Published in final edited form as:

Head Neck. 2012 May ; 34(5): 687–695. doi:10.1002/hed.21810.

Salivary Protein and SolCD44 Levels as a Potential Screening Tool for Early Detection of Head and Neck Squamous Cell Carcinoma

Elizabeth J. Franzmann, M.D.^{1,2}, Erika P. Reategui, M.S.³, Lutecia H. Mateus Pereira, Ph.D.¹, Felipe Pedroso, M.D.⁴, Debbie Joseph, M.D.⁵, Glenn O. Allen, MPH¹, Kara Hamilton, MPH⁶, Isildinha Reis, DrPH^{7,8}, Robert Duncan, Ph.D.⁷, W. Jarrard Goodwin, M.D.^{1,2}, Jennifer J. Hu, Ph.D.^{1,7}, and Vinata B. Lokeshwar, Ph.D.^{1,9}

¹Sylvester Comprehensive Cancer Center/University of Miami Leonard Miller School of Medicine, Miami, FL

²Department of Otolaryngology, University of Miami Leonard Miller School of Medicine, Miami, FL

³Department of Clinical Research Administration, University of Miami Leonard Miller School of Medicine, Miami, FL

⁴Department of Pediatrics, New York Presbyterian-Weill Cornell Medical College, New York, NY

⁵Department of Otolaryngology, University of Florida, Gainesville, FL

⁶Institute for Human Genomics, University of Miami Leonard Miller School of Medicine, Miami, FL

⁷Department of Epidemiology and Public Health, University of Miami Leonard Miller School of Medicine, Miami, FL

⁸Sylvester Comprehensive Cancer Center, Division of Biostatistics, University of Miami Leonard Miller School of Medicine, Miami, FL

⁹Department of Urology, University of Miami Leonard Miller School of Medicine, Miami, FL

Abstract

Background—Head and neck squamous cell carcinoma (HNSCC) is a devastating disease usually diagnosed at late stage when cure rates are **40%**. We examined a simple and inexpensive molecular tool that **may aid HNSCC detection**.

Methods—Building on prior findings that total protein levels are elevated in 102 HNSCC cases versus 84 controls, we further analyzed these levels with respect to important risk and demographic variables and compared results to solCD44. Using Multivariate Adaptive Regression Splines (MARS)-logit modeling and logistic regression, we determined whether total protein, solCD44 or the combination best identifies HNSCC.

Results—Combined higher levels of solCD44 and protein were significantly associated with HNSCC (OR=24.90 95% CI 9.04, 68.57, AUC=0.786). A model including protein plus **solCD44** resulted in a better area under the curve (0.796) than either marker alone.

Corresponding author: Elizabeth J. Franzmann, Department of Otolaryngology, Head and Neck Surgery Division, University of Miami, Leonard Miller School of Medicine, and Jackson Memorial Hospital, 1120 NW 14th Street, Clinical Research Bldg, Room 1513, Miami, FL 33136. Phone: 305-243-5955; Fax: 305-243-3383; efranzman@med.miami.edu..

Presented: Presented in abstract form at the Frontiers in Cancer Prevention Research Conference, Philadelphia, Pennsylvania, December 5-8, 2007.

Conclusions—Oral rinse levels of solCD44 and protein appear to hold promise for detection of HNSCC.

Keywords

soluble CD44; ELISA assay; protein; head and neck squamous cell carcinoma; upper aerodigestive tract

Introduction

Each year 50,000 people in the United States¹ and 600,000 people worldwide² are diagnosed with head and neck squamous cell carcinoma (HNSCC). HNSCC includes cancers of the oral cavity, larynx, and pharynx. Early stage disease (stages I and II) is less debilitating and 5-year relative survival rates reach 80% compared to less than 40% for late-stage disease (Stage III and IV) depending on site of disease.³ Unfortunately, over 60% of patients in the United States present with advanced stage disease.⁴

While vocal cord cancer patients develop hoarseness with early stage disease, for other sites, patients usually present at late stage since symptoms do not correlate with presence of disease.⁵ Because oral cavity cancer is the most common site for HNSCC and the mouth is easily **accessible**, one might expect screening examinations to be useful in early detection. Patton reviewed pertinent studies and found insufficient evidence to support the efficacy of such examinations in community-based screening.⁶ While a recent study from India shows a survival advantage for screening by oral cavity exam,⁷ sensitivity and specificity of physical exam are only 75%.⁸ Furthermore, this method is expensive, skill-dependent, and cannot detect occult disease. There are diagnostic kits available for oral cancer such as OralCDx Brush Biopsy (OralCDx Laboratories, Suffern, NY, USA), ViziLite Plus (Zila Pharmaceuticals, Phoenix, AZ, USA) and toluidine blue. Even though these tests can assist in the identification of abnormalities, they require an experienced professional to interpret the results and follow up visits for a more definitive test such as biopsy.⁹ These facts may explain why, despite the American Cancer Society and National Cancer Institute's emphasis that oral examinations could prevent many deaths, few individuals receive these exams.¹⁰⁻¹²

Since screening by physical exam has these limitations, researchers have turned to molecular markers for solutions. Much of the work on HNSCC tumor markers performed in serum has been replicated in whole unstimulated saliva or oral rinses.¹³⁻²⁰ Spafford *et al.*¹³ showed loss of heterozygosity or microsatellite instability in one of 23 markers in 79% of saliva samples from 44 HNSCC patients and in none of 43 healthy control subjects. In a cohort including 211 patients with HNSCC and 527 normal controls, methylation specific marker panels in salivary rinses improved detection when compared with single markers, including a panel with 35% sensitivity and 90% specificity and a panel with 85% sensitivity and 30% specificity.²⁰ Telomerase activity has also been studied using a PCR-based assay where activity was found in 80% of HNSCC patients and 5% of normals.²¹ At the RNA level, salivary transcriptome profiling¹⁶ and most recently microRNA²²⁻²⁷ have been used for HNSCC biomarker discovery. Wong's group, used microarray followed by quantitative PCR analysis of mRNA in saliva and showed that a panel of 4 markers detects HNSCC with 91% sensitivity and 91% specificity.¹⁶ High-throughput technology is also used to identify salivary microRNAs (miRNAs) as putative biomarkers for oral cancer detection. Most often miRNAs are tested in tumor samples yielding very high sensitivity and specificity^{22, 25} for disease prediction. Recently, Park *et al.*¹⁷ detected dozens of miRNAs in saliva and found miR-125a and miR-200a to be present at significantly lower levels ($p < 0.05$) in oral squamous cell carcinoma than in control patients.

