Predicting Cancer Development in Oral Leukoplakia: Ten Years of Translational Research[1]

J. Jack Lee,[2] Waun Ki Hong, Walter N. Hittelman, Li Mao, Reuben Lotan, Dong M. Shin, Steven E. Benner, Xiao-Chun Xu, Jin S. Lee, Vali M. Papadimitrakopoulou, Clarissa Geyer, Cherie Perez, Jack W. Martin, Adel K. El-Naggar, and Scott. M. Lippman

Departments of Biostatistics [J. J. L.], Clinical Cancer Prevention [S. M. L., X. C. X., C. P.], Thoracic/Head and Neck Medical Oncology [W. K. H., L. M., R. L., D. M. S., J. S. L., V. M. P., C. G.], Head and Neck Surgery [J. W. M.], Clinical Investigation [W. N. H.], and Pathology [A. K. E.], University of Texas M. D. Anderson Cancer Center, Houston, Texas 77030, and Bristol-Myer Squibb, Wallingford, Connecticut 06492 [S. E. B.]

ABSTRACT

Our 10-year translational study of the oral premalignant lesion (OPL) model has advanced the basic understanding of carcinogenesis. Although retinoids have established activity in this model, a substantial percentage of our OPL patients progress to cancer, especially after treatment is stopped. On the basis of our 10-year OPL study, we have developed the first comprehensive tool for assessing cancer risk of OPL patients. This cancer risk assessment tool incorporates medical/demographic variables, epidemiological factors, and cellular and molecular biomarkers.

Between 1988 and 1991, 70 advanced OPL patients were enrolled in a chemoprevention trial of induction with high dose isotretinoin (1.5 mg/kg/day for 3 months) followed by 9 months of maintenance treatment with either low dose isotretinoin (0.5 mg/kg/day) or β-carotene (30 mg/d; total treatment duration, 1 year). We assessed the relationship between cancer risk factors and time to cancer development by means of exploratory data analysis, logrank test, Cox proportional hazard model, and recursive partitioning.

With a median follow-up of 7 years, 22 of our 70 patients (31.4%) developed cancers in the upper aerodigestive tract following treatment. The overall cancer incidence was 5.7% per year. The most predictive factors of cancer risk are OPL histology, cancer history, and three of the five biomarkers we assessed (chromosomal polysomy, p53 protein expression, and loss of heterozygosity at chromosome 3p or 9p). In the multivariable Cox model, histology ($P = 0.0003$) and the combined biomarker score of chromosomal polysomy, p53, and loss of heterozygosity ($P = 0.0008$) are the strongest predictors for cancer development. Retinoic acid receptor β and micronuclei were not associated with increased cancer risk.

We have demonstrated a successful strategy of comprehensive cancer risk assessment in OPL patients. Combining conventional medical/demographic variables and a panel of three biomarkers can identify high risk patients in our sample. This result will need to be validated by future studies. With the identification of high risk individuals, more efficient chemoprevention trials and molecular targeting studies can be designed.

INTRODUCTION

Oral leukoplakia is a premalignant lesion that long has been considered to confer increased risk for the development of oral cancer (1, 2). Although the etiology of oral leukoplakia is not fully understood, these lesions often are associated with carcinogenic exposures, such as from use of tobacco, alcohol, or, particularly in Southeast Asia, betel nut (usually chewed) (3, 4).

The level of risk for malignant transformation of leukoplakia is associated with lesion histology. The overall malignant transformation rates for dysplastic lesions range from 11 to 36%, depending on the length of follow-up (2, 5). A recent report showed that proliferative verrucous leukoplakia has a malignant transformation rate as high as 70.3% (mean follow-up of 11.6 years) (6).

On the basis of strong data showing retinoid activity in preventing cancers of the upper aerodigestive tract (7–9), in 1988 we launched a chemoprevention trial of an induction phase of 3-month high dose isotretinoin (13-*cis*-retinoic acid) followed by a 9-month maintenance phase of low dose isotretinoin or β-carotene in subjects with OPLs.[3] Seventy patients enrolled in the trial. The efficacy of high dose induction and low dose maintenance with isotretinoin in this trial has been reported (10, 11).

During the course of this translational trial, we prospec-

Received 9/10/99; revised 2/4/00; accepted 2/7/00.

The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked *advertisement* in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

[1] Supported in part by grants from the National Institutes of Health (CA94026, CA16672, CA52051, DE11906), funds from the M. D. Anderson Cancer Center's Tobacco Initiative Research Program, the Margaret and Ben Love Professorship (S. M. L.), and the Stanley S. Schor Visiting Scholar program of Merck Research Laboratories (J. J. L.). Dr. Waun Ki Hong is an American Cancer Society Clinical Research Professor.

[2] To whom requests for reprints should be addressed, at the Department of Biostatistics, Box 213, University of Texas M. D. Anderson Cancer Center, 1515 Holcombe Boulevard, Houston, TX 77030. Phone: (713) 794-4943; Fax: (713) 745-4940; E-mail: jjlee@mdanderson.org.

[3] The abbreviations used are: OPLs, oral premalignant lesions; RAR-β, retinoic acid receptor β; LOH, loss of heterozygosity; FOM, floor of mouth; CI, confidence interval; CP, chromosomal polysomy.

tively collected tissue samples for analysis of biomarkers to characterize the molecular/cellular biology of leukoplakia, to assess correlations between biomarker expression and short term response (p53, RAR-β, and micronuclei) to evaluate the value of these biomarkers for predicting long term outcome (RAR-β; LOH at 3p, 9p, and chromosome polysomy) (11–19).

With accumulating data and follow-up of this pivotal chemoprevention trial, the main objective of the present report is to provide comprehensive cancer risk assessment tools for patients with oral leukoplakia, taking into account all collected variables, including medical-demographic variables, epidemiological factors, and cellular-molecular biomarkers. Our goal is to construct risk models to facilitate assigning appropriate interventions based on OPL patients' specific cancer risk or disease process. We also examined whether the short term intervention had an impact on preventing or delaying cancer. By identifying high risk leukoplakia patients, more efficient, better targeted chemoprevention studies can be designed with fewer patients and/or shorter duration. Increased knowledge of biomarkers and the effect of chemopreventive agents in the oral carcinogenesis pathway can help us in designing better mechanism-based prevention studies as well.

As translational chemoprevention study advances, with more and more marker information (*e.g.*, from chip technology) being gathered from smaller and smaller tissue specimens (biopsies and brushings) (20), the field urgently needs systematic statistical approaches to be able to analyze and integrate the explosion in technology and biomarker data. These approaches are needed for chemopreventive biomarker analyses and must be able to incorporate the real life issue of missing or expended tissue samples or noninformative results (*e.g.*, with LOH in the present study). Our study illustrates the statistical complexity of translational data and provides an example of a systematic framework-model for analyzing them.

