Quantitative Grading of Tissue and Nuclei in Prostate Cancer for Prognosis Prediction

William A. Christens-Barry and Alan W. Partin

Prostate cancer is evidenced by profound histological, cellular, and nuclear changes in the organization of the prostate. Histological assessment of prostate tissue taken from surgically removed tumors has traditionally relied on visual pathological interpretation. The principal classification scheme used in visual pathology of the prostate, Gleason grading, has proven successful in characterizing the state of disease but has had limited prognostic value. We are conducting studies that aim to provide quantitative measures of disease state to improve prognosis prediction. Our findings show that the orientational distribution of tumor cells, often assumed to be isotropic, can play a significant role in statistical studies of intranuclear DNA organization. Inclusion of anatomical factors in the selection of reference frames for measurements of intranuclear DNA can improve the statistical power of cytometric studies and may provide a unifying framework for relating histological, morphometric, and intranuclear descriptions of prostate tumors.
(Keywords: Image cytometry, Prostate histology, Texture analysis.)

THE PROSTATE: AGING AND NEOPLASTIC DISEASES

Throughout childhood and adolescence, the prostate, central to many urologic functions in men, grows because of continuous cell proliferation. During early adulthood, cell proliferation comes into balance with normal cell death, and prostate size stabilizes. The walnut-shaped prostate (Fig. 1) maintains a constant (about 3-cm) length throughout early adulthood and the middle years. The bulk of the prostate is organized into glands that consist of central luminal regions into which a surrounding layer of epithelial cells secrete fluids that add to the ejaculate. The glands, in turn, are connected by a ductal system that collects the glandular secretions and carries them toward the urethra. This system looks somewhat like a cluster of grapes joined by a branched system of stems. The tissue between the epithelial glands is surrounded by stroma that gives the prostate its structural strength.

During the sixth and seventh decades of life, 60 to 70% of men experience benign prostatic growth or neoplasia. This resumption of growth, although usually

