Human papillomavirus infections and Upper aero-digestive tract cancers: the ARCAGE study

Authors:

Devasena Anantharaman[1], Tarik Gheit[2], Tim Waterboer[3], Behnoush Abedi-Ardekani[1], Christine Carreira[4], Sandrine McKay-Chopin[2], Valerie Gaborieau[1], Manuela Marron[5], Pagona Lagiou[6], Wolfgang Ahrens[7,8], Ivana Holcátová[9], Franco Merletti[10], Kristina Kjaerheim[11], Renato Talamini[12], Lorenzo Simonato[13], Xavier Castellsague[14], Tatiana V. Macfarlane[15], Gary Macfarlane[15], Nalin Thakker[16], Ariana Znaor[17], Peter Thomson[18], Cristina Canova[19], David I. Conway[20,21], Claire M. Healy[22], Massimo Tommasino[2], Michael Pawlita[3], Paul Brennan[1*]

Author affiliations

1 Genetic Epidemiology group, International Agency for Research on Cancer, Lyon, France

2 Infections and Cancer Biology group, International Agency for Research on Cancer, Lyon, France

3 Division of Genome Modifications and Carcinogenesis, German Cancer Center (DKFZ), Heidelberg, Germany.

4 Molecular Pathology group, International Agency for Research on Cancer, Lyon, France

5 Institute of Medical Biostatistics, Epidemiology and Informatics (IMBEI), University of Mainz, Germany

6 Department of Hygiene, Epidemiology and Medical Statistics, University of Athens Medical School, Athens, Greece

7 BIPS- Institute for Epidemiology and Prevention Research, Bremen, Germany

8 Institute for Statistics, University Bremen, Germany

9 Charles University of Prague, 1st Faculty of Medicine, Institute of Hygiene and Epidemiology, Prague, Czech Republic

10 Unit of Cancer Epidemiology, CeRMS and University of Turin, Turin, Italy

10 Cancer Registry of Norway, Oslo, Norway

12 Unit of Epidemiology and Biostatistics, National Cancer Institute, IRCCS, Aviano, Italy

13 Department of Environmental Medicine and Public Health, University of Padova, Padova, Italy

14 Institut Català d'Oncologia, RTICC/CIBERESP, IDIBELL Hospitalet de Llobregat, Barcelona, Spain

15 School of Medicine and Dentistry, University of Aberdeen, Aberdeen, UK

16 School of Dentistry, University of Manchester, UK

17 Croatian National Cancer Registry, Croatian National Institute of Public Health, Zagreb, Croatia

18 Dental School, University of Newcastle on Tyne, Newcastle, UK

19 Respiratory Epidemiology and Public Health, National Heart and Lung Institute, Imperial College, London UK

20, 21 Dental School, University of Glasgow, Glasgow, UK; NHS NSS, ISD, Edinburgh, UK

22 Trinity College School of Dental Science, Dublin, Ireland

*Corresponding Author: Paul Brennan

Abstract

Background: Human Papillomavirus is causally implicated in a subset of cancers of the upper aero-digestive tract (UADT).

Methods: Associations between type-specific HPV antibodies were examined among 1496 UADT cancer cases and 1425 controls by estimating odds ratios in logistic regression analyses adjusted for potential confounders. Serological associations were validated in a tumor subset using HPV genotyping and p16 overexpression to identify active HPV infections.

Results: HPV16 L1 seropositivity was associated with increased risk of oral cavity and oropharyngeal cancer (OR= 1.94, 95% CI= 1.03- 3.65, OR= 8.60, 95% CI= 5.21- 14.20 respectively). In particular, HPV16 E6 antibodies were present in 30.2% of oropharyngeal cases and only 0.8% of controls (OR= 132.0, 95% CI= 65.29- 266.86). Combined seropositivity to HPV16 E6 and E7 was rare (1 of 1425 controls). HPV16 E6 serology could identify active HPV infection with 67% concordance. A HPV16 independent association was observed for HPV18 and oropharyngeal cancer (OR= 8.14, 95% CI= 2.21- 29.99 for HPV18 E6 seropositivity), and HPV6 and laryngeal cancer (OR= 3.25, 95% CI= 1.46- 7.24 for HPV6 E7 seropositivity).

Conclusions: These results confirm an important role for HPV16 infection in oropharyngeal cancer. HPV16 E6 antibodies are a promising marker to identify HPV16 related tumors, and indicate that at least 30% of oropharyngeal cancers are HPV16 related. These data indicate that HPV16 E6 serology could potentially replace HPV DNA testing in the identification of HPV related oropharyngeal cancer. Additionally, these results support a role for HPV18 in oropharyngeal cancer and HPV6 in laryngeal cancer.

Keywords: Human papillomavirus, upper aero-digestive tract cancer, multiplex serology, HPV16, HPV DNA, p16

Introduction

Cancers of the upper aero-digestive tract (UADT), comprising the oral cavity, pharynx, larynx and esophagus, contribute to over a million new cases each year worldwide. Annually, over 700,000 die of the disease (1). Tobacco and alcohol are known risk factors and HPV infection has been implicated in a subset of UADT cancers (2, 3). The high risk types, HPV16 and HPV18, have been commonly identified (>90% of HPV positive tumors), consistent with observations in cervical and other ano-genital cancers (4-6). However, unlike cervical cancer, not all HPV infections in UADT cancer are transcriptionally active.

The epidemiological, molecular and mechanistic association of HPV16 and UADT cancer is strongest for the oropharynx (7-11). Even so, large variation in HPV16 DNA prevalence (8-100%) is observed, possibly due to differences in study population, composition of tumor subsite, proportion of other known risk factors including smoking and alcohol, type of specimen assayed, and assay variability (5, 12). Conversely, HPV18 appears to be rare in oropharyngeal cancers (5). Recently developed serological methods simultaneously detect type-specific antibodies to multiple HPV proteins, including early viral oncoproteins E6 and E7 that are considered markers of ongoing HPV-related malignancy. The studies published so far have consistently reported an association between presence of HPV16 antibodies and risk of oropharyngeal cancer (4, 7, 13-15). Even so, the utility of serological markers in identifying biologically active HPV infections is poorly understood. Further, the contribution of HPV infections to non-oropharyngeal sites and particularly, of other mucosal HPV types remains unclear.

Using a large panel of markers of HPV infection in a large case- control study, we aimed to: (i) comprehensively evaluate the association between serological markers of oncogenic HPV infection and UADT cancer, (ii) estimate how this varies by subsite, (iii) validate the important serological associations using a subset of tumor biopsies, and (iv) clarify the true proportion of active HPV16 infections.

Materials and Methods

Study population

1496 cases and 1425 controls from the Alcohol-Related CAncers and GEnetic Susceptibility in Europe (ARCAGE) study were included in the present study. Details of the study have been described previously (16). Briefly, 2304 cases and 2227 controls were recruited across 10 European countries during 2002- 2005, using a standardized protocol in all centers (except France). Cases were histologically or cytologically confirmed primary cancers of the oral cavity, oropharynx, hypopharynx, larynx, esophagus and overlapping sites. A comparable group of hospital or population based controls were recruited in each center, frequency matched for age, gender and area of residence. All cases and controls underwent personal interviews to record lifestyle exposures. The study was approved by ethical review boards at the participating centers and by the IARC ethical review committee.

Laboratory methods

Serological methods: Plasma samples from 1496 cases and 1425 controls were tested for type-specific HPV antibodies using bead-based multiplex serology method described elsewhere (17). We report associations on 27 markers of mucosal HPV infection, including high risk types: HPV16 (L1, E1, E2, E4, E6, E7), HPV18 (L1, E6, E7), HPV31 (L1, E6, E7), HPV33 (L1, E6, E7), HPV45 (L1, E6, E7), HPV52 (L1, E6, E7) and low risk types: HPV6 (L1, E6, E7), HPV11 (L1, E6, E7). Additionally, we tested antibodies to cutaneous HPV types (HPV1, HPV5, HPV8 and HPV38) and non-HPV related antibodies (P53, P16, JCV, HHV and EBV) as specificity controls. Serology data were generated as continuous measures of mean fluorescence intensity which were dichotomized using cut-offs derived from earlier studies (18, 19). Briefly, the ratio of predefined and extrapolated cut offs in a bridging panel of sera were used for normalization and cut offs were set at mean plus 3 standard deviations.