PATIENTS AND METHODS

Eligibility Criteria and Clinical Study Design. Patients with advanced OPLs (defined as a dysplastic lesion; extensive, symptomatic hyperplastic lesion; or hyperplastic lesion in a high risk oral site, such as the ventral-lateral tongue or FOM were eligible for the trial. Patients with prior cancer were eligible if they had been disease free for at least 2 years at the time of study entry.

The study consisted of two phases. In the first phase, eligible and consenting patients were treated with a high dose isotretinoin induction regimen (1.5 mg/kg/day) for 3 months. Patients with lesion progression during the induction phase were removed from study and offered alternative treatments. Patients with nonprogressing or responding lesions entered the second phase of the study, in which they were randomized to receive a 9-month maintenance therapy with either low dose isotretinoin (0.5 mg/kg/day) or β-carotene (30 mg/day). A more detailed discussion on patient eligibility and study design can be found in our prior report (10).

Biomarker Measurements. The study design called for analysis of five biomarkers: p53; RAR-β; CP; LOH; and micronuclei. Detailed descriptions of laboratory procedures for these biomarker assessments were reported in our previous papers (12–18). As specified in the protocol, biopsies were taken from patients' primary index OPLs during scheduled clinic visits at baseline, 3 months, and 12 months. The biomarker measurements included in the present study were performed on all available and evaluable tissue samples. Sample evaluability was affected by several factors (*e.g.*, dropout, loss to follow-up, or patient refusal). Because each biopsy tissue block can be cut into only 20–30 sections of 4 μm, the need for histological evaluation and planned and unplanned biomarker analyses had exhausted certain tissue samples over the years. Inevaluable samples typically resulted from tangential cut of specimen, which produced insufficient epithelium cells in the basal and/or parabasal layers for analysis. To our knowledge, there was no systematic cause behind the absence of certain data. The majority of available samples are considered evaluable for biomarker analysis.

Statistical Analysis. The primary end point of the study and analysis we report here is time to cancer development. The Kaplan-Meier estimate was computed to estimate the probability of cancer-free survival. The logrank and Cox proportional hazards model was applied to analyze the effect of single and multiple covariates in predicting cancer development (21). A composite score of biomarkers was formed to evaluate the prognostic effect of multiple biomarkers. Exploratory data analysis using event charts (22) and scatter plot matrix was used to assist in the visualization of the association of multiple covariates and in model building (23). Recursive partitioning using RPART with exponential scaling for survival data were applied to provide an alternative method for classifying patients according to their cancer risk (24). All reported *P*s are two sided.

RESULTS

Accrual, Follow-up, and Cancer Development

Seventy patients were enrolled in the trial from January 1988 to March 1991 at the University of Texas M. D. Anderson Cancer Center. Fifty-nine of these patients had improved or stable lesions after the 3-month induction therapy and were randomized to receive low dose isotretinoin ($n = 26$) or β-carotene ($n = 33$) maintenance. On completion of the study in 1992, patients continued to be followed off protocol on a voluntary basis. Systematic follow-up by telephone and by chart review were performed in August 1997 and in September 1998, respectively. This report includes all of the events with the last follow-up as of December 31, 1998. By December 1998, a total of 22 patients (31.4%) developed cancers in the upper aerodigestive tract. The median follow-up was 7 years (range, 0.2–10.6 years) with 75% of patients having at least 5 years of follow up. The distribution of the time to cancer ranged from 0.2 to 10.1 years, with a mean of 4.0 years and a median of 3.1 years among patients who developed new cancers. The overall upper aerodigestive tract cancer incidence rate was 5.7% per year. New cancer sites included the tongue-FOM (9), buccal mucosa (3), gingiva (3), palate (3), lip (2), larynx (1), and esophagus (1). All of them were invasive carcinomas except one with *in situ* squamous cell carcinoma in the tongue. One other patient developed cancer, an anal squamous cell carcinoma that was not within the cancer end point defined for our study. The overall cancer-free survival and its 95% CI are shown in Fig. 1*A*.

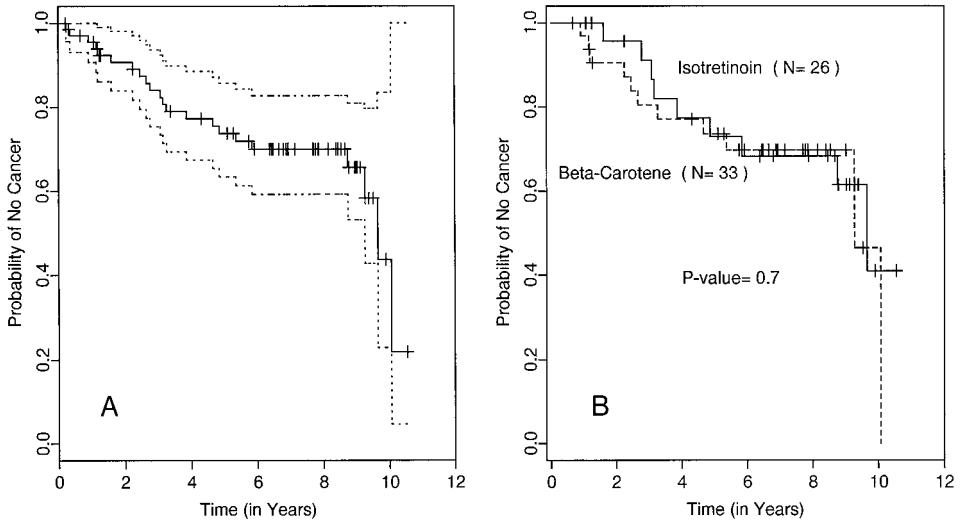

Fig. 1 *A*, cancer-free survival with 95% CI; *B*, cancer-free survival by treatment group.

