Journal of Cellular Biochemistry, Supplement 23:43-54 (1995)

Image Cytometry and Chemoprevention in Cervical Cancer

Branko Palcic, PhD, David M. Garner, PhD and Calum E. MacAulay, PhD

Cancer Imaging Department, Medical Physics Division, British Columbia Cancer Agency, Vancouver, British Columbia V5Z 1L3 Canada

Abstract Of the approximately 60 million Pap smears performed in the United States in 1995, about 8% or 5 million will show cytology that is "not negative" (ASCUS, AGCUS, LSIL, HSIL, *etc.*). Possibly 15% or about 0.7 million of these cases will have positive follow-up by repeated Pap smears, colposcopy or biopsy. More than 4 million will be false-positive smears based on the reference standard of biopsy or repeated smears. If no treatment or medical intervention was offered to the 0.7 million cytologically and histologically positive cases, perhaps 20,000 (3%) would develop into invasive cancer. Of the original 5 million cytologically "not negative" cases, fewer than 0.5% have the potential to develop into invasive cancer.

While considerable attention has been paid to false-negatives in Pap screening, the above considerations indicate that the cytological and histological criteria for assessing the malignant potential of "not negative" samples might benefit from some refinement. Until such refinement occurs, any chemoprevention studies in cervix face a formidable signal-to-noise problem—worse than 1:30.

This paper presents data from quantitative image cytometry of cervical smears for assessing the malignant potential of various "not negative" cases. We have approached this in two ways—by analyzing dysplastic cell nuclei and by analyzing the nuclei of cytologically normal cells growing in the vicinity of the neoplastic lesion. In both cases, nuclear features describing the distribution of the DNA in the cell nuclei (especially texture features) are the discriminating factors. Future research into the objective assessment of malignant potential of "not negative" cases is outlined.

Key words: Image cytometry, malignant potential, progression, regression

Although mass Pap screening programs have demonstrably reduced mortality from cervical cancer [1,2], there is considerable evidence that relatively few abnormal cases detected by screening would ever become invasive cervical cancer had they been left untreated. If the mapping between these abnormal cases and development of invasive cervical cancer is poor, then any chemoprevention trials aimed at reversing cervical precancers are especially difficult to design. A partial solution to this problem may come from automated image cytometry. Preliminary data support the hypothesis that the malignant potential of "abnormal" cervical specimens may be more objectively and accurately assessed based on quantitative measurements [3].

In this paper we will describe our efforts to try and gauge the malignant potential of current "abnormal" cases, estimate the required study and control group sizes necessary to measure the efficacy of a chemoprevention protocol, illustrate how quantitative image cytometry may be applied to the problem of assessing malignant potential, and describe some retrospective experiments we are planning to test the power of image cytometry for assessing malignant potential.

Address correspondence to Branko Palcic, PhD, Cancer Imaging Department, Medical Physics Division, British Columbia Cancer Research Center, 601 West Tenth Avenue, Vancouver, British Columbia V5Z 1L3 Canada.

THE MALIGNANT POTENTIAL OF "NOT NEGATIVE" CASES

It is difficult to estimate the malignant potential of "not negative" cervical cases. Typically, between 5–10% of Pap smears are initially screened as "not negative" [4,5]. The ASCUS (atypical squamous cells of undetermined significance) category alone "may be expected in no more than 5% of PAP findings" according to Kurman *et al.* [6] and supported by Kreiger *et al.* [5]. Thus, with 50–60 million Pap smears in the United States in 1995, about 5 ± 1 million will be called "not negative."

Most of those cases initially screened as "not negative" will be followed up by some combination of repeat smears, colposcopy, and biopsy. Some studies place the number of positively followed-up ASCUS cases at 15% [7], but others place it as high as 60% [8]. The number of positively followed-up LSIL (low-grade squamous intraepithelial lesions) has been measured at 50% [9]. If we *very* conservatively put the total positive follow-up at only 15%, then at least 0.7 million "not negative" cases will be treated. This estimate appears consistent with that of Morrow [10] in these proceedings.

How many of these 0.7 million or more cases have the potential to develop into invasive cervical cancer? Koss [4] put the number at 1 in 10 or less. One difficulty in making this estimate is that few, if any, baseline measurements were done before screening programs were implemented [11]. However, Pap screening has been credited with reducing the incidence of invasive cervical cancer by 50–80% [1,2,12]. In 1995, there will be fewer than 15,000 cases of invasive cervical cancer [13] with screening in place. Pap screening in the U.S. probably covers about 60–70% of the target "at-risk" population [14]. Of the 0.7 million cases of dysplasia, 25–50,000 invasive cervical cancers are estimated to result in 1995 if there was no Pap screening. This rate is between 1 in 15 and 1 in 30.

Another approach to making this estimate is to use the reported incidence rates (age standardized) from locations which have no screening program and apply these to the U.S. population of 100 million women at risk. The reported age-standardized rate of invasive cervical cancer in British Columbia in 1955 when screening started was 28.5 per 100,000 for women aged 20 and over [2]. Similarly, a Swedish study showed a rate of 25 invasive cancers per 100,000 unscreened women [1]. Although these rates include a prevalence component (rather than strictly reflecting incidence), applying them to the U.S. at-risk population yields an estimate of about 20,000 invasive cervical cancers in 1995 from the 0.7 million cases of dysplasia if treatment was withheld, a rate of about 1 in 35.

Another alternative focuses on carcinoma *in situ* as the precursor to invasive cancer. Dutch data [15] indicates that for women aged 65, virtually all CIS progresses, whereas for women at age 30, more than 75% *regress* to negative. Because of this, van Ballegooijen *et al.* [15] recommend that screening not start before the age of 30 or 35. The study of Stenkvist *et al.* [1] also indicated a progression rate from CIS at about 25% for women aged 30–39. Similarly, in British Columbia between 1955 and 1985, about 26,000 cases of CIS were detected and treated [2]. Yet, based on the incidence rate of invasive cervical cancer in 1955 and adjusting for the population changes to 1985, there would have been 7,000 invasive cervical cancers in British Columbia, which is also consistent with a progression rate of 25% for CIS. Thus, even considering only the CIS group of patients, it is unlikely that the malignant potential is much higher than 1 in 3 or 1 in 4.

