The stem cell factor HMGA2 is expressed in non-HPV-associated head and neck squamous cell carcinoma and predicts patient survival of distinct subsites

Kathrin Günther, Ronja Foraita, Juliane Friemel, Frauke Günther, Jörn Bullerdiek, Rolf Nimzyk, Dominique Nadine Markowski, Thomas Behrens, Wolfgang Ahrens

DOI
10.1158/1055-9965.EPI-16-0492

Published in
Cancer Epidemiology, Biomarkers & Prevention

Document version
Accepted manuscript

Online publication date
16 October 2016

Corresponding author
Kathrin Günther

Citation
Günther K, Foraita R, Friemel J, Günther F, Bullerdiek J, Nimzyk R, et al. The stem cell factor HMGA2 is expressed in non-HPV-associated head and neck squamous cell carcinoma and predicts patient survival of distinct subsites. Cancer Epidemiol Biomarkers Prev. 2017;26(2):197-205.

The stem cell factor *HMGA2* is expressed in non-HPV associated head and neck squamous cell carcinoma and predicts patient survival of distinct subsites

Kathrin Günther[1], Ronja Foraita[1], Juliane Friemel[1], Frauke Günther[1], Jörn Bullerdiek[2], Rolf Nimzyk[2], Dominique Nadine Markowski[2], Thomas Behrens[3,4], Wolfgang Ahrens[1]*

[1] Leibniz Institute for Prevention Research and Epidemiology – BIPS, Bremen, Germany

[2] Center for Human Genetics (ZHG), University of Bremen, Bremen, Germany

[3] Institute for Prevention and Occupational Medicine of the German Social Accident Insurance, Institute of the Ruhr-Universität Bochum (IPA), Bochum, Germany

[4] Protein Research Unit Ruhr within Europe (PURE), Ruhr-University Bochum, Germany

Running Title: HMGA2 expression and survival of HNSCC patients

Keywords: head and neck squamous cell carcinoma, HMGA2, gene expression, genetic biomarker, survival

Financial support: This project was funded by the German Research Foundation (DFG research grant AH 82/6-1 (W. Ahrens) and BU 592/11-1 (J. Bullerdiek)) and by the German Cancer Aid (No. 110276 (W. Ahrens, J. Bullerdiek)).

* Correspondence to:

Kathrin Günther, Leibniz Institute for Prevention Research and Epidemiology – BIPS, Achterstraße 30, D-28395 Bremen;

Tel.: 0049/421/218-56825

Fax: 0049/421/218-56821

E-Mail: kguenth@leibniz-bips.de

Disclosure of Potential Conflicts of Interest: No potential conflicts of interest were disclosed.

Words: Abstract: 250; Manuscript: 3295

Number of References: 47

Number of Tables: 3

Number of Figures: 3

Number of Supplementary Tables: 2

Background: The transcription factor high-mobility AT-hook 2 (*HMGA2*) is involved in stem cell renewal and is expressed in many tumor tissues. Head and neck squamous cell carcinomas (HNSCC) comprise tumors of the upper aerodigestive tract and are characterized by high recurrence rates that represent a challenge to patient management. The study addresses the potential of *HMGA2* as a molecular biomarker for HNSCC patient survival.

Methods: Patients with HNSCC of the larynx, pharynx, tonsils, or oral cavity were recruited in a hospital-based case-control study (n=202). Quantitative expression of HMGA2 in tumor tissues was measured by RT-PCR. In a 6-10 year follow-up, secondary cancers, vital status and cause of death were ascertained. The hazard ratios (HR) and 95% confidence intervals (CI) for overall, tumor-specific and progression-free survival were estimated by Cox proportional hazards with *HMGA2* expression level as the independent variable.

Results: High *HMGA2* expression in tumor tissues of HNSCC patients was significantly correlated with negative HPV status (p=0.01), and associated with shorter overall survival time. In Cox regression modelling, *HMGA2* expression yielded a risk increase for overall and tumor-specific death in subsets of HNSCC patients, i.e. laryngeal cancer patients (overall survival: HR=4.00; CI 95%:1.18-13.62) and in oral cancer patients (tumor-specific survival: HR=2.88; CI 95%:1.06-7.84), but not in patients with pharyngeal and tonsillar HNSCC.

Conclusions: *HMGA2* expression is associated with a risk increase for adverse outcomes in patients with HNSCC of the larynx and oral cavity.

Impact: The understanding of stem cell signaling in HNSCC may offer new strategies for cancer treatment.

Introduction

Head and neck squamous cell carcinomas (HNSCC) are epithelial tumors representing a heterogeneous disease entity, which encompasses a variety of tumors originating in the oral cavity, larynx and pharynx with differences in epidemiology, etiology and therapeutic approach. Over the past decades, HNSCC incidence rates increased in several countries in Eastern and Northern Europe and among females in Southern and Western Europe (1-3). Smoking, alcohol consumption and poor oral hygiene, the most easily preventable cancer causes, are associated with an elevated HNSCC risk (4-6). HNSCC induced by human papilloma virus (HPV) via the oncoproteins E6/7 and HNSCC caused by other factors (such as smoking and alcohol consumption) are two separate entities, with distinct etiologies, clinical characteristics, prognoses and a different epidemiology and molecular basis.

The molecular pathogenesis of HNSCC is not yet completely understood, a fact that impairs the improvement therapeutic approaches. Different genetic biomarkers have been proposed to identify patients who are at risk for recurrent tumors or aggressive disease progress. Most prominently, the cell cycle proteins cyclin D1, p21, p53 and MDM2 have been described (7, 8). Other targets comprise growth factor receptors like FGFR, VEGF and EGFR (9-11).

The architectonic transcription factor high mobility AT-hook 2 (HMGA2), is expressed during early embryogenesis and in cell differentiation (12-15). It has been attributed to the p53 cell cycle pathway (16, 17). Targeted inhibition of HMGA2 induces apoptosis and chemosensitization in p-53 mutant tumorspheres (18). In neuronal cells, HMGA2 was reported to regulate stemness (19). *HMGA2* expression in tumor tissues has been detected in many cancer sites, i.e. the digestive tract, lung, thyroid gland, urinary bladder, liver, testis and pancreas (20-26). In HNSCC of the oral cavity, *HMGA*2 expression was associated with poor survival in a small cohort of 42

patients of Asian origin (27).

No systematic analysis of HMGA2 expression in a larger cohort stratified by HNSCC subsites and exposure to risk factors has been done so far. Therefore the aim of this study was to investigate whether *HMGA2* expression in tumor tissue predicts recurrence-free, overall and tumor-specific survival in a larger group of German HNSCC patients. An understanding of these mechanisms may offer new strategies for cancer treatment.

Materials and Methods

Study design and population

As part of the European multicenter hospital-based case-control study ARCAGE (28, 29) a total of 287 cases were recruited in Germany from 2002 to 2005. The cancer sites comprised (ICD-10): oral cavity (C01-C06), tonsils (C09), pharynx (C10-C13) and larynx (C32). We had no patients with nasopharyngeal carcinomas (C11) in our study population. The patients were interviewed face-to-face about their socio-demographic characteristics, medical history and life-style factors including smoking and drinking history. Between 2011 and 2012 a mortality and morbidity follow-up was conducted. In the follow-up period (range: 6-10 years), all hospitalizations/new diagnoses (recurrence, secondary tumor) and ambulant examinations were recorded (pathological/ histological records). The primary endpoint was time from enrollment in the ARCAGE study to all-cause death, occurrence of metastasis or relapse (tumor progression) or last follow-up. Events were determined by local health departments and medical practitioners, death certificates were confirmed through the respective public health department or the Bremen mortality index (BreMI). The BreMI is an electronic database providing all information recorded on death certificates of Bremen citizens who have died since 01.01.1998.

BreMI is based on the law of the Bremen Cancer Registry and follows the example of the National Death Index (NDI) in the USA (30). One participant emigrated and was censored at date of emigration. All other information concerning the survival or causes of death are complete.