Table 1 Analysis of cancer risk by patient characteristics, medical variables, and demographic variables using the Cox proportional hazards model

Variable	Category	Cancer	*n*	Risk ratio	*P*
Sex	Female	15	37	1.33 (0.84, 2.09)[a]	0.22
	Male	7	33		
Age (yr)	>60	16	35	1.61 (1.01, 2.58)	0.05
	≤60	6	35		
Smoking	Current	11	39	0.98 (0.64, 1.50)	0.93
	Former, never	11	31		
Alcohol	Current	16	49	1.27 (0.79, 2.04)	0.32
	Former, never	6	21		
Chewing tobacco	Yes	1	10	0.57 (0.21, 1.56)	0.27
	No	21	60		
Site	Tongue/FOM	11	27	1.14 (0.74, 1.75)	0.56
	Others	11	43		
Cancer history	Yes	7	11	1.89 (1.17, 3.05)	0.009
	No	15	59		
Histology	Mod./Severe Dysp.	6	9	2.30 (1.39, 3.81)	0.001
	Hyp./Mild Dysp.[b]	16	61		
Erythroplakia	Yes	11	30	1.27 (0.83, 1.95)	0.27
	No	11	40		
Treatment	Low-dose isotretinoin	9	26	1.00 (0.46, 2.18)[c]	0.99
	β-Carotene	11	33	1.06 (0.77, 1.46)[c]	0.74
	Induction only	2	11		
3-month response	Response	11	39	0.85 (0.55, 1.30)	0.45
	No response	10	26		
12-month response	Response	5	24	0.58 (0.35, 0.98)	0.04
	No response	13	27		

[a] Numbers in parentheses, 95% CI.
[b] Hyp., hyperplasia; Mild, Mod., Severe Dysp., mild, moderate, and severe dysplasia.
[c] Compared with the Induction Only group.

Patient Characteristics, Medical, and Epidemiological Variables as Predictors for Cancer

Table 1 summarizes patients' demographic, medical, and epidemiological variables and their prognostic value for cancer development. Older patients had a higher risk of developing cancer ($P = 0.05$), with a risk ratio of 1.61 and a 95% CI of (1.01, 2.58). New cancers developed in 7 of 11 patients (63.6%) with a prior history of cancer compared with only 15 of 59 patients (25.4%) with no prior cancer history ($P = 0.009$; Tables 1 and 2). The cancer risk of OPLs with moderate or severe dysplasia was 2.3 times higher than the cancer risk of OPLs with hyperplasia or mild dysplasia ($P = 0.001$). In our sample, however, smoking status, alcohol use, and chewing tobacco did not predict cancer risk.

Nine of 26 (34.6%), 11 of 33 (33.3%), and 2 of 11 (18.2%) patients developed cancer in the three treatment groups with low dose isotretinoin, β-carotene, and induction only, respectively. The relatively smaller percentage of cancer development in the

Table 2 Prior cancer sites, leukoplakia sites, and new cancer sites

Prior cancer site	Leukoplakia site	New cancer site	*n*
Tongue/FOM	Tongue/FOM	Tongue/FOM	1
Tongue/FOM	Tongue/FOM	Gingiva	1
Tongue/FOM	Buccal mucosa	Buccal mucosa	1
Tongue/FOM	Buccal mucosa	Gingiva	1
Tongue/FOM	Gingiva	Buccal mucosa	1
Palate	Buccal mucosa	Palate	1
Palate	Tongue/FOM	Palate	1
NA	Tongue/FOM	Tongue/FOM	7
NA	Tongue/FOM	Esophagus	1
NA	Buccal mucosa	Buccal mucosa	1
NA	Buccal mucosa	Lip	2
NA	Buccal mucosa	Palate	1
NA	Buccal mucosa	Tongue/FOM	1
NA	Gingiva	Gingiva	1
NA	Palate	Larynx	1

induction only group was associated with a shorter follow-up period. Three patients in the induction only group were lost to follow-up within 1 year of registration, compared with no patients in the low dose isotretinoin group and one in the β-carotene group. Fig. 1*B* shows the cancer-free survival of the low dose isotretinoin and β-carotene groups. Although the overall time to cancer was not statistically significant between the two groups ($P = 0.70$), it appeared that cancer onset was relatively delayed in the low dose isotretinoin between years 1 and 3. With a smoothing band width of 0.8 year, the estimated cancer rates at years 1, 2, and 3 were 2.5%, 5.7%, and 9.6%, respectively, for the low dose isotretinoin group and 6.2%, 7.1%, and 10.4%, respectively, for the β-carotene group. As a reminder, patients in this study received only 1 year of treatment after registration.

Clinical response at 12 months, but not at 3 months, was statistically significantly predictive of cancer. Subjects showing continued response to the maintenance therapy developed fewer cancers than did nonresponders (risk ratio, 0.58; $P = 0.04$). Note that the status of 3-month and 12-month clinical response was not available in 5 and 19 patients, respectively.

Prior Cancer Sites, Leukoplakia Sites, and New Cancer Sites

Table 2 lists the prior cancer site, primary leukoplakia site, and new cancer site for the 22 study patients who developed new cancer. Of the seven patients with prior cancer history, only one had a new cancer that developed in the same location as the prior cancer and leukoplakia (tongue-FOM). Fifteen patients with no prior cancer history developed cancer. In nine of these patients, new cancer occurred in the leukoplakia site. The remaining patients had cancers develop in new sites. Among the 22 patients who developed cancers in the upper aerodigestive tract, 13 (59%) had new cancer in the same site as prior cancer or leukoplakia. However, the remaining nine patients (41%) developed new cancer in new sites away from the prior lesions, suggesting that field cancerization may play an important role in cancer development in leukoplakia patients.

Combined Predictive Effect of Histology and Prior Cancer History on Cancer Development

Table 1 shows that histology and prior cancer history are the two most important clinical-epidemiological predictors for cancer. Fig. 2 provides a graphic assessment of their effect on time to cancer using the interval event chart (22). Starting from the bottom of the figure, six of nine patients with moderate or severe dysplasia developed cancer. One of these six patients had a prior cancer (solid circle) which occurred ~10 years before registration. The events of 59 patients with hyperplasia or mild dysplasia were plotted in the upper part of the figure. Six of 10 patients with prior cancer history developed new cancer. Although patients with prior cancer were more likely to develop new cancers, it appears that time since the prior cancer did not correlate with time to new cancer. Furthermore, the relatively long time from prior cancer to new cancer diagnosis suggests that the new cancer is unlikely the result of locoregional recurrence (range, 6.0–26.6 years; median, 9.2 years; $n = 7$).

Fig. 2 shows that six patients developed cancer within 5 years but had no prior cancer or moderate/severe dysplasia. On the other hand, one patient with moderate/severe dysplasia and three patients with a history of prior cancer did not develop cancer with >5 years of follow-up. These findings indicate that prior cancer history and histology combined could not completely explain the cancer risk. The results call for a need to investigate the predictive effect of the molecular and cellular biomarkers to enhance the accuracy of cancer risk assessment.

