using well-established methods, with calibration and control cells included in each stain batch [7,8].

For the skin study, 64 female SENCAR mice were randomly assigned to four groups of 16 animals, and each group was further divided so that 4 animals were sacrificed at each of 5, 10, 15, and 20 weeks after beginning the carcinogen treatment. All animals were on the TD88268 diet from Teklad. In one experimental group the mice were shaved and the dorsal skin was painted with 0.1 of the carcinogen benzo(*a*)py-

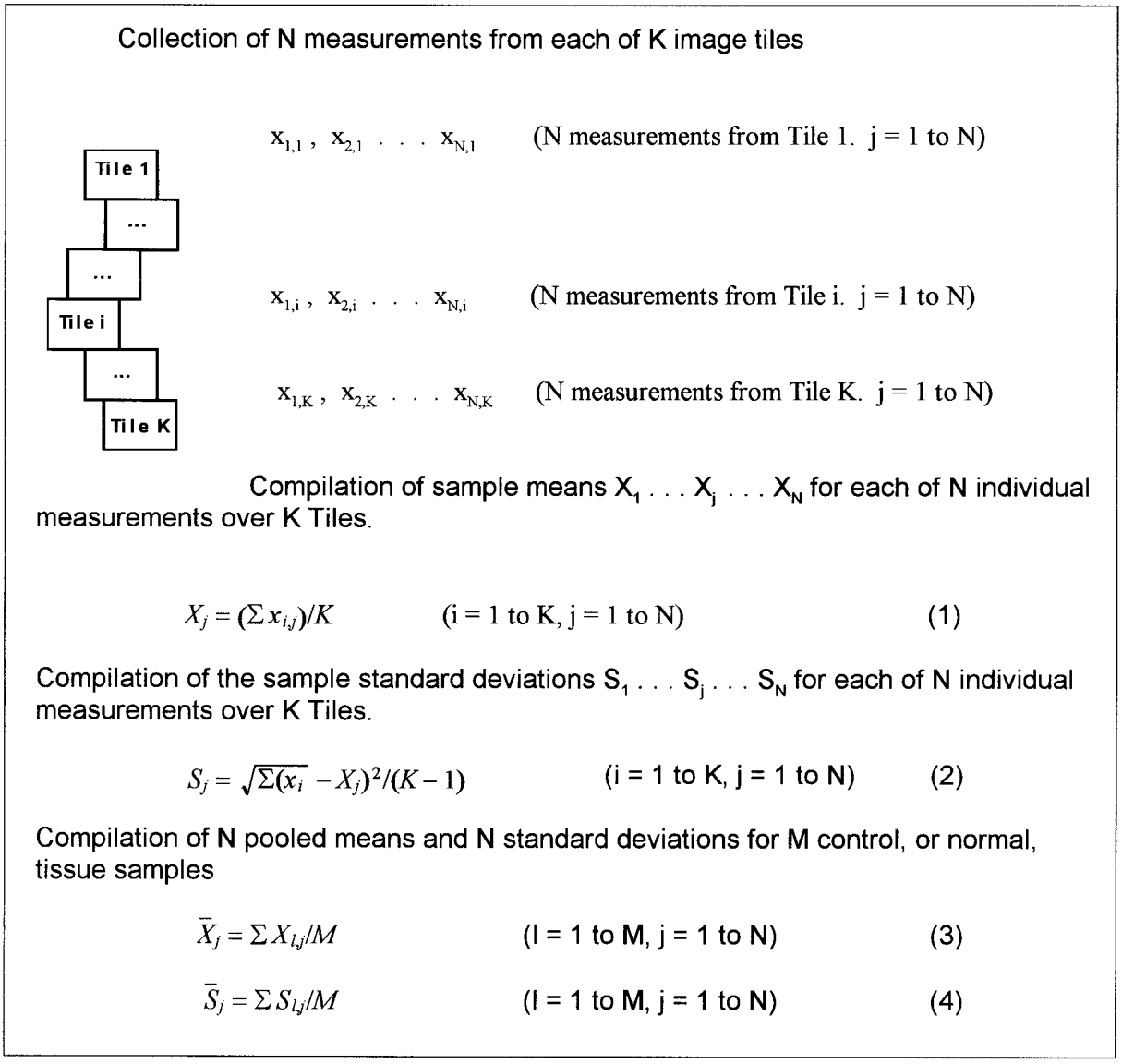


Fig. 3. Analysis procedure to determine the normal tissue pooled means and standard deviations, used to scale the individual measurements that make up the histologic grade.

rene (B[*a*]P) in spectrograde acetone 3 times per week until sacrifice. In another experimental group, 1 g/kg of the chemopreventive agent DFMO was added to the diet 1 week prior to B[*a*]P dosing and continued throughout the study. In another (placebo) group, the mice were shaved and the dorsal skin was painted only with acetone 3 times per week; in another (control) group the mice were shaved only. At sacrifice, the entire shaved area of the skin was removed, pinned on a small index card, and fixed in 10% neutral buffered formalin for 24 hours. It was then placed in 70% ethanol for 48 hours. Prior to embedding, the flattened skin was trimmed so that reproducible, comparable 2 cm^2 areas of skin from each animal were obtained. The selected sample of skin was bisected along the midline and embedded in paraffin with the cut surfaces oriented toward the face of the block. Sections were cut at 5 µm, and Feulgen stained as described above.

Image Data Acquisition

The CAS200™ image analysis equipment was used [9,10], with the attached automated x,y microscope stage Multiscan™ image recorder [11]. Image analysis was performed in two steps. The first step was interactive, and consisted of

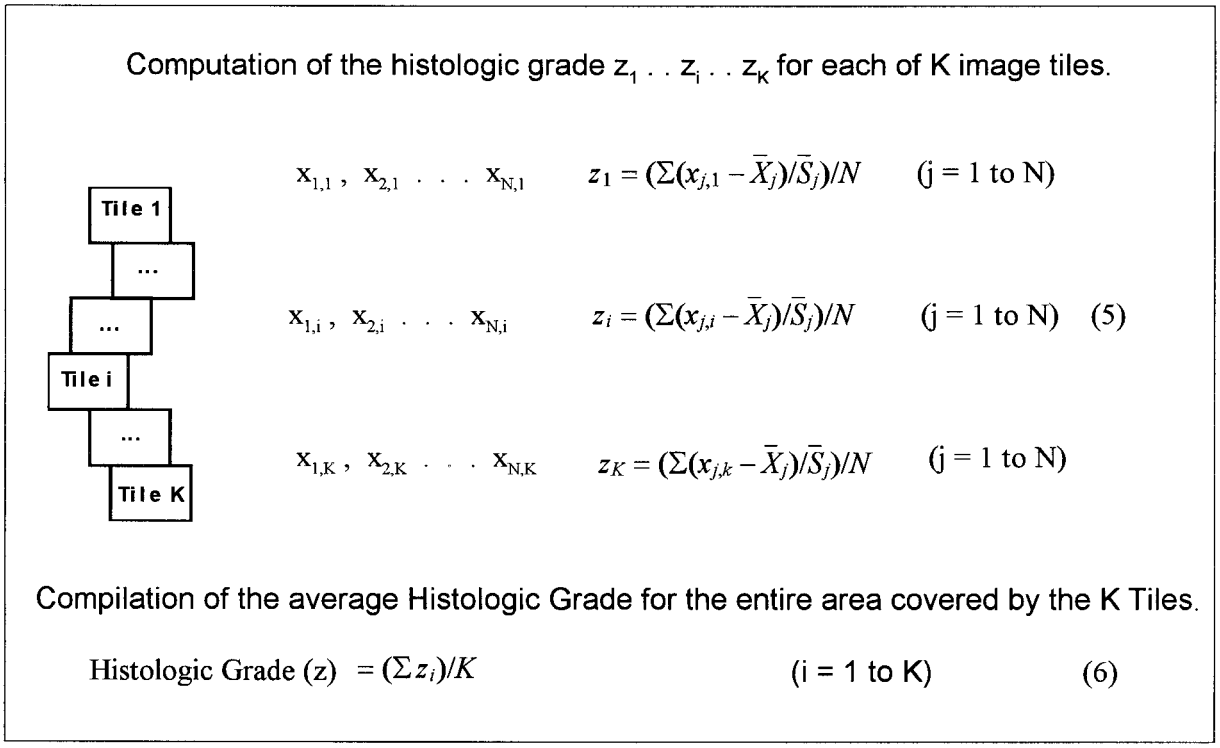


Fig. 4. Analysis procedure to determine the individual histologic grade for each image tile (equation 5) and the average histologic grade for the selected tissue segment represented by all of the image tiles, in standard deviation units of z, for suspected neoplastic tissue.

an operator choosing the region of interest in the tissue section by tracing it from a display on the system video monitor. The software allowed for large areas to be traced. In this instance, the traces consisted of outlining a region 146 µm wide following the basal epithelial layer. Image scanning and analysis were then automatically performed on the entire outlined region at high resolution. The scanning was performed at .57 by .34 µm pixel size with a 40x optical magnification. Each image field was 146 µm by 87 µm. The resulting scan was "tiled' together using the x,y positions of each field, which were stored during image acquisition as indicated in Figure 2, which also shows how software changed the digital magnification of the displayed images so that they could be viewed as a pathologist would, i.e., visually simulating different microscope objective lenses to allow for observation of a large area of tissue structure at low power, and smaller areas at higher power, at effective nuclear morphometric resolution.