Only patients for whom formalin-fixed, paraffin-embedded (FFPE) tumor tissue was available were included in the molecular biological analysis for the genetic biomarker *HMGA2* (n=202). All these samples were obtained for diagnostic purposes during the ARCAGE study (baseline). Our sample size allows the detection of a hazard ratio (HR) of 1.79 for patients with a high expression of *HMGA2* in the tumor compared to those with low expression of *HMGA2* at a statistical significance level of 5% with a power of 80%. We performed a sensitivity analysis that only included HPV negative tumors (n=159) to check whether the observed associations were influenced by HPV status. Only one HPV positive case with a high HMGA2 expression level was detected. That is the reason we could not adjust for HPV status.

Outcome measures

The vital status and the cause of death were determined for each participant. The vital status and current place of residence were determined by contacting the respective registration office. Death certificates were collected from the public health department or the local mortality index (30). Data on the clinical course of the tumor disease were collected in cooperation with the chief physicians of ear, nose and throat (ENT) hospital departments or the patients' general practitioners. Data from paper files, histological and pathological findings and discharge letters, electronic hospital files, mortality index records, general practitioners' and public health department questionnaires were transferred to a computer-based standardized form by qualified study nurses (28).

The study endpoints were overall survival, progression-free survival, and tumor-specific survival. Overall survival time was defined as time the patient stayed alive from the date of HNSCC diagnosis to the end of the study or to death. Progression-free survival time was defined as time alive from HNSCC diagnosis without any local recurrence or newly diagnosed metastasis. Tumor-specific survival was defined as time alive from HNSCC diagnosis to the end of the study or to death which was not related to the primary tumor.

If no further visits to health departments or medical practitioners were documented (loss to follow-up), the last day of a documented visit to a health department/medical practitioner was the time point of censoring. TNM classification was performed according to the 2010 International Union against Cancer (UICC) guidelines (31).

HMGA2 expression

Total RNA was isolated from the tissue using an RNeasy FFPE kit (Qiagen, Hilden, Germany) in a QIACube (Qiagen, Hilden, Germany) and measured in triplicates, as previously described (16, 32). For calibration and *HMGA2* quantification, *HPRT* was used as control gene, as has been recommended for head and neck tumors (33). The short length of both the amplicon of interest (*HMGA2* 61 Bp) and the control (*HPRT1* 81 Bp) are advantageous when FFPE tissue is examined.

Human papilloma virus detection

HPV-DNA was detected using the primer system GP5+/6+ developed by de Roda Husman et. al. (34). HPV type specific polymerase chain reaction (PCR) and primers were used as previously described (32) (HPV16 (ATATAAGGGGTCGGTGGACCG, GCAATGTAGGTGTATCTCCATGC) and HPV18 (AAGGATGCTGCACCGGCTGAA, CACGCACACGCTTGGCAGGTTT). The PCR was performed with a PCR Core

Kit^{PLUS} (Roche) according to the manufacturer's instructions. Prior to the PCR the reaction mixture was incubated with 0.5 U uracil-DNA-glycosylase (UNG) for 5 min at 20°C followed by thermal inactivation of UNG for 2 minutes at 95°C. The PCR was performed with an initial denaturation of 30 s at 98°C, 35 repeats of 20 s denaturation at 98°C, 15 s annealing at 55°C and 20 s elongation at 72°C, followed by a final elongation step for 3 minutes at 72°C. Unfortunately p16 expression was not assessed.

Lifestyle variables

Smoking behavior was quantified as pack-years (number of packs consumed (20 cigarettes) per day multiplied by the number of years the person had smoked). Alcohol drinking behavior was quantified as drink-years (drinks per day multiplied by the number of years of alcohol consumption). The definition of one alcoholic drink equivalent was 18ml of pure alcohol, corresponding to 330ml of beer, 150ml of wine or 36ml of hard liquor (35).

Statistical methods

We used the 75th quantile of each tumor site to dichotomize the logarithmized and normalized expression values of *HMGA2* into a low and high expression group, because on the one hand, we wanted to identify high risk patients and thus choose a high cut-off and on the other hand we wanted to obtain stable estimates. Because of the low number of cases we could not use the 90th percentile. That is why we decided to use the upper quartile as a conventional cut-off for HMGA2 expression a priori as has been done in previous biomarker studies (36, 37). A sensitivity analysis with other cut-offs (upper tertile, upper quintile) is now presented in supplementary Table 1.Two sided χ^2-tests and t-tests were performed to compare the frequency of

clinicopathological parameters between patients with low and high expression levels of *HMGA2*. Differences in survival between groups were assessed by the Kaplan-Meier method and the log-rank test. Cox proportional hazards regression models were applied to investigate the effects of low vs. high *HMGA* expression levels for the defined endpoints.

Stratification by tumor site (“oral cavity, “tonsils”, “pharynx”, “larynx”) was used to detect possible differences between HNSCC of distinct subsites, since tumor sites are known to differ regarding biological behavior and HPV status. HPV infection status of the tumor and tumor stage are established predictors for HNSCC patient survival. Therefore, we excluded HPV positive tumors (HPV 16 and 18) and stratified by tumor stage (“1/2”, “3/4”) in a sensitivity analysis. To determine whether the prognostic levels of HMGA2 expression are independent of clinicopathological parameters, multivariate Cox regression models were adjusted for sex, age (continuous), treatment (surgery, radio- and chemotherapy), tobacco (pack-years; continuous) and alcohol consumption (drink-years; continuous).

We tested the proportional hazard (PH) assumptions by including time-by-covariate interactions in the multivariate Cox proportional hazard models by assuming a significance value of $\alpha = 0.05$. The results indicated that tumor stage violated the PH assumption in all models. We also detected time-varying effects for the variable surgery in the subgroup analysis on tumor-specific survival. We therefore stratified all Cox models by tumor stage and where needed also by treatment (surgery). Stratification, however, led to a loss of power. This resulted in unstable effect estimates regarding HMGA2 and overall survival in the subgroup “oral cavity”. In general, p-values <0.05 were considered statistically significant and two-sided 95% confidence intervals (CI) were calculated. All statistical analyses were performed with SAS software (v9.3; SAS Institute Inc., Cary, NC, USA).

Results

Patient characteristics and lifestyle factors

Table 1 shows the distribution of patient characteristics by tumor site. Of a total of 202 HNSCC patients, 62 had tumors of the oral cavity (30.7%), tonsils (12.9%), pharynx (27.7%), or larynx (28.7%).

Fifty-five percent of HNSCC patients for whom complete information was available for TNM staging according to UICC (n=153) were classified as TNM stage 3 and 4. This does not reflect the high number of locally limited tumors, because many cases lacked values referring to distant metastasis (M-stage). A total of 24 HPV-associated HNSCC (11.9%) were detected. The prevalence of HPV infection differed substantially with a maximum in tonsillar tumors (42.3%) and the minimum in laryngeal HNSCC (3.5%). Tobacco and alcohol consumption were highest among pharyngeal cancer patients (45.96 pack years/98.96 drink years).

Association of expression levels of HMGA2 with the clinicopathological characteristics of HNSCC

HMGA2 expression was dichotomized into low and high, based on the 75th quantile of logarithmized expression values. Table 2 shows the relationship to expression status, clinicopathological features and patient characteristics. HMGA2 expression levels did not differ with regard to age, sex, tumor site, tumor stage, and smoking habits (Table 2). A statistically significant association between HMGA2 expression and negative HPV status was observed ($p<0.01$). HMGA2 expression levels also differed by cumulated amount of alcohol drinking and N status, but these differences were not statistically significant.

HMGA2 expression and patient survival

Shorter overall survival was observed in patients with high *HMGA2* expressing HNSCC, regardless of the subsite (p=0.02; Figure 1). The median survival times were 63.5 months in the low expression group and 21.4 months in the high expression group.