Biomarkers as Predictors for Cancer, One Covariate Case

CP. Baseline tissue samples were available from 40 patients for the analysis. Due to the skewed distribution (range, 0.2–39.5; mean, 6.6; median, 3.0), we chose 3 as the cutoff for dichotomizing the CP into the low and high groups. Table 3 shows that 13 of 20 (65%) patients with high polysomy developed cancer whereas only 5 of 20 (25%) patients had cancer in the low polysomy group. The risk ratio was 1.85 with a 95% CI of 1.05–3.25 ($P = 0.03$).

p53 Protein. Because there was no p53 modulation during treatment and worsening histology associated with higher p53 accumulation in the parabasal layer (12), we used the baseline p53 status in the parabasal layer as the marker for predicting cancer risk. In addition, we have chosen 0.2 as a cutoff for dichotomizing the p53 labeling index (12, 15). Among the 52 evaluable patients, 6 of 10 (60%) in the high p53 group developed cancer whereas only 10 of 42 patients (24%) had cancer in the low p53 group ($P = 0.11$; Table 3). Analysis based on the principal component analysis on multiple p53 measurements gave similar results (data not shown).

LOH at 3p or 9p. In 37 evaluable patients, 8 of 19 (42%) patients with 3p or 9p LOH and 2 of 18 (11%) patients without LOH developed cancer. Our initial report showed that 3p or 9p LOH was associated with higher cancer risk (17). However, with longer follow-up, this significant association was somewhat weakened ($P = 0.09$; Table 3).

RAR-β Expression. RAR-β expression was measured at baseline, 3 months, and 12 months after registration. Due to the change of RAR-β status (up-regulated to 90% expression at 3 months) (13) in our retinoid-based trial, we used the

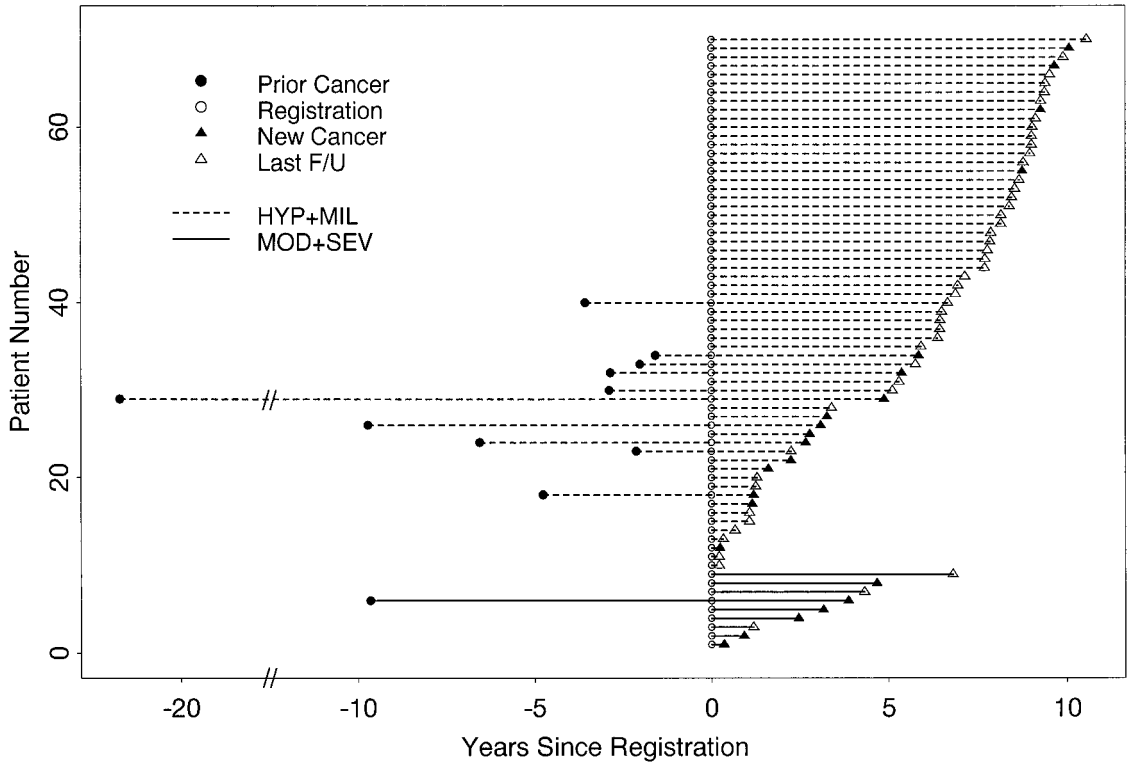

Fig. 2 Interval event chart in years since registration. Patients were sorted by histology first, and then by the time from registration to cancer or last follow-up (*F/U*). *MOD+SEV*, moderate plus severe dysplasia; *HYP+MIL*, hyperplasia plus mild dysplasia.

Table 3 Analysis of cancer risk by cellular and molecular biomarkers using the Cox proportional hazards model, one covariate

Variable	Category	Cancer	*n*	Risk ratio	*P*
CP	>3%	13	20	1.85 (1.05, 3.25)[a]	0.03
	≤3%	5	20		
p53 in parabasal layer	>0.2	6	10	1.53 (0.91, 2.59)	0.11
	≤0.2	10	42		
LOH in 3p or 9p	Yes	8	19	1.94 (0.89, 4.23)	0.09
	No	2	18		
Last measured RAR-β	Loss	3	10	0.91 (0.26, 3.21)	0.88
	Present	15	46		
Last measured micronuclei	>2	6	18	0.91 (0.56, 1.47)	0.70
	≤2	16	45		

[a] Numbers in parentheses, 95% CI.

last measured RAR-β status as the covariate for modeling cancer development. Consistent with our prior report (19), we found that RAR-β expression measured at the last follow-up in the trial was not a predictor for long term cancer risk ($P = 0.88$; Table 3).

Micronuclei. Similar to the RAR-β expression, micronuclei were also measured at baseline, 3 months, and 12 months after registration. Because the baseline micronuclei were reduced after the induction of isotretinoin treatment (16), we have chosen the last measured micronuclei as the predictor for cancer. Table 3 showed that there was no difference in cancer risk between the low and high micronuclei groups ($P = 0.70$).