Proteomic analysis of saliva has provided good results when studying cytokines as candidate biomarkers for HNSCC. Cytokines are soluble proteins involved in immune and inflammatory responses and have been studied in serum²⁸ and saliva^{15, 18} for immuno-based biomarker discovery. More recently, a new screening tool using multiplexed immunobead-based technology (Luminex Corporation, Austin, TX, USA) has been tested in serum and saliva of HNSCC patients.^{19, 29} Linkov *et al.*, utilizing serum samples, identified a panel comprised of 25 markers including IL-8, EGF and EGFR that offered 84.5% sensitivity and 98% specificity.²⁹

While these studies show promise, in many cases they do not fulfill the ideal characteristics of a screening test, since they are relatively expensive to perform, require considerable expertise, and are not widely available. Furthermore, no molecular screening method for HNSCC has been validated in large studies.

Our data show that a simple, inexpensive oral rinse test shows promise for detecting HNSCC. We have found that the marker soluble CD44 (solCD44) was able to distinguish HNSCC from controls with 62-79% sensitivity and 88-100% specificity.^{30, 31} While in our prior work we tested total protein levels as a potential normalizer for solCD44 levels,³¹ we were surprised to find that total protein levels were, in fact, elevated in cancer cases compared to controls. In this report we further investigate salivary protein levels in oral rinses to 1) determine the relationship between total protein levels and important risk and demographic variables, 2) compare results for total protein and solCD44 and 3) investigate whether the combination of salivary protein plus solCD44 is a better marker for HNSCC than either marker alone.

Materials and Methods

Subject Characteristics

One hundred and two HNSCC patients and 84 controls (**15 healthy volunteers** and 69 patients with benign diseases of the upper aerodigestive tract, 59% smokers) were enrolled according to the protocol approved by the University of Miami Institutional Review Board. **We used the same cohort as previously described.**³¹ To enroll a control population with tobacco and alcohol exposure similar to the HNSCC patient population, control subjects from otolaryngology clinics were approached if they answered “yes” to tobacco or alcohol use on the clinic intake questionnaire. Control patients were excluded if they had a potentially malignant condition or if final diagnosis of their condition was unknown. One control patient was excluded when, in follow-up, the patient developed a dysplastic lesion. Healthy control subjects were volunteers from healthcare and research fields. All HNSCC patients had biopsy proven newly diagnosed or recurrent squamous cell carcinoma. We included all stages and sites except nasopharynx, since nasopharyngeal carcinoma tends to behave differently compared to HNSCC in other sites. Five HNSCC patients had cervical lymph node disease but no identified mucosal primary. Subjects who were known to be pregnant or infected with the human immunodeficiency virus were excluded.

Oral Rinse Collection

The oral rinse collection procedure was previously described.³¹ Five milliliters of normal saline was placed in the subject’s mouths. Patients were asked to swish for five seconds, gargle for five seconds and then deposit the oral rinse into a specimen cup. Saliva was placed on ice for transport and stored at –80 degrees.

Protein Assay

Total protein concentration in oral rinse specimens was determined in our prior study³¹ using the BioRad Protein Assay (BioRad Hercules, CA, USA), as previously described.^{30, 31} For each specimen, protein estimation was carried out at 3 dilutions and each dilution was assayed in triplicate to calculate the average protein concentration (mg/ml).

solCD44 Assay

solCD44 concentration in oral rinse specimens was determined in our previous study.³¹ Briefly, we measured levels of solCD44 using an ELISA assay (Bender MedSystems, Vienna, Austria) that recognizes all CD44 normal and variant isoforms. The test involves a sandwich-type ELISA using a monoclonal anti-solCD44 antibody. Sample concentrations were determined by a standard curve. We optimized the test for oral rinses by preparing the standards in a synthetic saliva matrix (Salimetrics, State College, PA, USA) diluted 1:5 in normal saline (since patients swished and gargled with 5-cc saline) and used a sample diluent (Salimetrics, State College, PA, USA) developed for saliva samples. Samples were vortexed, centrifuged at 3,000×g and the supernatant was used for study. We performed the test at full, 1:2 and 1:4 dilutions. All sample assays were performed in triplicate.

Statistical Analysis

In our previous work we analyzed solCD44 with respect to important risk and demographic variables. While we showed that higher total protein levels were associated with case status, we did not investigate total protein as a potential marker for HNSCC in combination with solCD44.

In this work, we analyze the association of solCD44 and total protein levels and 1) risk factor and demographic covariates within and between the two groups (cancer cases and benign disease controls), and 2) tumor characteristics within the cancer group. The inclusion of solCD44, which was previously reported using a somewhat different analysis,³¹ permits comparison of the two markers with respect to the important risk factors and demographics.

