Cytometry Part A 60A:81–89 (2004)

Exploratory Analysis of Quantitative Histopathology of Cervical Intraepithelial Neoplasia: Objectivity, Reproducibility, Malignancy-Associated Changes, and Human Papillomavirus

Martial Guillaud,[1] Dennis Cox,[2] Karen Adler-Storthz,[3] Anais Malpica,[4] Gregg Staerkel,[4] Jasenka Matisic,[1] Dirk Van Niekerk,[1] Neal Poulin,[1] Michele Follen,[3] and Calum MacAulay[1]*

[1]Department of Cancer Imaging, British Columbia Cancer Agency, Vancouver, British Columbia
[2]Department of Statistics, Rice University, Houston, Texas
[3]Department of Gynecology, Obstetrics and Reproductive Sciences, University of Texas Health Science Center, Houston, Texas
[4]Department of Pathology, University of Texas M.D. Anderson Cancer Center, Houston, Texas
[5]Biomedical Engineering Center, University of Texas M.D. Anderson Cancer Center, Houston, Texas

Received 31 March 2003; Revision Received 8 October 2003; Accepted 26 January 2004

Background: Background: As part of a project to evaluate emerging optical technologies for cervical neoplasia, our group is performing quantitative histopathological analyses of biopsy specimens from 1,190 patients. Objectives in the interim analysis are (a) quantitatively assessing progression of the neoplastic process of cervical intraepithelial neoplasia (CIN)/squamous intraepithelial lesions (SIL), (b) detecting malignancy-associated changes (MACs), and (c) phenotypically measuring human papillomavirus (HPV) detected by DNA testing.
Methods: The diagnostic region of interest (ROI) from immediately adjacent sections were imaged, and the basal lamina and surface of the superficial layer were delimited. Nonoverlapping quantitatively stained nuclei were selected from 1,190 samples with histopathological characteristics of normal (929), koilocytosis (130), CIN 1 (40), CIN 2 (23), and CIN 3/carcinoma in situ (CIS) (68). A fully automatic procedure located and recorded the center of every nucleus in the region of interest (ROI). We used linear discriminant analysis to assess the changes between normal and CIN 3/CIS.
Results: Scores computed from the cell-by cell features and the clinical grade of CIN/SIL were highly correlated, as were those of the architectural features and the clinical grade of CIN/SIL. We found even higher correlations between a combination of cell-by-cell and architectural scores, and clinical grade. Using these scores, we found MACs in the normal biopsy specimens from patients with high-grade CIN/SIL. Furthermore, the same scores correlated with the molecular detection of HPV.
Conclusions: Quantitative histopathology can be used in large clinical trials as an objective and reproducible measure of CIN/SIL. Detectable phenotypic changes correlate well with CIN/SIL neoplastic progression. It can also be used to infer the presence of CIN/SIL (MACs) and molecular changes associated with increased risk of cancer development (high-risk HPV). © 2004 Wiley-Liss, Inc.

Key terms: papillomavirus, human; cervical intraepithelial neoplasia; lesions, squamous intraepithelial; histopathology, quantitative; phenotype

Several prominent studies have indicated that systems that attempt to define subcategories of high- and low-grade lesions are neither reproducible nor comparable among institutions or different observers (1,2). These discrepancies are caused by the subjective nature of visual interpretation and the difficulties in expressing and teaching the set of rules and techniques that constitutes the art of clinical pathology (3). The purpose of quantitative pathology is to reduce the uncertainty of subjective visual evaluation by using quantitative optical imaging to render an objective, quantitative, and reproducible representation of pathologic assessment of cervical preneoplasia. Semi-automated analysis of histological material from the

Contract grant sponsor: National Cancer Institute; Contract grant number: 3POlCA82710–04.
*Correspondence to: Calum MacAulay, British Columbia Cancer Research Centre, 601 West Tenth Avenue, Vancouver, British Columbia, V5Z 1L3 Canada.
E-mail: cmacaula@bccancer.bc.ca
Published online 26 May 2004 in Wiley InterScience (www.interscience.wiley.com).
DOI: 10.1002/cyto.a.20034

cervix has been tested for many years (4–7). Only recently has the computer power required to perform these forms of analysis in a cost-effective or timely manner become available. Samples in which the nuclei have been quantitatively stained have shown some promise to provide quantitative information about the amount and distribution of DNA in the nuclei of normal and abnormal cells, which appears to be correlated with genetic changes (LOH, HPV infection, CIN grade, and the future development of the lesions in the organ). However all these results have only been demonstrated in limited small study sets to date. In a recent paper, Baak (8) discusses the reasons preventing the implementation of quantitative pathology in routine clinic. Reproducibility, quality control, and cost-benefit ratio in addition to additional diagnostic and prognostic information should be investigated.

Quantitative pathology, unlike automated cytology, is a semi-automatic procedure. An important effort should be made to reduce the intervention of the technologist and to make the analysis as automated as possible. Other issues are more specific to the application, such as the sampling, the type of staining, and the normalization issues. Recently, Keenan et al. (9) found that quantitative analysis of cervical biopsy specimens stained with hematoxylin and eosin (H&E) could provide discriminant power to separate the different grades of dysplasia. The question then arises regarding the need for using Feulgen-stained specimen, which requires additional cutting and staining steps. Stoichiometric stains allow assessment of the ploidy status of the specimen (10), but its biological validity in tissue sections remains a highly controversial issue (11). What is more important is that with stoichiometric staining, the chromatin organization texture features can be measured, unlike with H&E staining. These texture features have been shown to be potentially useful in different studies and for different types of tissue (12). Unlike quantitative DNA specific stains, H&E stains both cytoplasm and the nucleus, reducing greatly the contrast between the nucleus and surrounding cytoplasm, making nuclear segmentation in sectioned material very complicated.

Cervical cancer is an important cause of morbidity and mortality in women in both the developing and developed world (13). In our clinical trials, our group is performing quantitative histopathological analysis of biopsies from 1,000 asymptomatic patients undergoing screening and 800 patients who have been diagnosed. The present study reports an interim analysis of our quantitative pathology data with the following objectives: (a) to assess quantitatively the progression of the neoplastic process of CIN/SIL, (b) to detect malignancy-associated changes (MACs), and (c) to conduct a phenotypic measure of human papillomavirus (HPV) detected by molecular testing.

