Available online at www.sciencedirect.com

SCIENCE DIRECT®

Gynecologic Oncology 94 (2004) 296–306

Gynecologic Oncology

www.elsevier.com/locate/ygyno

Quantitative histopathology and chromosome 9 polysomy in a clinical trial of 4-HPR

Jose-Miguel Yamal,[a,b] Dennis Cox,[b] Walter N. Hittelman,[c] Iouri Boiko,[d] Anais Malpica,[e] Martial Guillaud,[f] Calum MacAulay,[f] Michele Follen,[g,*] and Anne-Therese Vlastos[g,h]

[a] *The Department of Biostatistics, The University of Texas M. D. Anderson Cancer Center, Houston, TX 77030, USA*
[b] *The Department of Statistics, Rice University, Houston, TX 77005, USA*
[c] *The Department of Experimental Therapeutics, The University of Texas M. D. Anderson Cancer Center, Houston, TX 77030, USA*
[d] *The Department of Pathology, The University of Texas Health Science Center, Houston, TX 77030, USA*
[e] *The Department of Pathology, The University of Texas M. D. Anderson Cancer Center, Houston, TX 77030, USA*
[f] *The Department of Cancer Imaging, British Columbia Cancer Centre, British Columbia, Vancouver, Canada V5Z 1L3*
[g] *The Department of Gynecologic Oncology, The University of Texas M. D. Anderson Cancer Center, Houston, TX 77030, USA*
[h] *Départemente de Gynécologie et Obstétrique, Hôpitaux Universitaires de Genéve, Geneva, Switzerland*

Received 4 November 2003
Available online 20 June 2004

Abstract

Objective. This trial examined the use of 4-hydroxyphenyl-retinamide (4-HPR), demonstrated to be a potent inhibitor of carcinogenesis in vitro and in animal models, in patients with cervical intraepithelial neoplasia (CIN) grades 2 to 3. Quantitative pathology and chromosome 9 polysomy were used to understand the biology and quantify the clinical histopathologic changes observed.

Methods. Patients were randomized to 4-HPR or placebo for 6 months and followed for six more months. Cervical biopsies were obtained at baseline, 6 months, and 12 months; the biopsies were read blinded three times by the study pathologist. Feulgen-stained sections were also obtained and analyzed using computer-assisted image cytometry. Chromosome 9 polysomy was performed on tissue slices using in situ hybridization and measured quantitatively. Statistical analyses were carried out in S-Plus (Insightful Corporation, Seattle, WA) and R.

Results. The interim analysis, planned for 40 patients, was carried out on 39. The 6- and 12-month analyses showed a statistically significant difference between the two study arms. When code was broken, the 4-HPR-treatment arm was found to have fared less well than placebo. Analyses of Feulgen-stained sections provided a quantitative measure of the increase of DNA content and texture features. Chromosome 9 polysomy was also measured using image analysis. The changes observed were consistent with those of cells displaying cancerous changes, indicating a lack of response.

Conclusion. 4-HPR is not active at 200 mg/day. The interim analysis was helpful in directing the study; and, in this case, ending it. The intermediate endpoint biomarkers of quantitative histomorphometry and chromosome 9 polysomy yielded quantitative and repeatable results consistent with the findings of the clinical pathologist.

Keywords: Chemoprevention; Cervical intraepithelial neoplasia; 4-HPR; Intermediate endpoint biomarkers; Quantitative histopathology; Chromosome 9 polysomy

Introduction

4-hydroxyphenyl-retinamide (4-HPR) has been found to be a potent inhibitor of mammary carcinogenesis in the rodent [1]. Because of its wide applicability in cell cultures from 30 carcinoma cell lines, there is great promise for its clinical use (reviewed in Ref. [2]). The mechanism of action of 4-HPR is thought to be well understood as being mediated

* Corresponding author. The Center for Biomedical Engineering, The University of Texas M.D. Anderson Cancer Center, 1515 Holcombe Boulevard, Unit 193, Houston, TX 77030. Fax: +1-713-792-4856.
E-mail address: mfollen@mdanderson.org (M. Follen).

0090-8258/$ - see front matter © 2004 Elsevier Inc. All rights reserved.
doi:10.1016/j.ygyno.2004.04.030

by oxygen radicals and not by retinoic-acid receptors [3–6]. A comparatively nontoxic retinoid, 4-HPR is being studied in randomized clinical trials of preneoplastic precursor lesions in the breast, lung, bladder, prostate, and cervix [7].

Veronesi reported the results of a nonplacebo-controlled randomized trial in 2972 Italian women with stage I breast cancer in whom the endpoint of interest was the prevention of contra- or ipsilateral breast cancer. After 97 months of follow-up, no statistically significant differences were noted in the primary endpoints. Post hoc analysis showed a suggested benefit in premenopausal women and significantly decreased contralateral cancers in this group [8].

Kurie conducted a randomized trial of 4-HPR in patients with bronchial metaplasia. No statistically significant differences were noted in the histopathologic regression of lesions in the treated group compared with the placebo group. Similarly, there were no changes in the intermediate endpoint biomarkers under study, which included analysis of loss of heterozygosity at putative tumor suppressor loci on chromosomes 3p, 9p, and 17p [9].

This study reports the corroboration of quantitative histomorphometric and chromosome 9 polysomy analysis with the conventional histopathologic analysis in a placebo-controlled randomized clinical trial in patients with cervical intraepithelial neoplasia (CIN), grades 2 to 3(high-grade squamous intraepithelial lesions). Quantitative histomorphometry and chromosome 9 polysomy were carried out in this study as an intermediate endpoint biomarker of cervical carcinogenesis. The quantitative staining of samples has been shown to provide information about the amount and distribution of DNA in the nuclei of normal and abnormal cells, which can reduce the uncertainty of subjective visual evaluation. Chromosome 9 polysomy has been demonstrated to be a reliable biomarker of genetic instability. In head and neck cancers, genetic instability increased as lesions progressed from normal to hyperplasia to cancer. Studies in patients with head and neck cancers show that histologically normal epithelium demonstrates chromosome polysomy, whereas normal epithelium from control subjects does not [10]. The intent of the study was to understand the biology, quantify the biologic changes in tissue, and refute or confirm agreement with the clinical histopathologic results.