Patients most recently accrued, completed a questionnaire containing information on tobacco and alcohol exposure, race, ethnicity, gender, and socioeconomic status (SES). **We used t-tests to analyze data** for 2 categories and ANOVA for 3 or more categories. Tobacco was analyzed by two methods. In the first method we compared current, former and never smokers. A subject was categorized as a current cigarette smoker if they smoked any amount of cigarettes a day, a former smoker if they smoked at least 100 cigarettes in their lifetime, but smoked 0 cigarettes per day presently, and a never smoker if they never smoked at least 100 cigarettes in their lifetime. In the second method we compared those who ever used a tobacco product to those who never used a tobacco product. Subjects were categorized as ever tobacco users if they smoked at least 100 cigarettes or ever smoked tobacco in a pipe or cigar or chewed smokeless tobacco. Alcohol use was categorized by drinking level. Subjects who drank no alcoholic beverages in a month were nondrinkers. Those who drank 1-15 days per month and had 1-3 beverages on the days they drank were light drinkers. Those who drank 1-5 days per month and 4 or more beverages on the days they drank or 16-30 days per month and 1-3 beverages on the days they drank were categorized as moderate drinkers. Heavy drinkers drank 16-30 days per month and 4 or more alcoholic beverages on the days they drank.

Logistic regression analysis was used to assess the effect of total protein and solCD44 on risk for HNSCC using both continuous and categorical measurements. For continuous measurements, solCD44 data was log base-2 transformed since data were positively skewed to the right. For the categorical representation of data, protein levels were categorized as

high or low based on the median of the control group. A Multivariate Adaptive Regression Splines (MARS) model, as previously described³² was used to identify the optimal cut points for solCD44. MARS software (Salford Systems, San Diego, CA, USA) estimates spline-based models to construct a logistic regression analysis using Statistical Package for the Social Sciences (SPSS) software (IBM, Chicago, IL, USA). Model-derived estimates of sensitivity, specificity, and accuracy are reported.

Results

Table 1 summarizes the demographic and risk factor data for cases and controls as well as solCD44 and protein levels for each group (controls and HNSCC). We have previously described that both solCD44 and total protein levels are significantly elevated in HNSCC patients compared to controls.³¹ Within the control group, a significant association was found between mean protein level and gender ($p=0.004$), with higher protein levels found in males. Results for solCD44 similarly, showed higher levels for men, but did not reach significance. There was a trend towards higher solCD44 levels in current smokers and subjects with 6 or more teeth missing in the control group. Our results show that solCD44 levels in oral rinses are significantly higher in cases versus controls for the former smoker category ($p=0.01$) but not for the current smoker category ($p=0.87$). Since smoking is associated with increased HNSCC risk and this risk begins to decrease as early as 1 year following smoking cessation,³³⁻³⁵ we believe the finding of lower solCD44 levels in former smokers is related to their decreased risk for HNSCC compared to current smokers. Within the cancer group, there were no significant associations between protein and solCD44 levels and any of the risk factors or demographic variables evaluated. In general, protein and solCD44 levels were higher in cases than controls regardless of demographics or risk factor behaviors. Most instances where significance was not reached was likely due to very small numbers ($n<10$) of subjects within at least one of the categories analyzed. Analysis of maximum likelihood estimates showed no significant interaction between solCD44 and protein.

To investigate the effects of tobacco use further, we developed box plots of protein and solCD44 levels in five groups of study subjects with different case/control or smoking status: 1=healthy, non-smoking volunteer ($n=10$), 2=non-smoker controls **with benign disease** or quit > 6 months ago ($n=48$), 3=smoker controls **with benign disease** or quit < 6 months ago ($n=13$), 4=HNSCC patients non-smokers or quit > 6 months ago ($n=18$), 5= HNSCC patients who were smokers or quit < 6 months ago ($n=24$). A global test of the five means in Figure 1 shows a statistically significant difference for solCD44 ($p=0.0417$) and protein ($p=0.0001$). For both solCD44 and total protein, case smokers did not differ from case non-smokers and smoker controls did not differ from non-smokers. The data illustrate that the three groups of controls have very similar distribution of protein and solCD44 levels irrespective of presence of benign disease or history of smoking. Similarly, the two groups of cases also have a very similar distribution of protein and solCD44 levels. Therefore, combined data from three groups of controls and two groups of cases were used for case-control comparison.

We also evaluated whether the protein and solCD44 levels differ by tumor characteristics (Table 2). Five subjects with disease in the cervical lymph nodes and no identified primary site were excluded. Protein levels differed by tumor site ($p=0.048$), with the only significant pairwise comparison being between oropharynx and larynx/hypopharynx. Likewise solCD44 levels differed significantly by tumor site ($p=0.022$); in this case, the only significant pairwise comparison was between oral cavity and larynx/hypopharynx. The differences in protein and in solCD44 between oral cavity and oropharynx groups were not significant at the 5% significance level.

The categorical assessment revealed that higher levels of protein (0.4325 mg/ml) or solCD44 (8.84 ng/ml) was significantly associated with HNSCC (Table 3). HNSCC was significantly associated with high levels of both solCD44 (>14.56 ng/ml) and protein (0.4325 mg/ml) (OR=24.90; 95%CI=9.04, 68.57) (Table 3). The AUC (0.786) was better when the markers were combined in multivariate analysis compared to univariate analysis. Sensitivities and specificities derived from various predicted probability cut points based on the bivariate model were summarized in Table 3.

When continuous measurements were used, univariate and multivariate logistic regression resulted in ROC curves with AUC as shown in Table 4. In univariate analysis, for every two-fold increase in solCD44, there was a 2.20-fold (95%CI=1.66, 2.93) increased risk for HNSCC; for every unit increase of protein, there was a 5.82-fold (95%CI=2.80, 12.13) increased risk for HNSCC. In bivariate analysis, the adjusted OR for solCD44 and protein was 1.86 (95%CI=1.38, 2.50) and 3.10 (95%CI=1.47, 6.52), respectively. The model including solCD44 and protein presented the best AUC value at 0.796. Sensitivities and specificities derived from various predicted probability cut points based on the bivariate model were summarized in Table 4. Best accuracy was obtained using a predictive probability cutpoint of 0.47, with corresponding estimates of sensitivity and specificity 80.4 and 65.5, respectively.