During analysis, multiple measurements were made on each high-resolution image tile. These measurements were maintained in association with each tile, and additionally were used as described below to construct a "learning reference set," using normal control tissue from regions with similar histologics. A histologic grade was calculated using this reference set for each tile, which was made available to operator inquiry by selection from a list of image tiles. In addition, as described below, the measurement associated with each tile was used to construct various graphical displays relating the tissue structure and individual regions to their histologic grade.

Calculation of the Histologic Grade

We characterized neoplastic growth quantitatively with multiple measurements on each tissue image tile and combined these measurements into a single number, i.e., a histologic grade, for each tile, and used parameters from the distributions of these measurements to represent the histologic grade for the entire region considered.

Three basic types of measurements characterized each image tile: (1) measurements related to the overall DNA content of the image, (2) Markovian texture measurements (related to analysis of differences in optical density be-

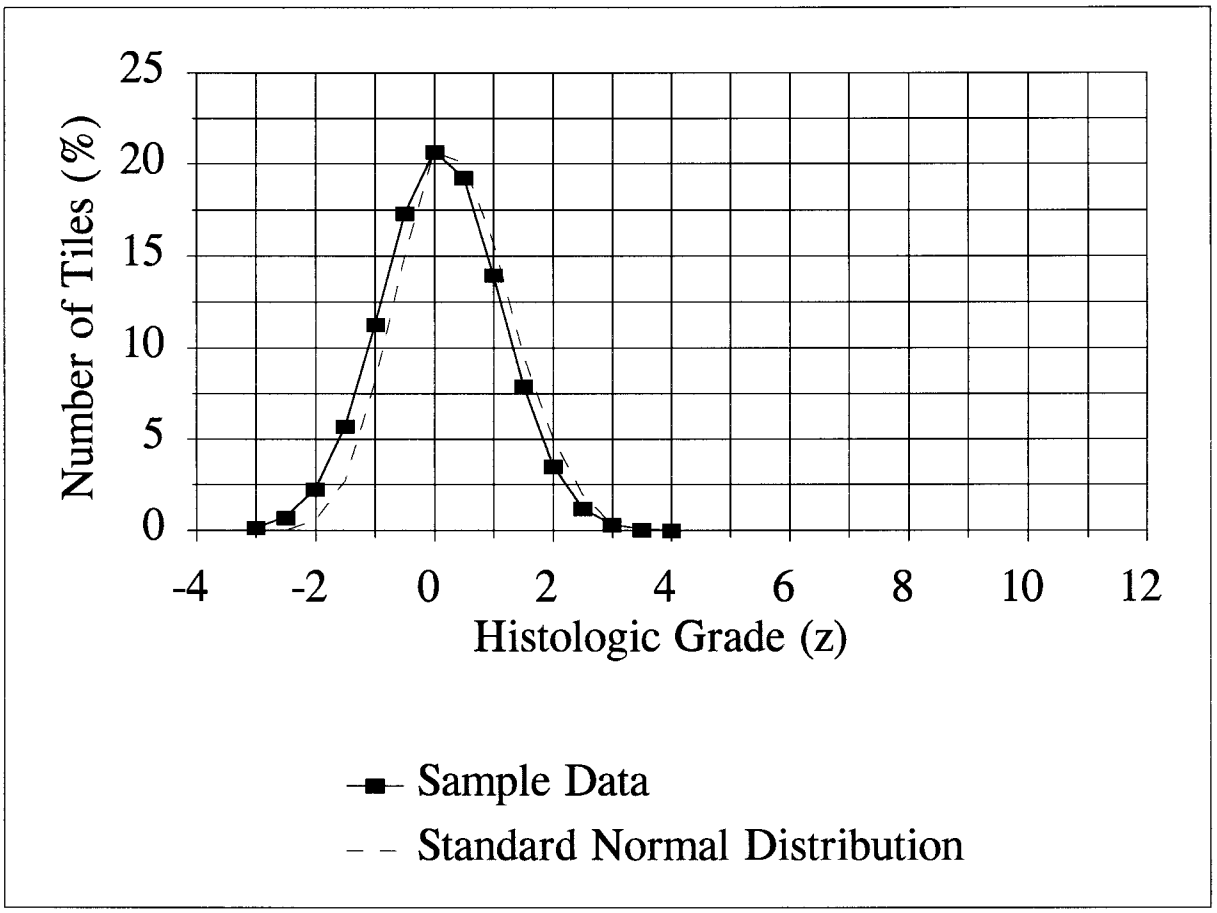


Fig. 5. Distribution of individual tile histologic grade z values from a cross-section of rat esophagus from a normal untreated control. The distribution of z value measurement data is shown superimposed upon a standard normal (Gaussian) distribution scaled to match the height of the measured z value distribution at zero.

tween a reference pixel and other pixels at defined distances from the reference pixel) [10,12] to characterize the overall association and interrelationship of optical density textures in each image tile, and (3) counts of individual event texture measurements per tile to identify increases in fine chromatin texture. Approximately 100 different types and combinations of these three main categories of measurements were evaluated for discriminatory effect and suitability, separately for the two assays developed. In each case, a small number of measurements were finally selected that adequately represented each preinvasive neoplastic process. The selection criteria were a combination of the following: a high histologic grade at the endpoint of the neoplastic process, and a generally monotonically increasing histologic grade from zero to the endpoint over the period of preinvasive neoplastic growth; a generally Gaussian distribution of histologic grade for the distribution of samples taken from individual control (or normal) tissues representing the different time points of the experiment; and measurements that visually agreed with the histological interpretation of the tissue section areas in each image tile. In evaluations of this type, point texture measurements were displayed where they occurred in individual tiles so they could be visually confirmed to represent chromatin patterns in the tissue. Additionally, the software enabled "point and click" review of individual tiles to examine abnormally high or low tile grades as reliable indications of an authentic neoplastic process, and listed all tile grades for "point and click" review by bringing up the low-resolution image of the tissue section visually as well as a high-resolution image of the individual tile in question. Specific individual measurements that contributed to the grade were listed, and optionally overlaid on the image.

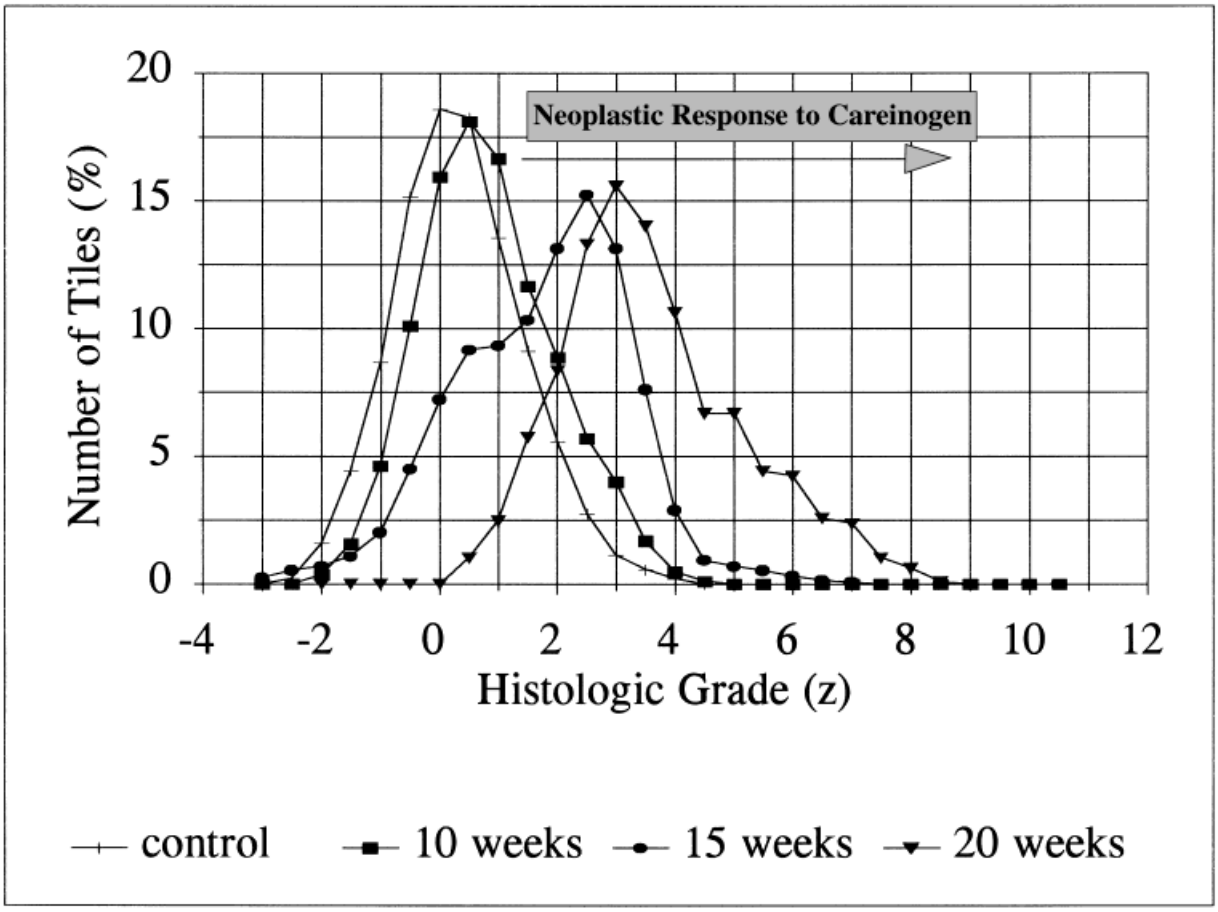


Fig. 6. Distribution of z values from longitudinal strip cross-sections of mouse skin, from control tissue, compared to animals treated for progressively longer time periods with the carcinogen B[*a*]P.