MATERIALS AND METHODS

The study reported here was conducted in Houston at the University of Texas M.D. Anderson Cancer Center and in Canada at the British Columbia Cancer Agency in Vancouver. Women 18 years and older were enrolled in the study and an informed consent was obtained from each. The Institutional Review Board at M.D. Anderson Cancer Center and the Internal Review Board at the British Columbia Cancer Agency, University of British Columbia, approved the protocols. Patients were asked to participate in a study of reflectance and fluorescence spectroscopy and quantitative cytohistopathology.

Entrance criteria for the screening study required that participants have a medical history free of abnormal Pap smear findings. Entrance criteria for the diagnostic study required that patients be referred with an abnormal Pap smear and usually have lesions on their colposcopic examination. Each patient has specimens collected for hybrid capture (Gaithersburg, MD) and for quantitative polymerase chain reaction (PCR) for HPV DNA and RNA for HPV 16, 18, and other high-risk types. In either study, if an abnormal region is detected during colposcopy, spectroscopy was performed and a biopsy was taken from the abnormal area, and spectroscopy was performed if a biopsy was taken from a normal area. If the colposcopy findings were normal, spectroscopy was performed and biopsy specimens from two normal areas were obtained.

Pathology

All biopsy specimens are fixed in buffered formalin and embedded in paraffin blocks. Three adjacent sections are cut at 4 μm and stained with H&E. These sections are used for histopathological interpretation. The adjacent sections are stained using the stoichiometric Thionin-Feulgen staining procedure.

The first pathology review was done by one of the gynecological pathologists on clinical duty. A second blinded review was performed by one of our study pathologists (A.M., G.S., J.M., D.V.N.). In the event of discrepancies between the first two readings, the slide was read a third time by study pathologists as a consensus review During the second review, the pathologist took a digital picture to record the region of interest (ROI).

HPV Testing

Cervical samples were tested for HPV using CR and the Hybrid Capture II method (Digene, Beltsville, MD). A nucleic acid hybridization microplate assay with signal amplification for the chemiluminescent detection of HPV DNA, the Hybrid Capture II test identifies both low-risk HPV types (6, 11, 42, 43, 44) and high-risk HPV types (16, 18, 31, 33, 35, 39, 45, 51, 52, 56, 58, 59, 68). Following the methods of Manos et al. (14), we then analyzed samples for HPV DNA using MY9 and MY 11, consensus HPV primers that amplify a 450-base pair region of the Ll open reading frame of at least 28 different HPV types. PCR products were resolved, transferred to nylon membranes (Bio-Rad, Hercules, CA), and hybridized to ^{32}P-labeled consensus HPV 16 and HPV 18 probes. The membranes, a separate one for each probe, were inserted in X-ray cassettes overnight at −80°C before autoradiographing. DNA extracted from HPV 18-positive HeLa cells, HPV 16-positive Caski cells, and a negative control without DNA were used as controls in PCR and subsequent hybridization.

