Different cellular p16^{INK4a} localisation may signal different survival outcomes in head and neck cancer

N Zhao[1,2], M-K Ang[3], X-Y Yin[1,4], MR Patel[4], K Fritchie[5], L Thorne[1,5], KL Muldrew[5], MC Hayward[1], W Sun[2], MD Wilkerson[1], BS Chera[1,6], T Hackman[4], AM Zanation[4], JE Grilley-Olson[1,7], ME Couch[4], WW Shockley[4], MC Weissler[1,4], CG Shores[1,4], WK Funkhouser[1,5], AF Olshan[8] and DN Hayes*[,1,7]

[1]Lineberger Comprehensive Cancer Center, University of North Carolina, Chapel Hill, NC, USA; [2]Department of Biostatistics, Gillings School of Global Public Health, University of North Carolina, Chapel Hill, NC, USA; [3]Department of Medical Oncology, National Cancer Centre Singapore, Singapore; [4]Division of Otolaryngology, School of Medicine, University of North Carolina, Chapel Hill, NC, USA; [5]Division of Pathology and Laboratory Medicine, School of Medicine, University of North Carolina, Chapel Hill, NC, USA; [6]Department of Radiation Oncology, School of Medicine, University of North Carolina, Chapel Hill, NC, USA; [7]Division of Hematology and Oncology, School of Medicine, University of North Carolina, Chapel Hill, NC, USA; [8]Department of Epidemiology, Gillings School of Global Public Health, University of North Carolina, Chapel Hill, NC, USA

BACKGROUND: Recently, the management of head and neck squamous cell carcinoma (HNSCC) has focused considerable attention on biomarkers, which may influence outcomes. Tests for human papilloma infection, including direct assessment of the virus as well as an associated tumour suppressor gene p16, are considered reproducible. Tumours from familial melanoma syndromes have suggested that nuclear localisation of p16 might have a further role in risk stratification. We hypothesised p16 staining that considered nuclear localisation might be informative for predicting outcomes in a broader set of HNSCC tumours not limited to the oropharynx, human papilloma virus (HPV) status or by smoking status.

METHODS: Patients treated for HNSCC from 2002 to 2006 at UNC (University of North Carolina at Chapel Hill) hospitals that had banked tissue available were eligible for this study. Tissue microarrays (TMA) were generated in triplicate. Immunohistochemical (IHC) staining for p16 was performed and scored separately for nuclear and cytoplasmic staining. Human papilloma virus staining was also carried out using monoclonal antibody E6H4. p16 expression, HPV status and other clinical features were correlated with progression-free (PFS) and overall survival (OS).

RESULTS: A total of 135 patients had sufficient sample for this analysis. Median age at diagnosis was 57 years (range 20–82), with 68.9% males, 8.9% never smokers and 32.6% never drinkers. Three-year OS rate and PFS rate was 63.0% and 54.1%, respectively. Based on the p16 staining score, patients were divided into three groups: high nuclear, high cytoplasmic staining group (HN), low nuclear, low cytoplasmic staining group (LS) and high cytoplasmic, low nuclear staining group (HC). The HN and the LS groups had significantly better OS than the HC group with hazard ratios of 0.10 and 0.37, respectively, after controlling for other factors, including HPV status. These two groups also had significantly better PFS than the HC staining group. This finding was consistent for sites outside the oropharynx and did not require adjustment for smoking status.

CONCLUSION: Different p16 protein localisation suggested different survival outcomes in a manner that does not require limiting the biomarker to the oropharynx and does not require assessment of smoking status.

British Journal of Cancer (2012) **107,** 482–490. doi:10.1038/bjc.2012.264   www.bjcancer.com
Published online 26 June 2012
© 2012 Cancer Research UK

Keywords: p16 cellular localisation; tissue microarray; HPV infection

Molecular Diagnostics

Head and neck squamous cell carcinoma (HNSCC) diagnoses constitute ~3–5% of all cancers with an estimate of 49 000 new cases and 11 000 deaths in 2010 in the United States (National Cancer Institute, 2005; Jemal *et al*, 2010). Recent epidemiological data suggest an increasing incidence rate among younger people who are often nonsmokers and nondrinkers (Schantz and Yu, 2002; Shiboski *et al*, 2005; Curado and Hashibe, 2009; Marur *et al*, 2010; Patel *et al*, 2011), which are frequently attributable to human papilloma virus (HPV) infection (Franceschi *et al*, 1996; El-Mofty and Lu, 2003; Dahlstrand *et al*, 2004; Furniss *et al*, 2007;

Chaturvedi *et al*, 2011). Human papilloma virus-positive tumours are typically found in the oropharynx and have better response to treatment (Fakhry *et al*, 2008) and better disease outcome (Hafkamp *et al*, 2008; Ang *et al*, 2010). There is significant consensus that knowledge of patient HPV status will increasingly have a role in the management of this disease.

However, assessment of risk in the context of HPV infection has ongoing challenges. Perhaps chief among these is the fact that the diagnostic tests for the infection have limitations, and second, that smoking appears to degrade the favourable outcomes in patients with HPV-associated cancers for reasons that are unclear. There are two broad categories of assays for HPV. In the first category are tests for the virus itself including PCR, IHC and *in situ* hybridisation. Alternatively, HPV status can be assessed indirectly through the p16 biomarker, which is generally highly expressed

in the setting of HPV infection. Detection of HPV directly suffers from a variety of limitations including both false positives and false negatives depending on the setting for reasons that have been extensively reviewed (Shroyer and Greer, 1991; Ha et al, 2002; Termine et al, 2008; Stevens et al, 2011). Recently, large clinical trials have addressed the false positive concern primarily by assessing HPV only in the oropharynx, assuming that most positive tests outside the oropharynx would be false positives. The concern for false negatives has frequently been addressed with the addition of the biomarker p16, which is highly correlated with HPV infection because it is generally believed that HPV in situ hybridisation is less sensitive and more specific than p16 staining (Begum et al, 2007; Schache et al, 2011; Stevens et al, 2011). In fact, recent studies have consistently shown favourable correlation between the two biomarkers, with nearly all HPV-positive samples also staining for p16 (Begum et al, 2007). Interestingly, however, there is also a consistent pattern of p16-positive, HPV-negative oropharynx tumours in the range of ~20% (Ang et al, 2010). Strikingly, however, p16-negative, HPV-positive tumours are rare. Most commonly, the p16-positive, HPV-negative case has been attributed to a failed test of HPV, such as the presence of an HPV subtype not assessed by the assay. Such an explanation fails to address the fact that p16 is frequently positive in HNSCC outside the oropharynx, where HPV infection has generally been classified as a rare event. Interestingly, p16 positivity within the oropharynx appears to be at least as good a marker of favourable outcome, independent of whether samples also stained for HPV (Reimers et al, 2007; Ang et al, 2010). Yet outside the oropharynx, p16 has only infrequently been reported as a favourable marker (Harris et al, 2010b).