Materials and methods

Tissue materials

The placebo-controlled randomized clinical trial is reported elsewhere [11]. Briefly, patients with high-grade squamous intraepithelial dysplasia, confirmed by colposcopy and biopsy, signed consent forms and were randomized to 4-HPR or placebo for 6 months. 4-HPR was given at 200 mg/day for 6 months, with a 3-day drug holiday monthly. This dose was selected by the National Cancer Institute because of Phase I data showing nyctalopia and recovery of retinol levels with a 3-day monthly drug holiday. After a period of 6 months, the colposcopy exam and biopsy were repeated. In case of progression of the disease, code was broken and a patient under placebo received the drug. If already on the drug, the patient was treated by standard treatment loop electrosurgical excision procedure (LEEP). Baseline biopsies were compared with 6- and 12-month biopsies. If a cervical lesion was present at the 12-month visit, a loop excision of the entire transformation zone of the cervix was carried out, removing all the tissue at risk (Fig. 1).

Histological analysis

A blinded review of the hematoxylin- and eosin-stained biopsies was carried out by the study pathologist on the three separate occasions of the study biopsies. Discrepancies between the first and second readings were resolved in the third review. The principal investigator and pathologist were blinded to the treatment arm and assigned the patients to four categories at the 6- and 12-month endpoints: complete response, partial response, no change, and progression. Complete response was defined as no cervical intraepithelial neoplasia (CIN) in the loop excision specimen. Partial response was defined as regression to CIN grades 1 to 2. No change was defined as the continued presence of CIN 2 to 3. Progression was defined as the progression from CIN 2 to 3 to squamous carcinoma in situ. For some analyses, response categories were collapsed, so that responders were those with partial and complete responses, and nonresponders were those with no change and progressive disease. Again, all disease status was based on clinical histopathologic readings read three times by an investigator blinded to outcome. Discrepant readings were resolved during subsequent readings.

Quantitative histopathology

Additionally 4-μm sections were cut and stained with Feulgen. The details of the Feulgen stain are published elsewhere [12,13]. The Feulgen-stained sections were reviewed and mapped with the hematoxylin–eosin sections. The areas that were mapped were measured with the Cytosavant (Oncometrics, Vancouver, BC, Canada) (Fig. 2), an automated image cytometer that includes a 12-bit Microimager 1400 digital camera (picture elements 6.8 × 6.8 μm). The Feulgen-stained nuclei were imaged with monochromatic light at 600 nm (10-nm bandwidth) using a 20 by 0.75NA Plan Apo lens. The effective pixel size is 0.34 × 0.34 μm. Nuclei were located using a grey level thresholding procedure and a refined individual cell-focusing algorithm. The segmentation of the object was standardized by an algorithm that places the edge at the highest local grey level. The sections were measured using the SLICE program, and the results were analyzed blinded to treatment arm using S-Plus and R. A subset of 126 nuclear features is measured

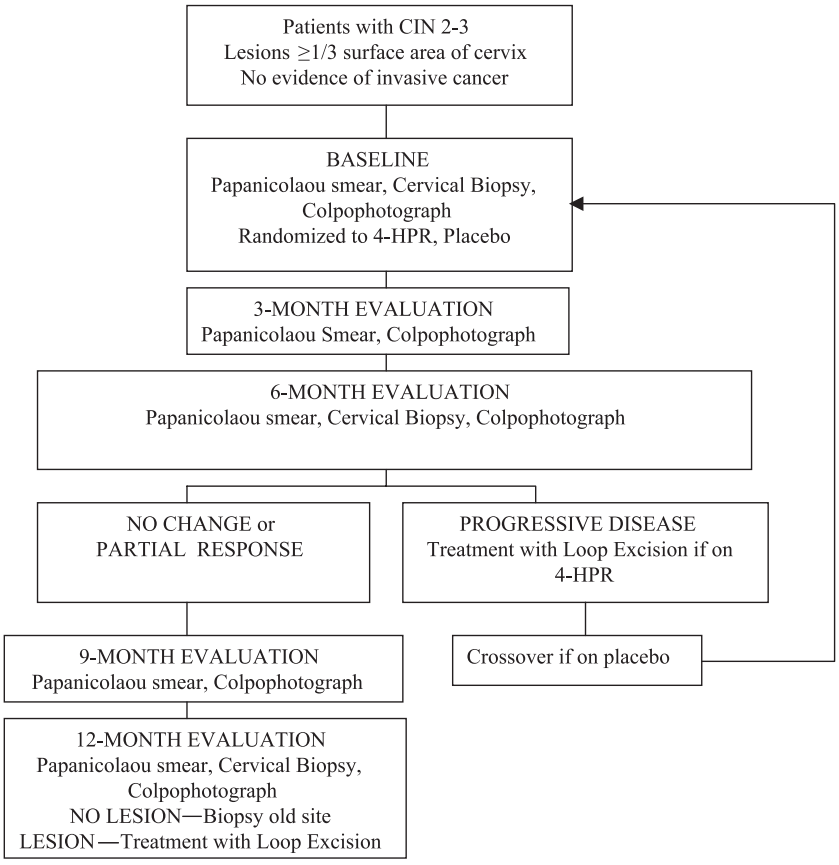

Fig. 1. Schema of the study and evaluation for phase II randomized clinical trial of 4-HPR and placebo in CIN 2 to 3. The colpophotograph is a picture of the cervix taken under magnification during colposcopy.

with the Cytosavant [12]. The 126 features can be further classified as (1) morphometric: describing the size, shape, and boundary irregularities of the nucleus; (2) photometric: describing the distribution of optical density within the stained nucleus, including the integrated optical density, (proportional to the DNA content); (3) discrete texture: relying on the division of the nucleus into regions of high, medium, and low-chromatin condensation states, and describing the spatial distribution and photometric properties of these regions; (4) Markovian texture: features that characterize the distribution of grey-level intensity values between adjacent pixels in the image; (5) run-length texture: features that describe the length of contiguous regions with constant grey-level intensity; and (6) fractal texture: which measures the surface area of a three-dimensional plot of optical density vs. the position in the nuclear image. The slides were scanned for 3–4 h by the histopathologist to detect all diagnostic cells in the mapped areas. Lymphocytes were used to normalize the DNA content per cell.

Chromosome in situ hybridization (CISH)

Tissue sections (4 μm) from paraffin blocks of leukoplakia biopsies were placed on xylene-coated slides. The slides were placed overnight on a slide warmer at 65°C, then dewaxed in xylene, and cleared in 100% ethanol. The slides were then treated with 1 mg/ml Rnase in 2× SSC and digested with 0.4% pepsin (Sigma-Aldrich, Inc., St. Louis, MO) in 0.2 N HCI, as described previously [10].