Tumor tissue analyses: 150 snap- frozen tumor tissues were identified that included 125 cases with a high *a priori* expectation for HPV infection based on serology (all HPV16 L1, E6 and E7 positives), tumor site (all cancers of the oropharynx and overlapping sites) and other characteristics (all women, young male never smokers) and 25 cases with a low *a priori* expectation included cancers of the oral cavity or larynx among male smokers. 30 cases were excluded based on pathological evaluation that aimed to confirm tumor histology and record cellular features. The p16 expression was qualitatively evaluated using the CINtec Histology $P16^{INK4a}$ Kit (9511, mtmlabs) following manufacturer's instructions. Expression was scored based on the percentage and intensity of nuclear or cytoplasmic staining. A combined score of 4 or greater was considered positive for $p16^{INK4a}$ overexpression (20). DNA extraction from biopsies was performed using the Qiagen BioRobot EZ1 (Qiagen, Hilden, Germany). HPV genotyping using the Type-Specific E7 PCR bead-based multiplex assay (TS-E7-MPG) was performed to detect all high risk HPV types (HPV16, 18, 26, 31, 33, 35, 39, 45, 51, 52, 53, 56, 58, 59, 66, 68a, 68b, 73, 82) and three low risk HPV types (HPV6, 11, 70) (21-23). Briefly, the reporter fluorescence was quantified using Luminex reader 200 and cutoffs were computed by adding 5 to 1.1 x the median background value expressed as median fluorescence intensity (MFI).

Statistical analysis

The association between serological markers of HPV infection and UADT cancer was examined by calculating the odds ratios and corresponding 95% confidence intervals using unconditional logistic regression models adjusted for age, sex, level of education (finished primary school, finished secondary school or university degree), pack years of tobacco smoking (never, <20, 20- 39, 40-59, 60-79 and ≥80), number of alcoholic drinks consumed per day (never, <1, 1-2, 3-4, 5-6, ≥7) and country. To examine the effect modification by smoking in the association of HPV16 seropositivity and UADT cancer, we included an appropriate interaction term in the model. Since homologous proteins of closely related HPV types could potentially cross-react, we performed antigen specific sensitivity analysis following exclusion of individuals positive for phylogenetically related HPV types (HPV16, HPV31, HPV33 and HPV52 from the alpha7 species, HPV18 and HPV45 from alpha9 species and HPV6 and HPV11 from alpha10 species). We assessed the concordance between serological and tumor markers of HPV16 infection based on a previous algorithm

(24). Active HPV16 infections were defined as cases positive for HPV16 DNA and overexpressing surrogate marker p16. Inactive infections were HPV16 DNA positive/p16 negative and HPV unrelated tumors were all HPV16 DNA negatives. The HPV16 attributable fraction was determined based on serology using the formula: p (ec) x ((OR-1)/OR), where p (ec) represents the proportion of exposed (seropositive proportion for type specific marker) among cases and OR were derived from logistic regression model adjusted for age, sex, level of education and country (25, 26). Statistical significance was set at P<0.05. All statistical analyses were performed using STATA statistical software, version 11 and all reported p-values are two sided.

Results

Table 1 shows the characteristics of the study population. Of the 1496 cases, 24% were diagnosed with oral cavity cancer, 22% oropharyngeal cancer (123 were tonsillar), 35% laryngeal/ hypopharyngeal cancer and 13% esophageal cancer. Compared to controls, cases attained lower levels of education and were more often smokers. As expected, smoking and alcohol consumption were strong risk factors for UADT cancer. A clear dose-response relationship was observed between increasing pack years of smoking, number of alcoholic drinks consumed per day and the risk of UADT cancer overall and for each subsite (data not shown). Cancer stage was ascertained for 81% of cases.

HPV16 antibodies and UADT cancer

Figure1 shows the association between HPV16 L1, E6, E7 antibodies and the risk of UADT cancer. HPV16 L1 antibodies were detected in 2.6% of controls, 5.6% of all UADT cancer cases, 4.5% of oral cavity cancers, 13.7% of oropharyngeal cases and were associated with risk of cancer at these sites (OR= 2.92, 95% CI= 1.91- 4.47; OR= 1.94, 95% CI= 1.03- 3.65 and OR= 8.60, 95% CI= 5.21- 14.20, respectively). HPV16 L1 antibody associated UADT cancer risk estimates were higher among never smokers compared to former or current smokers (p- heterogeneity= 0.01). Anti-HPV16 E6 antibodies were rare in the general population (0.8%) and were associated with an 18 fold increased UADT cancer risk (OR= 18.44, 95% CI= 9.72- 34.98). The observed association was consistent across gender and smoking groups. Analysis stratified by tumor site reflected the highest risk for oropharyngeal cancer (OR= 132.00, 95% CI= 65.29- 266.86). Anti-HPV16 E7 antibodies were also associated with UADT cancer (OR= 2.86, 95% CI= 2.06- 3.96) and conferred higher risk estimates among never and former smokers compared to current smokers (p- heterogeneity <0.001). Analysis by tumor site revealed that the associations were driven by oropharyngeal cases (OR= 9.00, 95% CI= 6.06- 13.36 for oropharyngeal cancer; OR= 1.39, 95% CI= 0.95-2.03 for non-oropharyngeal sites, supplementary table5). Combined seropositivity to HPV16 E6 and E7 was very rare among controls (1/1425) and was associated with increased risk of UADT cancer at all sites, with a >800-fold increased risk of oropharyngeal cancer. Additionally, HPV16 E1, E2 and E4 antibodies were examined. HPV16 E1 and E2 antibodies were associated with increased UADT cancer risk, specifically oropharyngeal cancer (table2 and supplementary table1, respectively). HPV16 E2 antibodies were associated with esophageal cancer in this study (OR= 3.31, 95% CI= 1.47- 7.43, data not shown).

Non-HPV16 antibodies and UADT cancer

Antibodies to high risk types HPV18 (L1, E6 and E7), HPV31 (L1 and E7), HPV33 (L1, E6 and E7), HPV45 (L1, E6 and E7), and HPV52 (L1 and E7) and low risk type HPV11 L1 were associated with oropharyngeal cancer (table2). To test for a HPV16-independent association, if any, we excluded all HPV16 L1, E6 and E7 seropositives. While the HPV11 L1 association was consistent, only the associations between HPV18 L1 and E6 remained robust (OR= 2.37, 95% CI= 1.06- 5.32 and OR= 8.14, 95% CI= 2.21- 29.99, respectively). We found associations for HPV6 (L1 and E7) and UADT cancer that appeared to be driven by laryngeal cases. A HPV16 independent effect was observed for HPV6 E7 and laryngeal cancer (supplementary table2).

Sensitivity analysis

The observed results were insensitive to varying definitions of seropositivity, either as continuous data or following doubling of calculated thresholds (data not shown). Additionally, the important associations, including HPV16, HPV18 and HPV6, did not change substantially upon exclusion of phylogenetically related, homologous and potentially cross reacting proteins (supplementary table3).

Tumor tissue analysis

Since HPV16 antibodies formed the principal associations observed in the present study and the results appeared to be driven by the oropharyngeal cases, we prioritized a subset of 120 tumors for HPV genotyping and p16 expression to establish the concordance between HPV serology and HPV16 DNA in the corresponding tumor. We identified 47 tumors that were positive for any HPV DNA (39.2%). Of these, 44 were positive for HPV16 (93.6%), 2 for HPV31 (4.2%), 6 for HPV33 (12.8%), 2 for HPV35 positive (4.2%) and one low risk HPV66 positive (2.1%) (Figure2). Eight multiple infections involving HPV16 were identified including HPV31 (n=2) and HPV33 (n=6). Three non-HPV16 positive cases were identified (two HPV35 and one HPV66). The largest proportion of HPV16 positives were observed for oropharynx followed by oral cavity, larynx, esophagus and overlapping topologies (43%, 32%, 11%, 7% and 7% respectively, data not shown). An algorithm for detecting active HPV infection in head and neck cancer tissue has been proposed that includes an initial test for p16 overexpression followed by HPV DNA detection. Only cases positive at both stages are judged to have an active HPV tumor (24). In all, 9 tumors overexpressed the p16 protein, 7 were cancers of the oropharynx, and one each of larynx and esophagus. Of the nine p16 positive tumors, 6 were concurrently positive for HPV16 DNA indicating an active HPV tumor; all six were oropharyngeal cancers. 37 cases were negative for p16, although positive for HPV DNA, indicating the presence of inactive HPV infection. We observed a strong correlation with HPV16 E6 serology and measures of active HPV infection. Seropositivity was observed in 4 of six active infections (figure 3, supplementary table 4). Based on HPV16 E6 serology, 29.5% of oropharyngeal cancers can be attributed to HPV16 infection. Confirmation of active HPV infection was provided in 17% of oropharyngeal cancers, none of the cancers at other sites of UADT appeared to be HPV16 related (table3).

Discussion

In this large case-control study, we examined the associations between 27 serological markers of mucosal HPV infection and the risk of UADT cancer. Among the various HPV types assessed, strong associations were observed between HPV16 antibodies and oropharyngeal cancer. HPV16 E6 seropositivity was strongly correlated with active HPV infection. Additionally, we found associations for HPV18 antibodies and oropharyngeal cancer, and HPV6 and laryngeal cancer.

