Novel nomograms for survival and progression in HPV+ and HPV− oropharyngeal cancer: a population-based study of 1,542 consecutive patients

Christian Grønhøj Larsen[1], David H. Jensen[1], Amanda-Louise Fenger Carlander[1], Katalin Kiss[2], Luise Andersen[2], Caroline Holkmann Olsen[6], Elo Andersen[3], Emilie Garnæs[1], Finn Cilius[4], Lena Specht[5], Christian von Buchwald[1]

[1]Department of Otorhinolaryngology, Head and Neck Surgery and Audiology, Rigshospitalet, University of Copenhagen, Copenhagen, Denmark

[2]Department of Pathology, Rigshospitalet, University of Copenhagen, Copenhagen, Denmark

[3]Department of Oncology, Herlev Hospital, University of Copenhagen, Copenhagen, Denmark

[4]Centre for Genomic Medicine, Rigshospitalet, University of Copenhagen, Copenhagen, Denmark

[5]Department of Oncology, Rigshospitalet, University of Copenhagen, Copenhagen, Denmark

[6]Department of Pathology, Roskilde Hospital, Roskilde, Denmark

Correspondence to: Christian von Buchwald, **email:** christian.von.buchwald@regionh.dk

ABSTRACT

Background: No study has combined tumour and clinical covariates for survival to construct an individual risk-profile for overall survival (OS), time to progression (TTP), and survival after progression (SAP) in patients with HPV+ and HPV− oropharyngeal squamous cell carcinoma (OPSCC). Based on the largest-to-date, unselected, population-based cohort of patients diagnosed with OPSCC, we performed a comprehensive analysis of long-term OS, TTP, and SAP and constructed novel nomograms to evaluate patients' prognoses.

Results: At a median follow-up of 4.0 years (range: 0.8–15.8 yrs.), 690 deaths were recorded. The 5-year OS, TTP, and SAP for the HPV+/p16+ subgroup were 77%, 82%, and 33, vs. 30%, 66%, and 6% for the HPV−/p16− group ($P < 0.01$). 376 patients failed to maintain disease control with a median TTP of 13 months in the HPV+/p16+ subgroup vs. 8.5 months in the HPV−/p16− subgroup ($P < 0.05$). HPV combined with p16 status remained one of the most informative covariates in the final Cox regression model for OS, TTP, and SAP.

Methods: We included all patients diagnosed with OPSCC ($n = 1,542$) between 2000–2014 in Eastern Denmark. Survival rates were estimated by the Kaplan-Meier method. A multivariate Cox regression model was used to construct predictive, internally validated nomograms.

Conclusion: The HPV+/p16+ subgroup had improved OS, TTP, and SAP compared with other combinations of HPV and p16 after adjusting for covariates. Nomograms were constructed for 1-, 5- and 10-year survival probability. Models may aid patients and clinicians in their clinical decision making as well as in counselling, research, and trial design.

INTRODUCTION

In the Western world, the main risk factor for developing oropharyngeal squamous cell carcinoma (OPSCC) is infection with high-risk human papillomavirus (HPV); while, a smaller proportion is related to the high consumption of alcohol and smoking tobacco. Patients with HPV-associated OPSCCs have improved survival, which may be related to a different mutational profile [1, 2], histopathology [3], and clinical features [4]. The first reports of improved survival for HPV-positive (HPV+) OPSCCs were published in the early 2000's [5, 6].Shortly after, p16 (a surrogate-marker for HPV-infection) was demonstrated to have prognostic value [7, 8]. The importance of HPV and p16 has since been confirmed, and the combination (e.g. HPV/p16-status) has shown better prognostication [9]. However, long-term results are missing for overall survival (OS), time to progression (TTP), and survival after progression (SAP), especially their relationship to clinico-pathological characteristics from a large, population-based cohort.

Progression is a strong predictor of survival in patients with OPSCCs. The subgroup of patients who are HPV and p16-positive have a lower risk of recurrence and improved SAP compared with patients without HPV/p16 [10–12]. These findings are based on selected patients from small cohort-studies examining only p16 or HPV without a detailed account of progression. Therefore, HPV+ OPSCC may have an increased OS due to a better response to salvage therapy (i.e. SAP). A nomogram is a graphic demonstration of a statistical model for calculating the cumulative effect of weighted variables on the probability of a particular outcome, and enables continuous estimation of the probability of specific outcomes (i.e. death or progression). Furthermore, they can combine multiple independent variables while considering the prognostic weight for each variable when calculating the probability of an outcome. Although nomograms have been developed to predict clinical end-points for patients with several types of malignancies, to the best of our knowledge, this is the first study addressing OS, TTP, and SAP in OPSCCs [13–17].

To avoid selection-bias, we aimed to include all patients diagnosed with OPSCC in the Eastern part of Denmark, which has an approximate population size of 2.4 million inhabitants. The cohort is well described and includes detailed information on both tumour (HPV, p16, stage, site) as well as clinical characteristics (follow-up, treatment, progression) with virtually no patients lost to follow-up. The primary aim was to identify tumour- and patient-factors associated with OS, TTP, and SAP. Due to our unique, complete population-based cohort spanning 15-years, we were able to construct and validate reliable, predictive nomograms.

RESULTS

Population demographics

A total of 1,542 patients diagnosed with OPSCC between 2000–2014 were included, and their characteristics are summarized in Table 1 with univariate hazard ratios for OS. The majority of patients were males (72%), with a median age of 60 years at diagnosis, and typically presented with tumours in an advanced nodal stage, but with small primary tumours (Table 1). Of the total cohort, 54% were stage III or IV, 61% had T1 or T2 primary tumours, and 60% had N2 or N3 nodal classification (Table 1). The median follow-up for patients alive at the last date of follow-up was 4.0 years (range, 0.8 to 15.8 years). Tobacco use was high with a median number of pack-years of 27 (mean 30). Subsequently, we examined which of the variables were correlated (Figure 1D). It was observed that the subgroup of patients with an HPV+/p16+ tumour correlated with: a shorter smoking history, higher N-stage, better performance status, lower T-stage, and younger age (Figure 1D).