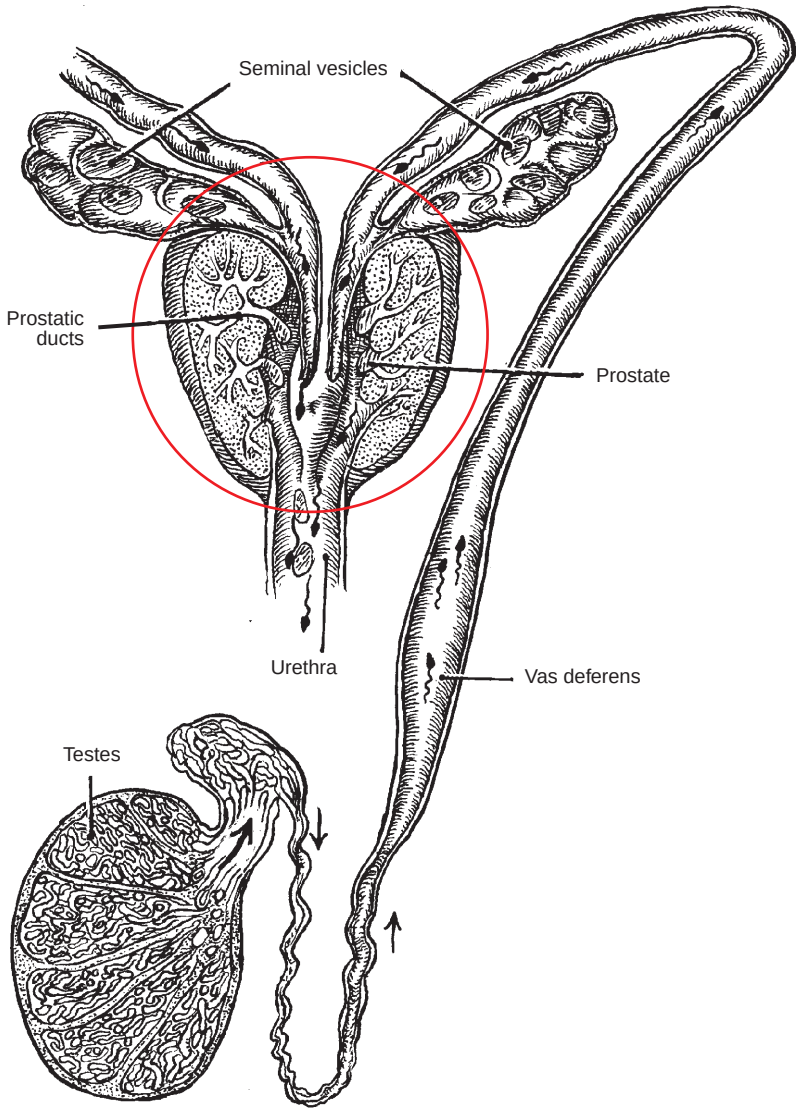

Figure 1. Anatomy of the male reproductive system. Glands within the prostate produce secretions that are collected by a ductal system and added to seminal fluid. (Reprinted from Walsh, P. C., and Worthington, J. F., *The Prostate: A Guide for Men and the Women Who Love Them*, p. 9, The Johns Hopkins University Press, Baltimore, MD, 1995, by permission.)

benign, can cause moderate levels of annoyance and discomfort (e.g., irritative or obstructive voiding symptoms) in some men. Others may experience more significant consequences leading to pronounced urinary problems and infection or benign prostatic hyperplasia. This condition is not itself life threatening and can be managed through a range of pharmaceutical and surgical approaches.

A different, more troubling prospect faced by many men in their 60s and 70s is the development of adenocarcinoma of the prostate (CaP), or prostate cancer, the single most frequent form of cancer in men, affecting nearly one in nine. During 1997, about 370,000 men in the United States will be diagnosed with CaP; of these, 44,000 are expected to die from this disease. CaP causes death principally by metastasizing, that is, the primary tumor expands and spreads to distant (secondary) sites. DNA provides the genetic program that defines all the proteins that are synthesized within cells and ultimately regulates the processes by which this synthesis takes place. Specific genes (proto-oncogenes and oncogenes) have been implicated in the "transformation" of cells to a state of hyperproliferation (cancer), which disrupts the cooperation among cells of an organ or tissue. Highly invasive cells (metastases) can break away from the tumor site and be carried by the lymphatic or circulatory systems to distant organs. Metastasis is the most troubling aspect of cancers because localized treatment or surgical removal of a primary tumor is frequently inadequate to prevent the formation of secondary tumors elsewhere in the body. To date, no complete cure for advanced CaP is available, and because the disease is often fatal, an urgent need exists to improve diagnosis, prognosis prediction, and disease outcome.

Our understanding of the factors that play a role in the etiology of prostate cancer or its evolution from a localized malignancy to a life-threatening metastatic disease is not clear. Clinical research arises from the need to diagnose cancer, to define (or "stage") the status of neoplastic disease in a patient at a given time, to predict the likely course or "prognosis" of the disease, and finally to develop treatments that have a high probability of success for a patient at a given stage.

Researchers in APL's Milton S. Eisenhower Research and Technology Development Center recently received funding from the National Cancer Institute to collaborate with a group of scientists and physicians in the Department of Urology at the Johns Hopkins Medical Institutions (JHMI). The effort is focused on the identification of cytometric markers that correlate with stage and prognosis. In particular, optical microscopy of cancer cell nuclei has indicated that measurements of nuclear shape (nuclear morphometry) and investigation of the distribution of histone protein/DNA complexes (known as chromatin) within the

nucleus can be used to characterize the disease state and predict its progression. We are applying these techniques of image cytometry and quantitative analysis to problems of nuclear shape, chromatin DNA distribution, and histology of both the normal prostate and prostate cancer. Our broad goal is to improve patient outcome by developing accurate assessment techniques for use in clinical disease management.

LABORATORY INVESTIGATIONS

Our current studies mainly concentrate on understanding and describing two structural aspects of prostate cancer: (1) the characterization of tissue architecture and organ function and (2) a statistical description of nuclear chromatin DNA in normal and diseased prostate tissue. A quantitative description of these structures can help us predict prognoses so that appropriate curative measures can be selected for individual patients.

Pathological Grading of Prostate Tissue

The most reliable basis for establishing appropriate follow-up therapy for CaP is provided by the visual appearance of tissue specimens. Of several pathological grading schemes, Gleason grading[1] is the most clinically useful (Fig. 2). Gleason grading can accurately predict prognosis using only the visual assessment of thin sections of tissue that have been stained and fixed following removal of the prostate (radical prostatectomy) or from biopsy material. However, training pathologists to achieve reproducible Gleason grading is difficult, and numerous studies have shown an undesirable level of inter- and intra-observer variability (see Epstein[2] for a lucid explanation of the use of Gleason grading in prostate pathology). We are developing quantitative descriptors of tissue organization that can be used to augment Gleason grading with an objective algorithmic technique amenable to automation.