HPV16 L1 antibodies are considered markers of previous exposure (27, 28). In interpreting the association between L1 antibodies and cancer, it is important to note that capsid seropositivity represents a mixed group of current and past infections in a subset of individuals who seroconvert. It is interesting that even so, such antibodies are consistently associated with oropharyngeal cancer (4, 7, 14, 15). Our results on HPV16 seroprevalence are comparable with that of Herrero et al (4). It can be argued that the presence of capsid antibodies may reflect systemic exposure from any mucosal HPV infection. This is however, unlikely given that we observed consistent associations across gender and included first primary cancers of the UADT. The risk estimates for HPV16 L1 related cancer was higher among never smokers supporting the notion that HPV-related and smoking-related cancers of the UADT follow distinct etiologies (8). HPV16 E6 and E7 antibodies are regarded as markers of current HPV related malignancy and low E6, E7 antibody prevalence in the general control population in previous studies support this view (4, 7, 14, 15). Consistently, HPV16 E6 seroprevalence was rare (0.8%) in this study, while HPV16 E7 was more common (4.6%), possibly due to assay limitations. In cervical cancer, high levels of E6 and E7 antibodies are associated with advanced stage of the disease (29). In this study, presence of E6 and/or E7 antibodies were associated with late stage at diagnosis of UADT cancer and loco-regional metastasis. A similar trend was observed for oropharyngeal cancer (Supplementary table6). A distinct profile has been described for HPV positive oropharyngeal cases (8). Based on HPV16 E6 seropositivity, we observed a similar risk profile; HPV16 E6 seropositive UADT cases were more likely to be men, never smokers and light drinkers (data not shown). Contrarily, antibodies to E1, E2 and E4 are less well understood. E1 and E2 proteins are expressed in episomal viral infection and often disrupted during viral integration into the host genome (30). This is believed to result in the loss of E6/ E7 repression; an essential and important event in the course of evolution of a HPV related malignancy. In support, serum antibody levels against HPV16 E2 were increased among patients diagnosed with cervical *carcinoma in situ* compared to controls but not in frank malignancies (31). In the present study, such antibodies were associated with oropharyngeal cancer risk. These results are supported by a previous study involving 40 oropharynx cases and 50 cancer free controls (32).

Several studies have examined the association between HPV16 and esophageal cancer; however the results have been inconsistent (33, 34). In the present study, we observed inconsistent associations based on HPV16 serology. Seropositivity to L1 and E4 showed a trend towards reduced risk, E1 and E7 appeared to increase risk while E2 and E6 were associated with increased risk (data not shown). Even though 3 cases were HPV16 DNA positive, p16 overexpression was not observed, implying inactive infections. Notably, all three HPV16 DNA positive esophageal cancers were seronegative for all HPV16 antibodies examined. Based on these results, we conclude that although

HPV16 infections are common (37.5%), any contribution to esophageal squamous carcinoma is unlikely. The observed site-specific differences in the association of HPV16 markers could reflect differences in viral load, viral states (episomal or integrated), site-specific immune differences or true non-causal associations. Even though HPV16 DNA was observed in non-oropharyngeal sites of UADT, none of the tumors overexpressed p16 protein, indicating inactive infections or HPV unrelated etiology, such as smoking. Further, among HPV16 DNA positive tumors, the proportion of HPV16E seropositivity (E1, E2, E4, E6 and E7) was lower for non-oropharyngeal cases compared to cancers of the oropharynx although the proportion of L1 seropositivity did not differ, indicating that although infection rates at heterogeneous UADT sites are likely to be similar, HPV16 related cancer rates are lower for non-oropharyngeal sites. These conclusions are supported by serological associations for HPV16 that appear to principally driven by oropharyngeal cases.

Among other mucosal HPV antibodies examined, we observed strong associations for HPV18, particularly L1 and E6 antibodies and oropharyngeal cancer risk. Antibody cross-reactivity did not appear to influence these associations. Although no HPV18 DNA was identified in this case series, it is important to note that the series was preferentially selected based on HPV16 serology. We conclude that the contribution of HPV18 to UADT cancer though unresolved, is likely to be limited since no associations were observed for non-oropharyngeal sites, and the HPV18 related oropharyngeal cancer fraction was small (7%). We did not observe a HPV16-independent association for HPV31, HPV33, HPV45, HPV52 antibodies and oropharyngeal cancer. Sensitivity analyses performed following exclusion of homologous proteins of related HPV types were inconclusive due to limited power. Even though the case series was not selected to represent these HPV types, of the 120 tumors analyzed, we did not identify any HPV45 or HPV52 DNA. Given the low seroprevalence of these markers (1% for HPV45 E6, E7 and 2% for HPV52), the contribution of these types to oropharyngeal cancer, if any, is likely to be small. Serological analysis indicated a 7-fold increased risk of UADT cancer with HPV31 E7, albeit in the presence of HPV16. This is supported by tumor analyses where we found both the HPV31 DNA positive cases were concomitantly HPV16 positive. Even so, the correlation between HPV31 E7 serology and HPV31 positive tumor was moderate (50%, data not shown). Similarly, all six HPV33 DNA positive tumors were concurrently positive for HPV16 as indicated by serology. Again, the type specific correlation of HPV33 E6 and E7 antibodies and HPV33 DNA was poor (17% and 29% respectively). Given that serology supports a role of 5% of HPV31 and almost 6% of HPV33 in oropharyngeal cancer, the contribution of these types warrants further investigation. The causal association, if any, will be difficult to disentangle given the concurrence with HPV16. Low risk mucosal HPV types such as HPV6 and HPV11 are associated with benign laryngeal papillomas, a rare disease that occasionally undergo malignant transformation (35). In this study, HPV6 L1 and E7 were associated with laryngeal cancer independent of HPV16 serology. The proportion of HPV6 related cancers (E6 laryngeal seropositivity= 3.5%) though small, is consistent with previous estimates indicating larger focused studies will be required to clarify such rare events.

In this study, nearly 94% of all HPV positive tumors were positive for HPV16 DNA, consistent with earlier reports (4, 9, 36, 37). Other types found were HPV31, HVP33, HPV35 and HPV66 together contributing a total of 23%, higher

than previous estimates (5, 12). Presence of HPV16 DNA, although necessary, is not sufficient to establish causality as it includes a subset of transient infections. In this study, we found 37 HPV DNA positive/ p16 negative tumors indicating that the majority of HPV16 infections (84%) are inactive. Interestingly, >60% of these were smokers (all oropharyngeal cases), suggestive of a smoking related etiology. Viral transcription and the production of viral oncoprotein E7 leads to the upregulation of p16 via the retinoblastoma pathway (30). A previous study found 100% sensitivity for p16 as a surrogate marker to identify HPV16 related cancers of the head and neck (24). We used the combination of p16 overexpression and HPV16 DNA positivity to identify biologically active infections. It can be argued that HPV independent mechanisms could result in the upregulation of p16. We found three p16 positive and HPV16 DNA negative cases (2.5%), lower than the previous estimates (12). It is also plausible that infection by non-HPV16 types could up regulate p16. In this study, we did not observe p16 overexpression among any of the HPV16 negative cases, although 3 of the 6 active infections were concomitantly positive for HPV31 or HPV33. It has been suggested that using HPV DNA presence as a marker could lead to overestimation of HPV related head and neck cancer and serology might reflect a more accurate estimate (4, 9). In agreement with this speculation, we found approximately 53% prevalence of HPV16 DNA among oropharyngeal cancers while based on serology, we found nearly 30%. HPV16 E6 serology identified a higher proportion of active HPV related cancers in the present study (67%) than previously published (40%) (24). One possible explanation is the higher threshold (s) used to define seropositivity. The observed concordance was further strengthened upon inclusion of additional markers such as HPV16 E2 or E4 (from 67% to 83%), warranting further investigation on the panel of serological markers that can accurately identify HPV driven cancers.

Our study has several limitations. First, the reported significance levels were not adjusted for multiple testing. However, even under the assumption of complete independence, important associations (particularly HPV16, HPV18 and HPV6) were robust. Second, reverse causality is a concern in the interpretation of these results given that blood samples were drawn at diagnosis. However, our results are concordant with the prospective case-control study that found 2-fold increased head and neck cancer risk with HPV16 capsid seropositivity (38). Third, some of the associations could be affected by other undiagnosed HPV related malignancies. Since the associations were consistent across men and women and the cases were first primary cancers of the UADT, this appears unlikely. Fourth, we restricted our analyses to the alpha papillomavirus family, although this could potentially underestimate the role of HPV in UADT cancer, literature indicates that these constitute known carcinogenic types. Our study has several strengths: first, the study was large enough to allow examination of site-specific associations of HPV infection. Second, the results of the sero-epidemiologic study were validated in a tumor subset that confirmed the concordance between serology and presence of an active HPV infection in the tumor. Additionally, our results were robust upon various sensitivity analyses.