In addition to the complex story involving tumour site (oropharynx) and the biomarkers p16 and/or HPV is the fact that risk is also modified by smoking (Ang et al, 2010). Patients with greater smoking histories appear to have their favourable outcomes significantly tempered relative to nonsmoking HPV/p16-positive oropharynx cases for reasons that are not explained by the biomarker staining alone. Ang et al (2010) documented at least 30% chance of death at 3 years for HPV-positive patients with positive smoking histories. There is little question that HPV-positive/p16 positive nonsmoking patients have more favourable outcomes. However, in patient populations with high or modest smoking rate, it is still valuable to assess patients' survival beyond HPV status. A biomarker that more precisely captures the biology of both smoking and tumour site, and that unifies the frequent discrepancies between HPV staining and p16 staining would be welcome. Recently, our group reported that p16 staining was prognostic in a set of young patients with HNSCC who were confirmed HPV negative by PCR and in situ hybridisation, (Harris et al, 2010b), leading us to question whether p16 alone could be extended to evaluate risk outside the oropharynx.

Smoking and HPV infection are two important etiologies of p16 alteration in HNSCC. In HPV-infected patients, the protein RB1 is inactivated by viral oncoprotein E7, leading to a high and nuclear localised p16 expression (Andl et al, 1998; Wiest et al, 2002; Li et al, 2004; Marur et al, 2010). In contrast, in situations where p16 is retained but altered in function by mutation or other genetic events, we may still observe modest to high p16 expression, but with abnormal cellular localisation. In many additional smoking patients, p16 can be lost via more deleterious genetic or epigenetic changes, such as homozygous deletion, nonsense mutation, or perhaps methylation and gene silencing. On the basis of these aetiologic differences, we expected to observe distinct patterns in p16 IHC staining. Similar hypotheses of p16's role in prognosis have been tested in other tumour types. For example, in high-grade astrocytoma, a study has shown that nucleus-located p16 is associated with better disease outcome while cytoplasmic p16 indicates worse patients' survival (Arifin et al, 2006). In other tumour types, including endometrial cancers, melanoma and astrocytomas (Emig et al, 1998; Salvesen et al, 2000; Straume et al,

2000; Milde-Langosch et al, 2001; Ghiorzo et al, 2004; Arifin et al, 2006), reports also exist where p16 localisation is associated with disease outcomes. As p16 protein functions as a cell cycle inhibitor in the nucleus, we proposed that nuclear p16 staining and cytoplasmic p16 staining may have a distinct prognostic effect in HNSCC. We tested this hypothesis in a population-based patient cohort – the Carolina Head and Neck Cancer Study (CHANCE).

METHODS

Study population

The CHANCE was a population-based case–control study of incident HNSCC conducted from 2002 to 2006 in 46 counties in Central and Eastern North Carolina (Divaris et al, 2010). The subcohort of 143 patients from this study who were treated at UNC (University of North Carolina at Chapel Hill) hospitals and had banked tissue available were eligible. Patients with cancers of all head and neck subsites except nasopharynx (oral cavity, oropharynx, larynx and hypopharynx) were included. Treatment decisions were recommended by the UNC Head and Neck multidisciplinary team and based on patient age, tumour extent, site, comorbidities and performance status. Clinical information was extracted from patient charts. Patients who received complete medical care at UNC were followed by retrospective review of the medical record for outcomes including relapse and death. Patients who had follow-up in local institutions outside UNC were followed by requesting medical records from the local institution or in cases where there was no return of information from the outside institution, patients deaths were queried from the Social Security Death Index and local obituaries in compliance with the CHANCE study protocols. Patients without sufficient tumour sample for p16 staining were excluded, leaving 135 patients in the analysis. An independent UNC TMA cohort was available for validation that our group has reported on previously (Harris et al, 2010a). Both studies were approved by the Institutional Review Board of the UNC.

Tissue microarray

Tissue microarrays (TMAs) were constructed using core samples from formalin-fixed paraffin-embedded tumour blocks. Hematoxylin- and eosin-stained slides were reviewed by two pathologists (KF and WF) to confirm the original diagnosis. About 1-mm microarray blocks were constructed on a manual tissue microarrayer-1 from Beecher Instruments (Sun Prairie, WI, USA) in triplicate. Sequential 4-μm sections were cut from each TMA. Sectioned slides were coated in paraffin and stored at 4 °C until staining. A second confirmatory tissue resource was also used for the current analysis the construction and results of which have been previously reported (Harris et al, 2010a). Briefly, a TMA (designated young nonsmoking oral cavity cohort (YNOCC)) was constructed in a similar manner as above that included a cohort of 42 HNSCC between the age of 18 and 39. Processing of tissue and reagents is otherwise consistent with the current methods.

p16 IHC staining

p16 IHC staining was carried out in the Bond Autostainer (Leica Microsystems Inc, Norwell, MA, USA) according to the manufacturer's IHC protocol. Slides were put in a 60° oven to remove excess paraffin. Slides were then placed in the autostainer and dewaxed in Bond Dewax solution (AR9222) and hydrated in Bond Wash solution (AR9590). Antigen retrieval was performed for 30 min at 100 °C in Bond-Epitope Retrieval solution 1 (pH 6.0, AR9961). Slides were then incubated with p16^{INK4a} antibody (mouse monoclonal anti-p16 antibody (MAB4133), Chemicon International Company/Millipore Corporation, Temecula, CA,

USA) for 15 min. Antibody detection was performed using the Bond Polymer Refine Detection System (DS9800). Stained slides were dehydrated and coverslips added. Immunohistochemical staining was performed in the Translational Pathology Lab at UNC. After completion of IHC, slides are stored at room temperature in our laboratory and a virtual scanned copy of all TMA slides will be kept for further reference.