The stratification by T stages (Figure 2) showed that patients with low *HMGA2* expression and locally limited tumors had the best prognosis (p<0.001). In the latter, the survival time was 82.8 months compared to 55.6 months in the group with locally limited tumors and high *HMGA2* expression. The shortest overall survival was seen in patients with advanced tumors and high HMGA2 expression (13 months).

Subsite-specific Kaplan-Meier curves revealed a significant survival benefit for patients with HNSCC of the oral cavity and the larynx compared to patients with tumors in other subsites (Figure 3).

Table 3 summarizes the results of the Cox proportional hazard regression analysis. The 1.5-fold risk increase in the total of 202 patients with HMGA2 positive HNSCC was observed without adjustment for confounders. The analysis stratified by tumor site showed that the associations varied between localizations, corroborating the Kaplan-Meier results (Figure 3). In the Cox regression model stratified by tumor stage and fully adjusted for age, sex, tumor site, treatment, tobacco and alcohol consumption (Table 3), an up to 4-fold increased risk for the adverse outcome (endpoint overall survival) was seen in patients with high *HMGA2* expressing HNSCC of the larynx (HR=4.00; CI 95%:1.18-13.62). The risk for reduced overall survival was also increased for tumors of the oral cavity (HR=1.99; CI 95%: 0.92-4.29), but not statistically significant. The sensitivity analysis with other cut-offs (upper tertile, upper

quintile) for high and low HMGA2 expression levels is presented in supplementary Table 1. The resulting effect estimates varied but did not change their direction.

Patients with HNSCC of the oral cavity and the larynx showed a risk of earlier death due to the tumor (tumor-specific survival), in accordance with the overall survival results (Table 3). Recurrence-free survival was also negatively associated with HMGA2 expression status (progression-free survival), although the association did not reach statistical significance. On the contrary, in patients with pharyngeal HNSCC an inverse association between HMGA2 expression and the risk of earlier death was detected (HR= 0.39; CI 95%: 0.17-0.89). A sensitivity analysis excluding HPV-positive tumors confirmed the increased risk of earlier death in patients with high *HMGA2* expression. Effect estimates show only minor changes after removal of HPV+ participants (suppl. Table 2).

Discussion

In this follow-up study of 202 incident HNSCC cases, the stem cell factor *HMGA2*, a p53 modulating candidate molecular biomarker of tumor aggressiveness, was negatively associated with survival of patients with HNSCC of the larynx.

For many different cancer types, previous studies demonstrated that higher HMGA2 expression, possibly through p53 mutagenesis effects, is associated with a more aggressive tumor type, e.g. gastric cancer (38, 39), colon cancer (40), pancreatic cancer (25), and carcinomas of the oral cavity (27, 41). A study by Miyazawa and co-authors (27) showed that HMGA2 expression was associated with poor survival in 42 patients with HNSCC of the oral cavity. In that study the 5-year survival rate was 100% among patients who did not show high expression of HMGA2.

Chang and co-workers (42) suggested that *HMGA2* could be a potential prognostic tissue biomarker for HNSCC of the oral cavity. Their analysis indicated that *HMGA2* expression was an independent predictor of overall survival, disease-specific survival, and disease-free survival (42). The multivariate analysis was adjusted for age, sex, overall stage, and perineural invasion.

In our study of 202 HNSCC patients involving four different subsites, we confirmed that HMGA2 predicts overall survival in HNSCC of the larynx independent of covariates like age, sex, treatment, alcohol and tobacco consumption. Interestingly, HMGA2 was predominantly expressed in non-HPV associated HNSCC, which supports the notion that HPV- and non-HPV-associated tumorigenesis differ dramatically from each other. HNSCC is a heterogeneous group of tumors with different clinical characteristics, which may therefore have distinct pathways of tumorigenesis. Biological heterogeneity in HNSCC has been suggested by molecular analyses that have uncovered distinct classes of HNSCC with unique mRNA expression patterns or differences in DNA copy number alterations. It may be speculated that topical noxae like alcohol and tobacco smoke have a more direct effect on the mucosal tissue of the oral cavity. The elevated cell turnover caused by topical noxae might explain why a stem cell gene like HMGA2 is upregulated. In fact, HMGA2 expression could identify a subgroup of HNSCC, characterized by a stem cell phenotype of tumor cells with higher cell turnover.

As demonstrated by T stage specific Kaplan-Meier curves with and without HMGA2 expression, the HMGA2 phenotype can already be diagnosed in early tumor stages, indicating a strong impact on post-operative patient management. HMGA2 is independently associated with tumor-specific survival in HNSCC of the oral cavity and the larynx. The 10-fold increased risk of tumor-specific death in patients with

laryngeal HNSCC has to be interpreted cautiously considering the wide confidence interval and due to the small number of cases in this group. As has been reported, patients with HPV-positive HNSCC (which are 42% of the tonsillar HNSCC) are known to have a longer survival per se (43), which could explain why HMGA2 is of limited prognostic value in patients with tonsillar HNSCC. Furthermore the small sample size of tonsillar HNSCC resulted in unstable effect estimates in the multivariate analysis. The reason for the inverse, i.e. protective, effect of HMGA2 expression resulting in longer recurrence-free survival in patients with pharyngeal HNSCC remains unclear, although we cannot rule out a random finding. Otherwise, the result implies that HNSCC of the pharynx differ in terms of tumorigenesis pathways from other HNSCC subsites, regardless of their epithelial origin. Finally, it is worthy of note that the classical clinicopathological factors (TNM stage, local extent of the tumor disease, lymph node involvement and surgical resection of the tumor) in patients with HNSCC are still the most useful independent prognostic factors. A limitation of this study is that the stratified analysis yielded small sample sizes and data on TNM stage were missing in 25% of all cases. Only FFPE tissues were available. RNA isolated from archived material can be degraded and formalin fixation can cause chemical modification in the RNA, like cross-linking to proteins and addition of monomethylol groups. Therefore, qRT-RCR performed on FFPE tissue consistently obtains lower CT values than for fresh frozen tissues with the same input RNA (44). Nevertheless, experiments comparing qRT-PCR results with fresh frozen tissue and FFPE tissue suggest that mRNA expression levels derived from FFPE tissue reflect the actual expression levels in the original tissue samples, regardless of the variable effects of fixation (45, 46). In addition, we could not investigate the effect of tumor stage on survival in multivariate models since the proportional hazard assumptions were not fulfilled for tumor stage.

In contrast to previous studies, we did not use immunohistochemistry to detect *HMGA2* (over-)expression, because sensitivity and specificity of HMGA2 immunohistochemistry have been an issue e.g. in mesenchymal tumors (47). Since our study employed quantifications of *HMGA2* expression relative to the non-regulated house-keeping gene HPRT, cut-off points were defined by data-driven quartiles to discriminate between low and high *HMGA2* expression levels.The major limitation of our study is the small sample size for some variables leading to insufficient power to detect smaller associations. Despite small the sample size, the findings are relatively consistent across clinicopathological parameters. Further, the importance of our findings in providing new insight into potential preventive and treatment strategies for HPV-negative HNSCCs outweighs the limitations.

Our findings indicate that HMGA2 plays a major role in non-HPV associated HNSCC, especially of the the larynx. The extensive adjustments for confounders in the Cox regression model we used support the validity of our findings. The role of HMGA2 in stem cell renewal facilitates sustained proliferative signaling in cancer cells and may identify a stem cell phenotype within the heterogeneous group of HNSCC. The previously reported anti-tumorigenic effects achieved through targeted HMGA2 inhibition (18), could offer new strategies for cancer treatment in HMGA2 expressing early stage HNSCC.

Acknowledgement

The authors would like to thank the patients and their families for their participation and gratefully acknowledge the excellent technical assistance of Beate Schütte, Anja Bergmann and Marina Resnikov (Leibniz Institute for Prevention Research and

Epidemiology, BIPS) and Sabrina Dorschner (Center for Human Genetics). We are also grateful to our clinical colleagues in hospitals and primary care who supported this study. We also thank the team of the Bremen Mortality Index (BreMI) and Carola Lehmann for her technical support.