The histopathological hallmark of CaP is the absence or diminution of well-formed glands, which is apparent in the varying size, shape, and spatial organization of the glandular lumina. These lumina serve as collection areas for the secretions of epithelial cells organized around the interior margins of the gland. In addition, with progression of CaP, the secretory cells themselves become disordered and fail to form glands.

For our study, surgically removed tumors were sectioned to a 4-μm thickness and stained with hematoxylin and eosin (H&E), which are used to provide contrast between cytoplasmic and nuclear components, respectively. Brightfield images of these sections at magnifications of 10, 20, and 40× were digitized using a red/green/blue color model. Digital image processing allows segmentation of glandular lumina (using color spectrum content, statistical properties of the luminal color image, and Sobel and Laplacian-of-Gaussian filtering), cells (based on identification of cell membranes), and nuclei (using local contrast and color spectrum content).

Analysis of the stained tissue sections was based on three measures: (1) histopathological order (e.g., the glandular luminal area, perimeter, and the number and form factor), which describes the overall shape of each gland and the organization of the prostate into discrete glandular regions; (2) Voronoi analysis and Delaunay tessellation neighbor analysis (Fig. 3), which characterize the effective use of gland volume; (3) and neighborhood graphs based on the distribution of epithelial and stromal cell nuclei, which describe the proximity of secretory cells to the ductal system. Our recent work indicates that increasing Gleason grade correlates with

- Decreasing quantitative measures of order in prostate tissue, based on a model that relates glandular functionality to the size and shape of the lumina
- Broadening of the distribution of luminal areas as measured by the coefficient of variation (COV) of this distribution[3]
- Decreasing organization of secretory cells around the luminal perimeter as measured by the density of nuclei per length of the luminal boundary

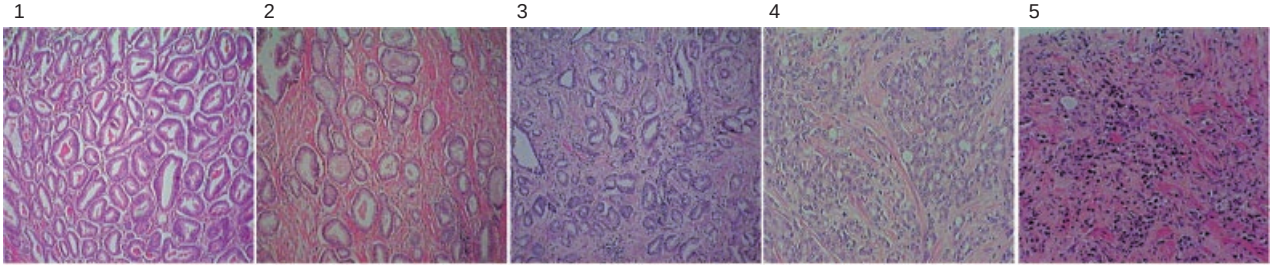

Figure 2. Gleason grading, the most reliable visual assessment scheme for grading tumors, is based on the histological appearance of prostate tissue. Each region of a specimen is assigned a Gleason grade from 1 to 5, with higher grades corresponding to worsening stages of cancer. A Gleason score (2–10) is then formed based on the sum of the two most predominant grades. Shown are hematoxylin- and eosin-stained sections with Gleason grades (left to right) 1 through 3 (at 10×) and 4 and 5 (at 20×). Reduced organization of cells (identified by dark nuclei) into functional glands and decreasing size and order of glandular lumina (white regions) are apparent at higher Gleason grades. Panel-to-panel color differences reflect typical staining variability.

