Abstract Number: 17

Risk Stratification System for Oral Cancer Screening

Lutécia H Mateus Pereira[1], Isildinha M Reis[2,3], Erika Reategui[4], Claudia Gordon[2], Sandra Saint-Victor[4], Robert Duncan[3], Carmen Gomez[5], Stephanie Bayers[6],Penelope Fisher[4], Aymee Perez[2], W Jarrard Goodwin[2,4], Jennifer J Hu[2,3], Elizabeth Franzmann[2,4]

[1]Miami-Dade College, Miami, FL, USA; [2]Sylvester Comprehensive Cancer Center, University of Miami Leonard Miller School of Medicine, Miami, FL, USA; [3]Department of Public Health Sciences, University of Miami Leonard Miller School of Medicine, Miami, FL; [4]Department of Otolaryngology, University of Miami Leonard Miller School of Medicine, Miami, FL, USA; [5]Department of Pathology, University of Miami Leonard Miller School of Medicine, Miami, FL, USA; [6]MacNeal Hospital, Berwyn, IL, USA

Background

- Each year 50,000 individuals in the United States and 500,000 worldwide are diagnosed with head and neck squamous cell carcinoma (HNSCC).
- The main risk factors are tobacco and alcohol use and human papillomavirus (HPV) infection.
- Early detection tests are needed because the majority of patients present in late stage when cure rates reach only 40%.
- Salivary soluble CD44 (solCD44), a tumor initiating marker and total protein may be useful diagnostic markers for HNSCC.
- Since minority patients and those of low socioeconomic status (SES) suffer disproportionately from this disease, such a test should be simple, noninvasive, and inexpensive so that patients with limited access to expert diagnosis might also have access to screening.

Methodology

- A case-control, hospital-based design with 150 oral cavity (OC) and oropharyngeal (OP) cases and 150 frequency-matched controls was used to determine whether soluble CD44 (CD44) and total protein levels in oral rinses were associated with oral cancer independent of age, gender, race, ethnicity, tobacco and alcohol use, and socioeconomic status (SES).
- The performance of the marker panel developed using data from the hospital-based case-control study, was evaluated in 150 participants from a target screening community previously determined to be at elevated risk for oral cancer due to poverty and smoking.
- Levels of solCD44 (normal and variant isoforms) were measured using a sandwich ELISA assay (eBioscience), with previously published modifications (1-3). We performed the DC protein assay (Bio-Rad Laboratories) according to the manufacturer's protocol using saliva samples prepared as previously published. Each sample was tested in duplicate and the technician was blinded to disease status.
- Patient groups were compared with respect to the distribution of potentially important categorical covariates. Logistic regression analysis and multivariate recursive partitioning was used to assess the association between markers and the risk for oral cancer. Odds ratio (OR) estimates were reported with corresponding 95% confidence interval (95%CI) and area under the curve (AUC) of the receiver operating characteristic curve (ROC) for fitted models. Kaplan-Meier and Cox regression models were used to evaluate progression-free survival (PFS) and overall survival (OS). Hazard ratio (HR) estimates and corresponding 95%CI are reported (SAS version 9.2 and R package).