In conclusion, the majority of the HPV16 infections in the UADT appear to be inactive. HPV16 E6 antibodies are promising markers to identify HPV16 related tumors, and indicate that at least 30% of oropharyngeal cancers are HPV16 related. In the light of highly consistent observations in the prospective setting and given the increasing

proportion of oropharyngeal cancers in Europe, it will be important to determine if HPV16 E6 serology can replace HPV DNA testing, particularly given the substantial overestimation DNA testing suffers. Larger focused studies will be required to clarify the appropriate algorithm to accurately identify HPV related UADT cancers.

Acknowledgements

We would like to thank all the patients and their families for their participation. We are also grateful to the support of many clinicians and staff of the hospitals, interviewers, data managers, pathology departments, and primary care clinics. G.J.M. and T.V.M. partly worked on ARCAGE study while at the University of Manchester. We acknowledge the help of Dr. Ann-Marie Biggs and Professor Martin Tickle in study conducted in the Manchester center and Professor Phil Sloan and Professor Nalin Thakker who in addition coordinated sample collection and processing for all the UK centers. The ARCAGE study was supported by the grant from European commission's 5th Framework Program contract QLK1-2001-00182. This project was partly funded by the Health General Directorate of the French social affairs and health ministry. The work done in this study was supported in part by a grant from the European Commission's 7th Framework Program contract (FP7-HEALTH-2011-282562).

References

1. Ferlay J, Shin HR, Bray F, Forman D, Mathers C, Parkin DM. Estimates of worldwide burden of cancer in 2008: GLOBOCAN 2008. *Int J Cancer*. 2010 Dec 15;127(12):2893-917.
2. IARC Monographs on the Evaluation of Carcinogenic Risks to Humans: A Review of Human Carcinogens: Biological Agents Lyon: IARC; 2012.
3. Leemans CR, Braakhuis BJ, Brakenhoff RH. The molecular biology of head and neck cancer. *Nat Rev Cancer*. [Research Support, Non-U.S. Gov't

Review]. 2011 Jan;11(1):9-22.
4. Herrero R, Castellsague X, Pawlita M, Lissowska J, Kee F, Balaram P, et al. Human papillomavirus and oral cancer: the International Agency for Research on Cancer multicenter study. *J NatlCancer Inst*. 2003;95(23):1772-83.
5. Kreimer AR, Clifford GM, Boyle P, Franceschi S. Human papillomavirus types in head and neck squamous cell carcinomas worldwide: a systematic review. *Cancer epidemiology, biomarkers & prevention : a publication of the American Association for Cancer Research, cosponsored by the American Society of Preventive Oncology*. [Review]. 2005 Feb;14(2):467-75.
6. Roden R, Wu TC. How will HPV vaccines affect cervical cancer? *Nat Rev Cancer*. [Research Support, N.I.H., Extramural

Research Support, Non-U.S. Gov't

Review]. 2006 Oct;6(10):753-63.
7. D'Souza G, Kreimer AR, Viscidi R, Pawlita M, Fakhry C, Koch WM, et al. Case-control study of human papillomavirus and oropharyngeal cancer. *N Engl J Med*. 2007 May 10;356(19):1944-56.
8. Gillison ML, D'Souza G, Westra W, Sugar E, Xiao W, Begum S, et al. Distinct risk factor profiles for human papillomavirus type 16-positive and human papillomavirus type 16-negative head and neck cancers. *J Natl Cancer Inst*. 2008 Mar 19;100(6):407-20.
9. Gillison ML, Koch WM, Capone RB, Spafford M, Westra WH, Wu L, et al. Evidence for a causal association between human papillomavirus and a subset of head and neck cancers. *J NatlCancer Inst*. 2000;92(9):709-20.
10. Jung AC, Briolat J, Millon R, de Reynies A, Rickman D, Thomas E, et al. Biological and clinical relevance of transcriptionally active human papillomavirus (HPV) infection in oropharynx squamous cell carcinoma. *International journal of cancer Journal international du cancer*. [Research Support, Non-U.S. Gov't]. 2010 Apr 15;126(8):1882-94.
11. Smith EM, Ritchie JM, Summersgill KF, Klussmann JP, Lee JH, Wang D, et al. Age, sexual behavior and human papillomavirus infection in oral cavity and oropharyngeal cancers. *International journal of cancer Journal international du cancer*. [Research Support, U.S. Gov't, Non-P.H.S.

Research Support, U.S. Gov't, P.H.S.]. 2004 Feb 20;108(5):766-72.
12. Marur S, D'Souza G, Westra WH, Forastiere AA. HPV-associated head and neck cancer: a virus-related cancer epidemic. *Lancet Oncol*. 2010;11(8):781-9.
13. Zumbach K, Hoffmann M, Kahn T, Bosch F, Gottschlich S, Gorogh T, et al. Antibodies against oncoproteins E6 and E7 of human papillomavirus types 16 and 18 in patients with head-and-neck squamous-cell carcinoma. *Int J Cancer*. 2000 Mar 15;85(6):815-8.
14. Smith EM, Ritchie JM, Pawlita M, Rubenstein LM, Haugen TH, Turek LP, et al. Human papillomavirus seropositivity and risks of head and neck cancer. *IntJ Cancer*. 2007;120(4):825-32.
15. Ribeiro KB, Levi JE, Pawlita M, Koifman S, Matos E, Eluf-Neto J, et al. Low human papillomavirus prevalence in head and neck cancer: results from two large case-control studies in high-incidence regions. *IntJ Epidemiol*. 2011.
16. Lagiou P, Georgila C, Minaki P, Ahrens W, Pohlabeln H, Benhamou S, et al. Alcohol-related cancers and genetic susceptibility in Europe: the ARCAGE project: study samples and data collection. *Eur J Cancer Prev*. [Multicenter Study

Research Support, Non-U.S. Gov't]. 2009 Feb;18(1):76-84.
17. Waterboer T, Sehr P, Michael KM, Franceschi S, Nieland JD, Joos TO, et al. Multiplex human papillomavirus serology based on in situ-purified glutathione s-transferase fusion proteins. *Clin Chem*. [Comparative Study

Research Support, Non-U.S. Gov't]. 2005 Oct;51(10):1845-53.

18. Clifford GM, Shin HR, Oh JK, Waterboer T, Ju YH, Vaccarella S, et al. Serologic response to oncogenic human papillomavirus types in male and female university students in Busan, South Korea. *Cancer epidemiology, biomarkers & prevention : a publication of the American Association for Cancer Research, cosponsored by the American Society of Preventive Oncology*. [Research Support, Non-U.S. Gov't]. 2007 Sep;16(9):1874-9.
19. Michael KM, Waterboer T, Sehr P, Rother A, Reidel U, Boeing H, et al. Seroprevalence of 34 human papillomavirus types in the German general population. *PLoS Pathog*. [Research Support, Non-U.S. Gov't]. 2008 Jun;4(6):e1000091.
20. Koo CL, Kok LF, Lee MY, Wu TS, Cheng YW, Hsu JD, et al. Scoring mechanisms of p16INK4a immunohistochemistry based on either independent nucleic stain or mixed cytoplasmic with nucleic expression can significantly signal to distinguish between endocervical and endometrial adenocarcinomas in a tissue microarray study. *J Transl Med*. 2009;7:25.
21. Gheit T, Landi S, Gemignani F, Snijders PJ, Vaccarella S, Franceschi S, et al. Development of a sensitive and specific assay combining multiplex PCR and DNA microarray primer extension to detect high-risk mucosal human papillomavirus types. *J Clin Microbiol*. [Evaluation Studies]. 2006 Jun;44(6):2025-31.
22. Schmitt M, Bravo IG, Snijders PJ, Gissmann L, Pawlita M, Waterboer T. Bead-based multiplex genotyping of human papillomaviruses. *J Clin Microbiol*. [Evaluation Studies