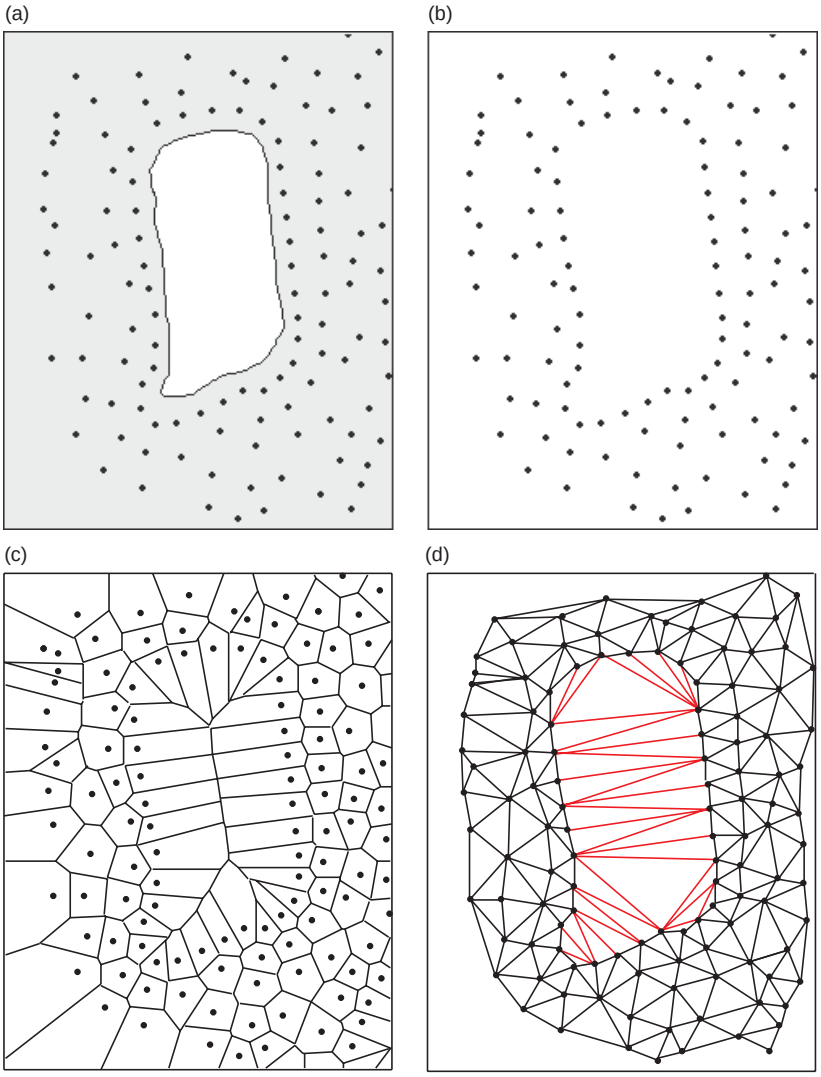

Figure 3. Identification of Voronoi cells and Delaunay tessellations from H&E-stained prostate tissue. (a) Image indicating cell nuclei (dark dots), cytoplasm (gray), and luminal region (white). (b) Segmented and thresholded image showing locations of cell nuclei. (c) Voronoi cell diagram of nuclei. (d) Delaunay graph constructed from Voronoi cells; edges within lumen (red) are subsequently removed using a mask based on the luminal region in Fig. 3a. Analysis of the distribution of Voronoi cell sizes and Delaunay tessellation edge lengths can be used to evaluate the organization of prostate tissue into functional glands.

DNA Image Cytometry

After much long-term research, staff at the JHMI Brady Urological Institute have combined Gleason grading with intranuclear DNA descriptors to create a model with great predictive power in classifying cells according to their disease state.[4–6] Because Feulgen staining of nuclei is stoichiometric with DNA content, we have recorded high-magnification (40×) images of Feulgen-stained tissue (Fig. 4). These images are used for statistical analysis of DNA in individual nuclei to study prognosis prediction after surgery.

In an effort to relate nuclear chromatin distribution to CaP prognosis, we are using a statistical image processing technique (see the boxed insert) broadly applied in remote sensing to statistically analyze the "textural" content of an image. Statistical texture is a visual property that describes characteristic statistical features of an object or region that may lack well-defined structural definition. In remote sensing applications, texture methods are used to distinguish crops or sea states on the basis of second-order statistics. Alternatively, subregions of an image with differing structural details but with stationary statistical properties can be assigned to a common class based on their measured textures.

For our application we used data acquired from Feulgen-stained tissue blocks taken from patient tumors removed during radical prostatectomy surgery by Patrick C. Walsh at the Johns Hopkins Hospital. Several innovations have improved our ability to retrospectively predict patient prognosis.