The measurements listed in Tables I and II were used in the preinvasive neoplastic histologic grade for mouse skin and rat esophagus assays. Although these measurements were computed from each image tile, many of them serve to characterize the chromatin structure of the nuclei in that tile. Others characterize the nuclei but also reflect differences in histologic structure of the preinvasive neoplastic basal cell layer and immediately derived cells in individual tiles compared to the normal histologic structure of the basal cell layer in individual tiles acquired from normal control animals.

In order to address the issue of a common scale, and to normalize the contribution of the variance of normal tissue, all of the above measurements were made on individual tiles from the entire selected tissue area from age-matched control animals. The sample mean and standard deviation of each variable was calculated and averaged to obtain a pooled mean and standard deviation for each normal control group. This procedure is detailed in Figure 3, with accompanying equations 1, 2, 3, and 4. As shown in Figure 4, the histologic grade for each tile was then computed by first normalizing each measurement and then averaging the contribution of each of the measurements after normalization. This process is defined by equation 5. Normal control tissues processed in this manner essentially resulted in distributions with zero mean and unit standard deviations, while tissue undergoing neoplastic progression resulted in distributions with means shifted from zero to higher grades, and with higher variances. Various descriptive parameters of these distributions, e.g., the mean of the distribution of image tile histologic grades (computed by equation 6), the percent of tissue area with histologic grades greater than two standard deviations, etc., were used to characterize overall tissue samples. Additionally, as mentioned above, focal areas of dysplasia were easily represented graphically by plotting the z values of equation 5.

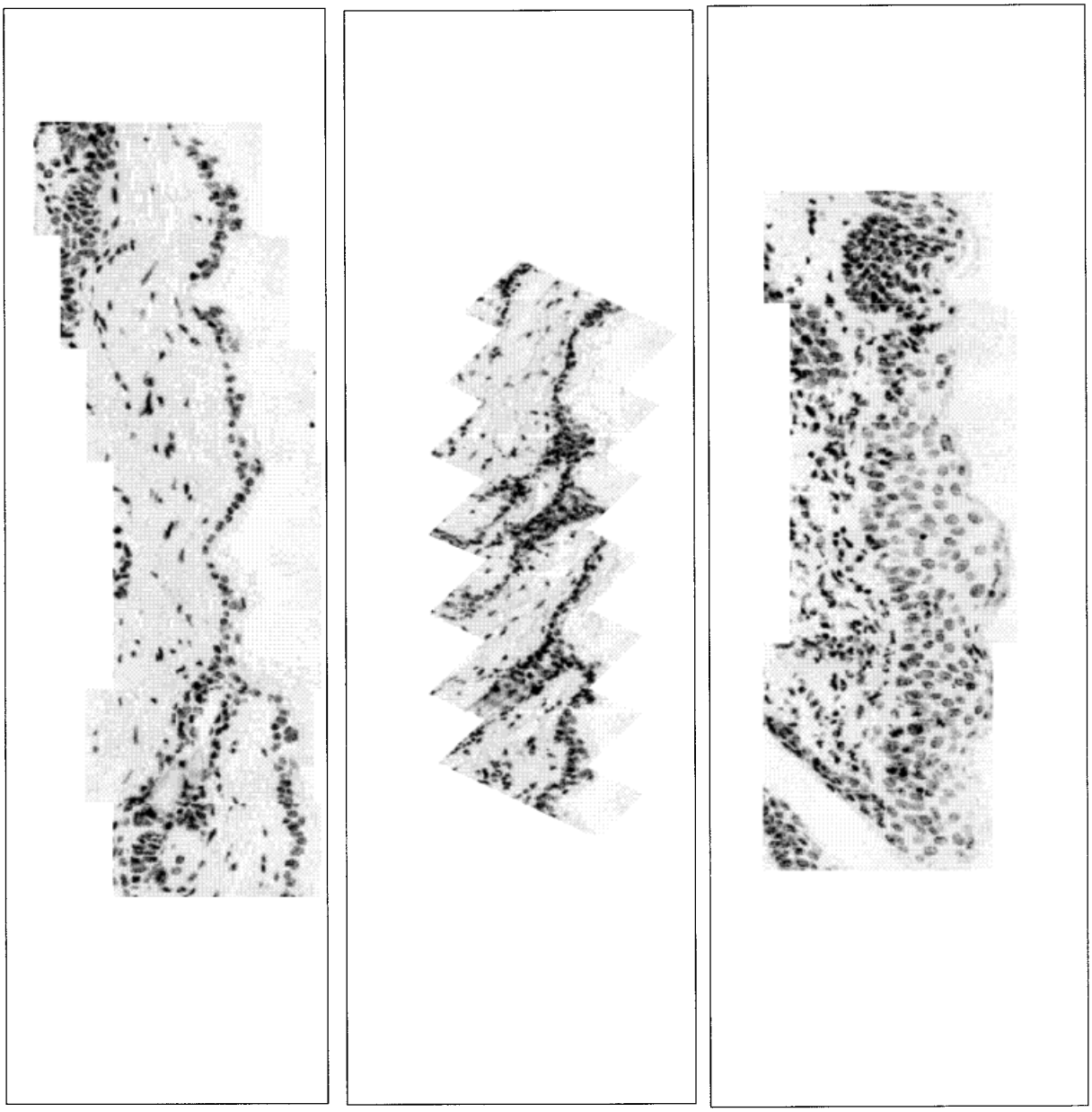
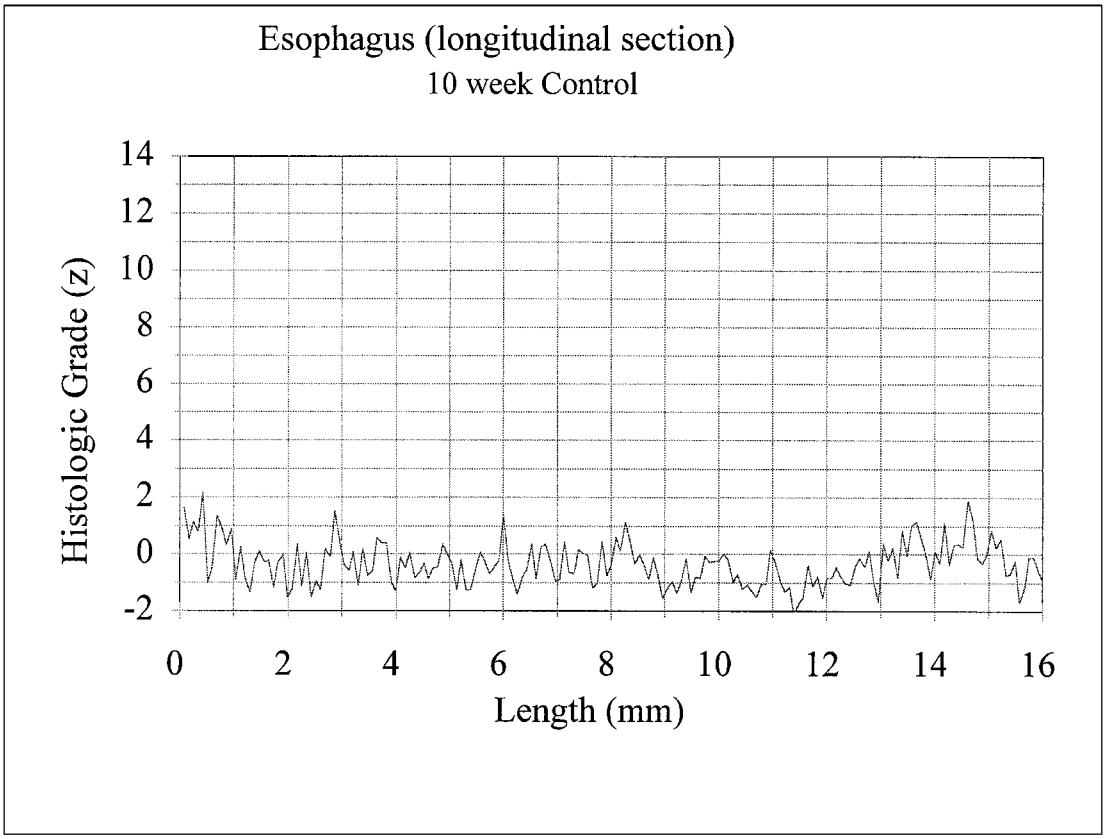


Fig. 7. The visual histologic appearance of individual skin biopsies from the same tissues that the distributions of Figure 6 were derived from.

The quantitative histologic grade defined by equation 5 for individual small areas of the histologic section represented by an image tile, or defined by equation 6 for a complete histologic section structure, is in standard deviation units represented by the symbol z. Normal tissue components tend to be at mean z values of zero, rarely exceeding values of 2 z, and very rarely exceeding values of 3 z. Histologic grades above 3 z represent abnormal nuclear morphometry and abnormal histologic structure.