Research Support, Non-U.S. Gov't]. 2006 Feb;44(2):504-12.
23. Schmitt M, Dondog B, Waterboer T, Pawlita M, Tommasino M, Gheit T. Abundance of multiple high-risk human papillomavirus (HPV) infections found in cervical cells analyzed by use of an ultrasensitive HPV genotyping assay. *J ClinMicrobiol*. 2010;48(1):143-9.
24. Smeets SJ, Hesselink AT, Speel EJ, Haesevoets A, Snijders PJ, Pawlita M, et al. A novel algorithm for reliable detection of human papillomavirus in paraffin embedded head and neck cancer specimen. *IntJ Cancer*. 2007;121(11):2465-72.
25. Rothman K GS, Lash T. Modern Epidemiology. 3rd ed. ed. Philadelphia: Lippincott Williams & Wilkins; 2009.
26. Miettinen OS. Proportion of disease caused or prevented by a given exposure, trait or intervention. *Am J Epidemiol*. 1974 May;99(5):325-32.
27. Clifford GM, Gallus S, Herrero R, Munoz N, Snijders PJ, Vaccarella S, et al. Worldwide distribution of human papillomavirus types in cytologically normal women in the International Agency for Research on Cancer HPV prevalence surveys: a pooled analysis. *Lancet*. [Research Support, Non-U.S. Gov't]. 2005 Sep 17-23;366(9490):991-8.
28. Dillner J. The serological response to papillomaviruses. *Semin Cancer Biol*. [Review]. 1999 Dec;9(6):423-30.
29. Zumbach K, Kisseljov F, Sacharova O, Shaichaev G, Semjonova L, Pavlova L, et al. Antibodies against oncoproteins E6 and E7 of human papillomavirus types 16 and 18 in cervical-carcinoma patients from Russia. *IntJ Cancer*. 2000;85(3):313-8.
30. Moody CA, Laimins LA. Human papillomavirus oncoproteins: pathways to transformation. *Nat Rev Cancer*. [Research Support, N.I.H., Extramural

Review]. 2010 Aug;10(8):550-60.
31. Rocha-Zavaleta L, Jordan D, Pepper S, Corbitt G, Clarke F, Maitland NJ, et al. Differences in serological IgA responses to recombinant baculovirus-derived human papillomavirus E2 protein in the natural history of cervical neoplasia. *Br J Cancer*. [Research Support, Non-U.S. Gov't]. 1997;75(8):1144-50.
32. Anderson KS, Wong J, D'Souza G, Riemer AB, Lorch J, Haddad R, et al. Serum antibodies to the HPV16 proteome as biomarkers for head and neck cancer. *Br J Cancer*. 2011 Jun 7;104(12):1896-905.
33. Sitas F, Egger S, Urban MI, Taylor PR, Abnet CC, Boffetta P, et al. InterSCOPE Study: Associations Between Esophageal Squamous Cell Carcinoma and Human Papillomavirus Serological Markers. *Journal of the National Cancer Institute*. 2012 Jan 18;104(2):147-58.
34. Syrjanen KJ. HPV infections and oesophageal cancer. *J Clin Pathol*. [Review]. 2002 Oct;55(10):721-8.
35. Herrero R. Chapter 7: Human papillomavirus and cancer of the upper aerodigestive tract. *J Natl Cancer Inst Monogr*. [Review]. 2003(31):47-51.
36. Schwartz SM, Daling JR, Doody DR, Wipf GC, Carter JJ, Madeleine MM, et al. Oral cancer risk in relation to sexual history and evidence of human papillomavirus infection. *Journal of the National Cancer Institute*. [Research Support, Non-U.S. Gov't

Research Support, U.S. Gov't, P.H.S.]. 1998 Nov 4;90(21):1626-36.

37. Smith EM, Hoffman HT, Summersgill KS, Kirchner HL, Turek LP, Haugen TH. Human papillomavirus and risk of oral cancer. *The Laryngoscope*. [Research Support, U.S. Gov't, Non-P.H.S.

Research Support, U.S. Gov't, P.H.S.]. 1998 Jul;108(7):1098-103.

38. Mork J, Lie AK, Glattre E, Hallmans G, Jellum E, Koskela P, et al. Human papillomavirus infection as a risk factor for squamous-cell carcinoma of the head and neck. *NEnglJ Med*. 2001;344(15):1125-31.

Table1: Demographic and lifestyle characteristics of the study group

Description	*Serological series**		*Tumor subset^*
	Controls (n= 1425)	**Cases** (n= 1496)	**Cases** (n= 120)
Country			
Czech Republic	185 (13)	187 (12)	16 (13)
Germany	187 (13)	188 (13)	-
Greece	167 (12)	224 (15)	53 (44)
Italy	462 (32)	440 (29)	51 (43)
Ireland	16 (1)	33 (2)	-
Norway	168 (12)	162 (11)	-
United Kingdom	112 (8)	119 (8)	-
Spain	82 (6)	89 (6)	-
Croatia	46 (3)	54 (4)	-
Age group (in years)			
≤55	494 (35)	498 (32)	38 (32)
56 - 65	456 (32)	551 (37)	39 (33)
≥66	475 (33)	447 (30)	43 (36)
Gender			
Men	1059 (74)	1190 (80)	81 (68)
Women	366 (26)	306 (20)	39 (32)
Smoking status			
Never	516 (36)	172 (12)	32 (27)
Former	475 (33)	361 (24)	16 (13)
Current	434 (31)	963 (64)	72 (60)
Alcohol consumption[£]			
Never	172 (12)	89 (6)	8 (6)
Former	134 (9)	206 (14)	18 (15)
Current	1118 (79)	1201 (80)	94 (78)
Level of education attained[£]			
Finished primary	444 (31)	630 (42)	70 (58)
Finished secondary	834 (59)	798 (53)	40 (33)
University degree	147 (10)	66 (4)	10 (8)
Cancer site			
Oral cavity		366 (24)	42 (35)
Oropharynx		324 (22)	36 (30)
Hypopharynx/ Larynx		529 (35)	16 (13)
Esophagus		200 (13)	8 (7)
Overlapping [¥]		77 (5)	18 (15)
Stage[£]			
I & II		530 (35)	48 (40)
III & IV		686 (46)	67 (56)

*cases and controls assayed by bead-based multiplex serology for type specific markers of HPV infection

^ cases included in the analysis following multiplex HPV genotyping and p16 expression testing

[£]Numbers do not add up to the total due to due to missing data, information on alcohol consumption was missing for one control, education level data was missing for 2 cases, stage missing for 280 cases

[¥] includes cancers of overlapping topologies and non-specified cancers of the head and neck

Table2: HPV type-specific antibodies and oropharyngeal cancer risk

HPV antibody	Controls (N=1395) Seropositive (%)^	Oropharynx cancer (N= 321)	OR (95% CI)$	p-value	Controls (N=1288) Seropositive (%)^	Oropharynx cancer (N= 198)	OR (95% CI)$	p-value
Antibodies to mucosal high risk HPV types								
HPV16¥								
E1	31 (2.2)	69 (21.2)	22.63 (13.63- 37.57)	**<1E-07**				
E2	29 (2.1)	81 (25.2)	30.65 (18.56- 50.64)	**<1E-07**				
E4	85 (5.9)	41 (12.7)	2.59 (1.68- 4.00)	**1.7E-05**				
	Unstratified analyses				**Excluding HPV16 L1, E6 and E7 positives**			
HPV18								
L1	50 (3.6)	21 (6.5)	2.32 (1.33- 4.04)	**0.003**	32 (2.3)	9 (4.5)	2.37 (1.06- 5.32)	**0.04**
E6	7 (0.5)	11 (3.4)	8.16 (2.81- 23.66)	**0.0001**	6 (0.4)	7 (3.5)	8.14 (2.21- 29.99)	**0.002**
E7	5 (0.4)	8 (2.5)	9.31 (2.75- 31.50)	**0.0003**	3 (0.2)	2 (1.0)	4.81 (0.65- 35.45)	0.12
HPV31								
L1	51 (3.7)	16 (5.0)	1.90 (1.03- 3.51)	**0.04**	38 (2.7)	3 (1.5)	0.92 (0.27- 3.12)	0.89
E6	17 (1.2)	7 (2.2)	1.71 (0.66- 4.41)	0.27	16 (1.1)	2 (1.0)	0.67 (0.14- 3.28)	0.62
E7	14 (1.0)	53 (16.5)	32.33 (16.93- 61.74)	**2.5E-06**	10 (0.7)	1 (0.5)	0.77 (0.09- 6.57)	0.81
HPV33								
L1	37 (2.7)	19 (5.9)	2.73 (1.46- 5.09)	**0.002**	26 (1.9)	8 (4.0)	1.97 (0.78- 4.96)	0.15
E6	7 (0.5)	14 (4.4)	12.95 (4.90- 34.28)	**2.5E-07**	4 (0.3)	0		
E7	21 (1.5)	62 (19.3)	26.48 (15.18- 46.20)	**2.5E-06**	15 (1.1)	2 (1.0)	0.62 (0.11- 3.56)	0.59
HPV45								
L1	41 (2.9)	15 (4.7)	1.89 (0.99- 3.61)	**0.05**	29 (2.1)	5 (2.5)	1.19 (0.42- 3.40)	0.75
E6	11 (0.8)	7 (2.2)	3.50 (1.28- 9.57)	**0.02**	10 (0.7)	2 (1.0)	1.79 (0.32- 9.86)	0.50
E7	10 (0.7)	6 (1.9)	4.29 (1.47- 12.50)	**0.008**	10 (0.7)	2 (1.0)	2.55 (0.53- 12.27)	0.24
HPV52								
L1	33 (2.4)	15 (4.7)	2.81 (1.43- 5.50)	**0.003**	20 (1.4)	4 (2.0)	1.94 (0.61- 6.20)	0.26
E6	11 (0.8)	4 (1.2)	1.48 (0.41- 5.31)	0.55	10 (0.7)	0		
E7	25 (1.8)	28 (8.7)	7.79 (4.27- 14.19)	**1.9E-11**	20 (1.4)	4 (2.0)	2.11 (0.65- 6.82)	0.21
Antibodies to mucosal low risk HPV types								
HPV6								
L1	223 (16.0)	65 (20.2)	1.17 (0.84- 1.63)	0.35	187 (13.4)	46 (23.2)	1.53 (1.02- 2.30)	**0.04**
E6	10 (0.7)	1 (0.3)	0.36 (0.04- 2.95)	0.34	9 (0.6)	0		
E7	18 (1.3)	5 (1.6)	1.10 (0.38- 3.22)	0.86	16 (1.1)	3 (1.5)	1.35 (0.34- 5.39)	0.67
HPV11								
L1	72 (5.2)	30 (9.3)	1.85 (1.15- 2.98)	**0.01**	22 (1.6)	2 (1.0)	1.83 (0.96- 3.48)	**0.06**
E6	22 (1.6)	2 (0.6)	0.36 (0.08- 1.71)	0.20	11 (0.8)	1 (0.5)	0.53 (0.10- 2.86)	0.46
E7	12 (0.9)	4 (1.2)	1.33 (0.37- 4.83)	0.66	198 (14.2)	25 (12.6)	0.29 (0.02- 3.40)	0.32