CHEMOPREVENTION TRIAL DESIGN FOR LOW DISEASE PROGRESSION PROBABILITY

The low malignant potential of cervical dysplasia has a serious impact on the design of chemoprevention trials, making very large study groups necessary.

Sample sizes required to detect an effect of the chemoprevention agent can be determined using standard statistical methods, for example, Fleiss [16]. For this problem, it is appropriate to use the two-tailed test for equal sample sizes (placebo and test groups). The calculation requires specification of the Type I error, α, the test significance which is the probability of declaring that the chemoprevention is effective when in fact it is not; and Type II errors, β, (the test power is $1-\beta$) which is the probability of declaring that the chemoprevention makes no difference, when in fact it does (it could be significantly better or

TABLE I. Number of Subjects Required for Each Group (Control and Test) for Various Progression to Regression Ratios and Chemoprevention Efficacies*

Regression: Progression Ratio	Number of Subjects					
	Chemoprevention Efficacy (% responding)					
	100	90	75	50	25	10
100:1	6	9	14	28	72	220
10:1	9	12	18	38	118	526
1:1	27	36	56	138	574	3,590
1:2	47	62	98	249	1,080	6,995
1:5	106	142	225	582	2,601	17,209
1:10	206	274	437	1,136	5,135	34,233
1:30	602	802	1,282	3,354	15,271	102,327

* Significance, $\alpha = 0.01$; power, $1-\beta = 0.95$

worse than the placebo). It is usual for a drug trial to set α low to 0.01 or 0.02; and to set β to about 4α, about 0.05 or 0.10 [16].

Table I and Figure 1 show the number of cases required for each group (*i.e.*, the placebo group and the study group) for a particular drug efficacy and ratio of progression to regression. Here, "drug efficacy" refers to the percentage of test subjects detectably responding. Table I and Figure 1 are calculated for a significance, α, of 0.01 and a power, $1-\beta$, of 0.95. Even at a progression to regression ratio of 1:5 (similar to CIS), typically 500 subjects per group are required to detect that a drug is effective for 75% of patients.

Table I and Figure 1 indicate that any successful chemoprevention trial requires good estimates of progression to regression rates and a large number of subjects in both control and test groups. Any method that increases the progression to regression ratio will significantly reduce the sample sizes required.

HOW QUANTITATIVE IMAGE CYTOMETRY CAN HELP

Many of the potential techniques of image cytometry, flow cytometry, and other cyto- and histometric methods have been described by other authors at these proceedings. These include DNA ploidy and texture measurements, immunohisto- and cytochemical markers of specific receptors or specific phenotype expressions, and various genetic and genotypical markers. All of these hold exciting potential as surrogate endpoint biomarkers.

We take the view that any chemoprevention trials that could be practically and ethically implemented in the United States for cervical cancer will *not* be applied to patients with carcinoma *in situ*, or even HSIL, but will, instead, be restricted to ASCUS, LSIL, and other very low-grade states. If this assumption is correct, the preceding statistical analysis clearly indicates the need not only for an endpoint marker, but also for a "startpoint" marker, and appropriate monitoring metrology. If the number of cases that must be followed is indeed large, then it requires a metrology method that is effective and economical when applied to large numbers of subjects. For this reason, we confine our comments to automated image cytometry.

Our hypothesis [3] is that it may be possible to assess the malignant potential of Pap smears and other types of single cell preparations by image cytometry if it is:

Automated. A statistically significant sample of perhaps several thousand cell nuclei per sam-