Biotin-labeled α satellite DNA (0.8 ng/μl), specific for the pericentromeric region of chromosome 9 (Oncor, Gaithersburg, MD), was mixed in a hybridization solution composed of 60% formamide in 2× SSC, 5% dextran sulfate, and 1 mg/ml salmon-sperm DNA. The hybridization solution was placed on the tissue section, and the two were denatured together at 93°C for 6 min and incubated at 37°C overnight. The next day, the sections were washed in 50% formamide in 1× SSC (pH 7.0) at room temperature twice for 15 min each, followed by three washings in 0.1× SSC at 37°C for 10 min each. The slides were then treated with a 3% BSA-blocking solution for 10 min and incubated with avidin (Vector, Burlingame, CA) and, subsequently, with antiavidin (Vector). These steps were repeated to amplify the signal, and then an avidin–biotin peroxidase complex solution was applied. Each incubation step was carried out at 37°C for 30 min, followed by a wash with PBS. The hybridization signals were developed with 50 mg of $NiCl_2$ in 100 ml of PBS and counterstained with Giemsa [10].

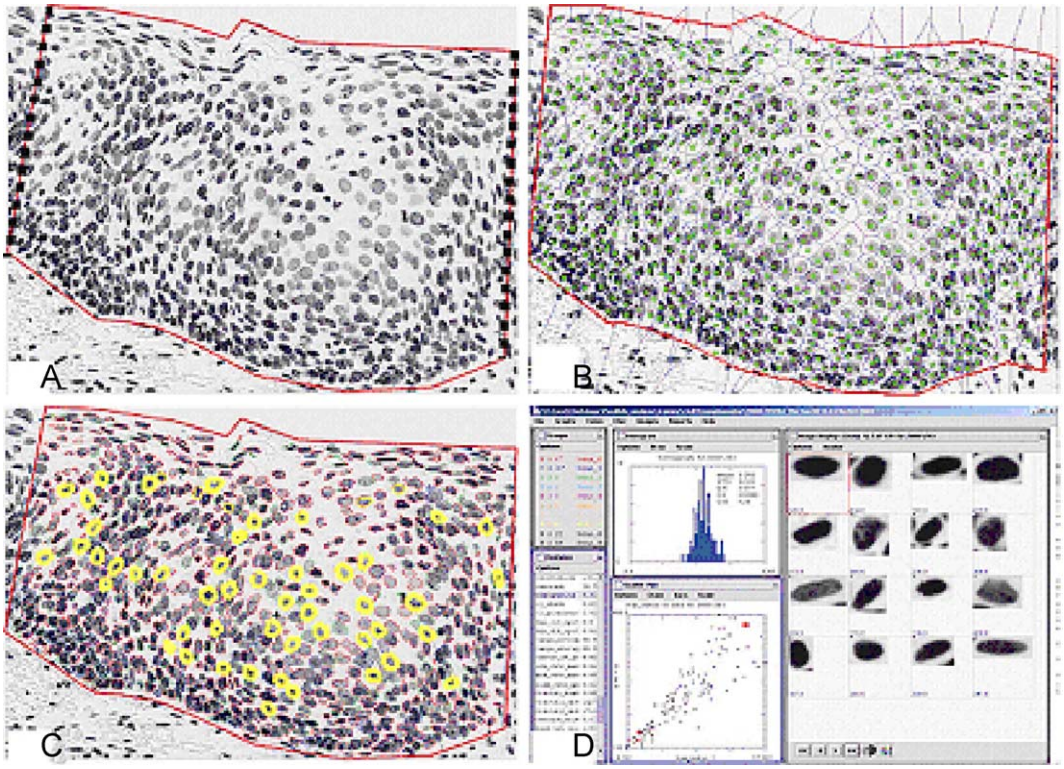

Fig. 2. Steps of the semiautomated analysis of cervical lesions: (A) definition of the return on investment; (B) automatic locations of the positions of the nuclei; (C) automatic segmentation of the nuclei in the intermediate layers of the epithelium; (D) snapshot of the interface of our software used for post-analysis of the morphometric analyses (quality control).

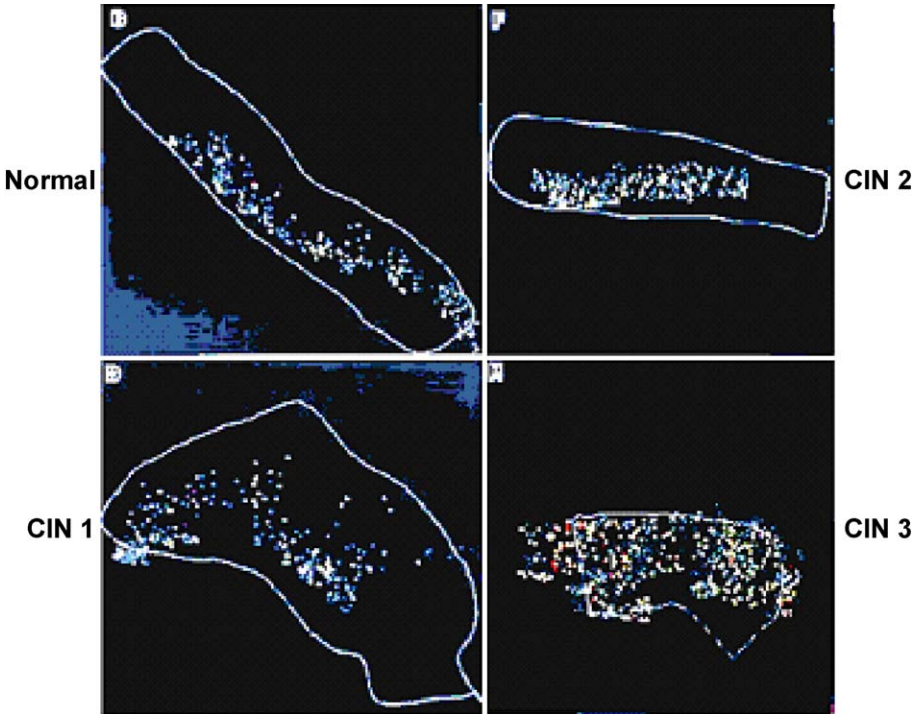

Fig. 3. Chromosome 9 polysomy in normal cervix, CIN1, CIN2, and CIN3. Cells with three and four copies of chromosome 9 are represented by red dots.