N represents number of cases and controls included in the analysis following exclusion of subjects missing details on smoking and alcohol (3 cases, 30 controls)

^ represents corresponding HPV type specific seropositivity

$ Odds ratios were adjusted for age, sex, level of education, smoking pack years and number of alcoholic drinks consumed per day, as appropriate

¥HPV16 antibodies to E1, E2 and E4 were included as additional markers of HPV16 infection, hence analysis were not stratified by HPV16 L1, E6 and E7.

Table3: Proportion of UADT cancers attributable to HPV based on serology and tumor analyses

Cancer site	HPV16 infection			
	E6 serology*		Active infection^	
	positive/total	AF[£]	positive/total	AF[$]
UADT cancer	119/1496	7.3%	6/120	5%
Oropharynx	98/324	29.5%	6/36	17%
Oral cavity	4/366	0.8%	0/42	-
Larynx[¥]	8/529	2.8%	0/16	-
Esophagus	5/200	7.4%	0/8	-

*proportion of cases in each category positive for HPV16 E6 antibodies

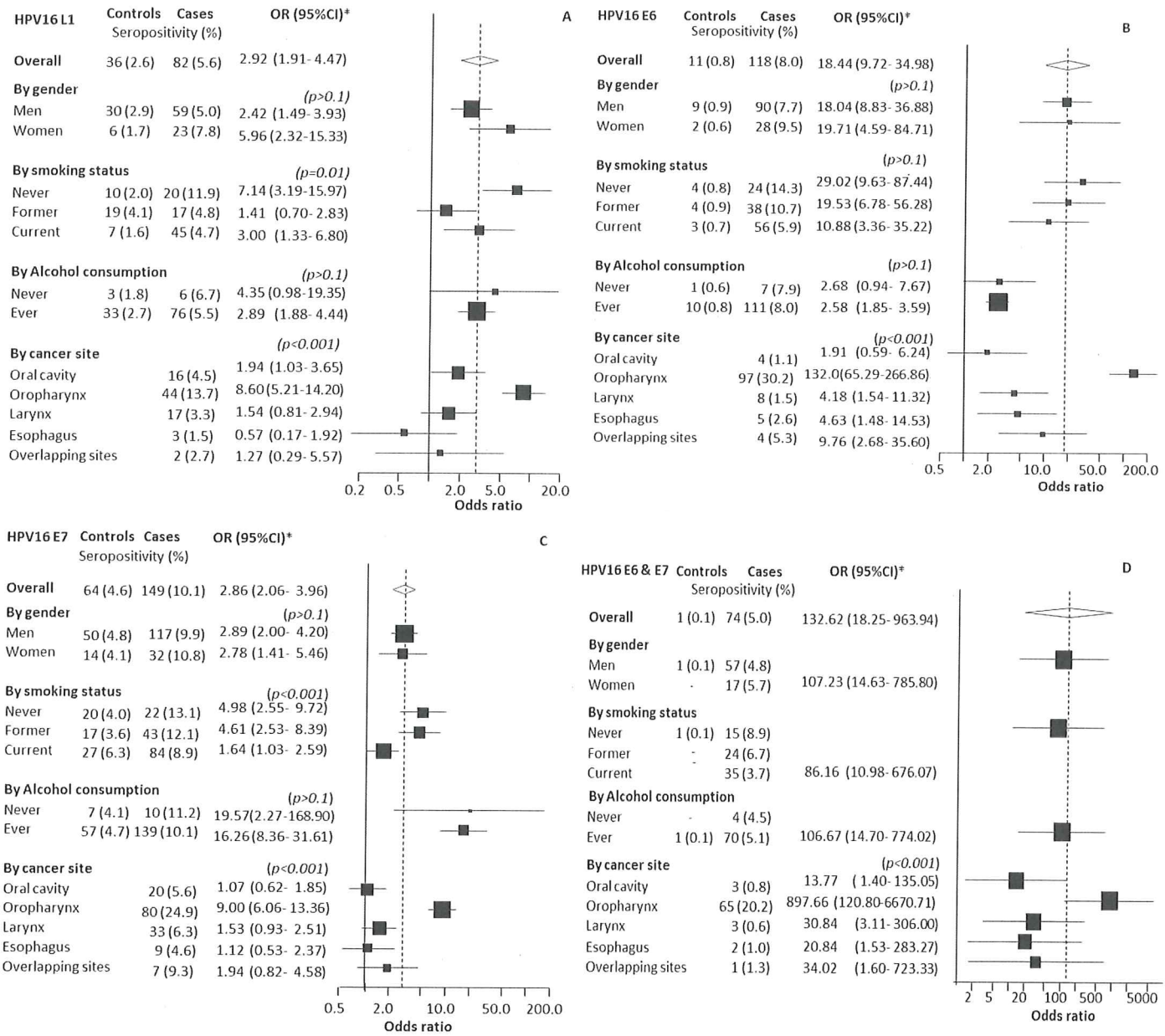
^HPV16 DNA positive and p16 overexpressing tumors were considered active infections

[£] Attributable fraction (AF) for HPV16 E6 serology was estimated by the formula pec* (OR-1)/OR, where pec= proportion of exposed among cases, OR= odds ratio derived from model adjusted for age, sex, level of education and country

[$] indicates proportion of tumors of active HPV16 infections (overexpressing p16 and HPV16 DNA positive) by total tumors in each category

[¥]Includes larynx and hypopharynx cases

Figure1: HPV16 L1, E6 and E7 antibodies and the risk of UADT cancer

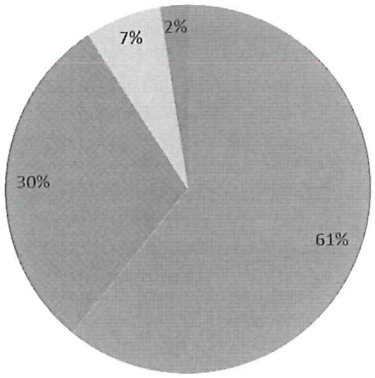


Odds ratios derived from logistic regression analyses adjusted for age, sex, level of education, country, smoking pack years and number of alcoholic drinks consumed per day, as appropriate. Corresponding antibody negatives were considered reference. 54 subjects were missing data on smoking pack years or frequency of alcohol consumption, a total of 1472 cases and 1395 controls were included in the analysis. P indicates p-value for heterogeneity. Larynx subgroup includes laryngeal and hypopharygeal cancer cases, overlapping sites group include cancers overlapping head and neck sites and non-specified cancers of the head and neck.