Biomarkers as Predictors for Cancer, Multiple Covariates Case

Fig. 3 shows the scatter plot matrix of the follow-up time (time to cancer or lost to follow-up), cancer status, CP, p53, and LOH. The scatter plot matrix revealed that there was a weak to moderate correlation between CP and p53 expression (Pearson's $r = 0.61$, Spearman's $\rho = 0.23$) but no correlation between polysomy and LOH (Spearman's $\rho = -0.07$) or p53 and LOH (Spearman's $\rho = 0.07$). The second row of the plots indicates that patients with cancer were inclined to have higher CP, higher p53, and LOH, consistent with the result presented in Table 3. Fig. 3 also shows that patients having two or three of the high

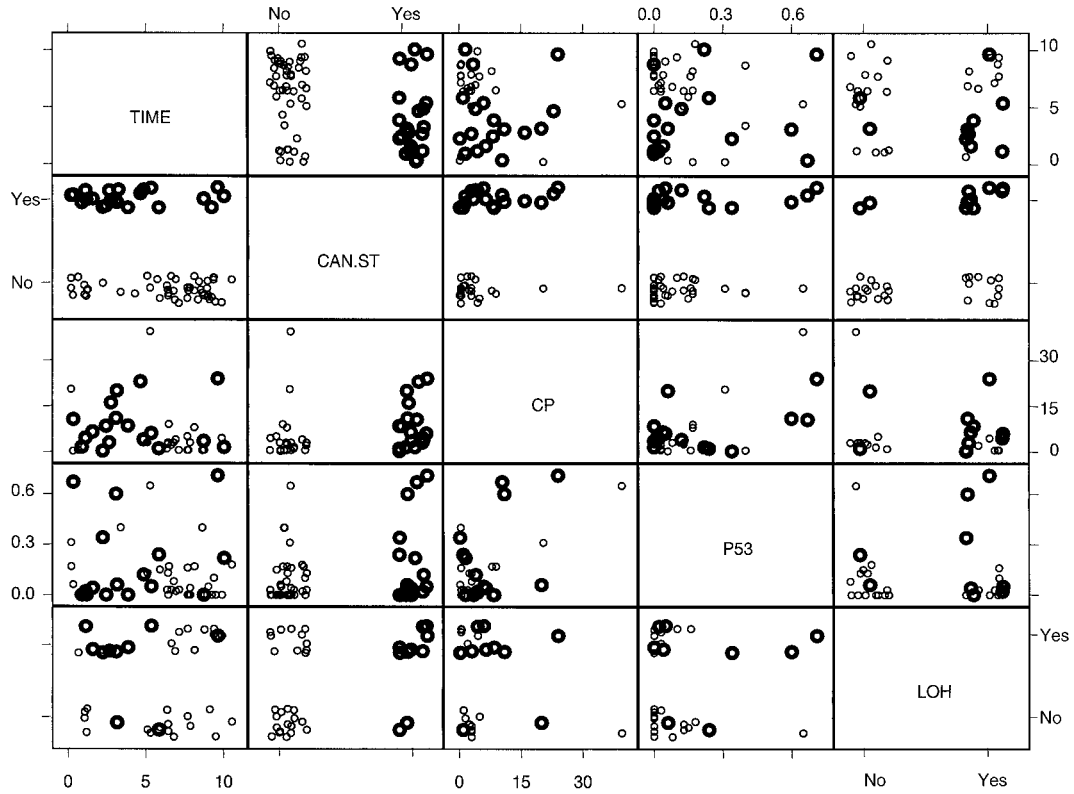

Fig. 3 Scatter plot matrix of time, cancer status (*CAN.ST*), CP, p53, and LOH at 3p or 9p. Each panel was one scatter plot of a pair of variables corresponding to the labels on the *x*-axis and *y*-axis. Complete data were available for time and cancer status ($n = 70$). $n = 40$, 52, 37, 37, 25, and 30 for CP, p53, LOH, CP × p53, CP × LOH, and p53 × LOH, respectively. Small random variations were added to the values of the discrete variables, *i.e.*, cancer status and LOH to break the ties for better presentation of the data. Patients who developed cancer were marked by *larger, darker circles,* and patients without cancer were marked by *smaller, lighter circles.*

risk factors were more likely to develop cancer. (Note: Better visualization can be achieved by "brushing" or highlighting points with certain features interactively on a computer monitor.) Because of the limited sample size and missing biomarker values, however, only 24 patients had complete data on these 3 variables, restricting the use of conventional regression analysis with multiple covariates. To overcome this limitation, we devised a combined score, denoted as CP.p53.LOH, to capture the collective information contained in these three markers. CP.p53.LOH was computed as the sum of the three indicators, one for each of the three biomarkers. Each indicator was assigned a value of either 0 or 1, denoting either low or high risk marker value, respectively. If a biomarker value was missing, the indicator was set as 0. Nine patients with missing CP, p53, and LOH information were removed from the analysis. With this combined score, CP.p53.LOH has a risk ratio of 2.27 (95% CI 1.41–3.66) and was highly significant in predicting cancer ($P = 0.0008$; Table 4).

Table 4 Analysis of cancer risk by combined biomarker score and patient characteristics using the Cox proportional hazards model, multiple covariate

Model	*n*	Variable	Risk ratio	*P*
1	61	CP.p53.LOH	2.27 (1.41, 3.66)[a]	0.0008
2	61	CP.p53.LOH	2.02 (1.21, 3.35)	0.007
		Cancer history	2.90 (1.02, 8.22)	0.046
3	61	CP.p53.LOH	2.41 (1.44, 4.03)	0.0008
		Histology	2.68 (1.57, 4.59)	0.0003
4	61	CP.p53.LOH	2.07 (1.17, 3.63)	0.012
		Histology	2.49 (1.45, 4.28)	0.001
		Cancer history	2.38 (0.78, 7.22)	0.13

[a] Numbers in parentheses, 95% CI.

Combined Biomarker Score and Patient Characteristics as Predictors for Cancer, Multiple Covariates Case

When the combined biomarker score and cancer history were entered in the Cox regression analysis, model 2 in Table 4 showed that CP.p53.LOH remained highly significant in predicting cancer. Cancer history remained significant ($P = 0.046$) for cancer risk but was not as significant as it was when used as a single predictor ($P = 0.009$; Table 1). The result indicated that after incorporating the biomarker information, the predictive value of cancer history was not as important as before. In other words, a major part of the information contained in cancer history for predicting new cancer was associated with high CP, high p53 expression, and

LOH. Model 3 in Table 4 shows the joint predictive effect of CP.p53.LOH and histology. Both variables remained highly significant (CP.p53.LOH, $P = 0.0008$; histology, $P = 0.0003$), suggesting that the histology and the combined biomarker score jointly were important in predicting cancer. Fig. 4 plots the probability of cancer incidence by histology and CP.p53.LOH, showing that among patients with hyperplasia-mild dysplasia lesions, the cancer incidence rates were 1 of 22 (4.5%), 6 of 22 (27.3%), 4 of 7 (57.1%), and 2 of 2 (100%) for the combined biomarker score of 0, 1, 2, and 3, respectively. The cancer incidence rates for patients with moderate-severe dysplasia were 1 of 3 (33.3%) if the combined biomarker score was 0 and 5 of 5 (100%) if patients had at least 1 biomarker in the high risk range. The same trend held in 24 patients with complete biomarker information (data not shown). Although the sample size in several subgroups was small, the trend shown in Fig. 4 illustrates the value of using both the histology and biomarker information in modeling the cancer risk.