Ethics

We obtained written informed consent from all patients. The study was approved by the ethics committee of the medical association of Bremen.

Disclaimer

The funders played no role in the design or conduction of the study, collection, management, analysis or interpretation of the data, preparation, review, or approval of the manuscript, or the decision to submit the manuscript for publication.

References

1. Rothenberg SM, Ellisen LW. The molecular pathogenesis of head and neck squamous cell carcinoma. The Journal of clinical investigation. 2012;122:1951-7.
2. Ferlay J, Shin HR, Bray F, Forman D, Mathers C, Parkin DM. Estimates of worldwide burden of cancer in 2008: GLOBOCAN 2008. International journal of cancer Journal international du cancer. 2010;127:2893-917.
3. Torre LA, Bray F, Siegel RL, Ferlay J, Lortet-Tieulent J, Jemal A. Global cancer statistics, 2012. CA: a cancer journal for clinicians. 2015;65:87-108.
4. Becher H, Ramroth H, Ahrens W, Risch A, Schmezer P, Dietz A. Occupation, exposure to polycyclic aromatic hydrocarbons and laryngeal cancer risk. International journal of cancer Journal international du cancer. 2005;116:451-7.
5. D'Souza G, Kreimer AR, Viscidi R, Pawlita M, Fakhry C, Koch WM, et al. Case-control study of human papillomavirus and oropharyngeal cancer. The New England journal of medicine. 2007;356:1944-56.
6. Ramroth H, Dietz A, Ahrens W, Becher H. Occupational wood dust exposure and the risk of laryngeal cancer: a population based case-control study in Germany. American journal of industrial medicine. 2008;51:648-55.
7. Denaro N, Lo Nigro C, Natoli G, Russi EG, Adamo V, Merlano MC. The Role of p53 and MDM2 in Head and Neck Cancer. ISRN Otolaryngol. 2011;2011:931813.
8. Thomas GR, Nadiminti H, Regalado J. Molecular predictors of clinical outcome in patients with head and neck squamous cell carcinoma. International journal of experimental pathology. 2005;86:347-63.
9. Grandis JR, Tweardy DJ. TGF-alpha and EGFR in head and neck cancer. Journal of cellular biochemistry Supplement. 1993;17F:188-91.
10. Goke F, Franzen A, Hinz TK, Marek LA, Yoon P, Sharma R, et al. FGFR1 Expression Levels Predict BGJ398 Sensitivity of FGFR1-Dependent Head and Neck Squamous Cell Cancers. Clinical cancer research : an official journal of the American Association for Cancer Research. 2015;21:4356-64.
11. Munger K, Scheffner M, Huibregtse JM, Howley PM. Interactions of HPV E6 and E7 oncoproteins with tumour suppressor gene products. Cancer surveys. 1992;12:197-217.
12. Liu F, Chau KY, Arlotta P, Ono SJ. The HMG I proteins: dynamic roles in gene activation, development, and tumorigenesis. Immunologic research. 2001;24:13-29.
13. Melillo RM, Pierantoni GM, Scala S, Battista S, Fedele M, Stella A, et al. Critical role of the HMGI(Y) proteins in adipocytic cell growth and differentiation. Molecular and cellular biology. 2001;21:2485-95.
14. Schwanbeck R, Manfioletti G, Wisniewski JR. Architecture of high mobility group protein I-C.DNA complex and its perturbation upon phosphorylation by Cdc2 kinase. The Journal of biological chemistry. 2000;275:1793-801.
15. Lanahan A, Williams JB, Sanders LK, Nathans D. Growth factor-induced delayed early response genes. Molecular and cellular biology. 1992;12:3919-29.
16. Markowski DN, Helmke BM, Belge G, Nimzyk R, Bartnitzke S, Deichert U, et al. HMGA2 and p14Arf: major roles in cellular senescence of fibroids and therapeutic implications. Anticancer research. 2011;31:753-61.
17. Wei JJ, Wu J, Luan C, Yeldandi A, Lee P, Keh P, et al. HMGA2: a potential biomarker complement to P53 for detection of early-stage high-grade papillary serous carcinoma in fallopian tubes. The American journal of surgical pathology. 2010;34:18-26.
18. Ji Q, Hao X, Meng Y, Zhang M, Desano J, Fan D, et al. Restoration of tumor suppressor miR-34 inhibits human p53-mutant gastric cancer tumorspheres. BMC cancer. 2008;8:266.
19. Nishino J, Kim I, Chada K, Morrison SJ. Hmga2 promotes neural stem cell self-renewal in young but not old mice by reducing p16Ink4a and p19Arf Expression. Cell. 2008;135:227-39.
20. Chang ZG, Yang LY, Wang W, Peng JX, Huang GW, Tao YM, et al. Determination of high mobility group A1 (HMGA1) expression in hepatocellular carcinoma: a potential

prognostic marker. Digestive diseases and sciences. 2005;50:1764-70.
21. Flohr AM, Rogalla P, Bonk U, Puettmann B, Buerger H, Gohla G, et al. High mobility group protein HMGA1 expression in breast cancer reveals a positive correlation with tumour grade. Histology and histopathology. 2003;18:999-1004.
22. Franco R, Esposito F, Fedele M, Liguori G, Pierantoni GM, Botti G, et al. Detection of high-mobility group proteins A1 and A2 represents a valid diagnostic marker in post-pubertal testicular germ cell tumours. The Journal of pathology. 2008;214:58-64.
23. Sarhadi VK, Wikman H, Salmenkivi K, Kuosma E, Sioris T, Salo J, et al. Increased expression of high mobility group A proteins in lung cancer. The Journal of pathology. 2006;209:206-12.
24. Fusco A, Fedele M. Roles of HMGA proteins in cancer. Nature reviews Cancer. 2007;7:899-910.
25. Piscuoglio S, Zlobec I, Pallante P, Sepe R, Esposito F, Zimmermann A, et al. HMGA1 and HMGA2 protein expression correlates with advanced tumour grade and lymph node metastasis in pancreatic adenocarcinoma. Histopathology. 2012;60:397-404.
26. Yang GL, Zhang LH, Bo JJ, Hou KL, Cai X, Chen YY, et al. Overexpression of HMGA2 in bladder cancer and its association with clinicopathologic features and prognosis HMGA2 as a prognostic marker of bladder cancer. European journal of surgical oncology : the journal of the European Society of Surgical Oncology and the British Association of Surgical Oncology. 2011;37:265-71.
27. Miyazawa J, Mitoro A, Kawashiri S, Chada KK, Imai K. Expression of mesenchyme-specific gene HMGA2 in squamous cell carcinomas of the oral cavity. Cancer research. 2004;64:2024-9.
28. 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. European journal of cancer prevention : the official journal of the European Cancer Prevention Organisation. 2009;18:76-84.
29. Ahrens W, Pohlabeln H, Foraita R, Nelis M, Lagiou P, Lagiou A, et al. Oral health, dental care and mouthwash associated with upper aerodigestive tract cancer risk in Europe: The ARCAGE study. Oral Oncol. 2014;50:616-25.
30. Giersiepen K, Brunings-Kuppe C, Lehmann C. [The Bremen mortality index]. Bundesgesundheitsblatt, Gesundheitsforschung, Gesundheitsschutz. 2004;47:451-6.
31. Sobin LH GM, Wittekind C. TNM Classification of Malignant Tumours, 7th Edition: Wiley-Blackwell, Chichester, UK; 2009.
32. Friemel J, Foraita R, Gunther K, Heibeck M, Gunther F, Pflueger M, et al. Pretreatment oral hygiene habits and survival of head and neck squamous cell carcinoma (HNSCC) patients. BMC oral health. 2016;16:33.
33. Lallemant B, Evrard A, Combescure C, Chapuis H, Chambon G, Raynal C, et al. Reference gene selection for head and neck squamous cell carcinoma gene expression studies. BMC Mol Biol. 2009;10:78.
34. de Roda Husman AM, Walboomers JM, van den Brule AJ, Meijer CJ, Snijders PJ. The use of general primers GP5 and GP6 elongated at their 3' ends with adjacent highly conserved sequences improves human papillomavirus detection by PCR. The Journal of general virology. 1995;76 (Pt 4):1057-62.
35. Macfarlane TV, Macfarlane GJ, Oliver RJ, Benhamou S, Bouchardy C, Ahrens W, et al. The aetiology of upper aerodigestive tract cancers among young adults in Europe: the ARCAGE study. Cancer causes & control : CCC. 2010;21:2213-21.
36. Akcakanat A, Zhang L, Tsavachidis S, Meric-Bernstam F. The rapamycin-regulated gene expression signature determines prognosis for breast cancer. Mol Cancer. 2009;8:75.
37. Noman AS, Uddin M, Rahman MZ, Nayeem MJ, Alam SS, Khatun Z, et al. Overexpression of sonic hedgehog in the triple negative breast cancer: clinicopathological characteristics of high burden breast cancer patients from Bangladesh. Scientific reports. 2016;6:18830.
38. Kong D, Su G, Zha L, Zhang H, Xiang J, Xu W, et al. Coexpression of HMGA2 and Oct4 predicts an unfavorable prognosis in human gastric cancer. Medical oncology.

