Image Morphometric Nuclear Grading of Intraepithelial Neoplastic Lesions with Applications to Cancer Chemoprevention Trials


Bacus Laboratories, Inc., Lombard, Illinois 60148 [J. W. B., J. V. B.]; Chemoprevention Program, Division of Cancer Prevention, National Cancer Institute, NIH, Bethesda, Maryland 20892 [C. W. B., G. J. K.]; and University of Texas M. D. Anderson Cancer Center, Houston, Texas 77030-4095 [M. F., V. K., S. M. L.]

Abstract
A new image morphometric method of nuclear grading is described and assessed in the context of the evaluation of histological samples from ductal carcinoma in situ of the breast and cervical intraepithelial neoplasia. The method results in a continuous scaled variable, or nuclear grading scale, expressed in SD units from measured normal nuclei from breast or cervix. For a given histological preinvasive neoplastic lesion, the mean nuclear grade of measured nuclei was shown to be analogous to the histopathological grade of the same lesion assigned subjectively by the pathologist. In a chemoprevention trial of the effect of difluoromethylornithine given for 1 month to subjects with cervical intraepithelial neoplasia grade 3, pathologists could see no difference in 14 histological sections taken before and after difluoromethylornithine treatment. However, the image morphometric method detected a systematic effect of lowered mean nuclear grade and a decrease in the variability of nuclear grade expression. Twelve of 14 samples showed a lower posttreatment mean nuclear grade (P < 0.05), and 13 of the 14 samples showed a decrease in the SD of their nuclear grade distributions (P < 0.01). This study demonstrates the use of image morphometric nuclear grading in a chemoprevention setting. It may be very useful in supplementing the pathologist’s histopathological grading by providing objective, quantitative assessments.

Introduction
Visual histological and cytological grading systems for cancer diagnosis and prognosis have a long history (1). They have been described for cancers of most organ systems, including breast cancer (2,3), cervical cancer (4), and prostate cancer (5, 6). Computer-assisted image morphometric analysis is being developed to quantitate and standardize these grading systems. It has been used predominately to evaluate the prognosis of established cancers (7–10). However, some investigators have used these methods to assess preinvasive neoplastic progression and have proposed quantitative indexes of progression (11–13).

During the last 10 years, some methods of quantitative microscopic image analysis have become accepted in the clinical laboratory (14). Assays have been developed for DNA ploidy (15, 16), estrogen receptors (17, 18), proliferation status (19, 20), and oncoprotein expression (21). Methods of morphometric grading for the evaluation of preinvasive neoplasia have yet to be accepted in the same sense and are still under development. The methods reported here are aimed at rigorously applying image morphometric analysis to quantitate the progression of preinvasive neoplasia, and especially the use of these methods in evaluating the effect of chemopreventive agents.

The natural history of intraepithelial neoplasia, called equivalently preinvasive neoplasia, precancer, dysplasia, or CIS1 [equivalent to very severe dysplasia (22)] has been reviewed (23–27). In brief, intraepithelial neoplasia generally consists of multiple foci of an abnormal clonal expansion of neoplastic cells that progress independently. In many preinvasive neoplastic lesions, e.g., skin (28), oral leukoplakia (29), larynx (30, 31), lung (32), uterine cervix (33, 34), esophagus (35), stomach (36), and colon (37, 38), the degree of abnormal morphological variation of cells and their nuclei, seen either in histological sections or cytology smears, is used to estimate the risk of further progression to invasive disease and clinical cancer. Fig. 1 illustrates the morphological progression of intraepithelial neoplasia from low grade to high grade, manifested by increasing size of the lesion and a greater degree of cytonuclear morphological aberration.

Conventional clinical trials of chemopreventive agents based on the end point of cancer incidence reduction usually require hundreds to thousands of subjects and many years of observation to complete. The need to screen large numbers of candidate chemopreventive agents in clinical trials has made it necessary to conduct trials with intermediate end point biomarkers that correlate with cancer incidence reduction, measured in histological sections of biopsied tissue and in cytological smears, so that the trials may be shorter (1 year or less) and may require fewer subjects (100 or fewer). We have developed a method using image morphometric analysis to measure intermediate end point biomarkers and their response to candidate chemopreventive agents with improved objectivity, sensitivity, and precision both in animal chemoprevention models and in Phase I and Phase II human clinical trials. We reported this method previously using image morphometric histological
grading of precancerous lesions for use in animal chemoprevention studies (39). We now report the assessment of this image morphometric method for the nuclear grading of human premalignant lesions. Specifically, we assessed this imaging method in the chemopreventive context of premalignant lesions of the breast and cervix.