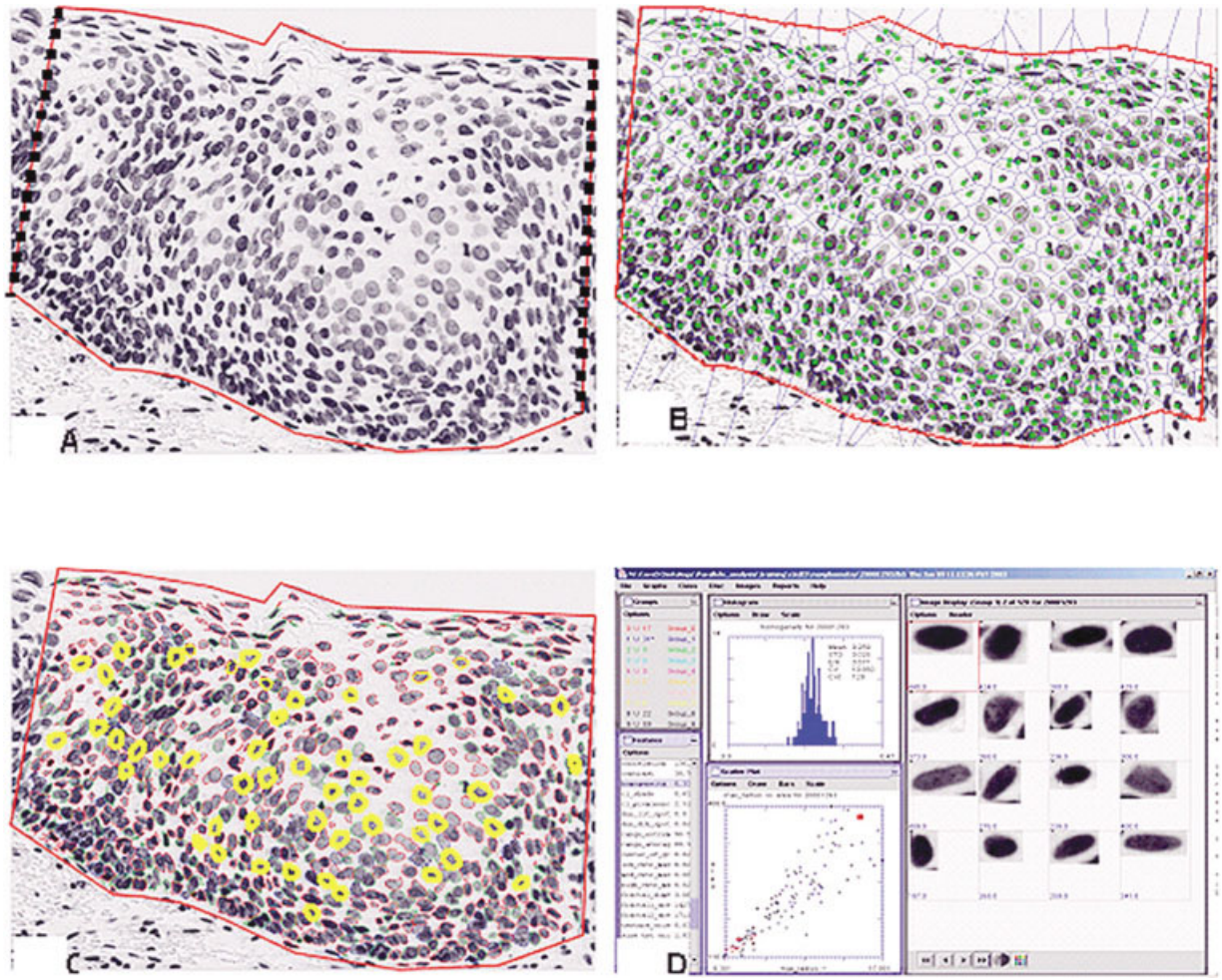

FIG. 1. Steps of the semi-automated analysis of cervical lesions. **A:** Definition of the region of interest (ROI). **B:** Automatic locations of the positions of the nuclei. **C:** Automatic segmentation of the nuclei in the intermediate layers of the epithelium. **D:** Snapshot of the interface of our software used for postanalysis of the morphometric analyses (quality control). [Color figure can be viewed in the online issue, which is available at www.interscience.wiley.com].

Samples. Two thousand quantitatively stained sections have been analyzed with the imaging system. Only samples containing squamous epithelium are included here. Of these, a final histopathological interpretation is complete on 1,190 samples, which is the subset of samples used for the studies reported here. These include the following diagnostic groups: (a) normal (929 samples); (b) koilocytotic samples (120); (c) CIN 1 (40), (d) CIN 2 (23), and (e) CIN 3/CIS (68 samples). The full set of samples will form the basis of a final report; this report is considered an interim analysis.

Image analysis. Image analysis was performed using a modified version of the Cyto-Savant automated quantitative system (Vancouver, BC, Canada). This system uses a 12-bit double-correlated sampling Microlmager 1400 digital camera (pixels 6.8 μm) placed in the primary image plane of the microscope. This software was designed for semi-automatic cellular and architectural analysis of tissue sections. Thionin-Feulgen-stained nuclei were measured with a monochromatic light at a wavelength of 600 nm (±10 nm) using a ×20 0.75-NA Plan APO objective lens.

With a printout of the diagnostic area on hand, a cytotechnologist located the same area in the Feulgen-stained slide as the ROI of the H&E slide. As illustrated in Figure 1, the cytotechnologist delineates the basal membrane and the surface of the superficial epithelium (Fig. 1A). These two surfaces define the ROI for the analysis software. Automatic detection of the nuclei (Fig. 1B) is performed for architectural analyses. A thresholding algorithm is used to separate the objects (nuclei) from the background (Fig. 1C) based on pixel intensity (15). A manual correction of the nuclear segmentation is made for touching objects. Autofocusing and edge relocation algorithms are applied to the nuclei to place the edge of the object precisely and automatically along the contour of highest local gray-level gradient (15). The digital gray-level images of these nuclei are stored in a gallery (Fig. 1D).

Table 1
Categories and Features

Category	Features
	Cytometric features
Morphometry (43)	
Size (4)	Area, mean_radius, variance_radius, maximum_radius
Shape (5)	Eccentricity, sphericity, elongation, compactness, inertia_shape
Boundaries (34)	Low_freq_fft, Freq_low_fft, Harm01-32_fft
Photometric (5)	DNA_index OD_max, OD_Var, OD_skew, OD_kurt
Discrete texture (20)	Low, medium, and high DNA amount Low, medium, and high DNA area Low, medium, high and medium_high DNA compactness Low, medium, high and medium_high DNA average distance Low, medium, and high density object Low, medium, and high centre mass Low_vs_medium, low_vs_high and low_vs_medium–high DNA
Markovian texture (7)	Entropy, energy, contrast, correlation, homogeneity, cl_shade, cl_prommence
Non-Markovian texture(5)	Density_light_spots, density_dark_spots, center_of_gravity, range_extreme, range_average,
Fractal texture (3)	Fractal_area1, fractal_area2, fractal_dimension
Run length texture (20)	Short_runs_mean, short_run_stdv, short_run_min, short_run_max long_runs_mean, long_run_stdv, long_run_min, long_run_max gray_level_mean, gray_level_stdv, gray_level_min, gray_level_max run_length_mean, run_length_stdv, run_length_min, run_length_max run_percent_mean, run_percent_stdv, run_percent_min, run_percent_max
	Architectural features
Entropy	
Features derived from Voronoi polygons (18)	Area (mean), area (standard deviation), area (skewness), area (kurtosis), area disorder, perimeter (mean), perimeter (standard deviation), perimeter (skewness), perimeter (kurtosis), roundness factor (mean), roundness factor (standard deviation), roundness factor (skewness), roundness factor (kurtosis), roundness factor heterogeneity, number of sides (mean), number of sides (standard deviation), number of sides (skewness), number of sides (kurtosis)
Features derived from Delaunay graph (4)	Nearest-neighbor distance (mean), nearest-neighbor distance (standard deviation), Delaunay nearest-neighbor distance (mean), Delaunay nearest-neighbor distance (standard deviation)
Features derived from the minimum spanning tree (MST) (7)	Percentage of nuclei with one connected nucleus, percentage of nuclei with two connected nuclei, percentage of nuclei with more than two connected nuclei, length of the MST edge (mean), length of the MST edge (standard deviation)

Quality control. The cytotechnologist manually reviews each object in an image gallery of all the selected cells (Fig. 1D) and removes any object that does not fulfill the minimum requirements (e.g., bad mask, out of focus, pale nucleus, pyknotic nucleus).