2014;31:130.
39. Lee J, Ha S, Jung CK, Lee HH. High-mobility-group A2 overexpression provokes a poor prognosis of gastric cancer through the epithelial-mesenchymal transition. International journal of oncology. 2015;46:2431-8.
40. Wang X, Liu X, Li AY, Chen L, Lai L, Lin HH, et al. Overexpression of HMGA2 promotes metastasis and impacts survival of colorectal cancers. Clinical cancer research : an official journal of the American Association for Cancer Research. 2011;17:2570-80.
41. Pallante P, Sepe R, Puca F, Fusco A. High mobility group a proteins as tumor markers. Frontiers in medicine. 2015;2:15.
42. Chang KP, Lin SJ, Liu SC, Yi JS, Chien KY, Chi LM, et al. Low-molecular-mass secretome profiling identifies HMGA2 and MIF as prognostic biomarkers for oral cavity squamous cell carcinoma. Scientific reports. 2015;5:11689.
43. O'Rorke MA, Ellison MV, Murray LJ, Moran M, James J, Anderson LA. Human papillomavirus related head and neck cancer survival: a systematic review and meta-analysis. Oral oncology. 2012;48:1191-201.
44. Farragher SM, Tanney A, Kennedy RD, Paul Harkin D. RNA expression analysis from formalin fixed paraffin embedded tissues. Histochem Cell Biol. 2008;130:435-45.
45. Godfrey TE, Kim SH, Chavira M, Ruff DW, Warren RS, Gray JW, et al. Quantitative mRNA expression analysis from formalin-fixed, paraffin-embedded tissues using 5' nuclease quantitative reverse transcription-polymerase chain reaction. J Mol Diagn. 2000;2:84-91.
46. Kalmar A, Wichmann B, Galamb O, Spisak S, Toth K, Leiszter K, et al. Gene expression analysis of normal and colorectal cancer tissue samples from fresh frozen and matched formalin-fixed, paraffin-embedded (FFPE) specimens after manual and automated RNA isolation. Methods. 2013;59:S16-9.
47. Dreux N, Marty M, Chibon F, Velasco V, Hostein I, Ranchere-Vince D, et al. Value and limitation of immunohistochemical expression of HMGA2 in mesenchymal tumors: about a series of 1052 cases. Mod Pathol. 2010;23:1657-66.

Table 1: Characteristics of incident HNSCC patients (n=202) by tumor site

Characteristic	All n=202 (100.00%)	C01-C06 (oral cavity) n=62 (30.69%)	C09 (Tonsils) n=26 (12.87%)	C10-C13 (Pharynx) n=56 (27.72%)	C32 (Larynx) n=58 (28.71%)
	Mean ±SD	Mean ±SD	Mean ±SD	Mean ±SD	Mean ±SD
Age (years)	58.17±8.66	56.06 ±9.09	58.08 ±10.04	57.05 ±7.53	61.55 ±7.73
Pack-years (tobacco)	45.01 ±26.89	39.73 ±22.80	42.02 ±38.21	45.96 ±29.99	44.02 ±24.81
Drink-years (alcohol)	95.65 ±114.65	85.24 ±112.70	95.88 ±144.16	98.96 ±112.54	72.74 ±99.06
Survival (months)	54.23 ±37.17	53.09 ±37.84	50.13 ±38.46	44.90 ±37.70	66.31 ±32.82
	n (%)	n (%)	n (%)	n (%)	n (%)
Sex		xc			
Male	169 (83.66)	51 (82.26)	23 (88.46)	43 (76.79)	52 (89.66)
Female	33 (16.34)	11 (17.74)	3 (11.54)	13 (23.21)	6 (10.34)
TNM stage[1]					
1/2	41 (20.30)	12 (19.35)	4 (15.39)	5 (8.39)	20 (34.48)
3/4	112 (55.45)	30 (48.39)	19 (73.08)	43 (76.78)	20 (34.48)
Missing	49 (24.26)	20 (32.26)	3 (11.54)	8 (14.29)	18 (31.03)
T stage					
T1	47 (23.27)	21 (33.87)	6 (23.08)	5 (8.93)	15 (25.86)
T2	74 (36.63)	26 (41.94)	10 (38.46)	14 (25.00)	24 (41.38)
T3	28 (13.86)	8 (12.90)	4 (15.38)	13 (23.21)	3 (5.17)
T4	38 (18.81)	6 (9.68)	4 (15.38)	20 (35.71)	8 (13.79)
Missing	15 (7.43)	1 (1.61)	2 (7.69)	4 (7.14)	8 (13.79)
N status					
N-	65 (32.18)	23 (37.10)	5 (19.23)	7 (12.50)	30 (51.72)
N+	115 (56.93)	36 (58.06)	18 (69.23)	43 (76.79)	18 (31.03)
Missing	22 (10.89)	3 (4.84)	3 (11.54)	6 (10.71)	10 (17.24)
HPV status					
HPV-	159 (78.71)	54 (87.10)	15 (57.69)	43 (76.79)	47 (81.03)
HPV+	24 (11.88)	5 (8.06)	11 (42.31)	6 (10.71)	2 (3.45)
Missing	19 (9.41)	3 (4.84)	0 (0.00)	7 (12.50)	9 (15.52)
Surgery					
Yes	166 (82.18)	52 (83.87)	22 (84.62)	40 (71.43)	52 (89.66)
No/ unknown	36 (17.82)	10 (16.13)	4 (15.38)	16 (28.57)	6 (10.34)
Chemotherapy					
Yes	84 (41.58)	20 (32.26)	13 (50.0)	38 (67.86)	13 (22.41)
No/ unknown	118 (58.42)	42 (67.74)	13 (50.0)	18 (32.14)	45 (77.59)
Radiotherapy					
Yes	139 (68.81)	39 (62.90)	21 (80.77)	47 (83.93)	32 (55.17)
No/ unknown	63 (31.19)	23 (37.10)	5 (19.23)	9 (16.07)	26 (44.83)

[1] according to the International Union against Cancer (UICC)

