DISSERTATION

submitted to the

Combined Faculties for the Natural Sciences and for Mathematics

of the Ruperto-Carola University of Heidelberg, Germany

for the degree of

Doctor of Natural Sciences

presented by

Lea Schroeder, Master of Science

born in: Berlin, Germany

Oral examination: _____

Human Papillomavirus-Driven Neck Lymph Node Metastases from Oropharyngeal or Unknown Primary Squamous Cell Carcinoma

Referees: Prof. Dr. Martin Müller

Prof. Dr. Ralf Bartenschlager

Acknowledgements

This study was performed under the supervision of Dr. Michael Pawlita and Dr. Dana Holzinger between March 2014 and March 2017 in the division “Molecular Diagnosics of Oncogenic Infections” within the Research Program “Infection, Inflammation and Cancer” at the German Cancer Research Center (DKFZ) in Heidelberg.

For reviewing this thesis, I would like to thank Prof. Dr. Müller and Prof. Dr. Bartenschlager. Furthermore, I would like to thank both and Dr. Gerhard Dyckhoff for joining my thesis advisory committee and for their suggestions during the initial towards the final phase.

I would like to express particular thanks to Dr. Michael Pawlita for giving me the opportunity to work on this exciting project in his outstanding lab. I am grateful for the scientific guidance and stimulating discussions and for all kind of help and suggestions. I particularly appreciate the very constructive feedback to presentations and written reports as well as the time invested in proof reading of the manuscript and the thesis. Moreover, I am very grateful for the opportunity to join several international conferences to present the data and meet collaborators.

I was really lucky to get co-supervised by Dr. Dana Holzinger. I would like to thank her for initiating this project, for sharing her knowledge and teaching me. Her door was always open. Thanks for answering all my numerous questions and giving advice, and special thanks for her patience and encouragement during proof reading.

This study benefited greatly from the support of a very large number of national and international collaborators:

At Heidelberg University, I would like to thank Prof. Dr. Jochen Hess and all members of his research group “Experimental and Translational Head and Neck Oncology”, in particular Antje Schumann, for providing biopsies and the cryostat. Special thanks are due to Dr. Gerhard Dyckhoff for his valuable clinical input.

To Prof. Dr. Nikolaus Becker and Anne Byl from the NCT Cancer Registry and to Dr. Esther Herpel and her colleagues from the Institute of Pathology and the NCT Tissue Bank, in particular Dr. Carlota Lucena Porcel and Christiane Zgorzelski, for providing and processing samples, performing the immunohistochemistry and reading the slides.

At the DKFZ, I would like to thank in particular Dr. Manuel Wiesenfarth from the “Division of Biostatistics” for his advice and support, especially for his contribution to the survival analyses.

To Prof. Dr. Christoph Plass and the “Division of Epigenomics and Cancer Risk Factors” for the opportunity to perform the methylation analysis. Especially to Dr. Dieter Weichenhan for adapting the method and processing the data and to Oliver Mücke and Monika Helf for their experimental work.

And to Prof. Dr. Elisabeth Schwarz, who kindly supported the HPV integration analysis, to Ursula Klos for her great laboratory assistance and to Dr. Sasithorn Chotewutmontri for processing the sequencing data.

Thank you to Prof. Dr. Andreas Dietz and Dr. Gunnar Wichmann from the University Hospital in Leipzig, who made the serology study possible. To Maria Willner for collecting the patients’ information and tissue blocks and to Prof. Dr. Christian Wittekind and his colleagues from the Institute of Pathology, who allowed me to prepare sections in their laboratory.

Thank you to Dr. Laia Alemany and Dr. Xavier Castellsagué from the Catalan Institute of Oncology (ICO) and to Dr. Miquel Quer and Dr. Xavier León from the Hospital de la Santa Creu I Sant Pau in Barcelona for providing the Spanish cases and for their constructive feedback.

I am exceptionally grateful to Dr. Paolo Boscolo-Rizzo and Dr. Elisa Dal Cin from the Treviso Regional Hospital and the University of Padua for providing samples but also for their experimental assistance during the months they spent in Heidelberg. Thanks to Dr. Salvatore Romeo and to Dr. Roberta Bertorelle and her colleagues from the Istituto Oncologico Veneto in Padua for performing the TP53 sequencing.

In particular, I would like to acknowledge the help and support I received from Dr. Daniela Höfler, Dr. Lorena Baboci, Birgit Aengeneyndt, André Leischwitz, Astrid Hofmann, Ute Koch and Monika Oppenländer. I would also like to thank all colleagues, especially Julia Butt, Katrin Hufnagel, Juliane Schröter, Christy Varghese, Nicole Brenner and Dr. Angelika Michel, who created this convivial, inspiring and motivating working environment.

Finally, I wish to thank my family and Malte for their unconditional love and support, for believing in me and being continuous source of encouragement and motivation.

Contents

Summary

Patients with neck squamous cell carcinoma from unknown primary tumor (NSCCUP) present with lymph node metastases without evidence for a primary tumor. Most patients undergo an aggressive multimodal treatment, which induces severe toxicity. Primary tumors of NSCCUP can be hidden in the oropharynx. Human papillomavirus (HPV) is causally involved in a subgroup of oropharyngeal squamous cell carcinomas (OPSCC) associated with early lymph node metastasis and good prognosis. Detection of markers for HPV transformation in NSCCUP could allow focusing on the oropharynx in primary tumor search and could be of value for choice and extent of treatment.

In this retrospective multicenter study analyzing 180 NSCCUP cases from Heidelberg, Treviso and Barcelona, a substantial proportion (16%) was driven by HPV, mainly by HPV16 (89%). The prevalence of HPV-driven NSCCUP varied by geographical region, ranging from 10% in Barcelona to 20% in Heidelberg, and increased with year of diagnosis from 9% during 1988-2004 to 23% during 2005-2014 (p=0.007).

Compared to HPV mRNA as gold standard to identify HPV-driven tumors, sensitivity and specificity of HPV DNA, $p16^{INK4a}$ overexpression or the combination of both markers in NSCCUP ranged from 81-100% and 89-100%, respectively, with the lowest concordance for $p16^{INK4a}$ (kappa=0.7) and the highest for the combination (kappa=0.95). HPV seropositivity was a promising diagnostic marker for HPV-driven NSCCUP, since the detection of HPV antibodies in serum from NSCCUP patients had a sensitivity of 91% and specificity of 100% compared to HPV mRNA detected in the metastasis (kappa=0.93).

HPV-driven NSCCUP were molecularly different from non-HPV-driven NSCCUP, because they presented with a distinct DNA methylation pattern in five gene promoters and they did not harbor disruptive TP53 mutations, which were common in non-HPV-driven NSCCUP (52% vs. 0%, p=0.0002). Patients with HPV-driven, as well as HPV-seropositive NSCCUP had significantly better overall and progression-free survival rates ($p \leq 0.002$). Based on the observed survival benefit, HPV mRNA status assessment should be included in NSCCUP diagnosis. Besides an extended diagnostic work-up of the oropharynx in patients with HPV-driven NSCCUP, de-intensification of radiotherapy concentrating on the oropharynx appears a promising therapeutic strategy, the efficacy of which should be assessed in prospective trials.

Analysis of twelve pairs of HPV16-driven OPSCC and corresponding lymph node metastases revealed consistent presence of HPV DNA and mRNA. However, heterogeneity

was observed regarding HPV integration status and DNA methylation. Viral-cellular junctions identified in the primary tumor were present in only 43% of corresponding metastases, while new viral-cellular junctions were detected in 14%. Metastases had overall lower methylation levels in the five gene promoters included in the assessed HPV-associated methylation signature compared to primary tumors.

Zusammenfassung

Patienten mit Plattenepithelkarzinom von unbekanntem Primärtumor im Hals (NSCCUP) weisen Lymphknotenmetastasen ohne Anzeichen eines Primärtumors auf. Nichtauffindbare Primärtumoren können im Oropharynx verborgen sein. Eine Untergruppe von Plattenepithelkarzinomen im Oropharynx, die sich durch frühe Lymphknotenmetastasierung und eine gute Prognose auszeichnen, steht in kausalem Zusammenhang mit dem humanen Papillomvirus (HPV). Der Nachweis von HPV-Markern im NSCCUP könnte auf einen Primärtumor im Oropharynx hinweisen und somit eine ausführliche diagnostische Untersuchung dieser Region indizieren. Zudem könnte der HPV-Status ein entscheidendes Kriterium für die Wahl und das Ausmaß der Behandlung darstellen. Die meisten Patienten erhalten eine aggressive multimodale Therapie, welche schwere Nebenwirkungen hervorrufen kann.

Von den 180 NSCCUP Fällen aus Heidelberg, Treviso und Barcelona, welche in dieser retrospektiven Studie untersucht wurden, erwies sich ein beträchtlicher Anteil (16%) als HPV-getrieben, größtenteils durch HPV16 (89%). Der Anteil an HPV-getriebenen NSCCUP variierte zwischen den Zentren von 10% in Barcelona bis zu 20% in Heidelberg und stieg mit zunehmendem Diagnosejahr an, insgesamt von 9% zwischen 1988 und 2004 bis zu 23% zwischen 2005 und 2014 (p=0.007).

Gegenüber dem HPV mRNA-Status, der als Goldstandard für die Identifizierung von HPV-getriebenen Tumoren gilt, zeigten die Marker HPV DNA und $p16^{INK4a}$ Überexpression, sowie die Kombination dieser beiden Marker eine Sensitivität von 81% bis 100% und eine Spezifität von 89% bis 100%. Die Übereinstimmung mit HPV mRNA war für $p16^{INK4a}$ am geringsten (kappa=0.7) und für die Kombination am höchsten (kappa=0.95). Positiver HPV-Serostatus war ein vielversprechender Marker für HPV-getriebene NSCCUP, weil der Nachweis von HPV-Antikörpern im Serum von NSCCUP Patienten eine Sensitivität von 91% und eine Spezifität von 100% im Vergleich zum Nachweis von HPV mRNA in der Metastase aufwies.

HPV-getriebene NSCCUP unterschieden sich auf molekularer Ebene von nicht-HPV-getriebenen NSCCUP, weil sie ein bestimmtes DNA Methylierungsmuster in fünf Promotoren bevorzugt aufwiesen, aber keine TP53 Mutationen zeigten, die in nicht-HPV-getriebenen NSCCUP häufig vorkamen (52% vs. 0%, p=0.0002).

Patienten mit HPV-getriebenen, sowie HPV-seropositiven NSCCUP zeigten ein besseres Gesamtüberleben und Progressions-freies Überleben ($p \leq 0.002$). Der beobachtete Überlebensvorteil würde die Aufnahme der HPV mRNA-Statusbestimmung in die Standarddiagnostik rechtfertigen. Neben einer ausführlichen Untersuchung des Oropharynx

ist eine De-Intensivierung der Bestrahlung eine vielversprechende therapeutische Strategie, deren Wirksamkeit es in prospektiven Studien zu untersuchen gilt.

Paarweise Untersuchungen von zwölf HPV16-getriebenen Plattenepithelkarzinomen des Oropharynx und dazugehörigen Lymphknotenmetastasen ergab einen übereinstimmenden Nachweis von HPV DNA und mRNA in allen Paaren. Allerdings wurde Heterogenität zwischen den Primärtumoren und Metastasen hinsichtlich des Integrationsstatus der HPV DNA und zellulärer Methylierungsmuster beobachtet. So waren die im Primärtumor identifizierten Integrationsstellen nur in 43% der Metastasen nachweisbar, allerdings wurden in 14% der Metastasen neue Integrationsstellen detektiert. Der Methylierungsgrad war in den fünf Promotoren des HPV-assoziierten Methylierungsmusters in den Metastasen insgesamt niedriger als in den Primärtumoren.

1 Introduction

1.1 Human papillomavirus

1.1.1 Classification

Papillomaviruses are small (50-60 nm diameter), non-enveloped double-stranded DNA viruses [1]. Among them, 210 types are known to infect the mucosa or the skin of humans [152, International Human Papillomavirus Reference Center], thus called mucosal or cutaneous human papillomaviruses (HPVs). HPVs are classified into five genera: alpha-, beta-, gamma-, mu- and nu-HPVs [1]. The genus alpha is subdivided into 14 species and contains all mucosal and some cutaneous HPV types. HPV types are defined by the sequence of the L1 gene that is at least 10% dissimilar from any other type [2]. Mucosal HPV types are further grouped into high-risk types, which are associated with cancer, and low-risk types that are mainly associated with genital warts and non-malignant lesions [3]. Most high-risk mucosal HPV types are included in species 9 (HPV16, 31, 33, 35, 52, 58) and species 7 (HPV18, 39, 45, 59, 68, 70) of the genus alpha [4].

1.1.2 HPV infection and life cycle

Mucosal HPV types infect epithelial cells, more precisely the basal cells in stratified squamous epithelia, in the mucosa of the anogenital region and the upper respiratory tract [4]. The virus binds to heparan sulfate proteoglycans present on the cell surface [5] and is internalized via clathrin-dependent endocytosis [6]. The HPV genome enters the nucleus, where it is maintained in episomal form at low copy numbers and replicates with the host cell [7]. Upon differentiation, HPV-infected cells leave the basal layer towards the suprabasal layers, where the viral genome replicates to more than 1,000 copies per cell and expression of late genes is initiated, finally leading to assembly and release of viral particles from the epithelial surface (Figure 1) [8].

HPV is sexually transmitted through direct genital, anal and oral contacts [9]. Genital HPV infection is cleared by the immune system in most cases. Clearance takes on average 8-14 months for high-risk types [10]. However, genital HPV infection with high-risk types may become persistent, which is associated with the risk of progression to premalignant lesions and finally carcinoma [10]. Upon persistent infection, the episomal HPV genome may integrate into the host cell. HPV integration is considered an important driver of carcinogenesis [11] and is reported for up to 80% of cervical carcinomas [12] and for 0-100% of HPV16 DNA-positive oropharyngeal carcinomas [13-18].

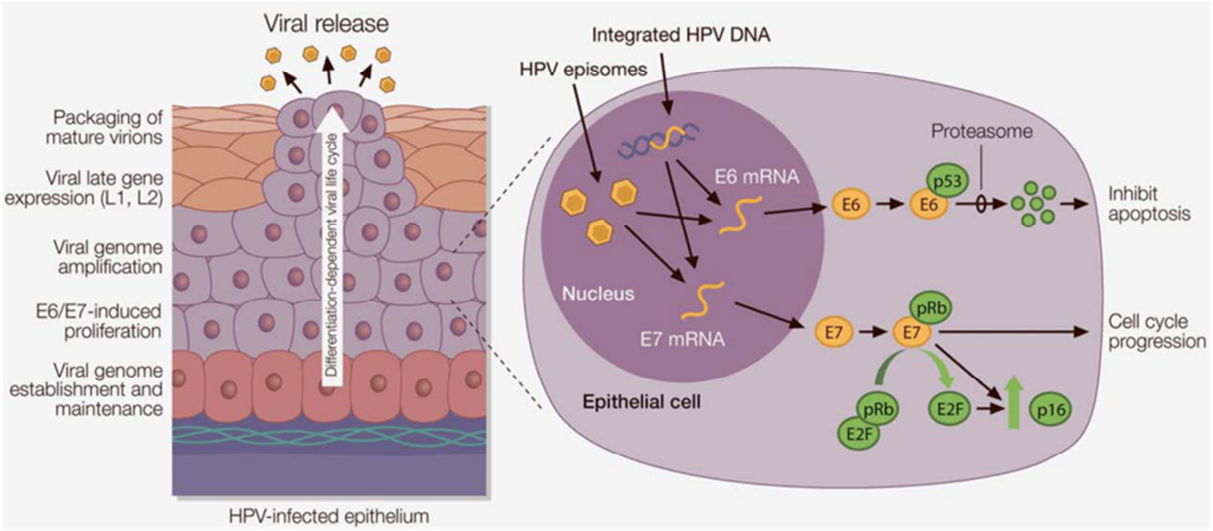


Figure 1: HPV life cycle and oncogenic activity. HPV infects basal cells in stratified epithelia and replicates in suprabasal layers. New infectious virions are released from cells at the surface. Expression of the oncogenes E6 and E7 leads to inhibition of apoptosis and to cell cycle progression through their interaction with the tumor suppressors p53 and pRb, respectively. Source: [19].

1.1.3 Genomic organization and HPV proteins

The circular HPV genome consists of approximately 8,000 base pairs (bp) [4] and contains eight open reading frames (ORFs) on the same DNA strand [1]. It is divided into three regions: (1) the non-coding region, termed long control region (LCR), (2) the early region coding for E1, E2, E4, E5, E6 and E7 proteins, and (3) the late region encoding the structural proteins L1 and L2 (Figure 2). The early genes are expressed from the p97 promoter located in the LCR, whereas the late genes are expressed from the p670 promoter located in the E7 region [8].

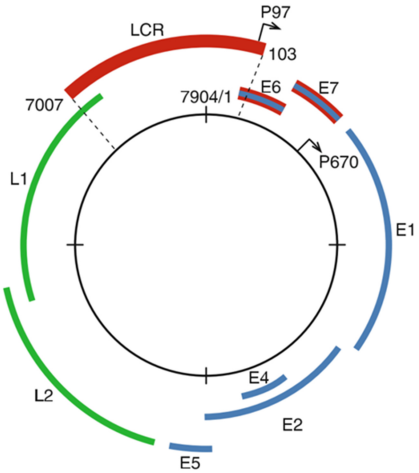


Figure 2: HPV16 genome organization. The HPV16 genome consists of 7904 base pairs and comprises three functional regions: the long control region (LCR, red), the early region (red/blue, blue) and the late region (green). Modified from [4].

The early proteins E1, E2, E4 and E5 play important roles in viral replication (Table 1), while E6 and E7 act as oncoproteins to deregulate cellular functions like cell cycle, apoptosis and differentiation [8]. The late proteins include the major (L1) and minor (L2) capsid proteins that self-assemble into viral capsids consisting of 72 L1 pentamers and a variable number of L2 proteins [20].

Table 1: HPV proteins and their function (modified from [9]).

Protein	Function	References
E1	Adenosine triphosphate(ATP)-dependent helicase; binding to origin of replication; unwinding viral DNA	[21]
E2	Binding to LCR and recruiting E1; regulating late gene expression	[22, 23]
E4	Expressed as fusion protein E1^E4; inducing cell cycle arrest in G_2, virus assembly and release	[24-26]
E5	Enhancing epidermal growth factor signaling	[27]
E6	Inducing degradation of the cellular tumor suppressor p53	[28, 29]
E7	Inducing degradation of the cellular tumor suppressor pRb	[30]
L1	Major capsid protein mediating HPV entry by binding to heparan sulfate proteoglycans on epithelial cells	[5]
L2	Minor capsid protein mediating transfer of the HPV genome to the nucleus of the infected cell	[31]

1.1.4 HPV carcinogenesis

HPV has been shown to be associated with several anogenital cancers including cervical [32], vaginal [33], vulvar [34], anal [35] and penile [36], as well as oral cancers [9, 37, 38]. In order to replicate in suprabasal epithelial layers consisting of differentiating cells, in which DNA replication is normally suppressed, HPV needs E6 and E7 to prevent the cells from growth arrest and to inhibit apoptosis [39]. Inhibition of apoptosis is mediated by E6, which induces proteasomal degradation of the cellular tumor suppressor protein p53 by forming a complex with the E6-associated protein (E6AP), a cellular E3 ubiquitin ligase [28, 29]. In addition, E6 prevents cellular senescence by activating the human telomerase reverse transcriptase (hTERT) through the interaction with E6AP [40, 41]. The telomerase stabilizes telomere length, thereby allowing indefinite proliferation. Cellular proliferation is induced by E7 via disruption of the cyclin-dependent kinase (CDK) inhibitors p21WAF1/CIP1 and p27KIP1 [42-44] and via degradation of tumor suppressors from the Rb family (pRb, p107,

p130). Binding of E7 to pRb within a cullin 2 ubiquitin ligase complex, leads to degradation of pRb [30, 45] and to the release of E2F transcription factors [46].

1.1.5 HPV detection methods

Commonly applied methods for the detection of HPV DNA or mRNA in human tissue are polymerase chain reaction (PCR) and in situ hybridization (ISH). Immunohistochemistry (IHC) is used to detect overexpression of the surrogate marker $p16^{INK4a}$, which is triggered by the interaction of the viral E7 protein with cellular pRb (Figure 1). Transcriptionally active HPV, and thus clinically relevant infection, can only be detected by targeting HPV mRNA [47], which is technically challenging but possible [48-50]. While PCR and ISH can be designed to detect several HPV types within one reaction (multiplex), only type-specific PCR can distinguish detected types. In contrast, ISH and IHC allow HPV detection within the tissue context, but their interpretation is observer-dependent.

A meta-analysis, comparing those techniques with the detection of HPV mRNA, which is the gold standard to identify HPV-driven tumors in research [47, 51, 52], revealed highest pooled sensitivity for HPV detection by PCR (98%), followed by IHC (94%), and lowest sensitivity for ISH (85%) [53]. Pooled specificity is highest for ISH (88%), but lower for IHC (83%), since $p16^{INK4a}$ overexpression might be triggered by non-viral mechanisms, and for PCR (84%) due to the risk of cross-contamination and the high analytical sensitivity allowing the detection of few individual HPV-infected cells [47, 53].

HPV detection by PCR works well in frozen tissue, but is challenging in formalin-fixed paraffin-embedded (FFPE) tissue, which is often available in the clinic and for retrospective studies, because of formalin-induced cross-linking or potential degradation of nucleic acids. In contrast, ISH and IHC are well suited for FFPE tissue and for routine use in the clinic, since those are standard or easily transferable techniques for pathology laboratories. Suggested algorithms to identify HPV-driven head and neck cancers in clinical settings are $p16^{INK4a}$ IHC, if positive followed by HPV PCR [51] or ISH [47]. Only double positive (HPV DNA/$p16^{INK4a}$-positive) cases are considered HPV-driven. The combination of positivity by PCR and IHC increased pooled specificity to 96% with 93% sensitivity when compared to positivity for HPV mRNA in the meta-analysis [53].

For oropharyngeal tumors, an evolving diagnostic method without need of tumor tissue is the detection of HPV antibodies in serum or plasma, especially antibodies against the HPV16 E6 protein, which show strong correlation with presence of HPV DNA, mRNA and $p16^{INK4a}$ overexpression in tumor tissue [54]. Using this technique, high-risk HPV infection can even be detected more than ten years prior to tumor diagnosis [55].

1.2 Head and Neck Squamous Cell Carcinoma

1.2.1 Epidemiology

In the head and neck region, cancers arise in the oral cavity, the pharynx with its subdivisions nasopharynx, oropharynx and hypopharynx, in the larynx, the nasal cavity, paranasal sinuses and in salivary glands (Figure 3). More than 90% of head and neck cancers are squamous cell carcinomas (HNSCC) [56]. Head and neck cancer is the sixth most common cancer [57] with 686,328 estimated new cases in 2012 worldwide and 139,603 in Europe, comprising 61,416 cancers in the lip and oral cavity, 39,921 in the larynx, 34,094 in the oropharynx and hypopharynx and 4,172 in the nasopharynx (http://globocan.iarc.fr). The estimated age-standardized mortality rate is 9.0 (males) and 1.3 (females) cases per 100,000 (http://globocan.iarc.fr). For oropharyngeal squamous cell carcinomas (OPSCC), a significant increase in incidence was reported in recent years for North American and European countries [58].

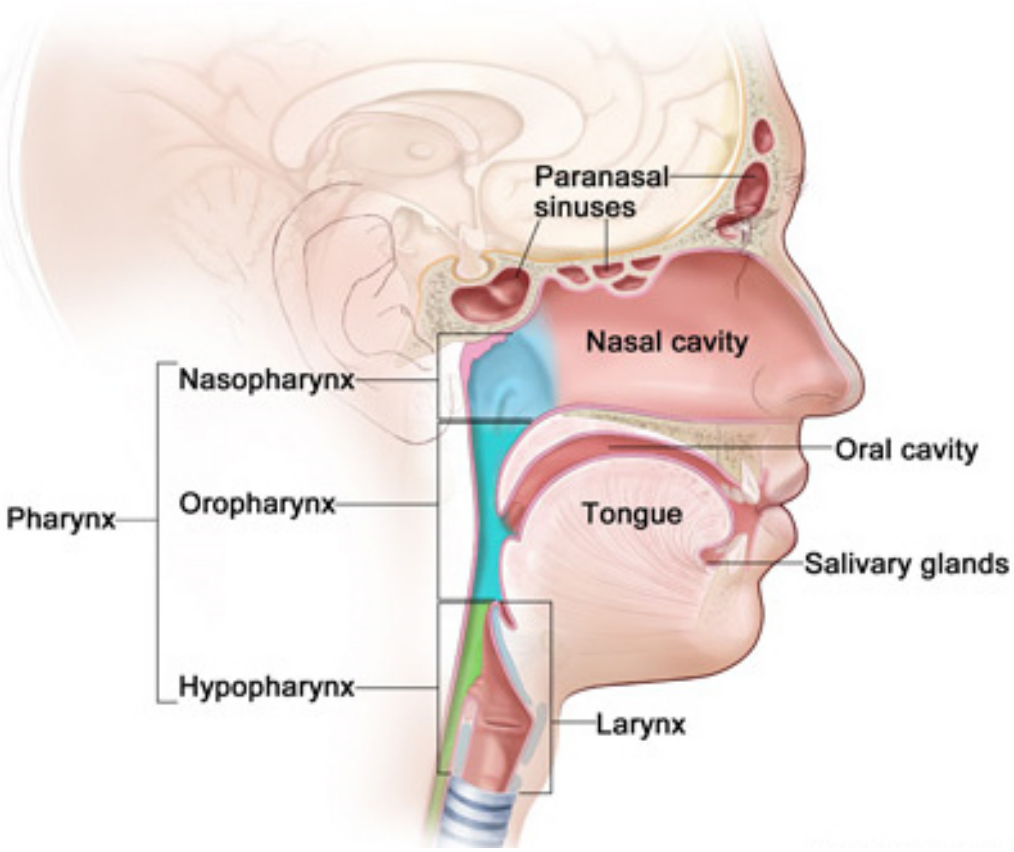


Figure 3: Head and neck cancer regions. Head and neck cancers arise in the pharynx (including nasopharynx, oropharynx and hypopharynx), larynx, oral and nasal cavity, paranasal sinuses and salivary glands. Source: https://www.cancer.gov/types/head-and-neck/head-neck-fact-sheet.

1.2.2 Risk factors

Major risk factors for head and neck cancer are tobacco and alcohol consumption. A pooled analysis found 35% of head and neck cancers to be attributed to the consumption of both tobacco and alcohol, 33% to tobacco alone and 4% to alcohol consumption alone [59]. Other risk factors are exposition to dust or industrial toxins, sun exposition for carcinomas in the lip, betel nut chewing for oral cavity carcinomas, radiation for carcinomas in salivary glands, poor mouth hygiene and maybe even alcoholic mouth wash for OPSCC [60].

In an US study population, OPSCC was significantly associated with a higher number of lifetime genital-sex and oral-sex partners, as well as with oral HPV16 infection [61], suggesting sexually transmitted HPV as a risk factor. Infection with high-risk HPV is causally associated with a subset of OPSCC [9, 62, 63] and Epstein-Barr virus (EBV) has been linked to nasopharyngeal carcinoma [64].

1.2.3 HPV prevalence in HNSCC

Among head and neck cancers, the highest HPV prevalence is reported for the oropharynx, particularly for the palatine tonsils followed by the base of tongue (53.9% and 47.8% HPV DNA-positive, respectively), whereas the HPV prevalence outside the oropharynx is low [65]. A retrospective study analyzing 281 HNSCC cases from the United Kingdom revealed 70% HPV DNA/$p16^{INK4a}$-positive OPSCC, but only 7.1% in the hypopharynx, 3.2% in the larynx and 4% in the oral cavity [66]. In an international study including 3,650 HNSCC cases from 29 countries, the HPV-driven proportion (HPV DNA/mRNA-positive) was 21.8% in the oropharynx, 15.1% in the unspecified pharynx, 5.9% in the nasopharynx, 3.9% in the oral cavity, 3.1% in the larynx and 2.4% in the hypopharynx [67]. HPV16 is the most common type, ranging from 51% of all HPV DNA-positive cases in the larynx up to 83% in the oropharynx [67].

The HPV prevalence varies not only by cancer site, but also by geographical region. A meta-analysis, including 148 studies with all together 12,163 HNSCC cases from 44 countries, revealed for the oropharynx the highest HPV DNA prevalence in North America (60.4%) and the lowest in South and Central America (14.9%) [65]. Within Europe, the highest HPV DNA prevalence is reported in Northern Europe (56.5%), followed by Central and Eastern (41.8%), Western (37.6%) and Southern Europe (24.2%) for the oropharynx [65].

1.2.4 HPV infection in the oropharynx

The high HPV prevalence in the oropharynx is probably due to facilitated HPV infection in the oropharynx, in particular in the tonsillar crypts representing a favored niche for HPV. The normal function of the crypts is to maximize the surface of the tonsils in order to foster the

transport of antigens taken up via the mouth to the lymphoid tissue [68]. The reticulated structure of the epithelium facilitates migration of immune cells like lymphocytes and antigen-presenting cells (Figure 4). This structure is favorable for HPV, because the virus can easily migrate through the epithelium to enter the basal cells, where it replicates. It may even play a role in cancer progression by promoting early invasion and metastasis [68].

Another favorable characteristic of the tonsillar epithelium is the expression of the programmed death-ligand 1 (PD-L1) on the membrane of epithelial cells observed in benign tonsils [69]. PD-L1 induces immune suppression by binding to the PD-1 receptor on T cells and turning those from an activated into an anergic state. The PD-L1/PD-1 pathway might therefore be important for persistence of HPV infection and immune resistance during malignant progression [69].

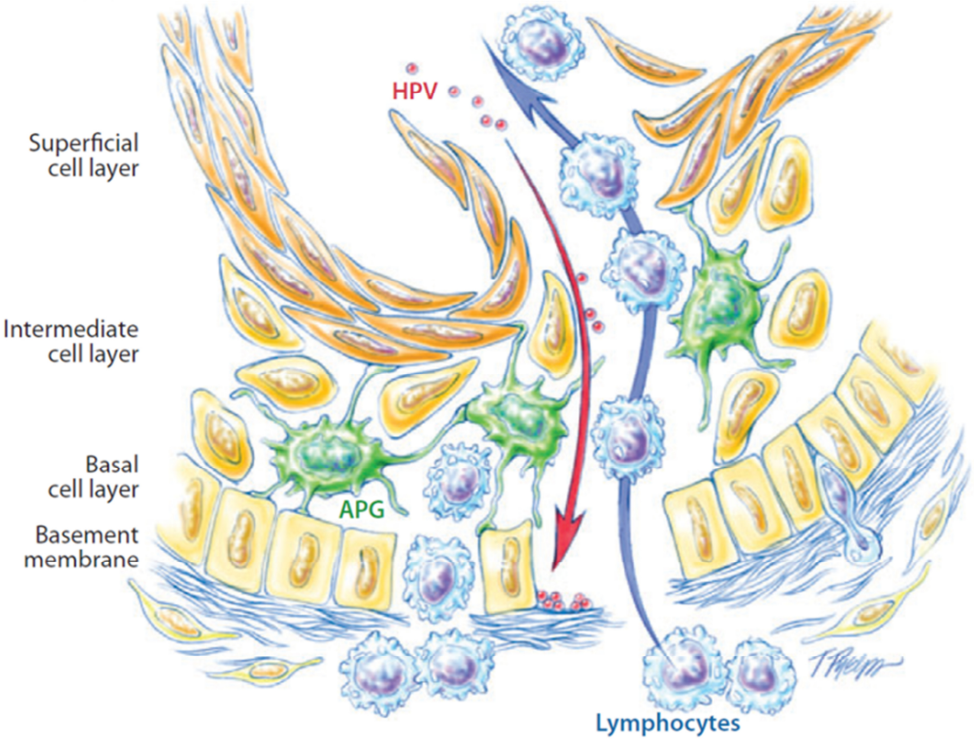


Figure 4: Reticulated epithelium in the tonsillar crypts. In the crypts of the palatine tonsils, the squamous epithelium is reticulated to allow migration of lymphocytes and cells from the antigen-presenting group (APG). This facilitates HPV infection and entry into the basal cells. Adapted from [68].

1.2.5 Characteristics of HPV-positive HNSCC

Epidemiological, molecular and clinical differences between HPV-positive and HPV-negative HNSCC have been identified [70]: Patients with HPV DNA-positive HNSCC report higher oral sex exposure, but they are more often nonsmokers and mild or moderate drinkers, whereas heavy smokers and drinkers present with HPV DNA-negative HNSCC [61, 71]. While

mutations in the TP53 gene encoding the p53 protein are common in HNSCC leading to loss of p53 function, HPV DNA-positive HNSCC mostly present with TP53 wild-type sequence since p53 is inactivated by the viral E6 protein [72, 73].

Additional analysis of HPV DNA-positive HNSCC for presence of HPV mRNA revealed a distinct group of HPV DNA-positive/mRNA-negative HNSCC, with molecular and clinical characteristics that were similar to HPV DNA-negative HNSCC [74]. Gene expression profiles and TP53 mutation rates in those HPV DNA-positive/mRNA-negative HNSCC were similar to HPV DNA-negative HNSCC, while both groups were significantly different from HPV-driven (mRNA-positive) HNSCC [75].

In HPV-driven HNSCC, the CDKN2A gene encoding $p16^{INK4a}$ is commonly upregulated leading to overexpression of $p16^{INK4a}$ [74]. A promoter methylation signature has been identified to be associated with HPV-driven OPSCC [76] and to correlate with better clinical outcome in HNSCC patients independent of the HPV status [77].

In previous studies, HPV DNA-positive HNSCC were more radiosensitive and those patients had a better survival than patients with DNA-negative HNSCC [78], which was also observed for patients with $p16^{INK4a}$-positive compared to $p16^{INK4a}$-negative HNSCC [79, 80]. Based on this survival benefit and the distinct characteristics, it is assessed in ongoing clinical trials whether HPV DNA-positive and/or $p16^{INK4a}$-positive OPSCC patients might benefit from de-intensified treatment [81]. However, HPV mRNA analysis revealed that only HPV-driven OPSCC patients show a better survival, while for HPV DNA-positive/mRNA-negative OPSCC the survival rate is as poor as for HPV DNA-negative OPSCC [49].

1.3 Neck Squamous Cell Carcinoma from Unknown Primary Tumor

1.3.1 Clinical presentation

Between 2% and 9% of all head and neck cancer patients initially present with an enlarged neck lymph node that turns out to be a metastasis from squamous cell carcinoma, but frequently without detectable primary site [82]. This syndrome is termed neck squamous cell carcinoma from unknown primary tumor (NSCCUP). The median size of the lymph node metastases is 5 cm, ranging from 2-14 cm [83], but pain is reported in only 9% of patients [84]. In 80-90% of patients multiple lymph nodes are affected, but only in 10% on both sides of the neck (bilateral vs. unilateral) [85]. Most neck lymph node metastases are located in lymph node level II (30-50%), followed by level I and III (10-20%), and level IV and V (5-10%) [86] (Figure 5).

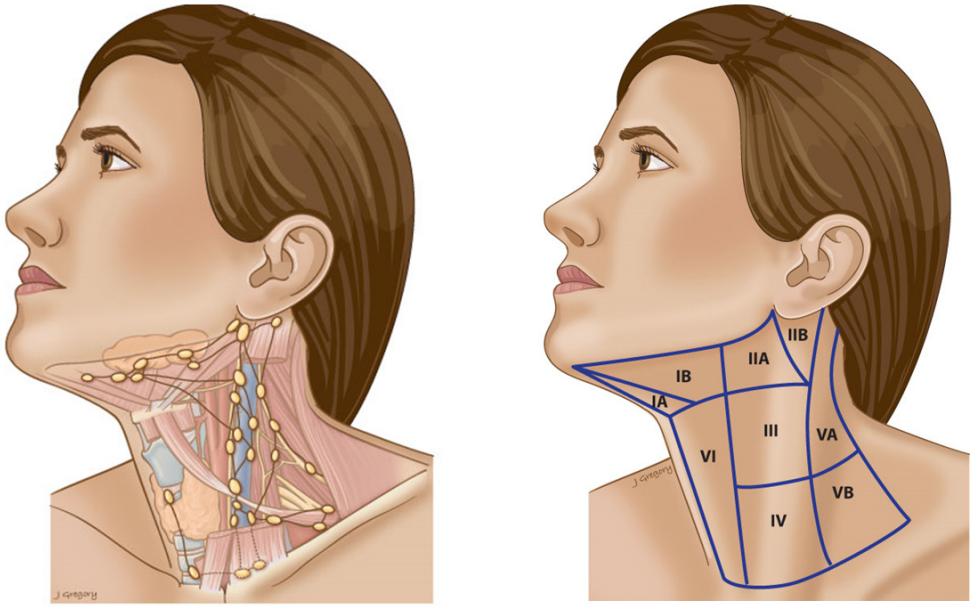


Figure 5: Lymph node levels of the neck. The neck lymph nodes (left) are grouped into different regions called levels (right). Source: http://www.headandneckcancerguide.org/adults/introduction-to-head-and-neck-cancer/neck-cancer/metastatic-lymph-nodes/anatomy/.

1.3.2 Diagnostics

Diagnostic procedures for patients with NSCCUP aim at assessing the lymph node involvement and searching for the primary site. Pathological staging is done by neck dissection to assess if only a single lymph node smaller than 3 cm (nodal stage N1) or larger (N2a) is affected, or multiple ipsilateral (N2b), bilateral or contralateral lymph nodes (N2c), or lymph nodes larger than 6 cm (N3) are affected. The position of the metastases, more precisely the lymph node level (Figure 5), is the best indicator for the localization of the primary tumor. Potential primary sites represent the organs draining into the metastatic lymph nodes (Table 2). Most metastases are reported in level II, where they may be derived from carcinomas in the nasopharynx, oropharynx or larynx, whereas primary tumors from metastases found in the supraclavicular level V would be expected to be located outside of the head and neck region.

Table 2: Location of neck lymph nodes and possible primary sites (modified from [86]).