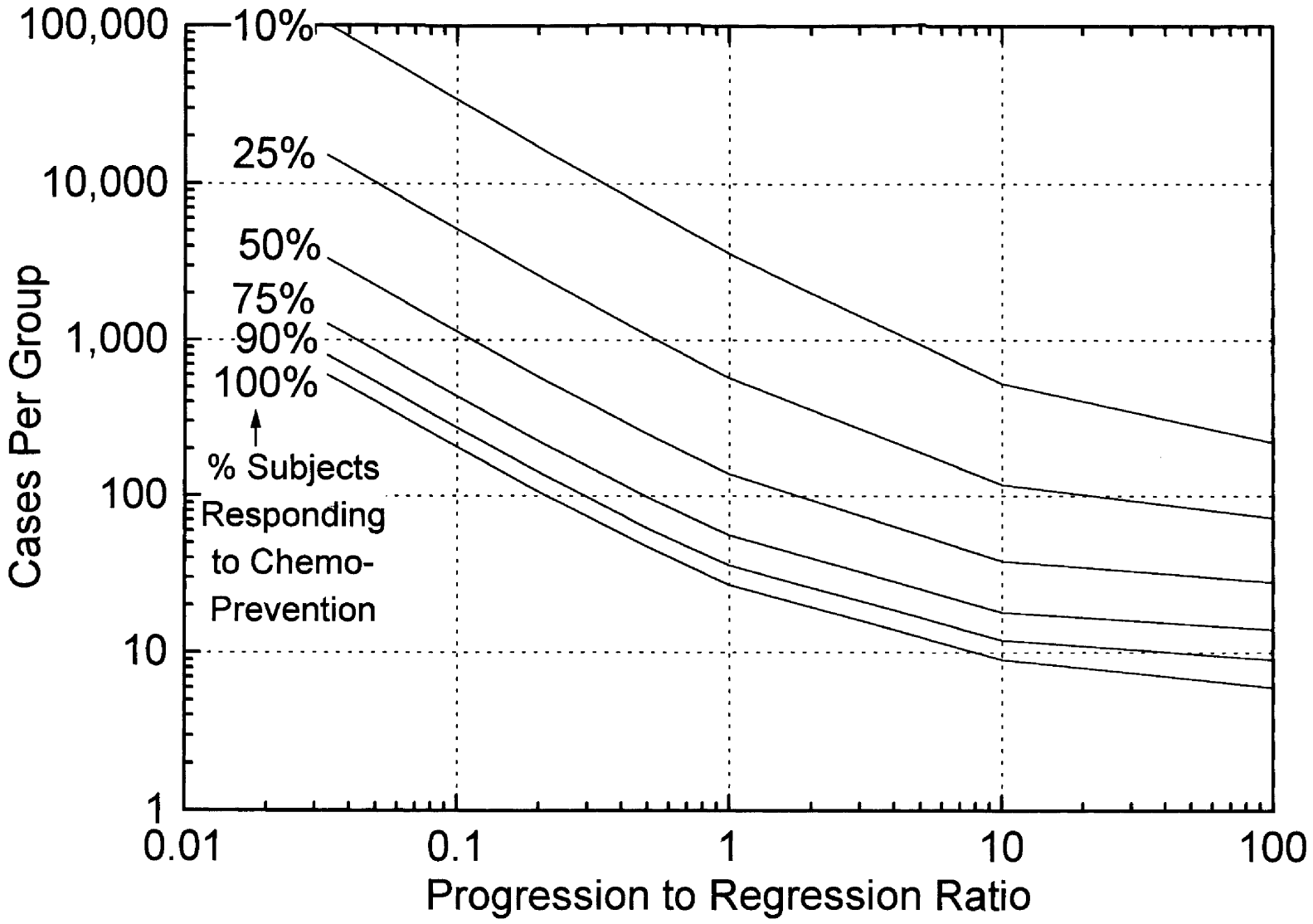


Fig. 1. The number of cases required in *each* group (control and test) for various chemoprevention efficacies and progression to regression ratios for cervical dysplasia. The test significance, $\alpha = 0.01$, and test power, $1-\beta = 0.95$.

ple must be measured [3]. This is absolutely impractical by manual or interactive means. Furthermore, high resolution imaging (see below) can *only* be performed consistently in terms of focus, segmentation, and unbiased cell selection if performed by a machine.

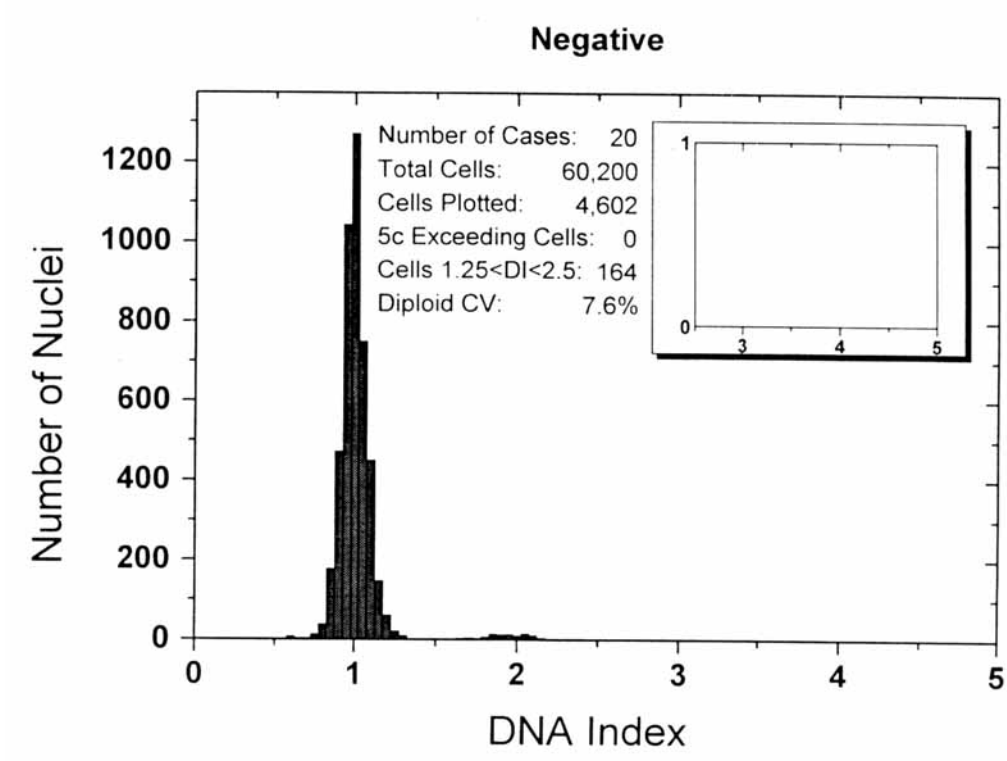
Quantitative for DNA. Hematoxylin-stained specimens have been very valuable throughout the twentieth century, but they have not been highly useful for the assessment of malignant potential. *Hematoxylin is not a specific stain for DNA* [17]. It seems likely that we have, over the past century, learned all that hematoxylin has to teach us. New techniques are required.