Feature calculation. Nuclear features are extracted from the digitized nuclear images of each selected cell. Table 1 identifies the features organized into different categories. One hundred three cytometric and 29 architectural features are calculated (16). Morphological features describe the nuclear size, shape, and boundary irregularities. The five photometric features estimate the absolute intensity and optical density of the nucleus and the intensity distribution characteristics. DNA amount is the raw measurement of the integrated optical density (IOD) from which all the photometric features are derived. The IOD norm is the mean value of the DNA amount of the reference population. The DNA index is the normalized measure of the integrated optical density of the object (i.e., the DNA amount divided by IOD norm).

There are 29 architectural features and 3 groups of architectural features: (a) Voronoi-based features, (b) Delaunay-based features, and (c) and features based on the minimum spanning tree (MST) (Table 1). The Voronoi-based features describe the shape and size of the Voronoi polygons. The Delaunay-based features measure the distances between neighboring cells. The MST-based features capture the spatial organization of the tree.

Normalization. Stain intensity variation is compensated by normalizing with the mean integrated optical density (IOD) of the sampled epithelial cells. For the architectural features one needs to normalize the features for size, shape of the ROI, and number of nuclei. A Monte Carlo procedure is used for this purpose. Five hundred random point distributions are generated according to a homogeneous Poisson process over the ROI with the total number of points for the random distributions fixed at the number of nuclei. All 29 features are computed for each random distribution, and the observed values of the features are normalized by subtracting the mean of the 500 randomly generated point distributions, and then dividing this by the standard deviation of these point distributions, which results in a Z-score for each feature.

Statistical Analysis

Features computed on a cell-by-cell basis were summarized by means and standard deviations to create sample-by sample level features. Linear discriminant analysis with stepwise selection was used to assess the diagnostic information on a sample-by-sample basis. The discriminant functions were constrained to use no more than three features by selection of the F to enter and the F to remove.

Table 2
Sample Sizes by Diagnostic Category

Group	Biopsy histology	Patient histology	Cytology	DNA-HPV	No. of biopsies
0	Negative	Negative	Negative	Negative	83
1	Negative	Negative	Negative	Positive	205
2	Negative	HGSIL			23
3	Negative				618
Subtotal					929
4	Koilocytosis	Koilocytosis	Unknown	Negative	19
5	Koilocytosis	Koilocytosis	Unknown	Positive	78
6	Koilocytosis				33
Subtotal					130
7	CIN 1	CIN 1		Negative	17
8	CIN 1	CIN 1		Positive	20
9	CIN 1	CIN 1		Unknown	3
Subtotal					40
10	CIN 2	CIN 2			23
Subtotal					23
11	CIN 3	CIN 3			68
Subtotal					68
Total					1,190

All training was done using only two groups: the samples of normal biopsy specimens from patients without any histologic or molecular abnormality and the CIN 3/CIS samples. Training refers to teaching the system on a learning set of slides that are classified by diagnostic category. The resulting linear discriminant scores were evaluated on the entire set of samples. Morphometric scores are linear discriminant scores of variables or features that are statistically significant, not collinear, that taken together, used as a predictor, can classify cells into separate categories. All statistical analyses were performed with the STATISTICA package produced by StatSoft (Tulsa, OK).

Our rationale for this approach to training the classifier is that the neoplastic process is a continuous process, but the gold standard of clinical pathology quantifies this continuum into ordered categories whose boundaries always involve a degree of uncertainty. By selecting the most extreme groups (samples from normal patients, i.e., patients with no concurrent dysplasia and no HPV) and CIN 3/CIS samples), we are confident of the diagnostic classes in the training sample. We then analyzed the resulting scores (including the intermediate diagnostic grades) to determine if their ordering is consistent with the clinical pathology. The features for the quantitative pathology are the same features as the quantitative cytology. The tissue can be additionally studied for architectural features.

RESULTS

Our objectives in the interim analysis were (a) quantitative assessment of the progression of the neoplastic process of CIN/SIL, (b) the detection of MACs, and (c) phenotypic measure of human papillomavirus (HPV) detected by molecular testing. Table 2 describes the sample sizes within each diagnostic category used in this analysis. The sample diagnostic categories were defined by the sample clinical histopathology, the patient histopathology (which is the worst case of the sample histopathology), and the patient1s molecular HPV test. The final categorization was a Cartesian product of each of these categorizations: the sample histopathology, the patient histopathology, and the HPV test result. Because we were interested in detecting subtle changes associated with increased cancer risk (changes that may not be reflected in conventional clinical histopathology), we refined the sample level histopathology to include patient level histopathology (for detection of MACs) and HPV test results. Further, within each sample level histopathologic grade, we considered the ordering defined by these additional

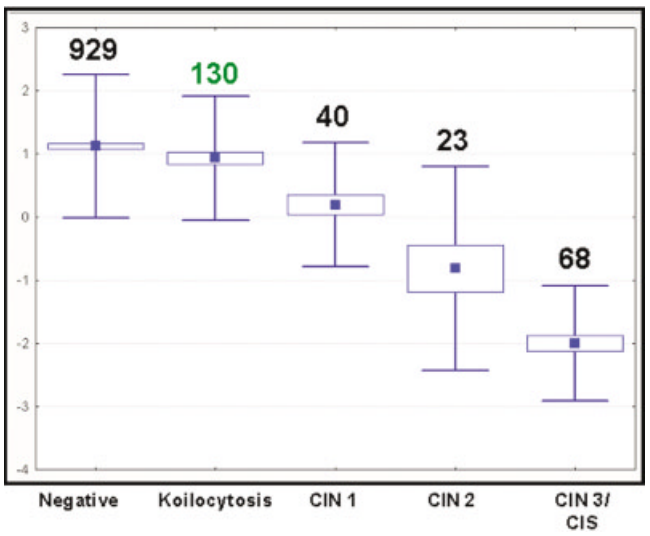

FIG. 2. Plot of morphometric score as a function of histopathology interpretation, generated from a linear discriminant analysis using three shape features: variance of the maximum radius, mean sphericity, and mean elongation. Mean (black square), standard deviation (vertical line with horizontal bars), and standard error (open rectangle). [Color figure can be viewed in the online issue, which is available at www.interscience.wiley.com].