Level	Involved neck lymph nodes	Possible primary sites
I	Submandibular	Mouth floor, lips, anterior part of the tongue
II	Jugulodigastric/upper jugular	Nasopharynx, base of tongue, tonsils, larynx
III	Middle jugular	Supraglottic larynx, inferior pyriform sinus, postcricoid region
IV	Inferior jugular	Hypopharynx, subglottic larynx, thyroid, esophagus
V	Supraclavicular	Lung, thyroid, breast, gastrointestinal system

The standard diagnostic work-up of patients with NSCCUP includes patients' history, physical examination, fine-needle aspiration biopsy (FNAB), imaging modalities like computed tomography (CT) and/or magnetic resonance imaging (MRI), as well as panendoscopy of the upper aerodigestive tract with biopsies from suspicious areas and from the tonsils and the base of tongue, which are common primary sites [86-88]. More recently, CT is combined with ^{18}F-fluorodeoxyglucose-positron emission tomography (PET/CT). Roh et al. revealed a significantly higher sensitivity of PET/CT vs. CT (87.5% vs. 43.7%) for the detection of primary tumors in 44 NSCCUP patients [89].

Patients benefit from localization of the primary tumor, especially in regions of favorable subsets, because they may be eligible for targeted treatment [90]. Mozet et al. reported that primary tumors are most often located in the head and neck region (60-80%), to a lesser extent in the lung (15-20%) and rarely in the gastrointestinal tract [91]. A study including 236 US patients with lymph node metastasis in level II and/or III revealed localization of 53.4% of the identified primary tumors in the tonsillar fossa, followed by 43.9% in the base of tongue [87]. Most NSCCUP patients undergo a unilateral or bilateral tonsillectomy, if not performed earlier because of tonsillitis, to reveal and directly resect suspicious tonsillar carcinomas. Contralateral or bilateral tonsillar carcinomas were identified after bilateral tonsillectomy in 23% to 25% (4/16 and 5/22, respectively) of NSCCUP patients presenting with unilateral metastases [92, 93]. The true incidence of synchronous bilateral tonsillar carcinoma is likely to be underreported and since bilateral tonsillectomy adds only minimal morbidity, this is recommended for NSCCUP patients [94].

In general, the identified primary tumors are much smaller than the lymph node metastases. Zengel et al. reported an average size of 1.1 cm (0.3-2.4 cm) for primary tumors in the tonsil and 0.8 cm (0.2-2.2 cm) in the base of tongue in contrast to large lymph node metastases with an average size of 3.1 cm (0.5-6.3 cm) for tonsillar and 2.7 cm (1.4-5.5 cm) for base of tongue carcinomas [95]. This significant difference in size might explain why the metastasis is diagnosed first, whereas the primary tumor sometimes cannot be detected by imaging, but only after taking biopsies. In about 10% of patients with NSCCUP, the primary site is revealed later at some time after initial diagnosis, usually within 2 years [96].

1.3.3 Treatment

Since no randomized trials have been performed yet, the choice of treatment for NSCCUP patients is based on evidence from small non-randomized retrospective studies, on clinical experience from HNSCC treatment and on institutional policy [86, 97]. It is suggested to treat NSCCUP similar to locally advanced head and neck cancer [98]. The guidelines of the European Society for Medical Oncology (ESMO) recommend neck dissection and/or bilateral

irradiation, and for advanced stages induction chemotherapy with platinum-based combination or chemoradiotherapy [99].

In a neck dissection, potentially affected lymph nodes are removed. While a radical neck dissection also targets the spinal accessory nerve, internal jugular vein and sternocleidomastoid muscle, a modified neck dissection spares those structures to preserve shoulder mobility [100]. A less invasive option is selective neck dissection removing only certain groups of lymph nodes in restricted levels [97].

Postoperative radiotherapy is particularly recommended for patients with multiple lymph nodes affected (N2b or higher) and/or extracapsular spread (ECS), meaning that the tumor penetrates the lymph node capsule, to include potential primary sites in the radiation field [96, 100]. For patients with poor general condition or comorbidities, or in case of risk of causing harm to adjacent structures, primary radiotherapy might be indicated [96]. A retrospective analysis of 106 NSCCUP patients revealed a lower regional recurrence rate and reduced appearance of the primary tumor in irradiated patients (9%) compared to patients that did not receive irradiation (32%, p=0.006), while overall and disease-free survival rates were similar [101].

However, more extensive treatment is associated with a higher risk for acute toxicity and late complications (Pavlidis 2009). High-grade acute skin toxicity was reported in 45% of irradiated patients, mucositis in 34-43%, dysphagia (impaired ability to swallow) in 35%, and high-grade late neck fibrosis and xerostomia (dryness in the mouth) in 27-34% and 34-61%, respectively [101, 102]. Late toxicity can be significantly decreased by using intensity-modulated radiation therapy to deliver high dose to lymph nodes, but minimized doses to organs at risk like salivary glands, while reaching similar overall and disease-free survival rates [103].

Additional use of chemotherapy might be indicated to reduce the risk of distant metastases. Patients with advanced nodal disease had less distant metastases after chemoradiotherapy compared to patients treated without chemotherapy (36% vs. 59%) [104]. In contrast, no benefit regarding overall and progression-free survival and loco-regional control, but significantly increased high-grade acute (59% vs. 25%) and late toxicity (47% vs. 14%) was reported for patients treated with concurrent chemoradiotherapy compared to radiotherapy [105].

1.3.4 Outcome and prognostic factors

Among all patients with carcinoma from unknown primary, patients with NSCCUP represent a favorable sub-set with 5-year survival rates ranging from 36% to 75% [96]. It has been

shown that subsequent discovery of the primary tumor in NSCCUP patients is associated with better overall and disease-free survival with a hazard ratio of 0.13 (0.02-0.82) and 0.25 (0.07-0.91), respectively [90]. In addition, several clinical and epidemiological parameters were identified as prognostic factors. Park et al. revealed the total volume and the ratio of metastatic lymph nodes as independent prognostic factors for disease-free survival with hazard ratios of 13.08 ($p=0.004$) and 17.28 ($p<0.001$), respectively [106]. Huang et al. identified increased nodal stage, age at diagnosis and surgical treatment compared to non-surgical treatment as prognostic factors with hazard ratios of 5.85 ($p=0.01$), 1.08 ($p<0.0001$) and 0.4 ($p=0.04$), respectively [107]. Extracapsular spread was an independent prognostic factor for overall survival ($p=0.013$) in a long-term follow-up study including 26 NSCCUP patients [108]. NSCCUP patients with low p53 expression had a significantly better 5-year overall survival than those with high p53 expression (69% vs. 14%) [109]. But the major prognostic factor discussed in the literature is HPV [110].

1.3.5 HPV as a biomarker in NSCCUP

The major challenges of the NSCCUP syndrome are how to find the primary site and how to treat the patients without knowing the origin of the metastasis and thus the site and aggressiveness of the tumor. Since a subset of OPSCC has been shown to be causally associated with HPV [9, 62, 63, 111, 112], whereas the HPV prevalence is low outside of the oropharynx [65-67], detection of HPV in the lymph node metastases of NSCCUP patients might be an indication for an occult HPV-driven oropharyngeal primary tumor [110]. Furthermore, NSCCUP patients with an occult HPV-driven OPSCC might have a better prognosis based on the survival benefit previously reported for patients with HPV-driven OPSCC [49].

Recently, we have systematically reviewed previous knowledge about the role of HPV in NSCCUP [113]. In the 18 evaluated studies including altogether 659 NSCCUP patients, the HPV prevalence defined by positivity for both HPV DNA and $p16^{INK4a}$ overexpression varied between 0% and 85%, with an average of 36% (Table 3). Seven studies showed that HPV may predict an oropharyngeal localization of the primary tumor [114-120].

Univariate survival analysis was performed in eleven studies (Table 4), and revealed in six studies a significantly better prognosis of patients with HPV-positive compared to HPV-negative NSCCUP [90, 109, 115, 118, 121, 122]. However, no significant survival difference was found in the remaining five studies [123-127], likely due to small study populations. The median number of NSCCUP cases studied was 34, ranging from 10 to 68 (Table 3).

Multivariate survival analysis revealed a better overall survival of patients with HPV-positive NSCCUP with hazard ratios between 0.29 (p=0.03, adjusted for p53, gender, age and smoking) and 0.71 (p=0.009, adjusted for age, smoking, alcohol and non-keratinizing morphology) [109, 122, 128], and a better disease-free survival with a hazard ratio of 0.29 (p=0.03, adjusted for ECS and p53) [115], whereas one study did not reveal a significant difference [126].

Table 3: HPV prevalence in NSCCUP (modified from [113]).

Country	Cases	HPV positivity (%)			Reference
	(n)	DNA	$p16^{INK4a}$	both	
USA	68	-	95.6	-	Graboyes, 2014 [129]
Germany	63	52.4*	57.1	36.5	Tribius, 2012 [123]
Denmark	60	21.7*	21.7	18.3	Jensen, 2014 [122]
Germany	59	22.0*/**	22.0	22.0	Vent, 2013 / Straetmans, 2014 [118, 124]
Korea	58	53.4**	50.0	43.1	Park, 2012 [115]
Sweden	50	40.0*	42.0	36.0	Sivars, 2014 [109]
USA	39	64.1**	74.1	64.1	Davis, 2014 [90]
USA	35	-	74.3	-	Keller, 2013 [121]
Japan	33	18.2**	24.2	18.2	Kobayashi, 2014 [117]
USA	34	-	47.1	-	Perkins, 2012 [125]
Japan	27	37.0*	51.9	37.0	Yasui, 2014 [119]
Germany	26	80.8*	80.8	80.8	Zengel, 2012 [95]
USA	25	28.0**	44.0	28.0	Compton, 2011 [126]
Netherlands	20	0.0*/**	0.0	0.0	Straetmans, 2014 [124]
USA	17	58.8**	58.8	58.8	Demiroz, 2014 [127]
Germany	13	92.3*	84.6	84.6	Weiss, 2011 [120]
USA	13	23.1**	15.4	15.4	Chenevert, 2012 [130]
USA	10	30.0**	30.0	30.0	Begum, 2007 [116]

both=positive for HPV DNA and $p16^{INK4a}$ overexpression; *DNA detected by PCR; **DNA detected by ISH

Table 4: Survival of NSCCUP patients in relation to HPV status (modified from [113]).

Reference	HPV positivity (HPV+)	Cases (n)	Survival of HPV+ vs. HPV-
Tribius (2012)	DNA and $p16^{INK4a}$	63	not significant
Jensen (2014)	DNA and $p16^{INK4a}$	60	OS HR=0.09, p=0.02; PFS HR=0.23, p=0.04
Park (2012)	DNA (OS), $p16^{INK4a}$ (DFS)	58	4-year OS 80.8% vs. 52.7%, p<0.05; 4-year DFS 85.0% vs. 56.9%, p=0.02
Straetmans (2014)	DNA and $p16^{INK4a}$	51	not significant
Sivars (2014)	$p16^{INK4a}$	50	5-year OS 76.2% vs. 37.9%, p=0.007; 5-year DFS 85.7% vs. 62.1%, p=0.03
Davis (2014)	DNA and/or $p16^{INK4a}$	39	OS (HPV-) HR=10.3, p<0.05; DSS (HPV-) HR=6.7, p=0.03
Vent (2013)	$p16^{INK4a}$	37	5-year OS 69% vs 33%, p<0.05
Keller (2013)	$p16^{INK4a}$	35	5-year OS 92% vs. 30%, p<0.0001; 5-year DSS 92% vs. 60%, p=0.09
Perkins (2012)	$p16^{INK4a}$	34	not significant
Compton (2011)	DNA and $p16^{INK4a}$	25	not significant
Demiroz (2014)	DNA and $p16^{INK4a}$	17	not significant

DNA=HPV DNA positivity (PCR or ISH), HR=hazard ratio, OS=overall survival, DSS=disease-specific survival, DFS=disease-free survival, PFS=progression-free survival,

While in the majority of studies, the combination of HPV DNA detection and $p16^{INK4a}$ overexpression was used to identify HPV-positive NSCCUP, three studies only assessed $p16^{INK4a}$ overexpression and none considered presence of HPV mRNA, which is the gold standard for the identification of HPV-driven tumors [47, 49]. In this review, we pointed out the potential of HPV as a biomarker in NSCCUP and the need for studying larger series of NSCCUP patients from various regions using adequate HPV detection methods to determine the HPV-driven proportion and to clarify the prognostic impact of HPV.

1.4 Aim of this thesis

Previous studies give evidence for HPV being a promising biomarker in patients with NSCCUP indicating oropharyngeal localization of the occult primary tumor and a better prognosis, but those studies are limited by small sample size and HPV markers analyzed. Aim of this study was to investigate the role of HPV in NSCCUP by determining the proportion of HPV-driven NSCCUP in a large multicenter study. Besides enlarging the power by increasing the number of analyzed cases, the retrospective setting and the inclusion of study centers in three European countries (Germany, Italy and Spain) allowed the analysis of time trends, as well as regional differences in HPV prevalence.

In order to identify cases with actively transcribed and thus clinically relevant HPV, presence of HPV mRNA was assessed. In addition, the HPV mRNA status was compared to presence of HPV DNA and $p16^{INK4a}$ overexpression, which are the HPV markers widely used in clinical settings, to evaluate the concordance between the markers and the importance of HPV mRNA testing in NSCCUP. Furthermore, the potential of detecting HPV antibodies in serum from NSCCUP patients as a novel HPV detection method was evaluated by comparing the HPV serostatus with the molecular HPV status in the patients' tumor tissue.

The HPV status was correlated with the patients' outcome in order to assess the prognostic impact of HPV detected in lymph node metastases from NSCCUP patients. NSCCUP cases driven by HPV were compared to non-HPV-driven cases regarding epidemiological, clinical and molecular differences in order to assess whether they represent a biologically distinct entity with a distinct risk profile. On the molecular level, the TP53 mutation status was assessed as a prognostic marker, as well as a prognostic promoter methylation signature known to be associated with HPV-driven OPSCC.

Another aim of this thesis was to assess potential heterogeneity regarding HPV status and HPV-associated markers between lymph node metastases and primary tumors. Therefore, pairs of HPV-driven OPSCC and corresponding neck lymph node metastases were compared, since HPV-driven metastases in NSCCUP patients are expected to be mainly derived from occult HPV-driven OPSCC. In particular, the stability of integrated HPV upon metastasis formation and an HPV-associated DNA methylation signature were investigated.

2 Materials

2.1 Reagents and solvents

Acetic acid, 100%	Merck, Darmstadt
Aceton	Sigma-Aldrich, Steinheim
Agarose NEEO ultra-quality	Carl Roth, Karlsruhe
Bromphenol blue (BPB) marker	Serva Electrophoresis, Heidelberg
Casein	Sigma-Aldrich, Steinheim
Dideoxynucleuotides (ddNTP)	Carl Roth, Karlsruhe
Eosin	Carl Roth, Karlsruhe
Ethanol, absolute	Sigma-Aldrich, Steinheim
Ethidium bromide	Sigma-Aldrich, Steinheim
Ethylenediaminetetraacetic acid (EDTA), pH 8.0, UltraPure™, 0.5M	Invitrogen, Carlsbad, CA, USA
Eukitt	O. Kindler GmbH, Freiburg
Glycerol, 100%, anhydrous	Carl Roth, Karlsruhe
Hydrochloric acid (HCl), 37%	Riedel-de Häen, Seelze
MagNA Pure DNA Tissue Lysis Buffer	Roche, Applied Science, Mannheim
MagNA Pure LC RNA Isolation Tissue Lysis Buffer	Roche, Applied Science, Mannheim
Mayer's Haemalaun Solution	AppliChem GmbH, Darmstadt
$MgCl_2$	Thermo Fisher Scientific
2-(N-Morpholino) ethanesulfonic acid (MES)	Sigma-Aldrich, Steinheim
Phosphate buffered saline (PBS), powder without Ca^{2+}, Mg^{2+}	Biochrom GmbH, Berlin
RNase AWAY spray	Thermo Fisher Scientific, Dreieich

Sarcosyl (N-Lauroylsarcosine), 20%	Sigma-Aldrich, Steinheim
Sodium dodecyl sulfate (SDS)	Serva Electrophoresis, Heidelberg
Streptavidin-R-Phycoerythrin (Strep-PE)	Invitrogen, Carlsbad, CA, USA
Tetramethylammonium chloride (TMAC)	Sigma-Aldrich, Steinheim
Tris-EDTA (TE) buffer, RNase-free, 1x (10 mM Tris-HCl, 1 mM EDTA, pH 8.0)	Acros Organics, Geel, Belgium
Tris-HCl, pH 8.0, RNase-free, 1M	Jena Bioscience, Jena
Tween20®, RNase-free	Serva Electrophoresis, Heidelberg
Water, DNase/RNase-free	Invitrogen, Carlsbad, CA, USA
xMAP™ Sheath Fluid	Luminex Corporation, Austin, TX, USA
Xylene	Fisher Scientific, Pittsburgh, PA, USA

2.2 Buffers

6x BPB DNA-loading buffer

0.25% (w/v) bromophenol blue (BPB), 30% (v/v) glycerol, H_2O ad 100 ml

Coupling wash buffer I (Luminex)

50 µl Tween20®, H_2O ad 250 ml

Coupling wash buffer II (Luminex)

2.5 ml of SDS (10%), H_2O ad 250 ml

Detection solution DNA/RNA (Luminex)

2M TMAC, 75mM Tris-HCl, pH 8.0, 6mM EDTA, pH 8.0, 1.5% (w/v) Sarcosyl

DNA hybridization solution (Luminex)

0.15M TMAC, 75mM Tris-HCl, pH 8.0, 6mM EDTA, pH 8.0, 1.5% (w/v) Sarcosyl

Hybridization wash buffer (Luminex)

0.02% Tween20®, 1x PBS, pH 7.4

MES coupling buffer, 0.1M (Luminex)

4.88 g MES, H_2O ad 250 ml, adjust to pH 4.5 with NaOH (5M)

2.3 Enzymes

AmpliTaq Gold DNA Polymerase (5 U/µl)	Thermo Fisher Scientific, Dreieich
Proteinase K, recombinant, PCR grade	Roche, Applied Science, Mannheim
RNase-free DNase (QIAGEN)	QIAGEN, Hilden
Taq DNA Polymerase, native	Thermo Fisher Scientific, Dreieich

2.4 Antibodies

Biotinylated goat anti-human IgG	Dianova, Hamburg
Mouse anti-human p16, Clone: G175–405	BD Pharmingen, San Jose, CA, USA

2.5 Commercial kits and reagents

DNA Clean & Concentrator-5 Kit	Zymo Research, Freiburg
EZ DNA Methylation Kit	Zymo Research, Freiburg
GeneJet Gel-Extraction Kit	Thermo Fisher Scientific, Dreieich
HotStarTaq Plus Master Mix Kit	QIAGEN, Hilden
LightCycler® 480 Probes Master	Roche, Applied Science, Mannheim
LightCycler® 480 RNA Master Hydrolysis Probe Kit	Roche, Applied Science, Mannheim
MagNA Pure 96 Cellular RNA Large Volume Kit	Roche, Applied Science, Mannheim
MagNA Pure 96 DNA and Viral NA Large Volume Kit	Roche, Applied Science, Mannheim
Multiplex PCR Kit and PCR Plus Kit	QIAGEN, Hilden
Nextera DNA Sample Preparation Kit	Illumina, San Diego, CA, USA
OncoE6™ Oral Test	Arbor Vita Corporation, Fremont, CA, USA
OptiView DAB IHC Detection Kit	Roche, Applied Science, Mannheim
PureLink FFPE Total RNA Isolation Kit	Invitrogen, Carlsbad, CA, USA

Qiagen Multiplex PCR Plus Kit	QIAGEN, Hilden
QIAquick PCR Purification Kit	QIAGEN, Hilden
QuantiTect Virus Kit	QIAGEN, Hilden

2.6 Consumables

Conductive Filtered Tips (50 µL), QIAgility	QIAGEN, Hilden
Cover slips	Thermo Fisher Scientific, Dreieich
Disposable protective coats (Foliodress)	Hartmann, Heidenheim
Eppendorf tubes, safe-lock (1.5, 2 ml)	Eppendorf, Hamburg
Falcon-tubes (15, 50 ml)	Greiner Bio One, Frickenhausen
Filter tips (0.1-10 µl, 2-20 µl, 20-200 µl)	Starlab, Hamburg
Filter tips (200-1000 µl)	nerbe plus GmbH, Winsen
Jung Tissue Freezing Medium (Tissue Tec)	Leica Microsystems, Nussloch
LightCycler®480 Multiwell Plate 96, white	Roche, Applied Science, Mannheim
LightCycler®480 Sealing Foil	Roche, Applied Science, Mannheim
Litter plastic bags	nerbe plus GmbH, Winsen
Low-retention tubes	Kisker GbR, Steinfurt
Microscope slides, SuperFrost Plus	Thermo Fisher Scientific, Dreieich
Microtome Sectioning Blades, 819	Leica Microsystems, Nussloch
Microtome Sectioning Blades, C35 Feather	pfm medical, Köln
Multiscreen 96-well wash plates	Millipore, Bedford, MA, USA
Pipette tips for motor pipette	Biozym, Hessisch Oldendorf
Reagent Reservoir 4870 (50 ml)	Corning Incorporated, Corning, NY, USA
Sealing foil	Steinbrenner, Wiesenbach
SeroMap™ beads	Luminex Corporation, Austin, TX, USA

Tubes, graduated, flat-base, 5 ml	QIAGEN, Hilden
PCR plates, 96-well, skirted	Thermo Fisher Scientific, Dreieich

2.7 Laboratory devices

ABI PRISM 3730 genetic analyzer	Applied Biosystems, Foster City, CA, USA
Agarose gel electrophoresis chamber including gel tray and combs	Renner, Dannstadt
Centrifuge, Heraeus Fresco 17	Thermo Fisher Scientific, Dreieich
Centrifuge, Heraeus Megafuge 40R	Thermo Fisher Scientific, Dreieich
8-channel pipette 0.5-10 µl, 20-200 µl	Brand, Roskilde, Denmark
8-channel-motor pipette Precision®	Biozym Diagnostik, Hessisch Oldendorf
Cryostat	Leica Microsystems, Nussloch
Freezer -20°C, fridge 4°C	Liebherr, Bulle, Switzerland
Freezer -80°C	Forma Scientific, Ohio, USA
Gilson pipettes (2 µl, 10 µl, 20 µl, 200 µl 1000 µl)	Gilson International, Limburg
Ice machine	Hoshizaki, Amsterdam, The Netherlands
Incubator, Heraeus	Thermo Fisher Scientific, Dreieich
Light microscope	Leica Microsystems, Nussloch
Luminex 100/200 Analyzer	Luminex Corporation, Austin, TX, USA
MagNA Pure 96 (MP96) System	Roche, Applied Science, Mannheim
Mastercycler	Eppendorf, Hamburg
Micro wave	Sharp, Osaka, Japan
Mini centrifuge	neoLab, Heidelberg
Mini plate spinner	Woodbridge, NJ, USA
Mini-tube rotator	Fisher Scientific, Pittsburgh, PA, USA

NanoDrop 1000	Peqlab Biotechnologie GmbH, Erlangen
Pasteur pipettes	Sigma-Aldrich, Steinheim
pH-Meter	inoLab®, WTW, Weilheim
Pipetboy	Integra Biosciences, Fernwald
QIAgility	QIAGEN, Hilden
Shaker Labnet (at 37°C)	Woodbridge, NJ, USA
Shaker UNIMAX 1010	Heidolph, Schwabach
Thermomixer heating block	Eppendorf, Hamburg
Vacuum-wash station	Millipore, Bedford, MA, USA
VENTANA BenchMark ULTRA	Roche, Applied Science, Mannheim
Vortex Genie	Scientific Industries, Bohemia, NY, USA

2.8 Oligonucleotides

Primers and probes were purchased from Sigma-Aldrich (Hamburg). Sequences of primers and probes for the E6*I mRNA assay (Table 5) and the multiplex HPV genotyping assay (Table 6) are indicated in 5' to 3' direction. Sequences of primers and TaqMan probes used for the HPV DNA real-time quantitative PCR and HPV RNA pattern detection, as well as the probes for the multiplex HPV genotyping assay are not presented for reason of intellectual property protection.

Table 5: Primers and probes used in the E6*I mRNA assay

HPV	Forward primer	Backward primer	Probe
16	GTGTACTGCAAGCAACAGTTA	TCCAGATGTCTTTGCTTTTCTTCA	GCGACGTGAGGTGTATTAAC
18	TGTATATTGCAAGACAGTATT	GCTGGATTCAACGGTTTCTGG	ACTTACAGAGGTGCCTGCG
31	ATGAACTAAGATTGAATTGTG	TTCTTCTGGACACAACGGTCT	TACTGCAAAGGTGTATAACG
33	TACAGTGCGTGGAATGCAAA	TACGTCGGGACCTCCAACACG	ACGATCTGAGGGCGCTGTG
35	GTATACTGCAAACAAGAATTA	ACTGGACACAGCGGTTTTTGA	GCGGAGTGAGGTGTATTACA
52	GTGCAGTGCAAAAAAGAGCTA	CAGGTCGGGGTCTCCAACACT	ACGAAGAGAGGGCGCTGTTC
53	GTGTTCTGCAAGAAGGCATTG	TCTGGTGTCAACGGATGTTGA	AGCGTCAGAGGTGCTACAGA

Table 6: Primers used in the multiplex HPV genotyping assay

Target	Primer	Direction	Sequence
Cellular beta-globin	MS3	Forward	AATATATGTGTGCTTATTTG
	Bio-MS10	Backward	AGATTAGGGAAAGTATTAGA
HPV L1	GP5+	Forward	TTTGTTACTGTGGTAGATACTAC
	BSGP5+-2	Forward	TTTGTTACTGTTGTIGATACTAC
	BSGP5+-3	Forward	TTTGTTACTGTTGTIGATACCAC
	BSGP5+-4	Forward	TTTGTTACTTGTGTIGATACTAC
	BSGP5+-5	Forward	TTTTTAACTGTTGTIGATACTAC
	BSGP5+-6	Forward	TTTGTTACTGTGGTAGACACTAC
	BSGP5+-7	Forward	TTTGTTACAGTIGTAGACACTAC
	BSGP5+-8	Forward	TTTGTTACAGTIGTAGATACCAC
	BSGP5+-9	Forward	TTTGTTACTGTGGTAGATACCAC
	Bio-GP6+	Backward	GAAAAATAAACTGTAAATCATATTC
	Bio-BSGP6+-b	Backward	GAAAAATAAATTGTAAATCATACTC
	Bio-BSGP6+-c	Backward	TCTGGTGTCAACGGATGTTGA

Bio=biotinylated

2.9 Software

EndNote X7.7.1	Thomson Reuters, New York City, NY, USA
GraphPad Prism® 6	Graphpad Software Inc, La Jolla, CA, USA
LightCycler® 480	Roche, Applied Science, Mannheim
Luminex 100 IS 2.3 SP1 Software	Luminex Corporation, Austin, TX, USA
Microsoft Office	Microsoft Deutschland GmbH, München
Nucleotide Blast NCBI (https://blast.ncbi.nlm.nih.gov/)	National Center for Biotechnology Information, Bethesda, MD, USA
R Studio (https://www.rstudio.com/)	Boston, MA, USA
R Version 3.2.3 (https://www.r-project.org/)	R core team
SAS 9.4, SAS Enterprise Guide 6.1	SAS Institute Inc, Cary, NC, USA

2.10 Patients and clinical specimen

NSCCUP patients treated at the ENT (Ear-Nose-Throat) clinic of the University Hospital in Heidelberg were identified from the laboratory (head Prof. Dr. Jochen Hess) database (n=30, diagnosed 1990-2011) and from the Cancer Registry and Tissue Bank of the National Center of Tumor Diseases (NCT) (n=19, 2009-2011). Patients from Spain (1990-2012) and Northern Italy (1988-2014) were identified by the collaborators (Dr. Laia Alemany and Dr. Paolo Boscolo-Rizzo) from the databases of the Hospital de Sant Pau in Barcelona (n=59), Treviso Regional Hospital (n=46), Montebelluna Hospital (n=15) and Trieste Hospital (n=11).

Patients presenting with neck lymph node metastases were eligible for the retrospective NSCCUP study, if no primary tumor was found during initial diagnostic work-up. The general work-up included physical examination, sonography of the neck and abdomen, CT (recently also PET/CT) and/or MRI scan of the head and neck, CT scan of the thorax, and panendoscopy under general anaesthesia including microlaryngoscopy, endoscopy of the oropharynx, hypopharynx, epipharynx, esophagus and trachea and bronchia [88]. Biopsies were taken from the nasopharynx and the base of tongue, and bilateral tonsillectomy was performed for all cases from Heidelberg and for most cases from the other centers. During follow-up examinations, in 13 cases a potential primary tumor was detected within three years after initial diagnosis.

FFPE tissues from lymph node metastases were available from 161 patients (Table 7). Among them were eleven patients from Heidelberg with additional fresh-frozen biopsies stored at -80°C. From 19 patients from Heidelberg only frozen biopsies were available.

NSCCUP patients from Leipzig (n=46) with at least one serum and up to 12 serial serum samples were identified by the collaborators from the University Hospital database (Table 7). FFPE blocks from 28 of those patients were available from the Pathology archive (Table 7). Due to limited access, all available blocks from HPV-seropositive NSCCUP (n=10) were selected and the maximum number of HPV-negative NSCCUP (n=18), but sectioning and molecular HPV analysis was performed unaware of the HPV serostatus.

From a previous HPV16 integration study performed by Dr. Dana Holzinger and Prof. Dr. Elisabeth Schwarz [Holzinger et al. in preparation] all patients with HPV16-driven OPSCC patients and lymph node metastasis at the time of diagnosis were selected in order to assess pairs of primary tumors (n=12, of which eight were re-analyzed for HPV markers including four that were re-analyzed for HPV integration) and metastases (n=20, 1-4 per patient). From those, either frozen biopsies (n=19) or FFPE tissues (n=9) were provided by the NCT Tissue Bank (Table 7).

Table 7: Clinical specimen included in this thesis.

Study center	Number	Specimen	Patients	Diagnosis
Heidelberg, Germany	30	FFPE tissues	NSCCUP	1997-2011
Heidelberg, Germany	30	Frozen biopsies	NSCCUP	1990-2011
Barcelona, Spain	59	FFPE tissues	NSCCUP	1990-2012
Treviso, Italy	46	FFPE tissues	NSCCUP	1988-2012
Montebelluna, Italy	15	FFPE tissues	NSCCUP	2003-2014
Trieste, Italy	11	FFPE tissues	NSCCUP	2003-2014
Leipzig, Germany	46	Serum	NSCCUP	2008-2016
Leipzig, Germany	28	FFPE tissues	NSCCUP	2008-2016
Heidelberg, Germany	19	Frozen biopsies (P/M)	OPSCC	1991-2008
Heidelberg, Germany	9	FFPE tissues (M)	OPSCC	1997-2008

FFPE=formalin-fixed paraffin-embedded, P=primary tumors, M=metastases, NSCCUP=neck squamous cell carcinoma from unknown primary, OPSCC=oropharyngeal squamous cell carcinoma

2.11 Ethics statement

All patients gave informed consent. The studies were approved by the local ethics committees (ethic votes: 421/AULSS9 in Treviso; PR077/11 in Barcelona; 201-10-12072010 and 202-10-12072010 in Leipzig; NCT Tissue Bank project approvals in Heidelberg: 1426, 1427, 1811, 1812, 1874, 1901, 2016, 2178 and 2200).

3 Methods

3.1 Tissue sectioning

FFPE tissue blocks were sectioned with 5 µm thickness according to the previously described sandwich method [49, 88, 131]. The first and the last section were stained by hematoxylin and eosin to verify the tumor content. In between, tissue ribbons were prepared for DNA and RNA extraction, followed by one section for $p16^{INK4a}$ staining and additional sections. In order to avoid cross-contamination between consecutively sectioned tissues, blades were used for only one tissue block, and before and after each block the microtome and all instruments were extensively cleaned with acetone, 70% ethanol and RNase AWAY spray (Thermo Fisher Scientific, Dreieich). Potential contamination was monitored by sectioning healthy mouse brain tissue at least once per day, which should be free of human beta-globin DNA and ubiquitin C mRNA in subsequent analysis.

Tissues from all study centers were processed according to this protocol. Tissue blocks from the University Hospital Leipzig were provided by the Institute of Pathology for in-house sectioning done by me in Leipzig. Samples from Heidelberg were sectioned by collaborators at the NCT Tissue Bank in accordance with their regulations. For samples from Barcelona, sectioning and DNA extraction were performed by collaborators at the Catalan Institute of Oncology (ICO, Barcelona, Spain).

From frozen biopsies, sections of 16 µm thickness yielding around 5 mg of tissue were cut, homogenized with pistils in liquid nitrogen and stored at -80°C prior extraction [49, 88]. As described for FFPE tissue, the first and the last section were prepared for hematoxylin and eosin staining. New blades and gloves were used for every biopsy. The cryostat was extensively cleaned with acetone, and mouse biopsies were included as negative controls. Sections for RNA extraction were prepared at a later time point using the same protocol, but only for HPV DNA-positive frozen biopsies (n=2) and few HPV DNA-negative controls (n=5).

3.2 Hematoxylin and eosin staining

Presence of tumor cells in the tissue was validated by hematoxylin and eosin (HE) staining. Tissue sections were stained for 10 min in Mayer's Haemalaun Solution (AppliChem GmbH, Darmstadt), rinsed with tap water for 10 min, and counterstained for 10 min in eosin (Carl Roth, Karlsruhe). After a short rinse with aqua bidest, sections were dehydrated for short time in 70%, 96% and 100% ethanol, and cleared in xylol before being covered with a cover slip using Eukitt mounting medium (O. Kindler GmbH, Freiburg).

FFPE tissue sections needed to be deparaffinized by short incubation in xylol and rehydration in 100%, 96% and 70% ethanol prior staining with Mayer's Haemalaun Solution. Tumor content was determined under the microscope. A subset and all ambiguous sections were evaluated by a pathologist, in addition. Cases with less than 10% tumor content were excluded.

3.3 Nucleic acid extraction

3.3.1 DNA extraction

Genomic DNA was released from FFPE tissue sections by incubation for 16 hours at 56°C in 200 µl proteinase K (1 mg/ml, Roche, Applied Science, Mannheim) solution consisting of 45 mM Tris-HCl, 0.9 mM EDTA and 0.45% Tween 20 in DNase/RNase-free water [88]. Samples were then incubated for 10 min at 72°C to inactivate the enzyme and were centrifuged to separate the paraffin layer from the aqueous phase containing the DNA. In order to optimize the quality of DNA extracted from FFPE by reversing potential formalin-induced cross-linking, an aliquot of extracted DNA was subsequently incubated for 20 min at 90°C, followed by centrifugation and transfer of the aqueous phase [88].

From frozen biopsies, DNA was extracted by incubating the sections in 200 µl MagNA Pure DNA Tissue Lysis Buffer (Roche) supplemented with proteinase K (0.5 mg/ml) for 3 hours at 55°C. The MagNA Pure 96 system (MP96, Roche) was then used to extract and elute DNA in 200 µl with the MP96 DNA and Viral NA Large Volume Kit and the protocol 'DNA tissue Large Volume 2.0' (Roche). Extracted DNA was stored at -20°C prior further analysis.

3.3.2 RNA extraction

RNA was extracted from FFPE tissue sections using the Pure-Link FFPE Total RNA Isolation Kit (Invitrogen, Carlsbad, CA) [88]. In addition to the manufacturer's instructions, DNase (Qiagen, Hilden) was added to the columns and incubated for 15 min prior to elution in 50 µl DNase/RNase-free water.

For RNA release from frozen biopsies, sections were lyzed in 300 µl MagNA Pure LC RNA Isolation Tissue Lysis Buffer (Roche) [88, Supplements]. Total RNA was extracted from 200 µl of lysate supernatant and was eluted in 50 µl using the MP96 with the MP96 Cellular RNA Large Volume Kit and the protocol 'Cellular RNA Large Volume' (Roche). Extracted RNA was stored at -20°C prior further analysis.

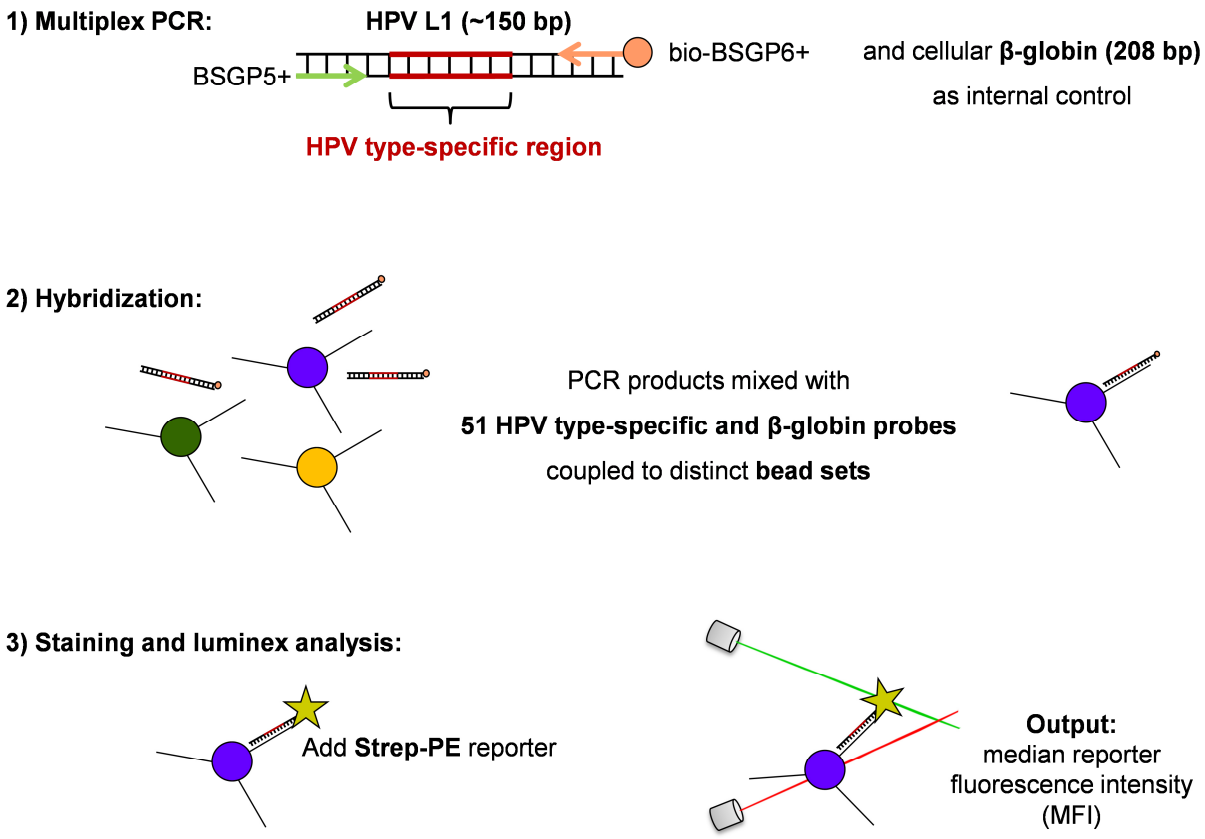


Figure 6: Detection of HPV DNA by multiplex HPV genotyping. The multiplex HPV genotyping (MPG) assay allows amplification of HPV L1 DNA from 51 mucosal HPV types by polymerase chain reaction (PCR) using broad-spectrum general-primers (BSGP), followed by hybridization to HPV type-specific probes coupled to distinct fluorescently-labeled beads. Antisense 5'-biotinylated primers (bio-BSGP6+) allow detection of PCR products by adding Streptavidin-R-Phycoerythrin (Strep-PE), which is measured by the green laser in the Luminex analyzer. The red laser detects the HPV type-specific bead set. Adapted from [132].