Discussion

Our previous work shows that solCD44 can distinguish HNSCC patients from controls with 62% sensitivity and 88% specificity.³¹ However, our prior work did not include analysis of total protein as a potential marker for HNSCC. Our results in this work indicate that subjects with high solCD44 and protein levels in the oral rinse are almost 25 times more likely to have cancer than those without these elevated levels. This association of the combined markers with cancer is much stronger than that of total protein alone (OR=5.00, 95%CI=2.55, 9.81) or solCD44 alone (OR=12.43, 95%CI=5.40, 28.61). This is supported by the ROC analysis which shows a higher AUC value of 0.796 with the combination of solCD44 and protein compared to either marker alone. Since the solCD44 and protein tests are very simple and inexpensive, this combination may be useful for early diagnosis of HNSCC.

This work also shows for the first time that protein levels were elevated regardless of tumor size or stage, indicating that these markers are present early in carcinogenesis. Furthermore, eighty-two percent of the 84 control subjects had benign diseases of the upper aerodigestive tract (UADT). Thus, the elevated protein level appears to distinguish cancer from benign disease. Hydration level does not appear to have a great impact on results for the oral rinse technique. Patients with small T1 lesions are not expected to have swallowing difficulties and thus should have similar hydration status as the control patients, yet the oral rinse protein levels of T1 patients were still elevated compared to control subjects.

Site appears to play an important role in protein and solCD44 levels, with larynx and hypopharynx having lower levels than oral cavity and oropharynx. However, marker levels are not different by tumor stage, nodal status, recurrence status, or history of second primary. Human papillomavirus (HPV) confers an increased risk to head and neck cancer, especially cancer of the oropharynx. While we did not have HPV status information for our patients, we did analyze results based on tumor site (Table 2). For solCD44, levels were significantly higher for oral cavity compared to larynx/hypopharynx. For protein, levels were significantly higher for oropharynx compared to larynx/hypopharynx. Therefore, while HPV may have an effect for protein, it does not appear to have an effect for solCD44.

Elevated protein levels in salivary rinses from HNSCC patients may be partly due to overexpression of numerous soluble proteins. Smith *et al.* showed global hypomethylation, in HNSCC tumor tissues compared to normal mucosa tissues.³⁷ Such hypomethylation would be expected to result in an overall increase in protein expression in HNSCC. Indeed, Hu *et al.* identified a group of five proteins that in combination were useful for detecting HNSCC.¹⁸ While global overexpression of tumor associated proteins is possible, it is also likely that tumors, being highly vascular, promote leaking of proteins such as albumin and IgG from the serum into the oral rinse as has been suggested by the work of Shpitzer *et al.*³⁸ The high levels of protein that we see, in the mg/ml range, suggest that this may be the case since tumor markers are generally expressed in the ng/ml range.

Saliva is becoming the biologic sample of choice for head and neck cancer screening.³⁶ It can be obtained simply and less invasively than blood and thus can be collected without special training. In general, there are three methods for collecting saliva. These include salivary rinses, whole saliva, and stimulated saliva. Our laboratory is studying soluble molecular markers for HNSCC using oral rinses. The patient swishes with 5 cc saline for 5 seconds, gargles the rinse for 5 seconds and then expectorates into a collection cup. While the oral rinse does contain whole saliva, it also results in a specimen that has had optimal contact with all regions of the UADT compared to whole saliva or stimulated saliva which only permits contact with the anterior oral cavity. Furthermore, concentrations of whole saliva components can vary greatly with hydration level. Since there is more contact with surfaces that contain secreted proteins, it is quite possible that the relative protein composition of the oral rinse would be different and perhaps more favorable than that of whole or stimulated saliva. This may explain why we see significantly higher total protein levels in HNSCC patients compared to controls when using the oral rinse, while we and other researchers did not see a significant difference when using other sampling methods.¹⁵

Ongoing work is investigating the associations between specific risk factors and demographics with protein and solCD44 levels. Our data show that protein levels are significantly higher in male controls compared to females. Further investigation is needed to verify this finding and to determine if other demographic and risk variables are associated with marker levels. Our data here suggest a trend towards association between higher solCD44 levels and current tobacco use and poor oral health in controls, though our analyses did not reach statistical significance. As can be inferred by the large standard deviation, most of the elevation seen in current smokers in the control group is due to a few subjects with very high levels. Thus, it is possible that the test is identifying very early disease, as has been described in previous work.³¹ This is not surprising since both tobacco use and poor oral health are risk factors for HNSCC.^{39,40}

Since approximately 20% of American still smoke and many more millions drink heavily or are exposed to HPV, a very simple and inexpensive test is needed to identify those individuals who are likely to have HNSCC or premalignant disease. Because it is simple, inexpensive, noninvasive and appears to be effective, we envision that a test including combined measurements of solCD44 or total protein would serve this purpose. Subjects with combined results above a predetermined cut-point would be referred to a specialist for examination and biopsy. Subjects with cancer or precancer would be treated with standard of care. Subjects without lesions could be followed closely with oral rinse tests and physical exam and encouraged to stop risky behavior such as smoking. Our prior work suggests that the test may identify disease before it becomes clinically visible.³¹ Since premalignant lesions are known to reverse with smoking cessation,⁴¹ even in the case of a positive rinse test and no identifiable lesion, the subject may be encouraged to modify their behavior in a positive manner. This may result in a reversal of marker levels and diminished risk of disease. Further work is underway to determine whether solCD44 levels change with

smoking cessation and whether subjects are likely to change risky behavior based on an oral rinse test result.