HPV *in situ* hybridisation

HPV *in situ* hybridisation was carried out in Ventana Benchmark XT autostainer. Slide deparaffinisation, conditioning, and staining with INFORM HPV III Family 16 Probe (B; Ventana Medical Systems, Tucson, AZ, USA) were performed on the autostainer according to the manufacturer's protocol. The probes have affinities to HPV subtypes 16, 18, 31, 33, 35, 39, 45, 51, 52, 56, 58 and 66. Slides were scored as positive for HPV if a punctate or diffuse pattern of signal was observed in the tumour nuclei.

p16 protein expression

p16 expression was assessed by pathologists who were blinded as to the clinical data for the patients. The CHANCE TMA and the YNOCC TMA were read by two pathologists (KF and LT, respectively), with any indeterminate scores evaluated by a third pathologist (WF). Digital images of cells were captured (magnification × 200) using the Aperio Scanscope (Aperio Inc., Vista, CA, USA). Tissue samples previously shown to be p16 over-expressors (endometrium) were used as a positive control for intensity scoring. Each sample was given a cytoplasmic intensity score and nuclear intensity score on a scale of 0–3, with intensity scored 0 equal to no staining; 1, faint or focal cytoplasmic staining; 2, moderate, diffuse staining; and 3, intense and diffuse staining. The percentage of tumour cells with positive nuclei was determined by scoring 10 microscopic fields of 100 tumour cells each. A semiquantitative percentage score was generated for the cytoplasm and nucleus staining for each specimen, ranging from 0 to 100. The TMA was constructed with the goal to obtain three cores per patient block. Not every block had sufficient tissues and some cases resulted with only one or two cores. For samples that had multiple cores, mean intensity or percentage scores across the cores were used as the final intensity or percentage score for that sample. A composite product score was calculated by multiplying the mean intensity score and mean percentage score in the cytoplasm or nucleus. Based on a bimodal distribution of the scores in oropharynx patients (dark grey in Figure 2), a nuclear product score of 100 was used as a cutoff for nuclear staining. The 75% percentile of cytoplasmic staining (133.4) was considered to be a cutoff for cytoplasmic staining. All samples that had high nuclear (HN) staining also had high cytoplasmic staining, resulting in three categories in total. Patients with a nuclear product score ⩾100 were considered high nuclear staining (HN). Patients with a product score at or above the 75th percentile of the cytoplasmic score (133.4) were considered high cytoplasmic staining (HC) if they were not in the HN group. Patients who failed to meet criteria either for high nuclear or high cytoplasmic score were categorised in the low staining group (LS). Based on this empirical separation, the patients were divided into three groups; high nuclear, high cytoplasmic staining (HN), high cytoplasmic, low nuclear staining (HC) and low nuclear and cytoplasmic staining (LS).

Statistical analysis

All statistical analysis was performed using the R 2.9.2 software (http://cran.r-project.org). Baseline characteristics of patients from each group (HN, HC and LS) were compared using Fisher's exact test for categorical variables and one-way analysis of variance for continuous variables. Overall survival was calculated as the time

from diagnosis date to death date or the last documented follow-up date. Progression-free survival (PFS) was defined as the time from diagnosis date to the date of disease progression or the last documented follow-up date or death date from any cause. Disease progression was defined as any documented tumour progression (local or distant) as indicated in the clinical record. All observations were censored at 60 months. Survival curves were calculated using the Kaplan–Meier method and compared nonparametrically using the log-rank test. Cox proportional hazard model was used to estimate the hazard ratio between different p16 staining groups, adjusting for patient alcohol consumption status, tumour stage, tumour site and HPV staining. All statistical tests were two sided with a significance level of 0.05 and all reported confidence intervals (CIs) were constructed at a two sided 95% confidence level.

RESULTS

Patient characteristics

A total of 143 patients were identified during the study period, of which 135 had sufficient tumour samples for p16 staining. The median follow-up time for these patients was 6.67 years, with only five patients lost to follow-up before 5 years. The baseline characteristics for these patients were summarised in Table 1. The median age of patients at diagnosis was 57 (range 20–82). In all, 68.9% of the patients were males, which is comparable to the national average (Ries *et al*, 2007). Most patients had smoking histories and/or alcohol use with only 12 (9%) never smokers and 44 (~30%) never drinkers. Furthermore, all of the 123 smokers, except 2, had smoked ⩾10 pack years. Approximately 30% of the patients received single-modality treatment with surgery or radiation alone. Other patients received a combination of different treatment methods. In all, 16 (11.9%) patients were detected as HPV positive, of which 14 had oropharyngeal tumours and the other 2 had tumours in the oral cavity.

p16 expression

In the sample set p16 showed baseline cytoplasmic and nuclear staining in at least one of the three cores for every patient. Examples of IHC images of p16 staining are shown in Figure 1. Overall, oropharyngeal cancers and HPV-positive cancers had stronger p16 staining in both the cytoplasm and nucleus compared with tumours of other types (Figure 2). The median nuclear product score was 22 in oropharyngeal tumour samples compared with 0 in non-oropharyngeal samples (permutation test of equal density *P*-value <0.001). The median cytoplasmic product score was 150 in oropharyngeal tumour samples compared with a median product score of 38 in non-oropharyngeal samples (permutation test of equal density *P*-value <0.001). A total of 9 patients had high nuclear and high cytoplasmic p16 staining (HN), 25 patients had HC and 101 had low p16 staining (LS). There was no significant difference in age, gender, smoking status, T stages and clinical stages between different staining groups. However, patients with high nuclear or cytoplasmic p16 staining have more oropharyngeal tumours and earlier nodal stage (N0–N1) compared with the low p16 staining group.

HPV *in situ* hybridisation

Table 2 summarised the distribution of tumour sites with respect to HPV positivity and smoking status. Overall, 16 of the 143 patients stained positively for HPV, with 14 of them having tumours in the oropharynx and 2 in the oral cavity. The HPV positivity rates were lower than some of the clinical trials and other university based reports (Chuang *et al*, 2008; Fakhry *et al*, 2008), due to, at least in part, the very high smoking rate in our study population (D'Souza *et al*, 2007; Ang *et al*, 2010). In all, 37%

Table I Patient characteristics by p16 staining

Characteristics	All patients (column %)	p16 Staining groups			P-values
		HN (column %)	HC (column %)	LS (column %)	
No. of patients	135	9	25	101	
Age					
Median	57	56	54	58	0.14
Range	20–82	20–66	34–79	24–82	
Gender					
Male	93 (68.9)	8 (88.9)	19 (76)	66 (65.3)	0.28
Smoking[a]	123 (91.1)	7 (77.8)	22 (88)	94 (93.1)	0.16
Mean pack years (s.d.)[a]	39.8 (25.9)	41.4 (39.1)	38.0 (26.0)	40.0 (24.7)	0.93
Alcohol[a]	91 (67.4)	6 (66.7)	18 (72)	67 (66.3)	0.90
T stage[a]					
T1–T2	65 (48.1)	4 (44.4)	10 (40)	51 (50.5)	0.65
T3–T4	70 (51.9)	5 (55.6)	15 (60)	50 (49.5)	
Nodal stage[a]					
N0–N1	79 (58.5)	4 (44.4)	8 (32)	67 (66.3)	0.004
N2–N3	56 (41.5)	5 (55.6)	17 (68)	34 (33.7)	
Stage					
Stage I–II	43 (31.9)	1 (11.1)	5 (20)	37 (36.6)	0.12
Site					
Oropharynx	38 (28.1)	7 (77.8)	15 (60)	16 (15.8)	<0.001
Larynx	35	1 (11.1)	1 (4)	33 (2.7)	
Oral cavity	54 (40)	1 (11.1)	6 (24)	47 (46.5)	
Hypopharynx	8 (5.9)	0	3 (12)	5 (5.0)	
HPV positive	16 (11.9)	3 (33.3)	10 (40.0)	3 (3.0)	<0.001