Analysis of chromosome signals

The levels of chromosome hybridization signals in lymphocytes infiltrating the submucosa served as controls for determining the hybridization efficiencies in the epithelial layers. Chromosome copy numbers were assessed separately in the basal, parabasal, and superficial cell layers to determine whether the detected degree of genetic instability was influenced by cell maturation and location in the epithelium. From 100 to 900 cells were counted in each predetermined region at ×1000 magnification. The Cell Index for each region was calculated by dividing the total number of signals detected by the total number of nuclei scored. The chromosome Polysomy Index was defined as the percentage of scored nuclei exhibiting three or more chromosome copies. Tissue sectioning results in the truncation of nuclear material and leads to an underrepresentation of chromosome copy number. To avoid bias, nuclei exhibiting no signals were included in all calculations.

The Magiscan Image Analysis system (Joyce Loebl, Ltd., Dukesway, England) was used to record the spatial distribution of chromosome counts in the epithelium. The relative *x*- and *y*-coordinates and detected chromosome copy number of each scored nucleus were recorded in list mode. This permitted the preparation of a genetic map of the tissue section where each nucleus is represented by a dot, the color of which represents the chromosome copy number detected for that cell (Fig. 3) [10].

Statistical analysis

The details of the statistical analysis used in the clinical trial are reported elsewhere [11]. Fisher's exact tests and nonparametric tests were used to compare variables showing that specimens with sufficient tissue for analysis were not significantly different from those with insufficient tissue for analysis. As with all studies of biomarkers, histologic blocks may or may not have sufficient material for analysis. The most important part of the methods is ascertaining that no bias exists in the samples that are sufficient for analysis compared to those that are insufficient for analysis.

Data were generated from the quantitative histopathological data cell by cell and analyzed on a patient level. Mean differences were used to compare pre- and posttreatment patients stratified by placebo and 4-HPR groups. Tests of statistical significance included the Wilcoxon rank-sum test, and Westfall and Young's Multiple Comparison Procedure was used to correct for multiplicity. Those that were statistically significant were further studied using the Alternating Conditional Expectation (ACE) analysis.

The Polysomy Index and Cell Index are both measures of chromosome 9 polysomy. An arcsin transformation was applied to percentages within a given patient at baseline and 6 months. A two-sample *t* test was performed on the differences within a given patient. Those that were statistically significant were further studied using the ACE analysis.

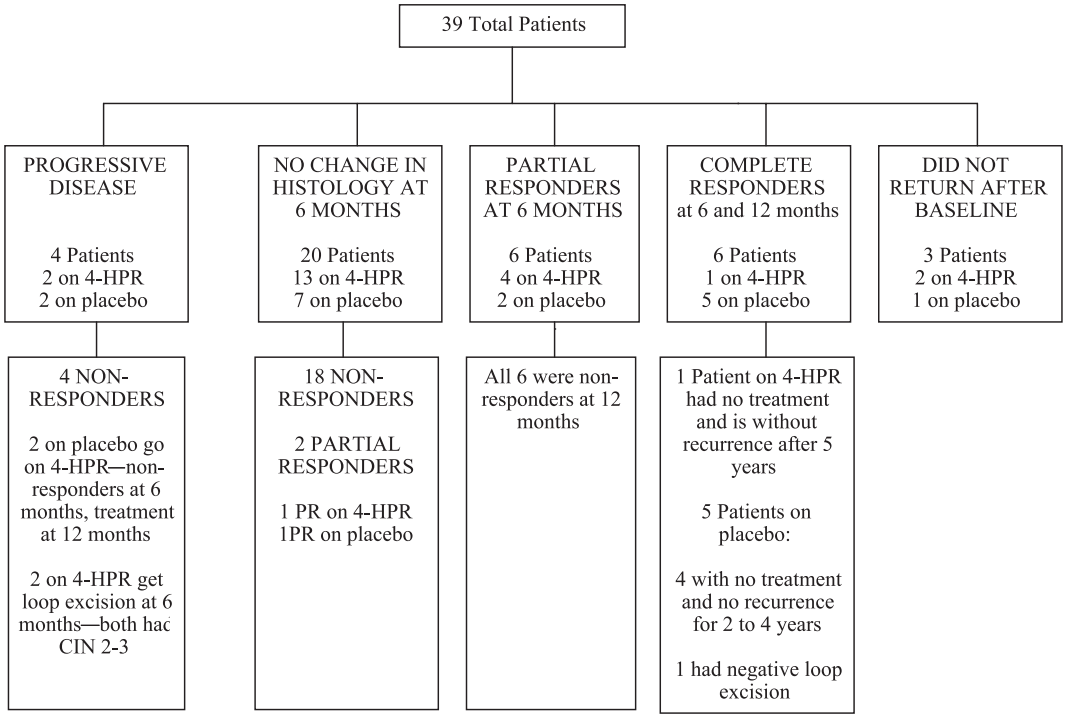

Fig. 4. Schema of patient outcomes for all patients in the randomized clinical trial of 4-HPR and placebo.

S-Plus and R were used to analyze both the quantitative pathology feature and the Polysomy Index and Cell Index; the features blinded to treatment arm and outcome status [14]. Westfall and Young's Multiple Corrections Procedure [15] was used to account for multiple testing.

To explore possible relationships between variables from the quantitative histology and variables in the chromosome 9 polysomy analysis, we used the ACE method for finding optimal transformations [16]. This data-analytic method finds transformations of each variable such that the pair of transformed variables has maximal correlations. Additionally, one can impose a monotonicity constraint on the transformations. To test the statistical significance of the result, we used a randomization-based method [17]. The S-Plus implementation of the ACE algorithm outputs the estimated maximal r^2. We randomized the assignment of the paired values of the two variables (the individual values were unchanged, but we "shuffled" the values of one variable so they were randomly paired with the value of the other variable). For each randomized data set, we computed maximal r^2 from the ACE algorithm. The proportion of r^2 from the randomized data that equaled or exceeded the r^2 from the real data gave a P value for the significance of the relationship between the two variables.

Results

The interim clinical trial data showed no statistically significant differences at 6 months; however, histopathologic review showed one group faring less well than the other. The 12-month analysis of data showed statistically significant differences, consistent with the histopathologic review that the same group identified at 6 months was faring much less well. When code was broken, the treatment arm (4-HPR) was found not to be performing as well as placebo. Analysis of confounders provided no explanation. Figs. 1 and 4 describe the conduct and the results of the clinical trial. The 6-month response rates were 25% for 4-HPR and 44% for placebo ($P = 0.25$). The 12-month response rates were 14% for 4-HPR and 50% for placebo ($P = 0.04$) [13]. The histopathologic analysis showed that 4-HPR did not regress lesions at 6 months, and 6 months after therapy may still be retarding the natural regression.