Materials and Methods

Breast DCIS. We used histological sections from 120 patients given a diagnosis of breast DCIS by a panel of five expert breast pathologists (40). The panel used a recently updated system of grading breast DCIS (41). Of the 120 cases, there was initial unanimous agreement by all five pathologists in 16 cases of nuclear grade 3, 24 cases of nuclear grade 2, and in none of the cases of nuclear grade 1. We used only the histological sections, for which there was immediate unanimous agreement, to determine the optimal nuclear features to be measured in setting up a continuous grading scale for breast DCIS. Five nonneoplastic surgical specimens of normal breast tissue were also used.

CIN. We used histological sections from a group of 15 CIN cases and 4 normal cases to determine the best nuclear morphological features to be used for a nuclear grading scale of CIN. To test the effectiveness of the final image morphometric grading method, we evaluated histological sections from a second group of 14 CIN-3 patients who had been given oral DFMO over a range of doses between 0.06 and 1.0 g/M2/day for 1 month (42). An inclusion criterion for this DFMO study was that patients must have a CIN-3 lesion involving more than one-third of the cervical epithelium. After pretreatment biopsy and administration of the DFMO for 1 month, the posttreatment biopsies were taken as distant as possible from the pretreatment and administration of the DFMO for 1 month, the posttreatment biopsies were taken as distant as possible from the pretreatment.

For both the breast DCIS and CIN studies, adjacent microtome sections of paraffin blocks were prepared; one was stained with H&E, and the other was stained with the Feulgen stain for DNA. The section thicknesses were nominally 5 μm in the DCIS specimens and 4 μm in the CIN specimens. The Feulgen DNA staining of histological sections was done using the DNA ploidy analysis kit associated with the CAS 200 Image Analysis System, which has quality control procedures that provide for reproducibility of absorbance and chromatin texture measurements between laboratories, including standardization with known DNA samples and the use of calibration cells to control for variations in section thickness, light source intensity, and other factors (15, 43). The most typical lesions were outlined with a marker pen on each H&E-stained slide by the panel of five pathologists in the case of the breast DCIS slides. The marked outlines provided easy reference to guide image analysis of nuclei on the Feulgen-stained slides. In the case of the CIN slides, the entire lesion was used.

Instrumentation. The H&E-stained slides were scanned by the BLISS Image Analysis System to obtain WebSlides for documentation and confirmation of the areas that were analyzed. Nuclear images were extracted from the Feulgen-stained histological sections as ILM files using the CAS200 Quantitative DNA software program (15). Fig. 2 illustrates this process. Nuclear features were measured on these individual nuclei using the Nuclear Grade software program (Bacus Laboratories, Inc.). The approach used was to characterize neoplastic growth quantitatively with multiple measurements on each nuclear image and to combine these measurements into a single number for each nucleus, i.e., a nuclear grade per nucleus. Parameters from the distributions of these measurements, e.g., the mean and SD, were then used to represent the nuclear grade for the subpopulation of cells considered. Four features were combined in the breast DCIS studies and five in the CIN studies.

Selection of Nuclear Morphological Features to Be Measured. The nuclear morphological feature measurements used to characterize each nucleus were selected from the list given in Table 1 (44). These feature measurements are of four basic categories: (a) measurements related to the overall DNA content of the nucleus; (b) measurements of nuclear dimension such as nuclear area, nuclear shape (circularity), maximum nuclear diameter, and others; (c) Markovian texture measurements that summarize differences in absorbance between a reference pixel and other pixels at defined distances from the reference pixel for the entire nucleus; and (d) counts per nucleus of defined individual point texture measurements used to measure alterations in fine chromatin texture. For nuclear grading of both breast DCIS and CIN, ~100 different types and combinations of measurements from these four categories were evaluated for discriminatory effect and suitability. The criteria used for selecting nuclear morphological features were the following: (a) the features produced a Gaussian distribution of nuclear grades in normal cell samples; (b) they produced a monotonically increasing nuclear grade scores from lower to higher pathological grades, with the highest nuclear grade at the end point of the neoplastic spectrum; and (c) their measurements visually agreed with expert human subjective assessment.

![Image of morphological progression of neoplasia from its intraepithelial beginning to invasive and metastatic disease.](https://example.com/image.png)
of the pathological grade of individual nuclei. The individual point texture measurements were displayed where they occurred in individual nuclei so that they could be confirmed as being characteristic of the chromatin pattern in the nucleus. All nuclear grades were listed for “point and click” review of the corresponding nuclear images. Specific feature measurements that contributed to the nuclear grade could be overlaid on the nucleus if desired.