Abbreviations: HC = high cytoplasmic, low nuclear staining; HN = high nuclear, any cytoplasmic staining; LS = low nuclear, low cytoplasmic staining. [a]Numbers do not sum to the total because of missing data.
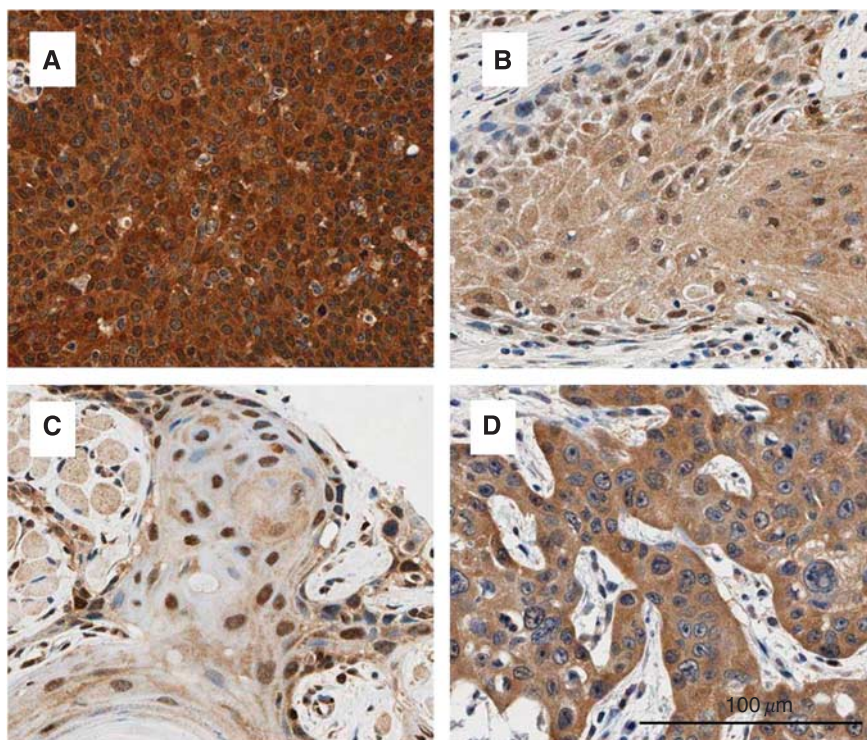

Figure I Representative examples of p16 immunostaining in HNSCC. IHC staining for p16 expression of HNSCC was evaluated by product scores in different cellular compartments separately. From the above left: (**A**) p16 high expression in both the nuclei and cytoplasm; (**B**) p16 low expression in both the nuclei and cytoplasm; (**C**) high nuclear expression and modest cytoplasmic staining (however, by our scoring this still qualified at the lowest end of 'high cytoplasmic'); and (**D**) high cytoplasmic expression and low nuclear expression.
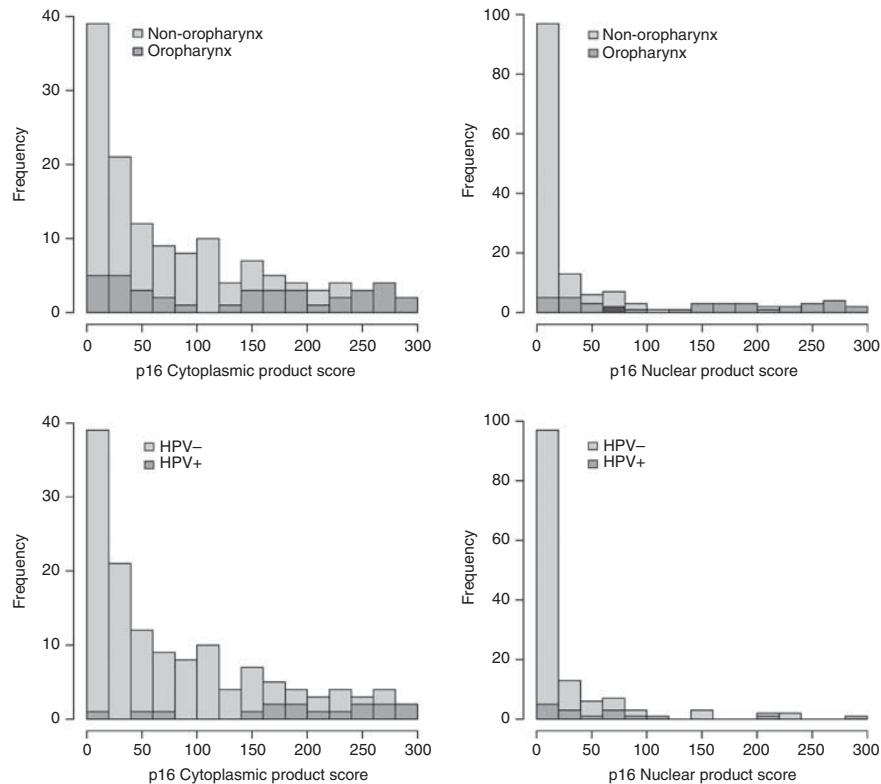
Molecular Diagnostics

Figure 2 Distributions of p16 staining product scores.

Table 2 p16 expression by smoking status and tumour site

p16 Expression	HN	HC	LS
Smokers			
HPV negative	4 (1 OC, 3 OP)	14 (5 OC, 1 LA, 3 HY, 5 OP)	92 (40 OC,33 LA, 3 HY, 16 OP)
HPV positive	3 (OP)	8 (OP)	2 (1 OC, 1OP)
Nonsmokers			
HPV negative	2 (OP)	1(OC)	6 (5 OC, 1 OP)
HPV positive	0	2 (OP)	1 (1 OC)

Abbreviations: HC = high cytoplasmic, low nuclear staining; HN = high nuclear, any cytoplasmic staining; HPV = human papillomavirus; HY = hypopharynx; LA = larynx; LS = low nuclear, low cytoplasmic staining; OC = oral cavity; OP = oropharynx.