Univariate and multivariate analysis of factors influencing overall survival

In the OS analysis, 1,524 patients (18 were lost to follow-up) were included, with a total of 690 deaths. In the total cohort, 1,336 patients were treated with curative regimes; of these patients, 543 died during follow-up. The subgroup of HPV+/p16+ had a markedly better prognosis compared with the other combinations of HPV/p16 (Table 1 and Figure 1A) ($p < 0.05$). Other factors significantly associated with a poorer OS in the univariate analysis were: pack years of smoking, treatment, tumour location, and T, N, and M classification (Table 1). We subsequently constructed a multivariate Cox regression model to examine which factors independently predicted OS. The HPV/p16 status was an independent predictor for OS, even when adjusted for T-stage, N-stage, treatment, smoking history, age, and performance status (Supplementary Table S1). To better illustrate any excess mortality in the HPV+/p16+ and HPV-/p16- subgroups, Kaplan-Meier survival curves were compared with the mortality rates of the background population matched for age, gender, and calendar period (Figure 1E and 1F). Interestingly, patient with stage T1-T2 tumours in the HPV+/p16+ subgroup had almost no excess mortality compared with the background population (Figure 1E), which was in stark contrast to the HPV−/p16− subgroup (Figure 1F).

Table 1: Patient characteristics and their relationship to overall survival

Covariate	n	95% CI	
Gender	1542		
Female	424	(ref)	
Male	1118	1.1 (CI: 0.97-1.4)	
Age, cont.	1542	1.03 (CI: 1.02-1.04)	
Year of diagnosis, cont.	1542	0.95 (0.93-0.97)	
Pk years, cont.	1395	1.012 (CI: 1.010-1.015)	
Pk years, cat.	1395		
0 pk-years	380	(ref)	
1-10 pk years	104	1.2 (CI: 0.80-1.8)	
11-20 pk years*	135	1.4 (CI: 1.0-2.0)	
>20 pk years**	776	2.7 (CI:2.1-3.3)	
HPV/p16 Status	1502		
HPV+/p16+	826	(ref)	
HPV+/p16-**	76	4.1 (CI: 3.0-5.6)	
HPV-/p16+**	93	2.7 (CI: 2.0-3.6)	
HPV-/p16-**	507	4.9 (CI: 4.1-5.9)	
Tumor location	1542		
NSTSCC	355	(ref)	
STSCC**	764	0.38 (CI: 0.32-0.45)	
BOT**	423	0.54 (CI: 0.45-0.66)	
Treatment	1468		
RT	751	(ref)	
RT and chemotherapy **	585	0.5 (CI: 0.5-0.7)	
Palliative**	96	5.5 (CI: 4.3-6.9)	
No treatment**	36	7.8 (CI: 5.5-11.0)	
T-stage	1489		
T1	297	(ref)	
T2*	654	1.4 (CI: 1.1-1.7)	
T3**	361	2.8 (CI: 2.2-3.7)	
T4**	177	5.2 (CI: 3.9-6.8)	
N-stage	1493		
N0	299	(ref)	
N1	271	0.9 (CI: 0.7-1.1)	
N2	800	0.9 (CI: 0.7-1.1)	
N3**	123	1.5 (CI: 1.2-2.0)	
M-stage	1488		
M0	1455	(ref)	
M1**	33	5.1 (CI: 3.5-7.4)	
TNM-stage	1491		
Stage I	48	(ref)	
Stage II	150	1.3 (CI: 0.7-2.2)	
Stage III	312	1.4 (CI: 0.8-2.3)	
Stage IV*	981	1.8 (CI: 1.1-2.9)	
Center	1501		
Copenhagen	814	(ref)	
Herlev**	687	0.72 (CI: 0.62-0.84)	
Performance score	1157		
0	772	(ref)	
1**	283	2.9 (CI: 2.4-3.5)	
2**	71	6.5 (CI: 4.9-3.5)	
3**	27	10.2 (CI: 6.7-15)	
4**	4	241 (CI: 82-714)	

Hazzard ratio

Footnote: The hazard ratios for age, year of diagnosis, and pack years represents the hazard ratio per year increase. RT: radiotherapy. BSCC: Base of tongue squamous cell carcinoma. STSCC and NSTSCC: Specified- and non-specified tonsillar squamous cell carcinomas.

Univariate and multivariate analysis of factors influencing time to progression

In total, 376 patients experienced a progression, with 153 (19%) in the HPV+/p16+ subgroup vs. 182 (36%) in the HPV-/p16- subgroup ($P < 0.01$). The subgroup of HPV+/p16+ had a significantly longer TTP compared with the other combinations of HPV/p16 (Figure 1B) ($P < 0.05$). The median TTP for the total cohort was 9.7 months. The median TTP was 13 months for the HPV+/p16+ subgroup and 8.5 months ($P < 0.05$) for the HPV−/p16− subgroup. Median time to loco-regional progression was
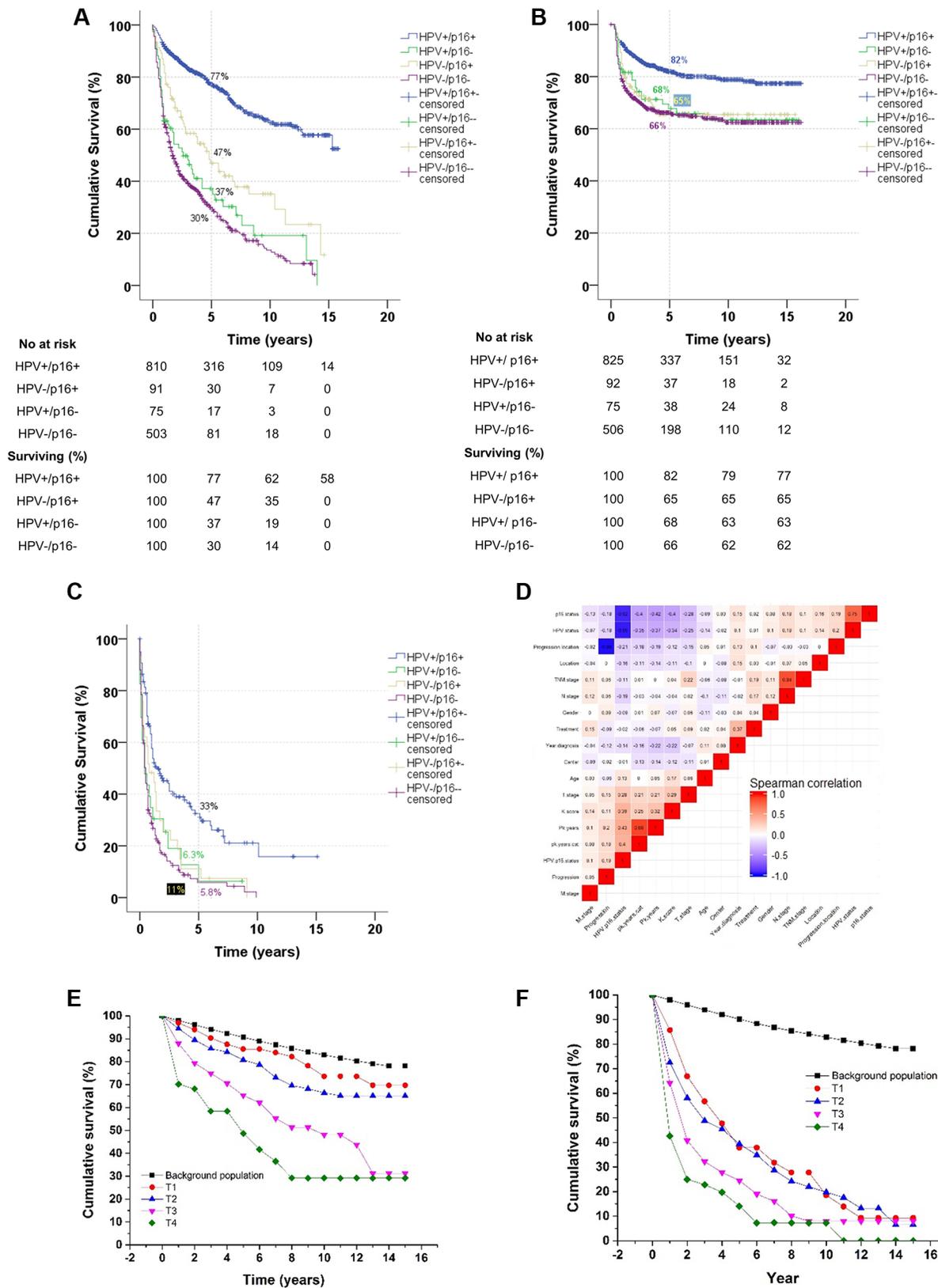
Figure 1: Kaplan-Meier curves for (**A**) overall survival, (**B**) time to progression, and (**C**) survival after progression. Correlation heatmap between variables (**D**). Comparison between survival in an age- and gender-matched background population and the HPV+/p16+ subgroup (**E**) and the HPV-/p16- subgroup (**F**) based on T classification of tumours.