Table 3
Accuracy of Classification Scores

Classification	Total accuracy score (%)
Morphometric	94.7
Densitometric	90.7
Discrete texture	81.5
Nucleometric	98.5
Architecture	92.5
Nucleo-architecture	100

refinements so that a normal sample from a patient with dysplasia is considered higher risk than a normal sample from a patient exhibiting no dysplasia. The most extreme groups selected for training consisted of 83 samples from normal patients (i.e., patients with no concurrent dysplasia and no HPV), and 68 CIN 3/CIS samples.

Figure 2 shows a clear trend of the morphometric score with sample histology. Except for the normal and koilocytosis diagnostic categories, there are clearly statistically significant differences between the group means and a monotonic trend. The overall classification accuracy for the two extreme groups used for training was 94.7%, indicating that there is very good separation based on these features (Table 3). Similar trends were observed in the densitometric score (Fig. 3A), which had a correct classification rate of 90.7%, and the discrete texture score (Fig. 3B), except this correct classification is somewhat lower (81.5%). The results shown in Figure 3C are based on combining the scores from Figures 2 and 3 to create an overall nucleometric score, which has an overall correct classification rate of 98.5%. This shows that combining the information from all three feature subsets provides additional discriminatory power.

From the architectural analysis, we observed that as the histopathological grade increases, the cellular density in the ROI significantly increases (data not shown). In contrast, as the histopathological grade increases, the epithelial thickness tends to decrease (Fig. 4). These observations are consistent with the increase in nucleus-to-cytoplasm ratio that has been documented in cytology

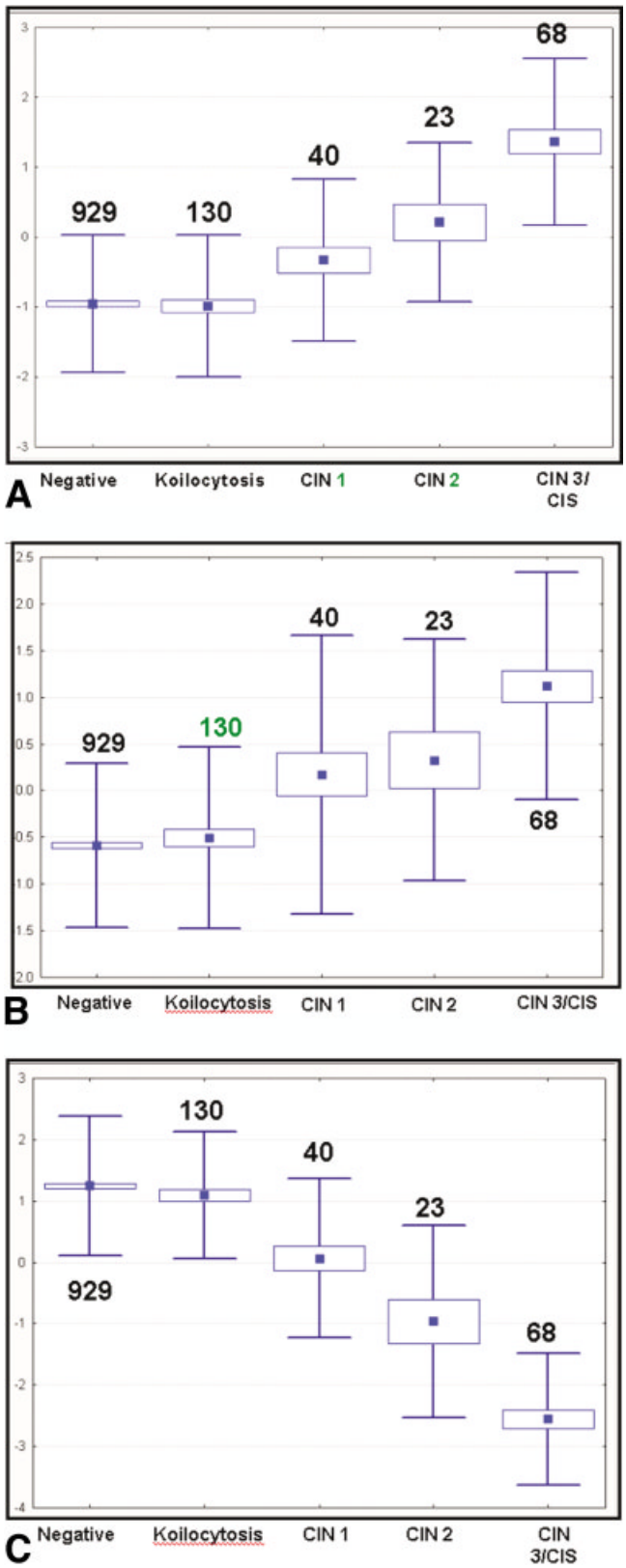

FIG. 3. **A:** Plot of densitometric score as a function of histopathology interpretation, generated from a linear discriminant analysis using three nuclear optical density features: variance across the cells in the sample of the DNA index, variance of the cells1 optical density (ODVAR), and the mean of the cells1 maximum optical density (ODMAX). Mean (black square), standard deviation (vertical line with horizontal bars), and standard error (open rectangle). **B:** Plot of discrete texture score as a function of histopathology interpretation, generated from a linear discriminant analysis using three discrete texture features, variance across the cells in the sample of the medium-high density average distance, DNA index, variance of the cell's ODVAR and the mean of the cell's ODMAX, the variance of the low-density center mass and the mean of the cell's low-density object number. Mean (black square), standard deviation (vertical line with horizontal bars), and standard error (open rectangle). **C:** Plot of the nucleometric score as a function of histopathology interpretation, generated from a linear discriminant analysis using the three scores from Figs. 2, 3A,B. Mean (black square), standard deviation (vertical line with horizontal bars), and standard error (open rectangle). [Color figure can be viewed in the online issue, which is available at www.interscience.wiley.com].