(14/38) of oropharyngeal tumours were stained HPV positive in this study, comparable to previous reports such as RTOG 9003 (Gillison *et al*, 2012), which reported 39% HPV positive in oropharyngeal cancers. Human papilloma virus-positive staining outside oropharyngeal tumours was rare, which is consistent with the general acceptance of a low rate of HPV infection outside the oropharynx (Begum *et al*, 2007). The vast majority of these HPV-positive patients were heavy smokers: 13 of the 16 HPV-infected patients had long histories of smoking, with a minimum of 18 pack years. Human papilloma virus infection has been strongly associated with both cytoplasmic and nuclear p16 positivity. All but three HPV-positive patients were categorised as having high nuclear or high cytoplasmic p16 expression.

Survival analysis

In the full cohort, the 3-year OS was 63.0% (95% CI: 55.3%–71.7%) and the 3-year PFS rate was 54.1% (95% CI: 46.3%–63.2%). Only one death occurred in the HN group during the follow-up.

In the LS group, the 3-year OS and PFS was estimated as 65.3% (95% CI: 56.7%–75.3%) and 54.5% (95% CI: 45.6%–65.1%) using the Kaplan–Meier method. The 3-year OS and PFS was estimated as 40% (95% CI: 24.7%–64.6%) and 36% (95% CI: 21.3%–60.7%), respectively, in the HC group (Figure 3). The 3-year OS and PFS in the HN group was 100% with CI not evaluable. Both OS and PFS results were significantly different between staining groups with a log-rank test P-values of 0.006 and 0.009, respectively. There is no significant difference in OS or PFS between the HPV-positive group and HPV-negative group ($P = 0.509$ and 0.434, respectively).

Cox proportional hazard model was used to assess the relationship between each variable with OS and PFS (Table 3). p16 expression status was significantly associated with both OS and PFS. The HN group had the best OS outcome and the lowest hazard ratio compared with the other groups. Similar results were obtained for PFS, although the difference was not statistically significant. Using the HC group as a reference, the hazard ratio for OS was 0.50 (95% CI: 0.29–0.88) for the LS group and 0.10 (95% CI: 0.013–0.75) for the HN group. Similarly, the hazard ratio for PFS was 0.61 (95% CI: 0.35–1.04) in the LS group and 0.09 (95% CI: 0.012–0.67) in the HN staining group. If we consider local recurrence and distant recurrence separately, the 3-year local recurrence rates were 24% and 26.7% for the HC and LS group, respectively, and the 3-year distant recurrence rate was 16.0% and 10.9% for the HC and LS group, respectively. The HN group had no recurrence during 3 years of follow-up. When nuclear staining and cytoplasmic staining were considered separately for their association with OS or PFS, high nuclear staining was significantly associated with PFS (HR = 0.13, 95% CI: 0.018–0.96) and insignificantly associated with OS (HR = 0.17, 95% CI: 0.024–1.24). Cytoplasmic staining was not significantly associated with either OS or PFS. In addition to p16 staining status, T3–T4 tumour stage was significantly associated with increased risk of mortality (P-value = 0.009). Nodal stages showed borderline significance in affecting OS (P-value = 0.07).

A

Overall survival
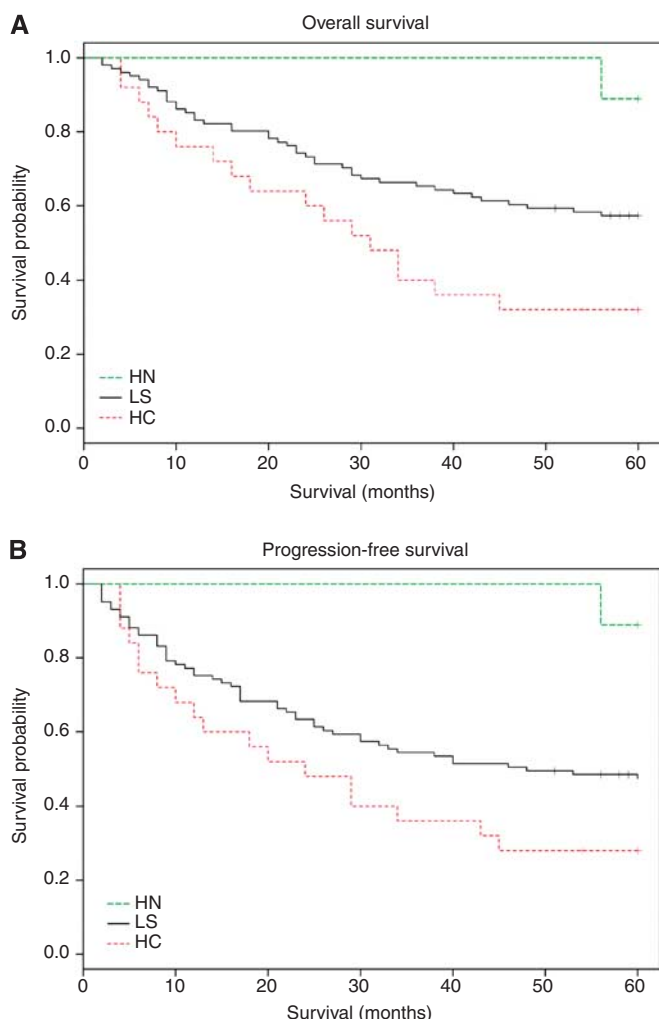

B

Progression-free survival

Figure 3 Kaplan–Meier estimates of overall survival (**A**) and progression free survival (**B**) according to p16 expression in whole study population. All survival estimates were censored at 60 months. Abbreviations: HC = high cytoplasmic, low nuclear staining; HN = high nuclear, any cytoplasmic staining; LS = low nuclear, low cytoplasmic staining. '+' censored observations.

No variable tested except p16 expression status showed significant association with PFS.

Multivariable Cox proportional hazard model showed that p16 expression status was still significantly associated with both OS and PFS (Table 4) after adjusting for tumour site, nodal stage, tumour stage, HPV staining and drinking pattern. Both the LS group and the HN staining group had significantly lower hazard than the HC staining group. Subset analysis was carried out for oropharynx patients: after controlling for tumour stages, HPV staining and drinking status, the hazard ratio of OS for LS and HN groups are 0.40 ($P = 0.18$) and 0.12 ($P = 0.06$), respectively, and the hazard ratio of PFS for LS and HN groups are 0.61 ($P = 0.43$) and 0.12 ($P = 0.06$), respectively, using the HC group as reference. Subset analysis for other tumour sites was not conducted because of the small number of patients.