7.6 months for the HPV+/p16 subgroup and 7.4 months for the HPV−/p16− group. The corresponding figures for distant progression were 18 and 11 months, respectively, ($P < 0.05$). The univariate analysis of factors influencing TTP included: pack years of smoking, HPV/16 status, and T, N, and M classification of tumours (Supplementary Table S2). A multivariate Cox regression model identified HPV/p16 status as an independent predictor of TTP with a hazard ratio of 3.0 (95% CI: 2.3–3.8) in the HPV−/p16− subgroup compared with the HPV+/p16+ subgroup, after adjustment (Supplementary Table S3).

Univariate and multivariate analysis of factors influencing survival after progression

The median SAP was 13 months in the HPV+/p16+ subgroup vs. 6 months in the HPV-/p16- subgroup ($P < 0.05$). The HPV+/p16+ status also independently influenced SAP positively with a hazard ratio of 2.2 (95% CI: 1.6–2.9) compared with the HPV–/p16– subgroup (Supplementary Table S4), when adjusted for other significant covariates. Performance status and progression location also independently influenced SAP (Supplementary Table S4).

Predictive nomograms

Nomograms (Figure 2A and 2B) were constructed to predict survival using the independent covariates identified in the multivariate Cox regression models. The median total points for the 1,523 patients used to fit the multivariate Cox model for OS was 101 (range, 0–421). The nomogram is used by totalling the points identified on the top scale for each independent covariate which is identified on the total points scale to identify the estimated median survival time (years) and the probability of 1-, 5- and 10-year survival.

The bias-corrected concordance index for this nomogram was 0.79 based on the bootstrap validated Cox model. The calibration curve (Supplementary Figure S1) illustrates how the predictions from the nomogram compare with actual outcomes. The corresponding information for the nomogram for TTP can be found in Supplementary Figure S2.

Prognostic index

A simplified prognostic index was constructed for the 5-year OS using seven independent prognostic factors. The index score is based on the sum total of factors, with points given for each of the following: treatment, HPV/p16-status, age, pack years, T-status, N-status, and performance status. Risk is assigned as follows: index score 0 to 4, low ($\geq$ 80% probability of surviving 5 years); index score 5 to 7, intermediate (< 80% and > 20% probability of surviving 5 years); and index score 8 or greater, high risk ($\leq$ 20% probability of surviving 5 years) (Table 2).

DISCUSSION

This study included the largest-to-date collection of 1,542 OPSCCs from a consecutive, population-based, non-selected cohort and assessed OS, TTP, and SAP as well as their relation to multiple clinico-pathological characteristics. From these results, we constructed and internally validated nomograms to predict OS and TTP. Furthermore, we established a simplified prognostic index based on the independent covariates for OS for use in categorizing patients into low-, intermediate-, and high-risk groups. Besides being useful for interpretation of the underlying Cox model, the nomogram combines independent prognostic factors and considers the importance of each on the probability of survival and progression. The prognostic models we present may facilitate discussions in clinical settings and aid in identifying low-risk patients that could be candidates for de-escalation therapy, as well as high-risk patients eligible for new clinical trials. In particular, HPV+/p16+ patients with T1-T2 tumours should be considered candidates for de-escalation therapy, as we demonstrated their survival is similar to the background population, and this would avoid the morbidity associated with therapy. Notably, at least nine de-escalation treatment trials are on-going or finishing [18]. Our nomograms are likely to be applicable to these and future trials, as we reported similar 5-year survival or progression rates as studies in North America;[19–21] Western [22–24], Southern [25], and Northern Europe [6, 26]; Australia [27]; and China [28].

An estimated 10% of all head and neck SCCs are p16+, but HPV– caused by alternative cellular misconfigurations leading to p16-overexpression.[29, 30] Therefore, it is suboptimal to include patients in de-escalation trials based on evaluation of a single biomarker (i.e. HPV or p16 alone) due to the risk of misclassification of tumours and thereby misallocation of patients with an undesired prognosis. [31, 32] One of the main findings in this study includes the identification of HPV+/p16+ as an important and independent predictor for improved OS, TTP, and SAP. Importantly, even though the subgroup of HPV+/p16+ patients was more likely to have a shorter smoking history and a better performance status, even when adjusted for these covariates, HPV+/p16+ was a significant and strong predictor for OS. The calibration for this model was robust and could explain 79% of the observed variability in OS in the cohort. Our findings are in accordance with similar studies addressing survival and progression in HPV+ or p16+; although, we present the first results of a population-based cohort with long-term follow-up. While another smaller study in a region with low HPV-prevalence (below 20%) constructed nomograms for OS and TTP in OPSCC, this study was not population-based and did not address SAP.[33] Furthermore, this study did not convincingly demonstrate which patients were lost to inclusion and follow-up, resulting in less robust external validity.