In summary, in this study of 102 HNSCC and 84 controls, oral rinse protein levels are significantly associated with HNSCC. High levels of solCD44 and protein are nearly 25 times more likely to have HNSCC than subjects without levels in this range. Our statistical models suggest that the combination of solCD44 and total protein will improve sensitivity and specificity of the HNSCC detection test compared to either marker alone. Our ongoing work will examine these findings further in a larger matched case-control study.

Acknowledgments

Grant Support: Grant support from the Flight Attendant Medical Research Institute [12536]; National Cancer Institute [R03CA107828, R01CA118584]; Women's Cancer Association and Sylvester Comprehensive Cancer Center.

References

1. American Cancer Society. Cancer Facts & Figures 2009. American Cancer Society; Atlanta: 2009.
2. Parkin DM, Bray F, Ferlay J, Pisani P. Global cancer statistics, 2002. *CA Cancer J Clin.* 2005; 55(2):74–108. [PubMed: 15761078]
3. Ries, LAGYJ.; Keel, GE.; Eisner, MP.; Lin, YD.; Horner, M-J., editors. SEER Survival Monograph: Cancer Survival Among Adults: U.S. SEER Program, 1988-2001, Patient and Tumor Characteristics. National Cancer Institute; Bethesda, MD: 2007. SEER Program, NIH Pub. No. 07-6215
4. Forastiere A, Koch W, Trotti A, Sidransky D. Head and neck cancer. *N Engl J Med.* 2001; 345(26): 1890–900. [PubMed: 11756581]
5. Dolan RW, Vaughan CW, Fuleihan N. Symptoms in early head and neck cancer: an inadequate indicator. *Otolaryngol Head Neck Surg.* 1998; 119(5):463–7. [PubMed: 9807070]
6. Patton LL. The effectiveness of community-based visual screening and utility of adjunctive diagnostic aids in the early detection of oral cancer. *Oral Oncol.* 2003; 39(7):708–23. [PubMed: 12907211]
7. Sankaranarayanan R, Ramadas K, Thomas G, et al. Effect of screening on oral cancer mortality in Kerala, India: a cluster-randomised controlled trial. *Lancet.* 2005; 365(9475):1927–33. [PubMed: 15936419]
8. Sankaranarayanan R, Mathew B, Jacob BJ, et al. Early findings from a community-based, cluster-randomized, controlled oral cancer screening trial in Kerala, India. The Trivandrum Oral Cancer Screening Study Group. *Cancer.* 2000; 88(3):664–73. [PubMed: 10649262]
9. Wong DT. Towards a simple, saliva-based test for the detection of oral cancer 'oral fluid (saliva), which is the mirror of the body, is a perfect medium to be explored for health and disease surveillance'. *Expert Rev Mol Diagn.* 2006; 6(3):267–72. [PubMed: 16706730]
10. Neville BW, Day TA. Oral cancer and precancerous lesions. *CA Cancer J Clin.* 2002; 52(4):195–215. [PubMed: 12139232]
11. National Cancer Institute/US. Oral Cancers: research report. US Department of Health and Human Services, Public Health Service, National Institutes of Health; 1991. NIH publication no. 92-2875
12. Weinberg MA, Estefan DJ. Assessing oral malignancies. *Am Fam Physician.* 2002; 65(7):1379–84. [PubMed: 11996421]
13. Spafford MF, Koch WM, Reed AL, et al. Detection of head and neck squamous cell carcinoma among exfoliated oral mucosal cells by microsatellite analysis. *Clin Cancer Res.* 2001; 7(3):607–12. [PubMed: 11297256]
14. Boyle JO, Mao L, Brennan JA, et al. Gene mutations in saliva as molecular markers for head and neck squamous cell carcinomas. *Am J Surg.* 1994; 168(5):429–32. [PubMed: 7977967]