Independent confirmation in second cohort

Using data from the YNOCC TMA, we were able to obtain p16 staining on an additional 42 samples, with 30 from the oral cavity, 6 from the oropharynx, 5 from the larynx and 1 from the hypopharynx. This is a cohort of younger patients who were

diagnosed between the age of 20 and 39, with 23 males, 29 with smoking history (median pack year 14.5) and 18 with alcohol consumption history. Previously we had reported a favourable overall outcome for those patients in the cohort who were p16 positive. At that time, we had not evaluated the independent contribution of nuclear staining to outcomes. In this study, we evaluated those patients by the same product score cutoff as an independent validation. The patients were then grouped using the same criteria for this study: 14 patients were placed in the HN group, 4 patients in the HC group and 24 patients in the LS group. Although P-values are not statistically significant owing to the small sample size, strikingly, the HN staining group had superior PFS compared with the other two groups, with similar magnitude to our observations in the CHANCE data set. The hazard ratio of having a recurrence in the HN group and LS group are 0.38 (95% CI: 0.092–1.62) and 0.71(95% CI: 0.20–2.52) compared with the HC staining group ($P = 0.34$).

DISCUSSION

The management of squamous cell carcinoma of the head and neck appears to be at a crossroads, with the possibility that the field may change long held treatment standards based on observations related to the staining for the biomarkers HPV and p16. Pivotal studies have documented significantly improved outcomes for patients staining positively for these markers, yet a closer look at how these biomarkers relate to each other has stimulated researchers to look for the mechanisms behind the beneficial outcome association. First, it is clear that mechanisms in addition to HPV infection itself are at work as evidenced by the modulation of risk caused by smoking. There is also at least circumstantial evidence that alterations of p16, independent of HPV, may convey some of the favourable prognoses seen in HNSCC patients that cannot simply be ascribed to false negative HPV assays. Evidence from tumours outside the head and neck lead us to consider nuclear localisation of p16 as a novel biomarker. In this report, the results comparing nuclear localisation of p16 to cases where p16 is excluded from the nucleus warrant further study. Furthermore, the results may help suggest a mechanistic role for this biomarker that go beyond an empiric view of p16 as a proxy for HPV of use limited to the oropharynx.

To consider p16 status (as indicated by p16 staining) as a mechanistic marker requires a review of the ways that p16 is altered in cancer. In the case of HPV, p16 overexpression is a result of expression of HPV-derived oncoproteins E6 and E7 and can functionally inactivate the p53 and pRb tumour-suppressor protein, resulting in a downregulation of p53, pRb and a strong upregulation of p16 at the molecular level (Andl et al, 1998; Wiest et al, 2002; Li et al, 2004; Marur et al, 2010). One could think of p16 expression in the context of HPV infection as a proxy for multiple genotypes that would generally be considered favourable for cancer prognosis (p53 wild type (WT), Rb WT and p16 WT). However, in the more common setting of tumours, p16 is lowly expressed, possibly by less favourable genetic or epigenetic changes, such as homozygous deletion of p16, nonsense mutation, or perhaps methylation and gene silencing. In those situations, where there are more deleterious mutations such as loss of Rb or perhaps amplification of cyclin D1 (common in HNSCC), the tumours can express high levels of p16 with no inhibition of cell cycling. In these situations, nuclear trafficking might be altered and high p16 expression might indicate particularly unfavourable cancer biology. Smoking could be the means of inactivation of genes downstream of p16 without requiring p16 loss as the disease modifying event associated with worse outcome. To evaluate such an explanation, we attempted to sequence p16 and other targets in the current sample set but were unsuccessful because of the quality of the DNA in these paraffin-embedded specimens.

Molecular Diagnostics

Table 3 Univariate analyses of prognostic factors for overall or PFS

Characteristics	PFS					OS				
	No. of events	PYs	HR	95% CI	*P*-value	No. of events	PYs	HR	95% CI	*P*-value
Age (years)										
≥57/<57	38/34	228/197	0.97	0.61–1.55	0.91	33/28	254/224	1.03	0.62–1.70	0.92
Smoker/nonsmoker	66/6	385/41	1.19	0.52–2.74	0.69	56/5	433/45	1.19	0.48–2.97	0.71
Drinking	53/19	270/154	1.55	0.92–2.62	0.10	45/16	305/174	1.61	0.91–2.84	0.10
Site										
Larynx	18	111	1.17	0.61–2.25	0.63	14	134	1.01	0.49–2.10	0.97
Oral cavity	29	166	1.27	0.71–2.29	0.42	26	179	1.40	0.74–2.64	0.30
Hypopharynx	7	15.4	2.74	1.14–6.60	0.02	6	20	2.70	1.04–6.99	0.04
Oropharynx	18	131	1.0 (Reference)			15	145	1.0 (Reference)		
T stage										
T3–T4/T1–T2	42/30	202/221	1.49	0.93–2.38	0.10	39/22	220/258	2.02	1.20–3.41	0.009
N stage										
N2–N3/N0–N1	31/41	165/259	1.16	0.73–1.86	0.52	30/31	177/301	1.61	0.97–2.65	0.07
Stage										
Late(III–IV)/early(I–II)	52/20	281/144	1.29	0.77–2.17	0.33	47/14	309/169	1.79	0.98–3.25	0.06
p16: Combined nuclear and cytoplasmic staining										
HN	1	45	0.09	0.012–0.67	0.067	1	45	0.10	0.013–0.75	0.025
LS	53	319	0.61	0.35–1.04	0.019	43	365	0.50	0.29–0.88	0.017
HC	18	61	1.0 (Reference)			17	68	1.0 (Reference)		
HPV										
±	7/65	55/368	0.73	0.34–1.60	0.44	6/22	60/418	0.75	0.32–1.75	0.51

Abbreviations: CI = confidence interval; HC = high cytoplasmic, low nuclear staining; HN = high nuclear, high cytoplasmic staining; HR = hazard ratio; LS = low nuclear, low cytoplasmic staining; OS = overall survival; PFS = progression-free survival; PYs = person-years.

Table 4 Multivariate analysis of prognostic factors for survival or PFS

	PFS			OS		
	HR	95% CI	*P*-values	HR	95% CI	*P*-values
T3–T4/T1–T2	1.32	0.77–2.26	0.31	1.72	0.94–3.12	0.08
N2–N3/N0–N1	0.96	0.55–1.68	0.90	1.24	0.68–2.28	0.48
Drinking	1.37	0.76–2.49	0.29	1.21	0.64–2.31	0.56
Site						
Larynx	1.29	0.57–2.89	0.54	1.34	0.54–3.30	0.53
Oral cavity	1.35	0.66–2.78	0.41	1.65	0.75–3.60	0.21
Hypopharynx	1.73	0.66–4.56	0.27	1.69	0.59–4.84	0.33
Oropharynx	1 (Reference)			1 (Reference)		
p16 Staining						
HN	0.092	0.01–0.71	0.02	0.10	0.01–0.78	0.03
LS	0.475	0.24–0.95	0.03	0.37	0.18–0.75	0.01
HC	1 (Reference)			1 (Reference)		
HPV positives	0.65	0.24–1.81	0.41	0.54	0.179–1.61	0.27

Abbreviations: CI = confidence interval; HC = high cytoplasmic, low nuclear staining; HN = high nuclear, high cytoplasmic staining; HR = hazard ratio; LS = low nuclear, low cytoplasmic staining; OS = overall survival; PFS = progression-free survival.