Table 2: Prognostic index for overall survival

Points	0	1	2	3	
Treatment	RT and chemotherapy	RT	Palliative	NA	No treatment
Age	0–44	45–64	65-89	90 and above	NA
HPV/p16	HPV+/p16+	HPV–/p16+	HPV–/p16– or HPV+/p16–	NA	NA
T classification	T1	T2-T3	T4	NA	NA
N classification	N0	N1-N2	N3	NA	NA
Pack years	0–59	59–179	180 and above		
Performance status	0–1	2–3	NA	NA	4

Footnote: It is possible to obtain between 0 and 19 points. Low-risk is from 0–4 points (≥ 80 % risk of being alive after 5 years), medium risk 5–7 points (between < 80 and > 30% chance of being alive after 5 years), and high-risk > 7 points (≤ 30% change of being alive after 5 years).
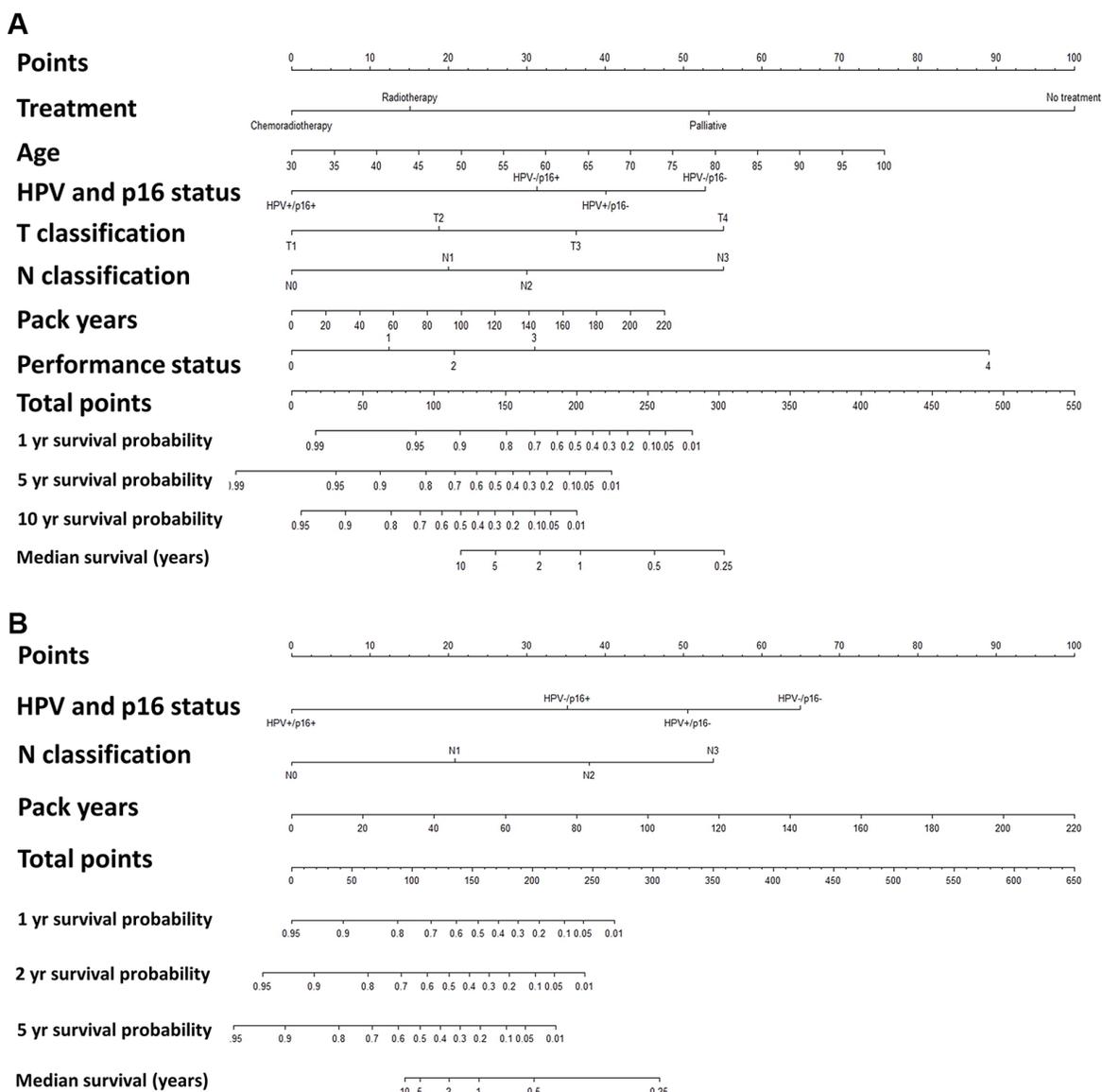
Figure 2: Nomogram for overall survival (A) and time to progression (B). The nomogram is used by totalling the points identified on the top scale for each independent covariate. The total points scale is used to identify the estimated median survival time (years) and the probability of 1-, 5-, and 10-year survival.

Our nomograms support the latest AHR-stage classification proposed by O'Sullivan *et al.* [34] and comply with the notion that T1–T2 and N0–N2 show significantly better OS, TTP, and SAP compared with ≥ T3 and ≥ N3 tumours, indicating that down-staging of N-classified OPSCCs is reasonable. Noticeably, our findings, in contrast to O'Sullivan *et al*, include both HPV and p16 status and patients with distant disease. Our models encourage further studies to better understand whether all N-classified tumours are eligible for down-staging. Although our findings rely on a population-based cohort and selection-bias is minimized, the nomograms have certain limitations. It would be most appropriate to apply the model as the last step in a clinical setting, and the models are not necessarily applicable for changing patients' decisions not to accept further treatment since we do not know her/his response to therapy. With respect to accuracy, the CIs at the various predicted probabilities of recurrence should be considered if using these nomograms in clinical settings. Although this model is internally validated, it could be strengthened by external validation, e.g. in similar population-based cohorts. Future work might focus on validating our results and incorporating additional prognostic factors including robotic, curative surgery, and specific salvage treatment for relapsed disease as well as further outcome measures such as time to treatment or to pathological evaluation.