RESULTS

In general, histologic grade is quantitated over a predetermined region of interest by subdividing that region into smaller tiles and measuring the morphometric characteristics of each tile. The quantitative characteristics of the



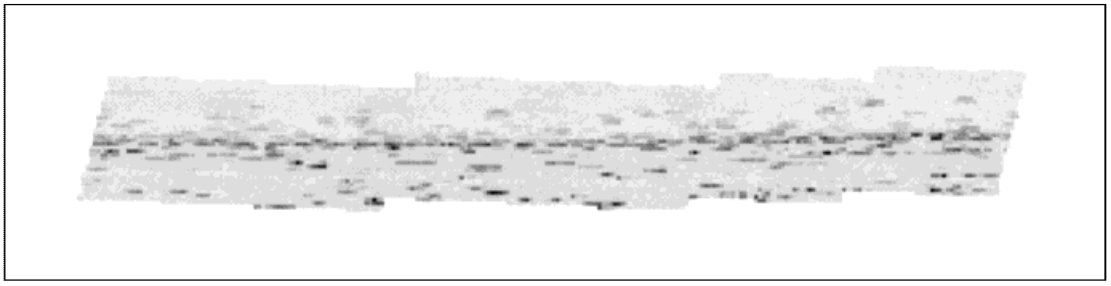
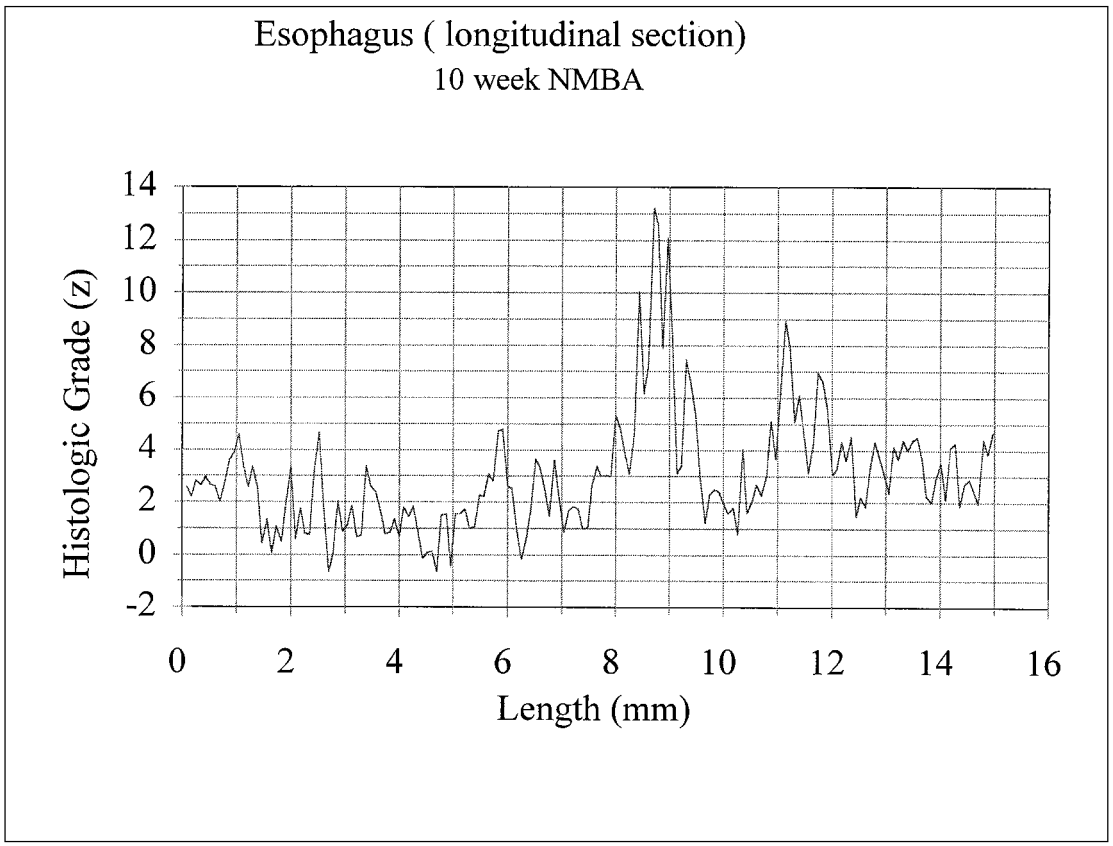


Fig. 8. A normal control tissue from rat esophagus. In the graph, the y axis is the histologic grade of each tile, and the x axis is the linear distance along the basal layer of the esophagus section. The actual tissue section, corresponding to the graph, is much longer. Only an illustrative section is shown at a higher resolution image of the basal layer.

larger predetermined region can then be defined by parameters of the distribution of the individual tile measurements. Figure 5 is an example of the distribution of individual tile measurements from a cross-section of untreated control rat esophagus tissue after analysis. In this case, the distribution of individual z values were compared to a normal Gaussian distribution with zero mean and unit standard deviation scaled to measure the height of the measured distribution at z equal to zero. This distribution was typical of all of the control animal tissue section distributions from mouse skin and rat esophagus, and indicated that the transformation of the measurement scale to zero mean and unit standard deviation resulted in essentially Gaussian distributions for normal, or control, tissue.



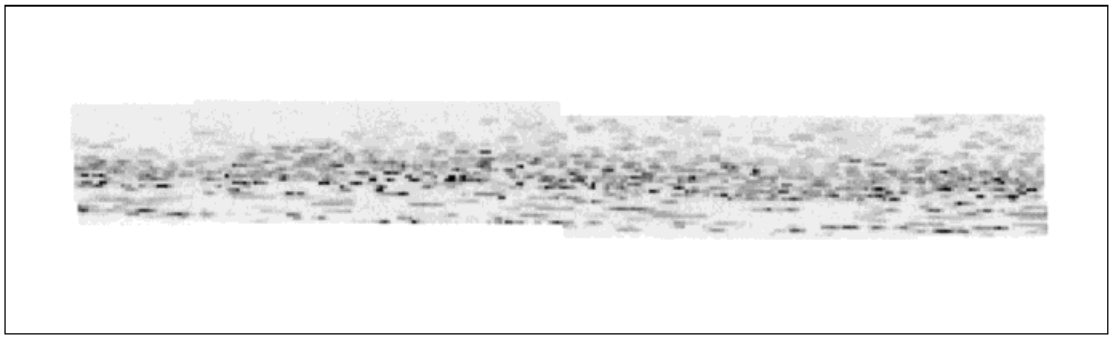


Fig. 9. The same type of plot from an esophagus section as that shown in Figure 8, from a rat after 20 weeks of treatment with NMBA. This representation further presents the characteristics of the tissue architecture as it relates to the quantitative histologic in the specimen region, and contributes to the potential of further quantitative definitions of "focal areas of dysplasia."

Figure 6 shows a similar distribution of z values from a cross-section of mouse skin from control tissue, compared to distributions of z values from animals treated for progressively longer time periods, further illustrating the results of the method. Notice that histologic grade measurements from this normal control tissue were also essentially distributed in a Gaussian manner, but as exposure to carcinogen continued, a population of tissue elements developed with progressive deviations from normal. A progressive increase in the central tendency of the distributions was shown, as well as a progressive increase in their variance. Figure 7 shows the correlation of the visual histologic appearance of individual skin biopsies from the distributions of Figure 6, taken from different animals after progressively longer B[*a*]P treat-