When CP.p53.LOH, histology, and cancer history were entered in the Cox regression, model 4 indicated that the cancer history no longer was significant ($P = 0.13$) but CP.p53.LOH and histology remained significant (Table 4). The likelihood ratio test showed that model 4 did not provide significant improvement over model 3 by adding prior cancer history ($P = 0.13$). Therefore, the best model for predicting cancer risk in our data contains the combined biomarker score and histology presented in model 3.

Sensitivity Analysis on Patients with Missing Biomarker Information

We also performed the sensitivity analysis by assigning missing biomarkers a value of 0.5. The results are consistent with models 1–4 in Table 4. Specifically, if 0.5 was assigned to missing biomarkers, the combined biomarker score was still significant in predicting cancer development as a single covariate ($P = 0.002$) and along with histology (combined biomarker score, $P = 0.002$; histology, $P = 0.001$).

Recursive Partitioning for the Classification of Cancer Risk, Multiple Covariates Case

In the above analyses, we determined that histology, history of cancer, CP, p53, and LOH are important predictors for cancer development. We have also applied recursive partitioning to construct an alternative classification model for cancer risk using these five covariates (Fig. 5). Except for p53, four of the five covariates were chosen in the model. The classification tree started with node 1 at the top where 22 of 70 total patients developed cancer. The standardized event rate, which is a special case of the event rate in the Poisson model for censored data with exponential scaling, was set to 1 for the entire sample in node 1. The recursive partitioning model chose histology for the first split. Patients with hyperplasia or mild dysplasia were placed in node 2 where 16 of 61 patients had cancer. The remaining patients with moderate or severe dysplasia formed node 3 where 6 of 9 patients had cancer. The standardized event rates were 0.8 and 2.6 in nodes 2 and 3, respectively. Patients in node 2 were further split into two groups according to their prior

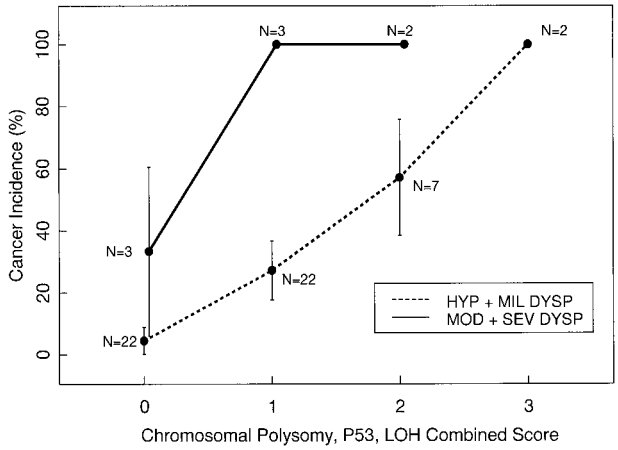

Fig. 4 Cancer incidence rate (±1 SE) by the composite score of CP, p53, and LOH at 3p or 9p. *HYP,* hyperplasia; *MIL,* mild; *MOD,* moderate; *SEV,* severe; *DYSP,* dysplasia.

cancer history to nodes 4 and 5. For patients with hyperplasia or mild dysplasia and no history of cancer (node 4), LOH was the next variable chosen for classification. Finally, CP with a cutoff value of 4.25 was selected to split patients in node 7 to nodes 8 and 9. Nodes 6, 8, 9, 5, and 3 are terminal nodes with increasing standardized event rates. The observed cancer incidence rates in these groups were 5.3, 13.3, 33.3, 60, and 66.7%, respectively. Note that eight patients were not shown in the terminal nodes due to missing biomarker information. The recursive partitioning algorithm uses all available information at each split. Therefore, variables with higher predictive power and less missing values are more likely to be chosen in earlier steps. The results show that patients with either moderate/severe dysplasia or history of cancer had a high risk for new cancer. In the remaining patients, *i.e.,* patients with hyperplasia-mild dysplasia and no histology of cancer, LOH and CP can provide additional information to classify patients according to their cancer risk.

DISCUSSION

The ability to identify oral leukoplakia patients at increased risk of cancer development is critical for improving control of oral cancer. Once identified, the highest risk individuals could be offered more aggressive treatment options and more intensive follow-up. In our study, 31.4% of leukoplakia patients developed cancer during a 10-year follow-up period. This result is consistent with reports of malignant transformation in OPLs (5, 6). Time to cancer development in our study ranged from 2.8 months to 10.1 years. Consistent with field carcinogenesis, 41% of new cancers developed in sites distinct from prior sites of cancer or leukoplakia. Therefore, leukoplakia is more than just a premalignant lesion but also is a marker for increased cancer risk throughout the upper aerodigestive tract.

After 10 years of translational assessments, we have identified clinical and molecular-cellular factors that appear to indicate leukoplakia patients at higher risk of malignant transformation. Although smoking and alcohol consumption are risk factors for OPLs, they were not associated with increased cancer risk in our sample. In our patients, the factors of prior cancer

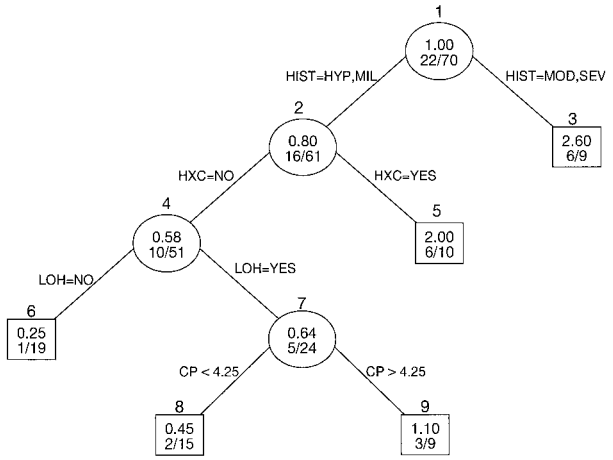

Fig. 5 Recursive partitioning tree for time to cancer development. *HIST*, histology; *HYP*, hyperplasia; *MIL, MOD, SEV*, mild, moderate, or severe dysplasia; *HXC*, history of cancer; *LOH*, LOH at 3p or 9p.

history and/or moderate-to-severe dysplastic OPLs apparently were associated with higher risk of cancer development, results agreeing with medical knowledge common in the field and many published reports (2, 5, 6). These cancer history and pathological risk factors, however, are not very sensitive cancer predictors. Therefore, we assessed the predictive effect of five cellular or molecular biomarkers that may complement the medical history and pathological factors and provide a more accurate risk profile.