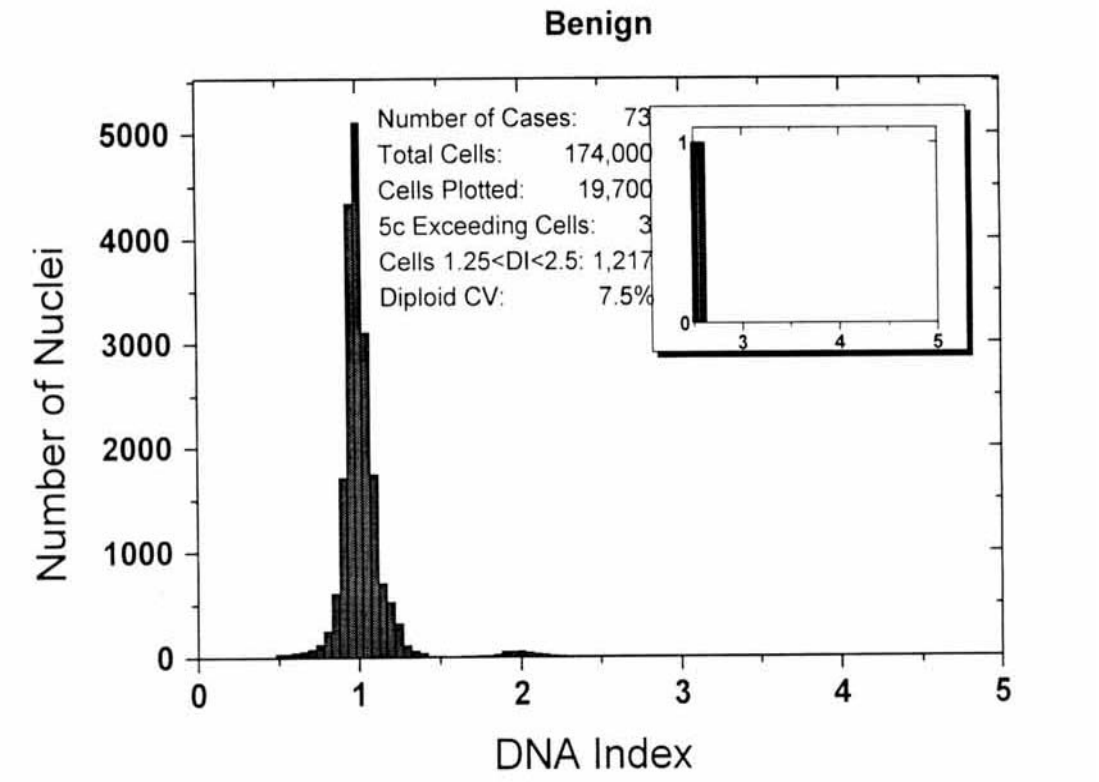
High resolution. Both high spatial and photometric resolutions are required to measure the subtle changes in nuclear texture of normal cells found growing adjacent to a neoplastic lesion [3]. We have data being prepared for publication that support the hypothesis that such measurements are of both diagnostic and prognostic significance.

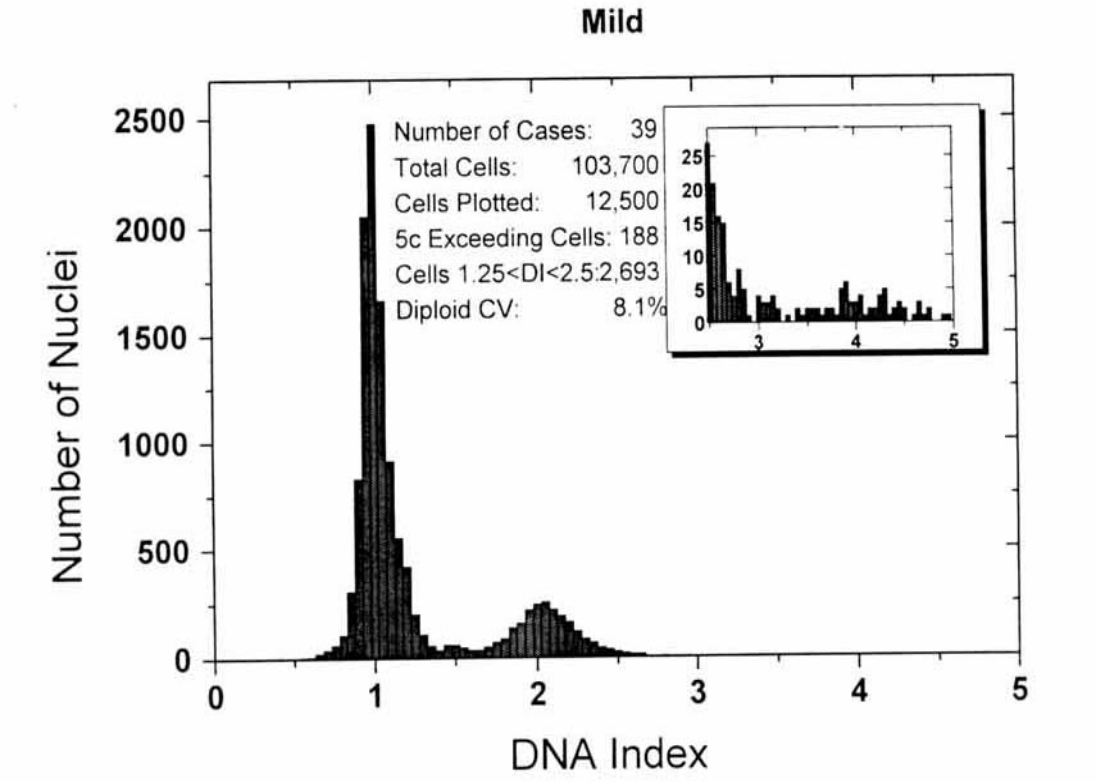
Preliminary data in both lung and cervix [3] indicate that quantitative image cytometry can give some assessment (typically 80% accurate) of the malignant potential of a lesion based on the analysis of a single sample. However, in the context of cervical cancer, which appears to be a very slowly developing disease [11], it seems far more interesting to assess malignant potential by performing objective measurements along a time base. This is also a requirement for monitoring any chemoprevention trial.

The first step in monitoring disease progression and regression along a time base is to establish an objective metric. Figure 2 illustrates the progression in DNA ploidy from negative to invasive cancer (these are composite plots, summing many slides and hundreds of thousands of cell nuclei). Figure 3 illustrates the changes in various nuclear texture features also along this progression, both for "normal" intermediate cells and for malignant cells.