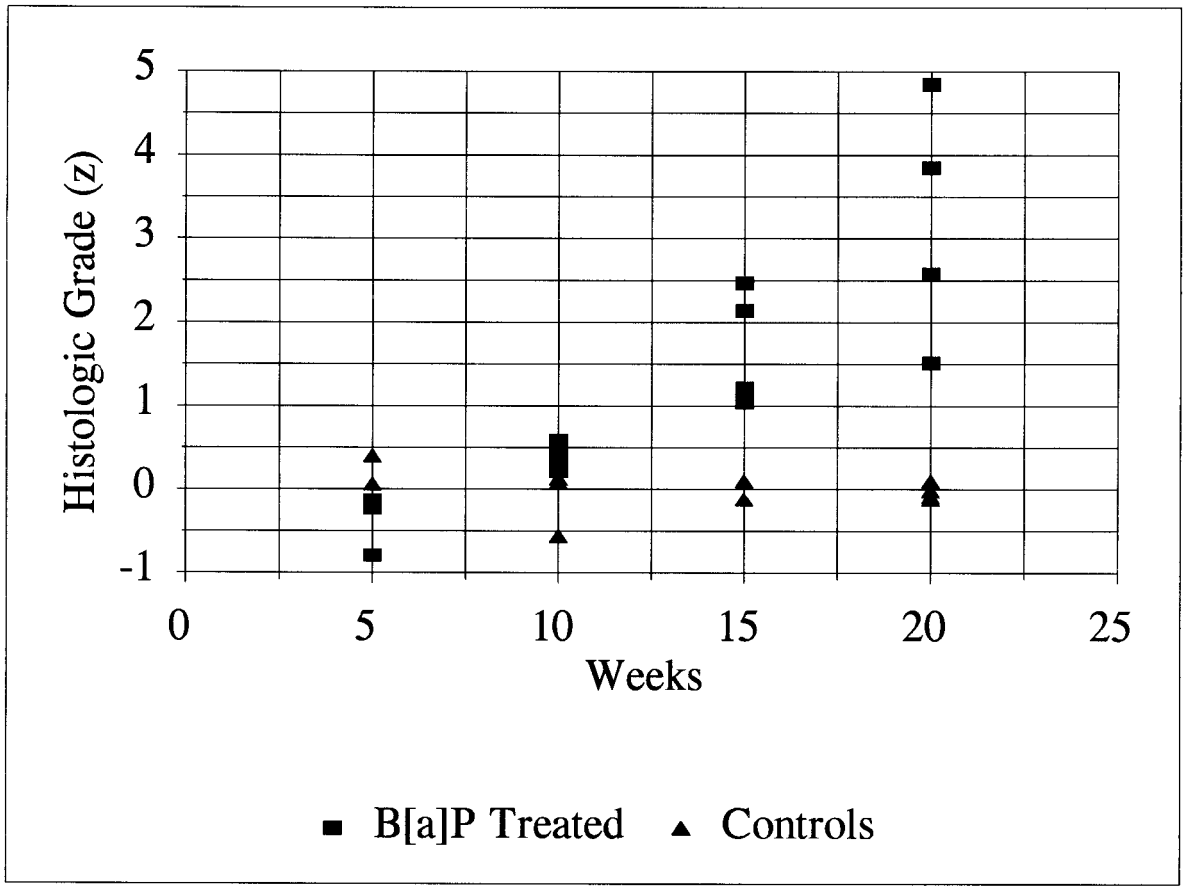


Fig. 10. The preinvasive neoplastic response in the SENCAR mouse skin animal model to B[*a*]P. The plot is organized to characterize the neoplastic response as a function of time. In this experiment, there were four animals in each group, and the average of the histologic grade, as computed by equation 6, for each animal was plotted (■). The average of the histologic grade for the individual normal control tissue samples was indicate by (▲). The differences from normal were significant at 15 and 20 weeks ($P < .05$).

ments. The progression of neoplastic expression is visually apparent, and together with knowledge of the measurements, visually confirms the shape and appearance of distributions from animals with increased exposures.

In both of these animal models, the region of interest was defined by the area around the basal layer. This special case resulted in a linear scan pattern, i.e., alignment of image tiles. The results could thus also be expressed graphically as shown in Figures 8 and 9. Figure 8 is a plot of z values from a cross-section of normal esophagus tissue, where the y axis is the histologic grade (z_i from equation 5) of each tile, and the x axis is the linear distance along the basal layer of the esophagus section. Figure 9 is the same type of plot from rat esophagus after 20 weeks of treatment with NMBA. This representation further presents characteristics of the tissue architecture as it relates to quantitative histology in the specimen region, and contributes to the potential of further quantitative definitions of "focal areas of dysplasia."

Figure 10 shows the neoplastic response to B[*a*]P in the SENCAR mouse skin animal model. The plot is organized to characterize the neoplastic response as a function of time for the complete experiment. This experiment had four animals in each group, plotting the individual mean value for histologic grade for each one. A separate group of pooled means and standard deviations was compiled for each of the 5-, 10-, 15-, and 20-week controls. When individual treated animals were analyzed from a specific time period, the corresponding pooled mean and standard deviation for that group, for each measured variable, was used to calculate a normalized value for that variable according to the procedure detailed in Figure 4 and equation 5. The mean value for each case, computed from

equation 6, was averaged from one complete section of approximately 300 image tiles taken longitudinally from the back of each animal. The data indicated an increasing histologic grade with exposure to B[*a*]P as a function of time. Similar to and showing the same trend as individual cases shown in Figure 6, this plot more succinctly summarized the entire study. As expected, some variance in the data may have been due to individual animal response to B[*a*]P application, i.e., development of neoplasia, or to variation in preparation and measurement techniques. However, the differences from normal were significant at 15 and 20 weeks ($P < .05$). Since the data were normalized (normal tissue is generally grouped at zero on the histologic grade scale, which is in units of standard deviation from normal), histologic grades (i.e., mean values of z) taken from tissues at 20 weeks that are greater than three standard deviations from normal are clearly significantly different.

Figure 11 shows the results of testing the effect of the chemopreventive agent DFMO in mouse skin on similar cohorts of four mice per group who were also treated with B[*a*]P. The upper plot represents the average neoplastic response from Figure 11, and the lower plot represents the average values for the experimental group treated with B[*a*]P and also with DFMO. The control data means were not plotted, since they were by definition all at a zero mean value. A clear tendency towards B[*a*]P modulation was significant at 20 weeks ($P < .05$). The mouse skin experiment also consisted of a placebo group of animals treated only with acetone instead of B[*a*]P. Figure 12 shows the results of the same type of analysis performed on this group. Each plotted point represents the mean of four animal assays. In general, placebo-

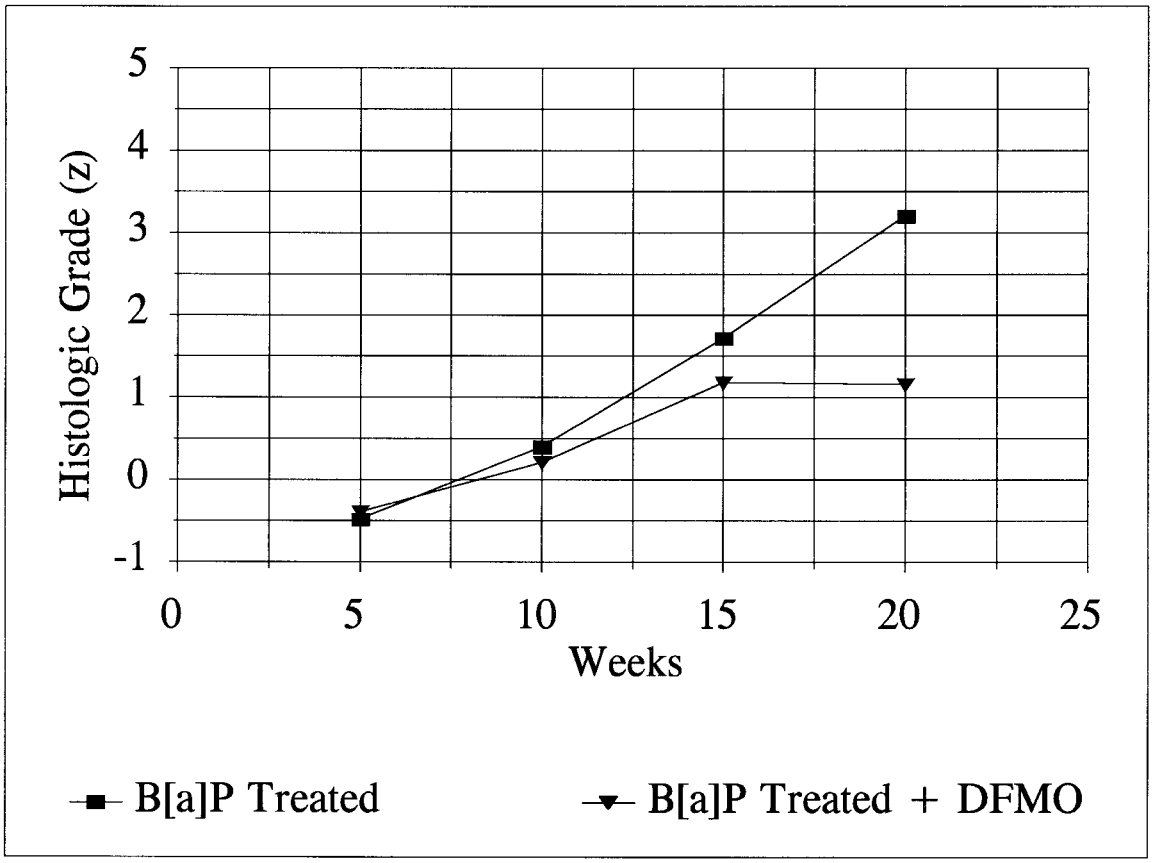


Fig. 11. Results of testing the effect of the chemopreventive agent DFMO to similar cohorts of 4 mice per group who were also treated with B[*a*]P. The upper plot (■) represents the average neoplastic response from Figure 10, and the lower plot (▼) represents the average values for the experimental group treated with B[*a*]P and also with DFMO. The control data means were not plotted, since they were all at a zero mean value. There was a clear tendency towards modulation of the effect of B[*a*]P, which was significant at 20 weeks ($P < .05$).

treated animals responded with some hyperplasia, with an average trend at a maximum of about .5 z, although no group was significantly different from normal.

Figure 13 shows the results plotted in a similar fashion from analysis of the neoplastic response to the carcinogen NMBA in the rat esophagus model, and its modulation by the chemopreventive PEITC; this experiment had six animals in each group. The mean histologic grade for each animal was computed from a sequence of approximately 300 image tiles from one complete section taken longitudinally along the length of the esophagus. The average value for each group of six was plotted using the mean histologic grade for each animal. The group data indicated an increasing histologic grade with time after exposure to NMBA. This plot was also very similar to Figure 11 for the mouse skin. Expected variance in this data may have been due to individual animal response to NMBA, i.e., development of neoplasia, or to other preparation or measurement techniques. However, the differences from normal, i.e., the neoplastic response represented by the upper plot, were significant for all time periods ($P < .05$). The lower plot shows the effect of the chemopreventive agent, PEITC. Significant differences in modulation, i.e., a retarded neoplastic response, was obtained at 10 and 15 weeks ($P < .05$). The reduction in neoplastic response at 20 weeks was attributed to the briefness of NMBA administration (indicated in Fig. 13), which allowed for recovery after 15 weeks. Figure 14 summarizes the distribution data in a slightly different manner using a three standard deviation cutoff level; the average group percentage of tissue tiles greater than this cutoff value was plotted as a function of time instead of the average group mean values, as in Figure 13. Each data point in this plot represents the mean from six esophagi, and was from the same cases plotted in Figure 13.