3.4 HPV DNA detection

3.4.1 Multiplex HPV genotyping (MPG)

HPV DNA was amplified by PCR and detected by hybridization to probes coupled to beads (Luminex Corp., Austin, Texas) as previously described [88, 132-134]. For the PCR, 5 µl of extracted DNA was used with the Multiplex PCR Kit (Qiagen, Hilden) and a broad-spectrum general-primer (BSGP5+/6+) mix that generates HPV L1 amplimers of ~150 bp and allows simultaneous amplification of 51 mucosal HPV types (Figure 6) [133]. PCR products were hybridized to HPV type-specific probes coupled to distinct fluorescently-labeled bead sets. After staining with Streptavidin-R-Phycoerythrin (Strep-PE, Invitrogen), bound PCR products were detected with the green laser of the Luminex analyzer, while the bead set representing the HPV type was detected with the red laser (Figure 6). The net median reporter fluorescence intensity (MFI) was calculated for each set after measuring at least 100 beads

per set. Samples with net MFI >5 for any HPV type were defined HPV DNA-positive for that type [134]. Human beta-globin was co-amplified to validate DNA integrity. Samples were invalid, if neither HPV nor beta-globin was detected.

3.4.2 HPV16 DNA real-time quantitative PCR

A real-time quantitative PCR (qPCR) was used to assess the quality of extracted DNA by the number of detectable cell equivalents (human beta-globin copies) and to determine the viral load (HPV copies per cell). In a total volume of 10 µl, 1 µl of extracted DNA was mixed with beta-globin and HPV16 E6 primers (0.5 µM), probes (0.1-0.2 µM) and 5 µl of LightCycler® 480 Probes Master (Roche) and PCR-grade water. Quantification was done based on beta-globin and HPV16 standard curves. Two standards were used in 10-fold dilutions: 0.1-100 ng/µl human placenta DNA (for beta-globin) and $1\text{-}10^6$ HPV16 copies diluted in 50 ng/µl human placenta DNA.

3.5 HPV RNA detection

3.5.1 E6*I mRNA detection

E6*I mRNA of HPV was detected by reverse transcription PCR (RT-PCR) using the QuantiTect Virus Kit (Qiagen) followed by hybridization to probes coupled to beads as previously described [48, 88]. This assay is available for twelve carcinogenic (HPV16, 18, 31, 33, 35, 39, 45, 51, 52, 56, 58 and 59; IARC Group 1 carcinogens) and eight probably/possibly carcinogenic HPV types (HPV26, 53, 66, 67, 68, 70, 73 and 82; IARC Group 2A/B carcinogens) [135].

For the RT-PCR, 1 µl of RNA was used. Short HPV E6*I (~65 bp) amplicons were amplified with HPV type-specific primers, and cellular ubiquitin C (85 bp) was co-amplified to assess RNA integrity. Samples were valid if positive for HPV E6*I mRNA and/or ubiquitin C. HPV DNA-positive samples were tested for mRNA of the HPV types detected by MPG. In addition, all FFPE samples were tested for ubiquitin C and HPV16 mRNA and DNA-invalid samples with high $p16^{INK4a}$ expression were additionally tested for HPV18, HPV33 and HPV35 mRNA [88].

3.5.2 HPV RNA pattern detection

In frozen biopsies, three spliced HPV16 transcripts (E6*I, E1^E4 and E1C) and the cellular ubiquitin C transcript were assessed by reverse transcription quantitative PCR analysis (RT-qPCR) [136]. A triplex RT-qPCR was used to quantify E6*I, E1^E4 and ubiquitin C, while the low abundant E1C transcript was quantified in a separate RT-qPCR. The two assays were performed using 2 µl (or 4 µl for E1C) of RNA with the LightCycler® 480 RNA Master

Hydrolysis Probe Kit (Roche) as previously described [136]. Samples were measured in duplicates.

Copy numbers per RT-qPCR were calculated based on standard curves obtained from 10-fold dilution of in vitro-generated transcripts (1-10^6 copies). Samples were classified as HPV16 RNA pattern-positive, if E6*I was increased relative to E1^E4 (E6*I/E1^E4 ratio>0.095) and/or if E1C was present [136].

3.6 $p16^{INK4a}$ immunohistochemistry

The cellular protein $p16^{INK4a}$ was detected in FFPE sections by IHC using the mouse anti-human $p16^{INK4a}$ antibody G175–405 (BD Pharmingen, San Jose, CA, USA) with the VENTANA BenchMark ULTRA system and the OptiView DAB IHC Detection Kit (Roche). Staining was evaluated unaware of the HPV status. Results were compared with an independent evaluator (Dr. Dana Holzinger) and discordant cases were discussed. $p16^{INK4a}$ overexpression was defined by moderate to strong diffuse nuclear and cytoplasmic staining of >25% of tumor cells in the section [88, 137]. For additional assessment of the commonly used 70% cut-off value, cases with $p16^{INK4a}$ expression in >25% but ≤70% of tumor cells were identified.

As internal staining control, background $p16^{INK4a}$ expression in non-cancerous cells present in lymph nodes, such as lymphocytes, was evaluated in $p16^{INK4a}$-negative cases. In case of no visible expression (n=5), $p16^{INK4a}$ staining was repeated and Ki-67 staining was performed on additional sections. Cases without Ki-67 expression (n=3) were invalid.

3.7 HPV16/18 E6 oncoprotein detection

The commercially available OncoE6™ Oral Test (Arbor Vita Corporation, Fremont, CA, USA) was used to detect the HPV E6 oncoprotein in HPV-driven frozen biopsies. This assay is a lateral diffusion immunoassay with E6 protein in lysates captured by a monoclonal antibody fixed to the strip and followed by staining. Sections were prepared as described above. According to the manufacturer's instructions, the homogenized and conditioned tissue lysate was loaded on a test strip. After washing and developing, a purple internal control line appeared on the test strip, and two additional lines only appeared, if E6 from HPV16 and/or HPV18 was present in the tissue. Intensity of the bands was evaluated by comparing with a reading guide showing bands with signal strength ranging from 1 to 5. As negative control, a lysate from a non-HPV-driven tonsillar carcinoma was used.

3.8 HPV Serology

HPV antibodies were detected in serum samples by bead-based multiplex serology. As previously described, HPV antigens were expressed in *Escherichia coli* as glutathione S-transferase (GST) fusion proteins to be coupled to glutathione-casein beads [138]. Each HPV antigen was coupled to a distinct fluorescently-labeled bead set. The bead mixture was combined with diluted serum to allow binding of HPV antibodies present in the serum. All sera were diluted 1:100 (first experiment), while serial serum samples from HPV-seropositive patients were additionally tested in 1:1,000 and 1:10,000 dilution (second experiment).

Bound antibodies were detected using a biotinylated anti-human immunoglobulin G antibody. After staining with Strep-PE, MFI values were determined for each antigen in each sample by measuring at least 100 beads per set with the Luminex analyzer [138]. Presence or absence of antibodies was defined based on antigen-specific cut-off values [55] (Table 8). MFI values >1000 were referred to as high antibody levels. Besides HPV antigens, also major capsid proteins VP1 from human polyomaviruses (BK virus, JC virus, Merkel cell polyomavirus) were included (Table 8).

Table 8: Antigens included in serology assay.

Organism	Virus	Antigens	Cut-off values 1:100 (MFI)
Papillomavirus	HPV16	E6, E7, E1, E2, E4**, L1	1000, 548, 200, 679, 876, 422
	HPV18	E6, E7, E1, E2, L1	243, 789, 200, 600, 394
	HPV31	E6, E7, L1	890, 200, 712
	HPV33	E6, E7, L1	253, 500, 515
	HPV35	E6, E7, L1	250, 500, 552
	HPV45	E6*, E7*, L1*	249, 200, 368
	HPV52	E6*, E7*, L1*	271, 200, 547
	HPV58	E6*, E7*, L1*	250, 200, 371
Polyomavirus	BK virus (BKV)	VP1	400/250 (1:100/1000)
	JC virus (JCV)	VP1**	250 (1:1000)
	Merkel cell polyomavirus (MCV)	VP1**	250 (1:1000)

*included only in first experiment, **included only in second experiment (HPV-seropositive patients)

Single sera of patients with serial samples were excluded, if antibody patterns were evidently different from the other sera of the same patient. Single nucleotype polymorphism (SNP)-genotyping [139] was performed to compare the sequence of the DNA extracted from the doubted serum with the DNA from tissue, if available, or from other serum samples of the

same patient. DNA was extracted from 40-50 µl of serum filled up with PBS to 500 µl using the MP96 (Roche) with the MagNA Pure 96 DNA and Viral NA Large Volume Kit and the protocol 'Viral NA Universal Large Volume' (Roche). DNA was eluted in 50 µl and sent for analysis (Multiplex Cell Line Authentication, service offered by Multiplexion, Heidelberg). In case of less than 90% DNA sequence identity, samples were assumed to be from different donors.

3.9 HPV16 integration analysis

3.9.1 TEN16

HPV integration sites were identified in frozen biopsies from HPV16-driven OPSCC primary tumors (n=5) and lymph node metastases (n=14) by using the TEN16 (Tagging, Enrichment and Next-generation sequencing of HPV16) strategy as previously described [12]. This analysis was performed in collaboration with Prof. Dr. Elisabeth Schwarz (DKFZ). For tagging with ten distinct barcodes, nine pools with two DNA samples (25 ng each) and one single DNA sample were prepared. DNA fragmentation and adaptor tagging was done by Nextera reaction (Illumina, San Diego, CA, USA) and was followed by purification of fragmented DNA and blocking of free 3'-OH termini (Figure 7). HPV16 DNA was enriched by Multiplex PCR using an HPV forward primer mix and a barcoded Nextera adapter primer, followed by pooling of all reactions into a single DNA sample and DNA purification (Figure 7). Prior sequencing, a fragment size selection was performed by loading the sample on an agarose gel with the bromphenol blue (BPB) marker and cutting the part, where fragments with a size between 200 bp and 500 bp would be expected. DNA was extracted from the cut gel, followed by ethanol precipitation to reach a final DNA concentration of 2 mg in 40 µl. Sequencing was done by the High Throughput Sequencing Unit (DKFZ) on an Illumina HiSeq2000 system.

Analysis of sequencing data included sorting into barcode groups and alignment against the HPV16R genome (NC_001526.3) and the human genome GRCh38 reference sequence using the tools provided by the National Center for Biotechnology Information (NCBI). This resulted in a list of candidates for viral-cellular junctions (VCJs), which were then validated by PCR and gel electrophoresis.

DNA samples from sections with less than 10% tumor content had to be excluded later, as well as cases, in which SNP-genotyping [139] revealed different sequences in the metastasis versus the primary tumor, indicating that the tissues were not from the same patient (Multiplex Cell Line Authentication, service offered by Multiplexion). Thus, one primary tumor and three metastases were excluded.

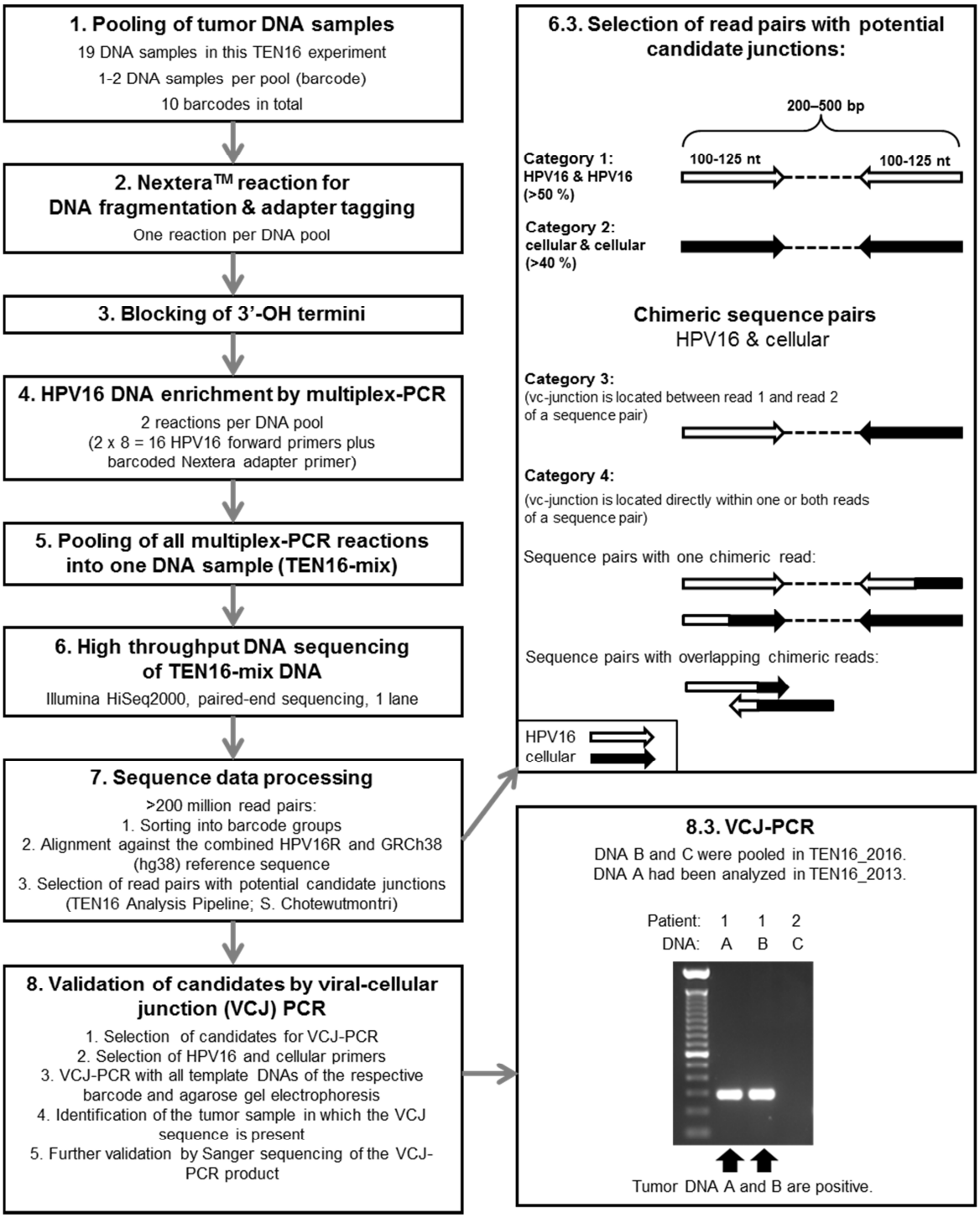


Figure 7: TEN16 work-flow. The TEN16 (Tagging, Enrichment and Next-generation sequencing of HPV16) strategy allows simultaneous detection of HPV16 integration sites in multiple samples. Pooled DNA (1) is fragmented and tagged with a Nextera adapter (2). Enrichment of HPV16 DNA is done by PCR using a mixture of the Nextera adapter primer and HPV primers covering the genes E1, E2 and E5 (4). Sequencing of a single pooled DNA sample (5, 6) and subsequent data processing (7) revealed four categories of potential candidate junctions: both reads contain only HPV sequences (category 1), or only cellular sequences (category 2), or a read pair consisting of one HPV sequence and one cellular sequence (category 3), or the viral-cellular junction located within one or both reads (category 4). Candidates from category 3 and 4 were selected for validation by PCR (8). Primary tumors analyzed in this study (TEN16_2016), were already included in a previous study (TEN16_2013). Adapted from [12].

3.9.2 Validation PCR

All VCJ candidates with at least 20 read pairs were validated by viral-cellular junction PCR (VCJ-PCR). Cellular and HPV primers were designed using the HUSAR sequence analysis package (DKFZ). For each barcode, both pooled DNA samples were analyzed for all detected candidate VCJs (Figure 7). Purified PCR products were sent for sequencing (GATC, Konstanz). Positive VCJs were validated by VCJ-PCR in all DNA samples from the patient, including the DNA extracted from the primary tumor, from all metastases and the DNA prepared for the previous study (TEN16_2013).

3.10 Methylation analysis

DNA methylation in the promoters of the five genes ALDH1, OSR2, GATA4, GRIA4 and IRX4 was quantified by PCR using bisulfite converted DNA, followed by matrix-assisted laser desorption/ionization time-of-flight mass spectrometry (MassARRAY) as previously described [76]. This analysis was performed in collaboration with Dr. Dieter Weichenhan (DKFZ). In order to allow application of this method to FFPE material, new primers were designed to shorten the amplicons, but still cover the informative CpG units of the five gene promoters of interest. One amplicon each was designed for OSR2 (174 bp), GRIA4 (180 bp) and IRX4 (158 bp) and two amplicons each for ALDH1 (119 bp and 155 bp, respectively) and GATA4 (131 bp and 126 bp, respectively).

Bisulfite conversion of genomic DNA was performed using the EZ DNA Methylation Kit (Zymo Research, Freiburg) according to the manufacturer's instructions. All DNA samples with sufficient quality, implying that beta-globin was detectable by MPG, and with enough volume left were included. As starting material 500-1000 ng of DNA (quantified by PicoGreen) were used and were eluted in 2x 30 µl M-Elution Buffer. In case of lower DNA concentration, samples were eluted in 2x 15 µl (300-500 ng DNA) or 2x 10 µl (<300 ng). The five gene promoters of interest were amplified by PCR and methylation was assessed by MassARRAY.

Percent methylation values were obtained for all CpG units. If methylation levels were not measurable in more than 66% of the CpG units in one gene, this gene was not considered in this particular sample. For each gene, the mean value of all CpG units was calculated, which was either below (low) or above (high) the pre-defined cut-off value (ALDH1: 22.7%, OSR2: 7.8%, GATA4: 17%, GRIA4: 16.4%, IRX4: 12%) [76]. A methylation score (MS, range 0-5) was determined for each patient by the number of gene promoters with a methylation level according to a previously identified pattern (ALDH1 low, OSR2 low, GATA4 high, GRIA4 high, IRX4 high methylation) [76].

3.11 TP53 sequencing

TP53 mutations were searched by analyzing exons 4 to 10, the gene region where the majority of mutations are localized [140], in collaboration with Dr. Roberta Bertorelle (Institute Oncologico Veneto, Padua, Italy). All DNA samples from HPV-driven biopsies and non-HPV-driven biopsies matched for gender, smoking status and age at diagnosis with ≥50% tumor content and with sufficient DNA quality (beta-globin detected by MPG) were included. DNA (100-200 ng) was amplified by PCR with exon-specific primers (0.2 µM, sequences from IARC protocol, 2010), 1.5-2.5 mM $MgCl_2$ and AmpliTaq Gold DNA polymerase (1 U/sample, Thermo Fisher Scientific). After initial denaturation at 95°C for 10 min, 40 cycles were run comprising a denaturation step (95°C for 1 min), an annealing step (exon-specific, 52-59°C for 90 sec) and an extension step (72°C for 90 sec), followed by final extension at 72°C for 7 min. PCR products (177-353 bp) were then sequenced by fluorescent capillary electrophoresis (ABI PRISM 3730 genetic analyzer, Applied Biosystems, Foster City, CA, USA) and sequences were compared with the NCBI reference sequence NC_000017.10. Mutations were confirmed by additional independent PCR and sequencing and were grouped into disruptive (stop mutations, frameshift mutations or non-conservative mutations occurring within the DNA binding domain) and non-disruptive mutations (all other mutations) as previously described [72, 75].

3.12 Statistics

Patient characteristics and follow-up data obtained from clinical charts were provided by the NCT Cancer Registry and collaborators in Heidelberg, Treviso, Barcelona and Leipzig. As previously described, overall survival (OS) time was calculated from the date of diagnosis to date of death (event) or end of follow-up (censored) [88]. Progression-free survival (PFS) time was calculated from the date of diagnosis to new lymph node or distant metastasis (event), malignancy in the head and neck region or outside (event), death (event) or end of follow-up without progression (censored). The Kaplan–Meier method was used to assess survival distributions. Curves were compared using log-rank tests.

To assess the effect of HPV on OS and PFS and adjust for possible confounders, multivariable Cox proportional hazard regression models were fitted in collaboration with Dr. Manuel Wiesenfarth (DKFZ). Besides HPV status, models included age at diagnosis (continuous), gender, N stage (categorized into 1/2a, 2b and 2c/3, according to the categories for OPSCC in the AJCC Cancer Staging Manual, 7th edition), extracapsular spread (presence vs. absence), alcohol and tobacco consumption (current or former vs. never) and treatment (multimodal vs. single). Only patients with complete clinical data could

be included. In the multicenter study, models were stratified by country and the proportional hazards assumption was met after modeling a time-dependent treatment effect. While multimodal treatment comprised postoperative radiotherapy, chemotherapy and chemo-radiotherapy in the multicenter study, only postoperative chemoradiotherapy was comprised in the serology study to meet the proportional hazards assumption. In the serology study, the Firth's penalized likelihood method was applied to mitigate convergence problems arising from the fact that none of the female patients experienced an event.

P values below 0.05 were considered statistically significant. For statistical analyses R version 3.2.3 and SAS version 9.4 was used. Sensitivity of a marker to identify HPV-driven NSCCUP (true positive rate) and specificity, as the probability for a negative result in non-HPV-driven NSCCUP (true negative rate), were calculated with 95% exact Clopper-Pearson confidence interval using an online tool (https://www.medcalc.org/calc/diagnostic_test.php). The degree of marker agreement was quantified by Cohen's kappa values that were calculated using a second online tool (http://www.graphpad.com/quickcalcs/kappa2/). Figures were generated using GraphPad Prism® version 6.

4 Results

4.1 Multicenter NSCCUP study

4.1.1 Study population

For retrospective HPV analysis, 180 NSCCUP patients from Heidelberg (n=49), Barcelona (n=59) and Treviso (n=72, including patients from Montebelluna and Trieste) were eligible. All patients presented with neck lymph node metastasis, but no primary tumor could be identified during initial diagnostic work-up. The median age at diagnosis was 62 years, ranging from 36 to 88 years, and there was strong male predominance (88%) (Table 9) [88]. Current or former tobacco (84%) and alcohol consumption (69%) were common, with 64% of patients consuming both tobacco and alcohol, but the proportion of tobacco and alcohol consumers decreased with increasing year of diagnosis (p=0.006 and 0.008, respectively, chi squared test for trend in proportion) [88]. The majority of patients presented with advanced nodal disease (73% stage N2b, N2c or N3) and extracapsular spread (71%). All cases received surgical treatment comprising neck dissection, except for two cases treated by primary radiotherapy. While surgery was the only treatment for 39 cases (22%), it was followed by chemotherapy in two cases, by radiotherapy in 52 cases (30%) and by combined chemoradiotherapy in 79 cases (45%).

The median year of diagnosis was 2004 for all 180 patients and for all patients from Treviso, while patients from Barcelona were diagnosed earlier (median: 2002) and patients from Heidelberg more recently (median: 2008). Patients from Barcelona presented with a significantly higher N stage (39% N2c/3 vs. 26%, p=0.006) and were more frequently treated by chemoradiotherapy (81% vs. 41% in Heidelberg and 19% in Treviso, p<0.0001, Pearson chi squared test, Table 9). Regarding age, gender, extracapsular spread, tobacco and alcohol consumption, no significant differences between the three study centers were observed (Table 9).

4.1.2 Prevalence of HPV-driven NSCCUP

Two molecular HPV markers (HPV DNA and E6*I mRNA) and the surrogate marker $p16^{INK4a}$ were analyzed to assess the HPV status. HPV-driven cases were defined by presence of HPV mRNA in combination with at least one additional marker, meaning HPV DNA and/or $p16^{INK4a}$ overexpression. Based on this algorithm, 28 (15.6%) of 180 NSCCUP were HPV-driven [88]. The prevalence of HPV-driven NSCCUP varied from 10.2% (6/59) in Barcelona to 16.7% (12/72) in Treviso and up to 20.4% (10/49) in Heidelberg and increased significantly with year of diagnosis (p=0.004, chi squared test for trend in proportion), in total from 8.5%

(early period: 1988-2004, n=93) to 23.0% (late period: 2005-2014, n=87; Figure 8) [88]. This increase was present in all three centers, although to a different extent: while in Barcelona only a small increase from 8.3% to 13.0% was observed, the HPV-driven proportion approximately tripled in Heidelberg (10.0% to 28.0%) and Treviso (8.1% to 25.7%).

Table 9: Characteristics of NSCCUP patients from Heidelberg, Barcelona and Treviso [88].

	Heidelberg (n=49)	Barcelona (n=59)	Treviso (n=72)	Total (n=180)	p†
Gender					
Male	43 (88%)	52 (88%)	63 (88%)	**158 (88%)**	1.0
Female	6 (12%)	7 (12%)	9 (13%)	**22 (12%)**	
Age, range (median)	43-85 (61)	36-81 (61)	39-88 (63)	**36-88 (62)**	0.7
Sampling period					
1988-2004	20 (41%)	36 (61%)	37 (51%)	**93 (52%)**	0.1
2005-2014	29 (59%)	23 (39%)	35 (49%)	**87 (48%)**	
Tobacco					
Ever	33 (73%)	52 (88%)	63 (88%)	**148 (84%)**	0.07
Never	12 (27%)	7 (12%)	9 (13%)	**28 (16%)**	
N/A	4			4	
Alcohol					
Ever	30 (68%)	43 (73%)	48 (67%)	**121 (69%)**	0.7
Never	14 (32%)	16 (27%)	24 (33%)	**54 (31%)**	
N/A	5			5	
N stage					
1, 2a	21 (45%)	12 (20%)	14 (19%)	**47 (26%)**	**0.006**
2b	14 (30%)	24 (41%)	39 (54%)	**77 (43%)**	
2c, 3	12 (26%)	23 (39%)	19 (26%)	**54 (30%)**	
N/A	2			2	
Extracapsular spread					
Yes	27 (63%)	43 (75%)	51 (72%)	**121 (71%)**	0.4
No	16 (37%)	14 (25%)	20 (28%)	**50 (29%)**	
N/A	6	2	1	9	
Treatment					
Surgery only	8 (18%)	10 (17%)	21 (29%)	**39 (22%)**	**<0.001**
+RT	15 (34%)	0	37 (51%)	**52 (30%)**	
+CRT	18 (41%)	47 (81%)	14 (19%)	**79 (45%)**	
Other*	3 (7%)	1 (2%)	0	**4 (2%)**	
N/A	5	1		6	
Median follow-up, years	1.7	3.3	2.9	**2.5**	

N/A=not available, +RT=postoperative radiotherapy, +CRT=postoperative chemoradiotherapy, *including radiotherapy only (n=2) and postoperative chemotherapy (n=2),
†Pearson chi squared test (age categories: ≥62 vs. <62), Fisher's exact test for treatment
Statistically significant values are displayed in bold.

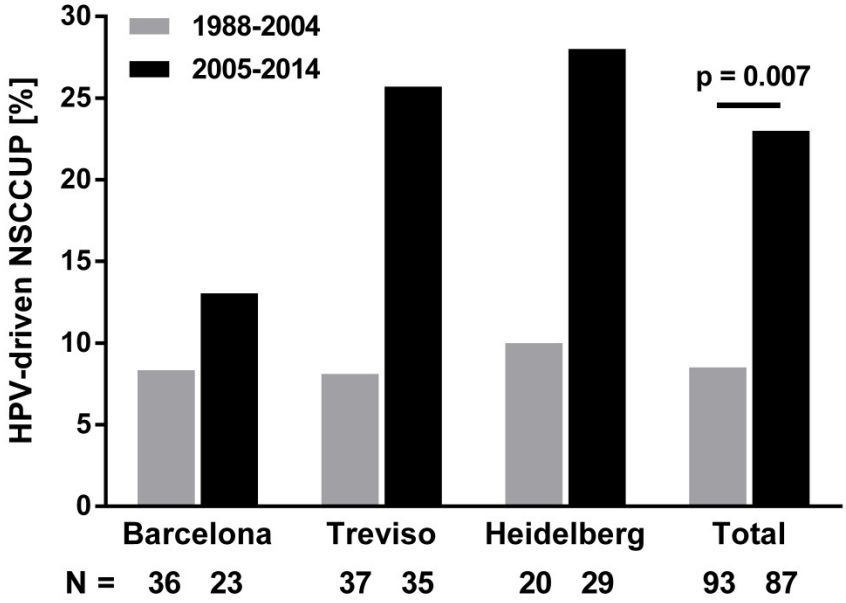


Figure 8: Time trend of HPV-driven NSCCUP. The proportion of HPV-driven NSCCUP diagnosed in the early period 1988-2004 (grey) is compared to the late period 2005-2014 (black) for all 180 cases and by center. The number of cases (N) analyzed in each period is depicted below the graph. Pearson chi squared test for trend in proportion was used to calculate the p value. Published in [88].

HPV16 was by far the most common HPV type, since 25 (89%) of 28 HPV-driven NSCCUP were driven by HPV16 [88]. The remaining three cases were driven by HPV18, HPV33 and HPV35, respectively (4% each). DNA from HPV90 was detected in one case. Since HPV90 is not known to be a (probable) high-risk type, no assay was available for detecting HPV90 E6*I mRNA. But this case did not show $p16^{INK4a}$ overexpression and was thus considered to be non-HPV-driven. In another non-HPV-driven case, DNA from both HPV52 and HPV53 was detected, but no mRNA of those types and no $p16^{INK4a}$ overexpression was detected.

4.1.3 HPV detection and marker concordance

Based on the defined algorithm, all 28 HPV-driven NSCCUP were positive for HPV mRNA. In addition, DNA of the same HPV type was detected in all cases. $P16^{INK4a}$ overexpression was observed in 25 of the 26 HPV-driven NSCCUP with FFPE tissue available for IHC (Figure 10, 11). One case positive for both HPV16 DNA and mRNA showed no $p16^{INK4a}$ staining signal in the tumor islands, while faint signals were detectable in the surrounding non-cancerous cells.

In the two HPV16-driven frozen biopsies, $p16^{INK4a}$ expression could not be evaluated, but transformation-specific HPV16 RNA patterns (E6*I/E1^E4 ratio>0.095 and E1C) were present. In addition, the HPV16 E6 protein could be detected in lysates from these two biopsies with intermediate to high signal strength (2 and 4, respectively, on a scale ranging from 0 to 5, Appendix Figure A-1) using the commercially available OncoE6™ Oral Test (Arbor Vita Corporation, Fremont, CA, USA).

While DNA extracted from frozen biopsies was of sufficient quality, extraction from FFPE tissue was challenging, as 102/161 (63%) FFPE samples were DNA-invalid (beta-globin-negative by MPG) after standard extraction. DNA extraction was improved by incubating the extracted DNA for 20 min at 90°C, which resulted in an increased detection of beta-globin copies by qPCR (median: 1235 vs. 467 copies per µl DNA, p=0.002, paired student's t-test; Figure 9). Re-testing DNA-invalid samples by MPG after this additional incubation reduced the total number of invalid samples to 58/161 (36%). Two cases were HPV16 DNA-positive after re-testing, of which one was positive but the other one was negative for HPV16 mRNA and $p16^{INK4a}$ overexpression, while the remaining 42 samples were HPV DNA-negative.

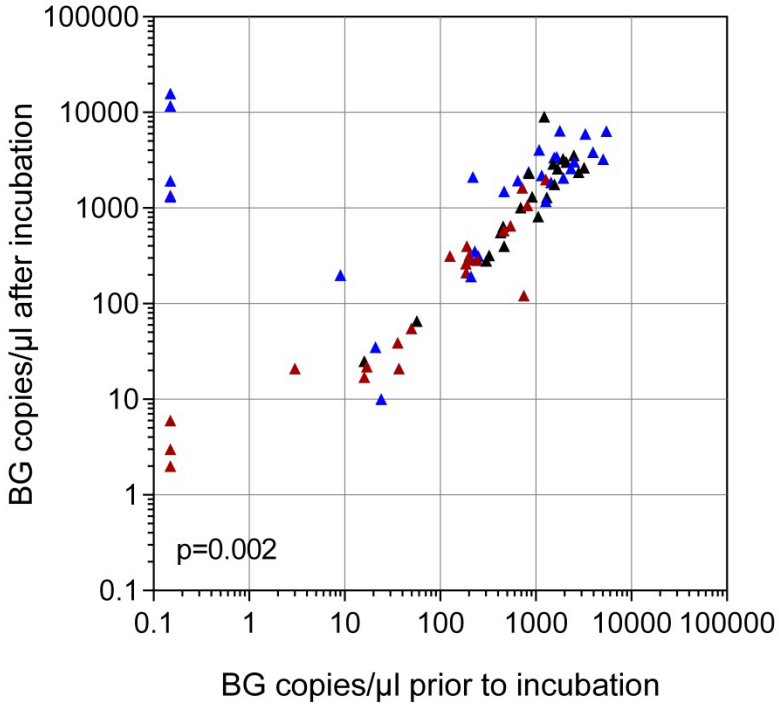


Figure 9: Beta-globin detection in DNA from FFPE tissue after standard vs. improved extraction. DNA extracted from FFPE tissue was tested for beta-globin (BG) by quantitative real-time PCR prior to and after additional incubation for 20 min at 90°C. BG-negative samples are set to 0.15 to be visible on log scale. DNA samples were tested by MPG with valid result (black), invalid result (red) or valid result only after incubation (blue). Paired student's t-test was used to calculate the p value.

The proportion of DNA-invalid cases was significantly increased among the cases diagnosed between 1988 and 2004 (44/93, 47%) compared to those diagnosed between 2005 and 2014 (14/87, 16%, p<0.0001, Pearson chi squared test). However, all 180 cases were mRNA-valid. Only three cases (diagnosed in 1993, 2002 and 2004, respectively) were $p16^{INK4a}$-invalid, because of no visible $p16^{INK4a}$ expression in surrounding non-cancerous cells and no signal in the subsequent Ki-67 staining. Overexpression of $p16^{INK4a}$ was equally common in HPV mRNA-negative cases that were DNA-negative (7/72, 10%) or DNA-invalid (6/55, 11%, p=0.8, Pearson chi squared test).

Among the 102 cases with valid result for HPV mRNA, DNA and $p16^{INK4a}$ expression, all three markers were concordantly positive or negative in 90 (88%) cases. The discordant cases comprised seven cases (7%) with only $p16^{INK4a}$ overexpression, three cases with only HPV DNA detected (3%), and a single case (1%) with HPV DNA and $p16^{INK4a}$ overexpression in the absence of HPV mRNA (Figure 10).

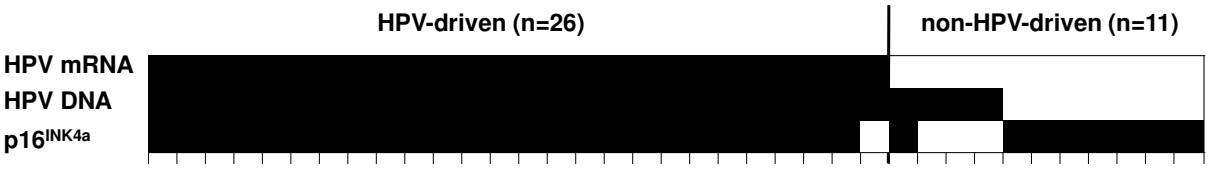


Figure 10: Marker distribution among the 37 cases positive for at least one marker with valid result for all markers. The 26 HPV-driven cases on the left are defined by presence (black box) of HPV mRNA with at least one additional marker. Not shown are the 65 cases negative for all three markers, DNA-invalid cases (n=58), and cases not tested for RNA (n=1) or $p16^{INK4a}$ (n=19, including 2 HPV-driven cases). Published in [88].

Table 10: HPV marker concordance in NSCCUP.

		HPV mRNA N		Sensitivity %	Specificity %	Cohen's Kappa
		+	-	(95% CI)	(95% CI)	(95% CI)
HPV DNA*	+	28	4	100 (88-100)	95 (88-98)	0.91 (0.82-1.00)
	-	0	76			
$p16^{INK4a}$ (>25%)	+	25	14	96 (80-100)	89 (83-94)	0.71 (0.58-0.85)
	-	1	117			
DNA/$p16^{INK4a}$ (>25%)**	+	25	1	96 (80-100)	99 (93-100)	0.95 (0.88-1.00)
	-	1	75			
$p16^{INK4a}$ (>70%)	+	21	9	81 (61-93)	93 (87-97)	0.70 (0.55-0.84)
	-	5	122			
DNA/$p16^{INK4a}$ (>70%)**	+	21	0	81 (61-93)	100 (95-100)	0.86 (0.74-0.98)
	-	5	76			

*only DNA-valid cases were considered, **positivity for both HPV DNA and $p16^{INK4a}$ overexpression
CI=confidence interval

When considering all cases with valid results for at least two markers, compared to HPV mRNA positivity as gold standard for HPV-driven cases, positivity for HPV DNA, $p16^{INK4a}$ overexpression and the combination of both had a sensitivity of 100%, 96% and 96%, and a specificity of 95%, 89% and 99%, respectively (Table 10). However, differences between the markers were not statistically significant, since 95% confidence intervals were overlapping (Table 10). The lowest concordance to HPV mRNA had $p16^{INK4a}$ overexpression (kappa=0.71). While in this study $p16^{INK4a}$ overexpression was defined by staining of more

than 25% of tumor cells (Figure 11), in other publications 70% has been applied as cut-off value. When this more stringent cut-off was explored, the sensitivity of $p16^{INK4a}$ overexpression to identify HPV mRNA-positive cases was lower and specificity was slightly higher (81% and 93% for the 70% cut-off value vs. 96% and 89% for the 25% cut-off value, respectively).