Figure 1
Flow Diagram for Case-Control Study

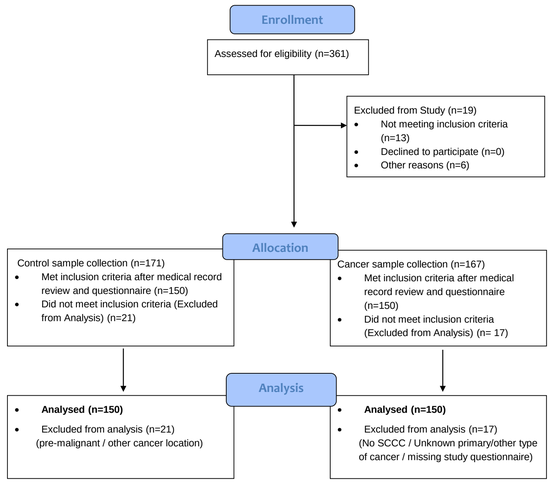


Table 1
Demographics and other characteristics of study subjects

Variable / Category	Hospital-based Cases (n=150) N	%	Hospital-based Controls (n=150) N	%	*P*	Community Controls (n=150) N	%	P (3 groups)
Site of enrollment								
JMH	80	53.3	71	47.3	0.299	NA	NA	NA
UM	70	46.7	79	52.7		NA	NA	
Liberty City	NA	NA	NA	NA		150	100.0	
Age, years								
<40	4	2.7	–	–	0.214	–	–	<.0001
40 |- <50	20	13.3	29	19.3		62	41.3	
50 |- <60	60	40.0	56	37.3		74	49.3	
60 |- <70	44	29.3	44	29.3		14	9.3	
≥70	22	14.4	21	14.0		–	–	
<60	84	66.0	85	56.7	0.449	136	90.7	<.0001
≥60	86	44.0	65	43.3		14	9.3	
Mean (SD)	58.6 (10.5)		58.5 (9.7)		0.887	51.2 (5.7)		
Median (range)	58 (28 - 88)		58.5 (40– 87)			50 (40 - 69)		<.0001
Gender								
Male	121	80.7	118	78.7	0.907	88	58.7	<.0001
Female	29	19.3	32	21.3		62	41.3	
Race								
White	123	82.6	118	79.7	0.534	–	–	<.0001
Black	26	17.4	30	20.3		150	100.0	
Asian/Other/Missing (1 case Other, 1 control Asian, and 1 control missing)	1	–	2	–		–	–	
Ethnicity								
Hispanic	77	51.3	93	62.0	0.062	1	0.7	<.0001
Non-Hispanic	73	48.7	57	38.0		149	99.3	
Education								
≤ Grade 12 or GED	77	52.0	57	38.0	0.015	115	76.7	<.0001
Some college or college graduate	71	48.0	93	62.0		35	23.3	
Refused/ Missing	2	–	–	–		–	–	
Employment								
Out-of/unable-to work	61	40.9	45	30.0	0.048	106	70.7	<.0001
Occupation with some income	88	59.1	105	70.0		44	29.3	
Refused/Missing	1							
Income								
$25,000 or less	86	67.2	80	58.4	0.139	132	100	<.0001
>$25,000	42	32.8	57	41.6		–	–	
Refused/Missing	22	–	13	–		18	–	
SES								
Low	100	66.7	90	60.0	0.231	150	100.0	<.0001
High	50	33.3	60	40.0		–	–	
Oral health score								
Poor/Fair	80	64.0	87	58.0	0.310	115	76.7	0.002
Good	45	36.0	63	42.0		35	23.3	
Missing	25							
Teeth removed								
None/1 to 5	86	58.9	92	63.0	0.301	29	19.6	<.0001
6 or more but not all	36	24.7	39	26.7		52	35.1	
All	24	16.4	15	10.3		67	45.3	
Missing	4		4			2		
Smoking status								
Never	33	22.0	32	21.3	0.889	6	4.0	<.0001
Ever	117	78.0	118	78.7		144	96.4	
Drinking habits								
Non-drinker/Mild	78	52.0	85	56.7	0.279	30	20.0	<.0001
Moderate	24	16.0	30	20.0		45	30.0	
Heavy	48	32.0	35	23.3		75	50.0	

Table 2 Prediction Models for Oral Cancer – Logistic Regression, all patients

Logistic Regression, all patients	Odds Ratio (95%CI)	*P*	AUC	Rescaled R^2
Univariate models (150 cases / 150 controls)				
log_2 solCD44	2.036 (1.552, 2.671)	<.0001	0.681	0.137
Protein	2.159 (1.288, 3.617)	0.003	0.590	0.042
Multivariable model * **(149 cases/148 controls)**				
log_2 solCD44	2.684 (1.797, 4.010)	<.0001	0.763	0.276
Protein	0.646 (0.301, 1.386)	0.262		

*Adjusted for age (p=0.132), race/ethnicity (p=0.004), age×race/ethnicity (p=0.006), gender (p=0.030), alcohol (p=0.032), gender×alcohol (p=0.020), smoking (p=0.527), and SES (p=0.042). Note: The inclusion of CD44 and protein improved significantly prediction, since model AUC=0.763 was statistically significant different than AUC= 0.686 for the reduced model excluding the markers and only including potential risk factors (p=0.003).

Supported by the following grants: NCI R01CA118584, NCI RO3 CA107828, 4BB-20 Bankhead-Coley, 10BG-02 Bankhead-Coley, Woman's Cancer Association, a gift from Vigilant Biosciences, Inc, Sylvester Comprehensive Cancer Center and University of Miami, Department of Otolaryngology.

Table 3
Logistics Regression Analysis of Risk Groups derived by Multivariate Recursive Partitioning

Univariate Model of Risk Groups based on CD44 and protein levels

Group size (n = cases + controls)	SolCD44 (ng/ml)	Protein (mg/ml)	Odds Ratio (95%CI)	Prediction	*P*	AUC	Rescaled R^2
102 = 29 + 73	<2.22	<1.23	Reference	control		0.722	0.227
5 = 4 + 1	<2.22	≥1.23	10.069 (1.079, 93.93)	case	0.043		
20 = 16 + 4	≥2.22 & <5.33	<0.558	10.069 (3.103, 32.672)	case	0.0001		
116 = 54 + 62	≥2.22 & <5.33	≥0.558	2.192 (1.247, 3.854)	control	0.006		
57 = 47 + 10	≥5.33		11.830 (5.279, 26.508)	case	<.0001		

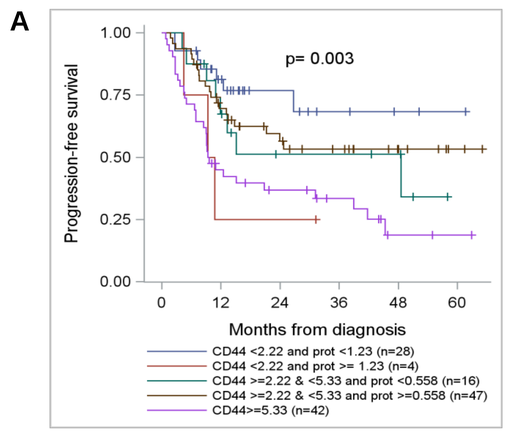
Note: AUC=0.722 for risk group (based on CD44 and protein) model is significantly different from AUC =0.681 for univariate model log_2 solCD44 (p=0.025).