Analysis for bias in specimens

Sufficient tissue for Feulgen stain mapping for quantitative histomorphometry was available for 26 women. First, an analysis was performed to show that these 26 patients were not statistically significantly different from the 13 from whom insufficient tissue was available (Table 1). Of the 26, 15 patients had been randomized to 4-HPR and 11 to placebo. Paired measurements for baseline and 6 months, and 6 and 12 months were available.

Nineteen samples had tissue available at baseline and 6 months for chromosome 9 polysomy analysis. In this sample, 12 were in the 4-HPR group and 7 were in the placebo group. Again, an analysis was performed to ensure that these patients did not differ substantially from the 20 who had tissue available for analysis.

Because the samples from the quantitative histopathology and chromosome 9 polysomy were to be compared, we established there were no significantly different number of responders than nonresponders in either group (Table 1).

Quantitative histopathology results

The DNA amount was the raw measurement of the "integrated optical density" from which all the photometric features were derived. To normalize the measure of "integrated optical density" of the object, the DNA amount of the cervical cells was divided by the mean value of the DNA in leukocytes measured in the cervical stroma (integrated optical density norm). This value, DNA index, is a well-accepted marker for the ploidy of cervical cells in the quantitative pathology literature [12,13]. In particular, the percent of cells that had a DNA index value of over 2.2 was considered as being aneuploid. Paired measurements for baseline and 6 months, and 6 and 12 months were available.

The proportions were transformed using the arcsin transformation to stabilize the variances and to enable us to take differences between the 6-month and baseline values for each patient. A standard two-sample t test resulted in a significant difference between the treated and the placebo patients ($P < 0.05$). The placebo cells abated back to a more normal state, while the treated cells progressed slightly in the cancerous direction. This was a statistically significant finding showing the biology did not reverse as expected with the medication.

Pair-wise analysis of each nuclear feature was done using a Wilcoxon rank-sum test to avoid the assumption of normality. In Table 2, we see that of the 114 nuclear features, 22 were significant at the 0.05 level, uncorrected for multiplicity. They are shown in Table 2, along with their means at baseline and at 6 months and the Wilcoxon rank-sum test P value [18]. It is striking that, of these significant variables, 20/22 suggested that the treated group had more cancerous characteristics (increased area, mean radius, etc.) than the placebo group. The histomorphometric features of cervical cells have been studied by many groups, and there is a consensus regarding the changes in features as cells move from normal to cancerous [12,13,18]. Westfall and Young's Multiple Comparison Procedure found only the variable OD skewness to be significantly different, with a higher value for the treated arm.

Fig. 5 is a scatter plot of the patients at baseline vs. 6 months. Fig. 5A shows the percentage of cells in patients with a DNA index over 2.2. In this graph, we see that more of the treated patients are in the upper left-hand part of the graph at 6 months; in other words, more of the treated

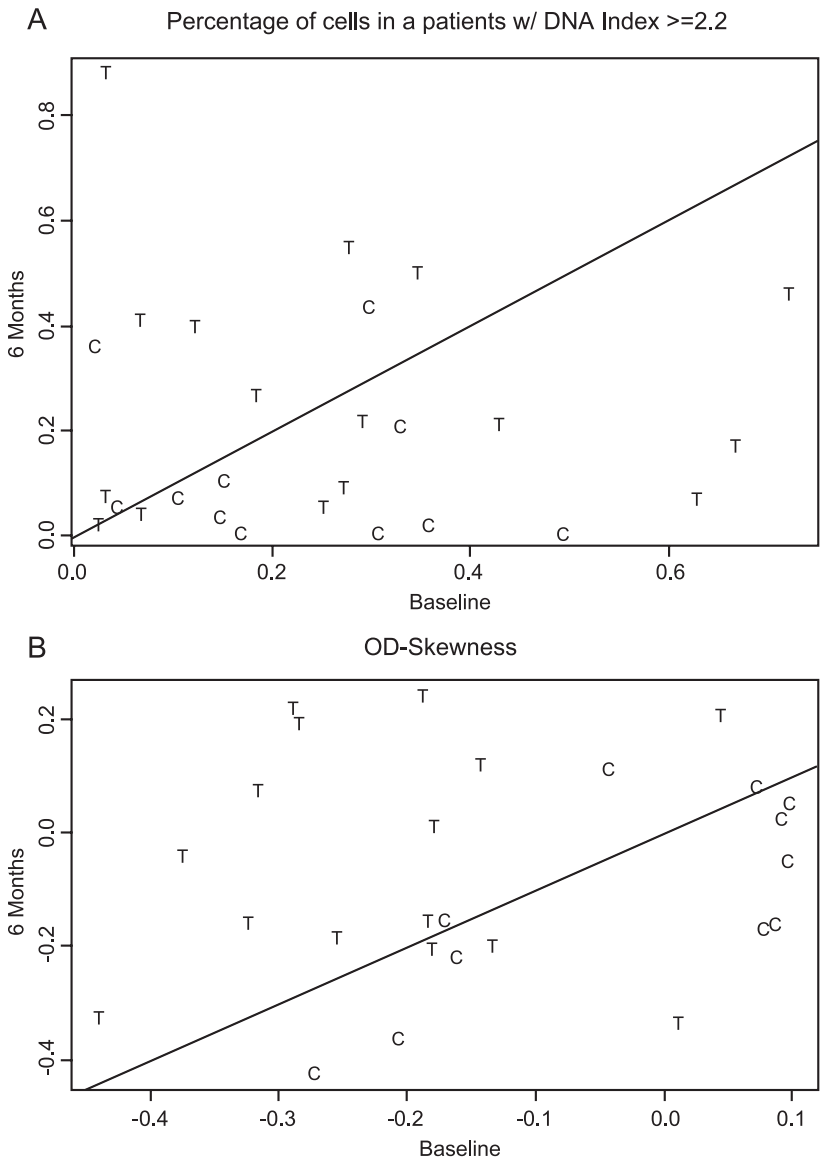

Fig. 5. (A) Scatter plot of significant histomorphometric variable "percentage of cells with a DNA index > 2.2" at baseline vs. 6 months into the trial. The treatment (4-HPR) and control patients are denoted by T and C, respectively. Patients above the line indicate progression towards cancer and patients below the line indicate regression towards normality. The regression present in the control appears to be inhibited in the treated group. (B) Scatter plot of significant histomorphometric variable Optical Density Skewness at baseline vs. 6 months into the trial. The treatment (4-HPR) and control patients are denoted by T and C, respectively. Patients above the line indicate progression towards cancer, and patients below the line indicate regression towards normality. The regression present in the control appears to be inhibited in the treated group.

patients have aneuploid cells at 6 months than the control patients, suggesting either the medication induced a unfavorable response or it prohibited the regression noted in the placebo group. In Fig. 5B, we are examining Optical Density (OD) Skewness, that variable which was significant after the Multiple Comparison Procedure. This variable refers to the amount of chromatin in a cell and how its measurement can be skewed when the texture of the chromatin is uneven. Optical density skewness is a measure of the inhomogeneity of the DNA stain in the nucleus and reflects changes in the chromatin organization. Again, patients above the line indicate those having more cancerous characteristics or progression towards cancer. The treated patients are in the upper left-hand part of the graph and the control patients are in the lower left-hand part of the graph. This shows that the chromatin of the control patients was more normal after 6 months of being on placebo than that of the treated patients.