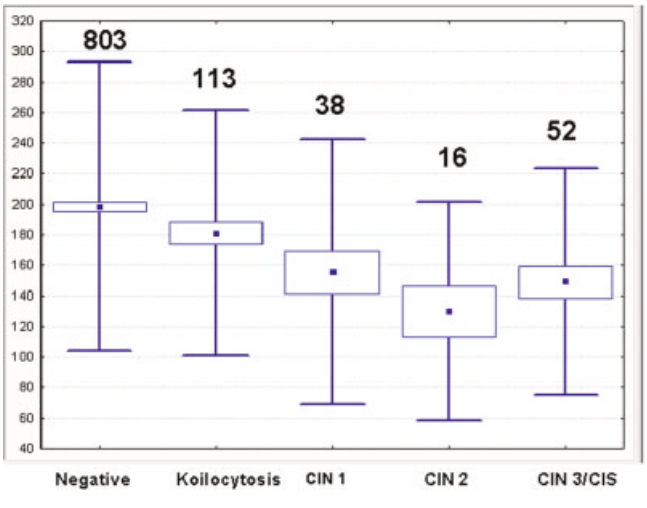

FIG. 4. Plot of epithelial layer thickness in the region of interest (ROI) as a function of histopathological interpretation. Mean (black square), standard deviation (vertical line with horizontal bars), and standard error (open rectangle). [Color figure can be viewed in the online issue, which is available at www.interscience.wiley.com].

(17). The architecture score shows clear separation in the mean for all histopathologic grades (Fig. 5), including the normal and koilocytosis diagnostic categories. The overall correct classification rate is 92.5%, which compares well with the nucleometric score, even though the architectural score makes no use of nuclear characteristics, only the locations. It should be noted that there are fewer cases in the CIN 3/CIS diagnostic category because 16 of the CIS samples had no basal membrane, which is required for the calculation of the architectural score.

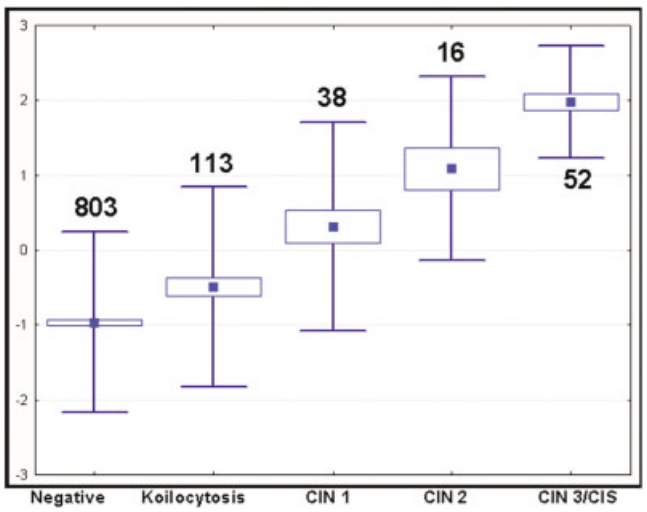

FIG. 5. Plot of architecture score as a function of histopathology interpretation, generated by linear discriminant analysis from three architectural features: minimum spanning tree (MST) mean branch length, mean of the Euclidean nearest distance, and SD Euclidean nearest distance. Mean (black square), standard deviation (vertical line with horizontal bars), and standard error (open rectangle). [Color figure can be viewed in the online issue, which is available at www.interscience.wiley.com].

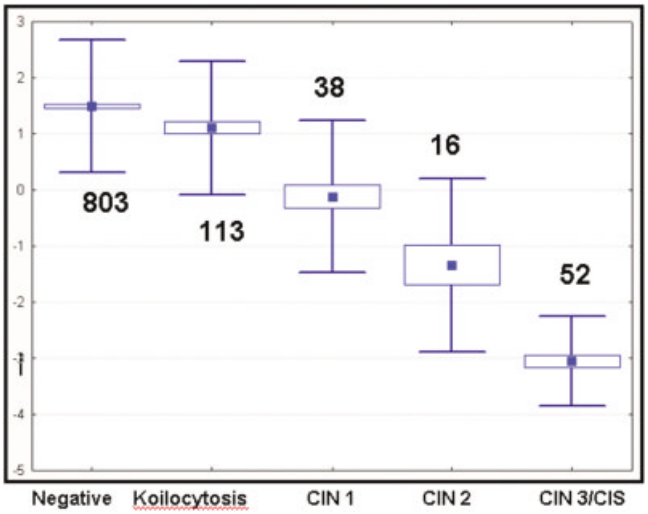

FIG. 6. Plot of the nucleo-architecture score as a function of histopathology interpretation, generated from a linear discriminant analysis using the nucleometric and architecture scores shown in Figs. 5 and 8. Mean (black square), standard deviation (vertical line with horizontal bars), and standard error (open rectangle). [Color figure can be viewed in the online issue, which is available at www.interscience.wiley.com].

The two summary features of the nucleometric and architecture scores were combined to form a linear discriminant score called the nucleo-architecture score. The trend in this score was clearly monotonic (Fig. 6), with significant differences among all the means and little overlap among the sample standard deviations. The overall correct classification rate is estimated to be 100%.

We also investigated the correlation between quantitative histopathology and other measures of cancer risk using only the nucleo-architecture score. Figure 7A shows normal samples from patients with no concurrent dysplasia or HPV infection compared with normal samples from patients with concurrent high-grade squamous intraepithelial lesions (HGSIL: CIN 2, CIN 3/CIS). There is good separation of the means ($P = 0.033$).

Figure 7B shows the comparison of nucleo-architecture scores for samples from patients with no concurrent dysplasia whose tests were negative for HPV with samples from patients with no concurrent dysplasia but whose tests were positive for HPV. A comparison of the nucleo-architectural scores from koliocytotic samples with and without molecularly determined HPV DNA showed no statistically difference in categories. The trend is in the right direction, although there is overlap in the standard errors. The *P*-value is not significant. The CIN 1 samples were subdivided into those that are positive and negative by a molecular HPV test. Similarly, no statistically significant difference was noted in the analogous results. Again, the results are consistent with the decreasing trend of this score with increasing cancer risk, but the *P*-value is not significant.