To our knowledge, no previous study has investigated how different p16 expression localisation can be related to disease outcomes in HNSCC despite evidence that differential staining patterns similar to what we describe have been shown to be relevant in other tumours, including endometrial cancers, melanoma and astrocytomas (Emig *et al*, 1998; Salvesen *et al*, 2000; Straume *et al*, 2000; Milde-Langosch *et al*, 2001; Ghiorzo *et al*, 2004; Arifin *et al*, 2006). Most strikingly, familial melanoma studies strongly support our hypothesis because of the associated point mutations and the failure to localise p16 to the nucleus (Ghiorzo *et al*, 2004). In this report, patients without the germline

variant displayed a combined nuclear and cytoplasmic staining. The authors demonstrated that p16 mutations in these melanoma patients may impair the cytoplasmic–nuclear shuttling similar to BRCA1 where BRCA1 is shifted to the cytoplasm because of a mutation in nuclear localisation signals and the HN2-terminal (Fabbro *et al*, 2004; Ghiorzo *et al*, 2004; Arifin *et al*, 2006).

This study has limitations and suggests that further evaluation of p16 nuclear staining is warranted. Most notably, the current study is relatively small and includes a large number of smokers. Similarly, because of the retrospective nature of the current study, patients are heterogeneous in stage, site, treatment and other

factors that might impact risk in ways that have not been appreciated. However, the prognostic effect of p16 localisation remained significant after controlling for these factors. The validation cohort provided extra support for our result. We do provide evidence regarding the use of p16 in nonsmokers with the YNOCC cohort, because this group does not include significant numbers of nonsmoking HPV-positive patients. However, because most HNSCC patients are still smokers despite the rising numbers of nonsmoking patients, these data are applicable to a larger portion of HNSSC patients. Finally, our cutoff for different p16

groups was based on the empirically observed distributions of p16 staining in oropharynx vs non-oropharynx samples. This cutoff was neither optimised nor cross-validated and cannot be directly used for clinical settings.

In conclusion, we have provided a preliminary investigation into the nuclear staining of p16 as a critical factor in the complex set of conditional biomarkers including HPV, smoking, oropharyngeal carcinomas and nonlocalised staining of p16. This biomarker, if validated, is already widely available and could potentially impact clinical care of HNSCC.

REFERENCES

Andl T, Kahn T, Pfuhl A, Nicola T, Erber R, Conradt C, Klein W, Helbig M, Dietz A, Weidauer H, Bosch FX (1998) Etiological involvement of oncogenic human papillomavirus in tonsillar squamous cell carcinomas lacking retinoblastoma cell cycle control. Cancer Res 58: 5–13

Ang KK, Harris J, Wheeler R, Weber R, Rosenthal DI, Nguyen-Tan PF, Westra WH, Chung CH, Jordan RC, Lu C, Kim H, Axelrod R, Silverman CC, Redmond KP, Gillison ML (2010) Human papillomavirus and survival of patients with oropharyngeal cancer. N Engl J Med 363: 24–35

Arifin MT, Hama S, Kajiwara Y, Sugiyama K, Saito T, Matsuura S, Yamasaki F, Arita K, Kurisu K (2006) Cytoplasmic, but not nuclear, p16 expression may signal poor prognosis in high-grade astrocytomas. J Neurooncol 77: 273–277

Begum S, Gillison ML, Nicol TL, Westra WH (2007) Detection of human papillomavirus-16 in fine-needle aspirates to determine tumor origin in patients with metastatic squamous cell carcinoma of the head and neck. Clin Cancer Res 13: 1186–1191

Chaturvedi AK, Engels EA, Pfeiffer RM, Hernandez BY, Xiao W, Kim E, Jiang B, Goodman MT, Sibug-Saber M, Cozen W, Liu L, Lynch CF, Wentzensen N, Jordan RC, Altekruse S, Anderson WF, Rosenberg PS, Gillison ML (2011) Human papillomavirus and rising oropharyngeal cancer incidence in the United States. J Clin Oncol 29(32): 4294–4301

Chuang AY, Chuang TC, Chang S, Zhou S, Begum S, Westra WH, Ha PK, Koch WM, Califano JA (2008) Presence of HPV DNA in convalescent salivary rinses is an adverse prognostic marker in head and neck squamous cell carcinoma. Oral Oncol 44: 915–919

Curado MP, Hashibe M (2009) Recent changes in the epidemiology of head and neck cancer. Curr Opin Oncol 21: 194–200

D'Souza G, Kreimer AR, Viscidi R, Pawlita M, Fakhry C, Koch WM, Westra WH, Gillison ML (2007) Case-control study of human papillomavirus and oropharyngeal cancer. N Engl J Med 356: 1944–1956

Dahlstrand H, Dahlgren L, Lindquist D, Munck-Wikland E, Dalianis T (2004) Presence of human papillomavirus in tonsillar cancer is a favourable prognostic factor for clinical outcome. Anticancer Res 24: 1829–1835

Divaris K, Olshan AF, Smith J, Bell ME, Weissler MC, Funkhouser WK, Bradshaw PT (2010) Oral health and risk for head and neck squamous cell carcinoma: the Carolina Head and Neck Cancer Study. Cancer Causes Control 21: 567–575

El-Mofty SK, Lu DW (2003) Prevalence of human papillomavirus type 16 DNA in squamous cell carcinoma of the palatine tonsil, and not the oral cavity, in young patients: a distinct clinicopathologic and molecular disease entity. Am J Surg Pathol 27: 1463–1470

Emig R, Magener A, Ehemann V, Meyer A, Stilgenbauer F, Volkmann M, Wallwiener D, Sinn HP (1998) Aberrant cytoplasmic expression of the p16 protein in breast cancer is associated with accelerated tumour proliferation. Br J Cancer 78: 1661–1668

Fabbro M, Savage K, Hobson K, Deans AJ, Powell SN, McArthur GA, Khanna KK (2004) BRCA1-BARD1 complexes are required for p53Ser-15 phosphorylation and a G1/S arrest following ionizing radiation-induced DNA damage. J Biol Chem 279: 31251–31258

Fakhry C, Westra WH, Li S, Cmelak A, Ridge JA, Pinto H, Forastiere A, Gillison ML (2008) Improved survival of patients with human papillomavirus-positive head and neck squamous cell carcinoma in a prospective clinical trial. J Natl Cancer Inst 100: 261–269