Considering the distributions of Figure 15, and the linear plots of Figures 8 and 9, increasing exposure to NMBA clearly causes a progres-

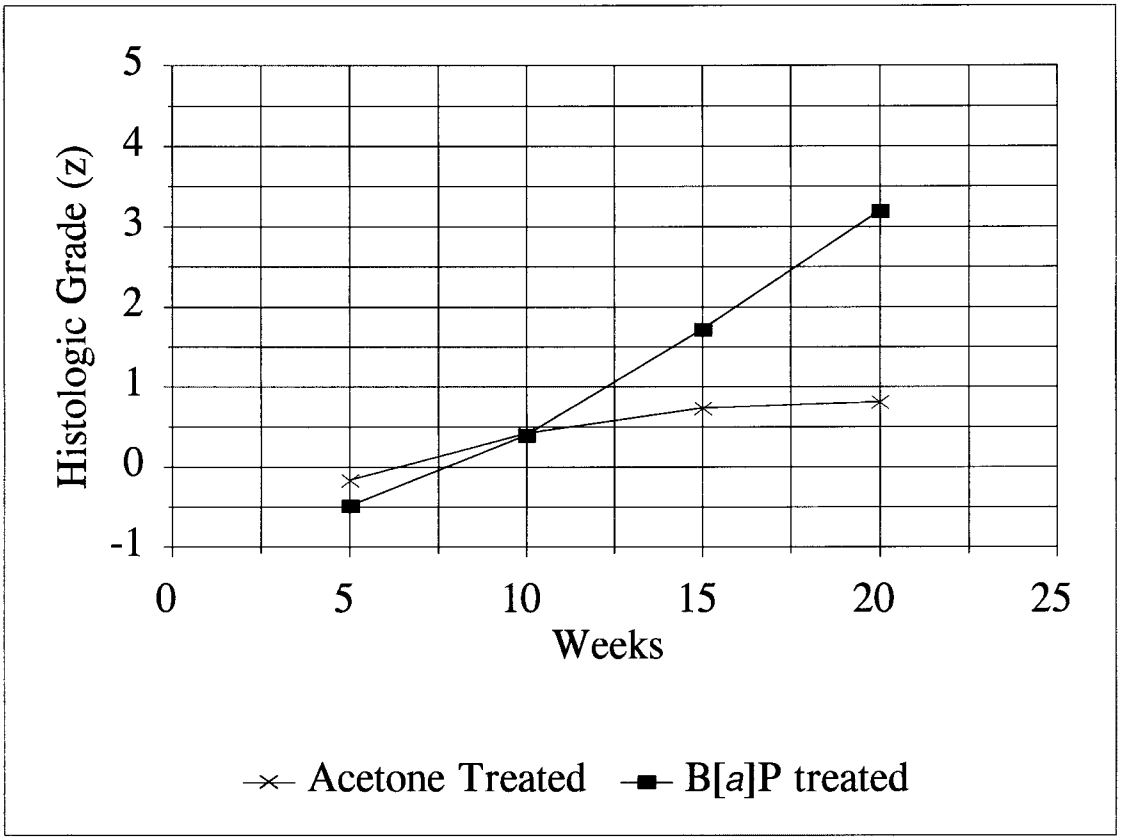


Fig. 12. Results of testing a placebo group of animals treated only with acetone instead of B[*a*]P. The upper plot (■) represents the average neoplastic response to B[*a*]P from Figure 10, and the lower plot (x) represents the average values for the experimental group treated with acetone.

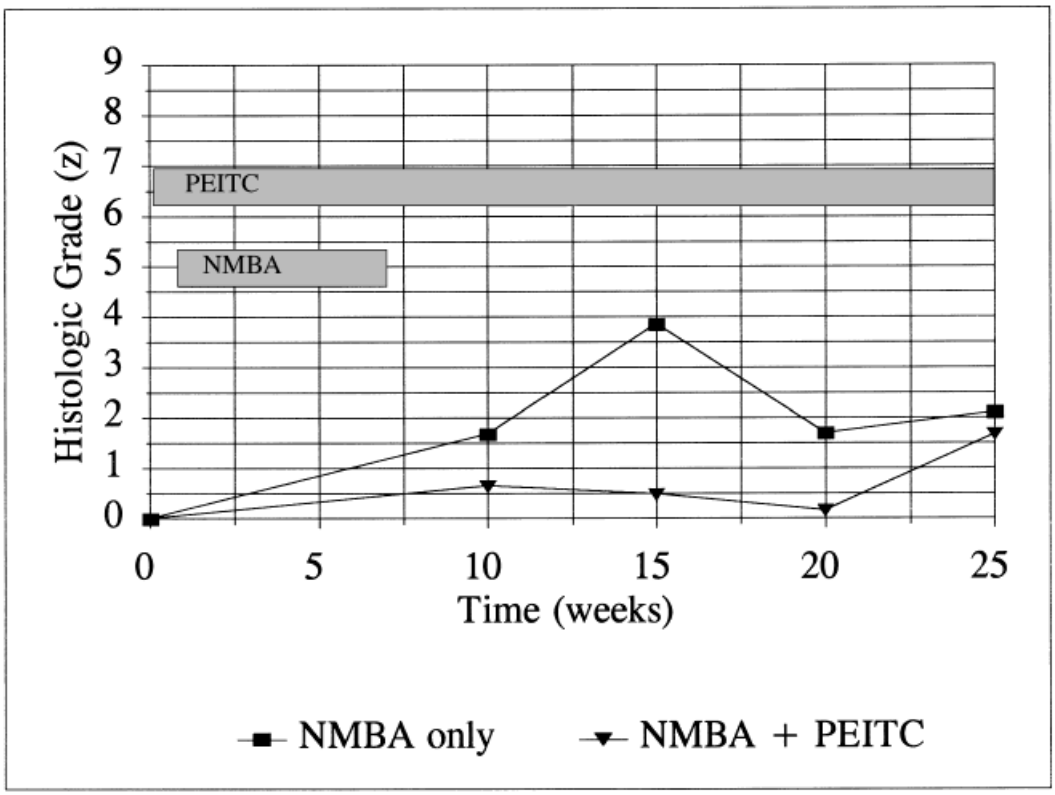


Fig. 13. The neoplastic response in the rat esophagus animal model to the carcinogen NMBA, represented by the upper graph (■), and its modulation by the chemopreventive PEITC, represented by the lower graph (▼). Each plotted point represents the average of the distribution means of 6 animals.

sive increase in the amount of aberrant tissue, or neoplastic progression in the tissue. After 15 weeks, 55% of the tissue was histologic grade of $>$ three standard deviations from normal. Also, PEITC effectively reduced the percentage of tissue greater than grade 3 from 55% to 5% at 15 weeks, decreasing the variance in the distribution of histologic grade to almost normal.

Figure 15 shows the distribution of z values from a longitudinal strip cross-section of rat esophagus from an animal treated with the carcinogen NMBA, compared to an animal treated with the carcinogen and also with the chemopreventive PEITC. A standard Gaussian distribution is shown as a reference. In a distribution form on an individual animal basis, these results show the same effect of the chemopreventive PEITC on reducing the preinvasive NMBA-induced neoplastic growth that was evident in the results shown in Figures 13 and 14.

DISCUSSION

During the last 10 years, interactive image cytometry has become an accepted testing method in pathology. Assays have been developed for DNA ploidy [7,8,13,14], estrogen receptor quantitation [15], proliferation status [16], and oncoprotein expression [17,18]. One of the strong features of this methodology is that it can be applied to tissue sections, simultaneously allowing for visual inspection and interaction as part of the measurement process. Since the development of neoplasia is defined by its visual and histologic tissue section presentation, image cytometry at the tissue level has the obvious potential of quantitating early neoplastic development and being used as a surrogate endpoint biomarker in cancer chemoprevention.