Chromosome 9 polysomy results

The Cell Index was counted as the total number of copies of chromosome 9 copies divided by the total number of cells counted per patient (average per cell). Most patients in the treated arm had Cell Index values that increased, while the control patients stayed relatively the same or decreased on

Table 1
Specimens with sufficient tissue for biomarker analysis demonstrating that no significant biases existed among available specimens compared to unavailable specimens in response and nonresponse

Measurement type	6 month status (nonresponse)	6 month status (response)	*P* value of rows	*P* value of rows
Quantitative histopathology				
Included	15	11	+	*
			0.15 +	
Not included	9	1	+	
				0.77*
Chromosome polysomy				
Included	10	9	+	*
			0.083 +	
Not included	14	3	+	

average. A Wilcoxon rank-sum test gave a significant *P* value for this difference ($P = 0.025$). The Polysomy Index showed the percentage of cells with three or four copies of chromosome 9 per patient. The Polysomy Index also showed an increase in the value for the treated group, while the value for the control group generally decreased. The increase was not statistically significant.

Figs. 6A and B show scatter plots of the Cell Index and Polysomy Index for the patient samples. Fig. 6A shows the chromosome 9 Cell Index. Again, the majority of the treated samples are in the left hand corner, indicating more genetic instability after 6 months of treatment, while more of the controls show less genetic instability after 6 months of treatment. Fig. 6B shows the chromosome 9 Polysomy Index, similar to the Cell Index. Most of the treated cases are in the upper left-hand corner, and most of the controls are in the lower right-hand corner, suggesting more genetic instability after 6 months of treatment and less genetic instability after 6 months of placebo.

The percentages were transformed with the arcsin transformation to stabilize the variances; then the paired two-sample *t* test gave a *P* value of 0.054. If we just look at the percentage of four copies of chromosome 9, then the paired two-sample *t* test *P* value is 0.034. While we were unable to establish significance using nonparametric tests, the arcsin transformation permitted us to look at the data in a different manner that establishes significance. This suggests that a larger data set is required to look at the variable adequately.

Correlations among quantitative histomorphometry and chromosome 9 polysomy

To explore possible relationships between variables from the quantitative histology and variables in the chromosome 9 polysomy analysis, we used the ACE method for finding optimal transformations [17]. The correlations among the DNA index, the percentage of cells with a DNA index over 2.2, optical density skewness, with chromosome 9Cell Index, the Polysomy Index, and the percentage of cells with four chromosomes were analyzed. High correlations were noted, affirming that from both the perspective of aneuploidy and genetic instability, the treated patients have worse histopathology at 6 months. These correlations are present in Table 3. This table shows that Cell Index and Polysomy Index are highly correlated with the DNA Index and the aneuploidy index. There was only a suggestion, one *P* value of 0.08, of significance with OD skewness. Few items

Table 2
Means and paired *P* values of significant variables at baseline and at 6 months

Feature	Treatment pre mean	Treatment post mean	Control pre mean	Control post mean	*P* value
OD skewness	− 0.2155	− 0.0392	− 0.0299	− 0.1211	0.001
fractal2.area	1155.23	1337.79	1245.69	909.38	0.004
run90.length	121.8852	147.5562	136.139	119.7772	0.004
run135.length	118.4471	146.1024	129.055	120.5111	0.005
runlength.mean	106.97	128.85	117.30	106.52	0.005
hiDNAcomp	2.705	3.9904	3.5969	3.1388	0.006
gray0.level	30.3556	38.8535	33.9321	30.9565	0.006
run45.length	93.2325	109.6601	100.2125	91.018	0.006
run0.length	94.3112	112.0793	103.8112	94.7565	0.008
gray45.level	30.3002	38.3808	33.3916	30.2908	0.013
runlength.stdv	23.5718	28.3945	25.2314	21.7873	0.013
fractal1.area	7129.59	7963.55	7470.69	5485.00	0.015
gray90.level	36.2373	45.6044	40.0471	35.7781	0.018
Harmon19.fft	0.3207	0.2898	0.297	0.3169	0.020
Area	376.3931	447.3927	393.1868	350.7978	0.024
high.den.obj	1.2307	1.8812	1.6717	1.6293	0.024
den.lit.spot	0.0029	0.0063	0.0055	0.0054	0.032
gray135.level	35.5498	45.2637	38.7318	35.7582	0.032
fractal.dimen	2.6482	2.5942	2.6045	2.608	0.036
mean.radius	10.5744	11.4943	10.7588	10.1564	0.041
max.radius	14.0465	14.8388	13.9941	12.8504	0.041
mhDNAcomp	1.7699	2.0423	1.8929	1.8538	0.041