Table 2: Characteristics of incident HNSCC patients by HMGA expression

	All tumor sites		
Characteristic	HMGA low n=148 (73.27%)	HMGA high n=54 (26.73%)	
	Mean ±SD	Mean ±SD	p-value
Age (years)	57.93 ±8.79	58.83 ±8.36	0.51
Pack-years (tobacco)	42.10 ±29.26	45.33 ±22.90	0.47
Drink-years (alcohol)	78.01 ±102.9	111.00 ±135.3	0.07
Survival (months)	58.25 ±36.36	43.22 ±37.48	**0.01**
	n (%)	n (%)	p-value
Sex			
Male	123 (83.11)	46 (85.19)	
Female	25 (16.89)	8 (14.81)	0.72
Tumor site			
C01-06 (oral cavity)	44 (30.41)	18 (31.48)	
C09 (tonsils)	19 (12.84)	7 (12.96)	
C10-C13 (pharynx)	41 (27.70)	15 (27.78)	
C32 (larynx)	46 (29.05)	12 (22.22)	0.99
TNM stage[1]			
1/2	35 (23.65)	6 (11.11)	
3/4	78 (52.70)	34 (62.96)	0.14
Missing	35 (23.65)	14 (25.93)	
T stage			
T1	37 (25.00)	10 (18.52)	
T2	57 (38.51)	17 (31.48)	
T3	18 (12.16)	10 (18.52)	
T4	26 (17.57)	12 (22.22)	0.52
Missing	10 (6.76)	5 (9.26)	
N status			
N-	54 (41.54)	11 (23.40)	
N+	61 (46.92)	29 (61.70)	0.09
Missing	15 (11.54)	7 (14.89)	
HPV status			
HPV-	114 (77.03)	45 (83.33)	
HPV+	23 (15.54)	1 (1.85)	**<0.01**
Missing	11 (7.43)	8 (14.88)	

[1] according to the International Union against Cancer (UICC)
Bold numbers indicate statistically significant p-values

Table 3: Cox regression models for overall, tumor-specific and progression-free survival time

	Univariate analysis				Multivariate analysis [1]		
	Overall survival	Tumor-specific survival	Progression-free survival		Overall survival	Tumor-specific survival	Progression-free survival
Variable	Hazard ratio* (95% CI)	Hazard ratio (95% CI)	Hazard ratio (95% CI)		Hazard ratio (95% CI)	Hazard ratio (95% CI)	Hazard ratio (95% CI)
All tumors (n=202)					**All tumors (n=198)**		
HMGA2 expression							
Low	1	1	1		1	1	1
High	**1.58 (1.07-2.33)**	1.56 (0.93- 2.63)	1.40 (0.96-2.05)		1.18 (0.77-1.80)	1.21 (0.69-2.12)	1.00 (0.66-1.51)
Sex							
Male	1	1	1		1	1	1
Female	0.67 (0.40-1.14)	0.60 (0.28-1.24)	0.66 (0.40-1.11)		0.65 (0.36-1.15)	0.53 (0.24-1.17)	0.60 (0.35-1.06)
Surgery							
No/unknown	1	1	1		1	1	1
Yes	**0.33 (0.22-0.49)**	**0.34 (0.20-0.58)**	**0.36 (0.24-0.53)**		**0.40 (0.24-0.66)**	**0.47 (0.25-0.88)**	**0.41 (0.25-0.66)**
Tumor site							
C01-06 (Oral cavity)	1	1	1		1	1	1
C09 (Tonsils)	0.94 (0.47-1.85)	0.94 (0.41-2.11)	0.82 (0.46-1.45)		0.96 (0.52-1.76)	0.79 (0.35-1.83)	0.71 (0.39-1.29)
C10-C13 (Pharynx)	1.55 (0.96-2.50)	1.55 (0.87-2.74)	1.17 (0.76-1.80)		0.92 (0.56-1.49)	1.00 (0.53-1.87)	0.99 (0.63-1.57)
C32 (Larynx)	**0.43 (0.23-0.81)**	**0.43 (0.20-0.92)**	**0.50 (0.31-0.81)**		**0.54 (0.31-0.93)**	0.47 (0.21-1.05)	**0.42 (0.25-0.71)**
Oral cavity (n=62)					**Oral cavity (n=62)[2]**		
HMGA2 expression							
Low	1	1	1		1	1	1
High	**2.14 (1.10-4.18)**	**2.57 (1.05-6.34)**	1.46 (0.77-2.77)		1.99 (0.92-4.29)	**2.88 (1.06-7.84)**	1.56 (0.75-3.24)
Tonsils (n=26)					**Tonsils (n=25)[2,3]**		
HMGA2 expression							
Low	1	1	1		1	1	1
High	2.64 (0.92-7.55)	1.99 (0.47-8.46)	2.54 (0.89-7.21)		1.04 (0.11-9.83)	3.48 (0.17-73.24)	1.24 (0.13-11.86)
Larynx (n=58)					**Larynx (n=56)[2]**		
HMGA2 expression							
Low	1	1	1		1	1	1
High	**2.44 (1.07-5.54)**	**3.66 (1.06-12.69)**	2.20 (0.99-4.92)		**4.00 (1.18-13.62)**	**10.12 (1.34-76.60)**	2.20 (0.69-6.99)
Pharynx (n=56)					**Pharynx (n=55)[2]**		
HMGA2 expression							
Low	1	1	1		1	1	1
High	0.62 (0.29-1.36)	0.59 (0.22-1.55)	0.62 (0.29-1.30)		0.50 (0.22-1.15)	0.47 (0.16-1.37)	**0.39 (0.17-0.89)**

[1] Hazard ratios: Cox regression model adjusted for age, sex, tumor site, treatment, tobacco and alcohol consumption, stratified by tumor stage (TNM)
[2] Hazard ratios: Cox regression model adjusted for age, sex, treatment, tobacco and alcohol consumption, stratified by tumor stage (TNM)
[3] Hazard ratios for tumor-specific survival: Cox regression model adjusted for age, sex, treatment, tobacco and alcohol consumption, stratified by tumor stage (TNM) and surgery
*Hazard ratios describe the risk of death, tumor progression or tumor specific death within the observational period
Bold numbers indicate statistically significant log rank tests and HR

Figure Legends:

Figure 1: Kaplan-Meier survival curve: overall (A), tumor-specific (B) and progression-free (C) survival (all tumor sites combined)

Figure 2: Kaplan-Meier survival curve: overall survival (all tumor sites combined, stratified by HMGA 2 expression level and T stage)

Figure 3: Kaplan-Meier survival curve: overall survival (stratified by tumor site) – A: oral cavity, B: tonsils, C: pharynx, D: larynx

Figure 1

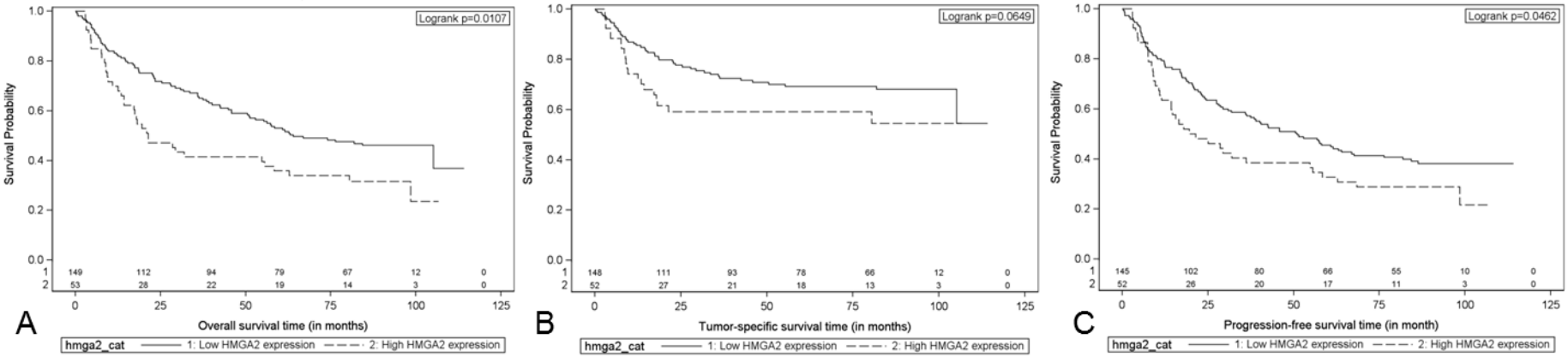

A
Logrank p=0.0107
Survival Probability
1.0
0.8
0.6
0.4
0.2
0.0
1 149 112 94 79 67 12 0
2 53 28 22 19 14 3 0
0 25 50 75 100 125
Overall survival time (in months)
hmga2_cat 1: Low HMGA2 expression 2: High HMGA2 expression
B
Logrank p=0.0649
Survival Probability
1.0
0.8
0.6
0.4
0.2
0.0
1 148 111 93 78 66 12 0
2 52 27 21 18 13 3 0
0 25 50 75 100 125
Tumor-specific survival time (in months)
hmga2_cat 1: Low HMGA2 expression 2: High HMGA2 expression
C
Logrank p=0.0462
Survival Probability
1.0
0.8
0.6
0.4
0.2
0.0
1 145 102 80 66 55 10 0
2 52 26 20 17 11 3 0
0 25 50 75 100 125
Progression-free survival time (in month)
hmga2_cat 1: Low HMGA2 expression 2: High HMGA2 expression