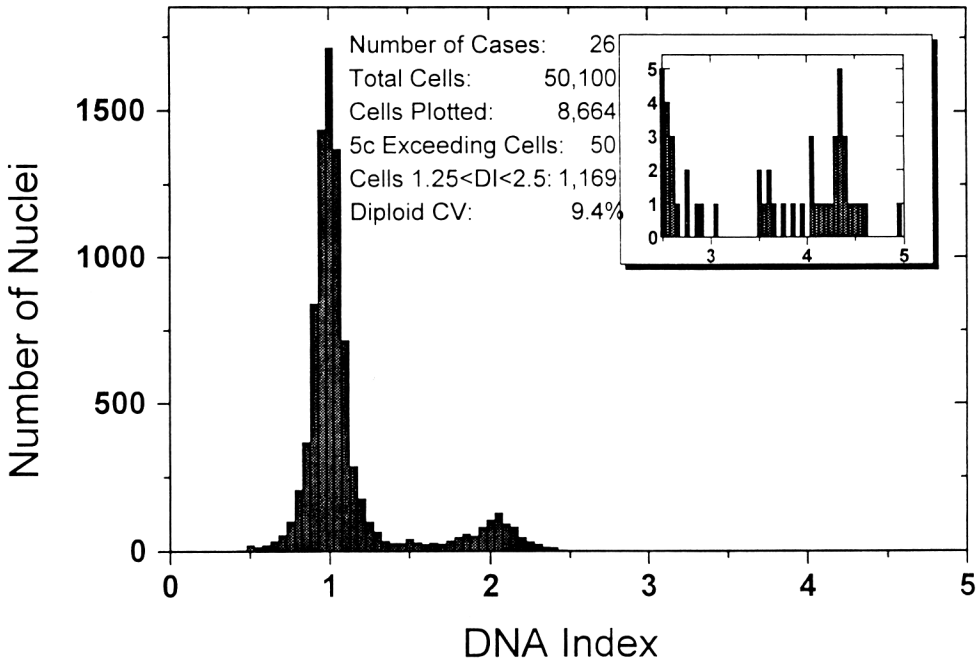
We are in the process of extending these measurements from 290 slides to 1,200–1,500 slides in order to build the foundation of an objective case classification system that can be used for time-based measurements of the malignant potential of Pap smears. The first of these measurements will be retrospective studies described below.



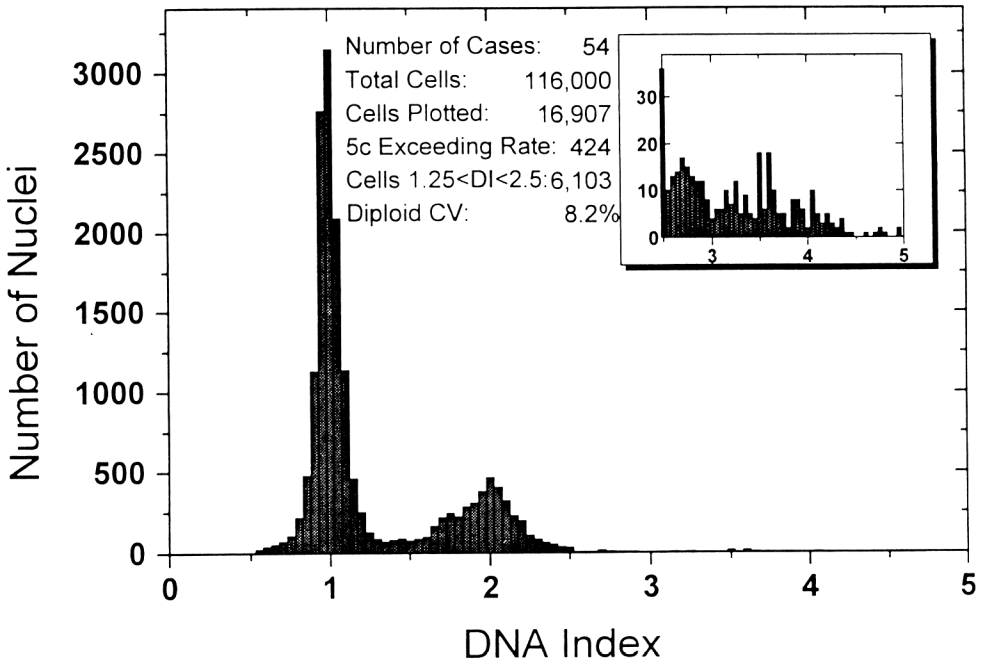

Benign
Number of Cases: 73
Total Cells: 174,000
Cells Plotted: 19,700
5c Exceeding Cells: 3
Cells 1.25<DI<2.5: 1,217
Diploid CV: 7.5%
Number of Nuclei
DNA Index
0
1000
2000
3000
4000
5000
0
1
2
3
4
5


Mild
Number of Cases: 39
Total Cells: 103,700
Cells Plotted: 12,500
5c Exceeding Cells: 188
Cells 1.25<DI<2.5:2,693
Diploid CV: 8.1%
Number of Nuclei
DNA Index
0
500
1000
1500
2000
2500
0
1
2
3
4
5