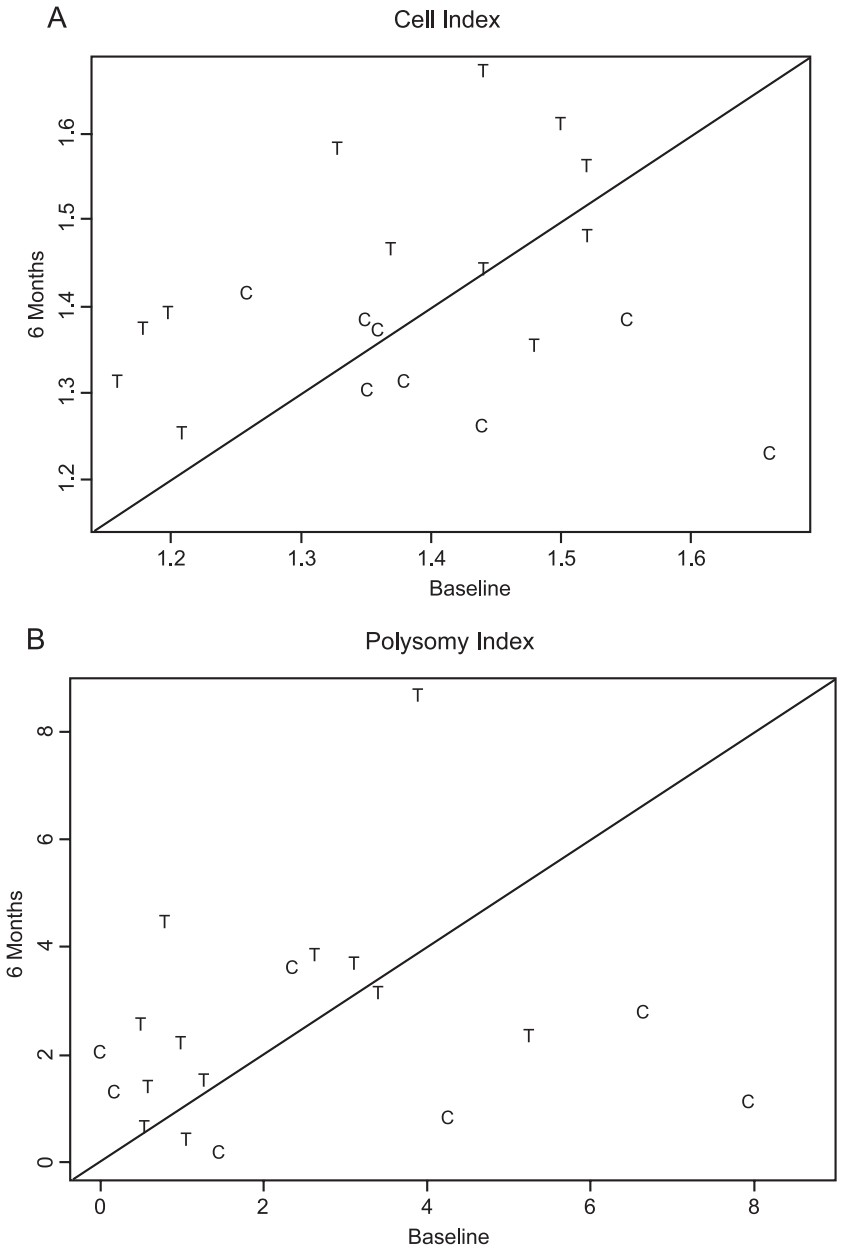

Fig. 6. (A) Scatter plot of significant chromosome polysomy variable Cell Index at baseline vs. 6 months into the trial. The treatment (4-HPR) and control patients are denoted by T and C, respectively. Patients above the line indicate progression towards cancer, and patients below the line indicate regression towards normality. The regression present in the control appears to be inhibited in the treated group. (B) Scatter plot of significant chromosome polysomy variable Polysomy Index at baseline vs. 6 months into the trial. The treatment (4-HPR) and control patients are denoted by T and C, respectively. Patients above the line indicate progression towards cancer, and patients below the line indicate regression towards normality. The regression present in the control appears to be inhibited in the treated group.

remained significant after permutation testing, but this is expected with data sets of these magnitudes.

Discussion

The intent of this study was to understand the biology, quantify the biologic changes in tissue, and refute or confirm agreement with the clinical histopathologic results. Previous work has shown that quantitative histomorphometry provides a reliable biologic marker of cancerous changes. In this study, the treated group had more cancerous biologic effects than the placebo group. The findings were statistically significant and agreed with the clinical histopathology. This shows that the pathway of carcinogenesis was not interrupted as it should have been by 4-HPR, that this could be quantified, and that it agrees with the clinical histopathologic findings.

Table 3
Correlations between the quantitative histopathology variables and the chromosome 9 variables

Quantitative histopathology variables	Chromosome 9 variables						
		Cell index			Polysomy index		
		Original *P*	ACE *P*	Permutation test *P*	Original *P*	ACE *P*	Permutation test *P*
	DNA Index	0.058	0.0004	0.085	0.09	0.007	0.231
	Percent DNA Index > 2.2	0.02	0.009	0.199	0.035	0.001	0.19
	OD skewness	0.49	0.1	0.76	0.285	0.077	0.75

The *P* value for the correlation between the original variables is calculated from a Spearman rank correlation test, the maximal correlation is from an Alternating Conditional Expectation (ACE) test, and the permutation test tests the significance of the relationship between the two variables.

Chromosome 9 polysomy has been demonstrated to be a reliable biomarker of genetic instability. In this study, chromosome 9 polysomy was found to be increased in the treated group showing that 4-HPR did not allow the normal regression that occurred in the placebo arm. Again, these results were quantifiable and correlated with the clinical histopathologic results. Moreover, the quantitative results of both the histomorphometry and the chromosome 9 polysomy agree with each other and demonstrate that the expected biology of regression did not occur in the treated arm.

4-HPR is inactive at 200 mg/day for 6 months in CIN 2 to 3. Moreover, 4-HPR may be inhibiting whatever natural regression occurs, as demonstrated in the placebo arm of this trial. While a very promising compound in the laboratory and in animal studies, 4-HPR may not be active at these oral doses in humans. While Phase I studies were carried out in patients at risk for breast cancer and other cancers (reviewed in Ref. [11]), they did not predict the correct dose level for this population and this precancerous lesion. Because of the ocular toxicity nyctalopia, the dose level was set at 200 mg/day for 6 months, with a 3-day drug holiday [19]. Data from Mehta et al. [20] suggest that this dose may lead to biologically relevant levels in the breast. The tissue levels have not been well documented in other organ sites. Nyctalopia is probably less common in younger patients. Each study needs it own phase I.

Interim analysis is helpful. Without this interim analysis, we would have continued to accrue patients in this trial. Clearly, the 12-month data showing a similar but more worrisome trend justified breaking the code. It is a very difficult decision to stop a study. Many resources have been invested. The investigators and patients are enthused. Ethics and patient safety must always be the highest priority. 4-HPR may have been altering the natural regression process that occurs at 6 and 12 months at this dose of 200 mg/day.

This quantitative histomorphometric and chromosome polysomy assessment agrees with the conventional pathologic analysis and demonstrates findings that the treatment arm was faring more poorly than the placebo arm at the cellular level. These data require sophisticated biostatistical analyses. If this were a random effect, the distribution of cancerous characteristics would have been randomly distributed between groups; it was not. The figures are illustrative of this and demonstrate the differences between placebo and treated arms. The correlation of the two measures of carcinogenesis was reassuring. Furthermore, the fact that both biomarkers confirmed the histopathologic results lends strong support for closing the study.

Probably the most interesting finding from this study is the high regression rate noted in the placebo arm, which was 44% at 6 months and 50% at 12 months. This was certainly a surprise to the investigator who designed the study (MF).