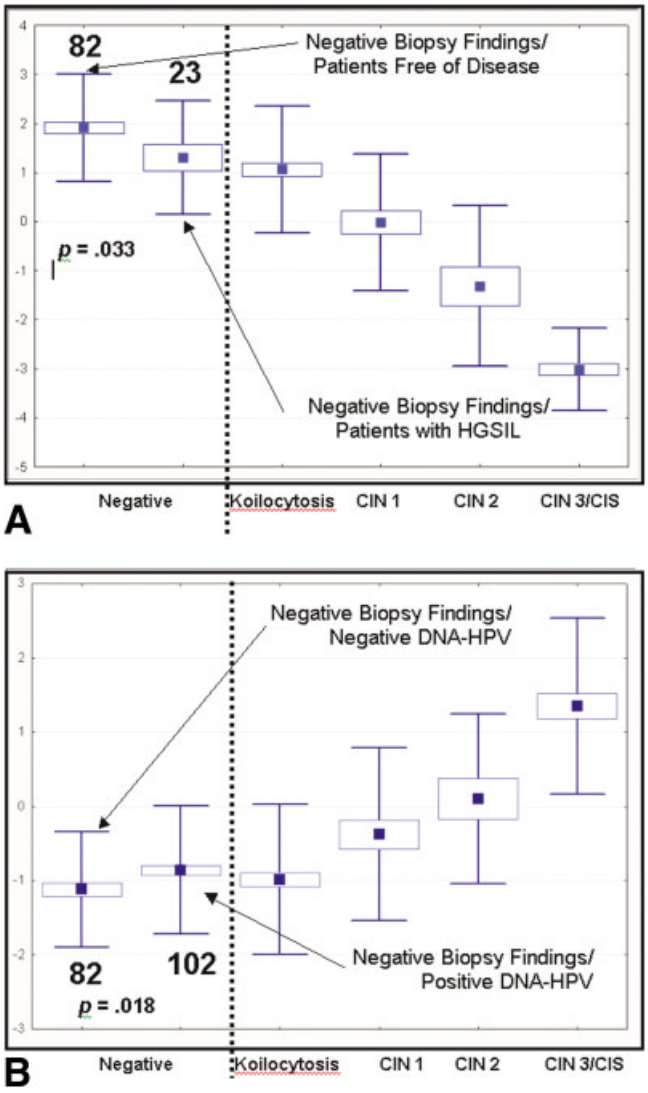

FIG. 7. **A:** Plot of the nucleo-architecture score as a function of histopathology interpretation, in which the negative group has been subdivided into the normal samples from patients without concurrent dyplasia or human papillomaviirus (HPV) infection and those normal samples from patients with high-grade squamous intraepithelial lesions (HGSIL). malignancy-associated change (MAC), mean (black square), standard deviation (vertical line with horizontal bars), and standard error (open rectangle). **B:** Plot of the densitometric score as a function of histopathology interpretation, in which patients with negative findings on biopsy have been divided into patients without concurrent dyplasia or HPV infection and those with HPV. Mean (black square), standard deviation (vertical line with horizontal bar and standard error (open rectangle). [Color figure can be viewed in the online issue, which is available at www.interscience.wiley.com].

DISCUSSION

In this interim analysis we have demonstrated that quantitative histopathology is a promising biomarker. For all the feature sets examined (morphometric, densitometric, discrete texture features, and architecture), we developed individual scores that demonstrated a high degree of correlation with clinical pathological interpretation. When these four scores were combined into a nucleo-architecture score, the results showed even better separation of the diagnostic groups.

While these findings are preliminary, we plan confirmation with a larger data set. We realize there is a possibility of overtraining in the development of each score. Nonetheless, training was carried out only with the 929 normal samples from patients with no concurrent dysplasia and no molecular evidence of HPV infection and the set of 68 CIN 3/CIS samples. Hence, all other samples from other patients may be reasonably regarded as independent of the training data, so the clear patterns that appear in Figures 2–7 with respect to intermediate diagnostic grades are reasonably valid. Additionally, the vast majority of normal samples were not used in training the discriminant scores. At the termination of the study, there will be data from 8,000 biopsies; thus, there will be a data set large enough to permit separation into independent training and validation sets so that accurate estimates of classification accuracy can be obtained.

Our preliminary results suggest that there is some discriminatory power to detect concurrent dysplasia in normal samples, that is, normal samples from patients with high-grade lesions (Fig. 7A). To our knowledge this is the first result on MACs in cervical histopathology, although MACs have been previously observed in the cervical cytology literature (19–21). These results agree with findings reported in the literature on cervical biomarkers (20) and the literature on MACs in other organ sites (22,23): lung (24–26), colon (17), stomach (27), and oral cavity (28).

Since samples from some of the same patients in the MACs group were included in the CIN 3/CIS training sample, the potential for patient-specific bias exists. When we have the final data set, which will be divided into training and test data subsets, we will be able to confirm these findings with test data independent of the training data. Furthermore, we can investigate other methods for detecting MACS using, for example, cell-by-cell classifiers or classifiers that are trained specifically to detect MACS.

We split each of the three lowest diagnostic categories (normal samples with no concurrent dysplasia, koilocytotic samples with no concurrent higher grade of dysplasia, and CIN 1 with no concurrent higher grade of dysplasia) into subgroups positive and negative for molecular evidence of high-risk HPV types. There was a consistent pattern of the scores in the HPV-positive samples indicating a higher risk of cancer. These findings need confirmation with studies of a larger sample size.

In conclusion, this preliminary study strongly indicates the power of quantitative histopathology for use in large-scale clinical trials as a biomarker of the progression of neoplasia. This study also shows the value of quantitative histopathology for evaluating the concurrent cervical dysplasia in patients (MACs), demonstrating that quantitative histopathology could be used to diagnose patients at a higher risk of progression. Finally, quantitative histopathology is a promising marker of the molecular risk of HPV. While larger sample sizes will be needed to confirm and characterize its value, studies with sufficient numbers

of patients are currently under way. Not only is quantitative pathology objective and reproducible, but also our results here concerning MACS and HPV indicate that it is able to unmask molecular markers.

ACKNOWLEDGMENTS

Special thanks go to Deanna Haskins, Anita Carraro, Tatiana Aleexenko, Louri Boiko, Nan Earle, and Trey Kell.