Figure represents UADT cancer associations with

A) HPV16 L1 antibodies; B) HPV16 E6 seropositivity; C) HPV16 E7 antibodies; D) combined seropositivity to both E6 and E7 antibodies

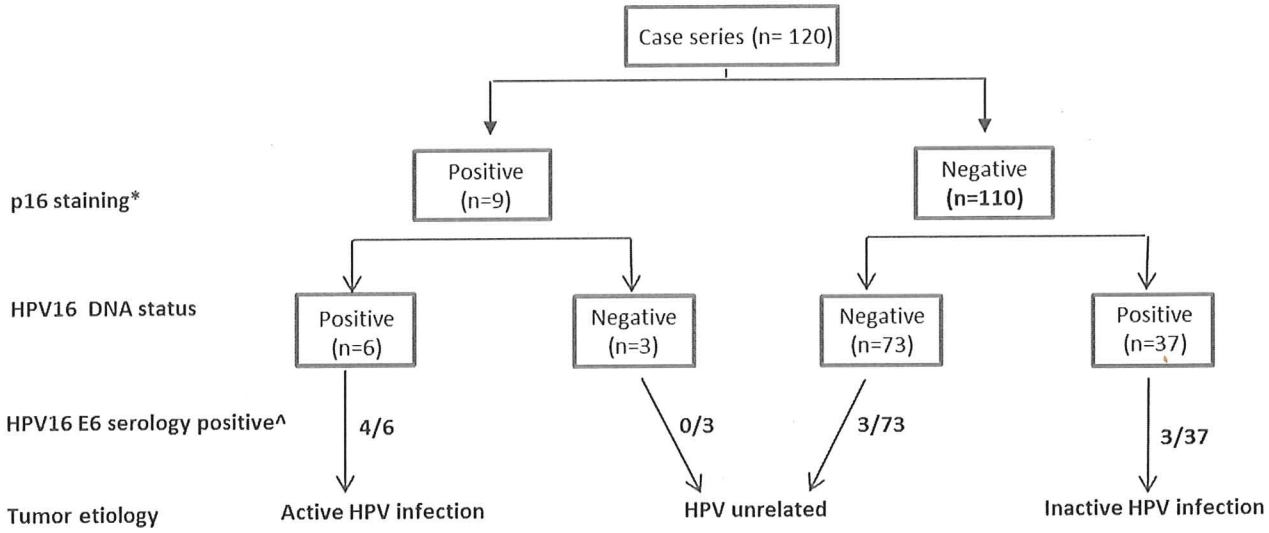
Figure2: HPV DNA presence and distribution in tumour tissues



A total of 120 tumors were examined.

- absence of HPV DNA indicates HPV unrelated tumors (n=73)
- indicate single infections with HPV16 (n=36)
- HPV31 (n=2) and HPV33 (n=6) DNA were each identified in the presence of HPV16 DNA as multiple infections including HPV16.
- Non-HPV16 related infections were of single type only, two positive for HPV35 DNA and one for HPV66 DNA.

Figure3: Schematic representation of concordance between markers of HPV16 infection



Overall 6 HPV16 DNA positive and P16 positive cases were identified (all six were cancers of the oropharynx). 6 of 110 p16 negative cases were positive for HPV16 E6 antibody, 3 were p16 negative/ HPV DNA positive and 3 were P16 negative/ HPV16 DNA negative.

*One case missing p16 staining data.

Supplementary table1: HPV seropositivity and UADT cancer risk

HPV type specific antibody	Controls Seropositive (%) ^ (n=1395)	UADT Cases (n=1472)	OR (95% CI)*
HPV16L1	36 (2.6)	82 (5.6)	2.92 (1.91- 4.47)
HPV16E1	31 (2.2)	85 (5.8)	4.04 (2.58- 6.32)
HPV16E2	29 (2.1)	109 (7.4)	5.54 (3.58- 8.57)
HPV16E4	83 (6.0)	97 (6.6)	1.24 (0.89- 1.72)
HPV16E6	11 (0.8)	118 (8.0)	18.44 (9.72- 34.98)
HPV16E7	64 (4.6)	149 (10.1)	2.86 (2.06- 3.96)
HPV18 L1	50 (3.6)	68 (4.6)	1.40 (0.94- 2.10)
HPV18 E6	7 (0.5)	24 (1.6)	3.50 (1.38- 8.83)
HPV18 E7	5 (0.4)	16 (1.1)	4.25 (1.46- 12.37)
HPV31 L1	51 (3.7)	57 (3.9)	1.32 (0.87- 2.01)
HPV31 E6	17 (1.2)	17 (1.2)	1.07 (0.51- 2.23)
HPV31 E7	14 (1.0)	72 (4.9)	7.15 (3.92- 13.07)
HPV33 L1	37 (2.7)	51 (3.5)	1.49 (0.93- 2.37)
HPV33 E6	7 (0.5)	24 (1.6)	3.88 (1.59- 9.48)
HPV33 E7	21 (1.5)	90 (6.1)	6.06 (3.66- 10.02)
HPV45 L1	41 (2.9)	49 (3.3)	1.32 (0.83- 2.07)
HPV45 E6	11 (0.8)	19 (1.3)	1.89 (0.86- 4.18)
HPV45 E7	10 (0.7)	15 (1.0)	2.29 (0.97- 5.39)
HPV52 L1	33 (2.4)	35 (2.4)	1.22 (0.72- 2.06)
HPV52 E6	11 (0.8)	14 (1.0)	1.20 (0.49- 2.93)
HPV52 E7	25 (1.8)	67 (4.6)	3.02 (1.83- 4.96)
HPV6 L1	223 (16.0)	289 (19.6)	1.08 (0.87- 1.33)
HPV6 E6	10 (0.7)	8 (0.5)	0.61 (0.23- 1.64)
HPV6 E7	18 (1.3)	35 (2.4)	1.64 (0.88- 3.07)
HPV11 L1	72 (5.2)	111 (7.5)	1.35 (0.97- 1.88)
HPV11 E6	22 (1.6)	32 (2.2)	1.50 (0.82- 2.77)
HPV11 E7	12 (0.9)	21 (1.4)	1.68 (0.77- 3.66)

^ represents the proportion of corresponding HPV antibody positives among cases and controls, excluding 54 subject missing smoking and alcohol exposure details

$ Odds ratios adjusted for age, sex, level of education smoking pack years and number of alcoholic drinks consumed per day, as appropriate. Respective seronegative groups were considered reference

Supplementary table2: HPV6 antibodies and laryngeal cancer risk

HPV6 antibody	Controls (n=1395)	Larynx cases* (n=521)	OR (95% CI)$	P-value	Controls (n=1288)	Larynx cases* (n=468)	OR (95% CI)$	P-value
	Seropositive (%)^				Seropositive (%)^			
	All larynx cancer cases				Excluding HPV16 L1, E6 & E7 positives			
	Larynx*, fully adjusted model$							
L1	223 (16.0)	110 (21.1)	1.33 (1.00- 1.77)	**0.05**	187 (14.5)	90 (19.2)	1.17 (0.85- 1.60)	0.35
E6	10 (0.7)	2 (0.4)	0.53 (0.11- 2.50)	0.42	9 (0.7)	2 (0.4)	0.46 (0.09- 2.41)	0.36
E7	18 (1.3)	18 (3.5)	3.16 (1.55- 6.45)	**0.002**	16 (1.2)	17 (3.6)	3.25 (1.46- 7.24)	**0.004**

* includes larynx and hypopharynx cancers

^ represents the proportion of corresponding HPV antibody positives among cases and controls, excluding subject missing smoking and alcohol exposure details

$ adjusted for age, sex, level of education, country, smoking (never, <20, 20- 39, 40-59, 60-79 and >80) and number of alcoholic drinks consumed per day (never, <1, 1-2, 3-4, 5-6, >7)

Supplementary table3: Sensitivity analysis based on phylogenetic relatedness

HPV type/ Antibody	Controls (n= 1425) Seropositive (%) ^	UADT cases (n= 1496)	OR (95% CI)*
Cross reactivity between HPV types: 16, 31, 33 & 52: alpha7 species			
L1 seropositivity			
None	1318 (93)	1340 (90)	Reference
HPV16 only	23 (1.6)	56 (3.7)	3.36 (1.95- 5.76)
HPV31 only	27 (1.9)	24 (1.6)	1.08 (0.59- 2.00)
HPV33 only	16 (1.1)	22 (1.5)	1.28 (0.62- 2.63)
HPV52 only	9 (0.6)	8 (0.5)	0.63 (0.21- 1.89)
E6 seropositivity			
None	1382 (97)	1344 (90)	Reference
HPV16 only	10 (1)	97 (6)	17.03 (8.66- 33.52)
HPV31 only	16 (1.1)	13 (0.9)	0.90 (0.39- 2.05)
HPV33 only	6 (0.4)	8 (0.5)	1.16 (0.35- 3.80)
HPV52 only	11 (0.8)	9 (0.6)	-
E7 seropositivity			
None	1311 (92)	1277 (85)	Reference
HPV16 only	56 (4)	72 (5)	1.31 (0.88- 1.95)
HPV31 only	11 (1)	10 (1)	0.70 (0.27- 1.79)
HPV33 only	15 (1)	18 (1)	1.40 (0.66- 2.95)
HPV52 only	21 (1)	32 (2)	1.64 (0.89- 3.04)
Cross reactivity between HPV types: 18 & 45: alpha9 species			
L1 seropositivity			
None	1360 (95)	1404 (94)	Reference
HPV18 only	24 (2)	41 (3)	1.50 (0.86- 2.60)
HPV45 only	15 (1)	21 (1)	1.33 (0.64- 2.75)
E6 seropositivity			
None	1408 (99)	1455 (97)	Reference
HPV18 only	6 (0.4)	22 (1)	4.64 (1.61- 13.34)
HPV45 only	9 (1)	17 (1)	2.14 (0.91- 5.04)
E7 seropositivity			
None	1411 (99)	1470 (98)	Reference
HPV18 only	4 (0.3)	11 (1)	2.72 (0.81- 9.15)
HPV45 only	9 (1)	10 (1)	1.45 (0.55- 3.81)
Cross reactivity between HPV types: 6 & 11: alpha10 species			
L1 seropositivity			
None	1175 (82)	1180 (79)	Reference
HPV6 only	176 (12)	204 (14)	0.95 (0.75- 1.21)
HPV11 only	21 (1)	23 (2)	0.97 (0.51- 1.83)
E6 seropositivity			
None	1393 (98)	1461 (98)	Reference
HPV6 only	9 (1)	3 (0.2)	0.28 (0.07- 1.12)
HPV11 only	22 (2)	27 (2)	1.40 (0.74- 2.67)
E7 seropositivity			
None	1395 (98)	1442 (96)	Reference