Moderate



Severe



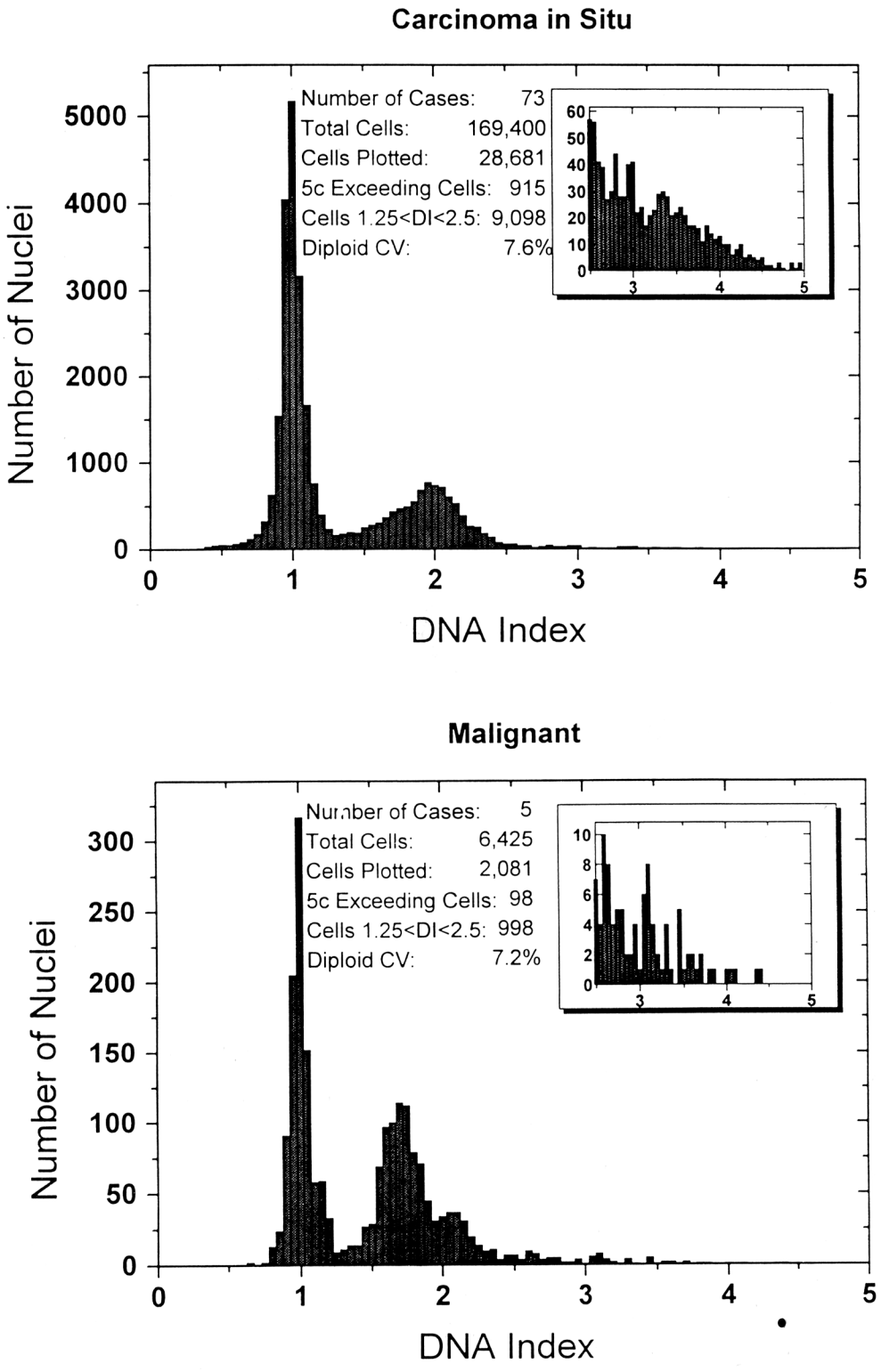
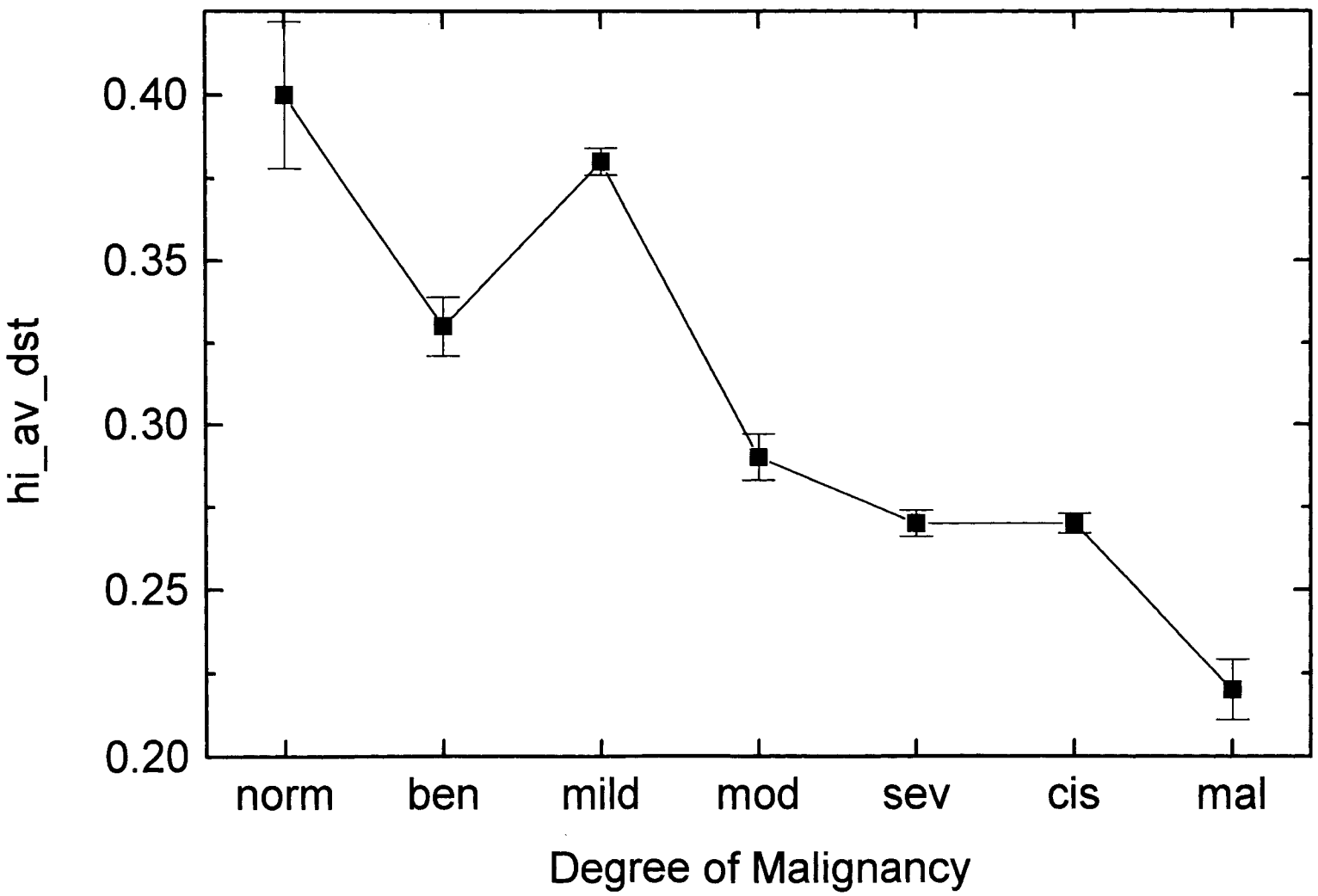