Our OPL data indicate that expressions within three of these five biomarkers were associated with higher cancer risk: high CP; high p53 protein accumulation in the parabasal layer; and LOH at 3p or 9p. The combined score of these three markers was more predictive of cancer risk than was the independent score of any single marker. Moreover, combining both phenotypic histopathology and genotypic biomarker information can provide better power in predicting cancer development. Therefore, the accuracy of cancer risk assessment apparently can be improved by using collective information derived from a panel of biomarkers combined with histological evaluation of the OPL. Histology and combined biomarker score also were independent predictors for cancer risk. The more abnormalities found in either histology or relevant biomarker expressions, the higher was the cancer risk. The same conclusions were substantiated by the multivariable Cox regression analysis and the recursive partitioning method. These results support the concept of multistep carcinogenesis and warrant further investigation to understand the underlying process.

Our biomarker model analysis led to a "statistically" diagnosed cancer in one of our patients (a 65-year-old male former smoker and current alcohol drinker). He was generally low risk [hyperplastic histology (tongue OPL) and no prior cancer history]. Nevertheless, our risk model indicated increased cancer risk (CP of 24%, parabasal p53 labeling index of 0.71, and LOH at both 3p and 9p). At the scheduled phone contact (September 1998), the patient (still asymptomatic) agreed to come in for an unscheduled clinic visit to have a biopsy-pathological procedure (December 1998), which revealed an invasive squamous cell carcinoma in the tongue. This cancer diagnosis (9.7 years after trial enrollment) resulted exclusively from our predictive model finding. The long delay of malignant transformation in some leukoplakia patients provides a window for treatment intervention, such as with chemoprevention. To avoid the possible bias introduced by the unscheduled follow-up clinic visit of this particular patient, we also considered that the patient was cancer free in September 1998 (the date of our systematic/scheduled phone contact/follow-up) and redid the analysis. The *P* values in Table 4 changed slightly, but all of the major findings remained the same.

In our study, the clinical study design allowed collection of tissue samples only during the 1-year treatment period. There were no prespecified follow-up visits or scheduled biopsies during the follow-up period. Therefore, we were limited in assessing longer term biomarker changes and in mapping out the carcinogenesis pathway. This limitation may explain in part why RAR-β and micronuclei expressions did not help predict cancer risk. Possibly, loss of RAR-β or increase in micronuclei precedes cancer development, but our study did not provide enough follow-up information to support or refute either possibility. Future chemoprevention trial designs should include scheduled follow-ups and biopsies during and after the treatment period. This will allow uniform follow-up and assessments of all patients to document patterns of biomarker changes over time. Long term follow-up of chemoprevention study patients will be essential to gain a full understanding of multistep carcinogenesis.

Increasing our ability to predict cancer development is prerequisite to the next step of cancer control, *i.e.*, to develop more effective and/or longer-term chemopreventive interventions within the carcinogenic pathway of individuals at increased cancer risk. Our molecular-cellular risk modeling approach dovetails with the most exciting new developments in chemoprevention, which involve molecular targeting approaches of agent development. Molecular targeting study is advancing rapidly in chemoprevention in general (25, 26) and in OPLs specifically, illustrated, *e.g.*, by research targeting p53 (27, 28).

This report does not attempt to present a validated, generalizable, definitive risk assessment model. Rather, it attempts to illustrate the utility of various statistical modeling approaches for gaining insight into cancer development under the constraints of the design, patient population, missing data, and so on, of a study. Given these limitations, the data set (collected by an experienced group of investigators from a prospective National Cancer Institute randomized trial) for our risk modeling still represents the most comprehensive and mature (10 years of collection) translational data of which we are aware. The use of event chart, scatter plot matrix, and recursive partitioning, along with the Cox model, illustrates rational steps for modeling the risks associated with cancer development. We recognize, however, that this relatively comprehensive data set still is very small, and so the present findings with respect to specifically predicting head and neck cancer must be validated by future studies. Notwithstanding these caveats, we believe that the current methodological and analytical approach contributes substantially to the future development of predictors of cancer, not

only in the head and neck, but in other sites as well. Our statistical modeling approach also can help address the growing translational chemoprevention problem of analyzing ever more biomarkers and biomarker data gathered (via chip technology, and the like) from much smaller tissue specimens (20, 29).

In summary, our report presents a successful strategy of comprehensive cancer risk assessment in OPL patients under the constraints of a translational chemoprevention trial. We demonstrated that both medical-demographic variables and a panel of three biomarkers can identify high risk patients. Our biomarker risk modeling approach currently is being tested for validation in an ongoing large scale long term National Cancer Institute-sponsored chemoprevention trial in patients with OPLs. Use of biomarkers to increase the sensitivity of histology in predicting cancer development will help in identifying high risk patients, including those with lower risk histology. With quantified cancer risk assessments, investigators and clinicians can offer the most appropriate, tailored cancer prevention strategies to each individual.

REFERENCES

1. Kramer, I. R., Lucas, R. B., Pindborg, J. J., and Sobin, L. H. Definition of leukoplakia and related lesions: an aid to studies on oral precancer. Oral Surg. Oral Med. Oral Pathol., *46:* 518–539, 1978.

2. Silverman, S. Jr. (ed.). Oral Cancer, Ed. 3. Atlanta, GA: American Cancer Society, 1990.

3. Slaughter, D. P., Southwick, H. W., and Smejkal, W. Field cancerization in oral stratified squamous epithelium: clinical implications of multicentric origin. Cancer (Phila.), *6:* 963–968, 1953.

4. Norton S. A. Betel: consumption and consequences. J. Am. Acad. Dermatol., *38:* 81–88, 1998.

5. Silverman, S., Jr., Gorsky, M., and Lozada, F. Oral leukoplakia and malignant transformation: a follow-up study of 257 patients. Cancer (Phila.), *53:* 563–568, 1984.

6. Silverman, S., Jr., and Gorsky, M. Proliferative verrucous leukoplakia: a follow-up study of 54 cases. Oral Surg. Oral Med. Oral Pathol. Oral Radiol. Endod., *84:* 154–157, 1997.