Laboratory Frames of Reference

The *a priori* assumption that nuclei are rotationally random is often used to justify the use of a common laboratory frame in which to construct co-occurrence matrices for all nuclei from a given tumor. Using such an approach, the distribution of values of a feature that is measured for the set of nuclei will, in general, be broad, even if all the nuclei are identical. The variance of the distribution of feature values has been found in many

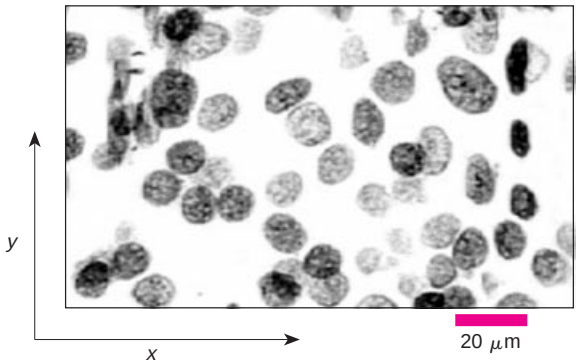

Figure 4. Feulgen-stained thin section (40×) in which nuclei are shown as dark, ovoid structures of varying eccentricity. The spatial variation of the intensity within a nucleus is a measure of chromatin DNA condensation and the concentration projected onto the *x*–*y* plane.

CO-OCCURRENCE MATRICES AND MARKOV TEXTURE FEATURES

A statistical description of the image is often the most powerful approach to segmentation and classification. This is especially true for images that lack exact structural exemplars, for example, sea surfaces or forests from aerial images. Here, no two regions or images are alike, yet the human eye can easily classify such regions by their statistical properties or textures. The co-occurrence matrix describes the second-order statistics of pixel pairs comprising an image or region. For such an image, many different co-occurrence matrices can be constructed, each differing in the spatial relationship of the pixel pairs that are used to construct it.[7] Thus, a particular co-occurrence matrix has a pair of parameters (a, b) describing this spatial relationship. For a given pair (a, b), each element $P_{ab}(i, j)$ of the co-occurrence matrix describes the joint probability that a pixel pair having gray scale values of i and j is separated by distances a and b in the x and y directions, respectively. For an image with a dynamic range of N (i.e., the intensities in the image range from 0 to $N - 1$), the co-occurrence matrix for a parameter pair (a, b) is constructed according to the following steps.

1. An $N \times N$ matrix is constructed, with all elements $P_{ab}(i, j) = 0$.
2. For a pixel pair $I(x, y) = i$ and $I(x + a, y + b) = j$, add 1 to $P_{ab}(i, j)$ and 1 to $P_{ab}(j, i)$.
3. Repeat step 2 for all pixel pairs $I(x, y)$ and $I(x + a, y + b)$.
4. Divide P_{ab} by an integer equal to twice the number of pixel pairs used in steps 2 and 3.

Several points regarding the co-occurrence matrix constructed in this manner should be noted. Because both $P_{ab}(i, j)$ and $P_{ab}(j, i)$ are incremented for each pixel pair meeting the separation criterion, the co-occurrence matrix is symmetric about the main diagonal. Division by twice the number of pixel pairs is just a normalization, so that the final matrix contains probabilities rather than pixel pair counts. An alternative definition of the co-occurrence matrix maintains each pixel pair as a directionally ordered pair: only element $P_{ab}(i, j)$ is incremented for the pixel pair $I(x, y) = i$, and $I(x + a, y + b) = j$. This little-used definition produces a co-occurrence matrix that is asymmetric about the main diagonal and is mathematically preferable since each element of the matrix is independent. Thus, the asymmetric definition preserves greater directional information about the underlying image. Moreover, the symmetric matrix can be recovered from the asymmetric form if desired.

The co-occurrence matrix can be large, in that it is indexed on the different gray scales present in an image. For an image with 8 bits of dynamic range, the co-occurrence matrix will be 256×256. Because use of a matrix of this size is often computationally intractable (several features require eigenvector calculation), it is often expeditious to perform a dynamic range reduction via a histogram equalization procedure. Last year we developed an iterative bifurcation procedure for dynamic range reduction with equalization that uses a maximum entropy method to improve upon traditional cumulative distribution function procedures.[8]

To characterize an image by its co-occurrence matrix, numerous quantitative features have been defined. For example, consider the values of the entries of a co-occurrence matrix: if the values along or near the main diagonal are large, one can conclude that pixel values vary slowly relative to the separation (a, b); if the values in the matrix are all roughly equal, then the image it describes has a random structure. If the second moment of the matrix about the main diagonal is large, the image it describes has large amounts of high contrast. Other commonly used features include correlation, contrast, entropy, sum moments, and difference moments.

studies to correlate with disease state and prognosis better than the mean values of these features. Yet, there has been no systematic examination of the features' orientational dependence. In addition, the degree to which the variance of features with good predictive power arises from the underlying spatial distribution of DNA in the nucleus—rather than from the rotational arrangement of cells in the tissue—has not been addressed. We have examined these issues in recent laboratory experiments.