Figure 2

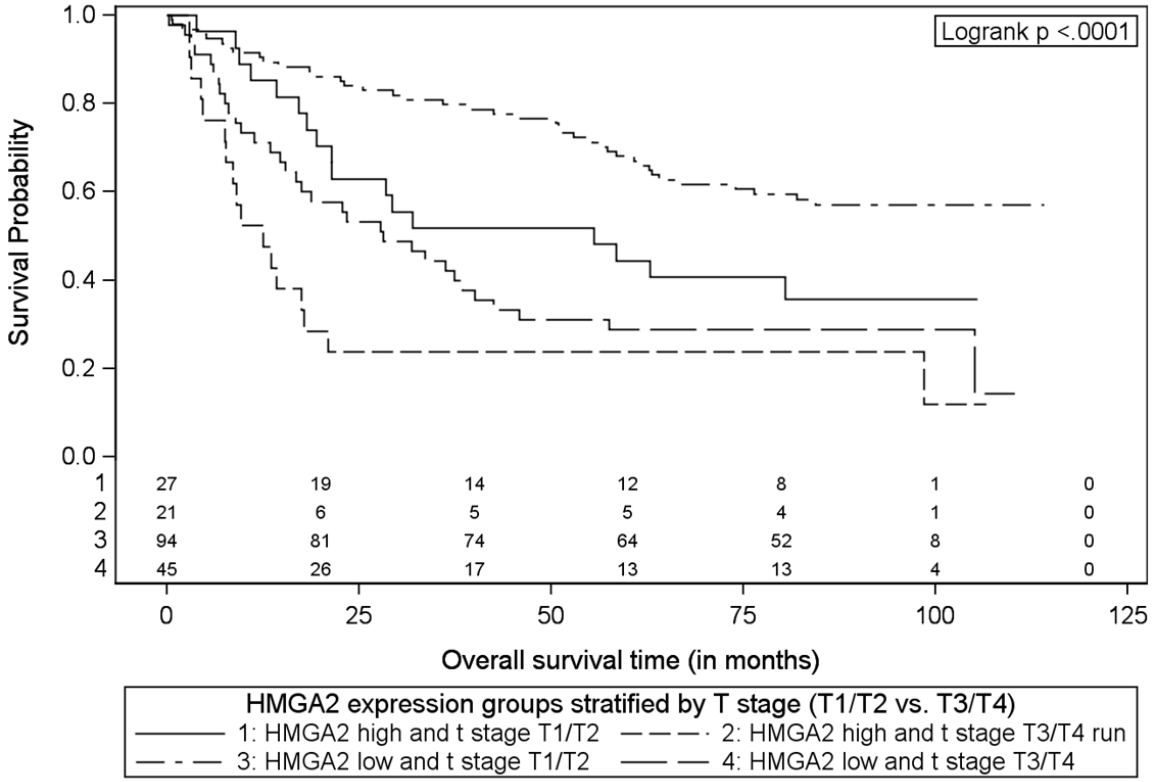
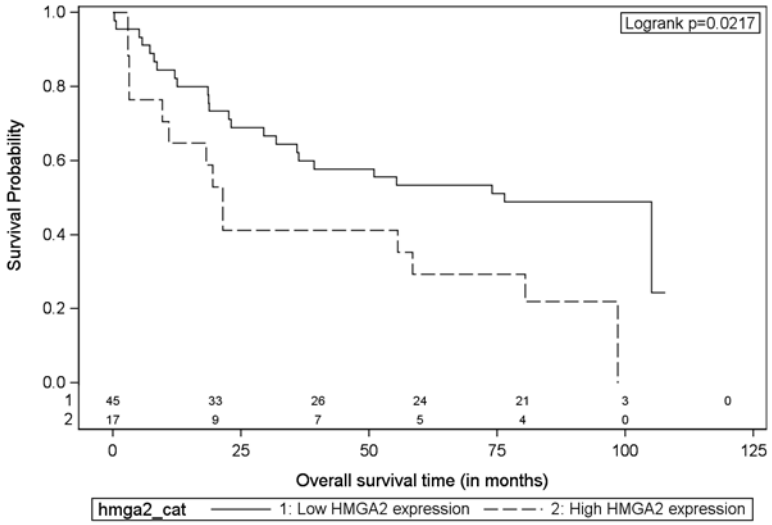
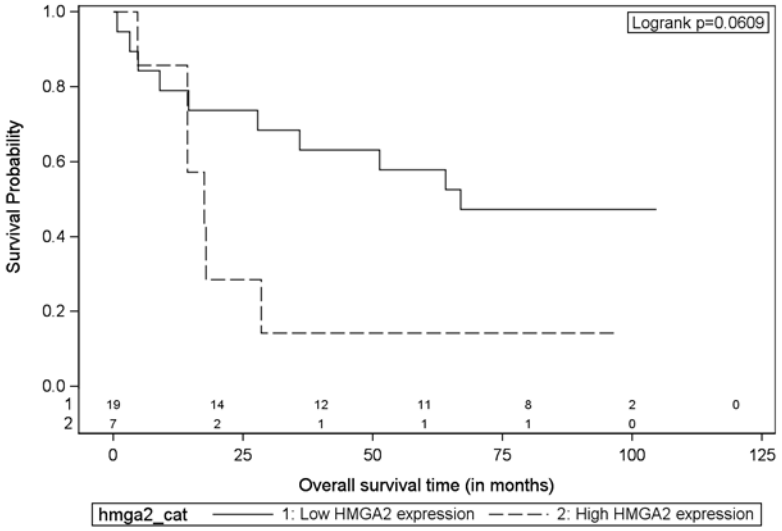

Logrank p <.0001
Survival Probability
1.0
0.8
0.6
0.4
0.2
0.0
1 27 19 14 12 8 1 0
2 21 6 5 5 4 1 0
3 94 81 74 64 52 8 0
4 45 26 17 13 13 4 0
0
25
50
75
100
125
Overall survival time (in months)
HMGA2 expression groups stratified by T stage (T1/T2 vs. T3/T4)
1: HMGA2 high and t stage T1/T2
2: HMGA2 high and t stage T3/T4 run
3: HMGA2 low and t stage T1/T2
4: HMGA2 low and t stage T3/T4