HPV6 only	18 (1)	33 (2)	1.73 (0.91- 3.29)
HPV11 only	11 (1)	19 (1)	1.83 (0.81- 4.13)

^ indicates number of seropositive individuals in the corresponding category negative for homologous protein in phylogenetically related HPV type

*Odds ratios were adjusted for age, sex, level of education, country, smoking pack years and alcohol consumption

Supplementary table4: Concordance between HPV16 serology and tumor markers of HPV16 infection

	Active HPV infection*			
HPV16 antibody status	**Positive** (n= 6)	**Negative** (n= 113)	**Sensitivity^**	**Specificity^**
HPV16 L1				
Positive	2	9		
Negative	4	104	33.3 (4.33- 77.7)	92.0 (85.4- 96.3)
HPV16 E6				
Positive	4	6		
Negative	2	107	66.7 (22.3- 95.7)	94.7 (88.8- 98.0)
HPV16 E7				
Positive	1	12		
Negative	5	101	16.7 (0.42- 64.1)	89.4 (82.2- 94.4)

*Active infection defined as cases positive for HPV16 DNA and P16 overexpression. One HPV16 DNA positive case was excluded as P16 data was unavailable

^ True positives were defined as cases positive for HPV16 DNA and overexpressing surrogate marker p16. True negatives were all p16 negative cases and p16 overexpressing and HPV16 DNA negative cases combined.

Supplementary table5: HPV16 antibodies and risk of oropharyngeal and non-oropharyngeal cancer

HPV16 Antibody	Controls Seropositivity (%)^	Oropharynx cases Seropositivity (%)^	OR (95% CI)*	Controls Seropositivity (%)^	Non-oropharynx cases Seropositivity (%)^	OR (95% CI)*
			HPV16L1			
Overall	36 (2.6)	44.0 (3.7)	8.60 (5.21- 14.20)	36 (2.58)	38 (3.30)	1.43 (0.87- 2.35)
Men	30 (2.9)	31.0 (2.4)	7.38 (4.15- 13.13)	30 (2.86)	28 (3.02)	1.18 (0.67- 2.09)
Women	6 (1.7)	13.0 (8.3)	18.29 (6.10- 54.85)	6 (1.74)	10 (4.44)	3.01 (1.02- 8.85)
≤55 years	12 (2.5)	16.0 (3.5)	8.98 (3.76- 21.45)	12 (2.48)	13 (3.50)	1.53 (0.63- 3.71)
56-65 years	17 (3.8)	19.0 (4.4)	5.92 (2.81- 12.47)	17 (3.78)	13 (3.16)	0.87 (0.39- 1.95)
≥65 years	7 (1.5)	9.0 (2.9)	15.67 (5.16- 47.59)	7 (1.52)	12 (3.25)	2.81 (1.03- 7.62)
Never smoker	10 (2.1)	14.0 (43.8)	49.45 (17.54- 139.46)	10 (2.01)	6 (4.41)	2.04 (0.70- 5.93)
Ever smoker	26 (2.9)	30.0 (0.4)	4.16 (2.37- 7.33)	26 (2.90)	32 (3.13)	1.07 (0.62- 1.84)
			HPV16E6			
Overall	11 (0.8)	97.0 (30.2)	132.00 (65.29- 266.86)	11 (0.79)	21 (1.82)	3.44 (1.59- 7.45)
Men	9 (0.9)	75.0 (30.0)	127.88 (58.30- 280.48)	9 (0.86)	15 (1.62)	3.10 (1.28- 7.51)
Women	2 (0.6)	22.0 (31.0)	139.28 (28.31- 685.34)	2 (0.58)	6 (2.67)	4.63 (0.91- 23.72)
≤55 years	5 (1.3)	35.0 (29.4)	100.03 (33.78- 296.23)	5 (1.03)	10 (2.70)	5.90 (1.88- 18.46)
56-65 years	3 (0.7)	36.0 (27.3)	108.33 (30.32- 387.05)	3 (0.67)	8 (1.95)	2.95 (0.72- 12.16)
≥65 years	3 (0.7)	26.0 (37.4)	310.23 (76.09- 1264.85)	3 (0.65)	3 (0.81)	1.68 (0.32- 8.72)
Never smoker	4 (0.8)	21.0 (65.6)	493.25 (106.43- 2286.0)	4 (0.80)	3 (2.21)	4.38 (0.93- 20.69)
Ever smoker	7 (0.8)	76.0 (26.3)	67.63 (30.06- 152.18)	7 (0.78)	18 (1.76)	2.50 (1.02- 6.16)
			HPV16E7			
Overall	64 (4.6)	80.0 (24.9)	9.00 (6.06- 13.36)	64 (4.59)	69 (5.99)	1.39 (0.95- 2.03)
Men	50 (4.8)	64.0 (25.6)	9.73 (6.20- 15.25)	50 (4.76)	53 (5.72)	1.33 (0.86- 2.06)
Women	14 (4.6)	16.0 (22.5)	6.64 (2.90- 15.24)	14 (4.06)	16 (7.11)	1.66 (0.76- 3.62)
≤55 years	30 (6.2)	33.0 (27.7)	8.89 (4.75- 16.62)	30 (6.20)	22 (5.93)	1.03 (0.55- 1.94)
56-65 years	19 (4.2)	31.0 (23.5)	8.56 (4.37- 16.74)	19 (4.22)	24 (5.84)	1.35 (0.69- 2.65)
≥65 years	15 (3.3)	16.0 (22.9)	12.97 (5.55- 30.30)	15 (3.25)	23 (6.23)	2.07 (1.02- 4.20)
Never smoker	20 (4.2)	13.0 (40.6)	23.95 (9.33- 61.49)	20 (4.02)	9 (6.62)	2.20 (0.94- 5.11)
Ever smoker	44 (4.9)	67.0 (23.8)	7.26 (4.72- 11.18)	44 (4.90)	60 (5.88)	1.32 (0.87- 2.00)

^ represents the proportion of corresponding HPV antibody positives among cases and controls, excluding subject missing smoking and alcohol exposure details

*adjusted for age, sex, level of education, country, smoking pack years and number of alcoholic drinks consumed per day, as appropriate

Supplementary table6: Antibodies to HPV16 E6 or E7, stage at diagnosis and loco- regional metastasis

Description	Overall		Oropharyngeal cancer	
	Seropositivity (%)	OR (95% CI)#	Seropositivity (%)	OR (95% CI) #
HPV16 E6 or E7 seropositivity				
Stage at diagnosis*				
Early (I - II)	46 (8.7)	1.0	14 (19.2)	1.0
Late (III - IV)	65 (9.5)	1.29 (0.86- 1.94)	29 (14.5)	0.82 (0.38- 1.76)
M stage^				
M0	87 (9.1)	1.0	33 (14.6)	1.0
M+	13 (18.3)	2.12 (1.09- 4.10)	7 (50.0)	5.53 (1.35- 22.57)
HPV16 E6 and E7 seropositivity				
Stage at diagnosis				
Early (I - II)	15 (2.9)	1.0	11 (15.1)	1.0
Late (III - IV)	45 (6.7)	3.06 (1.66- 5.64)	41 (20.8)	1.69 (0.74- 3.84)
M stage				
M0	54 (5.6)	1.0	46 (20.4)	1.0
M+	3 (4.2)	0.70 (0.20- 2.37)	3 (21.4)	1.72 (0.31- 9.70)

* Information on stage at diagnosis was missing for 280 UADT cases (18.7%), and 51 oropharyngeal cases (15.7%)

^ Metastasis information was missing for 451 UADT cases (30.2%), and 81 oropharyngeal cases (25%)

adjusted for age, sex, level of education, country, pack years of smoking and number of alcoholic drinks consumed per day