Numerical and Optical Distributions of Identical Nuclei

Feature variance that arises in a population of identical nuclei with varying orientations is demonstrated in Fig. 5. This population of nuclei is composed of duplicates of a single nucleus placed at angular increments of 30°. The COVs of several Markov textural features are graphed as a function of angle in Fig. 6; the variation in phase and amplitude for several such features is evident.

Because numerical rotation of an array can cause small interpolation errors,[9] we designed an optical experiment that avoids this effect. A Feulgen-stained slide was mounted at several different rotational angles on the microscope stage (Fig. 7) by our collaborators at UroCor, Inc. A fixed set of 16 nuclei was randomly selected for imaging and digitization at angles ranging from 0 to 180° using angular intervals of 30° (yielding 112 images). To understand the role of anatomical orientation, all 7 orientations of a specific nucleus were pooled (nuclear groups, G_N); to examine how variance arises in heterogeneous groups as a result of the choice of a common reference frame, all 16 nuclei at a specific orientation (e.g., at 60°) were pooled (rotational groups, G_R).

Analysis was carried out separately for each of the 16 nuclear and 7 rotational groups. The variance observed among nuclear groups, measured by pool COVs, was typically less than that found for rotational groups; the mean value of $COV(G_N)/COV(G_R)$ for all features was approximately 0.27, but for some features it was comparable in size, $COV(G_N)/COV(G_R) > 0.7$, to that

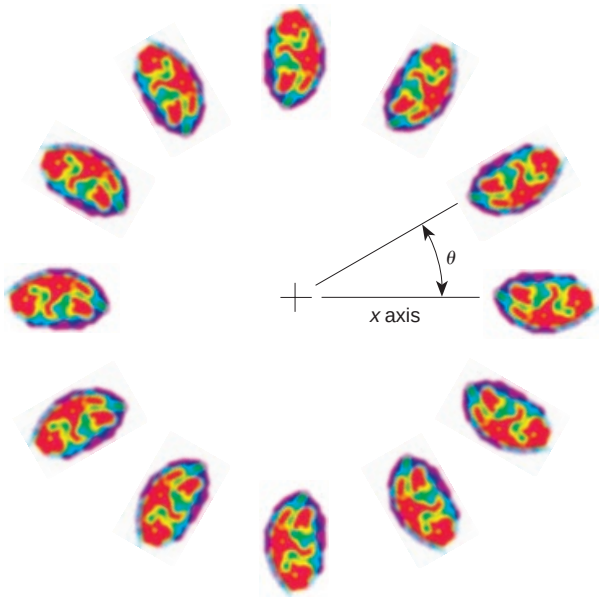

Figure 5. An ensemble of nuclei that is numerically generated by replication of a single nucleus and distributed at angular intervals of 30°. Measurements of textural features can be made using a fixed laboratory frame for all nuclei (x axis) or using a body-centered frame (based on the angle θ) for each nucleus. Statistics of pooled data for all nuclei in the ensemble vary with the use of a laboratory frame or the choice of different local (body-centered) reference frames.

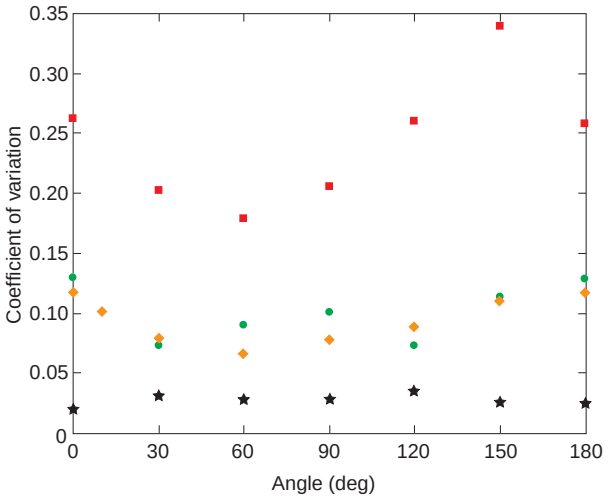

Figure 6. Values of the coefficient of variation (COV) of textural features as a function of angle for an ensemble of nonidentical nuclei: angular second moment (◆), contrast (■), overlapping correlation and difference moment (★), and inverse difference moment (●). Note that the changes in amplitude and phase with angle are similar for some features but different for others. Two factors contribute to the ensemble COV: (1) angular distribution of the nuclei in the tissue and (2) internal differences in chromatin texture.