15. St John MA, Li Y, Zhou X, et al. Interleukin 6 and interleukin 8 as potential biomarkers for oral cavity and oropharyngeal squamous cell carcinoma. *Arch Otolaryngol Head Neck Surg.* 2004; 130(8):929–35. [PubMed: 15313862]
16. Li Y, St John MA, Zhou X, et al. Salivary transcriptome diagnostics for oral cancer detection. *Clin Cancer Res.* 2004; 10(24):8442–50. [PubMed: 15623624]
17. Park NJ, Zhou H, Elashoff D, et al. Salivary microRNA: discovery, characterization, and clinical utility for oral cancer detection. *Clin Cancer Res.* 2009; 15(17):5473–7. [PubMed: 19706812]
18. Hu S, Arellano M, Boontheung P, et al. Salivary proteomics for oral cancer biomarker discovery. *Clin Cancer Res.* 2008; 14(19):6246–52. [PubMed: 18829504]
19. Arellano-Garcia ME, Hu S, Wang J, et al. Multiplexed immunobead-based assay for detection of oral cancer protein biomarkers in saliva. *Oral Dis.* 2008; 14(8):705–12. [PubMed: 19193200]
20. Carvalho AL, Jeronimo C, Kim MM, et al. Evaluation of promoter hypermethylation detection in body fluids as a screening/diagnosis tool for head and neck squamous cell carcinoma. *Clin Cancer Res.* 2008; 14(1):97–107. [PubMed: 18172258]
21. Califano J, Ahrendt SA, Meiningner G, Westra WH, Koch WM, Sidransky D. Detection of telomerase activity in oral rinses from head and neck squamous cell carcinoma patients. *Cancer Res.* 1996; 56(24):5720–2. [PubMed: 8971181]
22. Chang SS, Jiang WW, Smith I, et al. MicroRNA alterations in head and neck squamous cell carcinoma. *Int J Cancer.* 2008; 123(12):2791–7. [PubMed: 18798260]
23. Wong TS, Liu XB, Wong BY, Ng RW, Yuen AP, Wei WI. Mature miR-184 as Potential Oncogenic microRNA of Squamous Cell Carcinoma of Tongue. *Clin Cancer Res.* 2008; 14(9):2588–92. [PubMed: 18451220]
24. Kozaki K, Imoto I, Mogi S, Omura K, Inazawa J. Exploration of tumor-suppressive microRNAs silenced by DNA hypermethylation in oral cancer. *Cancer Res.* 2008; 68(7):2094–105. [PubMed: 18381414]
25. Avissar M, Christensen BC, Kelsey KT, Marsit CJ. MicroRNA expression ratio is predictive of head and neck squamous cell carcinoma. *Clin Cancer Res.* 2009; 15(8):2850–5. [PubMed: 19351747]
26. Avissar M, McClean MD, Kelsey KT, Marsit CJ. MicroRNA expression in head and neck cancer associates with alcohol consumption and survival. *Carcinogenesis.* 2009; 30(12):2059–63. [PubMed: 19901002]
27. Christensen BC, Moyer BJ, Avissar M, et al. A let-7 microRNA-binding site polymorphism in the KRAS 3' UTR is associated with reduced survival in oral cancers. *Carcinogenesis.* 2009; 30(6):1003–7. [PubMed: 19380522]
28. Hathaway B, Landsittel DP, Gooding W, et al. Multiplexed analysis of serum cytokines as biomarkers in squamous cell carcinoma of the head and neck patients. *Laryngoscope.* 2005; 115(3):522–7. [PubMed: 15744170]
29. Linkov F, Lisovich A, Yurkovetsky Z, et al. Early detection of head and neck cancer: development of a novel screening tool using multiplexed immunobead-based biomarker profiling. *Cancer Epidemiol Biomarkers Prev.* 2007; 16(1):102–7. [PubMed: 17220337]
30. Franzmann EJ, Reategui EP, Carraway KL, Hamilton KL, Weed DT, Goodwin WJ. Salivary soluble CD44: a potential molecular marker for head and neck cancer. *Cancer Epidemiol Biomarkers Prev.* 2005; 14(3):735–9. [PubMed: 15767360]
31. Franzmann EJ, Reategui EP, Pedroso F, et al. Soluble CD44 is a potential marker for the early detection of head and neck cancer. *Cancer Epidemiol Biomarkers Prev.* 2007; 16(7):1348–55. [PubMed: 17627000]
32. Smith TR, Levine EA, Freimanis RI, et al. Polygenic model of DNA repair genetic polymorphisms in human breast cancer risk. *Carcinogenesis.* 2008; 29(11):2132–8. [PubMed: 18701435]
33. Blot WJ, McLaughlin JK, Winn DM, et al. Smoking and drinking in relation to oral and pharyngeal cancer. *Cancer Res.* 1988; 48(11):3282–7. [PubMed: 3365707]
34. Schlecht NF, Franco EL, Pintos J, Kowalski LP. Effect of smoking cessation and tobacco type on the risk of cancers of the upper aero-digestive tract in Brazil. *Epidemiology.* 1999; 10(4):412–8. [PubMed: 10401876]

35. La Vecchia C, Franceschi S, Bosetti C, Levi F, Talamini R, Negri E. Time since stopping smoking and the risk of oral and pharyngeal cancers. *J Natl Cancer Inst.* 1999; 91(8):726–8. [PubMed: 10218516]
36. Arellano M, Jiang J, Zhou X, et al. Current advances in identification of cancer biomarkers in saliva. *Front Biosci (Schol Ed).* 2009; 1:296–303. [PubMed: 19482703]
37. Smith IM, Mydlarz WK, Mithani SK, Califano JA. DNA global hypomethylation in squamous cell head and neck cancer associated with smoking, alcohol consumption and stage. *Int J Cancer.* 2007; 121(8):1724–8. [PubMed: 17582607]
38. Shpitzer T, Bahar G, Feinmesser R, Nagler RM. A comprehensive salivary analysis for oral cancer diagnosis. *J Cancer Res Clin Oncol.* 2007; 133(9):613–7. [PubMed: 17479291]
39. Hashibe M, Brennan P, Benhamou S, et al. Alcohol drinking in never users of tobacco, cigarette smoking in never drinkers, and the risk of head and neck cancer: pooled analysis in the International Head and Neck Cancer Epidemiology Consortium. *J Natl Cancer Inst.* 2007; 99(10): 777–89. [PubMed: 17505073]
40. Guha N, Boffetta P, Filho V Wunsch, et al. Oral health and risk of squamous cell carcinoma of the head and neck and esophagus: results of two multicentric case-control studies. *Am J Epidemiol.* 2007; 166(10):1159–73. [PubMed: 17761691]
41. Banoczy J, Gintner Z, Dombi C. Tobacco use and oral leukoplakia. *J Dent Educ.* 2001; 65(4):322–7. [PubMed: 11336117]

FIGURE 1.

Boxplots showing protein and solCD44 levels by risk group

1= **Healthy volunteers** ($n=10$)

2= Non-smoker controls **with benign disease** or quit > 6 months ago ($n=48$)

3= Smoker controls **with benign disease** or quit < 6 months ago ($n=13$)

4= HNSCC patients non-smokers or quit > 6 months ago ($n=18$)

5= HNSCC patients who were smokers or quit < 6 months ago ($n=24$)

The solid line indicates the median for each group. Box edges mark the 25th and 75th percentiles of the observed values, and the T-bars indicate the 10th and 90th percentiles. The horizontal dashed line marks the median of group 1. Subjects were excluded if smoking status was unknown. Extreme outliers (very high levels) for group 5 were excluded.