Visual grading in cancer diagnosis and prognosis has had a long history [19]. This includes grading systems for prostate cancer [20,21], cervical cancer [22], and breast cancer [23,24], among others. More recently, with the advances in image cytometry referred to above, quantitative grading systems have been attempted using whole cell and especially nuclear imaging morphometric analysis [25–28]. This nuclear grading was more easily achievable technically

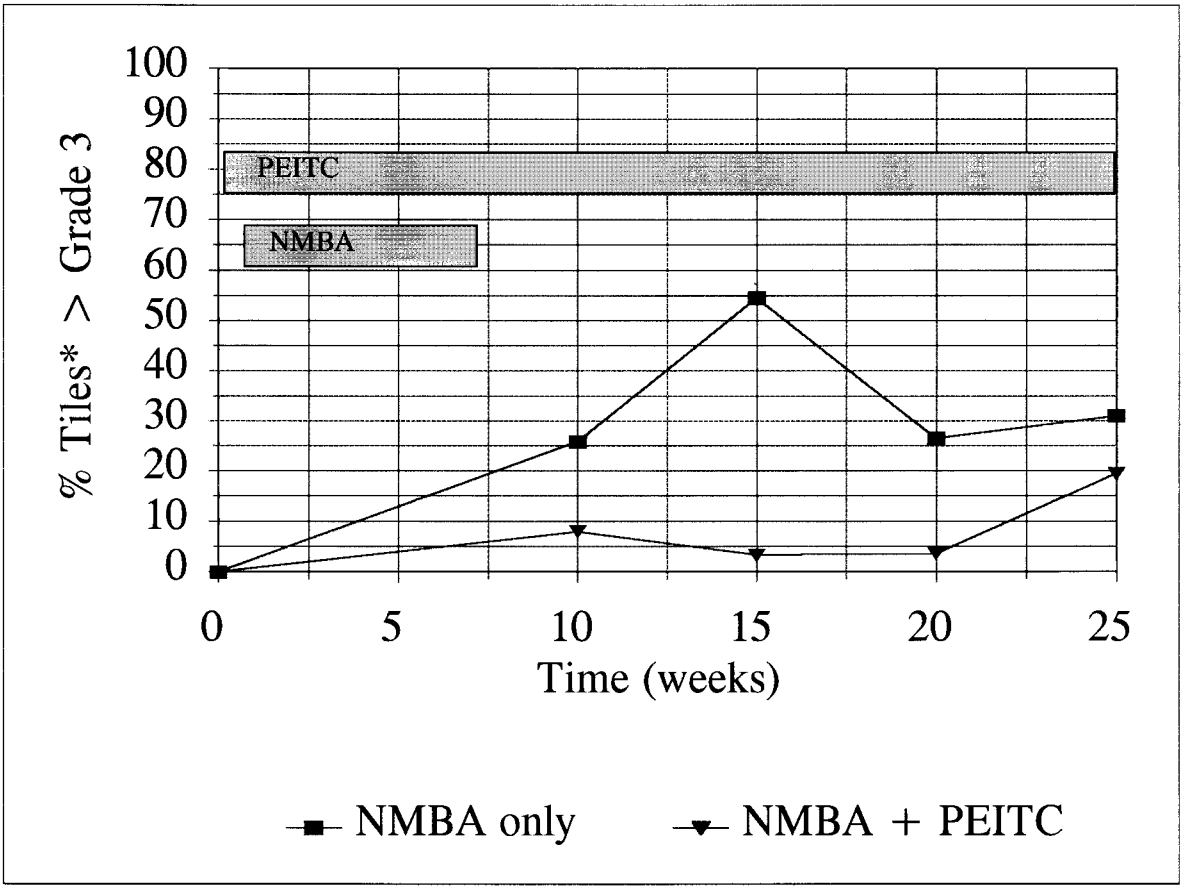


Fig. 14. The neoplastic response in the rat esophagus animal model to the carcinogen NMBA, represented by the upper graph (■), and its modulation by the chemopreventive PEITC , represented by the lower graph (▼). In this presentation a 3 standard deviation cutoff level was chosen for each individual distribution, and the average percentage of tissue tiles greater than this cutoff value was computed, instead of the mean of the distribution. Each data point in this plot represents the mean these percentages computed from 6 esophagus assays, and was from the same cases plotted in Figure 13.

and was consistent with the evolution of imaging techniques in cytometry. However, as pointed out by Boone et al. [1], preinvasive neoplastic growth clearly consists of two components, abnormal nuclear structure and the histologic extent and appearance of neoplastic growth extending from the basal cell layer. Currently, subjective inspection of the histologic appearance of sectioned and stained material, or "grading," by experienced pathologists is being used to evaluate neoplastic progression. This has well-known limitations of reproducibility, accuracy, and resolution of grading scale. This study presents a first attempt at the next step in the evolution of image cytometry by addressing the quantitation of the histology of neoplastic growth by tissue section grading on whole fields, and thereby extends image analysis and morphometry to the quantitation of preinvasive neoplastic progression. As mentioned above, this becomes very relevant in cancer chemoprevention, where quantitative evaluation of preinvasive neoplasia is needed.

The challenge of applying these image cytometric techniques to quantitating neoplastic growth for chemoprevention purposes was to provide a common scale for different tissue types, with potentially different multiparameter measurements characterizing the different tissues. Also, there was a need to compare experiments between different animal models with different chemopreventive agents, and to compare clinical trials with different organ site targets and chemopreventive agents. We wanted a method that would adjust itself to measurements of neoplastic growth for the natural variance of normal tissue or controls in animal experiments. We were able to substantially achieve the above requirements with measurement methods combining those used in remote

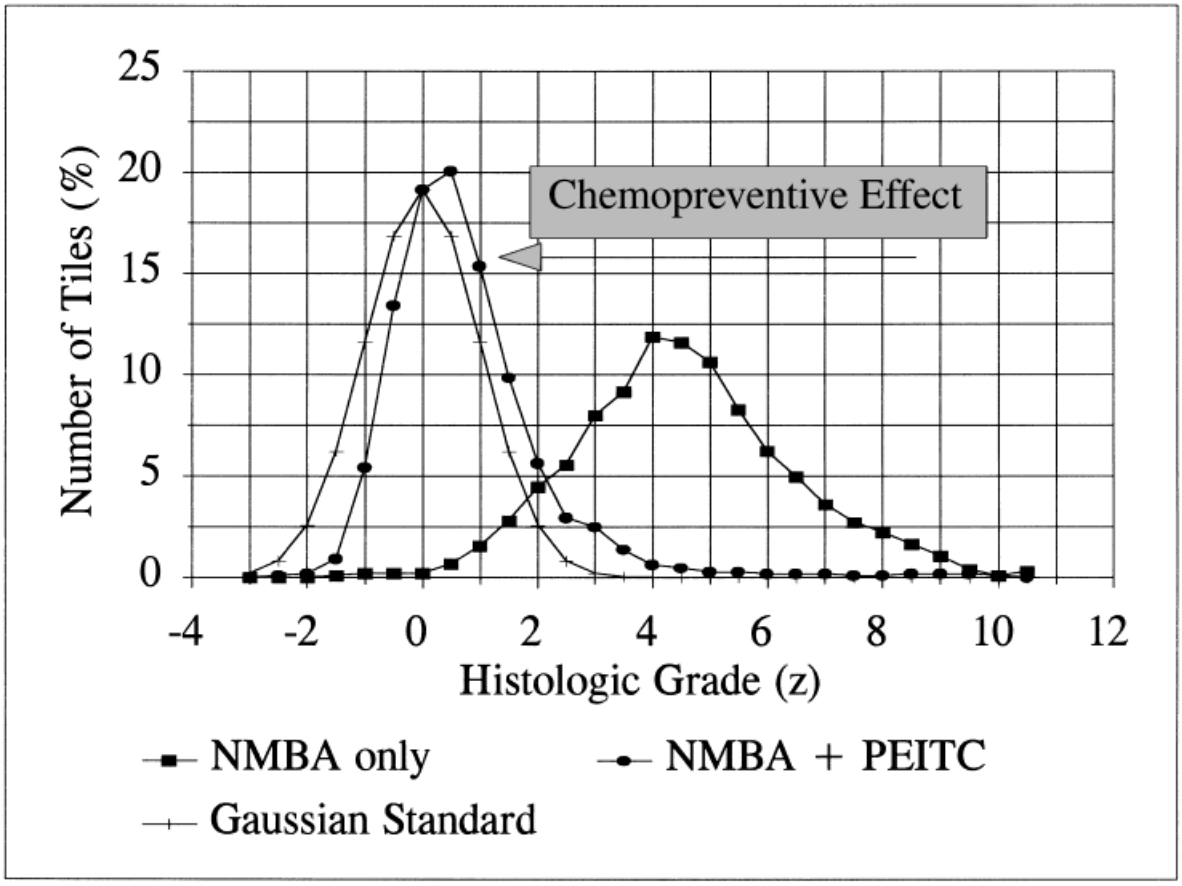


Fig. 15. Distribution of histologic grades computed from each individual image tile (equation 5) from longitudinal strip cross-sections of rat esophagus; from an animal treated with the carcinogen *N*-nitrosomethylbenzylamine (NMBA), compared to an animal treated with the carcinogen and also with the chemopreventive phenethyl isothiocyanate (PEITC). Both distributions are compared to a Gaussian standard distribution as a reference. The height of the distributions (y axis) was scaled to the number of tiles as a percentage of the total number of tiles analyzed for each animal.COMP: See hard copy for Tables I and II

sensing crop classification [29,30] and image cytometry [25–28] while handling the multiparameter measurements from each image tile with a method that had its roots in signal detection theory and psychophysical testing [31–34]. We have shown that, using two different animal models with different carcinogens and chemopreventive agents, results could be evaluated in a similar context. This same transferability of measurement scale comparison should also be useful in comparing different clinical trials with different organ site targets and chemopreventive agents.

The common scale for histologic grade, expressed in units of standard deviation, has a number of interesting aspects. First, it tends to automatically adjust the measurements of neoplastic growth for the natural variance of normal tissue, or control reference tissues. Also, it provides a theoretical link to psychophysical testing, signal detection theory, and ROC characteristics and methodology, where these basic normalization models are assumed [25,33,34]. In concert with the above analogies, this method allows prediction of ROC curves for detecting neoplastic response [33]. These are important considerations when comparing the visual interpretation of neoplastic progression, which is uniformly considered a difficult visual inspection task, to the performance of that same task by computerized image analysis.