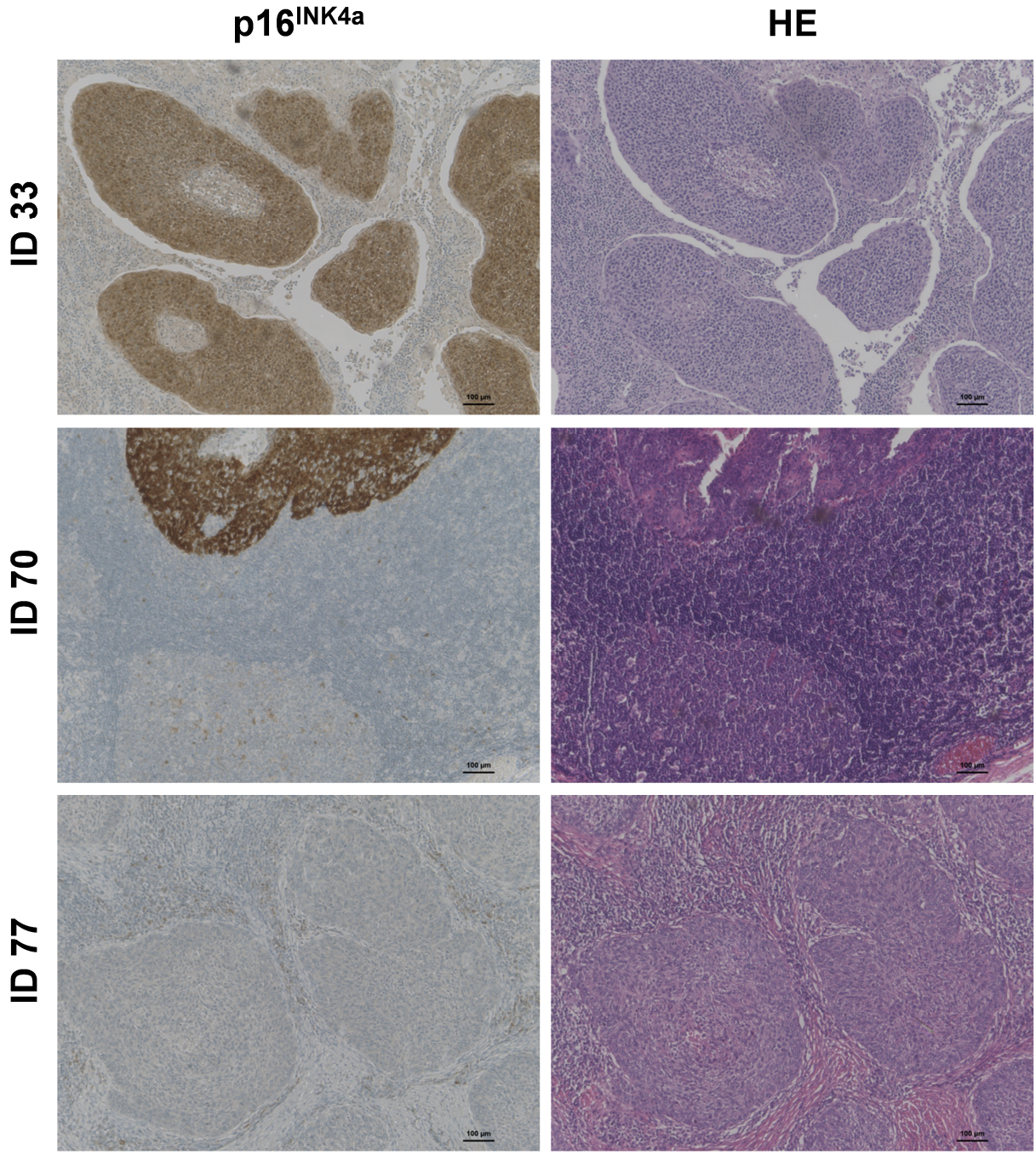


Figure 11: $p16^{INK4a}$ expression levels in FFPE tissue. FFPE tissue sections from three patients stained for $p16^{INK4a}$ (left column) and hematoxylin and eosin (HE, right column) are shown exemplarily for two HPV16 DNA-positive and mRNA-positive NSCCUP with $p16^{INK4a}$ expression in >70% (ID 33) or 25%-70% (ID 70) of tumor cells in comparison to an HPV mRNA-negative NSCCUP without $p16^{INK4a}$ expression (ID 77). Weak $p16^{INK4a}$ expression in surrounding non-cancerous cells served as staining control. Images were acquired with 10x magnification. The scale bars in the lower right corner represent 100 µm.

Table 11: Clinical parameters of NSCCUP patients in relation to HPV status [88].

	N	HPV-driven (n=28)	non-HPV-driven (n=152)	p†
Gender				
Male	158	21 (75%)	137 (90%)	**0.02**
Female	22	7 (25%)	15 (10%)	
Age, years (median)		39-79 (59)	36-88 (62)	0.2
Region				
Heidelberg	49	10 (36%)	39 (26%)	0.3
Treviso	72	12 (43%)	60 (39%)	
Barcelona	59	6 (21%)	53 (35%)	
Tobacco				
Ever	148	18 (64%)	130 (88%)	**0.002**
Never	28	10 (36%)	18 (12%)	
N/A	4		4	
Alcohol				
Ever	121	12 (44%)	109 (74%)	**0.003**
Never	54	15 (56%)	39 (26%)	
N/A	5	1	4	
N status				
1, 2a	47	12 (43%)	35 (23%)	**0.002**
2b	77	15 (54%)	62 (41%)	
2c, 3	54	1 (4%)	53 (35%)	
N/A	2		2	
Lymph node level				
II	36	9 (75%)	27 (47%)	0.07
I, III, IV, V	34	3 (25%)	31 (53%)	
N/A	110	16	94	
Extracapsular spread				
Yes	121	18 (64%)	103 (72%)	0.4
No	50	10 (36%)	40 (28%)	
N/A	9		9	
Treatment				
Surgery only	39	9 (32%)	30 (21%)	0.6
+RT	52	7 (25%)	45 (31%)	
+CRT	79	12 (43%)	67 (56%)	
Other*	4	0	4 (3%)	
N/A	6		6	

N/A=not available, +RT=postoperative radiotherapy, +CRT=postoperative chemoradiotherapy, *including radiotherapy only (n=2) and postoperative chemotherapy (n=2),
†Pearson chi squared test, student's t-test for age, Fisher's exact test for treatment
Statistically significant values are displayed in bold.

4.1.4 Risk factor assessment in HPV-driven vs. non-driven NSCCUP

Compared to patients with non-HPV-driven NSCCUP, those with HPV-driven NSCCUP were significantly more often females (25% vs. 10%, p=0.02), less frequently current or former tobacco (64% vs. 88%, p=0.002) and alcohol consumers (44% vs. 74%, p=0.003) and presented less frequently with far advanced nodal disease (4% vs. 35% N2c/3, p=0.002 Pearson chi squared test, Table 11) [88]. From 70 patients from Treviso the neck lymph node level was known, in which the lymph node metastasis was located. Most HPV-driven metastases were found in level II (9/12, 75%, p=0.07), while only 47% (27/58) of the non-HPV-driven metastases were found in level II (Table 11). In some patients, HPV-driven metastases were also found in level III (n=2) and IV (n=1). Regarding study center, age at diagnosis, extracapsular spread and treatment, no significant differences were observed between patients with HPV-driven and non-HPV-driven NSCCUP (Table 11).

4.1.5 Prognostic value of HPV

The impact of the HPV status on survival could be evaluated for 27 patients with HPV-driven and 137 with non-HPV-driven NSCCUP, for which all covariate information was available. Patients with HPV-driven NSCCUP had a significantly better overall and progression-free survival (p=0.002 and p=0.0006, respectively, log-rank test; Figure 12) [88]. Progression was reported for only eight (30%) of the 27 HPV-driven NSCCUP compared to 99 (72%) of the 137 non-HPV-driven NSCCUP, and only seven (26%) patients with HPV-driven NSCCUP died compared to 89 (65%) patients with non-HPV-driven NSCCUP.

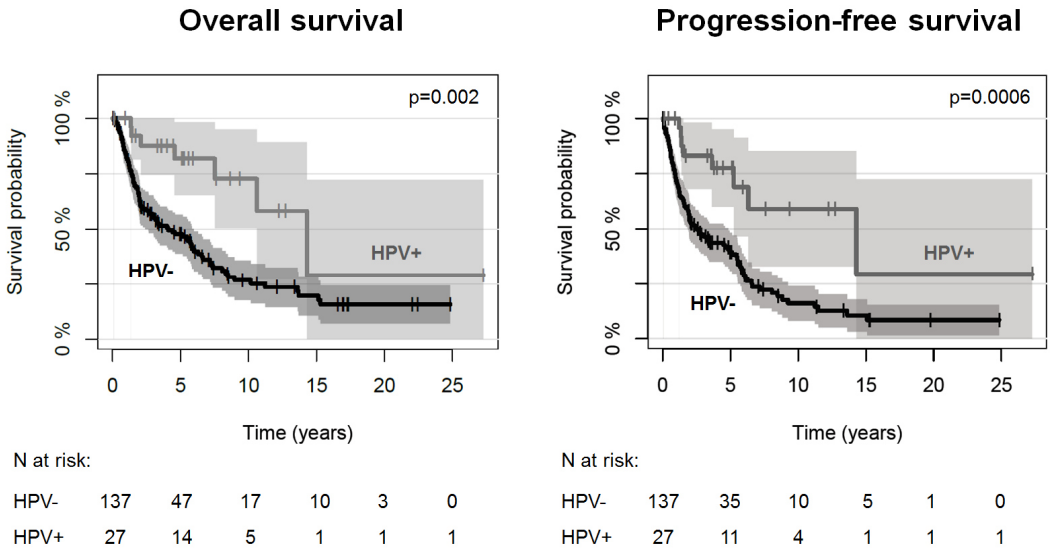


Figure 12: Survival of NSCCUP patients in relation to HPV status. Kaplan–Meier curves for overall survival (left) and progression-free survival (right) with 95% confidence interval (shaded areas) are compared for HPV-driven (HPV+, grey) versus non-HPV-driven (HPV-, black) patients. Only patients with all covariate information available were included (n=164). The numbers (N) at risk are depicted for each group below the graph. Log-rank test was used to calculate p values. Published in [88].

The prognostic role of HPV was confirmed by multivariable Cox regression analysis revealing significant hazard ratios of 0.30 for overall survival and 0.27 for progression-free survival (p=0.008 and p=0.003, respectively, Wald test; Table 12), when compared patients with HPV-driven versus non-HPV-driven NSCCUP [88]. Besides HPV status, also multimodal treatment was significantly associated with better progression-free survival, but a time-dependent effect of treatment needs to be considered. Unfavorable prognostic factors for both overall and progression-free survival were increased age, presence of extracapsular spread and advanced nodal stage with hazard ratios ranging from 1.03 to 2.69 (Table 12), while gender, tobacco and alcohol consumption had no significant effect on survival.

Table 12: HPV status and clinical parameters in relation to survival [88].

Multivariable analysis (n=164)*:	OS		PFS	
	HR (95% CI)	p†	HR (95% CI)	p†
HPV-driven (yes/no)	**0.30 (0.12 - 0.73)**	**0.008**	**0.27 (0.11 - 0.65)**	**0.003**
Gender (female/male)	0.71 (0.31 - 1.65)	0.4	0.48 (0.21 - 1.11)	0.09
Age at diagnosis	**1.03 (1.01 - 1.06)**	**0.003**	**1.03 (1.01 - 1.05)**	**0.01**
Tobacco (ever/never)	1.00 (0.48 - 2.10)	1.0	1.10 (0.53 - 2.26)	0.8
Alcohol (ever/never)	1.00 (0.57 - 1.75)	1.0	1.12 (0.66 - 1.91)	0.7
N stage				
2b vs. 1,2a	**2.02 (1.05 - 3.91)**	**0.04**	1.71 (0.93 - 3.14)	0.08
2c,3 vs. 1,2a	**2.31 (1.16 - 4.57)**	**0.02**	**2.69 (1.44 - 5.04)**	**0.002**
Extracapsular spread (yes/no)	**2.61 (1.41 - 4.85)**	**0.002**	**2.20 (1.26 - 3.83)**	**0.005**
Treatment (multimodal/single)	0.09 (0.01 - 1.46)	0.09	**0.12 (0.03 - 0.43)**	**0.001**
Treatment.Time**	2.18 (0.56 - 8.51)	0.3	**3.20 (1.25 - 8.18)**	**0.02**

OS=overall survival, PFS=progression-free survival, HR=hazard ratio, CI=confidence interval, single=surgery or radiotherapy only, multimodal=postoperative radiotherapy/chemotherapy/chemoradiotherapy
*stratified by country, **modelling of time-dependent treatment effect, †Wald test
Statistically significant values are displayed in bold.

4.1.6 HPV in patients with initial NSCCUP

In 13 patients, a squamous cell carcinoma was detected within three years after initial NSCCUP diagnosis that might represent the primary tumor (Table 13). In two of these patients, the initial NSCCUP was HPV-driven. Twelve tumors were detected in the head and neck region and one in the lung. Of the twelve patients with head and neck tumor, six (50%)

were only treated by neck dissection after NSCCUP diagnosis, but did not receive postoperative radiotherapy and/or chemotherapy.

Most patients (n=7) had a base of tongue carcinoma, among which one patient had presented with an HPV35-driven NSCCUP 18 months earlier (Table 13) and also the base of tongue carcinoma contained HPV35 DNA and mRNA and overexpressed $p16^{INK4a}$. In addition, serum samples were available from time of: NSCCUP diagnosis (month 0), a follow-up visit (month 14) and the base of tongue carcinoma diagnosis (month 18). High levels of HPV35 E7 antibodies (median fluorescence intensity >3000) and low to intermediate levels of HPV35 E6 antibodies (200-700) were present in all three sera.

One patient with subsequently detected carcinoma in the nasopharynx had an HPV16-driven NSCCUP. Since nasopharyngeal carcinomas are frequently associated with Epstein-Barr-Virus (EBV), the NSCCUP was additionally analyzed for presence of EBV DNA [141], but was found negative.

Table 13: Patients with carcinoma detected within three years after NSCCUP diagnosis.

Patient	Carcinoma	Time after NSCCUP diagnosis (months)	Postoperative Treatment	NSCCUP HPV status
ID 161	Base of tongue	10	No	No
ID 10	Base of tongue	11	CRT	No
ID 86	Base of tongue	18	No	HPV35
ID 13	Base of tongue	26	No	No
ID 31	Base of tongue	29	CRT	No
ID 166	Base of tongue	32	No	No
ID 49	Base of tongue	35	CRT	No
ID 150	Oral cavity	1	RT	No
ID 139	Retromolar trigone (oral cavity)	12	No	No
ID 132	Pyriform sinus (hypopharynx)	12	RT	No
ID 9	Nasopharynx	14	CRT	HPV16
ID 143	Helix (ear)	24	No	No
ID 85	Lung	29	CRT	No

RT=radiotherapy, CRT=chemoradiotherapy

4.2 HPV serology in NSCCUP patients

4.2.1 Study population

Serological samples (n=134) from 48 patients presenting with NSCCUP at the University Hospital in Leipzig between 2008 and 2016 (median: 2011) were collected and available for retrospective analysis to assess the potential of HPV antibodies as a diagnostic marker for HPV-driven NSCCUP and as a prognostic marker in patients with HPV-driven NSCCUP.

From three patients that had discordant HPV status by serology versus molecular HPV markers, DNA extracted from the serum and the metastasis was analyzed by SNP-genotyping [139] to ascertain that paired samples truly were derived from the same patient. This led to the exclusion of two patients due to sequence identity of only 57% and 73%, respectively, indicating that serum and tissue were not from the same donor. Thus, 128 sera from 46 patients were included in the analyses presented here.

The majority of patients was male (41/46, 89%), and the median age at diagnosis was 59 years (ranging from 38 to 80) (Table 14). Tobacco and alcohol consumption were common (36/46, 78% and 41/46, 89%, respectively). At presentation, the majority of patients had advanced nodal disease (29/45, 64% N2b/2c/3) and extracapsular spread (32/42, 76%).

Most patients (38/44, 86%) received surgical treatment, which was the only treatment in ten patients (23%) and was followed in six patients (14%) by radiotherapy and in 21 patients (48%) by chemoradiotherapy. Four patients (9%) received radiotherapy alone, one patient chemotherapy alone and one patient postoperative chemotherapy, while another patient refused treatment. One patient presented with a base of tongue carcinoma 15 months after surgical treatment, likely to be the primary tumor.

The serology study population from Leipzig and the population of the multicenter study (Chapter 4.1.1, Table 9) were similar regarding age at diagnosis, gender, tobacco consumption, N stage and extracapsular spread. However, alcohol consumption was more common in the Leipzig study population (89% vs. 69%, p=0.006, Pearson chi squared test) and fewer patients from Leipzig received postoperative radiotherapy (14% vs. 30%, p=0.0009), but were more often treated by radiotherapy alone (9% vs. 1%).

Table 14: Characteristics of NSCCUP patients from Leipzig in relation to HPV serostatus.

	Total (n=46)	Seropositive (n=13)	Seronegative (n=33)	p†
Gender				
Male	41 (89%)	11 (85%)	30 (91%)	0.6
Female	5 (11%)	2 (15%)	3 (9%)	
Age, years (median)	38-80 (59)	52-79 (66)	38-80 (58)	0.2
Tobacco				
Ever	36 (78%)	6 (46%)	30 (91%)	**0.003**
Never	10 (21%)	7 (54%)	3 (9%)	
Alcohol				
Ever	41 (89%)	11 (85%)	30 (91%)	0.6
Never	5 (11%)	2 (15%)	3 (9%)	
N stage				
1, 2a	16 (36%)	7 (54%)	9 (28%)	0.4
2b	22 (49%)	5 (38%)	17 (53%)	
2c, 3	7 (16%)	1 (8%)	6 (19%)	
N/A	1	1		
Extracapsular spread				
Yes	32 (76%)	6 (55%)	26 (84%)	0.09
No	10 (24%)	5 (45%)	5 (16%)	
N/A	4	2	2	
Treatment				
Surgery only	10 (23%)	3 (23%)	7 (23%)	0.7
+RT	6 (14%)	3 (23%)	3 (10%)	
+CRT	21 (48%)	5 (38%)	16 (52%)	
Other*	7 (16%)	2 (15%)	5 (16%)	
N/A	2		2	

N/A=not available, +RT=postoperative radiotherapy, +CRT=postoperative chemoradiotherapy, *including radiotherapy only (n=4), chemotherapy only (n=1), postoperative chemotherapy (n=1) and one untreated case, †Fisher's exact test, student's t-test for age
Statistically significant values are displayed in bold.

4.2.2 HPV antibody patterns

Out of 46 NSCCUP patients, 13 (28%) were HPV-seropositive as defined (i) for HPV16 and HPV18 by high antibody levels (MFI >1000) against E6 and/or presence of antibodies against at least three type-concordant early antigens (E6, E7, E1, E2, E4) [54, 55] and (ii) for the other HPV types (HPV31, 33, 35, 45, 52, 58) by presence of type-concordant E6 and E7 antibodies.

Most seropositive patients were seropositive for HPV16 (10/13, 77%), single patients were seropositive for HPV18 and HPV33, respectively, and one patient for both HPV31 and HPV35. High antibody levels against HPV16 E6 were present in nine of the ten HPV16-seropositive patients and HPV16 E6, E7, E1 and E2 antibodies were simultaneously present in seven patients. One HPV16-seropositive patient lacked E6 antibodies but had high antibody levels against HPV16 E7, E1 and E2 (Figure 13). HPV16 E4 and L1 antibodies were only detectable in three (30%) and five (50%) of the HPV-seropositive cases, respectively (Figure 13).

The HPV18-seropositive case had high antibody levels against HPV18 antigens E7, E1, E2 and L1, but lacked E6 antibodies (Figure 13). One case had high antibody levels against HPV31 E6, E7 and L1, but also against HPV35 E6 (low) and E7 (high). The HPV33-seropositive case had high antibody levels against HPV33 E6 and E7, but no L1 antibodies.

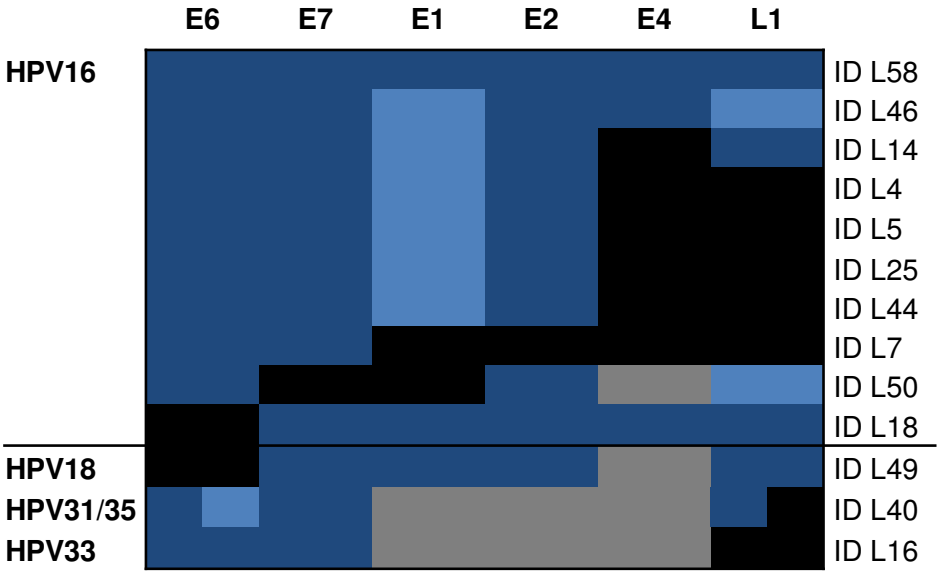


Figure 13: Antibody reactivity patterns against HPV antigens in 13 HPV-seropositive NSCCUP patients at diagnosis. Presence (light/dark blue = low/high levels) or absence (black) of antibodies against HPV16, HPV18, HPV31/35 or HPV33 proteins, measured in 1:100 serum dilution and based on pre-defined cut-offs [54], is depicted. High antibody levels (dark blue) were defined by MFI >1000. For some HPV types some antigens were not available (grey) and one HPV16-seropositive serum was not tested for E4 (grey).

In 16 (48%) out of the 33 HPV-seronegative patients, antibodies were detected against single (n=10), two (n=5, type-concordant only 1) or three (n=1, not type-concordant) HPV antigens (Table 15). Antibodies were mainly against HPV L1 (n=13), followed by E7 (n=6), E2 (n=2) and E6 (n=2).

Table 15: Individual antibody reactivity in HPV-seronegative NSCCUP patients.

Patient	HPV16	HPV18	HPV31	HPV33	HPV35	HPV45	HPV52	HPV58	N
ID L52	L1		L1		L1				3
ID L29	E2, L1								2
ID L24B				E7	E7				2
ID L17				L1	L1				2
ID L51					L1		L1		2
ID L34						E7		L1	2
ID L23		E2							1
ID L41		L1							1
ID L8				E7					1
ID L33				E7					1
ID L30					E6				1
ID L24					L1				1
ID L38						E6			1
ID L36							E7		1
ID L6								L1	1
ID L13								L1	1

Only HPV-seronegative patients with antibody reactivity against at least one (N≥1) HPV antigen are shown.

4.2.3 Diagnostic value of HPV seropositivity

From 28/46 (61%) NSCCUP patients with serum samples, FFPE tissues from the lymph node metastases were available for molecular HPV status assessment to be compared with the serological HPV status. Of these patients, nine were seropositive for HPV16 and one for both HPV31 and HPV35. The molecular HPV status was defined as described for the multicenter study (Chapter 4.1.2). Eleven (39%) of the 28 metastases were HPV-driven, nine (82%) by HPV16 and one each by HPV31 and HPV33. All HPV16-driven metastases were positive for HPV16 mRNA, DNA and $p16^{INK4a}$ overexpression and the patients were concordantly HPV16-seropositive (Table 16). Patients with HPV16-driven NSCCUP showed significantly increased antibody reactivity against the individual HPV16 antigens E6, E7, E1, E2 and L1 compared to patients with non-HPV16-driven NSCCUP ($p \leq 0.002$, Mann-Whitney U test, Figure 14).

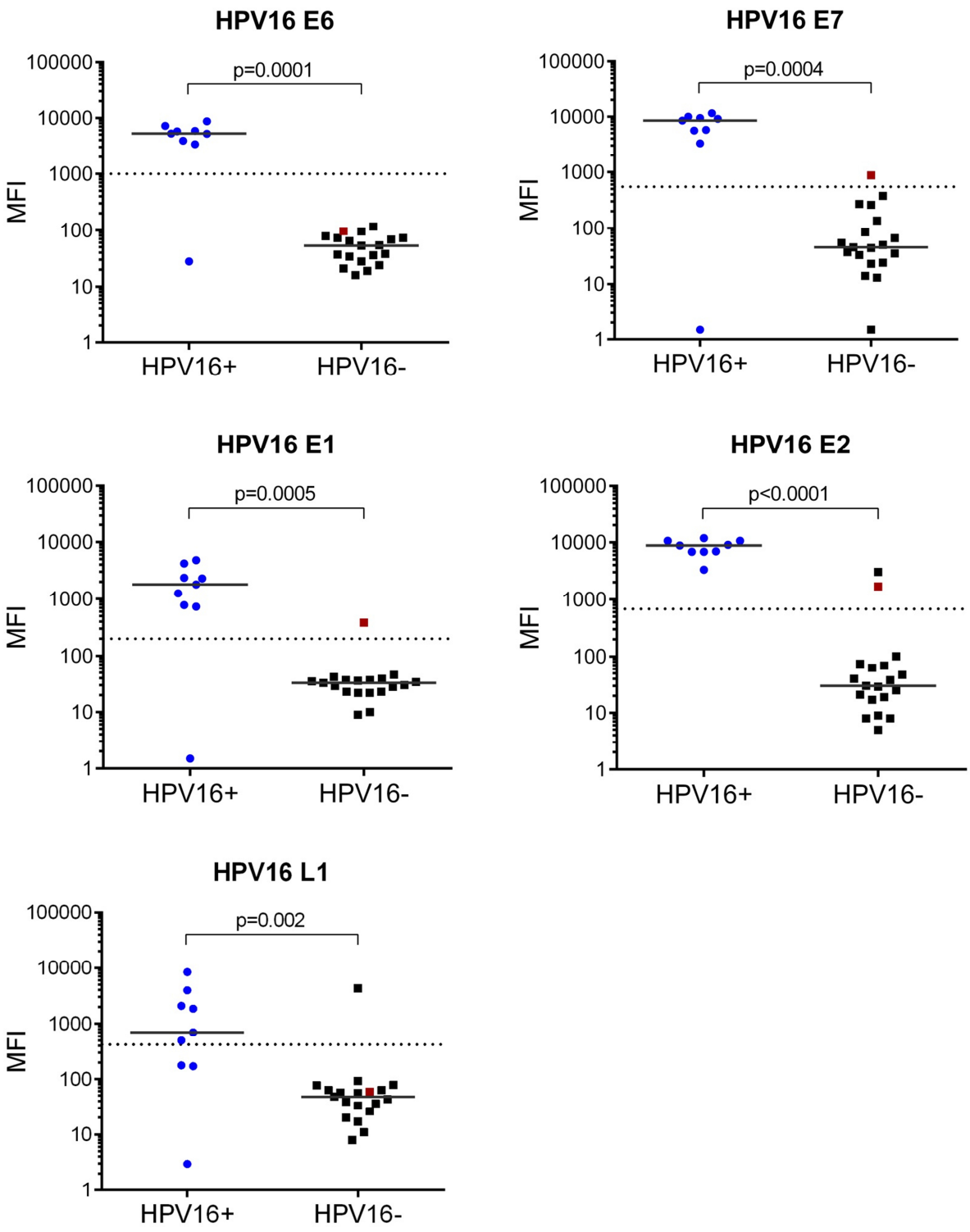


Figure 14: Antibody reactivity to individual HPV16 antigens in NSCCUP patient serum at time of diagnosis by molecular HPV status. Serum antibody reactivity at 1:100 dilution against HPV16 antigens is depicted as median fluorescence intensity (MFI) for patients with HPV16-driven (HPV16+) compared to non-HPV16-driven (HPV16-) NSCCUP. HPV16-driven NSCCUP are defined by presence of HPV16 mRNA together with HPV DNA and/or $p16^{INK4a}$ overexpression in the lymph node metastasis. Patients were defined HPV16-seropositive (blue, high HPV16 E6 antibody levels and/or antibodies against ≥3 early antigens), HPV31/35-seropositive (red, E6 and E7 antibodies of HPV31 and HPV35) or HPV-seronegative (black). The median of MFI levels (short solid line) and antigen-specific cut-off values (dotted line) are depicted. P values were calculated using Mann-Whitney U test.

The patient (ID L40) with HPV31-driven metastasis had high antibody levels against HPV31 E6, E7 (and also L1), but additional antibody reactivity against HPV35 E6 and E7 (Figure 13). In another patient (ID L24B) with HPV33-driven metastasis, the serological HPV status was not concordant with the molecular HPV status, since this case had antibodies against HPV33 E7, but not against HPV33 E6 (and L1), and was thus defined HPV-seronegative (Table 15).

In total, of the eleven cases with HPV-driven metastases, ten cases were seropositive for the same HPV type, but one case was HPV-seronegative. All 17 cases with non-HPV-driven metastases were HPV-seronegative, resulting in an overall sensitivity of 91% and specificity of 100% for identifying HPV-driven NSCCUP by HPV serology (Table 16). NSCCUP driven by HPV16 were identified with 100% sensitivity and specificity.

Table 16: Concordance of HPV serostatus and molecular HPV status in 28 NSCCUP patients.

		HPV mRNA* n		Sensitivity %	Specificity %	Cohen's Kappa
		+	-	(95% CI)	(95% CI)	(95% CI)
HPV serostatus	**+**	10	0	91 (59-100)	100 (81-100)	0.93 (0.78-1.00)
	-	1	17			
HPV16 serostatus	**+**	9	0	100 (66-100)	100 (82-100)	1.00 (1.00-1.00)
	-	0	19			

HPV serostatus was defined by presence of antibodies against HPV16/18 E6 (MFI >1000) and/or ≥3 early type-concordant antigens, or against type-concordant E6 and E7 of any other HPV type.
*Only positivity for HPV16 mRNA was considered for comparison with HPV16 serostatus.

4.2.4 Prognostic value of HPV antibodies

As observed for patients with HPV-driven NSCCUP in the multicenter study (Chapter 4.1.4, Table 11), tobacco consumption was significantly less common also among HPV-seropositive versus HPV-seronegative NSCCUP patients (46% vs. 91%, p=0.003, Fisher's exact test, Chapter 4.2.1, Table 14). HPV-seropositive patients presented with less advanced N stage (54% vs. 28% N1/2a) and less frequently with extracapsular spread (55% vs. 84%) compared to HPV-seronegative patients, while gender distribution, age at diagnosis, alcohol consumption and treatment were similar (Table 14).

HPV-seropositive NSCCUP patients showed significantly better overall and progression-free survival compared to HPV-seronegative patients (p=0.004 and p=0.02, log-rank test; Figure 15). Multivariable Cox regression analysis revealed for HPV-seropositivity hazard ratios of 0.09 (p=0.001, Wald test) and 0.03 (p=0.002) for overall and progression-free survival, respectively (Table 17). Another prognostic factor was multimodal treatment with a hazard ratio of 0.09 (p=0.007) and 0.02 (p=0.0007) for overall and progression-free survival,

respectively. Extracapsular spread was an unfavorable prognostic factor for progression-free survival with a hazard ratio of 6.64 (p=0.02), while no impact on survival was observed for gender, age at diagnosis, tobacco and alcohol consumption and N stage (Table 17).

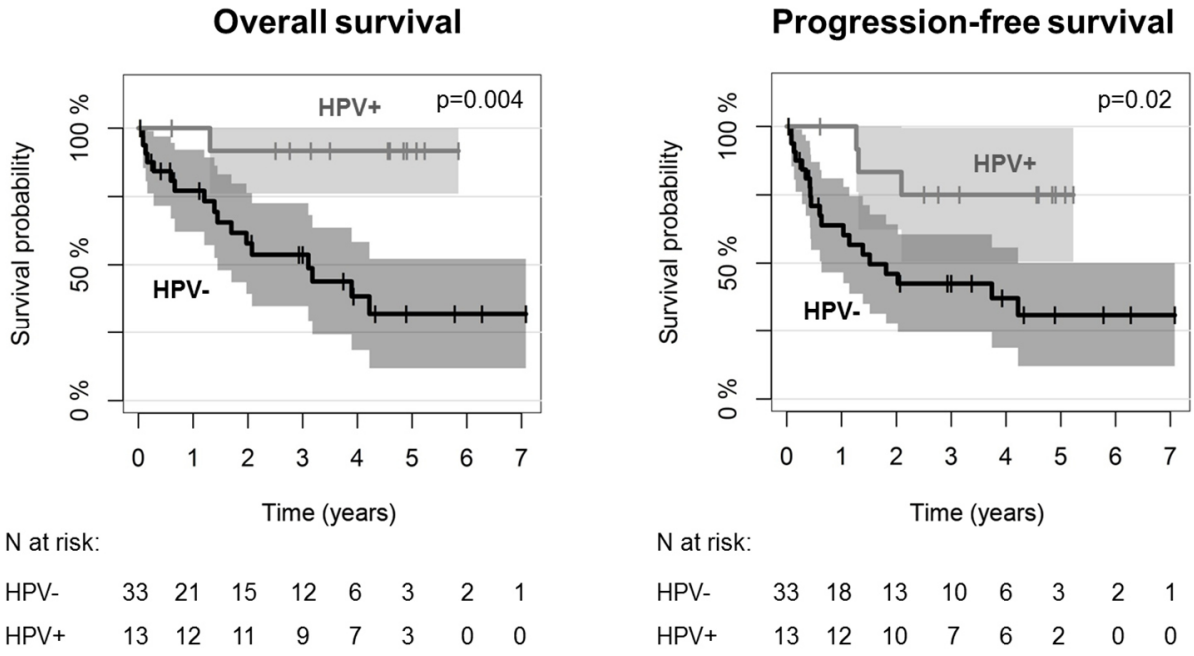


Figure 15: Survival of NSCCUP patients from Leipzig in relation to HPV serological status. Kaplan–Meier curves for overall (left) and progression-free survival (right) with 95% confidence interval (shaded areas) are compared for HPV-seropositive (HPV+, grey) versus HPV-seronegative (HPV-, black) cases. The numbers at risk are depicted below the graph. P values were calculated using log-rank test.

Table 17: HPV serostatus and clinical parameters in relation to survival.

Multivariable analysis (n=39):	**OS**		**PFS**	
	HR (95% CI)	**p†**	**HR (95% CI)**	**p†**
HPV serostatus (positive/negative)	**0.09 (0.01 - 0.42)**	**0.001**	**0.03 (0.002 - 0.18)**	**0.0002**
Gender (female/male)	0.23 (0.001 - 2.07)	0.2	0.26 (0.002 - 2.31)	0.3
Age at diagnosis	1.04 (0.98 - 1.13)	0.2	1.07 (0.99 - 1.16)	0.07
Tobacco (ever/never)	3.56 (0.60 - 24.2)	0.1	0.69 (0.10 - 4.65)*	0.7
Alcohol (ever/never)	0.55 (0.10 - 4.48)	0.5	1.64 (0.22 - 21.7)	0.7
N stage				
2b vs. 1,2a	3.24 (0.92 - 14.3)	0.07	1.25 (0.40 - 4.15)	0.7
2c,3 vs. 1,2a	0.37 (0.05 - 2.13)	0.3	0.54 (0.10 - 2.61)	0.4
Extracapsular spread (yes/no)	2.46 (0.52 - 22.4)	0.3	**6.64 (1.32 - 79.3)**	**0.02**
Treatment (multimodal/other)**	**0.09 (0.01 - 0.53)**	**0.007**	**0.02 (0.001 - 0.22)**	**0.0007**

Firth's penalized likelihood method has been applied to mitigate convergence problems due to the fact that none of the female patients experienced an event. Statistically significant values are displayed in bold.
OS=overall survival, PFS=progression-free survival, HR=hazard ratio, CI=confidence interval, multimodal= postoperative chemoradiotherapy, *PH assumption violated, **including one untreated case, †Wald test

4.2.5 HPV antibody levels in follow-up sera from seropositive NSCCUP

From five HPV-seropositive (3x HPV16, 1x HPV18 and 1x HPV33) NSCCUP patients between two and nine serial serum samples were available for follow-up periods of at least eight months and up to 60 months after treatment. HPV antibody levels in patients with HPV-driven OPSCC and NSCCUP are often very high and are thereby frequently in saturation, when measured at the routine 1:100 serum dilution. To better assess potential changes in antibody levels, all sera from HPV-seropositive patients with serial serum samples were additionally measured at 1:1,000 and 1:10,000 serum dilutions.

During follow-up, all patients remained HPV-seropositive, and HPV antibodies were still detectable at 1:10,000 serum dilution even five years after neck dissection (Figure 16). HPV antibody levels decreased within about one year irrespective of the HPV antigen, whereas antibody levels against the control antigens from the polyomaviruses BK (BKV), JC (JCV) and Merkel cell polyomavirus (MCV) did not (Figure 16). While in one patient (ID L7) HPV16 E6 and E7 antibody levels within eight months showed a reduction by 83% and 74%, respectively, in another case (ID L49) HPV18 E7, E1 and E2 antibody levels decreased by only 16%, 31% and 16%, respectively. In two further patients with longer follow-up, the main decrease in HPV antibody levels was observed within the first 18 months, ranging from 25% (ID L18, E2) to 52% (ID L16, E7), while within the following 18 months antibody levels decreased only by additional 14% and 10%, respectively.

One patient (ID L5) with HPV16-driven NSCCUP, who received surgical treatment but declined chemoradiotherapy, was diagnosed with an HPV16-driven base of tongue carcinoma 15 months after NSCCUP diagnosis. In this case, HPV16 antibody levels also decreased initially after surgical removal of the lymph node metastasis by 18% (E1), 34% (E6) and 36% (E7 and E2). However, relapse with new lymph node metastasis at month 29 was associated with an antibody level increase relative to the level measured immediately after surgical removal of the base of tongue carcinoma, not for E6 but for all three other HPV16 antigens (17% for E2, 19% for E1 and 57% for E7), reaching similar levels as measured at the time of NSCCUP diagnosis (81% for E2, 101% for E1 and 121% for E7) (Figure 16). In the following 18 months, antibody levels again decreased to 40%-50% (E6, E2, E1) or 85% (E7).