of the 16 nuclei at a single rotational angle that constitutes a rotational group. Thus, the variance of a group of nuclei measured at a single rotational angle can be partially accounted for by variation in angular orientations when the nuclei all have the same internal chromatin distribution. Consequently, the rotational

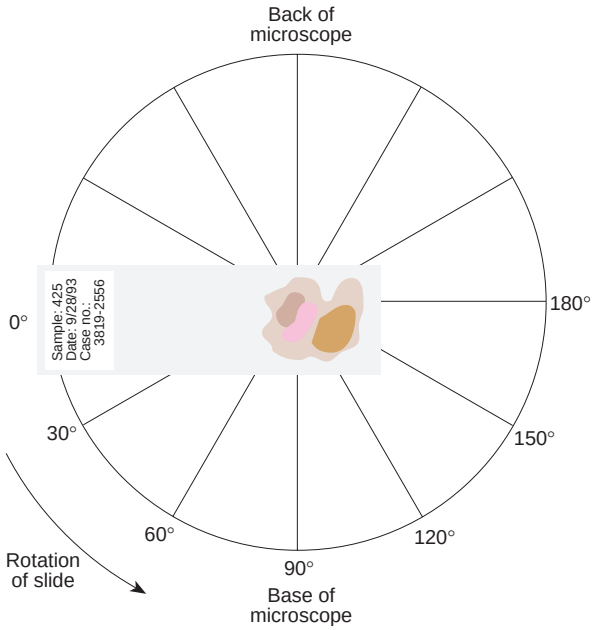

Figure 7. Diagram of orientations used in digitizing images of a Feulgen-stained tumor thin section mounted on a rotating microscope stage.

position of tissue on a slide must be controlled if results are to be compared among studies, and operators must agree on a standard method of orienting pathology slides on the microscope stage.

Body-Centered Reference Frames

We next examined the relationship between fundamental morphometric features (e.g., angular orientation of the major axis of the best-fitting ellipse to a nucleus) and measurements of chromatin texture. Using the same set of images, subgroups of nuclei were selected for a preliminary study. One subgroup (ϵ_+) consisted of nuclei with the highest eccentricities ($\epsilon_+ > 1.51$). For each of these nuclei, the co-occurrence matrix was calculated along the major axis of the best-fitting ellipse for that nucleus. A second subgroup (ϵ_0) of nuclei with low eccentricities ($\epsilon_- < 1.10$) served as a control. Because nuclei in the low-eccentricity group were approximately round, we could not establish the direction of the elliptical major axis. We therefore measured co-occurrence matrices for this group in a common reference frame, arbitrarily selected as the 0° direction in the laboratory (slide) frame. This technique corresponds to a traditional method of making measurements in DNA cytometry and allows both orientational and intranuclear variance to enter into the group statistics.[10]

The distribution of the angular second-moment features (Fig. 8) is typical of our findings for the 18 Markov textural features measured. The COV values for each group were significantly different: 8 features had no

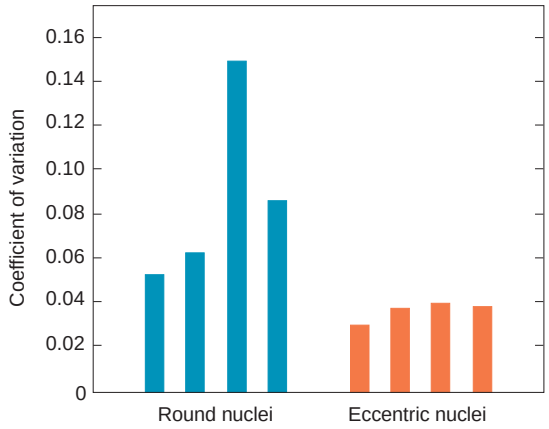

Figure 8. Variance of the angular second-moment feature in round nuclei (group ϵ_0) arbitrarily measured at 0° and eccentric nuclei (group ϵ_+), each measured in a reference frame defined by the direction of the elliptical major axis.

overlap in values between nuclei in groups ϵ_+ and ϵ_0, 4 features were partially overlapped with statistically significant differences in group means, and 6 features showed no significant difference in the means.