Franceschi S, Munoz N, Bosch XF, Snijders PJ, Walboomers JM (1996) Human papillomavirus and cancers of the upper aerodigestive tract: a review of epidemiological and experimental evidence. Cancer Epidemiol Biomarkers Prev 5: 567–575

Furniss CS, McClean MD, Smith JF, Bryan J, Nelson HH, Peters ES, Posner MR, Clark JR, Eisen EA, Kelsey KT (2007) Human papillomavirus 16 and head and neck squamous cell carcinoma. Int J Cancer 120: 2386–2392

Ghiorzo P, Villaggio B, Sementa AR, Hansson J, Platz A, Nicolo G, Spina B, Canepa M, Palmer JM, Hayward NK, Bianchi-Scarra G (2004) Expression and localization of mutant p16 proteins in melanocytic lesions from familial melanoma patients. Hum Pathol 35: 25–33

Gillison ML, Zhang Q, Jordan R, Xiao W, Westra WH, Trotti A, Spencer S, Harris J, Chung CH, Ang KK (2012) Tobacco smoking and increased risk of death and progression for patients with p16-positive and p16-negative oropharyngeal cancer. J Clin Oncol 30(17): 2102–2111

Ha PK, Pai SI, Westra WH, Gillison ML, Tong BC, Sidransky D, Califano JA (2002) Real-time quantitative PCR demonstrates low prevalence of human papillomavirus type 16 in premalignant and malignant lesions of the oral cavity. Clin Cancer Res 8: 1203–1209

Hafkamp HC, Manni JJ, Haesevoets A, Voogd AC, Schepers M, Bot FJ, Hopman AH, Ramaekers FC, Speel EJ (2008) Marked differences in survival rate between smokers and nonsmokers with HPV 16-associated tonsillar carcinomas. Int J Cancer 122: 2656–2664

Harris SL, Kimple RJ, Hayes DN, Couch ME, Rosenman JG (2010a) Never-smokers, never-drinkers: unique clinical subgroup of young patients with head and neck squamous cell cancers. Head Neck 32: 499–503

Harris SL, Thorne LB, Seaman WT, Neil Hayes D, Couch ME, Kimple RJ (2010b) Association of p16(INK4a) overexpression with improved outcomes in young patients with squamous cell cancers of the oral tongue. Head Neck 33(11): 1622–1627

Jemal A, Siegel R, Xu J, Ward E (2010) Cancer Statistics, 2010. CA Cancer J Clin 60(11): 277–300

Li W, Thompson CH, Cossart YE, O'Brien CJ, McNeil EB, Scolyer RA, Rose BR (2004) The expression of key cell cycle markers and presence of human papillomavirus in squamous cell carcinoma of the tonsil. Head Neck 26: 1–9

Marur S, D'Souza G, Westra WH, Forastiere AA (2010) HPV-associated head and neck cancer: a virus-related cancer epidemic. Lancet Oncol 11(8): 781–789

Milde-Langosch K, Bamberger AM, Rieck G, Kelp B, Loning T (2001) Overexpression of the p16 cell cycle inhibitor in breast cancer is associated with a more malignant phenotype. Breast Cancer Res Treat 67: 61–70

National Cancer Institute (2005) http://www.cancer.gov/cancertopics/factsheet/sites-types/head-and-neck

Patel SC, Carpenter WR, Tyree S, Couch ME, Weissler M, Hackman T, Hayes DN, Shores C, Chera BS (2011) Increasing incidence of oral tongue squamous cell carcinoma in young white women, age 18 to 44 years. J Clin Oncol 29: 1488–1494

Reimers N, Kasper HU, Weissenborn SJ, Stutzer H, Preuss SF, Hoffmann TK, Speel EJ, Dienes HP, Pfister HJ, Guntinas-Lichius O, Klussmann JP (2007) Combined analysis of HPV-DNA, p16 and EGFR expression to predict prognosis in oropharyngeal cancer. Int J Cancer 120: 1731–1738

Ries LAG, Young JL, Keel GE, Eisner MP, Lin YD, Horner M-J (eds) (2007) SEER Survival Monograph: Cancer Survival Among Adults: US SEER Program, 1988-2001. Patient and tumor characteristics. SEER Program, NIH Pub. no. 07-6215. National Cancer Institute: Bethesda, MD

Salvesen HB, Das S, Akslen LA (2000) Loss of nuclear p16 protein expression is not associated with promoter methylation but defines a subgroup of aggressive endometrial carcinomas with poor prognosis. Clin Cancer Res 6: 153–159

Schache AG, Liloglou, T, Risk, JM, Filia A, Jones TM, Sheard J, Woolgar JA, Helliwell TR, Triantafyllou A, Robinson M, Sloan P, Harvey-Woodworth C, Sisson D, Shaw RJ (2011) Evaluation of human papilloma virus diagnostic testing in oropharyngeal squamous

cell carcinoma: sensitivity, specificity, and prognostic discrimination. *Clin Cancer Res* **17**: 6262–6271

Schantz SP, Yu GP (2002) Head and neck cancer incidence trends in young Americans, 1973-1997, with a special analysis for tongue cancer. *Arch Otolaryngol Head Neck Surg* **128**: 268–274

Shiboski CH, Schmidt BL, Jordan RC (2005) Tongue and tonsil carcinoma: increasing trends in the US population ages 20–44 years. *Cancer* **103**: 1843–1849

Shroyer KR, Greer Jr RO (1991) Detection of human papillomavirus DNA by in situ DNA hybridization and polymerase chain reaction in premalignant and malignant oral lesions. *Oral Surg Oral Med Oral Pathol* **71**: 708–713

Stevens TM, Caughron SK, Dunn ST, Knezetic J, Gatalica Z (2011) Detection of high-risk HPV in head and neck squamous cell carcinomas:

comparison of chromogenic in situ hybridization and a reverse line blot method. *Appl Immunohistochem Mol Morphol* **19**(6): 574–578

Straume O, Sviland L, Akslen LA (2000) Loss of nuclear p16 protein expression correlates with increased tumor cell proliferation (Ki-67) and poor prognosis in patients with vertical growth phase melanoma. *Clin Cancer Res* **6**: 1845–1853

Termine N, Panzarella V, Falaschini S, Russo A, Matranga D, Lo Muzio L, Campisi G (2008) HPV in oral squamous cell carcinoma vs head and neck squamous cell carcinoma biopsies: a meta-analysis (1988-2007). *Ann Oncol* **19**: 1681–1690

Wiest T, Schwarz E, Enders C, Flechtenmacher C, Bosch FX (2002) Involvement of intact HPV16 E6/E7 gene expression in head and neck cancers with unaltered p53 status and perturbed pRb cell cycle control. *Oncogene* **21**: 1510–1517

Molecular Diagnostics