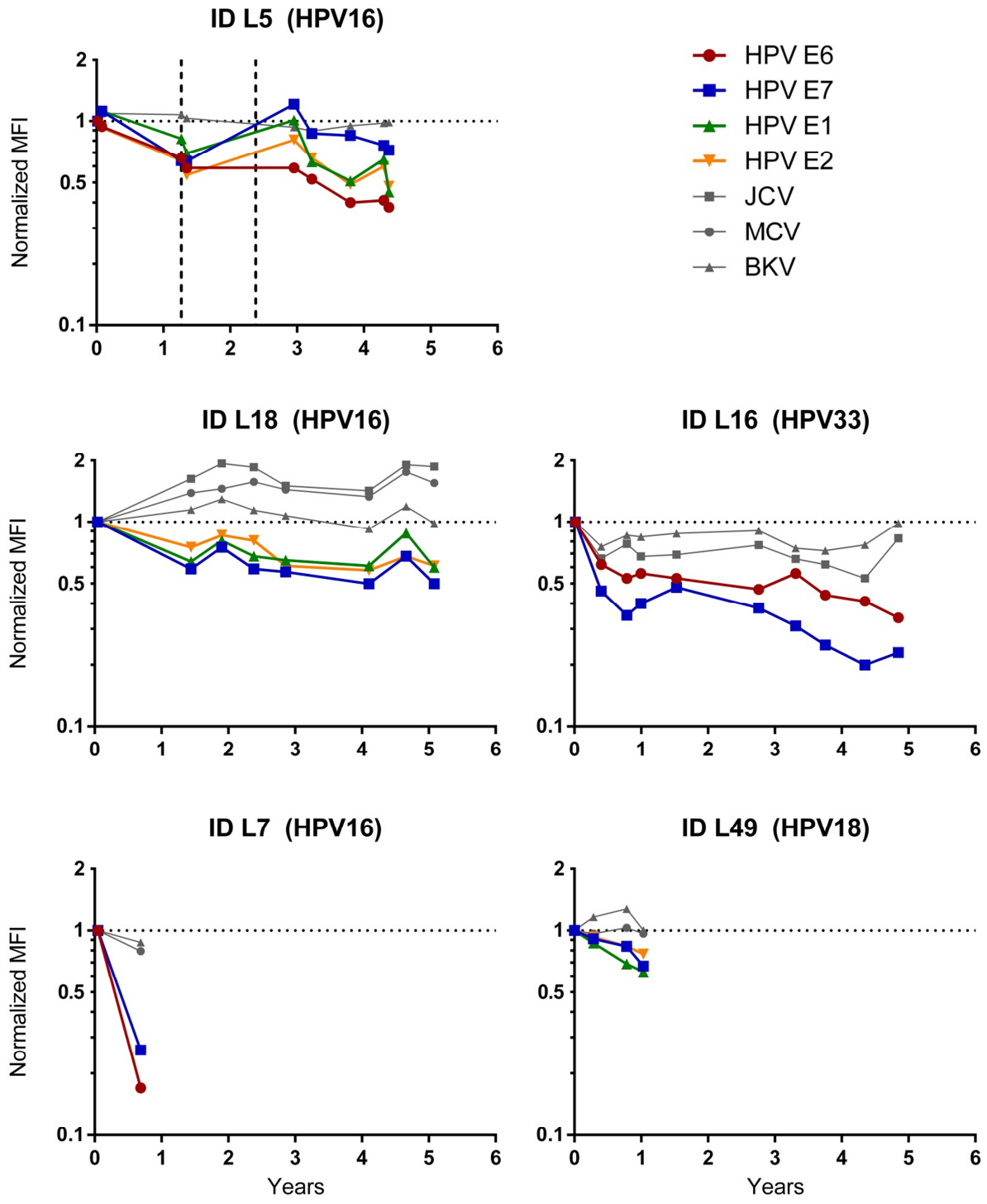


Figure 16: Antibody levels in follow-up sera from NSCCUP patients. HPV serum antibody levels of HPV-seropositive NSCCUP patients with up to nine follow-up sera collected within six years after neck dissection (time=0) measured at 1:1,000 (HPV16/18 E7 and E1, HPV33 E6) or 1:10,000 (HPV16 E6 and E2, HPV18 E2, HPV33 E7) serum dilution are depicted as normalized median fluorescence intensity (MFI) with 1 representing the value immediately after treatment (dotted line), and 0.5 a reduction of 50%. For sera collected within two consecutive days, the average value is depicted. Levels of JC virus (JCV), BK virus (BKV) and Merkel cell polyomavirus (MCV) control antigens were added (1:1,000, grey). Only antigens that were seropositive at time=0 are shown. HPV16 E4 and L1 were excluded due to rather low antibody reactivity (Figure 13) and weakest association with HPV-driven NSCCUP (Figure 14). The dashed vertical lines indicate subsequent discovery of the primary tumor (at month 15) and relapse (at month 29).

4.3 TP53 mutations in NSCCUP

4.3.1 TP53 mutation frequency

In order to investigate potential additional molecular differences between HPV-driven and non-HPV-driven NSCCUP, presence of tumor suppressor TP53 mutations was assessed in lymph node metastases from 70 NSCCUP patients from Heidelberg, Barcelona and Treviso by amplification and sequencing of the TP53 exons 4-8, which encode the DNA binding domain of p53. Exons 9 and 10 were additionally assessed in 57 patients.

Amplification and sequencing was successful for 397 (86%) of the 464 assessed exons. The success rate per exon ranged from 77% for exon 4 to 91% for exons 6 and 9 (Table 18). All seven exons (4-10) were successfully sequenced in 42 patients, and exons 4-8 in 45 patients. Of the 397 sequenced exons, 19 (5%) were mutated. The mutation frequency per exon varied between 0% for exon 9 and 10% for exon 7 (Table 18).

Only single TP53 mutations were detected in 19 patients (Table 19), but since in five of these patients one or two exons each could not be sequenced, additional mutations cannot be excluded. Out of the 19 identified TP53 mutations, 16 (84%) were disruptive, comprising missense mutations (n=7) that were all located within the DNA binding domain (amino acids 102-292), frameshift mutations (n=4, three resulting in missense mutations at a later amino acid position), as well as nonsense mutations (n=5, four within the DNA binding domain, Table 19).

Table 18: Mutation frequency per TP53 exon in NSCCUP tumor DNA.

TP53 exon	N assessed	N (%) sequenced	N mutated	% mutated of N sequenced
4	70	54 (77)	4	7
5	70	61 (87)	3	5
6	70	64 (91)	4	6
7	70	61 (87)	6	10
8	70	59 (84)	1	2
9	57	52 (91)	0	0
10	57	46 (81)	1	2
Overall	**464**	**397 (86)**	**19**	**5**

Table 19: TP53 mutations detected in NSCCUP tumor DNA.

TP53 exon	cDNA description	Protein description	Effect	Database report
4	c.182delA	p.D61fs	frameshift	
4	c.214_215delCGinsT	p.R72fs*50	frameshift/missense	
4	c.245_246insTG	p.P82fs*41	frameshift/missense	Cosmic
4	c.366G>A	p.V122V	synonymous	
5	c.451C>A	p.P151T	missense	IARC, Cosmic
5	c.455_456insC	p.P152fs*34	frameshift/missense	
5	c.535C>A	p.H179N	missense	IARC, Cosmic
6	c.584T>A	p.I195N	missense	IARC, Cosmic
6	c.585_586CC>TT	p.R196*	nonsense	IARC, Cosmic
6	c.586C>T	p.R196*	nonsense	IARC, Cosmic
6	c.660T>A	p.Y220*	nonsense	IARC, Cosmic
7	c.705C>T	p.N235N	synonymous	
7	c.707A>G	p.Y236C	missense	IARC, Cosmic
7	c.708C>A	p.Y236*	nonsense	IARC, Cosmic
7	c.712T>A	p.C238S	missense	IARC, Cosmic
7	c.743G>A	p.R248Q	missense	IARC, Cosmic
7	c.753C>T	p.I251I	synonymous	
8	c.833C>T	p.P278L	missense	IARC, Cosmic
10	c.1024C>T	p.R342*	nonsense	IARC, Cosmic

del=deletion; ins=insertion; fs=frameshift; *stop codon position;
IARC database: http://p53.iarc.fr/; Cosmic database: http://cancer.sanger.ac.uk/cosmic/gene/analysis?ln=TP53

4.3.2 TP53 mutations in relation to HPV status

NSCCUP patients with disruptive TP53 mutation in any exon (n=16) were compared to patients without disruptive mutation but with complete sequence of exon 4-9 (n=32) regarding HPV status and clinical parameters. None of the 17 patients with HPV-driven NSCCUP presented with disruptive TP53 mutation, but one patient had a non-disruptive mutation in exon 4. In contrast, disruptive mutations were found in 16/31 (52%) patients with non-HPV-driven NSCCUP (p=0.0002, Fisher's exact test).

Patients with non-HPV-driven NSCCUP without disruptive TP53 mutation presented with less extracapsular spread, but with significantly advanced N stage compared to patients with HPV-driven NSCCUP or disruptive mutation. Regarding age at diagnosis, gender, treatment, tobacco and alcohol consumption, no statistically significant differences were observed (Table 20).

Table 20: Characteristics of NSCCUP patients in relation to TP53 mutation status.

	Total (n=48)	HPV+ TP53 wt (n=17)	HPV- TP53 wt (n=15)	HPV- TP53 mut (n=16)	p†
Age, years (median)	**39-85 (58)**	39-79 (62)	40-85 (56)	42-80 (62)	0.5
Gender					
Male	**39 (81%)**	12 (71%)	12 (80%)	15 (94%)	0.2
Female	**9 (19%)**	5 (29%)	3 (20%)	1 (6%)	
Tobacco					
Ever	**38 (79%)**	11 (65%)	12 (80%)	15 (94%)	0.1
Never	**10 (21%)**	6 (35%)	3 (20%)	1 (6%)	
Alcohol					
Ever	**32 (67%)**	9 (53%)	12 (80%)	11 (69%)	0.3
Never	**16 (33%)**	8 (47%)	3 (20%)	5 (31%)	
N stage					
1, 2a	**19 (40%)**	9 (53%)	3 (20%)	7 (44%)	**0.04**
2b	**20 (42%)**	8 (47%)	6 (40%)	6 (38%)	
2c, 3	**9 (19%)**	0	6 (40%)	3 (19%)	
Extracapsular spread					
Yes	**26 (58%)**	11 (65%)	5 (36%)	10 (71%)	0.2
No	**19 (42%)**	6 (35%)	9 (64%)	4 (29%)	
N/A	**3**		1	2	
Treatment					
Surgery only	**9 (20%)**	4 (24%)	1 (8%)	4 (25%)	0.1
+RT	**18 (39%)**	4 (24%)	5 (38%)	9 (56%)	
+CRT	**19 (41%)**	9 (53%)	7 (54%)	3 (19%)	
N/A	**2**		2		

HPV+=HPV-driven; HPV-=non-HPV-driven; wt=wild-type sequence or non-disruptive mutation; mut=disruptive mutation; N/A=not available
†Fisher's exact test (age categories: ≥62 vs. <62)
Statistically significant values are displayed in bold.

4.3.3 Prognostic value of TP53 mutations

TP53 mutation status and HPV status were correlated with survival. Patients with HPV-driven NSCCUP without disruptive TP53 mutation (n=18) had the highest overall and progression-free survival rate (Figure 17). A significant difference was revealed for progression-free survival (p=0.04, log-rank test). Among the patients with non-HPV-driven NSCCUP (n=30),

those with disruptive TP53 mutation (n=15) showed slightly better survival than the patients without disruptive mutation within the first six years after diagnosis, but the difference was not statistically significant.

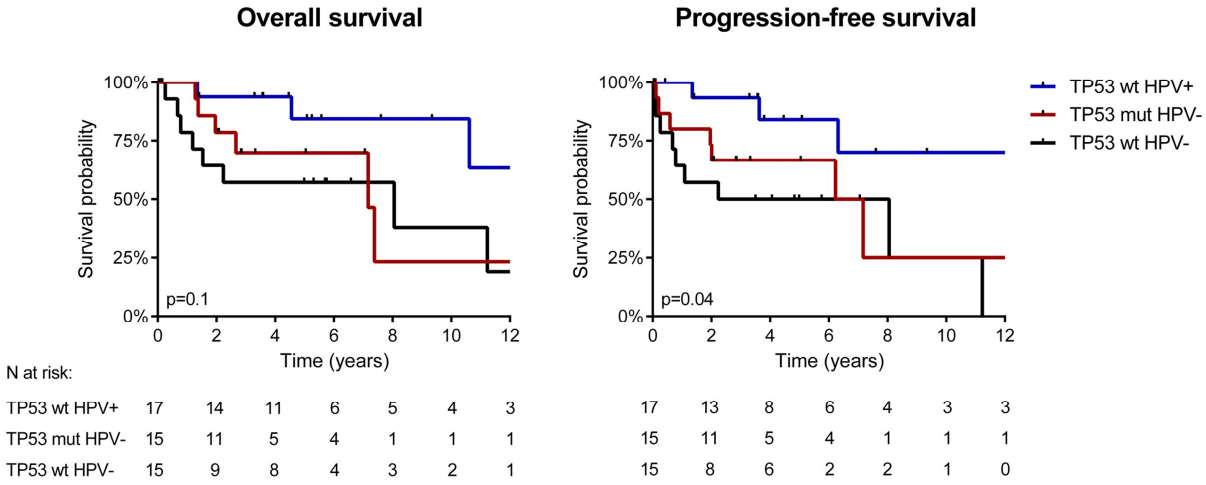


Figure 17: Survival of NSCCUP patients in relation to TP53 mutation status and HPV. Kaplan–Meier curves for overall (left) and progression-free survival (right) are compared for patients with non-HPV-driven NSCCUP with (TP53 mut HPV-, red) or without disruptive TP53 mutation (TP53 wt HPV-, black) and patients with HPV-driven NSCCUP without disruptive mutation (TP53 wt HPV+, blue). The numbers at risk are depicted below the graph. Log-rank test was used to calculate p values.

4.4 Methylation analysis of NSCCUP

4.4.1 HPV-associated methylation signature

In order to assess epigenetic differences between HPV-driven and non-HPV-driven NSCCUP as another potential prognostic marker, a pre-defined methylation signature was analyzed in metastatic lymph nodes from 103 NSCCUP patients, a subset of the series from Heidelberg, Barcelona and Treviso, in collaboration with Dr. Dieter Weichenhan (DKFZ). The methylation signature was defined by low methylation in the promoters of the genes ALDH1 and OSR2 and high methylation in the promoters of GATA4, GRIA4 and IRX4 and was previously shown to be associated with HPV-driven HNSCC and to predict better survival [76, 77]. Based on the methylation levels in those five promoters, a methylation score (MS, number of promoters with methylation level according to this signature, range 0-5) was calculated to group patients accordingly.

Tumor DNA from HPV-driven compared to non-HPV-driven NSCCUP was characterized by lower methylation levels in ALDH1 and OSR2 (p=0.4 and p=0.09, respectively; Mann-Whitney U test) and significantly higher methylation levels in GATA4, GRIA4 and IRX4 (p=0.0008, p=0.0005 and p=0.01, respectively, Figure 18). Of the HPV-driven NSCCUP,

76% (16/21) had a MS of ≥3, compared to only 32% (26/82) of the non-HPV-driven NSCCUP (p=0.0002, Pearson chi squared test). Thus, the methylation signature was significantly associated with HPV.

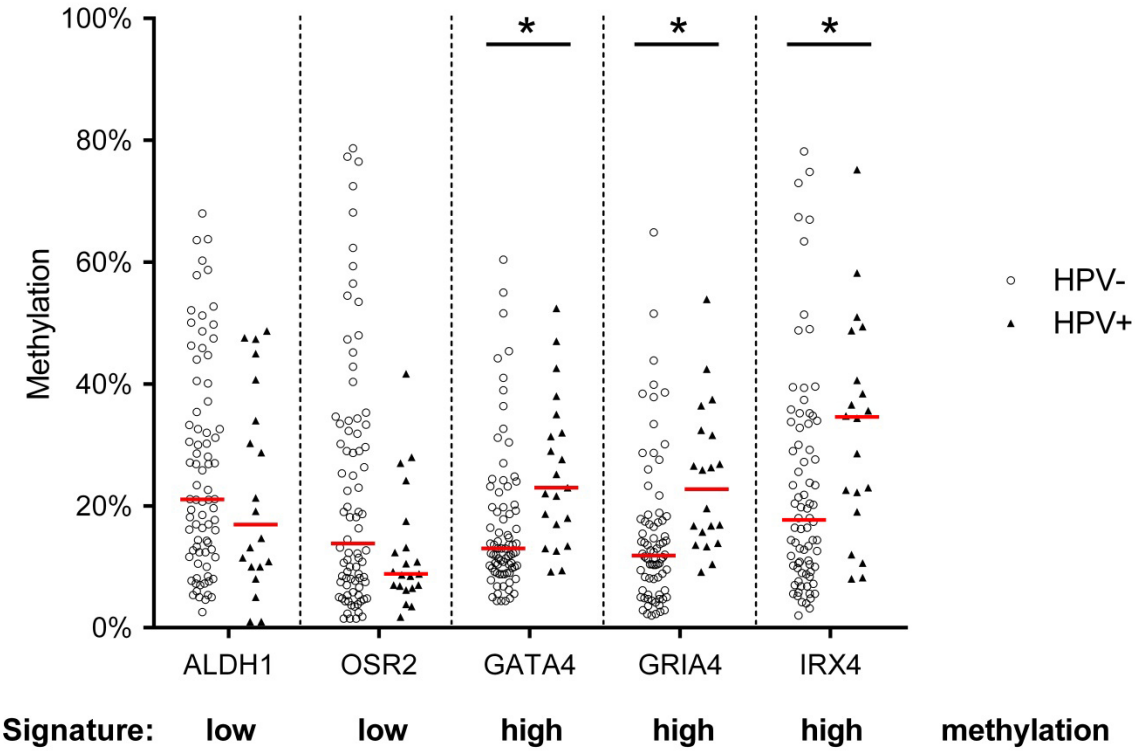


Figure 18: Methylation levels in five gene promoters of HPV-driven vs. non-HPV-driven NSCCUP. The percentage of methylation in the promoters of the five genes included in the pre-defined methylation signature is compared for non-HPV-driven (HPV-, open circles) vs. HPV-driven (HPV+, filled triangles) NSCCUP, defined by molecular HPV status (HPV+: presence of HPV16 mRNA together with HPV DNA and/or $p16^{INK4a}$ overexpression). The median methylation level is depicted for each group (red bar). An asterisk (*) indicates a significant difference (Mann-Whitney U test).

4.4.2 Prognostic value of methylation signature and HPV

Although the methylation signature was associated with HPV-driven NSCCUP, patients with MS≥3 did not have a significantly better overall or progression-free survival (p=0.6 and p=0.1, respectively, log-rank test; Figure 19, upper part) compared to patients with MS<3. But a survival benefit of patients with HPV-driven NSCCUP was observed (Figure 19, lower part), especially for the five patients presenting with NSCCUP with MS<3. However, when patients were regrouped into those with low (0-1), intermediate (2-3) or high MS (4-5), a trend for better overall and progression-free survival with increasing MS was observed (p=0.07 and p=0.03, respectively, Appendix, Figure A-2).

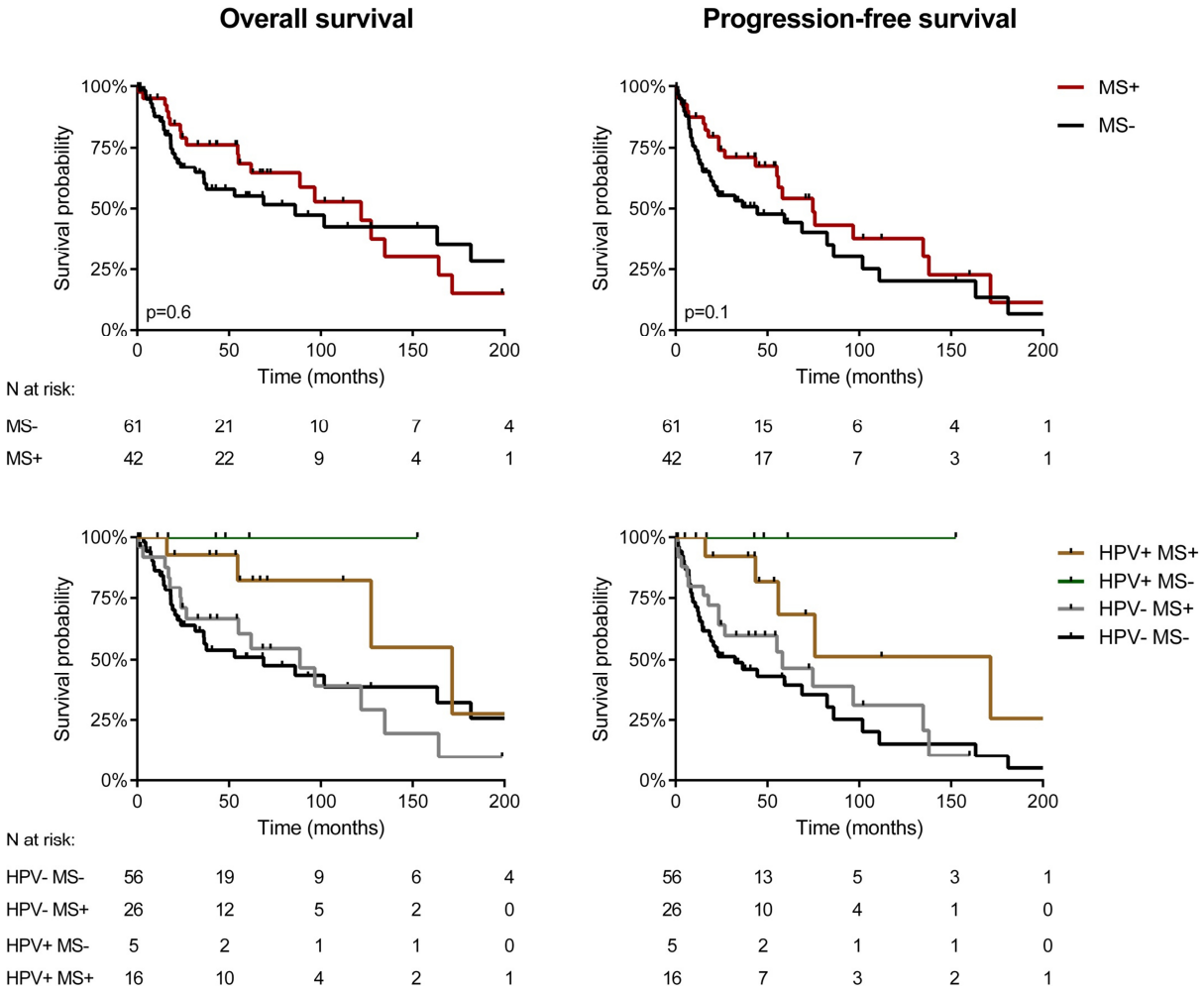


Figure 19: Survival of NSCCUP patients in relation to MS and HPV. Kaplan–Meier curves for overall (left) and progression-free survival (right) are compared for NSCCUP patients with at least three gene promoters (MS≥3, MS+, red) versus less than three promoters (MS<3, MS-, black) according to the pre-defined methylation signature (low methylation in the promoters of ALDH1 and OSR2, and high in GATA4, GRIA4 and IRX4). The lower panel shows survival of NSCCUP patients with MS≥3 (MS+) and MS<3 (MS-) by HPV status, meaning further subdivided into patients with HPV-driven (HPV+) and non-HPV-driven (HPV-) NSCCUP. The numbers at risk are depicted below the graph. Log-rank test was used to calculate p values. The time axis has been adapted to [76] for better comparison.

4.5 Paired analysis of OPSCC primary tumors and metastases

In order to assess potential heterogeneity between HPV-driven primary tumors and their corresponding lymph node metastases regarding HPV status and HPV-associated markers, pairs of primary tumors and metastases were compared. In particular, HPV-driven OPSCC were investigated, since those are the suspected source of HPV-driven lymph node metastases in NSCCUP patients.

4.5.1 HPV markers in primary tumors vs. metastases

Twelve HPV16-driven OPSCC primary tumors and corresponding lymph node metastases (n=20, 1-4 per patient) were available for paired HPV marker comparison. The three analyzed markers, HPV16 DNA (detected by MPG, and additionally by qPCR in frozen biopsies, if available), HPV16 mRNA (detected by E6*I assay and additionally by RNA pattern analysis in frozen biopsies, if available) and $p16^{INK4a}$ (only tested in five metastases with FFPE tissue available), were concordantly present in primary tumors and corresponding metastases.

However, the viral load (HPV16 E6 copies per tumor cell, assessed in frozen biopsies only) varied between primary tumors and corresponding metastases, as well as between metastases from the same patient (Table 21). Of the five analyzed metastases, two had a higher viral load compared to the primary tumor (factor 2-4) and three had a lower viral load (factor 2-6), while the intra-experimental and inter-experimental coefficient of variation, the ratio of standard deviation to mean, were very low (range: 0.1-2.2%).

Table 21: Viral load in primary tumors vs. corresponding metastases.

Patient	HPV16 E6 copies / tumor cell		
	Primary Tumor	Metastasis 1	Metastasis 2
661	1.6	3.2	-
993	12.6	2.2	-
1449	4.8	2.1	-
1553	0.2	0.8	0.1

4.5.2 HPV16 integration status in the primary tumors

The integration status of HPV16 DNA in the twelve primary tumors had been assessed in a previous study by TEN16 analysis (Tagging, Enrichment and Next-generation sequencing of HPV16, TEN16_2013, Holzinger et al, in preparation). In eight of the twelve primary tumors, a single (n=5), two (n=2), or four (n=1) viral-cellular junctions (VCJs) were identified, while four primary tumors had no integration. Since HPV16 DNA was concordantly present in the pairs, viral integration was assessed in lymph node metastases for comparison with the primary tumor.

4.5.3 Presence of viral-cellular junctions of the primary tumors in the metastases

Presence of all VCJ (n=13) previously identified in the eight primary tumors was assessed by VCJ-PCR in all available lymph node metastases (n=14, 1-4 per patient) from the same

patients. VCJs of the primary tumors were present in only 6/14 (43%) metastases (Table 22). In three patients (1425, 1464, 1650) all VCJ (four, two and two, respectively) were detected in their single metastasis, while the single VCJ of patient 1553 was detectable in only three out of four metastases (Figure 20, upper panel). The VCJs identified in the primary tumors of patients 661, 993, 1529 and 1643 were undetectable in all corresponding metastases (three, one, two and one, respectively).

Table 22: Validation of viral-cellular junctions of the primary tumors in corresponding metastases.

Patient	VCJ Primary Tumor	VCJ Metastasis			
	TEN16_2013	M1	M2	M3	M4
1425	1, 2, 3, 4	1,2,3,4	-	-	-
1464	1, 2	1,2	-	-	-
1553	1	no	1	1*	1*
1650	1, 2	1, 2	-	-	-
993	1	no	-	-	-
661	1	no	no	no	-
1529	1	no	no	-	-
1643	1	no	-	-	-

VCJ=viral-cellular junction, TEN16=Tagging, Enrichment and Next-generation sequencing of HPV16
*FFPE biopsies, in which integration sites were validated by VCJ-PCR

4.5.4 Identification of new viral-cellular junctions

A new TEN16 analysis (TEN16_2016) was performed in order to search for other VCJs that might be present only in the metastases. In addition, four previously assessed primary tumors, of which frozen biopsies were available for new sectioning, were re-analyzed to search for VCJs that might have been missed in the previous analysis (TEN16_2013, Holzinger et al, in preparation), since the processing of sequencing data has been further optimized. Furthermore, analysis of new sections from the same biopsy enables investigation of potential intratumoral heterogeneity.

All previously identified VCJs were confirmed in the re-analyzed primary tumors (Table 23). Additional 19 candidates for new VCJs were selected based on at least 20 read pairs from sequencing analysis. Validation VCJ-PCRs were performed for all 19 candidates, but only six (32%) could be confirmed, of which four were identified in primary tumors. In the primary

tumor of patient 1529, two additional VCJ were identified that were missed in the previous analysis (TEN16_2013), but were validated by VCJ-PCR using the DNA aliquot prepared for that analysis (Table 23). Additional two VCJ were identified in the primary tumor of patient 993 that could be validated by VCJ-PCR using the DNA extracted from the section prepared in 2016, but not in the DNA from the section prepared from the same biopsy in 2013, indicating intratumoral heterogeneity (Table 23).

Table 23: Identification of new viral-cellular junctions in primary tumors and metastases.

Patient	VCJ Primary Tumor		VCJ Metastasis			
	TEN16_2013	TEN16_2016	M1	M2	M3	M4
1553	1	1	**2**	1	1*	1*
1529	1, **2, 3**	1, **2, 3**	**2, 4**	w/o*	-	-
993	1	1, **2, 3**	w/o	-	-	-
661	1	1	w/o	w/o*	w/o*	-
339	w/o	na	w/o	w/o	-	-
1449	w/o	na	w/o	-	-	-
1639	w/o	na	w/o	-	-	-
1641	w/o	na	w/o	w/o	-	-

VCJ=viral-cellular junction, na=not assessed, w/o=without integration, TEN16=Tagging, Enrichment and Next-generation sequencing of HPV16, *FFPE biopsies, in which integration sites were validated by VCJ-PCR
New VCJ are displayed in bold.

However, new VCJs were identified in single metastases from two patients with known VCJ in their primary tumor. Metastasis 1 from patient 1529 had only one of the three VCJs present in the primary tumor, but one additional VCJ, that was exclusively present in this metastasis (Table 23). Metastasis 1 from patient 1553 was negative for the VCJ present in the primary tumor and in all three other metastases, but had another VCJ, which was not found in the primary tumor and in the other metastases by VCJ-PCR (Table 23, Figure 20). In four patients without VCJ in the primary tumor (339, 1449, 1639 and 1641), no VCJs were found in their corresponding metastases (Table 23).

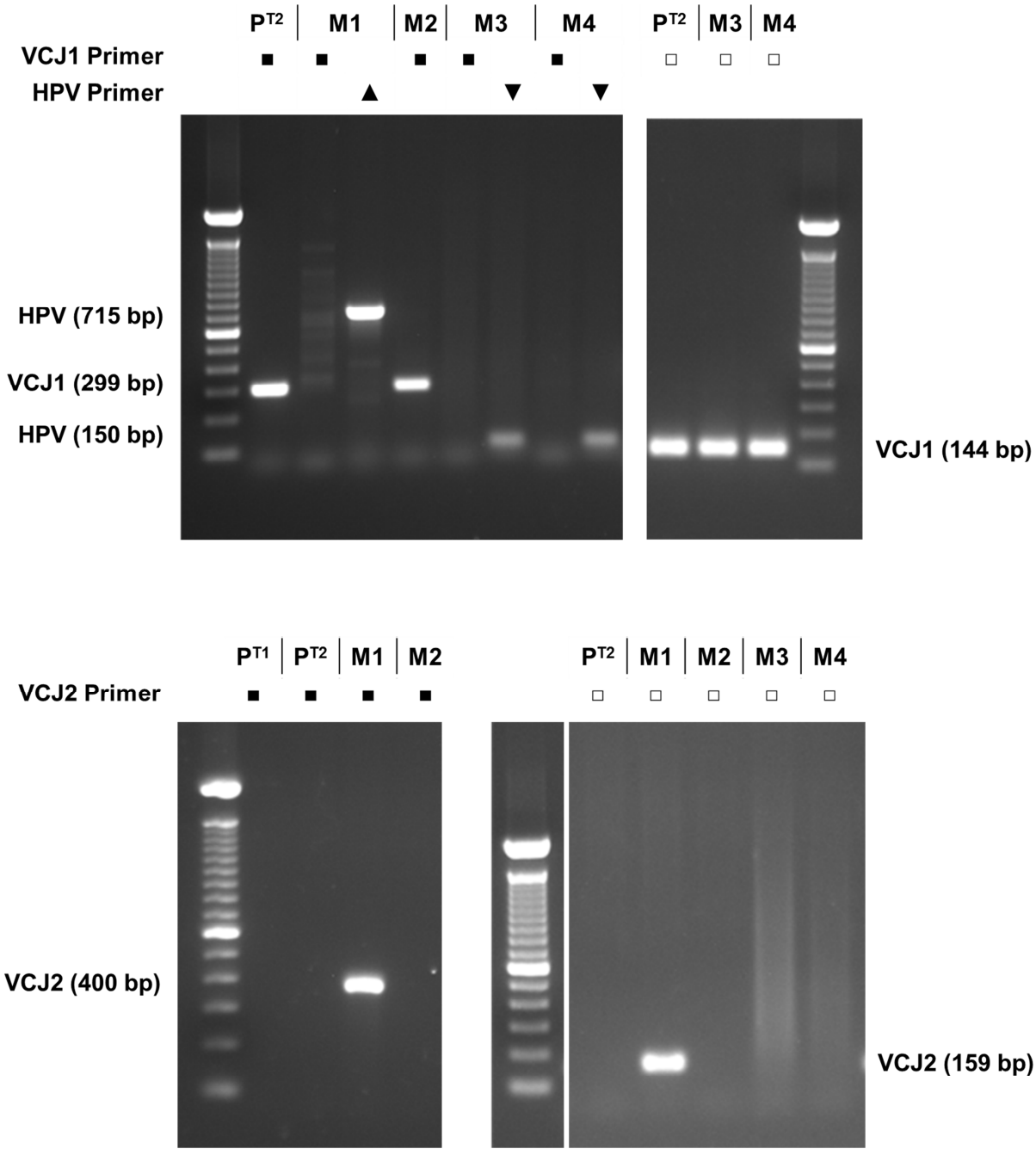


Figure 20: Validation of viral-cellular junctions by PCR (VCJ-PCR) in patient 1553. The two viral-cellular junctions identified by sequencing were validated by PCR in the primary tumor (P^{T2}: DNA from frozen sections prepared for TEN16_2016; P^{T1}: TEN16_2013) and metastases (M1-M4). VCJ primers (■) were re-designed to shorten the amplicons (□) for DNA from FFPE tissue (M3 and M4). HPV primers amplifying HPV16 E6/E7 (▲, for frozen biopsies) and L1 (▼, for FFPE tissues) were used as positive controls. In the upper right panel showing VCJ1 detection in M3 and M4, PCR products were reamplified in a second PCR using the same primers.

4.5.5 Localization of viral 3'-breakpoints

Sequence analysis of the VCJs revealed that most HPV16 3'-breakpoints were located in the HPV E1 gene (Figure 21). All patients had at least one 3'-breakpoint in E1 except for patient 1553, who had both 3'-breakpoints in E2. Four of eight patients (1464, 993, 1425, 1529) had additional 3'-breakpoints in the genes E4, E5 or L2 (Figure 21).

The localization of 3'-breakpoints affected viral transcription. Splicing of the transcripts E1C and E1^E4 was impaired in cases with VCJs located only upstream of the splice acceptor sites (at N2582 and N3358, respectively). Loss of those transcripts indicated absence of complete transcriptionally active HPV genomes in the tumors (primary tumor of 661, primary tumor and metastasis 2 of 1553), while detection of E1^E4 despite of upstream VCJs indicated presence of truncated and additional complete transcriptionally active HPV genomes (primary tumor of 1529, Table 24).

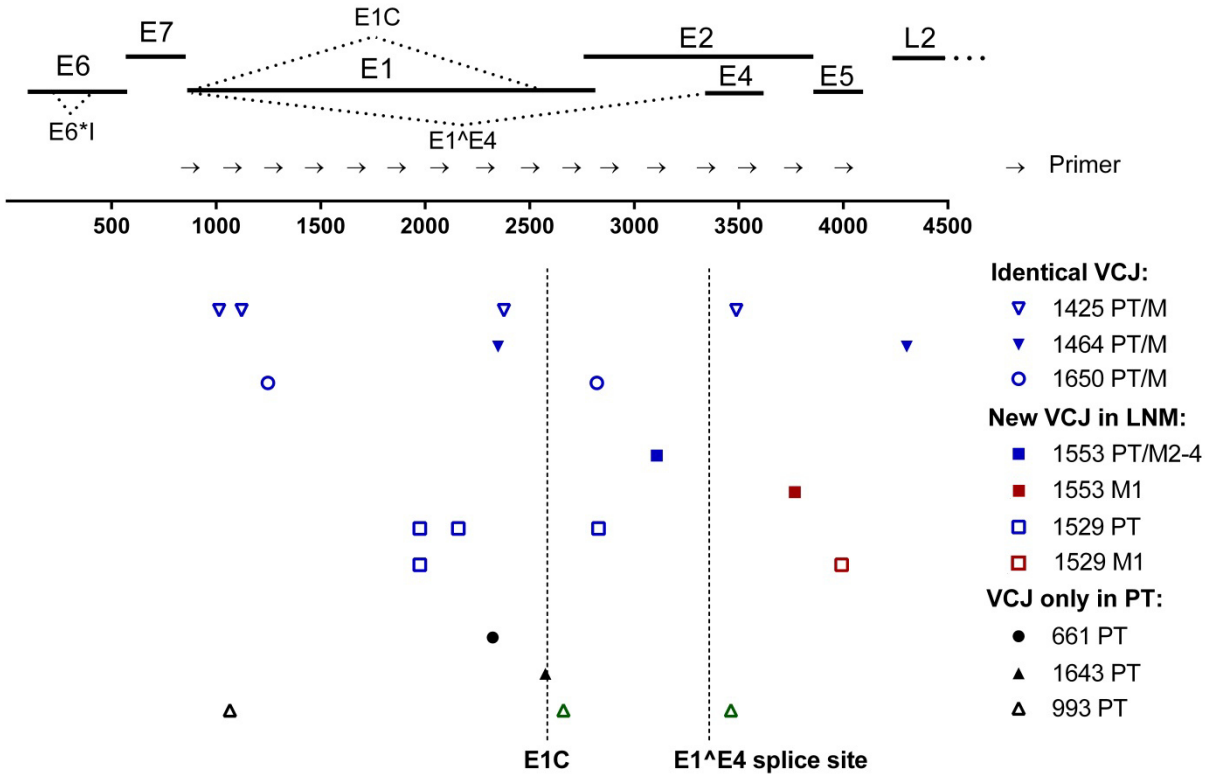


Figure 21: Localization of HPV16 3'-breakpoints. The localization of the viral 3'-breakpoints is represented for every patient. The early region of the HPV16 genome (N0-N4500) is depicted with the open reading frames of the early genes and the positions of the primers (arrows) used for the TEN16 analysis. The HPV16 E1C and E1-E4 splice acceptor sites (at N2582 and N3358, respectively) are depicted as dotted lines. Viral-cellular junctions (VCJs) were present in the pairs of primary tumor (PT) and metastasis (M, blue), only in the primary tumor (black), only in a part of the primary tumor (green) or only in the metastasis (red). The nucleotide positions refer to the HPV16 reference genome NC_001526.3 (National Center for Biotechnology Information).

Table 24: Detection of viral transcripts in relation to viral 3'-breakpoints.