MATERIALS AND METHODS

We included patients diagnosed with OPSCC in Eastern Denmark between 2000–2014. The reasons for exclusion are illustrated in Figure 3. We targeted all patients diagnosed with OPSCC in the Eastern part of Denmark cohort with detailed information on tumour (HPV, p16, stage, site) and clinical characteristics (follow-up, treatment, progression). Using the unique resident-code from the Danish Civil Registration System, we linked two national registries, the prospectively maintained Danish Head and Neck Cancer Group (DAHANCA) [35] database and the Danish Pathology Data Registry (DPDR) [36], to identify patients. Patient-data were retrieved from these databases as well as from medical records. The patient cohort from 2000 to 2010 has previously been described [37, 38].

Tumour-site, histology, tumour-grade, HPV-PCR, and p16 immunohistochemistry

An H&E-stained section of each tumour was reviewed by an expert head and neck pathologist. Based on the finding of tissue-specific structures and clinical information, tumours were divided into base of tongue squamous cell carcinoma (BSCC) and specified- and non-specified tonsillar squamous cell carcinomas (STSCC and NSTSCC). The latter was according to the certainty of

origin. The pathologist validated the diagnosis of OPSCC. The p16 staining was considered positive if there was a strong and diffuse nuclear and cytoplasmic reaction in more than 75% of the tumour cells [39]. The HPV DNA-PCR was performed with DNA isolated from FFPE tumour sections. Methods for immunohistochemistry and HPV-PCR have previously been described [37, 38].

Overall survival, time to progression, and statistics

The date of the last follow-up was the 1st of November 2015. The OS was defined as the time from diagnosis of OPSCC to death due to any cause. Data on vital status (death) stems from the real-time, national, patient-administrative *Green System (GSOpen)*. Progression was only considered if it was biopsy-verified. Data on progression stem from the DPDR. The TTP was defined as the time from diagnosis of OPSCC to time of progression at any site. In the analysis of TTP, patients were censored either at the last date of follow-up or at time of death. Patients alive at the last date of follow-up were censored in survival analyses for OS and SAP. The date of diagnosis was used to reflect the date of treatment, since the vast majority of Danish patients initiate treatment within 1 month [40, 41]. The SAP was defined as the time from progression to death due to any cause. Curative radiotherapy regimens consisted of 66–68 GY, divided into 33–34 fractions given 6 days a week. From 2007, stage III-IV patients were offered concomitant cisplatin, if tolerated. Covariates available for adjustment are described in Table 1. The following variables were coded as continuous variables in the survival analyses: age and year of diagnosis. The remaining variables were coded as categorical variables unless otherwise stated. Data on smoking (20 cigarettes per day for 1 year = 1 pack year) were retrieved from the DAHANCA database or medical files. Kaplan-Meier curves were used to illustrate survival differences and significant differences were assessed with log-rank tests. Hazard ratios for death, OS, and TTP were calculated by univariate Cox-regression with log-rank tests for each parameter (Table 1). To evaluate which covariates independently influenced survival, we performed multivariate cox regression analyses with the same factors used in the univariate analyses (Table 3 and Supplementary Tables S2 and S3) with backward elimination, until only significant factors remained. The stopping rule in the backward elimination of factors was based on Akaike's Information Criteria, with the statistics based on the pooled residual chi-square of the model with the R package rms and the function fastbw [42]. The models were internally validated with 200 bootstrappings. The difference in Somers' D between the training and test set was used to evaluate the optimism in the predictive accuracy (i.e. a measure of

the difference between observed and predicted values). Somers' D can range from –1 (all pairs disagree) to 1 (all pairs agree) and can be converted to the more well-known AUC via the following formula, AUC = Dyx/2 + 0.5. The calibration of the models was subsequently tested using the rms package in R. To test whether the assumption of proportional hazards was violated, we plotted Schoenfeld residuals for the univariate analyses [43]. Additionally, the multivariate models for survival were examined for violations of the proportional hazards assumption with the function cox.zph. None of the final models violated the proportional hazards assumption. To test for correlations between the covariates, we used Spearman's rank correlation (Figure 1D) [44].

The analysis comparing survival in the cohort with the background population was performed for each year of observation according to the Kaplan-Meier survival rates in the cohort. The survival analysis for the background population was performed for a gender and age-matched population in the same calendar-years as the cohort with mortality figures from the Danish National Department of Statistics (http:\\www.statistikbanken.dk/statbank5a/default.asp?w=1344). Missing data were left out of the analysis, and variables with missing values are illustrated in Table 1. A *p* value < 0.05 was considered significant. Data were analysed with SPSS version 23 (SPSS Inc., Chicago, IL, USA) and *R* statistics version 3.03.
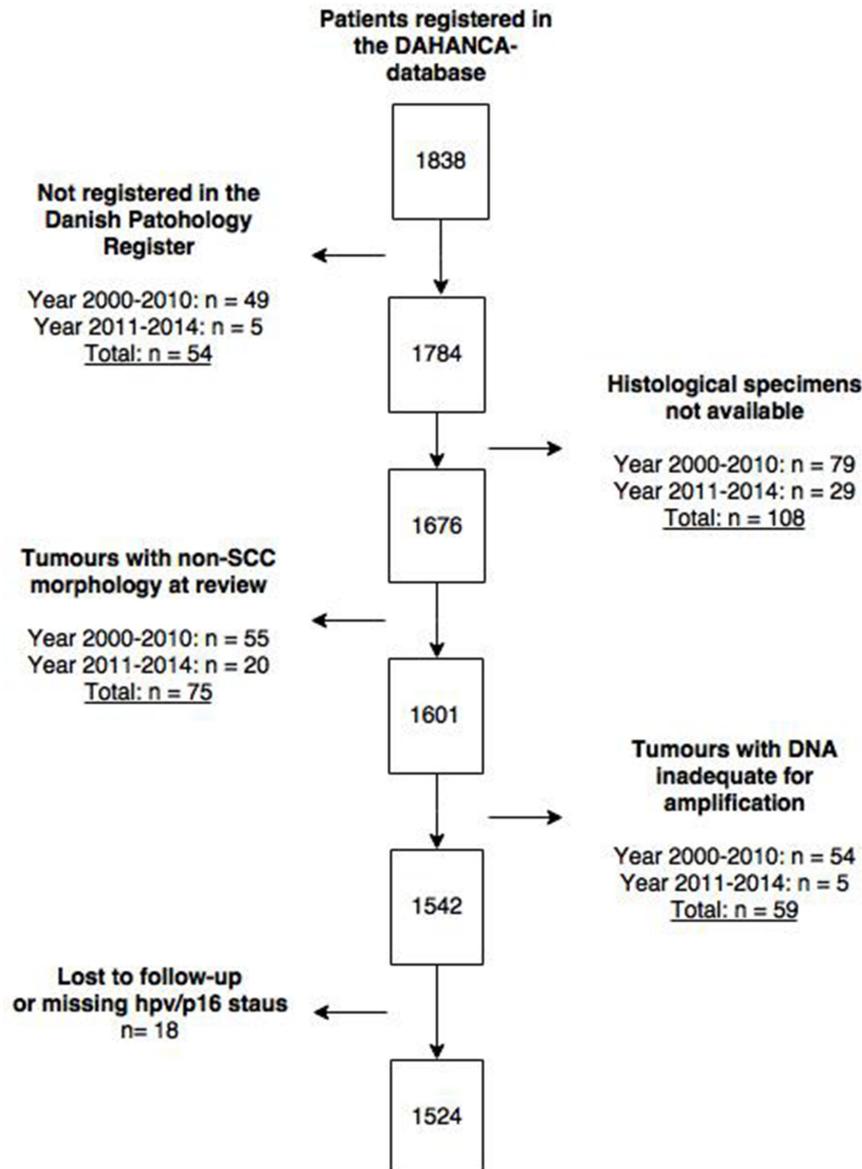
Figure 3: CONSORT flow-diagram