Z-Score Transformation of Nuclear Morphological Measurements. To create a common scale for nuclear grading equally applicable to different types of epithelia, we performed a Z-score transformation of the raw data measurements of each nuclear feature in each nucleus for both the breast DCIS and CIN lesions, as illustrated in Fig. 3. To do this, we first measured a set of morphological features on a sample of normal epithelial nuclei of the same tissue type as the intraepithelial neoplastic cells to be studied. The mean and SD of these measurements were then used to transform subsequent measurements on neoplastic nuclei into Z-scores scaled in SD units (Z-units). The process of Z-score transformation was carried out separately for each nuclear feature and for both breast DCIS and CIN lesions.

A similar procedure is used in multivariate analysis to allow the use of many different measurement scales in a multidimensional space. However, in that case, all of the data from usually a number of different classes, or categories, is lumped together. The method described here is different because it is based on only the normal nuclei. It accomplishes the same thing in a scale sense but additionally preserves a specific reference to normal cells and places them at zero on the nuclear grade measurement scale.

Others have proposed indexes of atypicality and progression (11–13). We feel that the method presented here has merit for a number of reasons. It automatically tends to adjust for differences in variances from one measurement to another and differences in variance because of the use of a different target tissues for evaluation, e.g., cervix versus breast. It also tends to adjust for sample preparation variances, such as sectioning and fixing, if the normal specimens are prepared in conjunction with the testing specimens. In all cases, the scale transformation adjusts the measurement scale for the variances in normal nuclei and imposes the scale adjustment to the measurement of neoplastic nuclei. In addition to the above, it provides a theoretical link to psychophysical testing, signal detection theory, and receiver operating characteristics and methodology, where these basic normalization models and this same z-score scale are assumed (45–47). It allows for the direct computation of the receiver operating characteristics curve for the discrimination of neoplastic from normal in a sampled population such as those shown in Figs. 4 and 6 (47).

The method of combining the Z-scores of selected nuclear morphological feature measurements into a single nuclear grade for each nucleus is illustrated in the lower part of Fig. 4. By this method, the nuclear grades of normal epithelia are always distributed about a mean value of zero. The individual feature Z-scores are weighted by a coefficient and summed to obtain the final nuclear grade. The weighting coefficients are obtained by Fisher linear discriminate analysis (48) of the \( n \)-dimensional mean differences of preinvasive neoplastic nuclei compared with the reference population of normal nuclei. For breast DCIS, the four nuclear features and their weighting coefficients used were: area, 0.74; DNA, 0.62; entropy, 0.21; and valley, 0.10. For CIN, the five nuclear features and their weighting coefficients used were: sum average, 0.59; circularity, 0.49; maximum diameter, 0.38; correlation, 0.37; and product moment, 0.35. The nuclear morphological feature with the greatest discriminating capacity has the highest weighting coefficient. In many nuclei, the contribution of morphological features with smaller weighting coefficients were still important because their measured values were high enough to override their smaller weights.

Fig. 5 illustrates the technique of Fisher linear discriminant analysis for the simplest case of only two nuclear morphological features, DNA and entropy, and indicates geometrically the concept of the Z-score axis derived from multidimensional measurements. The coordinates of each point in the figure are the values of DNA and entropy measured in one nucleus. The straight line (Z-axis line) through the cluster
of points represents the final continuous grading scale in Z-score SD units. A representative point is shown projected onto the Z axis line to give a nuclear grade of 9.00, measured from the origin to the foot of the perpendicular. The frequency distribution of nuclear grades in the upper part of Fig. 4 is the result of projecting all of the individual nuclear grades measured in a given histological section onto this linear discriminate Z-axis line. Various descriptive statistics, e.g., the mean, SD, percent exceeding normal limits, and others, may then be applied to describe the frequency distribution of nuclear grades.

**Results**

**Breast DCIS.** Fig. 6 shows the image morphometric nuclear grade distributions in histological sections from two patients with breast DCIS assigned pathological nuclear grades of 1 and 3, respectively. The distributions are superimposed for easy comparison. As described in “Materials and Methods,” the Z-score scale is in SD units. The nuclear grade distribution of tissue diagnosed as pathological grade 1 is shifted slightly to the right. The mean morphometric nuclear grade was 2.05, and the SD of the distribution was 1.2, as compared with a mean of 0.0 and SD of 1.0 expected for normal nuclei. By contrast, the nuclear grade distribution of tissue diagnosed as pathological grade 3 is shifted much farther to the right and has a greater SD. In this case, the mean morphometric nuclear grade was 7.50, and the SD of the distribution was 2.05. We concluded from the examination of many such distributions that the image morphometric nuclear grade for any given case could be adequately expressed by the mean and SD of the nuclear grade distributions. Because the distributions were reasonably Gaussian, they were suitable for parametric statistical testing, such as the t test to compare the mean cytometric nuclear grades between patients or cohorts of patients.