Patient	Tissue	Viral 3'-breakpoint (bp)			HPV Transcripts		
		N0-2582	N2583-3358	N3359-4500	E6*I	E1C	E1^E4
661	Primary Tumor	VCJ1	-	-	+	-	-
	Metastasis	-	-	-	+	+	+
1529	Primary Tumor	VCJ2, 3	VCJ1	-	+	+	+
	Metastasis 1	VCJ2	-	VCJ4	+	+	+
1553	Primary Tumor	-	VCJ1	-	+	+	-
	Metastasis 1	-	-	VCJ2	+	+	+
	Metastasis 2	-	VCJ1	-	+	-	-
993	Primary Tumor	VCJ1	VCJ2	VCJ3	+	+	+
	Metastasis	-	-	-	+	+	+

VCJ=viral-cellular junction, 2582=E1C splice acceptor site, 3358=E1^E4 splice acceptor site

4.5.5 Cellular localization of HPV16 integration sites

Regarding the cellular localization of the integrated HPV, clusters of up to three VCJs within 3 Mb of the cellular sequence were observed in 5/8 (63%) primary tumors. Those clustered VCJs were all (1650, 1425, 1464), partly (1529) or not (993) present in the corresponding lymph nodes (Table 25, further details in Appendix Table A-1). In four (66%) out of six patients with multiple VCJs, all VCJs were located on the same chromosome. In contrast, patient 1553 had VCJs on chromosomes 1 and 22, and patient 993 had two VCJs in close proximity on chromosome 9 and another one on chromosome 1 (Table 25).

While HPV16 integration occurred in most cases within the non-coding region upstream or downstream of cellular genes, the viral genome integrated in two patients within a cellular gene (CD24 in patient 1529 and RERE in patient 993, Table 25). Eleven (58%) out of 19 localized VCJs were in close proximity to cancer-related genes defined by the COSMIC database (catalogue of somatic mutations in cancer) and GeneRIF (Gene Reference into Function) in the NCBI (National Center for Biotechnology Information) database. Of those eleven VCJs, five (45%) were detectable in the primary tumor and in all metastases from the patient and seven (63%) at least in half of the metastases. In contrast, only two (25%) out of eight VCJs that were not located in proximity to cancer-associated genes were present in both the metastasis and the primary tumor (Table 25).

Table 25: Cellular localization of HPV16 integration sites in OPSCC primary tumors and metastases.

Patient	Chromosome Position	VCJ position in relation to cellular genes
1650	8q24.21	**POU5F1B**(+)__VCJ1(-)__VCJ2(+)__**MYC**(+)
1425	9q22.3	**XPA**(-)__VCJ1(+)__VCJ2(+)__**FOXE1**(+)__VCJ3(+)__C9orf156(-)
1553	1p13.3	**DENND2D**(-)__VCJ1(+)__CHI3L2(+)
	22q12.3	CSF2RB(+)__VCJ2(+)__TEX33(-)
1529	6q21	QRSL1(+)__VCJ1(+)__VCJ2(+) within **CD24**(-)__VCJ3(+)__VCJ4(+)__C6orf203(+)
661	5q14.3	**ARRDC3**(-)__VCJ1(-)__NRSF1(+)
993	1p36.23	VCJ1(-) within RERE(-)
	9p24.1	TMEM261(-)__VCJ2(+)__**PTPRD**(-)__VCJ3(+)
1643	20q13.13	**CEBPB**(+)__VCJ1(+)__**PTPN1**(+)
1464	1p21.2	GPR88(+)__VCJ1(-)__VCJ2(-)__VCAM1(+)

VCJ=viral-cellular junction (HPV integration site), (+/-)=DNA plus/minus strand;
Highlighted genes (bold) are cancer-associated genes defined by the COSMIC database (catalogue of somatic mutations in cancer, http://cancer.sanger.ac.uk/cosmic) and GeneRIF (Gene Reference into Function) in the NCBI (National Center for Biotechnology Information) database, accessed: 09/12/2016.
Highlighted VCJ (blue) were present in primary tumor and metastases (1553 VCJ1 in 3/4 metastases, 1529 VCJ2 in 1/2 metastases), or only in a metastasis (red). 1425 VCJ4 is located in repetitive sequences at several positions.

4.5.1 Methylation levels in primary tumors vs. metastases

Pairs of known OPSCC primary tumors (n=4) and corresponding neck lymph node metastases (n=10, 1-4 per patient) were analyzed to assess potential variability in methylation levels of the five gene promoters (described in chapter 4.4) between primary tumors and corresponding metastases.

In three (30%) metastases, methylation levels for all five promoters varied by less than 10% from the levels in the primary tumors. However, in other three (30%) metastases a variation of more than 25% was observed for the promoters of IRX1 and/or ALDH1 in comparison to the primary tumors (Figure 22). Overall, the metastases had significantly lower methylation levels compared to primary tumors (p=0.00003, paired student's t-test). A reproducibility analysis of six metastases tested in two independent assays revealed a mean inter-experimental deviation of only 1%, ranging from 0% to 3% (R^2=0.98).

The single methylation score (0 or 1) calculated for each promoter separately based on a promoter-specific cut-off value was identical in 40% (OSR2), 70% (ALDH1 and GRIA4) and 90% (GATA4 and IRX4) of pairs of primary tumors and metastases. For the total methylation score combining all five promoters in a sample the agreement between the pairs was 90%, since all primary tumors and all corresponding metastases except for one presented with MS≥3.

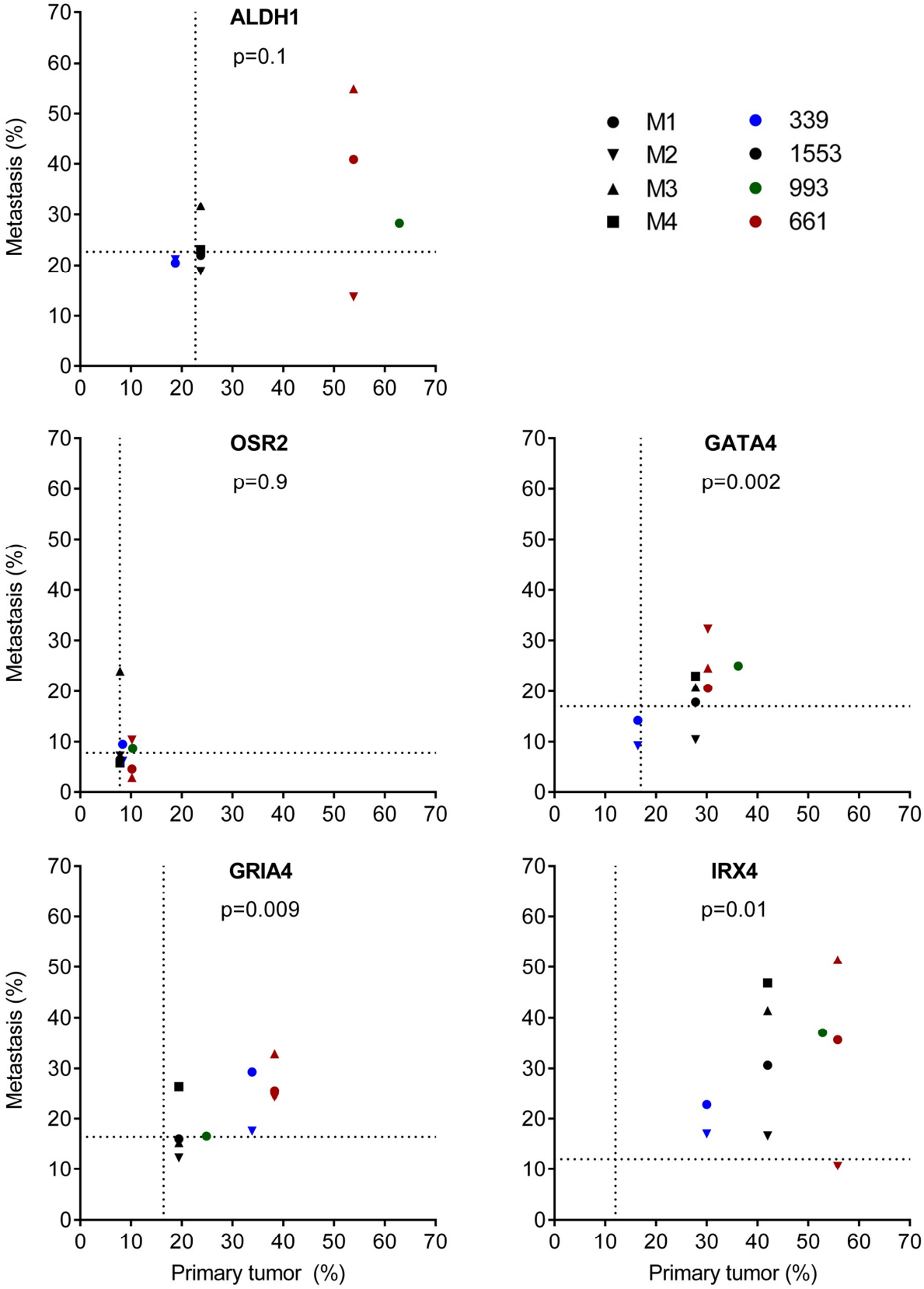


Figure 22: Methylation in the five gene promoters comparing OPSCC primary tumors and corresponding metastases. The percentage of methylation in the promoters of the five genes included in the pre-defined methylation signature is compared for primary tumors versus corresponding lymph node metastases. Promoter-specific cut-off values are represented as dotted line. Up to four metastases (M1-M4) are depicted for each patient (color). Paired student's t-test was used to calculate p values.

5 Discussion

5.1 Detection of HPV-driven NSCCUP

HPV mRNA detection is considered as gold standard to identify HPV-driven OPSCC patients [47, 49] and was performed in this study by HPV E6*I mRNA assay [48]. In comparison to HPV mRNA, HPV DNA had a sensitivity of 100%, but only 95% specificity (Chapter 4.1.3, Table 10). The reduced specificity of HPV DNA alone is supported by $p16^{INK4a}$ IHC data, since three out of the four HPV DNA-positive but mRNA-negative NSCCUP did not show $p16^{INK4a}$ overexpression. A recently published meta-analysis reported a similar sensitivity (98%), but lower specificity (84%) of HPV DNA to identify HPV-driven OPSCC [53]. Low specificity might be due to cross-contamination between different samples during sectioning. In the present study, the risk of cross-contamination was reduced to a minimum by following a very stringent, previously optimized sectioning protocol [49]. Aside from contamination, HPV DNA detected in the absence of HPV mRNA might be derived from past infection, or from recent oral HPV exposure and is probably biologically non-relevant [52]. As previously shown for OPSCC, HPV DNA detection alone is not sufficient to identify HPV-driven tumors [49, 51], and should therefore also not be used as a single marker in NSCCUP patients.

Although HPV16 was the most common HPV type, this study demonstrates the appropriateness of broad-range high-risk HPV testing in NSCCUP, because five (13%) of the 39 overall HPV-driven NSCCUP were driven by non-HPV16 types, including HPV18 (n=1), HPV31 (n=1), HPV33 (n=2) and HPV35 (n=1). Previous NSCCUP studies detected HPV18 (n=2), HPV33 (n=6) and in single cases HPV35, HPV58, HPV73 and HPV82, respectively, together with $p16^{INK4a}$ overexpression [95, 109, 114, 119, 123].

While frozen biopsies are preferable for the detection of nucleic acids, FFPE tissue is needed for $p16^{INK4a}$ expression analysis. In current literature, different cut-off values are applied to evaluate $p16^{INK4a}$ overexpression (reviewed in [53]). In the present study, NSCCUP were defined $p16^{INK4a}$-positive, if >25% of tumor cells were stained. Additional evaluation of the frequently used more stringent cut-off value of >70% showed that in nine (23%) NSCCUP with $p16^{INK4a}$ overexpression, between 25% and 70% of tumor cells were stained. Four of those were HPV-driven due to positivity for HPV mRNA and DNA. In comparison to HPV mRNA, sensitivity of $p16^{INK4a}$ overexpression using the 25% cut-off value was higher and specificity was slightly lower (96% and 89%, respectively) compared to the 70% cut-off value (81% and 93%, respectively, Chapter 4.1.3, Table 10).

Specificity of $p16^{INK4a}$ overexpression as a surrogate marker to identify HPV-driven HNSCC is generally low, because $p16^{INK4a}$ overexpression might be triggered not only by HPV transformation, but also by non-viral mechanisms [47]. Beadle and colleagues showed strong $p16^{INK4a}$ expression in 12.5% (3/24) of lymph node metastases from cutaneous HNSCC that were unassociated with HPV, since all three metastases were HPV DNA-negative [142]. Thus, assuming oropharyngeal localization of primary tumors only based on $p16^{INK4a}$ overexpression in the metastases might be misinterpretation [142]. In the present study, 14 (36%) of the 39 $p16^{INK4a}$-positive cases were HPV mRNA-negative. Therefore, $p16^{INK4a}$ overexpression should not be used as a single marker in NSCCUP patients.

From the three markers analyzed in this study, we can conclude that $p16^{INK4a}$ overexpression together with presence of HPV DNA [51] would be the best alternative to identify HPV-driven NSCCUP cases, if RNA analysis cannot be performed. This marker combination would increase specificity in the present study to 99% while maintaining 96% sensitivity, if applying the 25% $p16^{INK4a}$ cut-off value (Chapter 4.1.3, Table 10). With the 70% cut-off value specificity would further increase to 100%, but sensitivity would be reduced to 81%.

5.2 Nucleic acid detection in FFPE tissue

While HPV DNA could be easily detected in DNA extracted from frozen biopsies, FFPE samples turned out to be challenging. By MPG, neither HPV nor the cellular control gene beta-globin could be detected in 36% of FFPE samples included in the multicenter study even after improved DNA extraction. This high percentage of invalid samples demonstrates the importance of internal controls for DNA integrity to avoid false-negative results. The amplicon size is crucial for DNA detection. Of the DNA samples extracted from FFPE tissue, 59 were MPG-invalid, meaning negative for HPV (amplicon sizes around 150 bp) and beta-globin (208 bp amplicon), but in 22 (38%) of them between 100 and 1300 copies of beta-globin were detected by real-time qPCR using a 110 bp amplicon. Taking this into account, short amplicons (<200 bp) were designed for methylation analyses and for validation PCRs of viral-cellular junctions in HPV16 integration analyses in this study.

The HPV E6*I mRNA assay had been developed especially for RNA samples extracted from FFPE tissue and uses an amplicon of around 65 bp for HPV mRNA and a cellular ubiquitin C amplicon of 85 bp for validity analysis [48]. In this study, all samples were mRNA-valid, indicating a high technical sensitivity of the mRNA detection assay.

5.3 Prevalence of HPV-driven NSCCUP

5.3.1 Geographical differences

The prevalence of HPV-driven NSCCUP revealed by the multicenter retrospective study ranged from 10% to 20% and reached 28% in the HPV serology study. The lowest prevalence was measured in Barcelona (10%), which is in line with a previously reported low HPV prevalence in OPSCC patients. In another retrospective study, only 3% of OPSCC patients diagnosed between 1990 and 2009 in northern Spain were HPV DNA/p16^{INK4a}-positive [143]. Systematic reviews and a large international study report that Spain is among the countries with the lowest HPV prevalence (6-9%) in OPSCC [65, 67, 144].

The prevalence of HPV-driven NSCCUP was higher in the Treviso region (17%) and in Heidelberg (20%), but lower than reported by previous Italian (45%) and German studies (22%-85% DNA/p16^{INK4a}-positive) [118, 120, 123, 124]. This might be due to (i) early sampling in this study (starting 1988), (ii) identification of truly HPV-driven NSCCUP cases and (iii) the low proportion of patients initially presenting as NSCCUP, but later being diagnosed with tonsillar or base of tongue carcinoma. Those patients were overrepresented (60%, 100% and 100%, respectively) in two German studies reporting 81% and 85% HPV DNA/p16^{INK4a}-positivity, respectively [95, 120], as well as in an US study with 69 patients reporting the highest HPV prevalence (91% p16^{INK4a} overexpression and/or HPV ISH) [145]. The only previous study detecting HPV mRNA revealed a prevalence of 45% in 22 NSCCUP diagnosed between 2010 and 2012 in Rome [146]. Another German study analyzed 63 recently diagnosed NSCCUP patients (2002-2011), of which 37% were HPV DNA/p16^{INK4a}-positive [123]. The increased HPV prevalence in those studies compared to the present study might result from differences in the sampling period, since beside regional differences in the prevalence of HPV-driven NSCCUP also a time trend was observed in the present study (Chapter 4.1.2, Figure 8).

5.3.2 Increase in prevalence of HPV-driven NSCCUP over time

The relative prevalence of HPV-driven NSCCUP in recently diagnosed patients (2005-2014) was 2.5-fold increased, compared to those diagnosed earlier (1988-2004). The increase was observed for all three study centers ranging from 1.6-fold in Barcelona, to 2.8-fold in Heidelberg and up to 3.3-fold in Treviso, ruling out regional bias. This time trend might at least partly explain the regional differences observed in this study, because cases from Barcelona with the lowest HPV prevalence were diagnosed longer ago (median year of diagnosis: 2002), while cases from Heidelberg with the highest HPV prevalence were diagnosed more recently (median: 2008). The highest HPV prevalence (28%) was measured

by HPV serology in NSCCUP patients from Leipzig, who were diagnosed even more recently (2008-2016, median: 2011).

The increase in prevalence of HPV-driven NSCCUP in recent decades is in line with the increase observed for HPV-driven OPSCC [147]. A systematic review reported a steady increase of HPV prevalence in OPSCC from pre-1995 till 2013, more precisely from 28% to 50% (1.8-fold) in Europe, that was independent of the HPV detection method and whether FFPE or frozen tissue was analyzed [144]. However, it is discussed whether the HPV prevalence is absolutely increasing, or relatively increasing due to the decreasing proportion of smoking- and alcohol-related cancers in men in Western countries [68]. While the relative prevalence of HPV-driven NSCCUP cases increased in the present study (p=0.004), the proportion of tobacco and alcohol consumers among NSCCUP patients decreased in the same period (p=0.008 and p=0.006, respectively) [88]. However, even in the late sampling period, smoking was still very common (77%), and in both sampling periods the majority of patients with HPV-driven NSCCUP were also smokers (64%), pointing at a potential interaction between HPV and smoking [60].

5.4 Survival benefit of patients with HPV-driven NSCCUP

Although statistical power was limited due to the low number of events (death or progression) among patients with HPV-driven NSCCUP, the multicenter and the serology study revealed both a significantly better overall ($p<0.004$) and progression-free survival ($p<0.02$, Chapter 4.1.5, Figure 12 and Chapter 4.2.4, Figure 15) of patients with HPV-driven or HPV-seropositive NSCCUP, emphasizing the clinical relevance of the HPV status. This is in line with previous NSCCUP studies reporting significantly better 4-year or 5-year overall and progression-free survival rates (Chapter 1.3.5, Table 4) for HPV-positive patients defined by presence of HPV DNA and/or $p16^{INK4a}$ overexpression [109, 115, 121].

The survival benefit observed in the multicenter and the HPV serology study remained significant even after adjusting for potential confounders (Chapter 4.1.5, Table 12 and Chapter 4.2.4, Table 17) taking into consideration that patients with HPV-driven NSCCUP presented with significantly less risk factors, meaning less alcohol and tobacco consumption and less advanced N stage (Chapter 4.1.4, Table 11 and Chapter 4.2.1, Table 14). The hazard ratios calculated in both studies for overall survival (0.30 and 0.09, $p\leq0.001$, respectively) and progression-free survival (0.27 and 0.03, $p\leq0.003$, respectively) were similar or lower compared to previous studies. For overall survival, a Swedish study including 50 NSCCUP patients revealed a similar hazard ratio of 0.29 ($p=0.03$) for $p16^{INK4a}$-positive patients after adjusting for age at diagnosis, gender, tobacco consumption and p53 expression [109], while in a Danish study with 60 patients the hazard ratio was 0.71

(p=0.009) for $p16^{INK4a}$-positive patients after adjusting for age at diagnosis, tobacco and alcohol consumption and non-keratinizing morphology [122]. For disease-free survival, a Korean study analyzing 58 NSCCUP patients revealed for $p16^{INK4a}$-positive patients a hazard ratio of 0.29 (p=0.03) after adjusting for extracapsular spread and p53 expression, which is similar to the hazard ratio for progression-free survival in the presented multicenter study [115]. However, comparison to previous studies is limited due to the lack of HPV mRNA assessment.

5.5 Additional factors influencing survival of NSCCUP patients

Besides HPV status, multimodal treatment was revealed as independent prognostic factor, although in the multicenter study interpretation is complicated by the time-dependent effect implying that the hazard ratio varies over time. But also in the HPV serology study, a significant survival benefit of NSCCUP patients receiving postoperative chemoradiotherapy was observed (Chapter 4.2.4, Table 17).

Increased age at diagnosis, presence of extracapsular spread and advanced nodal stage were revealed as independent unfavorable prognostic factors (Chapter 4.1.5, Table 12 and Chapter 4.2.4, Table 17). Regarding prognosis, those should be considered in addition to HPV. This is in concordance with previous studies confirming age at diagnosis and extracapsular spread as unfavorable prognostic factors [115, 122]. In the updated TNM staging system (AJCC Cancer Staging Manual, 8th edition, January 2017) extracapsular spread is taken into consideration, since NSCCUP patients with extracapsular spread are classified into the highest nodal stage (N3b).

In addition, increased p53 expression was reported to be an unfavorable prognostic factor in NSCCUP patients [109, 115]. Expression levels were shown to be increased in HNSCC patients with missense TP53 mutations and reduced in patients with nonsense mutations [148]. TP53 mutation status was assessed in the multicenter study, but only in 70 patients and was therefore not included in the multivariate analysis. As described for OPSCC [72, 73], all HPV-driven NSCCUP presented without disruptive TP53 mutation, which might explain their survival benefit due to an increased radiosensitivity [149]. However, in the absence of HPV, patients with disruptive TP53 mutation showed a slightly better survival than those without disruptive mutation (Figure 17, Chapter 4.3.3), which might be due to the fact that patients without disruptive mutation presented with significantly advanced N stage. Sequencing of additional samples and multivariate survival analysis is required and the effect of the identified mutations on p53 expression levels should be investigated in order to get more insights into the role of TP53 in NSCCUP patients.

For gender, tobacco and alcohol consumption no significant effect on survival was observed (Chapter 4.1.5, Table 12 and Chapter 4.2.4, Table 17), which might be due to limited statistical power of the two studies. Although the multicenter study is the largest NSCCUP study published so far, confidence intervals are still large due to low number of events (death or progression) among females, never smokers and never drinkers with HPV-driven NSCCUP in the multicenter study (n=2, 2 and 4, respectively) and the HPV serology study (n=0, 1 and 0, respectively).

5.6 Epigenetic characteristics of HPV-driven lymph node metastases

The survival benefit of HPV-driven HNSCC is linked to epidemiological, molecular and clinical differences compared to non-HPV-driven HNSCC. On the molecular level, Kostareli et al. reported epigenetic differences in HPV-driven vs. non-HPV-driven HNSCC and identified a prognostic methylation signature [76, 77]. A survival benefit has been reported for OPSCC and other HNSCC patients presenting with high agreement to the signature [76, 77], but tumors outside the head and neck have not yet been studied. The multicenter study showed that this methylation signature is also present in HPV-driven NSCCUP (Chapter 4.4.1, Figure 18). However, it did not seem to have a prognostic role in NSCCUP patients. Among the analyzed NSCCUP, an unknown proportion of metastases might originated from primary tumors outside the head and neck region, such as from the lung. In those patients, the prognostic value of the methylation signature might be different.

Furthermore, a technical issue should be considered to influence the methylation analysis. Since methylation levels are measured in all cells of the tissue section, the tumor content should ideally approach 100%. Some samples had a lower tumor content (range: 10-100%), although the median was 90%. Non-malignant cells such as lymphocytes present in the section might change the percentage of methylation measured in the samples.

It has not been investigated before, whether the methylation signature and its prognostic value observed in OPSCC also apply to lymph node metastases from oropharyngeal primary tumors. In the present study, paired analysis of HPV-driven OPSCC and corresponding metastases revealed significantly reduced methylation levels in metastases (Chapter 4.5.1, Figure 22). This implies that the methylation score might not be directly transferable from primary tumors towards metastases. Instead, adjustment of the cut-off values might be necessary. Lower methylation might result in higher expression levels. For the gene ALDH1, lower methylation might explain the previous observation that in OPSCC ALDH1 expression was associated with advanced nodal stage (p=0.02) and was significantly increased in metastases compared to primary tumors (p=0.01) [150].

Taking into consideration those limitations of the analysis, in particular potentially imprecise cut-off values, we regrouped NSCCUP in those with high agreement to the pre-defined methylation signature (methylation in 4-5 promoters accordingly), low agreement (0-1 promoters) and an intermediate group (2-3 promoters) in order to assess differences between the two extremes (Appendix, Figure A-2). Indeed, separation of the curves was visible and a statistically significant difference for progression-free survival was revealed by log-rank test. Therefore, the intermediate group that is close to the applied cut-off value (≥3 promoters) might be at least partly misclassified in this study and might distort the survival analysis. However, HPV will probably remain the stronger prognostic marker in NSCCUP patients.

5.7 Stability of HPV markers and integration upon metastasis formation

While methylation levels might vary between primary tumors and metastases, presence of HPV DNA and E6*I mRNA was consistent between all pairs analyzed in this study. However, the viral load differed between primary tumors and corresponding metastases (Chapter 4.5.1, Table 21), but this might be biased by the tumor content, since an overestimated tumor content would result in underestimated viral load and vice versa. Although the tumor content was determined by an experienced pathologist, this is challenging due to presence of small lymphocytes in the lymph node metastases.

HPV integration site analysis revealed differences, since not all integration sites found in the primary tumors were present in the corresponding metastases (Chapter 4.5.4, Table 22). The fact that in one primary tumor two integration sites were only found in one, but not in another section of the same tumor, indicates intratumoral heterogeneity (Figure 23). A primary tumor might consist of several distinct subclones with different HPV integration sites, but metastatic cells would disseminate and expand from one subclone.

In patients with NSCCUP and HPV-driven OPSCC patients, metastasis formation is assumed to be an early event, as described for breast cancer by the early dissemination and parallel progression model [151]. If the HPV integration site of the primary tumor is not present in the corresponding metastasis, integration might either be a rather late event occurring after early formation of metastases, or more likely it might be unstable over time (Figure 23). If a random integration event was not beneficial for the viral transformation or proliferation of the cell, this specific integration site might get lost in the metastasis and even in the primary tumor over time. This hypothesis is supported by the observation that VCJs in close proximity to cancer-related genes were more stable than others (Chapter 4.5.5, Table 25). Further studies are needed in order to investigate, how HPV integration affects the neighboring cellular genes.

Specific HPV integration sites might be used to confirm origination of a metastasis from a primary tumor harboring the same integration site. However, different HPV integration status cannot exclude that the metastasis originated from that particular primary tumor.

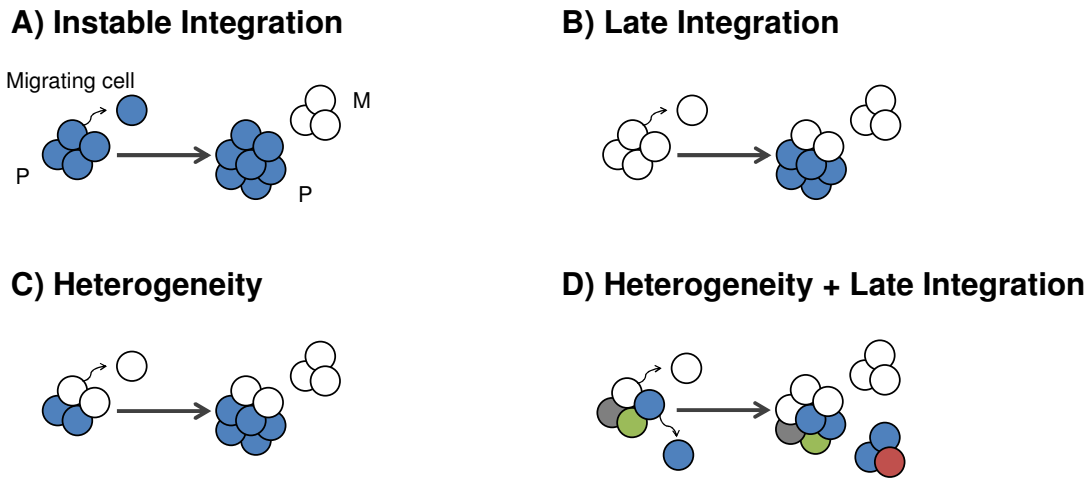


Figure 23: Scenarios for presence or absence of viral-cellular junctions in primary tumors and corresponding lymph node metastases. Three possible scenarios (A, B, C) might explain presence of viral-cellular junctions (blue, green, grey) only in the primary tumor (PT), but not in the corresponding lymph node metastasis (M). Scenario C assumes heterogeneity regarding HPV integration within the primary tumor and explains detection of another viral-cellular junction (red) only in the lymph node metastasis.

5.8 HPV detection methods suitable for clinical settings

With regard to the rising prevalence of HPV-driven NSCCUP and the prognostic value, HPV assessment is worthwhile to be included in the standard diagnostic work-up of NSCCUP patients. This requires a cost-effective, time-saving and easy-to-perform assay for clinical settings that is validated against the gold standard to identify HPV-driven HNSCC accepted in research settings, which is the detection of HPV mRNA.

5.8.1 Serological HPV status assessment

HPV serology might be particularly suitable for HPV status diagnostics in NSCCUP patients in clinical settings, because the HPV antibodies can be detected in few microliters of serum or plasma, which is routinely drawn for clinical purposes. Since the complex serological high-throughput assay used in this study needs to be further adapted for clinical settings, an HPV16 E6 antibody ELISA (enzyme-linked immunosorbent assay) has been developed aiming at comparable performance with more convenient handling in less specialized clinical laboratories [Viarisio et al, oral communication].

In the HPV serology study, it was demonstrated for the first time that HPV-driven NSCCUP patients produce HPV antibodies. Ten (22%) of the 46 NSCCUP patients had high

HPV16 antibody levels, which were in all cases in agreement with presence of HPV16 DNA, mRNA and $p16^{INK4a}$ overexpression. This is in line with the high sensitivity and specificity (96% and 98%, respectively) previously determined for HPV16 serology to identify HPV16-driven OPSCC and with the HPV16 antibody patterns observed in patients with HPV16-driven OPSCC [54].

For other high-risk HPV types, the serological HPV status was at least partly discordant in comparison to the molecular HPV status. One case with HPV31-driven metastasis had high antibody levels against HPV31 E6 and E7, but also showed cross-reactivity with HPV35 E6 and E7. Cross-reactivity was also seen for the ten HPV16-seropositive cases in this study against E7 of HPV31 (n=7), HPV33 (n=7) and HPV35 (n=6) and is expected due to the close phylogenetic relationship between those HPV types [152]. Another case with HPV33-driven metastasis had high antibody levels against HPV33 E7 but not against E6, and was thus defined HPV-seronegative. Absence of antibodies against HPV16 or HPV18 E6 despite of presence of antibodies against the other early proteins E7, E1 and E2, was shown for single patients with HPV-driven NSCCUP in this study (Chapter 4.2.1, Figure 13) and in a previous OPSCC study [54]. Analysis of further NSCCUP and OPSCC cases is required in order to optimize cut-off values and to better define the serology algorithms to identify cases driven by HPV types other than HPV16.

5.8.2 HPV E6 protein detection in FNAB

Another promising candidate for HPV status diagnostics in NSCCUP patients is the commercially available OncoE6^{TM} Oral Test, which was tested in this study. The fast and simple procedure compared to other HPV detection methods makes it suitable for clinical settings. In HPV16-driven frozen biopsies from two NSCCUP patients, detection of the HPV16 E6 protein using the OncoE6^{TM} Oral Test was demonstrated. A result could be obtained within three hours after sampling.

In a previous study, the assay detected HPV16 E6 in 32 (97%) out of 33 HPV16-driven OPSCC, but in none of the 24 non-HPV-driven OPSCC [Holzinger et al., in preparation]. In an ongoing study with the collaborators Dr. Paolo Boscolo-Rizzo and Dr. Elisa Dal Cin in Treviso, we are currently investigating, whether the OncoE6^{TM} Oral Test might be used to detect HPV16/18 E6 in fine-needle aspiration biopsies (FNAB) from lymph node metastases in OPSCC and NSCCUP patients with good concordance to molecular HPV markers and HPV antibodies. FNAB are routinely taken from patients with NSCCUP in most institutions and would represent a minimally invasive diagnostic method to identify patients with HPV-driven NSCCUP, which might be useful for guiding diagnostic work-up and treatment [110]. A previous study demonstrated detection of HPV16 by ISH together with

$p16^{INK4a}$ overexpression in FNAB from 10/19 (53%) metastases from oropharyngeal tumors, 0/46 (0%) from other primary sites and from 3/10 (30%) patients with NSCCUP, in which HPV might be an indicator for oropharyngeal primary tumor [116]. The HPV status in the FNAB of NSCCUP patients was reported to predicted an oropharyngeal primary tumor with 90-100% specificity, but only 50-70% sensitivity [110].

5.9 Potential clinical implications of HPV detection in NSCCUP

HPV status assessment might become clinically relevant, as it may help to target the search for the primary tumor and to target therapy in patients with HPV-driven NSCCUP, which may result in better prediction of the clinical course and may improve quality of life [88, 115-120].

5.9.1 Localization of the primary tumor

The multicenter study included 13 patients, in which a potential primary tumor was identified within three years after initial NSCCUP diagnosis (Chapter 4.1.6, Table 13), including two cases with an HPV-driven metastasis. One patient had an occult HPV-driven base of tongue carcinoma, which probably was missed when taking biopsies at the time of NSCCUP diagnosis, but was identified 18 months later. The other patient had an occult nasopharyngeal carcinoma diagnosed 14 months later (Chapter 4.1.6, Table 13). While the metastasis was EBV-negative, no tissue was available to determine the HPV or EBV status of the nasopharyngeal carcinoma. Although nasopharyngeal carcinomas are mainly associated with EBV, also HPV16 or HPV18 DNA/$p16^{INK4a}$-positive cases were reported in recent studies [153-156]. A large international archival HNSCC tissue study identified 5.9% (6/101) HPV-driven nasopharyngeal carcinomas [67]. However, oropharyngeal carcinomas might also extend to the nasopharynx due to close proximity, and might be wrongly diagnosed as nasopharyngeal carcinomas [156].

While in previous studies, patients initially presenting with HPV DNA/$p16^{INK4a}$-positive NSCCUP likely had primary tumors in the oropharynx [115, 118, 120], those occult HPV-driven OPSCC were underrepresented in this study. This might be due to the difficulty of identifying initial NSCCUP patients from clinic databases and the NCT cancer registry, since patients are reclassified after primary tumor diagnosis changing the ICD code (International Classification of Diseases) from C80 towards the ICD code describing the localization of the primary tumor. Nevertheless, an extended diagnostic work-up of the oropharynx might be indicated for patients with HPV-driven metastasis [114].

While occult tonsillar carcinomas can be efficiently resected by diagnostic tonsillectomy, even when histopathological analysis does not identify the primary tonsillar carcinoma, small occult base of tongue carcinomas can be frequently missed by routinely performed biopsies.

Resection of the base of tongue is hardly performed without complications and functional loss, such as impaired ability to swallow. Recently, an innovative tool, transoral robotic surgery (TORS), has been developed that facilitates resection of the base of tongue. The surgeon controls the instruments at a console allowing precise freedom of motion and improved video-assisted visualization [157]. In a US study, small occult base of tongue carcinomas were identified by TORS in nine (90%) out of ten NSCCUP patients [158]. The mean diameter of the carcinomas was 0.9 cm, the smallest was only 0.2 cm. Eight of the nine occult carcinomas were HPV16-positive, suggesting that this tool might be in particular beneficial for patients with HPV-driven NSCCUP.

5.9.2 De-intensified treatment

Besides indicating extended diagnostic work-up of the oropharynx, HPV status might also guide treatment decisions in NSCCUP patients. With regard to the significant survival benefit observed in the present study but also in published studies [90, 109, 115, 118, 121, 122] and taking into account the rather broad treatment and treatment-associated toxicity, it is discussed whether patients with HPV-driven NSCCUP might benefit from treatment de-intensification. In is currently assessed in patients with HPV-driven HNSCC, whether similar effects might be achieved with de-intensified treatment, while improving tolerance and post-treatment quality of life [81]. Toxicity might be minimized by focusing treatment to the oropharynx, or by reducing the radiation dose based on the increased radiosensitivity of HPV-driven HNSCC [159]. The phase II trial NCT01530997 found that de-intensification of chemoradiotherapy by using intensity-modulated radiation therapy and reducing the dose from 70 Gy to 60 Gy decreased toxicity, while maintaining a high pathologic complete response rate in HPV DNA-positive or $p16^{INK4a}$-positive OPSCC [160]. Randomized prospective treatment studies including NSCCUP patients have not yet been performed and will be difficult to conduct because of the low incidence. But since patients with HPV-driven NSCCUP can be strongly suspected to harbor occult OPSCC, evidence can be extrapolated from those studies.

The challenge of de-intensification treatment is to choose the appropriate subset of patients. Selecting patients with truly HPV-driven NSCCUP is essential. However, for several recent de-intensification trials patients were selected based on $p16^{INK4a}$ overexpression alone [81], which in the present study and in many OPSCC studies showed a rather low specificity involving the risk of undertreating the subgroup of patients with non-HPV-driven NSCCUP overexpressing $p16^{INK4a}$. Not only HPV status, but also unfavorable prognostic factors, such as extracapsular spread and advanced nodal stage, should be considered, as indicated in this study and in a few small published series [115, 122]. In patients with HPV-driven NSCCUP presenting with those characteristics, rather intensive treatment might be indicated.

5.10 HPV serology as follow-up marker

After treatment, patients are regularly invited for follow-up visits in order to check for relapse and particularly for outgrowth of a primary tumor not eradicated by treatment. The specific detection of HPV antibodies in patients with HPV-driven NSCCUP raised the interest in assessing whether kinetics of HPV antibody levels might be used as a marker to monitor successful tumor resection versus recurrence in NSCCUP patients with HPV-driven metastases. Indeed, in all five analyzed NSCCUP patients antibody levels decreased within the first 6-18 months after surgical removal of the metastasis by neck dissection and increased again in the single patient at the time of relapse (Chapter 4.2.5, Figure 16). This was consistent to the results obtained for patients with HPV-driven OPSCC, in which the main decrease in HPV16 antibody levels was observed within the first six months after treatment [Broglie et al., oral communication]. In the present study, follow-up serum samples drawn between three to nine months after treatment were only available for three patients, of which two showed a strong decrease in antibody levels (Chapter 4.2.5, Figure 16).