There are several important lessons from this carefully conducted clinical investigation. The colpophotographs do not show this regression. The cervical biopsies were read three times in blinded fashion by the histopathologist and have been reviewed by colleagues in the pathology department as part of routine care.

As this was an NCI-funded trial, the principal investigator (MF) reviewed all of the material with the pathologist in blinded fashion. All of the data were entered into the computer, verified, doubly entered, and analyzed. The treatment code was not broken until we were certain that one arm of the trial was faring more poorly than the other. The trial was stopped, yet all the biomarker analyses were still performed in blinded fashion. Thus, we are quite certain of the results. The only explanation for this high regression rate is the placebo effect plus the effect of being biopsied three times in a year.

In the review of natural history written with Mitchell et al. [21], we did note a lower rate of progression to cancer and a higher rate of regression in those patients followed with biopsies compared to those patients followed with cytology. Furthermore, having looked at the rate of regression in several other placebo-controlled chemoprevention trials of high-grade lesions, 40% is not that uncommon [22]. Admittedly, this does has implications for trial design. More patients with high-grade lesions are needed, and they must be followed for longer periods of time.

Intermediate endpoint biomarkers are surrogate endpoints of cancer incidence. As this report indicates, they have been useful in an interim analysis, helping confirm a puzzling clinical result and helping stop a clinical trial.

Acknowledgments

Funding from NCI Chemoprevention Contract #NO1-CN-25433A and Program Project Grant #P01-CA-82710 is gratefully acknowledged. Dr. Vlastos was a visiting scholar when this work was performed and was supported by the Swiss National Science Foundation, the Swiss Cancer League, the Cancer and Solidarity Foundation, and the Novartis Foundation.

References

[1] Moon RC, Thompson HJ, Becci PJ, Grubbs CJ, Gander RJ, Newton DL, et al. *N*-(4-hydroxyphenyl)retinamide, a new retinoid for prevention of breast cancer in the rat. Cancer Res 1979;39:1339–46.

[2] Follen M. A clinical trial of 4-hydroxyphenylretinamide for the treatment of high grade squamous intraepithelial neoplasia. Dissertation. University of Michigan, 2000.

[3] Oridate N, Suzuki S, Higuchi M, Mitchell MF, Hong WK, Lotan R. Involvement of reactive oxygen species in *N*-(4-hydroxyphenyl)retinamide-induced apoptosis in cervical carcinoma cells. J Natl Cancer Inst 1997;89:1191–8.

[4] Dmitrovsky E. *N*-(4-hydroxyphenyl)retinamide activation of a distinct pathway signaling apoptosis. [editorial] J Natl Cancer Inst 1997; 89:1179–81.

[5] Oridate N, Lotan D, Mitchell MF, Hong WK, Lotan R. Inhibition of proliferation and induction of apoptosis in cervical carcinoma cells by retinoids: implications for chemoprevention. J Cell Biochem Suppl 1995;23:80–6.

[6] Lotan R, Trizna Z, Zou CC, Lotan D, Mitchell MF, Zou CP. Additive inhibitory effects of retinoids and interferon-α on the growth of human cervical carcinoma cells: lack of association with HPV status or retinoid receptor expression. Int J Oncol 1996;8:5–13.

[7] Kelloff GJ, Boone CW, Crowell JA, Steele VE, Lubet RA, Doody LA, et al. New agents for cancer chemoprevention. J Cell Biochem 1996;26S:1–28.

[8] Veronesi U, De Palo G, Marubini E, Costa A, Formelli F, Mariani L, et al. Randomized trial of fenretinide to prevent second breast malignancy in women with early breast cancer. J Natl Cancer Inst 1999; 91:1847–56.

[9] Kurie JM, Lee JS, Khuri FR, Mao L, Morice JC, Lee JJ, et al. *N*-(hydroxyphenyl)retinamide in the chemoprevention of squamous metaplasia and dysplasia of bronchial epithelium. Clin Cancer Res 2000;6:2973–9.

[10] Kim J, Shin DM, El-Naggar A, Lee JS, Corrales C, Lippman SM, et al. Chromosome polysomy and histological characteristics in oral premalignant lesions. Cancer Epidemiol Biomarkers Prev 2001; 10:319–25.

[11] Follen M, Atkinson EN, Malpica A, West L, Hittelman WN, Lotan R, et al. A randomized clinical trial of 4-hydroxyphenylretinamide for high-grade squamous intraepithelial lesions of the cervix. Clin Cancer Res 2001;6:3356–65.

[12] Palcic B, Garner DM, MacAulay CE. Image cytometry and chemoprevention in cervical cancer. J Cell Biochem Suppl 1995;23:43–54.

[13] MacAulay C, Palcic B. Fractal texture features based on optical density surface area: use in image analysis of cervical cells. Anal Quant Cytol Histol 1990;12:394–8.

[14] Ihaka R, Gentleman R. R: a language for data analysis and graphics. J Comput Graph Stat 1996;5:299–314.

[15] Westfall P, Young S. Resampling based multiple testing. New York: Wiley; 1993.

[16] Breiman L, Friedman JH. Estimating optimal transformations for multiple regression and correlations (with discussion). J Am Stat Assoc 1985;80:580–619.

[17] Good P. Permutation tests: a practical guide to resampling methods for testing hypotheses. New York: Springer, 1994.

[18] Guillaud M, Cox D, Adler-Storthz K, Malpica A, Staerkel G, Matisic J, et al. Exploratory analysis of quantitative histopathology of cervical intraepithelial neoplasia: objectivity, reproducibility, malignancy-associated changes, and human papillomavirus. Cytometry [In press].

[19] Decensi A, Torrisi R, Polizzi A, Gesi R, Brezzo B, Rolando M, et al. Effect of the synthetic retinoid fenretinamide on dark adaptation and the ocular surface. J Natl Cancer Inst 1994;86:105–10.

[20] Mehta RG, Moon RC, Hawthorne M, Formelli F, Costa A. Distribution of fenretinide in the mammary gland of breast cancer patients. Eur J Cancer 1991;27:138–41.

[21] Mitchell MF, Hittelman WN, Hong WK, Lotan R, Schottenfeld D. The natural history of cervical intraepithelial neoplasia: an argument for intermediate endpoint biomarkers. Cancer Epidemiol Biomarkers Prev 1994;3:619–26.

[22] Follen M, Vlastos A, Meyskens Jr FL, Atkinson EN, Schottenfeld D. Why phase II trials in cervical chemoprevention are negative: what have we learned? Cancer Causes Control 2002;13:855–73.