7. Hong, W. K., Endicott, J., Itri, L. M., Doos, W., Batsakis, J. G., Bell, R., Fofonoff, S., Byers, R., Atkinson, E. N., Vaughan, C., *et al.* 13-*cis*-Retinoic acid in the treatment of oral leukoplakia. N. Engl. J. Med., *315:* 1501–1505, 1986.

8. Hong, W. K., Lippman, S. M., Itri, L. M., *et al.* Prevention of second primary tumors with isotretinoin in squamous cell carcinoma of the head and neck. N. Engl. J. Med., *323:* 795–801, 1990.

9. Hong, W. K., Lippman, S. M., Hittelman, W. N., and Lotan, R. Retinoid chemoprevention of aerodigestive cancer: from basic research to the clinic. Clin. Cancer Res., *1:* 677–686, 1995.

10. Lippman, S. M., Batsakis, J. G., Toth, B. B., Weber, R. S., Lee, J. J., Martin, J. W., Hays, G. L., Goepfert, H., and Hong,W. K. Comparison of low-dose isotretinoin with beta carotene to prevent oral carcinogenesis. N. Engl. J. Med., *328:* 15–20, 1993.

11. Papadimitrakopoulou, V. A., Hong, W. K., Lee, J. S., Martin, J. W., Lee, J. J., Batsakis, J. G., and Lippman, S. M. Low-dose isotretinoin *versus* β-carotene to prevent oral carcinogenesis: long-term follow-up. J. Natl Cancer Inst., *89:* 257–258, 1997.

12. Lippman, S. M., Shin, D. M., Lee, J. J., Batsakis, J. G., Lotan, R., Tainsky, M. A., Hittelman,W. N., and Hong, W. K. p53 and retinoid chemoprevention of oral carcinogenesis. Cancer Res., *55:* 16–19, 1995.

13. Lotan, R., Xu, X. C., Lippman, S. M., Ro, J. Y., Lee, J. S., Lee, J. J., and Hong, W. K. Suppression of retinoic acid receptor-β in premalignant oral lesions and its up-regulation by isotretinoin. N. Engl. J. Med., *332:* 1405–1410, 1995.

14. Xu, X. C., Zile, M. H., Lippman, S. M., Lee, J. S., Lee, J. J., Hong, W. K., and Lotan, R. Anti-retinoic acid (RA) antibody binding to human premalignant oral lesions, which occurs less frequently than binding to normal tissue, increases after 13-*cis*-RA treatment in vivo and is related to RA receptor β expression. Cancer Res., *55:* 5507–5511, 1995.

15. Shin, D. M., Xu, X. C., Lippman, S. M., Lee, J. J., Lee, J. S., Batsakis, J. G., Ro, J. Y., Martin, J. W., Hittelman, W. N., Lotan, R., and Hong, W. K. Accumulation of p53 protein and retinoic acid receptor β in retinoid chemoprevention. Clin. Cancer Res., *3:* 875–880, 1997.

16. Benner, S. E., Lippman, S. M., Wargovich, M. J., Lee, J. J., Velasco, M., Martin, J. W., Toth, B. B., and Hong, W. K. Micronuclei, a biomarker for chemoprevention trials: results of a randomized study in oral pre-malignancy. Int. J. Cancer, *59:* 457–459, 1994.

17. Mao, L., Lee, J. S., Fan, Y. U., Ro, Batsakis, J. G., Lipman, S., Hittelman, W., and Hong, W. K. Frequent microsatellite alterations at chromosome 9p21 and 3p14 in oral premalignant lesions and their value in cancer risk assessment. Nat. Med., *2:* 682–685, 1996.

18. Hittelman, W. N., Kim, H. J., Lee, J. S., Shin, D. M., Lippman, S. M., Kim, J., Ro, J. Y., *et al.* Detection of chromosome instability of tissue fields at risk: *in situ* hybridization. J. Cell. Biochem. Suppl., *25S:* 57–62, 1996.

19. Geyer, C., Papadimitrakopoulou, V., Lee, J. S., Lippman, S. M., Xu, X. C., Lee, J. J., Hong, W. K., and Lotan, R. A lack of association among the basal level of retinoic acid receptor β mRNA in oral premalignant lesions, its up-regulation by 13-*cis*-retinoic acid and cancer development during a long-term follow-up. Proc. Am. Soc. Clin. Oncol., *17:* 391, 1998.

20. Rosin, M. P., Epstein, J. B., Berean, K., Durham, S., Hay, J., Cheng, X., Zeng, T., Huang, Y., and Zhang, L. The use of exfoliative cell samples to map clonal genetic alterations in the oral epithelium of high-risk patients. Cancer Res., *57:* 5258–5260, 1997.

21. Cox, D. R., and Oakes, D. Analysis of Survival Data. New York: Chapman and Hall, 1984.

22. Lee, J. J., Hess, K. R., and Dubin, J. A. Extensions and applications of event charts. Am. Statistician, *54:* 63–70, 2000.

23. S-PLUS User's Guide. Seattle, WA: Data Analysis Products Division. MathSoft, Inc., 1997.

24. Therneau, T. M., and Atkinson, E. J. An introduction to recursive partitioning using the RPART routines. Mayo Clinic Biostatistics Technical Report 61, Rochester, MN, 1997.

25. Hong, W. K., and Sporn, M. B. Recent advances in chemoprevention of cancer. Science (Washington DC), *278:* 1073–1077, 1997.

26. Lippman, S. M., Lee, J. J., and Sabichi, A. L. Cancer chemoprevention: progress and promise. J. Natl. Cancer Inst., *90:* 1514–1528, 1998.

27. Rudin, C. M., Recant, W., Bennett, P., Sulzen, L., Kirn, D. H., and Vokes, E. E. Preliminary report: adenovirus ONYX-015 administered by mouthwash as a chemopreventive agent and for the treatment of oral dysplastic lesions. Proc. Am. Soc. Clin. Oncol., *18:* 445, 1999.

28. Clayman, G. L., El-Naggar, A. K., Lippman, S. M., Henderson, Y. C., Frederick, M., Merritt, J. A., Zumstein, L. A., Timmons, T. M., Liu, T. W., Ginsberg, L., Roth, J. A., Hong, W. K., Bruso, P., and Goepfert, H. Adenovirus-mediated p53 gene transfer in patients with advanced recurrent head and neck squamous cell carcinoma. J. Clin. Oncol., *16:* 2221–2232, 1998.

29. Mao L. Can molecular assessment improve classification of head and neck premalignancy? Clin. Cancer Res., *6:* 321–322, 2000.