Multivariable Model of Risk Groups

Group size (n)	SolCD44	Protein	Odds Ratio (95%CI)	Prediction	*P*	AUC	Rescaled R^2
102	<2.22	<1.23	Reference	control		0.790	0.325
5	<2.22	≥1.23	5.905 (0.591, 59.053)	case	0.131		
20	≥2.22 & <5.33	<0.558	11.860 (3.312, 42.472)	case	<.0001		
116	≥2.22 & <5.33	≥0.558	2.755 (1.483, 5.117)	control	0.001		
57	≥5.33		14.489 (5.973, 35.145)	case	<.0001		

Logistic regression model included CD44-protein risk groups (5 categories, p<0.0001), age (≥60 vs. <60, p=0.090), gender (p=0.017), race/ethnicity (p=0.001), alcohol (p=0.014), SES (p=0.092), and interaction age×race/ethnicity (p=0.029) and gender×alcohol (p=0.007). Smoking (ever vs. never, p=0.700) and teeth removed (6 or more or all vs. 5 or less, p=0.485) were tested for inclusion into model (AUC=0.791); they were removed since their inclusion did not improve model fit.
AUC: area under the ROC curve. Rescaled R^2: coefficient of determination measured the dispersion explained by model.
Odds ratios: 1-unit increase for continuous variables log2 CD44, protein, and age, unless specified categories; race/ethnicity (WNH and Black vs. WH), gender (Male v. Female), smoking and alcohol (Ever v. Never), and SES (high vs. low).

Figure 2
Kaplan-Meier Curves demonstrating significant differences in PFS and OS



Results

- There were no significant differences between cases and controls with respect to age, gender, race, SES, oral health, smoking history, alcohol habit or enrollment clinic (Figure 1, Table 1)
- CD44 and protein were higher in cases compared to controls at the p<0.05 level when risk factors and demographics were considered.
- In cases but not in controls, CD44 was significantly higher in subjects who were older, had worse gargle ability, and more teeth loss. CD44 and protein did not differ significantly by TNM status or HPV status.
- In univariate analysis, CD44 and total protein were significantly associated with cancer status The AUC was improved to 0.763 in a multivariable model including adjustments for important variables and their interactions (Table 2).
- Analysis using multivariate recursive partitioning resulted in a reference group and 4 risk groups for cancer (Table 3).
- Sensitivity was 80.7% for Stages I-IV and 85% for Stage I-II. Specificity was 48.7% for the hospital-based cohort. For the target screening group, specificity was 74% (n=150) after one baseline evaluation but reached 95% in subjects retested at one year (n=95).
- Based on multivariate analysis including tumor stage, age, gender, race and ethnicity, and SES hospital-based cases that had CD44 levels ≥ 5.33 ng/ml, had reduced PFS and OS compared with cases in the reference group. Kaplan-Meier curves shown in Figure 2.

Conclusions

- Salivary CD44 and total protein levels are significantly associated with oral cavity and oropharyngeal cancer.
- A larger study with longer follow-up period and attention to behavioral changes is needed to validate this simple, inexpensive and noninvasive test as a screening tool.

Literature:

1. Franzmann EJ et al. Soluble CD44 is a potential marker for the early detection of head and neck cancer. CEBP 2007;16(7):1348-55.
2. Franzmann EJ et al. Expression of tumor markers hyaluronic acid and hyaluronidase (HYAL1) in head and neck tumors. Int J Cancer 2003;106: 438-45.
3. Mateus Pereira LH et al. Salivary markers and risk factor data: A multivariate modeling approach for head and neck squamous cell carcinoma detection. Cancer Biomarkers 2011; 10 (5): 241-249.

Disclosure:

The University of Miami, Drs. Franzmann, Reis, Pereira and Duncan hold intellectual property used in the study and have potential for financial benefit from its future commercialization. Dr. Franzmann is Chief Scientific Officer, consultant, and equity holder in Vigilant Biosciences, Inc. University of Miami intellectual property used in this study has been exclusively licensed to Vigilant Biosciences, Inc.

Acknowledements:

We would like to thank members of the University of Miami, Division of Head and Neck Surgery, the Department of Family Medicine, the Sylvester Comprehensive Cancer, Center Disparities and Community Outreach Core, Liberty Square Community Center, Curley's House Food Bank and Dr. John Deo for their assistance with this work.