Finally, there are at least two immediate potential uses of this method in current chemoprevention activities. First, animal models such as the ones described here can be used to rapidly screen potential chemopreventive agents. For example, now that the dose response has been established in the SENCAR mouse skin

assay shown in Figure 10, various chemopreventive agents can be tested in a 20-week time period, with a significant reduction in the number of animals used. As an example, DFMO showed a significant reduction in the progression of neoplasia, i.e., a reduction of the cohort histologic grade from the neoplastic level of greater than three s.d. to approximately one s.d. ($P < .05$), not significantly different than the placebo-treated group at a histologic grade of just less than one s.d., using only 4 animals ($P > .05$). To achieve this same statistical accuracy of relative frequencies for animals with and without tumors, using tumor counting would require on the order of five times the number of animals per compound tested. This is both a savings in the initial cost per animal and in maintenance costs of the animals. A similar rationale would establish the esophagus model as a screening system since the NMBA dose response curve has now been established. In this case, a significant reduction in the development of neoplasia at all time intervals was seen using only six animals. At 15 weeks, neoplastic response was reduced from an average cohort-induced histologic grade of about 4 s.d. to about 0.5 s.d. ($P < .05$). The reduced level was not significantly greater than the control level ($P < .05$). To achieve this same statistical accuracy by the previous method of tumor counting would again require five times the number of animals per compound tested.

The second potential use of this method is in "before and after" testing in human clinical Phase II trials. In this type of experiment, a cohort of patients, e.g., a high-risk group suspected of having already developed high-grade preinvasive neoplasia, is administered a potentially useful chemopreventive agent to reduce neoplastic growth. Prior to administration of the chemopreventive agent, a tissue sample is obtained, and quantitatively graded as described above. At a later time, perhaps after successive administrations of the chemopreventive, a second sample is taken to assess the effectiveness of the treatment. An example of this type of study is now underway to assess the effectiveness of various chemopreventive agents in reducing cervical intraepithelial neoplasia (CIN) from high grade (CIN III) to a lower grade [35]. The testing method described here is being used in these studies. Reference to Figure 15 is helpful to visualize by analogy the type of comparison that could be made. The first sample may take the form of the distribution at the far right (a 15-week exposure to NMBA in this animal model experiment), and if the chemopreventive was effective, the distribution of the second sample would be shifted to the left, as shown in Figure 15 by the results of PEITC in modulating the neoplastic effect of the carcinogen. The advantage of the use of this type of testing as a surrogate endpoint in these clinical Phase II trials is that it doesn't rely on cancer incidence reduction, much longer, costlier, and harder to obtain, to evaluate the efficacy of the chemopreventive agent. Successful chemopreventive agents would, therefore, be available sooner to those at risk.

REFERENCES

1. Boone CW, Kelloff GJ, Steele VE (1992): Natural history of intraepithelial neoplasia in humans with implications for cancer chemoprevention strategy. Cancer Res 52:1651–1659.
2. Boone CW, Kelloff GJ, Freedman FS (1993): Intraepithelial and postinvasive neoplasia as a stochastic continuum of clonal evolution, and its relationship to mechanisms of chemopreventive drug action. J Cell Biochem 17G:14–25.
3. Poel WE (1959): Effect of carcinogenic dosage and duration of exposure on skin-tumor induction in mice. J Natl Cancer Inst 22:19–43.
4. Lijinsky W, Saavedra JE, Rueber MD, Singer SS (1982): Esophageal carcinogenesis in F344 rats by nitrosomethylamines substituted in the ethyl group. J Natl Cancer Inst 68:681–684.
5. Weeks CE, Herrmann AL, Nelson FR, Slaga TJ (1982): Difluoromethlornithine, an irreversible inhibitor of ornithine decarboxylase, inhibits tumor promotor-induced polyamine accumulation and carcinogenesis in mouse skin. Proc Natl Acad Sci USA 79:6028–6032.
6. Stoner GD, Morrissey DT, Heur YH, Daniel EM, Galati AJ, Wagner SW (1991): Inhibitory effects of phenethyl isothiocyanate on *N*-nitrosobenzylmethylamine carcinogenesis in the rat esophagus. Cancer Res 51:2063–2068.
7. Bacus JW, Bacus JV (1994): Quality control in image cytometry: DNA ploidy. J Cell Biochem (Suppl)19:153–164.
8. Taylor SR, Titus-Ernstoff L, Stitely S (1989). Central values and variation of measured nuclear DNA content in imprints of normal tissues determined by image analysis. Cytometry 10:382–387.
9. Bacus JW, Hernicz RS, Bacus JV (1995): Dual wavelength imaging microdensitometry for immunocytochemical assays in microscopy. Proceedings Society of Photo-Optical Instrumentation Engineers SPIE 2622: 704–711.
10. Bacus JW, Grace LJ (1987): Optical microscope system for standardized cell measurements and analyses. Appl Optics 26:3280–3293.

11. Bacus JV (1994): The CAS 200(tm) MultiScan(tm) automated pathology workstation. In Wied GL (ed): "The Compendium on the Computerized Cytology and Histology Laboratory." Chicago: Tutorials of Cytology, pp 360–367.
12. Dawson AE, Cibas ES, Bacus JW, Weinberg DS (1993): Chromatin texture measurement by Markovian analysis. Anal Quant Cytol Histol 15:227–235.
13. Auer GU, Caspersson TO, Wallgren AS (1980): DNA content and survival in mammary carcinoma. Anal Quant Cytol 2:161–165.
14. Bacus JW, Bacus JV (1994): A method of correcting DNA ploidy measurements in tissue sections. Mod Pathol 7:562–664.
15. Bacus S, Flowers J, Press M, Bacus JW, McCarty K (1988): Evaluation of estrogen recepter in primary breast carcinoma by computer-assisted image analysis. Am J Clin Pathol 90:233–239.
16. Berchuck A (1991): Determination of proliferation index in advanced ovarian cancer using quantitative image analysis. Am J Clin Pathol 99:736–740.
17. Bacus SS, Bacus JW, Slamon DJ, Press MF (1990): HER-2/*neu* oncogene expression and DNA ploidy analysis in breast cancer. Arch Pathol Lab Med 114:164–169.
18. Press MF, Pike MC, Chazin VR, Hung G, Udove JA, Markowicz M, Danyluk J, Godolphin W, Sliwkowski M, Akita R, Paterson MC, Slamon DJ (1993): Her-2/*neu* expression in node-negative breast cancer: Direct tissue quantitation by computerized image analysis and association of overexpression with increased risk of recurrent disease. Cancer Res 53:4960–4970.
19. Broders AC (1926): Carcinoma grading and practical application. Arch Pathol Lab Med 2:376–381.
20. Gleason DF (1966): Classification of prostatic carcinomas. Cancer Chemother Rep 50:125–128.
21. McNeal JE, Bostwick DG (1986): Intraductal dysplasia. A premalignant lesion of the prostate. Hum Pathol 17:64–71.
22. Richart RM (1968): Natural history of cervical intraepithelial neoplasia. Clin Obstet Gynecol 5:748–784.
23. Black MM, Speer FD (1957): Nuclear structure in cancer tissues. Surg Gynecol Obstet 105:97–102.
24. Bloom HJG, Richardson WW (1957): Histological grading and prognosis in breast cancer. Br J Cancer 11:359–377.
25. Bacus JW (1995): Cervical cell recognition and morphometric grading by image analysis. J Cell Biochem (Suppl)23:33–42.
26. Dawson AE, Austin RE, Weinberg DS (1991): Nuclear grading of breast carcinoma by image analysis. Classification by multivariate and neural network analysis. Am J Clin Pathol 95(Suppl 1):29–37.
27. Geisler JP, Wieman MC, Zhou Z, Miller GA, Geisler HE (1996): Markov texture parameters as prognostic indicators in endometrial cancer. Gynecol Oncol 62:174–180.
28. Veltri RW, Miller MC, Partin AW, Coffey DS, Epstein JI (1996): Ability to predict biochemical progression using gleason score and a computer-generated quantitative nuclear grade derived from cancer cell nuclei. Urology 48:685–691.
29. Haralick RM, Caspell F, Simonett DS (1970): Using radar imagery for crop discrimination: A statistical and conditional probability study. Remote Sensing Environ 1:131.
30. Haralick RM, Shanmugam K, Dinsteit I (1973): Textural features for image classification. IEEE Trans Systems Man Cybernetics SMC-3:610–615.
31. Swets JA (ed) (1964): "Signal Detection Theory and Recognition by Human Observers." New York: Wiley, 1964.
32. Swets JA (1986): Indices of discrimination or diagnostic accuracy: Their ROCs and implied models. Psychol Bull 99:100–117.
33. Swets JA (1986): Form of empirical ROCs in discrimination and diagnostic tasks: Implications for theory and measurement of performance. Psychol Bull 99:181–198.
34. Bacus JW, Wiley EL, Galbraith W, Marshall PN, Wilbanks GD, Weinstein RS (1984): Malignant cell detection and cervical cancer screening. Anal Quant Cytol 6:121–130.
35. NCI Contract NO-CN-65024: A Phase II Trial of 9-*cis*-retinoic acid as a chemopreventive agent in patients with cervical intraepithelial neoplasia (CIN) grade 2/3, 1997 (in progress).