Figure 3

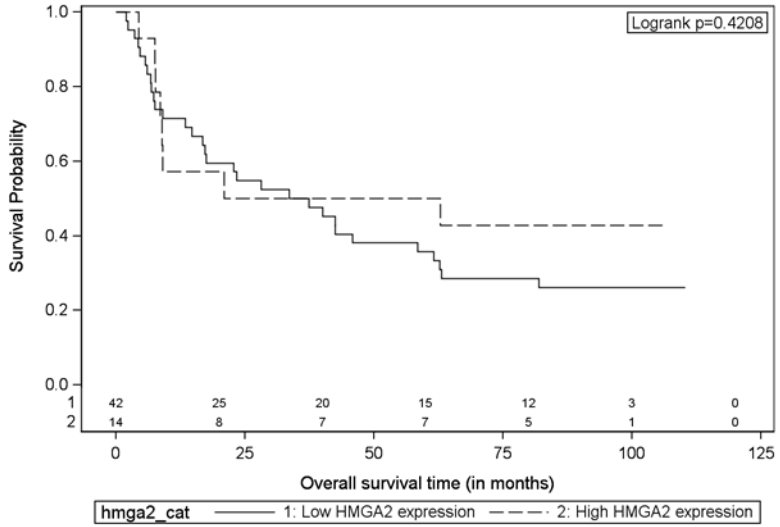
A: Oral Cavity



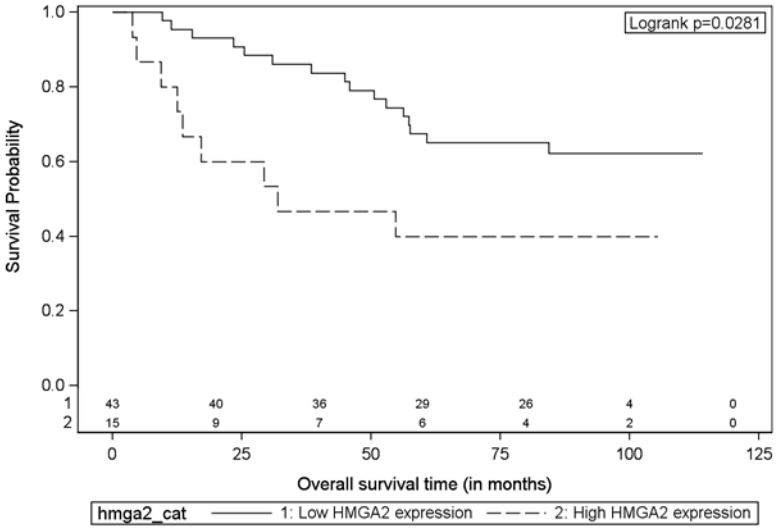
B: Tonsils



C: Pharynx



D: Larynx



Supplementary Table 1: Cox regression models for overall survival time. Comparison of different cut-offs for HMGA2 expression

		Multivariate analysis		
		Overall survival (HMGA2 cut-off: upper tertile)	Overall survival (HMGA2 cut-off: upper quartile)	Overall survival (HMGA2 cut-off: upper quintile)
Variable		Hazard ratio (95% CI)	Hazard ratio (95% CI)	Hazard ratio (95% CI)
All tumors (n=198)[1]				
HMGA2 expression				
Low		1	1	1
High		0.94 (0.62-1.41)	1.18 (0.77-1.80)	0.98 (0.61-1.56)
Oral cavity (n=62)[2]				
HMGA2 expression				
Low		1	1	1
High		1.63 (0.78-3.39)	1.99 (0.92-4.29)	1.54 (0.67-3.52)
Tonsils (n=25)[2,3]				
HMGA2 expression				
Low		1	1	1
High		0.84 (0.15-4.84)	1.04 (0.11-9.83)	6.65 (0.51-86.62)
Larynx (n=56)[2]				
HMGA2 expression				
Low		1	1	1
High		**5.20 (1.44-18.79)**	**4.00 (1.18-13.62)**	2.25 (0.65-7.78)
Pharynx (n=55)[2]				
HMGA2 expression				
Low		1	1	1
High		**0.33 (0.14-0.75)**	0.50 (0.22-1.15)	0.48 (0.20-1.16)

[1] Hazard ratios: Cox regression model adjusted for age, sex, tumor site, treatment, tobacco and alcohol consumption, stratified by tumor stage (TNM)
[2] Hazard ratios: Cox regression model adjusted for age, sex, treatment, tobacco and alcohol consumption, stratified by tumor stage (TNM)
[3] Hazard ratios for tumor-specific survival: Cox regression model adjusted for age, sex, treatment, tobacco and alcohol consumption, stratified by tumor stage (TNM) and surgery
*Hazard ratios describe the risk of death, tumor progression or tumor specific death within the observational period
Bold numbers indicate statistically significant log rank tests and HR

Supplementary Table 2: Cox regression models for overall, tumor-specific and progression-free survival time (only HPV-negative cases)

	Univariate analysis				Multivariate analysis [1]		
	Overall survival	Tumor-specific survival	Progression-free survival		Overall survival	Tumor-specific survival	Progression-free survival
Variable	Hazard ratio* (95% CI)	Hazard ratio (95% CI)	Hazard ratio (95% CI)		Hazard ratio (95% CI)	Hazard ratio (95% CI)	Hazard ratio (95% CI)
All tumors (n=159)					**All tumors (n=155)**		
HMGA2 expression							
Low	1	1	1		1	1	1
High	**1.71 (1.12-2.60)**	1.67 (0.97- 2.87)	1.44 (0.95-2.18)		1.20 (0.75-1.90)	1.11 (0.62-2.00)	1.00 (0.64-1.58)
Sex							
Male	1	1	1		1	1	1
Female	0.61 (0.33-1.12)	0.58 (0.26-1.29)	0.64 (0.36-1.15)		0.54 (0.27-1.05)	0.52 (0.22-1.23)	0.58 (0.31-1.09)
Surgery							
No/unknown	1	1	1		1	1	1
Yes	**0.32 (0.20-0.50)**	**0.27 (0.16-0.48)**	**0.38 (0.24-0.59)**		**0.41 (0.23-0.72)**	**0.40 (0.20-0.79)**	**0.48 (0.27-0.83)**
Tumor site							
C01-06 (Oral cavity)	1	1	1		1	1	1
C09 (Tonsils)	1.64 (0.85-3.15)	1.64 (0.69-3.90)	1.36 (0.71-2.59)		1.30 (0.61-2.76)	1.26 (0.48-3.32)	1.01 (0.49-2.12)
C10-C13 (Pharynx)	1.07 (0.65-1.74)	1.58 (0.86-2.90)	0.99 (0.62-1.58)		0.79 (0.46-1.36)	1.06 (0.53-2.09)	0.80 (0.47-1.34)
C32 (Larynx)	**0.54 (0.32-0.93)**	**0.43 (0.24-1.10)**	**0.47 (0.28-0.79)**		0.58 (0.32-1.04)	0.60 (0.26-1.39)	**0.44 (0.25-0.78)**
Oral cavity (n=54)					**Oral cavity (n=54)[2]**		
HMGA2 expression							
Low	1	1	1		1	1	1
High	**2.15 (1.08-4.29)**	2.52 (0.97-6.56)	1.55 (0.79-3.02)		2.01 (0.88-4.63)	2.73 (0.68-11.01)	1.68 (0.76-3.70)
Tonsils (n=15)					**Tonsils (n=14)[2,3]**		
HMGA2 expression							
Low	1	1	1		1	1	1
High	2.25 (0.67-7.60)	1.39 (0.30-6.42)	1.93 (0.61-6.17)		0.28 (0.03-2.82)	4.79 (0.16-143.56)	0.14 (0.01-1.92)
Larynx (n=47)					**Larynx (n=45)[2]**		
HMGA2 expression							
Low	1	1	1		1	1	1
High	**3.88 (1.58-9.51)**	**4.90 (1.40-17.09)**	**2.87 (1.16-7.10)**		**6.51 (1.54-27.63)**	**34.58 (2.39-500.32)**	2.96 (0.76-11.49)
Pharynx (n=43)					**Pharynx (n=42)[2]**		
HMGA2 expression							
Low	1	1	1		1	1	1
High	0.47 (0.19-1.17)	0.55 (0.20-1.47)	0.47 (0.20-1.10)		0.40 (0.15-1.09)	0.37 (0.12-1.21)	**0.35 (0.13-0.94)**

[1] Hazard ratios: Cox regression model adjusted for age, sex, tumor site, treatment, tobacco and alcohol consumption, stratified by tumor stage (TNM)
[2] Hazard ratios: Cox regression model adjusted for age, sex, treatment, tobacco and alcohol consumption, stratified by tumor stage (TNM)
[3] Hazard ratios for tumor-specific survival: Cox regression model adjusted for age, sex, treatment, tobacco and alcohol consumption, stratified by tumor stage (TNM) and surgery
*Hazard ratios describe the risk of death, tumor progression or tumor specific death within the observational period
Bold numbers indicate statistically significant log rank tests and HR