CONCLUSIONS

A broad concern among basic researchers, physicians, oncologists, and pathologists relates to the absence of a means of comparing the pictures afforded by the various descriptive techniques to each other. Our recent observations suggest that the measurements on the different spatial scales presented in this article could be related to each other in a hierarchical view. We are seeking to construct a picture that bridges these views based on the following observations:

1. Altered activation regimes of single genes implicated in oncogenesis have been noted to cause widespread changes in chromatin DNA texture.
2. Gene alterations resulting in oncogenesis often lead to altered chromosome copy number, accompanied by gross changes in nuclear morphology and size.
3. Prognosis models that use Gleason grading, nuclear shape factors, and Markov textural features prove more capable than models using any single measurement type alone.
4. The loss of orderly arrangement of epithelial and basal cells around luminal margins is visually prominent in higher Gleason grade tissues, which exhibit reduced numbers of effective prostatic glands.

REFERENCES

[1] Gleason, D. F., Mellinger, G. T., and the VA Cooperative Urological Research Group, "Prediction of Prognosis for Prostatic Adenocarcinoma by Combined Histological Grading and Clinical Staging," *Urology* **111**, 58–64 (1974).
[2] Epstein, J. I., *Prostate Biopsy Interpretation*, Raven Press, New York (1989).
[3] Christens-Barry, W. A., Partin, A. W., and Epstein, J. I., "Quantitative Analysis of Visual Cues in Gleason Grading in Prostate Cancer," *J. Urol.* **155**, 379a (1996).
[4] Partin, A. W., Walsh, A. C., Pitcock, R. V., Mohler, J. L., Epstein, J. I., and Coffey, D. S., "A Comparison of Nuclear Morphometry and Gleason Grade as a Predictor of Prognosis in Stage A2 Prostate Cancer: A Critical Analysis," *Urology* **142**, 1254–1258 (1989).
[5] Partin, A. W., Steinberg, G. D., Pitcock, R. V., Wu, L., Piantadosi, S., et al., "Use of Nuclear Morphometry, Gleason Histologic Scoring, Clinical Stage, and Age to Predict Disease-Free Survival Among Patients with Prostate Cancer," *Cancer* **70**, 161–168 (1992).
[6] Veltri, R. W., Partin, A. W., Epstein, J. I., Marley, G. M., Miller, M. C., et al., "Quantitative Nuclear Morphometry, Markov Texture Descriptors, and DNA Content Captured on a CAS-200 Image Analysis System, Combined with PCNA and Her-2-Neu Immunohistochemistry for Prediction of Prostate Cancer Progression," *J. Cell. Biochem.* **19**, 249–258 (1994).
[7] Haralick, R. M., Shanmugan, K., and Dinstein, I., "Textural Features for Image Classification," *IEEE Trans. Syst. Man. Cyb.* **3**, 610–621 (1973).
[8] Chen, M. H., Christens-Barry, W. A., and Partin, A. W., "Characterization of a Maximum Entropy Histogram Equalization Procedure," in *Proc. JHU/APL Symposium on Research and Development*, Baltimore, MD (Nov 1995).
[9] Rodenacker, K., "Invariance of Textural Features in Image Cytometry Under Variation of Size and Pixel Magnitude," *Analyt. Cell. Pathol.* **8**, 117–133 (1995).
[10] Irinopoulou, T., Rigaut, J. P., and Benson, M. C., "Toward Objective Prognostic Grading of Prostatic Carcinoma Using Image Analysis," *Analyt. Quant. Cytol. Histol.* **15**, 341–344 (1993).

ACKNOWLEDGMENT: The authors were supported in this work by funding provided by the National Cancer Institute under SPORE Grant 3P50-CA58236.

THE AUTHORS

WILLIAM A. CHRISTENS-BARRY is a physicist in the Physics, Modeling, and Applications Group of the Milton S. Eisenhower Research and Technology Development Center at APL. He received a B.S. in physics from the University of Delaware in 1979 and a Ph.D. in physics from the University of Alabama at Birmingham in 1987. After conducting postdoctoral studies at the University of Alabama Center for Biomolecular Spectroscopy on nucleic acid and protein crystal structures, Dr. Christens-Barry joined APL in 1988. He has studied optical properties of tissues, laser surgery techniques, dynamical and organizational properties of cancer cells, and statistical image processing applications in cancer research. His e-mail address is William.Christens-Barry@jhuapl.edu.

ALAN W. PARTIN earned his undergraduate degree from the University of Mississippi in 1983, a Ph.D. from The Johns Hopkins University in 1988, and an M.D. from the JHU School of Medicine in 1989. Dr. Partin is an Associate Professor of Urology at the Brady Urological Institute at JHMI. He has published widely on the value of prostate-specific antigen in the staging of prostate cancer and computerized analysis of nuclear morphometry in the prediction of metastatic potential. His e-mail address is apartin@welchlink.welch.jhu.edu.