Fig. 2. Composite ploidy plots for cervical samples of various degrees of neoplasia. Not all diploid cells (DNA index < 1.2) are plotted, but all non-diploid cells are plotted. The insert box shows the "5-c exceeding cells" (DNA index > 2.5) on a separate vertical scale for clarity.


hi_av_dst
0.40
0.35
0.30
0.25
0.20
norm
ben
mild
mod
sev
cis
mal
Degree of Malignancy

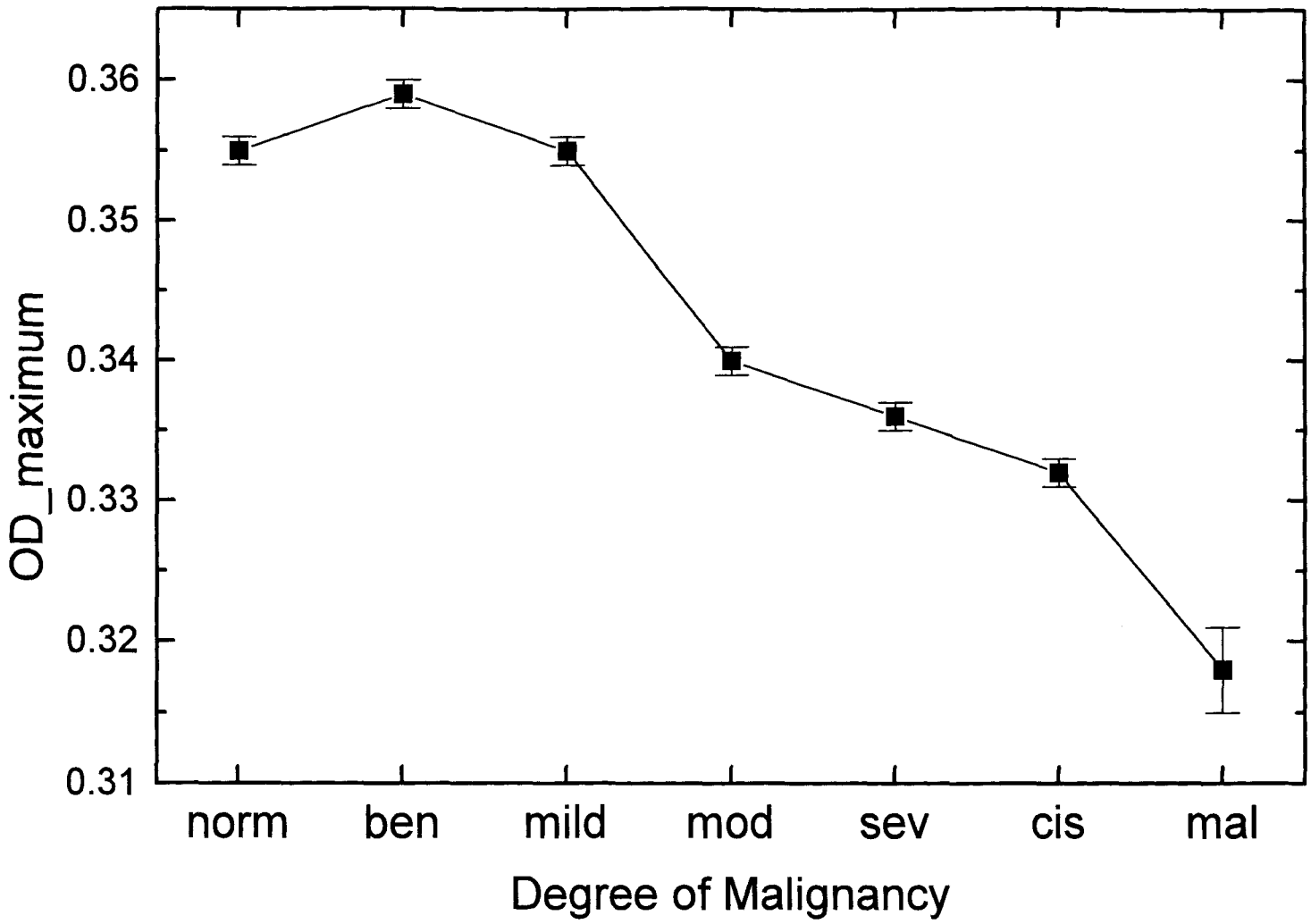


Fig. 3. Examples of the variation in some texture features as a function of malignancy grade for (a) "abnormal" cells (DNA index > 1.3) and (b) "normal" cells (DNA index < 1.2). These texture features are measured by computer assisted image analysis and *OD maximum* is a measure of the darkness of the darkest spot in the nucleus, normalized to eliminated slide to slide staining variations, and *hi_av_dst* a measure of the radial distribution of the dark areas: low values are close to the nucleus center, centroid high values are close to the nucleus edge.

PROPOSED RETROSPECTIVE STUDIES ASSESSING MALIGNANT POTENTIAL OF PAP SMEARS BY IMAGE CYTOMETRY

A crucial aspect of our DNA quantitative staining protocol [17] is that it may be applied to historical Pap-stained samples. Historic samples have their coverslips removed, have the Papanicolaou stain removed (usually very well), and are re-stained for DNA quantitation (the data illustrated in Figures 3a and 3b were measured on such slides). After image cytometer scanning, Papanicolaou stain can be reapplied to the specimens, returning them to a good approximation of their initial Pap-stained state. There are caveats to this, of course. DNA is slowly lost over time in specimens stored at room temperature. Twenty year-old samples should not be compared to 10 year-old samples; however, there is little problem in comparing 20 year-old samples to other 20 year-old samples.