However, in patients with HPV-driven NSCCUP, antibodies to HPV early proteins are generally produced at high levels and in patients followed-up over years after removal of the metastasis, antibody levels seem to be stable after the initial drop (Chapter 4.2.5, Figure 16). The patient with occult HPV-driven base of tongue carcinoma showed a decrease of HPV antibody levels during the first year after diagnosis and initial treatment that was similar to two patients without subsequent discovery of the primary tumor (Chapter 4.2.5, Figure 16). This is probably due to the fact that multiple metastatic ipsilateral lymph nodes (N stage 2b) were removed by neck dissection, while the occult base of tongue carcinoma must have been so small at that time that it was not detected by PET/CT and base of tongue biopsies. Together with the lymph nodes also plasma cells producing HPV antibodies might have been removed, resulting in a decrease in HPV antibody production despite presence of the HPV-driven base of tongue carcinoma.

While giving insights into HPV immune response and antibody stability, the results from the few cases analyzed are not strong enough to suggest the use of HPV antibody levels as follow-up marker in patients with HPV-driven NSCCUP. Analysis of additional cases and with closer follow-up time points is required to better characterize the HPV antibody decay. In particular, HPV antibody kinetics between treatment and relapse should be further studied.

5.11 Immune response against HPV

Presence of HPV antibodies indicates existence of an adaptive, HPV-induced immune response, which is discussed as one of the explanations for the better prognosis of patients

with HPV-driven HNSCC [19]. On the one hand, CD4+ and CD8+ T cells reacting to HPV16 E6 and E7 were previously found in six out of eight HPV DNA-positive OPSCC, in one out of two analyzed tumor-draining lymph nodes, and in the peripheral blood [161]. Furthermore, Ward et al. showed that the majority (84%) of 149 patients with HPV DNA/$p16^{INK4a}$-positive OPSCC had moderate or high levels of tumor-infiltrating lymphocytes, which were associated with a better disease-specific survival [162].

On the other hand, several immune escape strategies were described [19]. Critical components of the HLA class I antigen-processing machinery were shown to be downregulated in HPV-driven HNSCC tissue compared to adjacent normal squamous epithelium [163]. Moreover, the immune checkpoint regulator PD-L1 seems to play a role, but published results are controversial. While Lyford-Pike et al found significantly higher PD-L1 expression levels in HPV DNA/$p16^{INK4a}$-positive compared to HPV-negative HNSCC tissues (70% vs. 29%), a recent study reported similar PD-L1 expression in $p16^{INK4a}$-positive compared to $p16^{INK4a}$-negative OPSCC (71% and 61%, p=0.274) without significant association with survival [164]. Blockade of the PD-1/PD-L1 pathway is currently evaluated as promising strategy for immunotherapy [165]. In a pilot study in collaboration with Prof. Dr. Jochen Hess, PD-L1 expression levels will be analyzed in a subset of NSCCUP patients from the present study and will be correlated with HPV status in order to evaluate the potential use of immunotherapy in NSCCUP patients. A better understanding of the HPV immune response and immune escape mechanisms in NSCCUP patients is necessary in order to define prognostic markers and potential targets for immunotherapy.

5.12 Conclusion and outlook

The increasing HPV prevalence in NSCCUP patients suggests inclusion of HPV status assessment in the standard diagnostic work-up. If HPV mRNA detection is not feasible in clinical settings, detection of HPV DNA should be combined with $p16^{INK4a}$ expression, since the single markers showed low specificity and sensitivity to identify patients with HPV-driven NSCCUP. HPV serology might be a promising alternative, since the detection of HPV antibody patterns in serum from NSCCUP patients showed high sensitivity and specificity compared to the molecular HPV status in the metastasis. There is a need for simple and cost-effective assays particularly suited for clinical settings, which however need careful validation against the gold standard in prospective studies.

NSCCUP patients with HPV-driven lymph node metastases showed a better prognosis, which might at least partly be due to an HPV-induced immune response, absence of mutations in the TP53 tumor-suppressor gene and altered promoter methylation levels. Those patients should undergo an extended work-up of the oropharynx to reveal occult

primary tumors and they might benefit from de-intensified treatment, which needs to be investigated in clinical trials.

Although HPV markers were concordantly present in pairs of oropharyngeal primary tumors and lymph node metastases, differences were observed regarding HPV integration sites and promoter methylation levels, indicating potential changes upon metastasis formation. Further analysis would be necessary to investigate heterogeneity between primary tumors and corresponding metastases, as well as intratumoral heterogeneity.

,

Abbreviations

ATP	adenosine triphosphate
bp	base pairs
BS	broad-spectrum
°C	degree Celsius
CDK	cyclin-dependent kinase
CI	confidence interval
cm	centimeter
CT	computed tomography
DAB	3,3'-diaminobenzidine
DFS	disease-free survival
DKFZ	Deutsches Krebsforschungszentrum
DNA	deoxyribonucleic acid
DSS	disease-specific survival
E1C	N-terminal truncated E1
E6AP	E6-associated protein
EBV	Epstein-Barr virus
EDTA	Ethylenediaminetetraacetic acid
ELISA	enzyme-linked immunosorbent assay
ENT	Ear-Nose-Throat
FFPE	formalin-fixed paraffin-embedded
FNAB	fine-needle aspiration biopsy
g	gram
HE	hematoxylin and eosin

HNSCC	head and neck cancers are squamous cell carcinomas
HPV	human papillomavirus
HR	hazard ratio
HSV	herpes simplex virus
hTERT	human telomerase reverse transcriptase
IARC	International Agency for Research on Cancer
ICO	Institut Català d'Oncologia (Catalan Institute of Oncology)
IHC	immunohistochemistry
ISH	in situ hybridization
LCR	long control region
M	molar
MCV	Merkel cell polyomavirus
MFI	median fluorescence intensity
min	minutes
ml	milliliter
mM	millimolar
MP96	MagNA Pure 96
MPG	multiplex HPV genotyping assay
MRI	magnetic resonance imaging
mRNA	messenger ribonucleic acid
µl	microliter
µm	micrometer
n	number
N1234	Nucleotide sequence position 1234
N/A	not available

NCBI	National Center for Biotechnology Information
NCT	National Center of Tumor Diseases
ng	nanogram
nm	nanometer
NSCCUP	neck squamous cell carcinoma from unknown primary
OPSCC	oropharyngeal squamous cell carcinomas
ORF	open reading frame
OS	overall survival
PBS	phosphate buffered saline
PCR	polymerase chain reaction
PD-L1	programmed death-ligand 1
PET	positron emission tomography
PFS	progression-free survival
pRb	retinoblastoma protein
qPCR	quantitative PCR
RNA	ribonucleic acid
RT-qPCR	reverse transcription quantitative PCR analysis
SDS	sodium dodecyl sulfate
sec	seconds
Strep-PE	Streptavidin-R-Phycoerythrin
TEN16	Tagging, Enrichment and Next-generation sequencing of HPV16
TMAC	tetramethylammonium chloride
TORS	transoral robotic surgery
TP53	Tumor suppressor protein p53
Tris	tris(hydroxymethyl)aminomethane

URR	upstream regulatory region
US, USA	United States (of America)
VCJ	viral-cellular junction
vs	versus

Appendix

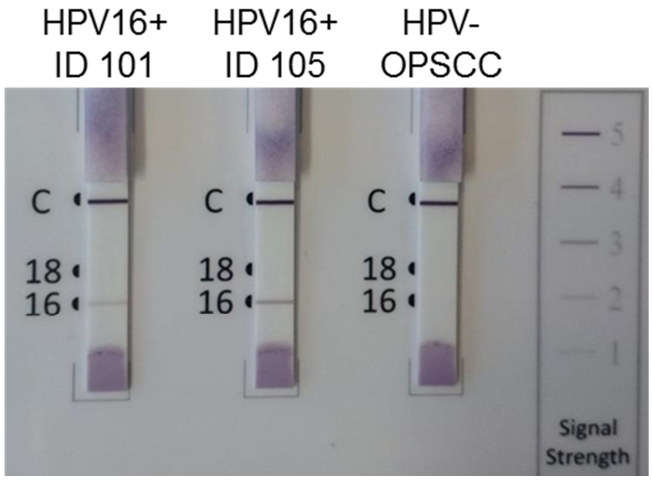
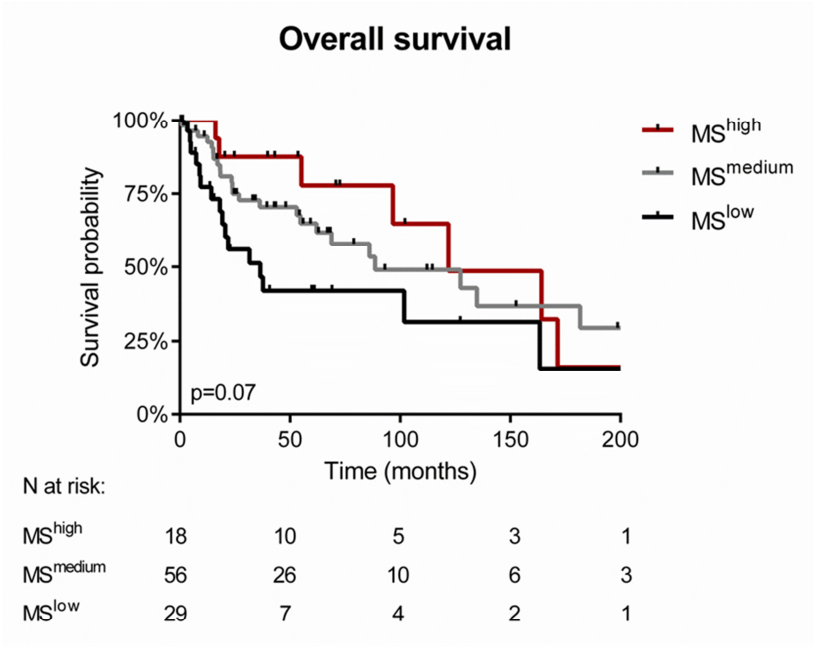


Figure A-1: HPV16 E6 protein detection

The commercially available OncoE6^{TM} Oral Test (Arbor Vita Corporation, Fremont, CA, USA) was used to detect HPV16 E6 in two HPV16-driven (HPV16+) frozen biopsies from NSCCUP patients (ID 101, ID 105). A lysate from a non-HPV-driven (HPV-) oropharyngeal squamous cell carcinoma (OPSCC) was included as negative control. Assay validity is indicated by visibility of the internal positive control line (C) at the top. The lowest purple line (16) indicates presence of the HPV E6 protein in the lysate. Signal strength was evaluated based on the reading guide template shown at the right.



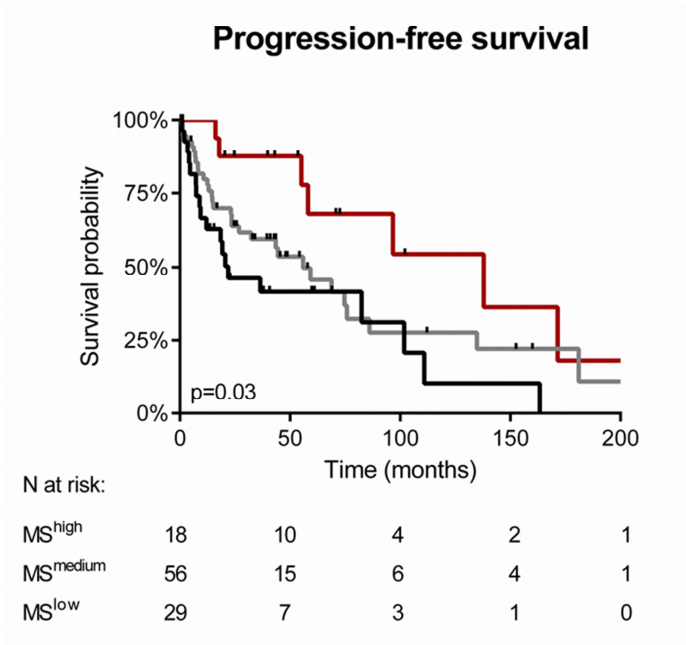


Figure A-2: Survival of NSCCUP patients in relation to MS (high vs. intermediate vs. low)

Methylation levels were measured in NSCCUP to group patients according to presence of a pre-defined promoter methylation signature (low methylation in the promoters of ALDH1 and OSR2, and high in GATA4, GRIA4 and IRX4 based on pre-defined cut-off values). Of the five promoters analyzed, 0-1 (MS^{low}, black), 2-3 (MS^{medium}, grey), or 4-5 (MS^{high}, red) were according to the methylation signature. Kaplan–Meier curves for overall (upper panel) and progression-free survival (lower panel) are compared for NSCCUP patients grouped accordingly. Log-rank test was used to calculate p values.

Table A-1: Structural organization of viral-cellular junctions in OPSCC patients.

Patient	VCJ	3'-breakpoint HPV16	Chromosome	Position (Strang) GRCh38
661	VCJ1	2323	5	91,997,752 (-)
993	VCJ1	1065	1	8,774,749 (-)
	VCJ2	2661	9	10,958,578 (-)
	VCJ3	3462	9	7,855,857 (+)
1425	VCJ1	3488	9	97,700,482 (+)
	VCJ2	1014	9	97,845,247 (+)
	VCJ3	1122	9	97,872,724 (+)
	VCJ4	2376	(21)	repetitive sequences at several positions
1464	VCJ1	4303	1	100,621,391 (-)
	VCJ2	2348	1	100,626,755 (-)
1529	VCJ1	2829	6	106,803,339 (+)
	VCJ2	1974	6	106,975,053 (+)
	VCJ3	2158	6	106,995,226 (-)
	VCJ4	3992	6	107,020,172 (+)
1553	VCJ1	3108	1	111,211,650 (+)
	VCJ2	3768	22	36,946,938 (+)
1643	VCJ1	2581	20	50,301,639 (+)
1650	VCJ1	1247	8	127,680,237 (-)
	VCJ2	2821	8	127,680,288 (+)

VCJ=viral-cellular junction, GRCh38=human reference genome (NCBI)

Table A-2: Characteristics of 180 NSCCUP patients included in the multicenter study.

		HPV Marker												OS		PFS			
ID	Center	Driven	DNA	RNA	p16	Diagnosis	Age	Gender	Tob	Alc	Treat	ECS	N	Years	Status	Years	Status	TP53	MS
1	B	0	0	0	0	13.09.90	62	0	1	1	S+R+C	1	3	1.37	1	1.37	1	N/A	N/A
2	B	0	invalid	0	0	20.03.91	56	0	1	1	S+R+C	1	3	0.99	1	0.99	1	N/A	N/A
3	B	0	invalid	0	0	27.08.91	81	0	1	1	S+R+C	1	2b	3.34	1	3.34	1	N/A	N/A
4	B	0	invalid	0	1	25.08.92	52	0	1	1	S+R+C	1	3	1.88	1	1.88	1	N/A	N/A
5	B	0	invalid	0	invalid	20.12.93	51	0	1	1	S+R+C	1	2b	4.60	1	4.60	1	N/A	4
6	B	0	invalid	0	0	03.01.94	44	1	0	0	R	0	3	22.04	0	19.77	0	N/A	N/A
7	B	0	invalid	0	0	15.10.93	61	0	1	1	S+R+C	1	3	5.56	1	5.56	1	N/A	N/A
8	B	0	invalid	0	0	02.11.94	67	0	1	0	S+R+C	0	1	3.27	1	3.27	1	N/A	N/A
9	B	1	1	1	1	10.02.94	55	0	1	1	S+R+C	1	2b	2.08	1	1.19	1	N/A	N/A
10	B	0	invalid	0	0	16.03.94	68	0	1	1	S+R+C	1	2b	2.01	1	1.25	1	N/A	N/A
11	B	0	invalid	0	0	28.07.95	66	0	1	0	S+R+C	0	2a	15.15	1	15.09	1	N/A	2
12	B	0	invalid	0	0	20.10.96	56	0	1	1	S+R+C	0	1	8.97	1	8.82	1	N/A	N/A
13	B	0	invalid	0	0	23.04.96	77	0	1	1	S	0	1	5.98	1	3.43	1	N/A	N/A
14	B	0	invalid	0	1	07.10.97	66	0	1	1	S+R+C	1	2b	1.02	1	0.92	1	N/A	N/A
15	B	0	invalid	0	0	30.03.98	56	0	1	1	S+R+C	1	2a	17.27	0	15.22	0	N/A	N/A
16	B	0	invalid	0	0	19.02.98	60	0	1	1	S+R+C	1	2b	0.82	1	0.82	1	N/A	N/A
17	B	0	invalid	0	0	25.02.98	42	1	1	0	S+R+C	1	3	17.33	0	15.27	0	N/A	N/A
18	B	0	invalid	0	0	11.05.98	65	0	1	0	S+R+C	1	1	7.42	0	7.42	0	N/A	N/A
19	B	0	invalid	0	0	18.11.98	44	0	1	1	S+R+C	1	2b	16.93	0	5.52	1	N/A	N/A
20	B	0	invalid	0	0	13.01.99	57	0	1	1	S+R+C	1	2b	0.91	1	0.91	1	N/A	N/A
21	B	0	invalid	0	0	29.03.99	63	0	1	1	S+R+C	1	2b	1.94	1	1.94	1	N/A	N/A
22	B	0	invalid	0	0	20.04.99	79	0	1	0	S	0	1	8.48	1	8.48	1	N/A	1
23	B	0	invalid	0	0	13.07.99	61	0	1	1	S	0	2b	16.56	0	13.33	0	N/A	3
24	B	0	invalid	0	0	21.06.99	62	0	1	1	S+R+C	1	2b	1.15	1	1.15	1	N/A	N/A
25	B	1	1	1	1	09.11.99	43	0	1	1	S	1	3	1.39	1	1.39	1	N/A	N/A
26	B	0	invalid	0	0	23.08.00	60	0	1	1	S+R+C	1	2b	3.04	1	3.04	1	N/A	0
27	B	0	invalid	0	0	06.06.01	72	0	1	1	S	1	3	0.18	1	0.18	1	N/A	N/A
28	B	0	invalid	0	invalid	26.09.02	68	0	1	1	S+R+C	1	2b	1.87	1	1.55	1	N/A	N/A
29	B	0	invalid	0	0	21.03.02	71	0	1	1	S+R+C	1	3	8.32	1	5.99	1	N/A	N/A
30	B	0	invalid	0	0	26.06.02	52	0	1	1	S+R+C	1	1	13.38	0	11.37	0	N/A	N/A
31	B	0	0	0	0	01.12.03	51	0	1	1	S+R+C	0	2b	4.02	1	2.40	1	N/A	N/A
32	B	0	invalid	0	0	13.09.04	62	0	1	1	S	1	2b	2.39	1	2.39	1	N/A	N/A
33	B	1	1	1	1	27.10.04	69	0	1	0	S+R+C	1	1	10.62	1	6.32	1	wt	3
34	B	0	invalid	0	0	29.10.03	68	0	0	1	S	1	3	0.44	1	0.44	1	N/A	N/A
35	B	0	invalid	0	0	07.02.05	77	0	1	0	S+R+C	1	2a	0.56	1	0.56	1	N/A	N/A
36	B	0	invalid	0	0	25.07.05	80	0	0	0	S+R+C	1	3	0.75	1	0.75	1	N/A	1
37	B	0	invalid	0	invalid	27.07.04	78	0	0	1	S+R+C	1	3	0.95	1	0.95	1	N/A	N/A
38	B	0	invalid	0	0	02.09.04	65	0	1	1	S+R+C	0	2b	5.81	1	5.12	1	N/A	N/A
39	B	0	0	0	0	25.02.05	58	0	1	1	S+R+C	1	3	1.95	1	1.95	1	N/A	N/A
40	B	1	1	1	1	21.05.05	47	1	1	0	S+R+C	1	2b	7.53	1	5.25	1	wt	N/A
41	B	0	1	N/A	0	02.08.06	70	0	0	1	N/A	N/A	2c	0.68	1	0.68	1	wt	N/A
42	B	0	invalid	0	0	04.04.06	36	0	1	1	S	N/A	3	1.28	1	1.28	1	N/A	N/A
43	B	0	1	0	0	19.04.06	61	0	1	1	S+R+C	1	3	1.12	1	1.12	1	N/A	N/A
44	B	0	0	0	0	11.04.06	65	0	1	1	S+R+C	1	2b	6.29	1	5.83	1	N/A	N/A
45	B	0	0	0	0	21.08.06	56	0	1	1	S+R+C	1	3	4.09	1	2.65	1	N/A	N/A
46	B	0	0	0	0	17.05.07	71	0	1	1	S+R+C	1	3	8.44	0	5.53	1	N/A	N/A
47	B	0	invalid	0	0	28.07.08	61	0	1	0	S+R+C	1	3	0.79	1	0.79	1	N/A	N/A
48	B	0	0	0	0	04.12.08	42	0	1	1	S+R+C	1	3	2.67	1	2.01	1	mut	N/A
49	B	0	0	0	0	23.02.09	60	0	1	1	S+R+C	1	3	5.84	1	5.84	1	N/A	N/A
50	B	0	0	0	0	29.04.09	56	1	1	1	S+R+C	0	1	6.58	0	4.83	0	wt	2
51	B	0	0	0	0	30.12.09	42	0	1	1	S+R+C	0	2b	5.31	0	3.50	0	wt	N/A
52	B	1	1	1	1	02.02.10	58	1	0	0	S+R+C	0	2a	5.25	0	3.79	0	wt	3
53	B	0	0	0	0	03.02.10	40	1	0	0	S+R+C	0	2b	5.70	0	4.06	0	wt	3
54	B	0	1	0	1	10.05.10	55	1	1	0	S+R+C	0	2b	5.69	0	3.42	0	N/A	2
55	B	0	0	0	0	08.09.10	75	0	1	0	S	1	3	0.07	1	0.07	1	N/A	3
56	B	0	0	0	0	02.07.10	73	0	1	1	S	1	2b	0.29	1	0.29	1	N/A	1
57	B	1	1	1	1	24.02.11	67	0	1	1	S+R+C	1	2b	4.56	1	3.63	1	wt	3
58	B	0	0	0	0	02.06.11	71	0	1	1	S+R+C	1	2b	4.42	1	0.69	1	N/A	2
59	B	0	0	0	0	06.03.12	75	0	1	0	S+R+C	1	3	2.00	1	0.52	1	N/A	3
60	H	1	1	1	1	17.03.09	67	0	1	1	S+R+C	1	2b	5.58	0	0.42	0	wt	3

		HPV Marker												OS		PFS			
ID	Center	Driven	DNA	RNA	p16	Diagnosis	Age	Gender	Tob	Alc	Treat	ECS	N	Years	Status	Years	Status	TP53	MS
61	H	0	0	0	0	15.06.09	70	0	1	0	S+R+C	0	2b	2.24	1	2.24	1	wt	3
62	H	1	1	1	1	25.09.09	52	0	0	0	S+R+C	1	2a	3.59	0	3.59	0	wt	5
63	H	0	0	0	0	03.12.09	50	0	1	0	S+R+C	1	2b	5.77	0	5.77	0	N/A	2
64	H	0	0	0	0	02.06.10	68	0	1	1	S+R	1	1	2.05	0	2.05	0	mut	2
65	H	0	0	0	0	08.06.10	66	0	1	0	S+R	0	2b	2.09	0	2.09	0	mut	2
66	H	0	0	0	0	12.05.10	65	0	1	1	S+R+C	1	2a	0.56	0	0.47	1	N/A	N/A
67	H	1	1	1	1	08.11.10	58	0	0	1	S	0	2a	5.09	0	5.09	0	wt	1
68	H	0	invalid	0	0	18.01.11	61	0	0	0	S+R+C	0	1	4.53	0	4.53	0	N/A	3
69	H	0	invalid	0	0	02.02.11	61	0	1	0	S	1	2a	1.24	1	1.24	1	N/A	N/A
70	H	1	1	1	1	15.02.10	55	0	0	N/A	S	0	2b	4.66	0	4.66	1	N/A	3
71	H	0	0	0	0	15.04.11	59	0	N/A	N/A	S+R	1	1	1.29	0	1.29	0	N/A	1
72	H	0	0	0	0	09.06.11	85	0	0	1	S+R+C	0	1	1.19	1	0.09	1	wt	1
73	H	0	0	0	0	12.07.11	69	0	1	1	S	1	2b	0.60	0	0.60	0	N/A	2
74	H	1	1	1	1	27.07.11	56	1	1	1	S+R+C	0	1	4.46	0	4.46	0	wt	5
75	H	1	1	1	1	08.08.11	56	0	0	0	S+R	0	2b	0.92	0	0.92	0	N/A	3
76	H	1	1	1	1	20.09.11	62	1	0	0	S+R	0	1	3.57	0	3.57	0	wt	2
77	H	0	invalid	0	0	25.08.11	51	0	1	1	S+R+C	1	2c	2.58	0	2.58	0	N/A	N/A
78	H	0	0	0	0	13.09.11	75	0	1	0	S+R+C	1	2b	0.31	0	0.31	0	N/A	N/A
80	H	0	invalid	0	0	22.08.97	53	0	0	1	S+R+C	1	3	1.49	1	1.49	1	N/A	4
81	H	0	invalid	0	1	01.02.99	62	0	1	1	S+R	1	2a	0.04	0	0.04	0	N/A	1
82	H	0	0	0	0	18.02.99	58	0	1	0	S+R+C	1	3	0.78	1	0.78	1	wt	1
83	H	0	invalid	0	0	19.10.05	55	0	1	1	S+R	1	2b	0.61	1	0.61	1	N/A	1
84	H	0	0	0	1	04.05.06	64	1	1	1	R+C	0	2c	1.21	1	1.21	1	N/A	2
85	H	0	invalid	0	0	16.10.06	48	1	1	1	S+R+C	0	2b	2.64	1	1.54	1	N/A	1
86	H	1	1	1	1	15.01.07	68	0	0	0	S	1	2a	8.64	0	1.52	1	N/A	N/A
87	H	0	invalid	0	0	08.02.08	69	0	1	1	S+C	1	3	2.06	0	2.06	0	N/A	4
88	H	0	invalid	0	0	18.06.08	83	1	0	0	S	0	2a	2.15	0	2.15	0	N/A	2
89	H	0	0	0	0	15.11.10	54	0	1	0	S+C	0	3	0.38	1	0.38	1	N/A	1
90	H	0	invalid	0	0	18.04.11	60	0	0	1	S+R+C	1	2c	0.41	1	0.16	1	N/A	1
93	H	0	0	N/A	N/A	19.05.93	53	0	1	1	R	N/A	2a	21.47	0	6.87	1	N/A	1
94	H	0	0	0	N/A	28.04.93	60	0	1	1	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
95	H	0	0	0	N/A	15.06.95	72	0	1	1	N/A	1	2a	7.76	0	3.71	1	N/A	2
96	H	0	0	0	N/A	20.06.95	66	0	1	1	N/A	1	3	0.14	0	0.14	0	N/A	2
97	H	0	0	N/A	N/A	22.01.96	59	0	1	1	S+R	1	2a	4.94	0	4.94	1	N/A	2
98	H	0	0	N/A	N/A	25.01.96	64	0	N/A	N/A	S+R	1	3	0.59	0	0.35	1	N/A	1
99	H	0	0	N/A	N/A	15.08.96	65	0	1	1	N/A	1	3	0.26	1	0.26	1	wt	3
101	H	1	1	1	N/A	04.09.97	77	0	0	0	S	1	2a	14.29	1	14.29	1	wt	4
102	H	0	0	N/A	N/A	27.01.98	43	0	1	1	S+R	0	2b	16.98	0	1.09	1	wt	2
103	H	0	0	N/A	N/A	06.03.98	75	0	N/A	N/A	R+C	N/A	2a	0.10	0	0.10	0	N/A	4
104	H	0	0	N/A	N/A	06.08.98	59	0	1	1	S+R	1	N/A	N/A	N/A	N/A	N/A	N/A	N/A
105	H	1	1	1	N/A	01.12.99	70	1	1	1	R+C	1	2b	1.35	1	1.35	1	wt	4
106	H	0	0	N/A	N/A	31.03.00	51	0	1	1	S+R	N/A	2a	2.83	0	2.83	0	mut	2
107	H	0	0	N/A	N/A	03.05.00	45	0	1	1	S+R	N/A	3	0.77	1	0.77	1	N/A	N/A
108	H	0	0	N/A	N/A	02.05.01	63	0	N/A	N/A	N/A	N/A	2a	0.03	1	0.03	1	N/A	N/A
109	H	0	0	N/A	N/A	28.07.04	55	0	1	1	S+R	1	2a	10.16	1	4.84	1	N/A	4
110	H	0	0	N/A	N/A	23.06.10	73	0	1	1	S	0	3	3.62	0	3.62	0	N/A	3
111	H	0	0	N/A	N/A	13.07.11	49	0	1	1	S+R+C	1	2b	0.04	0	0.04	0	wt	1
112	H	0	0	0	N/A	16.07.90	56	0	1	1	S+R+C	0	2b	15.29	1	6.10	1	N/A	N/A
116	Tre	0	invalid	0	1	15.08.98	70	0	1	1	S+R	1	2b	7.25	1	6.60	1	N/A	N/A
117	Tre	0	invalid	0	0	04.07.95	68	0	1	1	S	0	2b	13.68	1	11.50	1	N/A	4
118	Tre	0	invalid	0	0	12.05.98	62	0	1	1	S+R	1	2b	6.65	1	2.80	1	N/A	N/A
119	Tre	0	0	0	0	28.02.92	64	0	1	1	S+R	1	3	22.44	0	1.93	1	N/A	2
120	Tre	0	0	0	0	21.07.09	77	0	1	1	S+R	1	2b	5.04	0	5.04	0	mut	N/A
121	Tre	0	invalid	0	0	01.09.00	88	1	0	0	S+R	0	2b	13.62	1	13.62	1	N/A	0
122	Tre	0	0	0	0	01.08.00	56	0	1	1	S+R	1	3	1.54	1	0.08	1	wt	2
123	Tre	1	1	1	1	16.10.12	79	0	1	0	S+R+C	1	2a	3.30	0	3.30	0	wt	3
124	Tre	0	0	0	0	24.08.07	76	0	1	1	S+R	1	2b	3.15	1	0.60	1	N/A	1
125	Tre	0	0	0	0	13.03.89	72	0	1	1	S	1	3	1.99	1	0.89	1	N/A	2
126	Tre	0	0	0	0	31.08.06	76	0	1	0	S+R	1	2b	1.96	1	1.96	1	mut	3
127	Tre	1	1	1	1	20.07.88	57	0	1	1	S+R	1	2b	27.27	0	27.27	0	wt	4
128	Tre	0	0	0	0	10.09.89	48	0	1	0	S	0	2a	24.83	0	24.83	0	mut	2
129	Tre	1	1	1	1	23.05.08	51	0	1	0	S+R	1	2b	7.60	0	7.60	0	wt	N/A

		HPV Marker												OS		PFS			
ID	Center	Driven	DNA	RNA	p16	Diagnosis	Age	Gender	Tob	Alc	Treat	ECS	N	Years	Status	Years	Status	TP53	MS
130	Tre	1	1	1	0	09.03.10	48	0	1	0	S+R+C	1	2b	5.18	0	5.18	0	N/A	N/A
131	Tre	0	0	0	0	01.03.98	71	0	1	1	S+R	1	2b	7.17	1	7.17	1	mut	2
132	Tre	0	0	0	1	28.01.97	65	0	1	0	S+R	1	3	1.61	1	1.61	1	N/A	1
133	Tre	1	1	1	1	01.06.03	61	0	1	0	S	1	2a	12.72	0	12.72	0	N/A	2
134	Tre	0	invalid	0	0	11.08.98	67	0	1	1	S+R	1	3	1.52	1	1.22	1	N/A	2
135	Tre	0	0	0	0	08.09.92	58	0	1	1	S+R	1	3	7.38	1	6.23	1	mut	3
136	Tre	0	0	0	1	02.02.10	61	0	1	1	S+R+C	1	2b	4.97	0	4.97	0	N/A	N/A
137	Tre	0	0	0	0	27.08.13	57	1	1	0	S+R+C	0	2b	1.13	0	1.13	0	N/A	0
138	Tre	0	invalid	0	0	01.05.02	63	0	1	1	S	1	2b	1.61	1	1.61	1	N/A	N/A
139	Tre	0	0	0	0	18.12.89	68	0	1	1	S	1	2a	1.38	1	0.20	1	mut	N/A
140	Tre	0	0	0	0	21.08.07	53	1	1	0	S+R	1	2a	7.05	0	7.05	0	mut	N/A
141	Tre	0	0	0	0	05.01.89	47	0	1	1	S	1	2c	0.83	1	0.57	1	N/A	N/A
142	Tre	0	1	0	0	22.07.02	79	0	1	1	S+R	1	2b	2.02	1	0.94	1	N/A	N/A
143	Tre	0	invalid	0	1	16.07.01	60	0	0	0	S	0	2c	5.17	1	1.96	1	N/A	3
144	Tre	0	invalid	0	0	03.05.89	77	0	1	1	S	0	2a	0.48	1	0.48	1	N/A	N/A
145	Tre	0	0	0	0	06.03.01	52	0	1	0	S+R	0	2b	6.58	1	6.58	1	N/A	N/A
146	Tre	0	0	0	0	03.08.99	71	0	1	1	S	1	3	0.63	1	0.41	1	N/A	N/A
147	Tre	0	0	0	0	19.06.96	73	0	1	1	S+R	1	2b	5.47	1	5.47	1	N/A	N/A
148	Tre	0	0	0	0	01.02.04	56	1	1	1	S+R	1	3	10.61	0	9.24	1	N/A	1
149	Tre	0	invalid	0	1	01.12.02	72	0	0	0	S+R+C	1	3	3.47	1	0.08	1	N/A	N/A
150	Tre	0	0	0	0	12.09.00	47	0	1	0	S+R	1	2b	0.70	1	0.70	1	N/A	N/A
151	Tre	0	0	0	1	20.10.04	54	0	1	1	S+R	1	2b	0.68	1	0.45	1	N/A	2
152	Tre	0	0	0	0	19.06.12	77	0	1	0	S+R	1	2b	2.30	0	2.30	0	N/A	N/A
153	Tre	0	0	0	0	25.07.05	59	0	1	1	S+R	1	3	9.55	0	0.68	1	N/A	2
154	Tre	0	invalid	0	0	01.04.98	67	0	1	1	R+C	1	2c	0.76	1	0.08	1	N/A	N/A
155	Tre	0	0	0	0	08.07.02	67	0	1	0	S+R	1	3	12.08	0	4.48	1	N/A	N/A
156	Tre	0	0	0	0	07.11.90	79	0	1	0	S+R	1	2b	1.04	1	1.04	1	N/A	2
157	Tre	0	0	0	0	11.09.12	56	0	1	0	S+R+C	1	2b	1.90	0	1.90	0	N/A	1
158	Tre	0	0	0	0	09.12.92	53	0	1	1	S	0	2a	11.23	1	11.23	1	wt	3
160	Tre	1	1	1	1	22.01.10	63	0	1	0	S+R	1	2a	5.90	0	5.90	0	N/A	4
161	Tre	0	0	0	0	11.08.99	50	0	1	1	S	1	2b	1.72	1	1.72	1	N/A	1
162	Tre	0	invalid	0	0	04.07.03	74	1	1	1	S	1	3	1.42	1	0.58	1	N/A	3
163	M	0	0	0	0	10.07.06	56	0	1	1	S+R	0	2c	5.76	0	5.76	0	wt	1
164	M	0	0	0	0	22.07.14	76	0	1	1	S	0	2a	0.16	0	0.11	1	mut	3
165	M	0	0	0	0	02.08.11	76	0	0	0	S+R	1	2b	1.83	1	1.83	1	N/A	1
166	M	0	0	0	0	15.07.11	54	0	1	1	S	0	2a	3.02	1	2.70	1	N/A	2
167	M	0	0	0	1	16.01.08	75	0	1	1	S+R	0	2a	5.74	1	5.74	1	N/A	2
168	M	0	0	0	0	11.03.04	65	1	1	1	S+R	0	2b	8.06	1	8.06	1	wt	4
169	M	0	0	0	0	27.02.08	54	0	1	1	S+R	1	3	1.28	1	0.59	1	mut	2
170	M	0	0	0	0	02.02.04	63	0	1	1	S+R	0	2b	8.52	0	8.52	0	N/A	4
171	M	0	0	0	0	11.07.11	52	0	1	1	S+R+C	0	1	3.32	0	3.32	0	mut	4
172	M	0	invalid	0	0	16.02.12	58	0	1	1	S+R+C	1	2b	2.76	0	2.76	0	N/A	3
173	M	0	0	0	0	01.04.03	59	0	1	1	S+R	1	2c	1.51	1	1.00	1	N/A	1
174	M	0	0	0	0	07.07.09	58	0	1	1	S+R+C	1	3	4.99	0	4.99	0	wt	1
175	M	0	invalid	0	0	26.03.08	56	0	0	1	S+R	1	2b	6.05	0	6.05	0	N/A	4
176	M	1	1	1	1	17.02.12	66	1	0	0	S	0	2b	4.00	0	4.00	0	N/A	2
177	M	0	0	0	0	09.02.12	80	0	0	0	S+R	1	2b	2.86	0	2.86	0	mut	2
178	Tri	1	1	1	1	01.12.03	65	0	1	1	S	1	2b	12.22	0	12.22	0	wt	N/A
179	Tri	1	1	1	1	09.06.06	71	1	0	0	S+R	0	2b	9.35	0	9.35	0	wt	3
180	Tri	0	0	0	1	20.10.09	87	0	1	0	S	1	1	2.07	1	2.07	1	N/A	N/A
181	Tri	0	0	0	1	11.04.12	52	0	1	1	S+R	0	2b	3.39	0	3.39	0	N/A	1
182	Tri	0	0	0	0	15.10.12	68	0	0	1	S+R+C	1	2b	3.16	0	3.16	0	N/A	1
183	Tri	0	1	0	0	16.10.12	56	0	1	1	S	N/A	2b	0.00	1	0.00	1	mut	N/A
184	Tri	0	0	0	0	05.05.14	65	0	1	1	S+R+C	1	2b	1.27	1	1.27	1	N/A	3
185	Tri	0	0	0	0	09.05.14	63	1	1	1	S	1	2b	0.45	1	0.32	1	N/A	N/A
186	Tri	1	1	1	1	26.05.14	59	0	1	1	S+R	0	2b	1.70	0	1.70	0	N/A	4
187	Tri	1	1	1	1	22.08.14	56	0	1	1	S+R+C	1	2b	1.41	0	1.41	0	wt	2
188	Tri	1	1	1	1	21.10.14	39	0	1	1	S	0	2a	0.10	0	0.10	0	wt	4

Center: B=Barcelona, H=Heidelberg, Tre=Treviso, M=Montebelluna, Tri=Trieste; Driven: positive for HPV mRNA and HPV DNA and/or p16 overexpression (>25%); Age=age at diagnosis; Gender: 1=female, 0=male; Tob=tobacco consumption (1=current or former, 0=never); Alc=alcohol consumption (1=current or former, 0=never); Treat=treatment: S=surgery, R=radiotherapy, C=chemotherapy; ECS=extracapsular spread; N=N stage; OS=overall survival; PFS=progression-free survival; TP53: wt=wild-type or non-disruptive mutation, mut=disruptive mutation; MS=methylation score (0-5); N/A=not assessed

Table A-3: Characteristics of 46 NSCCUP patients from Leipzig analyzed by HPV serology.