Fig. 7 presents the mean image morphometric nuclear grades of all 40 cases of breast DCIS for which there was immediate unanimous agreement. There was little overlap of the continuously scaled mean morphometric nuclear grades compared with the pathological ordinal grades of normal, grade 2, and grade 3. In this collection of DCIS specimens, there were only nine cases of pathological grade 1, arrived at by discussion after initial grading. None of these cases were unanimously agreed upon prior to discussion. Because there were so few cases and no prior consensus in grade, these cases were not used to develop the model. They were, however, included for completeness in the full 120 cases presented in Fig. 8 illustrating progression. After the basis of the data from the cases of unanimous agreement of pathological grades 2 and 3, we concluded that suitable cutoff points corresponding to these two patho-

<table>
<thead>
<tr>
<th>Cell No.</th>
<th>Nuclear image</th>
<th>Raw measure</th>
<th>Z-scores for each nucleus</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>x₁</td>
<td>Z₁ = (x₁ − X)/SD</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>x₂</td>
<td>Z₂ = (x₂ − X)/SD</td>
<td></td>
</tr>
<tr>
<td>...</td>
<td>...</td>
<td>...</td>
<td></td>
</tr>
<tr>
<td>n</td>
<td>xₙ</td>
<td>Zₙ = (xₙ − X)/SD</td>
<td></td>
</tr>
</tbody>
</table>

\[ \bar{X} = \text{Mean of feature measurements in normal nuclei} \]
\[ \text{SD} = \text{Standard deviation of feature measurements in normal nuclei} \]
logical grade categories were 1.5 and 4.5, i.e., a mean morphometric nuclear grade of <1.5 was either grade 1 or normal, between 1.5 and 4.5 was grade 2, and >4.5 was grade 3.

Having devised a continuous morphometric nuclear grading method that adequately separated histopathological grading categories, we then evaluated all 120 breast DCIS cases. Their grades are shown in ranked order in Fig. 8. The figure fully demonstrates the continuous nature of the quantitative image morphometric nuclear grading scale. Mean nuclear grades increase sharply from pathological grade 1 to grade 2, increase more slowly and linearly to nuclear grade 3, and then increase sharply again to much higher mean morphometric nuclear grades.

Fig. 9 depicts the results of the mean morphometric nuclear grade analysis of 14 patients with CIN-3 compared with 4 normal subjects. Each plotted point represents the mean of a distribution of nuclear grades of 300 or more nuclei in a histological section from one subject. There was no overlap between the mean morphometric nuclear grades of subjects with CIN-3 and subjects with normal cervical epithelium, the measurement values of which cluster around zero. Whereas the subjective grades assigned by the pathologists were limited to three or four ordinal categories, the mean morphometric nuclear grade is a continuous variable that ranges from zero to any positive or negative value.

A subset of 14 patients with CIN-3 lesions from a pilot chemoprevention trial of oral DFMO (see “Materials and Methods”) was analyzed. To the pathologist, none of the histological sections from cervical biopsies of the 14 patients treated with DFMO for 1 month exhibited any histopathological change whatsoever after DFMO treatment. On the other hand, image morphometric nuclear grading of the same histological sections revealed a statistically significant difference after treatment with DFMO. Fig. 10 illustrates how the distribution of image morphometric nuclear grades in the posttreatment samples became left shifted toward normal compared with the distribution...
of morphometric nuclear grades in the pretreatment samples. In addition, treatment with DFMO was associated with a decrease in the variability of the distributions.

The complete data set for the mean morphometric nuclear grade and SD of nuclear grade distributions of all 14 patients is given in Tables 2 and 3, respectively. Table 2 shows that the mean nuclear grade of 12 of the 14 patients shifted toward normal after DFMO treatment. A paired sample \( t \) test comparing the average of mean morphometric nuclear grades of pre- and posttreatment samples rejects the null hypothesis at a significance level (two-tail) of \( P < 0.00942 \). Table 3 shows that the SD of mean nuclear grades of 13 of the 14 patients also decreased after DFMO treatment. A paired sample \( t \) test comparing the average SD of nuclear grades of pre- and posttreatment samples rejects the null hypothesis at a significance level (two-tail) of \( P < 0.0435 \).