The British Columbia Cancer Agency (BCCA) operates the centralized Pap screening service for this Canadian province of 3.5 million persons, and has done so since 1949. Computerized patient records span back to 1954, and all samples are retained in an archive for at least seven years ("positive" slides are kept "forever"). This Pap screening program presently handles just under 700,000 slides per year and has an active archive of more than four million slides [G.H. Anderson, personal communication].

Two time-based studies are proposed. The first study would involve the most scientifically interesting test of progression and regression, using invasive cervical cancer as the endpoint by examining historical slides from patients who developed cervical cancer. However, there may be too few cases to make a good study, since in British Columbia most women who develop invasive cervical cancer have never been previously Pap screened [18]. Also this test should have a set of control slides from patients who did not develop invasive cancer (age-matched both in patient age and slide age). In general, these patients will have been treated, except for cases of very low-grade dysplasia, which may confound the study. This study may thus be impossible, due to lack of sufficient historical control specimens.

The second study, more practical but scientifically less satisfying, would look only at very early dysplasia, effectively using the transition from LSIL to HSIL as the endpoint. Because most HSIL will also not develop into invasive cancer, this study does not address the complete issue of malignant potential, but it should address the earliest stages of the process. Until last year in British Columbia, colposcopy was not generally recommended until the patient had two "moderate" dysplasias (these are HSIL under Bethesda) in a row. In the past six years, there have been typically 30–50,000 mild and moderate dysplasia cases per year, so finding cases that progressed until medical treatment intervened and control cases that spontaneously regressed to negative should not be difficult [G.H. Anderson, personal communication].

In both studies, the question is: "Can image cytometry provide objective measurements that distinguish progressors from regressors?" The measurements can be made across several slides per patient spanning a time base. We believe the results of this kind of study could have beneficial effects on the design and implementation of any chemoprevention study for cervical cancer.

ACKNOWLEDGMENTS

We gratefully acknowledge the support of Drs. Jean LeRiche, George Anderson, and Jasenka Matisic of the BCCA Departments of Pathology and Laboratory Medicine. The support of the Science Council of British Columbia Health Development Fund and Xillix Technologies Corp., Richmond, British Columbia, are also gratefully acknowledged.

REFERENCES

1. Stenkvist B, Bergström R, Eklund G, Fox CH: Papanicolaou smear screening and cervical cancer. What can you expect? JAMA 252:1423–1426, 1984.
2. Anderson GH, Boyes DA, Benedet JL, LeRiche JC, Matisic JP, Suen KC, Worth AJ, Millner A, Bennett OM: Organization and results of the cervical cytology screening programme in British Columbia, 1955–1985. Br Med J 296:975–978, 1988.
3. Palcic B, MacAulay C: Malignancy associated changes: Can they be employed clinically? In Weid GL, Bartels PH, Rosenthal DL, Schenck U (eds): "Compendium on the Computerized Cytology and Histology Laboratory." Chicago: Tutorials on Cytology, 1994, pp 157–165.
4. Koss LG: The Papanicolaou test for cervical cancer detection. A triumph and a tragedy. JAMA 261:737–

743, 1989.
5. Kreiger P, Naryshkin S: Random rescreening of cytologic smears: A practical and effective component of quality assurance programs in both large and small laboratories. Acta Cytol 38:291–298, 1994.
6. Kurman RJ, Henson DE, Herbst AL, Noller KL, Schiffman MH: Interim guidelines for management of abnormal cervical cytology. JAMA 271:1866–1869, 1994.
7. Wilbur DC, Bonfiglio TA: Atypical squamous cells in cervical smears—resolving a controversy. Diagn Cytopathol 9:427–429, 1993.
8. Sidawy MK, Tabbara SO: Reactive change and atypical squamous cells of undetermined significance in Papanicolaou smears: A cytohistologic correlation. Diagn Cytopathol 9:423–427, 1993.
9. Hall S, Wu TC, Soudi N, Sherman ME: Low grade squamous intraepithelial lesions: Cytologic predictors of biopsy confirmation. Diagn Cytopathol 10:3–9, 1994.
10. Morrow CP, Cozen W: Perspectives on cervical cancer. Why prevent? J Cell Biochem (Suppl 23), 1995, this issue.
11. Walton RJ, Blanchet M, Boyes DA, Carmichael JA, Marshall KG, Miller AB, Thompson DW, Hill GB: Canadian Deputy Ministers' Task Force Report. Can Med Assoc J 114:1003–1033, 1976.
12. Kirby AJ, Speigelhalter DJ, Day NE, Fenton L, Swanson K, Mann EMF, MacGregor JE: Conservation treatment of mild/moderate cervical dyskaryosis: Long-term outcome. Lancet 339:828–831, 1992.
13. Reis LAG, Miller BA, Hankay BF, Korsay CL, Harras A, Edwards BK (eds): "SEER Cancer Statisitics Review, 1973–1991: Tables and Graphs," National Cancer Institute. NIH Pub. No. 94–2789, Bethesda MD, 1994.
14. Celentano DD, deLissovoy G: Assessment of cervical cancer screening and follow–up programs. Pub Health Rev 17:173–240, 1989/90.
15. van Ballegooijen M, Habbema JDF, van Oortmarssen GJ, Koopmanschap MA, Lubbe JThN, van Agt HME: Preventative Pap smears: Balancing costs, risks and benefits. Br J Cancer 65:930–933, 1992.
16. Fleiss JL: Determining sample sizes needed to detect a difference between two proportions. In "Statistical Methods for Rates and Proportions." New York: John Wiley & Sons Inc., 1981, pp 33–49.
17. Tezcan A, Garner DM, Lam P, Korbelik J, Palcic B: An investigation of the three nuclear stains: Thionin, gallocyanin and hematoxylin for automated quantitative image cytometry of specimens from the uterine cervix. Cytometry, 1995, (in press).
18. Anderson GH, Benedet JL, LeRiche JC, Matisic JP, Thompson JE: Invasive cancer of the cervix in British Columbia: A review of the demography and screening histories of 437 cases seen from 1985–1988. Obstet Gynecol 80:1–4, 1992.