Table 1

Demographic and risk factor data for cases and controls

	Controls			Cases			Group comparison		
	n=84 n (%)	Protein mean (SD)	SolCD44 mean (SD)	n=102 n (%)	Protein mean (SD)	SolCD44 mean (SD)	Protein P [†]	SolCD44 P [†]	
Total	84 (100.0)	0.55 (0.39)	9.31 (14.85)	102 (100.0)	1.07 (0.83)	24.44 (32.01)	<.001	<.001	
Age group									
60 y	51 (81.0)	0.52 (0.36)	10.29 (18.63)	20 (46.5)	0.81 (0.44)	15.93 (13.78)	0.006	0.22	
> 60 y	12 (19.0)	0.43 (0.24)	9.20 (4.52)	23 (53.5)	1.08 (0.64)	24.29 (29.25)	0.002	0.09	
Missing	21			59					
P *		0.38	0.84		0.11	0.25			
Race									
Non-White	10 (16.1)	0.50 (0.53)	6.08 (3.90)	3 (7.1)	1.37 (1.21)	8.77 (4.81)	0.09	0.34	
White	52 (83.9)	0.51 (0.30)	10.82 (18.40)	39 (92.9)	0.94 (0.50)	21.40 (24.43)	<.001	0.02	
Missing	22			60					
P *		0.97	0.42		0.21	0.38			
Gender									
Male	42 (54.6)	0.64 (0.41)	11.14 (20.14)	84 (84.0)	1.11 (0.85)	23.45 (31.32)	0.001	0.022	
Female	35 (45.4)	0.41 (0.24)	6.93 (4.60)	16 (16.0)	0.91 (0.71)	31.88 (36.95)	0.0005	0.0002	
Missing	7			2					
P *		0.004	0.230		0.375	0.340			
Education level									
More than high school	54 (85.7)	0.52 (0.36)	9.99 (16.91)	17 (39.5)	0.89 (0.50)	26.66 (31.20)	0.001	0.006	
No school - high school	9 (14.3)	0.41 (0.20)	10.64 (17.46)	26 (60.5)	1.00 (0.61)	16.31 (16.13)	0.008	0.38	
Missing	21			59					
P *		0.36	0.92		0.51	0.16			
Income level									
<35K	7 (12.1)	0.47 (0.28)	9.78 (3.71)	24 (64.9)	0.98 (0.58)	16.18 (15.97)	0.03	0.31	
>35K	51 (87.9)	0.47 (0.33)	8.11 (9.15)	13 (35.1)	0.73 (0.37)	25.25 (33.99)	0.02	0.002	
Missing	26			65					

	Controls		Cases		Group comparison			
	n=84 n (%)	Protein mean (SD)	SolCD44 mean (SD)	n=102 n (%)	Protein mean (SD)	SolCD44 mean (SD)	Protein <i>P</i> [†]	SolCD44 <i>P</i> [†]
<i>P</i> [*]		0.98	0.64		0.17	0.28		
Ever smoked cigarettes								
No	26 (41.3)	0.53 (0.39)	7.44 (4.54)	5 (11.9)	0.65 (0.35)	11.50 (9.04)	0.53	0.13
Yes	37 (58.7)	0.49 (0.30)	11.93 (21.58)	37 (88.1)	0.97 (0.56)	21.49 (24.94)	<001	0.08
Missing	21			60				
<i>P</i> [*]		0.65	0.30		0.23	0.38		
Cigarette use status								
Current	13 (20.6)	0.48 (0.31)	19.72 (34.88)	20 (47.6)	0.95 (0.47)	18.22 (17.46)	0.003	0.87
Former	24 (38.1)	0.49 (0.31)	7.71 (6.44)	17 (40.5)	0.98 (0.65)	25.33 (31.76)	0.003	0.01
Never	26 (41.3)	0.53 (0.39)	7.44 (4.54)	5 (11.9)	0.65 (0.35)	11.50 (9.04)	0.53	0.13
Missing	21			60				
<i>P</i> [*]		0.90	0.07		0.48	0.46		
Tobacco use								
Ever used	45 (71.4)	0.50 (0.30)	11.31 (19.69)	42 (97.7)	0.96 (0.57)	20.63 (23.73)	<001	0.049
Never used any	18 (28.6)	0.51 (0.43)	7.00 (4.21)	1 (2.3)	0.46	10.60	0.91	0.42
Missing	21			59				
<i>P</i> [*]		0.95	0.36		0.38	0.68		
Drinking Level								
Nondrinker	9 (14.8)	0.58 (0.31)	10.45 (6.47)	17 (42.5)	0.96 (0.63)	12.90 (7.08)	0.11	0.40
Light drinker	34 (55.7)	0.45 (0.35)	10.55 (21.11)	8 (20.0)	0.94 (0.38)	32.71 (41.61)	0.001	0.04
Moderate drinker	15 (24.6)	0.62 (0.35)	10.49 (13.04)	8 (20.0)	0.83 (0.53)	20.84 (20.51)	0.25	0.15
Heavy drinker	3 (4.9)	0.32 (0.14)	4.29 (3.17)	7 (17.5)	0.91 (0.57)	17.21 (20.52)	0.12	0.32
Missing	23			62				
<i>P</i> [*]		0.27	0.95		0.96	0.25		
Teeth removed								
None	35 (55.6)	0.49 (0.35)	8.96 (10.60)	9 (20.9)	0.72 (0.46)	11.84 (7.82)	0.11	0.45
1-5	19 (30.2)	0.43 (0.32)	6.16 (4.25)	16 (37.2)	1.08 (0.67)	28.04 (34.10)	0.001	0.009
6 or more	8 (12.7)	0.75 (0.29)	24.42 (40.46)	9 (20.9)	1.06 (0.57)	17.74 (8.13)	0.19	0.63

	Controls			Cases			Group comparison		
	n=84 n (%)	Protein mean (SD)	SolCD44 mean (SD)	n=102 n (%)	Protein mean (SD)	SolCD44 mean (SD)	Protein $P^{\ddagger}$	SolCD44 $P^{\dagger}$	
All	1 (1.6)	0.43	8.95	9 (20.9)	0.86 (0.43)	18.03 (19.41)	0.37	0.67	
Missing	21			59					
P^*		0.15	0.07		0.42	0.39			

* Difference in Protein or SolCD44 by selected characteristics within control or case groups.