Table 3: Independent covariates for overall survival in the multivariate Cox regression model

	Hazard ratio for death	Lower CI	Upper CI	*P*
T classification				
T1	Ref			
T2	1.55	1.13	2.11	0.0060
T3	2.32	1.69	3.19	< 0.0001
T3	3.59	2.52	5.11	< 0.0001
N classification				
N0	Ref			
N1	1.59	1.17	2.17	0.0034
N2	2.00	1.54	2.60	< 0.0001
N3	3.59	2.48	5.19	< 0.0001
Treatment				
Radiotherapy	Ref			
Chemoradiotherapy	0.7	0.6	0.9	0.0042
Palliative	2.4	1.7	3.5	< 0.0001
No treatment	7.2	4.2	12.1	< 0.0001
HPV/p16				
HPV+/p16+	Ref			
HPV+/p16–	2.5	1.8	3.7	< 0.0001
HPV–/p16+	2.1	1.4	3.0	0.0001
HPV–/p16–	3.4	2.7	4.3	< 0.0001
Age	1.03	1.01	1.04	< 0.0001
Pack years	1.005	1.002	1.008	0.0015
Performance score				
0	Ref			
1	1.3	1.1	1.7	0.0105
2	2.1	1.5	2.9	< 0.0001
3	1.6	0.9	2.8	0.0905
4	7.9	2.1	29.9	0.0024

Footnote: CI, 95% confidence interval. The hazard ratio for age and pack years represents the increase in hazard ratio per 1 year increase.

Validation and calibration of multivariate cox regression models

When performing multivariate analysis of survival, it is of interest to examine how well the models can explain the variation in survival observed in the cohort. The Somers' D for the final model for OS was 0.60, which corresponds to an AUC of 0.80. This was validated with bootstrapping to prevent possible over fitting using the Somers' D rank correlation for the predicted log hazard and observed survival time. The optimism in the model was 0.0064, and the expected bias-corrected AUC was 0.79, suggesting that the model was not over fitted. Finally, the models were evaluated for calibration accuracy in predicting the probability of surviving 5 years. This was performed with bootstrapping to estimate the optimism from the models for how the predicted 5-year survival estimates compared with the observed survival estimates (Supplementary Figure S1). The corresponding AUC and optimism in the multivariate Cox regression model for TTP were 0.68 and 0.010, respectively, and the calibration curve is shown in Supplementary Figure S2.

CONFLICTS OF INTEREST

The authors declare no conflicts of interest.

Author's contributions

CGL, DHJ, and CVB designed the study. CGL, EG, and ALC extracted data from files and registries. EG and ALC performed the laboratory analysis. KK, CHO, and LA performed the pathology review. DHJ performed the statistical analysis. CGL, DHJ, and CVB led the writing. All authors provided conceptual input, interpreted the findings, and contributed in significant ways to the final article.

FUNDING

REFERENCES

1. Lawrence MS, Sougnez C, Lichtenstein L, Cibulskis K, Lander E, Gabriel SB, Getz G, Ally A, Balasundaram M, Birol I, Bowlby R, Brooks D, Butterfield YSN, et al. Comprehensive genomic characterization of head and neck squamous cell carcinomas. Nature [Internet]. Nature Publishing Group; 2015; 517:576–82. doi: 10.1038/nature14129.

2. Hayes DN, Van Waes C, Seiwert TY. Genetic Landscape of Human Papillomavirus-Associated Head and Neck Cancer and Comparison to Tobacco-Related Tumors. J Clin Oncol. 2015; 33:3227–34. doi: 10.1200/JCO.2015.62.1086.

3. Lewis JS, Khan RA, Masand RP, Chernock RD, Zhang Q, Al-Naief NS, Muller S, McHugh JB, Prasad ML, Brandwein-Gensler M, Perez-Ordonez B, El-Mofty SK. Recognition of nonkeratinizing morphology in oropharyngeal squamous cell carcinoma - a prospective cohort and interobserver variability study. Histopathology [Internet]. 2012; 60:427–36. doi: 10.1111/j.1365–2559.2011.04092.x.

4. Marur S, D'Souza G, Westra WH, Forastiere AA. HPV-associated head and neck cancer: a virus-related cancer epidemic. Lancet Oncol. 2010; 11:781–9.

5. Gillison ML, Koch WM, Capone RB, Spafford M, Westra WH, Wu L, Zahurak ML, Daniel RW, Viglione M, Symer DE, Shah K V, Sidransky D, Gillison ML, et al. Evidence for a causal association between human papillomavirus and a subset of head and neck cancers. J Natl Cancer Inst [Internet]. 2000; 92:709–20. Available from http://www.ncbi.nlm.nih.gov/pubmed/10793107.

6. Mellin H, Friesland S, Lewensohn R, Dalianis T, Munck-Wikland E. Human papillomavirus (HPV) DNA in tonsillar cancer: clinical correlates, risk of relapse, and survival. Int J Cancer. 2000; 89:300–4. doi: 10.1002/1097-0215(20000520)89:3<300::AID-IJC14>3.0.CO;2-G [pii].