LITERATURE CITED

1. Ismail S, Colclough A, Dinnen J, Eakins D, Evans D, Gradwell E, et al. Reporting cervical intra-epithelial neoplasia (CIN): intra- and interpathologist variation and factors associated with disagreement. Histopathology 1990;16:371–376.
2. McCluggage W, Walsh M, Thornton C, Hamilton P, Date A, Caughley L, et al. Inter- and intra-observer variation in the histopathological reporting of cervical squamous intraepithelial lesions using a modified Bethesda grading system. Br J Obstet Gynaecol 1998;105:206–210.
3. Koss L. Diagnostic cytology and its histopathologic basis. Philadelphia: JB Lippincott, 1992.
4. Hanselaar A, Poulin N, Pahlplatz M, Garner D, MacAulay C, Matisic J, et al. DNA-Cytometry of progressive and regressive cervical intraepithelial neoplasia. Anal Cell Pathol 1998;16:11–27.
5. Poulin N, Boiko I, MacAulay C, Boone C, Nishioka K, Mittelman W, et al. Nuclear morphometry as a intermediate endpoint biomarker in chemoprevention of cervical carcinoma using alpha-difluoromethylornithine. Cytometry 1999;38:214–223.
6. Artacho-Perula E, Roldan-Villalobos R, Salas-Molina J, Vaamon de Lemos R. Histomorphometry of normal and abnormal cervical samples. Anal Quant Cytol Histol 1993;15:290–297.
7. Weyn B, Tjalma W, Van De Wouwer G, Van Daele A, Scheunders P, Jacob W, et al. Validation of nuclear texture, density, morphometry and tissue syntactic structure analysis as prognosticators of cervical carcinoma. Anal Quant Cytol Histol 2000;22:373–382.
8. Baak J. The framework of pathology: good laboratory practice by quantitative and molecular methods. J Pathol 2002;198:277–283.
9. Keenan S, Diamond J, McCluggage W, Bharucha H, Thompson D, Bartels P, et al. An automated machine vision system for the histological grading of cervical intraepithelial neoplasia (CIN). J Pathol 2000;192:351–362.
10. Steinbeck RG. Proliferation and DNA aneuploidy in mild dysplasia imply early steps of cervical carcinogenesis. Acta Oncol 1997;36:3–12.
11. Mairinger T, Gschwendtner A. Comparison of different mathematical algorithms to correct DNA-histograms obtained by measurements on thin liver tissue sections. Anal Cell Pathol 1996;11:159–171.
12. Palcic B. Nuclear texture: can it be used as a surrogate endpoint biomarker? J Cell Biochem 1994;19(suppl):40–46.
13. Parkin DM, Pisani P, Ferlay J. Estimates of the worldwide incidence of 25 major cancers in 1990. Int J Cancer 1999;80:827–841.
14. Manos M, Ting Y, Wright D, Lewis A, Broker T, Wolinsky S. Use of polymerase chain reaction amplification for detection of genital human papillomaviruses. Cancer Cells 1989;7:209–213.
15. MacAulay C, Palcic B. An edge relocation segmentation algorithm. Anal Quant Cytol Histol 1990;12:165–171.
16. Doudkine A, Macaulay C, Poulin N, Palcic B. Nuclear texture measurements in image cytometry. Pathologica 1995;87:286–299.
17. Montag AG, Bartels PH, Dytch HE, Lerma-Puertas E, Michelassi F, Bibbo M. Karyometric features in nuclei near colonic adenocarcinoma. Statistical analysis. Anal Quant Cytol Histol 1991;13:159–167.
18. Koutsky LA, Ault KA, Wheeler CM, Brown DR, Barr E, Alvarez FB, et al. A controlled trial of a human papillomavirus type 16 vaccine. N Engl J Med 2002;347:1645–1651.
19. Wied GL, Bartels PH, Bibbo M, Sychra JJ. Cytomorphometric markers for uterine cancer in intermediate cells. Anal Quant Cytol 1980;2:257–263.
20. Zahniser DJ, Wong KL, Brenner JF, Ball HG, Garcia GL, Hutchinson ML. Contextual analysis and intermediate cell markers enhance high-resolution cell image analysis for automated cervical smear diagnosis. Cytometry 1991;12:10–14.
21. Palcic B, MacAulay C, Shlien S, Treurniet W, Tezcan H, Anderson G. Comparison of three different methods for automated classification of cervical cells. Anal Cell Pathol 1992;4:429–441.
22. Nierbergs H, Zak R, Allen D, Reisman H, Clandy T. Systemic cellular changes in material from human and animal tissues in the presence of tumours. In: Transactions of the Seventh Annual Meeting of the International Society of Cytology Council. Northbrook, IL: International Society of Cytology, 1959. p 137.
23. Nierbergs H. Recent progress in the interpretation of malignancy-associated changes (MAC). Acta Cytol 1968;12:445–453.
24. MacAulay C, Lam S, Payne PW, LeRiche JC, Palcic B. Malignancy-associated changes in bronchial epithelial cells in biopsy specimens. Anal Quant Cytol Histol 1995;17:55–61.
25. Payne PW, Sebo TJ, Doudkine A, Garner D, MacAulay C, Lam S, et al. Sputum screening by quantitative microscopy: a reexamination of a portion of the National Cancer Institute Cooperative Early Lung Cancer Study. Mayo Clin Proc 1997;72:697–704.
26. Ikeda N, MacAulay C, Lam S, LeRiche J, Payne P, Garner D, et al. Malignancy associated changes in bronchial epithelial cells and clinical application as a biomarker. Lung Cancer 1998;19:161–166.
27. Doudkine A, MacAulay C, Garner D, Palcic B. Early lung cancer detection using image cytometry: where are the challenges? In: Eighth Congress of the European Society for Analytic Cellular Pathology; 2002 September 14–21; Crete, Greece. Amsterdam: ESACP; 2002. p 34.
28. Klawe H, Rowinski J. Malignancy associated changes (MAC) in cells of buccal smears detected by means of objective image analysis. Acta Cytol 1974;18:30–33.