### Discussion

These results demonstrate the effective application of an image morphometric nuclear grading system to breast DCIS and CIN. The system uses a continuous grading variable scaled in SD units that is the weighted sum of selected nuclear morphological feature measurements, where the weighting coefficients are obtained by Fisher linear discriminant analysis. In a given histological section, the mean of the frequency distribution of morphometric nuclear grades may be taken as representative of the pathological status of the tissue, analogous to a histopathological grade subjectively assigned by the pathologist to the same tissue. Because the nuclear grading system is scaled in SD units referred to normal tissue of the same type as that of the neoplastic nuclei being graded, nuclear grades of different neoplastic tissues may be compared on a uniform scale.

The image morphometric nuclear grading results for breast DCIS presented in Figs. 7 and 8 were obtained from measurements of nuclear size, nuclear DNA content, and two chromatin texture features related to the degree of chromatin granularity. These features are practically the same as those used by pathologists to estimate the histopathological nuclear grade. Because of this, morphometric nuclear grading was in good agreement with the pathological nuclear grade category assigned by the panel of pathologists. Image morphometric nuclear grading exhibited increased resolving power compared with pathologival feature measurements.
Tissue nuclear grading, as illustrated by the wide range of nuclear grade measurements in the DCIS-3 category shown in Fig. 7. The continuous nature of the morphometric nuclear grade is well illustrated in Fig. 8.

The modulation of image morphometric nuclear grades in histological sections from 14 patients with CIN-3 produced by oral administration of DFMO illustrates how the method permits statistically significant results to be obtained from a relatively small group of patients. Read by the pathologist, these same histological sections appeared to show no differences. By contrast, image morphometric nuclear grading revealed systematic changes after the administration of DFMO that were significant at the \( P < 0.05 \) level.

One explanation for the left-shift toward normal of the frequency distribution of nuclear grades observed after the administration of DFMO relates to the proliferation rates of cervical dysplasias. Richart (4) has shown that the tumor cells of CIN with higher histopathological grade (higher individual nuclear grades) have much faster proliferation rates (4, 23). The proliferation-suppressing effect of DFMO may have differentially suppressed these faster growing, higher grade tumor cells so that they became overtaken by more slowly growing, lower grade tumor cells. Another explanation may be that the repeat biopsy was the sole cause of the regression. In this regard, it should be noted that the repeat (posttreatment) biopsies were taken as distant as possible from the pretreatment site to avoid this. Whatever the explanation may be for this effect, these results were not presented to validate the chemopreventive in the pilot study. Further studies of a more complex design would be necessary for this. It is clear, however, that a subtle but quantifiable effect was measured, and that this is similar to that which might be obtained in a chemoprevention trial. The results were included to illustrate the usage of the quantitative nuclear grading method in a chemoprevention trial using small samples. This was especially important in this instance because the normal pathological examination of these cases was not sensitive enough to detect the subtle shift to a less severe dysplasia.

The subjective histopathological grading systems established for breast cancer (2, 3), cervical cancer (4), and prostate cancer (5, 6), among others, have well-known limitations of reproducibility, accuracy, and resolution of grading scale. The nuclear morphological changes quantitatively measured by our image morphometric nuclear grading system are similar to those pathologists use in assessing histopathological grade. Therefore, image morphometric nuclear grading may be very useful in supplementing the pathologist’s histopathological grading in a chemoprevention setting by providing objective, quantitative, and reproducible measures of nuclear morphometry.

References

References


Image Morphometric Nuclear Grading of Intraepithelial Neoplastic Lesions with Applications to Cancer Chemoprevention Trials


Updated version
Access the most recent version of this article at:
http://cebp.aacrjournals.org/content/8/12/1087

Cited articles
This article cites 41 articles, 5 of which you can access for free at:
http://cebp.aacrjournals.org/content/8/12/1087.full.html#ref-list-1

Citing articles
This article has been cited by 14 HighWire-hosted articles. Access the articles at:
/content/8/12/1087.full.html#related-urls

E-mail alerts
Sign up to receive free email-alerts related to this article or journal.

Reprints and Subscriptions
To order reprints of this article or to subscribe to the journal, contact the AACR Publications Department at pubs@aacr.org.

Permissions
To request permission to re-use all or part of this article, contact the AACR Publications Department at permissions@aacr.org.