7. Weinberger PM, Yu Z, Haffty BG, Kowalski D, Harigopal M, Sasaki C, Rimm DL, Psyrri A. Prognostic significance of p16 protein levels in oropharyngeal squamous cell cancer. Clin Cancer Res [Internet]. 2004 [cited 2016 Jan 4]; 10:5684–91. doi: 10.1158/1078-0432.CCR-04-0448.

8. Klussmann JP, Gultekin E, Weissenborn SJ, Wieland U, Dries V, Dienes HP, Eckel HE, Pfister HJ, Fuchs PG, Gültekin E, Weissenborn SJ, Wieland U, Dries V, et al. Expression of p16 protein identifies a distinct entity of tonsillar carcinomas associated with human papillomavirus. Am J Pathol [Internet]. 2003; 162:747–53. doi: 10.1016/S0002-9440(10)63871-0.

9. Garnaes E, Frederiksen K, Kiss K, Andersen L, Therkildsen MH, Franzmann M, Specht L, Andersen E, Norrild B, Kjaer SK, Christian von Buchwald. Both HPV DNA and p16 are necessary for accurate prognostication in tonsillar and base of tongue cancer: insights from a large population-based study. In review.

10. Trosman S, Al-Khudari S, Koyfman SA, Nwizu T, Greskovich JF, Scharpf J, Lamarre E, Khan M, Adelstein DJ, Burkey BB. Distant metastatic failure patterns in squamous cell cancer of the oropharynx (SCCOP) treated with chemoradiation: The impact of human papillomavirus (HPV). Int J Radiat Oncol Biol Phys. 2014; 88:471. doi: http://dx.doi.org/10.1016/j.ijrobp.2013.11.037.

11. McBride SM, Busse PM, Clark JR, Wirth LJ, Ancukiewicz M, Chan AW. Long-term survival after distant metastasis in patients with oropharyngeal cancer. Oral Oncol [Internet]. 2014 [cited 2016 Jan 5]; 50:208–12. doi: 10.1016/j.oraloncology.2013.10.020.

12. Guo T, Rettig E, Fakhry C. Understanding the impact of survival and human papillomavirus tumor status on timing of recurrence in oropharyngeal squamous cell carcinoma. Oral Oncol [Internet]. 2016; 52:97–103. doi: 10.1016/j.oraloncology.2015.10.016.

13. Kattan MW, Eastham JA, Stapleton AM, Wheeler TM, Scardino PT. A preoperative nomogram for disease recurrence following radical prostatectomy for prostate cancer. J Natl Cancer Inst. 1998; 90:766–71. doi: 10.1093/jnci/90.10.766.

14. Wierda WG, O'Brien S, Wang X, Faderl S, Ferrajoli A, Do KA, Cortes J, Thomas D, Garcia-Manero G, Koller C, Beran M, Giles F, Ravandi F, et al. Prognostic nomogram and index for overall survival in previously untreated patients with chronic lymphocytic leukemia. Blood. 2007; 109:4679–85. doi: 10.1182/blood-2005-12-051458.

15. Ying HQ, Deng QW, He BS, Pan YQ, Wang F, Sun HL, Chen J, Liu X, Wang SK. The prognostic value of preoperative NLR, d-NLR, PLR and LMR for predicting clinical outcome in surgical colorectal cancer patients. Med Oncol. 2014; 31:305. doi: 10.1007/s12032-014-0305-0.

16. Ganly I, Amit M, Kou L, Palmer FL, Migliacci J, Katabi N, Yu C, Kattan MW, Binenbaum Y, Sharma K, Naomi R, Abib A, Miles B, et al. Nomograms for predicting survival and recurrence in patients with adenoid cystic carcinoma. An international collaborative study. Eur J Cancer [Internet]. 2015; 51:2768–76. doi: 10.1016/j.ejca.2015.09.004.

17. Su D, Zhou X, Chen Q, Jiang Y, Yang X, Zheng W, Tao K, Wu J, Yan Z, Liu L, Wu S, Mao W. Prognostic nomogram for thoracic esophageal squamous cell carcinoma after

radical esophagectomy. PLoS One. 2015; 10. doi: 10.1371/journal.pone.0124437.

18. Masterson L, Moualed D, Liu ZW, Howard JEF, Dwivedi RC, Tysome JR, Benson R, Sterling JC, Sudhoff H, Jani P, Goon PKC. De-escalation treatment protocols for human papillomavirus-associated oropharyngeal squamous cell carcinoma: a systematic review and meta-analysis of current clinical trials. Eur J Cancer [Internet]. 2014; 50:2636–48. doi: 10.1016/j.ejca.2014.07.001.

19. 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, et al. Human papillomavirus and survival of patients with oropharyngeal cancer. N Engl J Med. 2010; 363:24–35. doi: 10.1056/NEJMoa0912217.

20. Fakhry C, Zhang Q, Nguyen-Tan PF, Rosenthal D, El-Naggar A, Garden AS, Soulieres D, Trotti A, Avizonis V, Ridge JA, Harris J, Le Q-T, Gillison M. Human papillomavirus and overall survival after progression of oropharyngeal squamous cell carcinoma. J Clin Oncol. 2014; 32:3365–73. doi: 10.1200/JCO.2014.55.1937.

21. Agoston ES, Robinson SJ, Mehra KK, Birch C, Semmel D, Mirkovic J, Haddad RI, Posner MR, Kindelberger D, Krane JF, Brodsky J, Crum CP. Polymerase chain reaction detection of HPV in squamous carcinoma of the oropharynx. Am J Clin Pathol. 2010; 134:36–41.

22. Hafkamp HC, Manni JJ, Haesevoets A, Voogd AC, Schepers M, Bot FJ, Hopman AH, Ramaekers FC, Speel EJ. Marked differences in survival rate between smokers and nonsmokers with HPV 16-associated tonsillar carcinomas. Int J cancerJournal Int du cancer. 2008; 122:2656–64.

23. Preuss SF, Weinell A, Molitor M, Stenner M, Semrau R, Drebber U, Weissenborn SJ, Speel EJM, Wittekindt C, Guntinas-Lichius O, Hoffmann TK, Eslick GD, Klussmann JP. Nuclear survivin expression is associated with HPV-independent carcinogenesis and is an indicator of poor prognosis in oropharyngeal cancer. Br J Cancer [Internet]. 2008; 98:627–32. doi: 10.1038/sj.bjc.6604192.

24. Junor E, Kerr G, Oniscu A, Campbell S, Kouzeli I, Gourley C, Cuschieri K. Benefit of chemotherapy as part of treatment for HPV DNA-positive but p16-negative squamous cell carcinoma of the oropharynx. Br J Cancer. 2012; 106:358–65.