‡ Difference in Protein or SolCD44 between controls and cases.

Table 2

Total protein (mg/ml) and solCD44 (ng/ml) by tumor characteristics for HNSCC cases

Tumor Characteristics		Protein (mg/ml)		solCD44 (ng/ml)*		
	N	Mean±SD	P	N	Mean±SD	P
Site						
Larynx/Hypopharynx	35	0.814 [†] ± 0.478	0.048	35	15.112 [‡] ± 9.838	0.022
Oral Cavity	38	1.095 ± 0.757		38	35.657 [‡] ± 45.921	
Oropharynx	24	1.249 [†] ± 0.819		24	21.389 ± 24.306	
T/N/M Classification						
T1	27	1.015 ± 0.592	0.74	27	26.00 ± 29.11	0.34
T2	28	0.955 ± 0.614		28	19.36 ± 18.69	
T3	19	1.186 ± 0.98		19	35.95 ± 56.55	
T4	23	1.017 ± 0.679		23	20.43 ± 21.58	
T/N/M stage						
I	21	0.994 ± 0.588	0.97	21	24.59 ± 31.47	0.78
II	19	0.982 ± 0.674		19	17.84 ± 15.15	
III	20	1.071 ± 0.621		20	27.38 ± 39.64	
IV	37	1.057 ± 0.83		37	26.87 ± 36.49	
Nodal status						
N0	65	1.044 ± 0.654	0.82	65	20.450 ± 23.849	0.15
N1,N2,N3	32	1.008 ± 0.803		37	31.460 ± 42.245	
Recurrence						
No	85	1.061 ± 0.714	0.28	87	23.636 ± 32.312	0.54
Yes	12	0.826 ± 0.598		15	29.127 ± 30.863	
2nd Primary						
No	90	1.017 ± 0.695	0.45	94	24.068 ± 32.366	0.69
Yes	7	1.224 ± 0.828		8	28.864 ± 29.064	
RT Therapy						
No	82	1.044 ± 0.711	0.68	85	23.003 ± 31.720	0.31
Yes	15	0.963 ± 0.669		17	31.649 ± 33.480	

* The solCD44 data were derived from a previous study.³¹

[†]Pairwise mean comparisons significant at $p < 0.05$ by Fisher's least-significant-difference test are indicated by the same symbol.

[‡]Pairwise mean comparisons significant at $p < 0.05$ by Fisher's least-significant-difference test are indicated by the same symbol.

Table 3
Categorical assessment of solCD44 (ng/ml) and total protein (mg/ml) as molecular predictors of HNSCC

Logistic regression models					
	Control n (%)	Case n (%)	OR (95% CI)	P	AUC*
Univariate					
Protein					
Low (<0.4325)	42 (50)	17 (17)	Reference		
High (≥0.4325)	42 (50)	85 (83)	5.00 (2.55, 9.81)	<0001	0.667
SolCD44					
Low (<8.84)	59 (70)	29 (28)	Reference		
Intermediate (8.84–14.55)	16 (19)	18 (18)	2.29 (1.02, 5.12)	0.044	0.751
High (≥14.55)	9 (11)	55 (54)	12.43 (5.40, 28.61)	<0001	
Bivariate[†]					
Protein					
SolCD44					
Low	34 (41)	10 (10)	Reference		
Intermediate	6 (7)	3 (3)	1.88 (0.81, 4.33)	0.141	0.786
High	2 (2)	4 (4)	8.86 (3.73, 21.05)	<0001	
Low	25 (30)	19 (18)	2.81 (1.34, 5.90)	0.006	
Intermediate	10 (12)	15 (15)	5.27 (1.86, 14.88)	0.002	
High	7 (8)	51 (50)	24.90 (9.04, 68.57)	<0001	
Examples of classification based on bivariate model					
	Predicted probability cutpoint		Sensitivity (%)	Specificity (%)	Accuracy (%)
Best accuracy	0.59		68.6	77.4	72.6
Better sensitivity	0.44		87.3	47.6	69.4
Better specificity	0.71		53.9	89.3	69.9

* AUC: area under receiver operating curve.

[†] Interaction protein x solCD44 was not significant ($p=0.946$, 2 degrees of freedom).

Table 4

Continuous measurement of solCD44 (ng/ml) and total protein (mg/ml) as molecular predictors of HNSCC

Logistic regression models				
	Odds Ratio [*] (95%CI)	P	AUC [†]	
Univariate				
Protein	5.82 (2.80, 12.13)	<.0001	0.733	
Log ₂ (solCD44)	2.20 (1.66, 2.93)	<.0001	0.776	
Bivariate [‡]				
Protein	3.10 (1.47, 6.52)	0.003	0.796	
Log ₂ (solCD44)	1.86 (1.38, 2.50)	<.0001		
Examples of classification based on bivariate model				
	Predicted probability cutpoint	Sensitivity (%)	Specificity (%)	Accuracy (%)
Best accuracy	0.47	80.4	65.5	73.7
Better sensitivity	0.43	83.3	59.5	72.6
Better specificity	0.58	64.7	78.6	71.0

* Odds ratios correspond to a two-fold increase in solCD44 level, and a unit increase in protein level, respectively.

[†] AUC: area under receiver operating curve.

[‡] Interaction protein x log₂ (solCD44) was not significant ($p=0.306$, 1 degree of freedom).