25. Rittà M, De Andrea M, Mondini M, Mazibrada J, Giordano C, Pecorari G, Garzaro M, Landolfo V, Schena M, Chiusa L, Landolfo S. Cell cycle and viral and immunologic profiles of head and neck squamous cell carcinoma as predictable variables of tumor progression. Head Neck [Internet]. 2009; 31:318–27. doi: 10.1002/hed.20977.

26. Dahlstrand HM, Lindquist D, Bjornestal L, Ohlsson A, Dalianis T, Munck-Wikland E, Elmberger G. P16(INK4a) correlates to human papillomavirus presence, response to radiotherapy and clinical outcome in tonsillar carcinoma. Anticancer Res. 2005; 25:4375–83.

27. Hong AM, Dobbins TA, Lee CS, Jones D, Harnett GB, Armstrong BK, Clark JR, Milross CG, Kim J, O'Brien CJ, Rose BR. Human papillomavirus predicts outcome in oropharyngeal cancer in patients treated primarily with surgery or radiation therapy. Br J Cancer. 2010; 103:1510–7.

28. Zhao D, Xu Q, Chen X, Fan M. Human papillomavirus as an independent predictor in oral squamous cell cancer. Int J Oral Sci [Internet]. 2009; 1:119–25. doi: 10.4248/IJOS.09015.

29. Wang H, Sun R, Lin H, Hu WH. P16INK4A as a surrogate biomarker for human papillomavirus-associated oropharyngeal carcinoma: Consideration of some aspects. Cancer Science. 2013. p. 1553–9. doi: 10.1111/cas.12287

30. Lingen MW, Xiao W, Schmitt A, Jiang B, Pickard R, Kreinbrink P, Perez-Ordonez B, Jordan RC, Gillison ML. Low etiologic fraction for high-risk human papillomavirus in oral cavity squamous cell carcinomas. Oral Oncol. 2013; 49:1–8.

31. Chung CH, Zhang Q, Kong CS, Harris J, Fertig EJ, Harari PM, Wang D, Redmond KP, Shenouda G, Trotti A, Raben D, Gillison ML, Jordan RC, et al. p16 protein expression and human papillomavirus status as prognostic biomarkers of nonoropharyngeal head and neck squamous cell carcinoma. J Clin Oncol [Internet]. 2014; 32:3930–8. doi: 10.1200/JCO.2013.54.5228.

32. Seiwert TY. Ties that bind: p16 as a prognostic biomarker and the need for high-accuracy human papillomavirus testing. J Clin Oncol [Internet]. 2014 [cited 2016 Apr 14]; 32:3914–6. doi: 10.1200/JCO.2014.57.9268.

33. Rios Velazquez E, Hoebers F, Aerts HJWL, Rietbergen MM, Brakenhoff RH, Leemans RC, Speel EJ, Straetmans J, Kremer B, Lambin P. Externally validated HPV-based prognostic nomogram for oropharyngeal carcinoma patients yields more accurate predictions than TNM staging. Radiother Oncol. 2014; 113:324–30. doi: 10.1016/j.radonc.2014.09.005.

34. O'Sullivan B, Huang SH, Su J, Garden AS, Sturgis EM, Dahlstrom K, Lee N, Riaz N, Pei X, Koyfman SA, Adelstein D, Burkey BB, Friborg J, et al. Development and validation of a staging system for HPV-related oropharyngeal cancer by the International Collaboration on Oropharyngeal cancer Network for Staging (ICON-S): a multicentre cohort study. Lancet Oncol [Internet]. Elsevier; 2016; doi: 10.1016/S1470-2045(15)00560-4.

35. Danish Head and Neck Cancer Group (DAHANCA).

36. Bjerregaard B, Larsen OB. The Danish Pathology Register. Scand J Public Health. 2011; 39:72–4. doi: 10.1177/1403494810393563.

37. Garnaes E, Kiss K, Andersen L, Therkildsen MH, Franzmann MB, Filtenborg-Barnkob B, Hoegdall E, Krenk L, Josiassen M, Lajer CB, Specht L, Frederiksen K, Friis-Hansen L, et al. A high and increasing HPV prevalence in tonsillar cancers in Eastern Denmark, 2000–2010: The largest registry-based study to date. Int J Cancer. 2014. doi: 10.1002/ijc.29254.

38. Garnaes E, Kiss K, Andersen L, Therkildsen MH, Franzmann MB, Filtenborg-Barnkob B, Hoegdall E, Lajer CB, Andersen E, Specht L, Joenson L, Frederiksen K, Friis-Hansen L, et al. Increasing incidence of base of tongue cancers from 2000 to 2010 due to HPV: the largest demographic study of 210 Danish patients. Br J Cancer [Internet]. Cancer Research UK; 2015; 113:131–4. doi: 10.1038/bjc.2015.198.

39. Grønhøj Larsen C, Gyldenløve M, Jensen DH, Therkildsen MH, Kiss K, Norrild B, Konge L, von Buchwald C. Correlation between human papillomavirus and p16 overexpression in oropharyngeal tumours: a systematic review. Br J Cancer [Internet]. 2014; 110:1587–94. doi: 10.1038/bjc.2014.42.

40. Sorensen JR, Johansen J, Gano L, Sørensen JA, Larsen SR, Andersen PB, Thomassen A, Godballe C. A "package solution" fast track program can reduce the diagnostic waiting time in head and neck cancer. Eur Arch Otorhinolaryngol [Internet]. 2014; 271:1163–70. doi: 10.1007/s00405-013-2584-z.

41. Lyhne NM, Christensen a., Alanin MC, Bruun MT, Jung TH, Bruhn M a., Jespersen JBB, Kristensen C a., Andersen E, Godballe C, Buchwald C, Bundgaard T, Johansen J, et al. Waiting times for diagnosis and treatment of head and neck cancer in Denmark in 2010 compared to 1992 and 2002. Eur J Cancer. 2013; 49:1627–33. doi: 10.1016/j.ejca.2012.11.034.

42. Harrell FE. Regression modeling strategies. With applications to linear models, logistic regression, and survival analysis. Springer Series in Statistics. 2001. XXIII + 568 p. doi: 10.1007/978-1-4757-3462-1.

43. Grambsch PM, Therneau TM. Proportional hazards tests and diagnostics based on weighted residuals. Biometrika. 1994; 81:515–26. doi: 10.1093/biomet/81.3.515.

44. Benjamini Y, Hochberg Y. Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing [Internet]. Journal of the Royal Statistical Society. Series B (Methodological). 1995. p. 289–300. doi: 10.2307/2346101.