	HPV status										OS		PFS	
ID	Serum	RNA	Diagnosis	Age	Gender	Tob	Alc	Treat	ECS	N	Years	Status	Years	Status
4	16	16	21.08.08	79	0	1	1	R	1	2b	1.31	1	1.31	1
5	16	16	19.10.10	67	0	0	1	S	1	2a	5.85	0	1.27	1
6	0	0	05.11.10	56	0	1	1	S+R+C	1	2a	5.78	0	5.78	0
7	16	N/A	09.12.10	68	0	0	1	S+R+C	1	2b	4.56	0	4.56	0
8	0	0	08.06.11	68	0	1	1	S+R+C	1	2b	1.45	1	0.64	1
10	0	0	22.02.10	38	0	1	1	S+R+C	1	2b	3.90	1	3.38	0
11	0	N/A	19.07.10	67	0	1	1	N/A	1	1	3.74	0	3.74	1
12	0	N/A	27.04.15	78	0	1	0	S	1	2b	0.13	1	0.13	1
13	0	0	25.09.13	52	1	1	1	S+R+C	1	2a	2.93	0	2.93	0
14	16	16	28.08.13	57	0	1	0	S+R	0	2a	3.15	0	3.15	0
15	0	N/A	21.06.12	52	0	1	1	S+R+C	1	2b	4.33	0	4.33	0
16	33	N/A	16.09.10	67	0	0	1	S+R+C	N/A	2a	4.85	0	4.85	0
17	0	N/A	26.05.10	48	0	1	1	S+R+C	1	2b	1.39	1	1.39	1
18	16	16	19.05.09	68	0	0	1	S+R+C	1	2b	5.08	0	5.08	0
19	0	0	26.11.08	66	0	0	1	S+R+C	0	2b	4.22	1	4.22	1
20	0	0	23.07.08	58	0	1	1	S+R+C	1	2b	2.08	1	2.03	1
23	0	0	07.07.11	48	0	1	1	S+R+C	1	2a	3.10	1	1.52	1
24	0	0	15.07.09	67	0	1	1	S	1	3	0.59	1	0.35	1
25	16	16	01.10.08	52	1	0	1	S+R+C	0	1	4.91	0	4.91	0
29	0	0	03.01.13	60	0	1	1	S	1	2a	0.28	1	0.28	1
30	0	N/A	31.08.11	53	0	0	1	S+R+C	0	2b	3.93	0	3.93	0
33	0	N/A	09.11.13	63	0	1	1	S+R+C	0	2b	3.00	0	3.00	0
34	0	0	03.02.14	59	0	1	1	S	1	2b	0.16	1	0.16	1
35	0	N/A	17.03.14	56	0	1	1	S+R+C	1	2a	1.97	1	1.81	1
36	0	N/A	12.06.14	59	1	1	1	S	0	N/A	2.07	0	2.07	0
37	0	N/A	18.06.14	79	0	1	1	S	1	2b	0.08	1	0.08	1
38	0	0	10.09.15	55	0	1	1	S+R	N/A	3	1.11	0	0.43	1
40	31/35	31	19.02.16	60	0	1	1	S+R	0	2a	0.61	0	0.61	0
41	0	N/A	24.02.16	79	0	1	1	S+R	1	2a	0.65	0	0.42	1
43	0	0	14.12.12	57	0	1	1	S+C	1	2c	3.18	1	1.14	1
44	16	16	14.03.14	63	0	1	1	S+R	0	1	2.50	0	2.50	0
45	0	0	03.07.13	59	0	1	1	S	1	2a	1.70	1	1.04	1
46	16	16	25.02.13	65	0	1	1	S	1	3	3.50	0	2.09	1
47	0	0	15.11.12	80	0	1	1	S+R	1	2b	0.66	1	0.44	1
49	18	N/A	16.02.12	57	0	0	0	S	N/A	2b	4.59	0	4.59	0
50	16	16	14.09.11	66	0	0	1	S+R+C	1	2b	5.23	0	5.23	0
51	0	N/A	13.07.10	79	0	1	1	C	N/A	2a	6.28	0	6.28	0
52	0	N/A	23.09.09	46	1	1	1	S+R+C	1	2b	7.08	0	7.08	0
53	0	N/A	07.09.09	58	0	1	1	S+R	1	3	0.58	0	0.58	0
56	0	0	12.08.08	52	0	1	1	N/A	1	3	0.08	1	0.08	1
57	0	N/A	07.08.08	73	0	0	0	S+R	1	2b	1.20	1	0.61	1
58	16	16	16.01.14	68	1	1	1	0	0	2a	2.76	0	2.76	0
59	0	0	13.06.16	55	0	1	0	S+R+C	1	2b	0.41	0	0.41	0
60	0	N/A	29.07.16	57	0	1	1	R	1	2b	0.03	0	0.03	0
61	0	0	29.08.16	59	0	1	1	S+R+C	1	2c	0.23	0	0.23	0
24B	0	33	02.11.09	46	0	1	1	S+R+C	0	1	4.89	0	4.89	0

Age=age at diagnosis; Gender: 1=female, 0=male; Tob=tobacco consumption (1=current or former, 0=never); Alc=alcohol consumption (1=current or former, 0=never); Treat=treatment: S=surgery, R=radiotherapy, C=chemotherapy; ECS=extracapsular spread; N=N stage; OS=overall survival; PFS=progression-free survival; N/A=not assessed

References

[1] Doorbar J, Egawa N, Griffin H, Kranjec C, Murakami I. Human papillomavirus molecular biology and disease association. Reviews in medical virology. 2015;25 Suppl 1:2-23.

[2] Bernard HU, Burk RD, Chen Z, van Doorslaer K, zur Hausen H, de Villiers EM. Classification of papillomaviruses (PVs) based on 189 PV types and proposal of taxonomic amendments. Virology. 2010;401:70-9.

[3] zur Hausen H. Papillomaviruses and cancer: from basic studies to clinical application. Nature reviews Cancer. 2002;2:342-50.

[4] Rautava J, Syrjanen S. Biology of human papillomavirus infections in head and neck carcinogenesis. Head and neck pathology. 2012;6 Suppl 1:S3-15.

[5] Shafti-Keramat S, Handisurya A, Kriehuber E, Meneguzzi G, Slupetzky K, Kirnbauer R. Different heparan sulfate proteoglycans serve as cellular receptors for human papillomaviruses. Journal of virology. 2003;77:13125-35.

[6] Day PM, Lowy DR, Schiller JT. Papillomaviruses infect cells via a clathrin-dependent pathway. Virology. 2003;307:1-11.

[7] Nguyen HP, Ramirez-Fort MK, Rady PL. The biology of human papillomaviruses. Current problems in dermatology. 2014;45:19-32.

[8] Tommasino M. The human papillomavirus family and its role in carcinogenesis. Seminars in cancer biology. 2014;26:13-21.

[9] IARC. Human papillomaviruses. IARC monographs on the evaluation of carcinogenic risks to humans / World Health Organization, International Agency for Research on Cancer. 2007;90:1-636.

[10] Stanley M. Immune responses to human papillomavirus. Vaccine. 2006;24, Supplement 1:S16-S22.

[11] Speel EJ. HPV Integration in Head and Neck Squamous Cell Carcinomas: Cause and Consequence. Recent results in cancer research Fortschritte der Krebsforschung Progres dans les recherches sur le cancer. 2017;206:57-72.

[12] Xu B, Chotewutmontri S, Wolf S, Klos U, Schmitz M, Durst M, et al. Multiplex Identification of Human Papillomavirus 16 DNA Integration Sites in Cervical Carcinomas. PloS one. 2013;8:e66693.

[13] Olthof NC, Speel EJ, Kolligs J, Haesevoets A, Henfling M, Ramaekers FC, et al. Comprehensive analysis of HPV16 integration in OSCC reveals no significant impact of physical status on viral oncogene and virally disrupted human gene expression. PloS one. 2014;9:e88718.

[14] Hafkamp HC, Manni JJ, Haesevoets A, Voogd AC, Schepers M, Bot FJ, et al. Marked differences in survival rate between smokers and nonsmokers with HPV 16-associated tonsillar carcinomas. International journal of cancer Journal international du cancer. 2008;122:2656-64.

[15] Hafkamp HC, Speel EJ, Haesevoets A, Bot FJ, Dinjens WN, Ramaekers FC, et al. A subset of head and neck squamous cell carcinomas exhibits integration of HPV 16/18 DNA

and overexpression of p16INK4A and p53 in the absence of mutations in p53 exons 5-8. International journal of cancer Journal international du cancer. 2003;107:394-400.

[16] Mooren JJ, Gultekin SE, Straetmans JM, Haesevoets A, Peutz-Kootstra CJ, Huebbers CU, et al. P16(INK4A) immunostaining is a strong indicator for high-risk-HPV-associated oropharyngeal carcinomas and dysplasias, but is unreliable to predict low-risk-HPV-infection in head and neck papillomas and laryngeal dysplasias. International journal of cancer Journal international du cancer. 2014;134:2108-17.

[17] Mooren JJ, Kremer B, Claessen SM, Voogd AC, Bot FJ, Peter Klussmann J, et al. Chromosome stability in tonsillar squamous cell carcinoma is associated with HPV16 integration and indicates a favorable prognosis. International journal of cancer Journal international du cancer. 2013;132:1781-9.

[18] Mellin H, Dahlgren L, Munck-Wikland E, Lindholm J, Rabbani H, Kalantari M, et al. Human papillomavirus type 16 is episomal and a high viral load may be correlated to better prognosis in tonsillar cancer. International journal of cancer Journal international du cancer. 2002;102:152-8.

[19] Andersen AS, Koldjaer Solling AS, Ovesen T, Rusan M. The interplay between HPV and host immunity in head and neck squamous cell carcinoma. International journal of cancer Journal international du cancer. 2014;134:2755-63.

[20] Buck CB, Cheng N, Thompson CD, Lowy DR, Steven AC, Schiller JT, et al. Arrangement of L2 within the papillomavirus capsid. Journal of virology. 2008;82:5190-7.

[21] Wilson VG, West M, Woytek K, Rangasamy D. Papillomavirus E1 proteins: form, function, and features. Virus genes. 2002;24:275-90.

[22] Desaintes C, Demeret C. Control of papillomavirus DNA replication and transcription. Seminars in cancer biology. 1996;7:339-47.

[23] Johansson C, Somberg M, Li X, Backstrom Winquist E, Fay J, Ryan F, et al. HPV-16 E2 contributes to induction of HPV-16 late gene expression by inhibiting early polyadenylation. The EMBO journal. 2012;31:3212-27.

[24] Nakahara T, Peh WL, Doorbar J, Lee D, Lambert PF. Human papillomavirus type 16 E1circumflexE4 contributes to multiple facets of the papillomavirus life cycle. Journal of virology. 2005;79:13150-65.

[25] Wilson R, Fehrmann F, Laimins LA. Role of the E1--E4 protein in the differentiation-dependent life cycle of human papillomavirus type 31. Journal of virology. 2005;79:6732-40.

[26] Doorbar J. The E4 protein; structure, function and patterns of expression. Virology. 2013;445:80-98.

[27] Pim D, Collins M, Banks L. Human papillomavirus type 16 E5 gene stimulates the transforming activity of the epidermal growth factor receptor. Oncogene. 1992;7:27-32.

[28] Scheffner M, Huibregtse JM, Vierstra RD, Howley PM. The HPV-16 E6 and E6-AP complex functions as a ubiquitin-protein ligase in the ubiquitination of p53. Cell. 1993;75:495-505.

[29] Scheffner M, Werness BA, Huibregtse JM, Levine AJ, Howley PM. The E6 oncoprotein encoded by human papillomavirus types 16 and 18 promotes the degradation of p53. Cell. 1990;63:1129-36.

[30] Huh K, Zhou X, Hayakawa H, Cho JY, Libermann TA, Jin J, et al. Human papillomavirus type 16 E7 oncoprotein associates with the cullin 2 ubiquitin ligase complex, which contributes to degradation of the retinoblastoma tumor suppressor. Journal of virology. 2007;81:9737-47.

[31] Yang R, Yutzy WHt, Viscidi RP, Roden RB. Interaction of L2 with beta-actin directs intracellular transport of papillomavirus and infection. The Journal of biological chemistry. 2003;278:12546-53.

[32] Walboomers JM, Jacobs MV, Manos MM, Bosch FX, Kummer JA, Shah KV, et al. Human papillomavirus is a necessary cause of invasive cervical cancer worldwide. The Journal of pathology. 1999;189:12-9.

[33] Daling JR, Madeleine MM, Schwartz SM, Shera KA, Carter JJ, McKnight B, et al. A population-based study of squamous cell vaginal cancer: HPV and cofactors. Gynecologic oncology. 2002;84:263-70.

[34] Madeleine MM, Daling JR, Carter JJ, Wipf GC, Schwartz SM, McKnight B, et al. Cofactors with human papillomavirus in a population-based study of vulvar cancer. Journal of the National Cancer Institute. 1997;89:1516-23.

[35] Daling JR, Madeleine MM, Johnson LG, Schwartz SM, Shera KA, Wurscher MA, et al. Human papillomavirus, smoking, and sexual practices in the etiology of anal cancer. Cancer. 2004;101:270-80.

[36] Daling JR, Madeleine MM, Johnson LG, Schwartz SM, Shera KA, Wurscher MA, et al. Penile cancer: importance of circumcision, human papillomavirus and smoking in in situ and invasive disease. International journal of cancer Journal international du cancer. 2005;116:606-16.

[37] zur Hausen H. Papillomaviruses in the causation of human cancers — a brief historical account. Virology. 2009;384:260-5.

[38] 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. Journal of the National Cancer Institute. 2003;95:1772-83.

[39] Grabowska AK, Riemer AB. The invisible enemy - how human papillomaviruses avoid recognition and clearance by the host immune system. The open virology journal. 2012;6:249-56.

[40] Klingelhutz AJ, Foster SA, McDougall JK. Telomerase activation by the E6 gene product of human papillomavirus type 16. Nature. 1996;380:79-82.

[41] Gewin L, Myers H, Kiyono T, Galloway DA. Identification of a novel telomerase repressor that interacts with the human papillomavirus type-16 E6/E6-AP complex. Genes & development. 2004;18:2269-82.

[42] Funk JO, Waga S, Harry JB, Espling E, Stillman B, Galloway DA. Inhibition of CDK activity and PCNA-dependent DNA replication by p21 is blocked by interaction with the HPV-16 E7 oncoprotein. Genes & development. 1997;11:2090-100.

[43] Jones DL, Alani RM, Munger K. The human papillomavirus E7 oncoprotein can uncouple cellular differentiation and proliferation in human keratinocytes by abrogating p21Cip1-mediated inhibition of cdk2. Genes & development. 1997;11:2101-11.

[44] Zerfass-Thome K, Zwerschke W, Mannhardt B, Tindle R, Botz JW, Jansen-Durr P. Inactivation of the cdk inhibitor p27KIP1 by the human papillomavirus type 16 E7 oncoprotein. Oncogene. 1996;13:2323-30.

[45] Dyson N, Howley PM, Munger K, Harlow E. The human papilloma virus-16 E7 oncoprotein is able to bind to the retinoblastoma gene product. Science (New York, NY). 1989;243:934-7.

[46] Chellappan S, Kraus VB, Kroger B, Munger K, Howley PM, Phelps WC, et al. Adenovirus E1A, simian virus 40 tumor antigen, and human papillomavirus E7 protein share the capacity to disrupt the interaction between transcription factor E2F and the retinoblastoma gene product. Proceedings of the National Academy of Sciences of the United States of America. 1992;89:4549-53.

[47] Westra WH. Detection of human papillomavirus (HPV) in clinical samples: evolving methods and strategies for the accurate determination of HPV status of head and neck carcinomas. Oral oncology. 2014;50:771-9.

[48] Halec G, Schmitt M, Dondog B, Sharkhuu E, Wentzensen N, Gheit T, et al. Biological activity of probable/possible high-risk human papillomavirus types in cervical cancer. International journal of cancer Journal international du cancer. 2013;132:63-71.

[49] Holzinger D, Schmitt M, Dyckhoff G, Benner A, Pawlita M, Bosch FX. Viral RNA patterns and high viral load reliably define oropharynx carcinomas with active HPV16 involvement. Cancer research. 2012;72:4993-5003.

[50] Schmitt M, Dalstein V, Waterboer T, Clavel C, Gissmann L, Pawlita M. Diagnosing cervical cancer and high-grade precursors by HPV16 transcription patterns. Cancer research. 2010;70:249-56.

[51] 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. International journal of cancer Journal international du cancer. 2007;121:2465-72.

[52] Boscolo-Rizzo P, Pawlita M, Holzinger D. From HPV-positive towards HPV-driven oropharyngeal squamous cell carcinomas. Cancer treatment reviews. 2016;42:24-9.

[53] Prigge ES, Arbyn M, von Knebel Doeberitz M, Reuschenbach M. Diagnostic accuracy of p16INK4a immunohistochemistry in oropharyngeal squamous cell carcinomas: A systematic review and meta-analysis. International journal of cancer Journal international du cancer. 2017;140:1186-98.

[54] Holzinger D, Wichmann G, Baboci L, Michel A, Hofler D, Wiesenfarth M, et al. Sensitivity and specificity of antibodies against HPV16 E6 and other early proteins for the detection of HPV16-driven oropharyngeal squamous cell carcinoma. International journal of cancer Journal international du cancer. 2017.

[55] Kreimer AR, Johansson M, Waterboer T, Kaaks R, Chang-Claude J, Drogen D, et al. Evaluation of human papillomavirus antibodies and risk of subsequent head and neck cancer. Journal of clinical oncology : official journal of the American Society of Clinical Oncology. 2013;31:2708-15.

[56] Sanderson RJ, Ironside JAD. Squamous cell carcinomas of the head and neck. BMJ : British Medical Journal. 2002;325:822-7.

[57] Vigneswaran N, Williams MD. Epidemiological Trends in Head and Neck Cancer and Aids in Diagnosis. Oral and maxillofacial surgery clinics of North America. 2014;26:123-41.

[58] Chaturvedi AK, Anderson WF, Lortet-Tieulent J, Curado MP, Ferlay J, Franceschi S, et al. Worldwide trends in incidence rates for oral cavity and oropharyngeal cancers. Journal of clinical oncology : official journal of the American Society of Clinical Oncology. 2013;31:4550-9.

[59] Hashibe M, Brennan P, Chuang SC, Boccia S, Castellsague X, Chen C, et al. Interaction between tobacco and alcohol use and the risk of head and neck cancer: pooled analysis in the International Head and Neck Cancer Epidemiology Consortium. Cancer epidemiology, biomarkers & prevention : a publication of the American Association for Cancer Research, cosponsored by the American Society of Preventive Oncology. 2009;18:541-50.

[60] Wittekindt C, Wagner S, Mayer CS, Klussmann JP. Basics of tumor development and importance of human papilloma virus (HPV) for head and neck cancer. GMS current topics in otorhinolaryngology, head and neck surgery. 2012;11:Doc09.

[61] 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.

[62] 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. Journal of the National Cancer Institute. 2000;92:709-20.

[63] Dalianis T. Human papillomavirus (HPV) and oropharyngeal squamous cell carcinoma. Presse medicale (Paris, France : 1983). 2014;43:e429-e34.

[64] IARC. Proceedings of the IARC Working Group on the Evaluation of Carcinogenic Risks to Humans. Epstein-Barr Virus and Kaposi's Sarcoma Herpesvirus/Human Herpesvirus 8. Lyon, France, 17-24 June 1997. IARC monographs on the evaluation of carcinogenic risks to humans / World Health Organization, International Agency for Research on Cancer. 1997;70:1-492.

[65] Ndiaye C, Mena M, Alemany L, Arbyn M, Castellsagué X, Laporte L, et al. HPV DNA, E6/E7 mRNA, and p16INK4a detection in head and neck cancers: a systematic review and meta-analysis. The Lancet Oncology. 2014;15:1319-31.

[66] Upile NS, Shaw RJ, Jones TM, Goodyear P, Liloglou T, Risk JM, et al. Squamous cell carcinoma of the head and neck outside the oropharynx is rarely human papillomavirus related. The Laryngoscope. 2014;124:2739-44.

[67] Castellsague X, Alemany L, Quer M, Halec G, Quiros B, Tous S, et al. HPV Involvement in Head and Neck Cancers: Comprehensive Assessment of Biomarkers in 3680 Patients. Journal of the National Cancer Institute. 2016;108.

[68] Pai SI, Westra WH. Molecular pathology of head and neck cancer: implications for diagnosis, prognosis, and treatment. Annual review of pathology. 2009;4:49-70.

[69] Lyford-Pike S, Peng S, Young GD, Taube JM, Westra WH, Akpeng B, et al. Evidence for a role of the PD-1:PD-L1 pathway in immune resistance of HPV-associated head and neck squamous cell carcinoma. Cancer research. 2013;73:1733-41.

[70] Lechner M, Fenton TR. The Genomics, Epigenomics, and Transcriptomics of HPV-Associated Oropharyngeal Cancer--Understanding the Basis of a Rapidly Evolving Disease. Advances in genetics. 2016;93:1-56.

[71] 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. Journal of the National Cancer Institute. 2008;100:407-20.

[72] Westra WH, Taube JM, Poeta ML, Begum S, Sidransky D, Koch WM. Inverse relationship between human papillomavirus-16 infection and disruptive p53 gene mutations in squamous cell carcinoma of the head and neck. Clinical cancer research : an official journal of the American Association for Cancer Research. 2008;14:366-9.

[73] Wiest T, Schwarz E, Enders C, Flechtenmacher C, Bosch FX. Involvement of intact HPV16 E6/E7 gene expression in head and neck cancers with unaltered p53 status and perturbed pRb cell cycle control. Oncogene. 2002;21:1510-7.

[74] 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. 2010;126:1882-94.

[75] Wichmann G, Rosolowski M, Krohn K, Kreuz M, Boehm A, Reiche A, et al. The role of HPV RNA transcription, immune response-related gene expression and disruptive TP53 mutations in diagnostic and prognostic profiling of head and neck cancer. International journal of cancer Journal international du cancer. 2015;137:2846-57.

[76] Kostareli E, Holzinger D, Bogatyrova O, Hielscher T, Wichmann G, Keck M, et al. HPV-related methylation signature predicts survival in oropharyngeal squamous cell carcinomas. The Journal of clinical investigation. 2013;123:2488-501.

[77] Kostareli E, Hielscher T, Zucknick M, Baboci L, Wichmann G, Holzinger D, et al. Gene promoter methylation signature predicts survival of head and neck squamous cell carcinoma patients. Epigenetics : official journal of the DNA Methylation Society. 2016:1-13.

[78] Fakhry C, Westra WH, Li S, Cmelak A, Ridge JA, Pinto H, et al. Improved survival of patients with human papillomavirus-positive head and neck squamous cell carcinoma in a prospective clinical trial. Journal of the National Cancer Institute. 2008;100:261-9.

[79] Ang KK, Harris J, Wheeler R, Weber R, Rosenthal DI, Nguyen-Tan PF, et al. Human papillomavirus and survival of patients with oropharyngeal cancer. The New England journal of medicine. 2010;363:24-35.

[80] Wang MB, Liu IY, Gornbein JA, Nguyen CT. HPV-Positive Oropharyngeal Carcinoma: A Systematic Review of Treatment and Prognosis. Otolaryngology--head and neck surgery : official journal of American Academy of Otolaryngology-Head and Neck Surgery. 2015.

[81] Mirghani H, Amen F, Blanchard P, Moreau F, Guigay J, Hartl DM, et al. Treatment de-escalation in HPV-positive oropharyngeal carcinoma: ongoing trials, critical issues and perspectives. International journal of cancer Journal international du cancer. 2015;136:1494-503.

[82] Jereczek-Fossa BA, Jassem J, Orecchia R. Cervical lymph node metastases of squamous cell carcinoma from an unknown primary. Cancer treatment reviews. 2004;30:153-64.

[83] Erkal HS, Mendenhall WM, Amdur RJ, Villaret DB, Stringer SP. Squamous cell carcinomas metastatic to cervical lymph nodes from an unknown head and neck mucosal site treated with radiation therapy with palliative intent. Radiotherapy and oncology : journal of the European Society for Therapeutic Radiology and Oncology. 2001;59:319-21.

[84] Grau C, Johansen LV, Jakobsen J, Geertsen P, Andersen E, Jensen BB. Cervical lymph node metastases from unknown primary tumours. Results from a national survey by the Danish Society for Head and Neck Oncology. Radiotherapy and oncology : journal of the European Society for Therapeutic Radiology and Oncology. 2000;55:121-9.

[85] Adams JR, O'Brien CJ. Unknown primary squamous cell carcinoma of the head and neck: a review of diagnosis, treatment and outcomes. Asian journal of surgery / Asian Surgical Association. 2002;25:188-93.

[86] Zhuang SM, Wu X-F, Li J-J, Zhang G-H. Management of lymph node metastases from an unknown primary site to the head and neck (Review). Molecular and clinical oncology. 2014;2:917-22.

[87] Cianchetti M, Mancuso AA, Amdur RJ, Werning JW, Kirwan J, Morris CG, et al. Diagnostic evaluation of squamous cell carcinoma metastatic to cervical lymph nodes from an unknown head and neck primary site. The Laryngoscope. 2009;119:2348-54.

[88] Schroeder L, Boscolo-Rizzo P, Dal Cin E, Romeo S, Baboci L, Dyckhoff G, et al. Human papillomavirus as prognostic marker with rising prevalence in neck squamous cell carcinoma of unknown primary: A retrospective multicentre study. European Journal of Cancer. 2017;74:73-81.

[89] Roh JL, Kim JS, Lee JH, Cho KJ, Choi SH, Nam SY, et al. Utility of combined (18)F-fluorodeoxyglucose-positron emission tomography and computed tomography in patients with cervical metastases from unknown primary tumors. Oral oncology. 2009;45:218-24.

[90] Davis KS, Byrd JK, Mehta V, Chiosea SI, Kim S, Ferris RL, et al. Occult Primary Head and Neck Squamous Cell Carcinoma: Utility of Discovering Primary Lesions. Otolaryngology--head and neck surgery : official journal of American Academy of Otolaryngology-Head and Neck Surgery. 2014.

[91] Mozet C, Wichmann G, Stumpp P, Dietzsch S, Dietz A. Zervikales CUP-Syndrom. Onkologe. 2013;19:44-51.

[92] Kothari P, Randhawa PS, Farrell R. Role of tonsillectomy in the search for a squamous cell carcinoma from an unknown primary in the head and neck. The British journal of oral & maxillofacial surgery. 2008;46:283-7.

[93] Koch WM, Bhatti N, Williams MF, Eisele DW. Oncologic rationale for bilateral tonsillectomy in head and neck squamous cell carcinoma of unknown primary source. Otolaryngology--head and neck surgery : official journal of American Academy of Otolaryngology-Head and Neck Surgery. 2001;124:331-3.

[94] Moualed D, Qayyum A, Price T, Sharma A, Mahendran S. Bilateral synchronous tonsillar carcinoma: a case series and review of the literature. European archives of oto-rhino-laryngology : official journal of the European Federation of Oto-Rhino-Laryngological Societies (EUFOS) : affiliated with the German Society for Oto-Rhino-Laryngology - Head and Neck Surgery. 2012;269:255-9.

[95] Zengel P, Assmann G, Mollenhauer M, Jung A, Sotlar K, Kirchner T, et al. Cancer of unknown primary originating from oropharyngeal carcinomas are strongly correlated to HPV positivity. Virchows Archiv : an international journal of pathology. 2012;461:283-90.

[96] Pavlidis N, Pentheroudakis G, Plataniotis G. Cervical lymph node metastases of squamous cell carcinoma from an unknown primary site: a favourable prognosis subset of patients with CUP. Clinical & translational oncology : official publication of the Federation of Spanish Oncology Societies and of the National Cancer Institute of Mexico. 2009;11:340-8.

[97] Strojan P, Ferlito A, Langendijk JA, Corry J, Woolgar JA, Rinaldo A, et al. Contemporary management of lymph node metastases from an unknown primary to the neck: II. a review of therapeutic options. Head & neck. 2013;35:286-93.

[98] Pavlidis N, Briasoulis E, Hainsworth J, Greco FA. Diagnostic and therapeutic management of cancer of an unknown primary. European journal of cancer (Oxford, England : 1990). 2003;39:1990-2005.

[99] Fizazi K, Greco FA, Pavlidis N, Daugaard G, Oien K, Pentheroudakis G. Cancers of unknown primary site: ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up. Annals of oncology : official journal of the European Society for Medical Oncology / ESMO. 2015;26 Suppl 5:v133-8.

[100] Cerezo L, Raboso E, Ballesteros AI. Unknown primary cancer of the head and neck: a multidisciplinary approach. Clinical & translational oncology : official publication of the Federation of Spanish Oncology Societies and of the National Cancer Institute of Mexico. 2011;13:88-97.

[101] Iganej S, Kagan R, Anderson P, Rao A, Tome M, Wang R, et al. Metastatic squamous cell carcinoma of the neck from an unknown primary: management options and patterns of relapse. Head & neck. 2002;24:236-46.

[102] Beldi D, Jereczek-Fossa BA, D'Onofrio A, Gambaro G, Fiore MR, Pia F, et al. Role of radiotherapy in the treatment of cervical lymph node metastases from an unknown primary site: retrospective analysis of 113 patients. International journal of radiation oncology, biology, physics. 2007;69:1051-8.

[103] Madani I, Vakaet L, Bonte K, Boterberg T, De Neve W. Intensity-modulated radiotherapy for cervical lymph node metastases from unknown primary cancer. International journal of radiation oncology, biology, physics. 2008;71:1158-66.

[104] Rödel RM, Matthias C, Blomeyer BD, Wolff HA, Jung K, Christiansen H. Impact of distant metastasis in patients with cervical lymph node metastases from cancer of an unknown primary site. The Annals of otology, rhinology, and laryngology. 2009;118:662-9.

[105] Chen AM, Farwell DG, Lau DH, Li BQ, Luu Q, Donald PJ. Radiation therapy in the management of head-and-neck cancer of unknown primary origin: how does the addition of concurrent chemotherapy affect the therapeutic ratio? International journal of radiation oncology, biology, physics. 2011;81:346-52.

[106] Park GC, Jung JH, Roh JL, Lee JH, Cho KJ, Choi SH, et al. Prognostic Value of Metastatic Nodal Volume and Lymph Node Ratio in Patients with Cervical Lymph Node Metastases from an Unknown Primary Tumor. Oncology. 2014;86:170-6.

[107] Huang CC, Tseng FY, Yeh TH, Wen YH, Hsu CJ, Ko JY, et al. Prognostic factors of unknown primary head and neck squamous cell carcinoma. Otolaryngology--head and neck surgery : official journal of American Academy of Otolaryngology-Head and Neck Surgery. 2008;139:429-35.

[108] Lanzer M, Bachna-Rotter S, Graupp M, Bredell M, Rucker M, Huber G, et al. Unknown primary of the head and neck: A long-term follow-up. Journal of cranio-maxillo-facial surgery : official publication of the European Association for Cranio-Maxillo-Facial Surgery. 2015;43:574-9.

[109] Sivars L, Nasman A, Tertipis N, Vlastos A, Ramqvist T, Dalianis T, et al. Human papillomavirus and p53 expression in cancer of unknown primary in the head and neck region in relation to clinical outcome. Cancer medicine. 2014.

[110] Sivars L, Tani E, Nasman A, Ramqvist T, Munck-Wikland E, Dalianis T. Human Papillomavirus as a Diagnostic and Prognostic Tool in Cancer of Unknown Primary in the Head and Neck Region. Anticancer research. 2016;36:487-93.

[111] Andl T, Kahn T, Pfuhl A, Nicola T, Erber R, Conradt C, et al. Etiological involvement of oncogenic human papillomavirus in tonsillar squamous cell carcinomas lacking retinoblastoma cell cycle control. Cancer research. 1998;58:5-13.

[112] Snijders PJ, Cromme FV, van den Brule AJ, Schrijnemakers HF, Snow GB, Meijer CJ, et al. Prevalence and expression of human papillomavirus in tonsillar carcinomas, indicating a possible viral etiology. International journal of cancer Journal international du cancer. 1992;51:845-50.

[113] Boscolo-Rizzo P, Schroeder L, Romeo S, Pawlita M. The prevalence of human papillomavirus in squamous cell carcinoma of unknown primary site metastatic to neck lymph nodes: a systematic review. Clinical & experimental metastasis. 2015.

[114] Begum S, Gillison ML, Ansari-Lari MA, Shah K, Westra WH. Detection of human papillomavirus in cervical lymph nodes: a highly effective strategy for localizing site of tumor origin. Clinical cancer research : an official journal of the American Association for Cancer Research. 2003;9:6469-75.

[115] Park GC, Lee M, Roh JL, Yu MS, Choi SH, Nam SY, et al. Human papillomavirus and p16 detection in cervical lymph node metastases from an unknown primary tumor. Oral oncology. 2012;48:1250-6.

[116] Begum S, Gillison ML, Nicol TL, Westra WH. Detection of human papillomavirus-16 in fine-needle aspirates to determine tumor origin in patients with metastatic squamous cell carcinoma of the head and neck. Clinical cancer research : an official journal of the American Association for Cancer Research. 2007;13:1186-91.

[117] Kobayashi K, Saito Y, Omura G, Ando M, Sakamoto T, Yamasoba T, et al. Clinical Features of Human Papilloma Virus-Related Head and Neck Squamous Cell Carcinoma of an Unknown Primary Site. ORL. 2014;76:137-46.

[118] Vent J, Haidle B, Wedemeyer I, Huebbers C, Siefer O, Semrau R, et al. p16 expression in carcinoma of unknown primary: diagnostic indicator and prognostic marker. Head & neck. 2013;35:1521-6.

[119] Yasui T, Morii E, Yamamoto Y, Yoshii T, Takenaka Y, Nakahara S, et al. Human papillomavirus and cystic node metastasis in oropharyngeal cancer and cancer of unknown primary origin. PloS one. 2014;9:e95364.

[120] Weiss D, Koopmann M, Rudack C. Prevalence and impact on clinicopathological characteristics of human papillomavirus-16 DNA in cervical lymph node metastases of head and neck squamous cell carcinoma. Head & neck. 2011;33:856-62.

[121] Keller LM, Galloway TJ, Holdbrook T, Ruth K, Yang D, Dubyk C, et al. p16 status, pathologic and clinical characteristics, biomolecular signature, and long term outcomes in unknown primary carcinomas of the head and neck. Head & neck. 2013.

[122] Jensen DH, Hedback N, Specht L, Hogdall E, Andersen E, Therkildsen MH, et al. Human papillomavirus in head and neck squamous cell carcinoma of unknown primary is a common event and a strong predictor of survival. PloS one. 2014;9:e110456.

[123] Tribius S, Hoffmann AS, Bastrop S, Gorogh T, Haag J, Rocken C, et al. HPV status in patients with head and neck of carcinoma of unknown primary site: HPV, tobacco smoking, and outcome. Oral oncology. 2012;48:1178-84.

[124] Straetmans J, Vent J, Lacko M, Speel EJ, Huebbers C, Semrau R, et al. Management of neck metastases of unknown primary origin united in two European centers. European archives of oto-rhino-laryngology : official journal of the European Federation of Oto-Rhino-Laryngological Societies (EUFOS) : affiliated with the German Society for Oto-Rhino-Laryngology - Head and Neck Surgery. 2014.

[125] Perkins SM, Spencer CR, Chernock RD, Haughey BH, Nussenbaum B, Adkins DR, et al. Radiotherapeutic management of cervical lymph node metastases from an unknown primary site. Archives of otolaryngology--head & neck surgery. 2012;138:656-61.

[126] Compton AM, Moore-Medlin T, Herman-Ferdinandez L, Clark C, Caldito GC, Wang XI, et al. Human papillomavirus in metastatic lymph nodes from unknown primary head and neck squamous cell carcinoma. Otolaryngology--head and neck surgery : official journal of American Academy of Otolaryngology-Head and Neck Surgery. 2011;145:51-7.

[127] Demiroz C, Vainshtein JM, Koukourakis GV, Gutfeld O, Prince ME, Bradford CR, et al. Head and neck squamous cell carcinoma of unknown primary: neck dissection and radiotherapy or definitive radiotherapy. Head & neck. 2014;36:1589-95.

[128] Keller F, Psychogios G, Linke R, Lell M, Kuwert T, Iro H, et al. Carcinoma of unknown primary in the head and neck: comparison between positron emission tomography (PET) and PET/CT. Head & neck. 2011;33:1569-75.

[129] Graboyes EM, Sinha P, Thorstad WL, Rich JT, Haughey BH. Management of human papillomavirus-related unknown primaries of the head and neck with a transoral surgical approach. Head & neck. 2014.

[130] Chenevert J, Seethala RR, Barnes EL, Chiosea SI. Squamous cell carcinoma metastatic to neck from an unknown primary: the potential impact of modern pathologic evaluation on perceived incidence of human papillomavirus-positive oropharyngeal carcinoma prior to 1970. The Laryngoscope. 2012;122:793-6.

[131] Halec G, Holzinger D, Schmitt M, Flechtenmacher C, Dyckhoff G, Lloveras B, et al. Biological evidence for a causal role of HPV16 in a small fraction of laryngeal squamous cell carcinoma. British journal of cancer. 2013;109:172-83.

[132] Schmitt M, Bravo IG, Snijders PJ, Gissmann L, Pawlita M, Waterboer T. Bead-based multiplex genotyping of human papillomaviruses. Journal of clinical microbiology. 2006;44:504-12.

[133] Schmitt M, Dondog B, Waterboer T, Pawlita M. Homogeneous amplification of genital human alpha papillomaviruses by PCR using novel broad-spectrum GP5+ and GP6+ primers. Journal of clinical microbiology. 2008;46:1050-9.

[134] 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. Journal of clinical microbiology. 2010;48:143-9.

[135] IARC. Human papillomaviruses. IARC monographs on the evaluation of carcinogenic risks to humans / World Health Organization, International Agency for Research on Cancer. 2012;100B:255-314.

[136] Hofler D, Bohmer G, von Wasielewski R, Neumann H, Halec G, Holzinger D, et al. HPV16 RNA patterns defined by novel high-throughput RT-qPCR as triage marker in HPV-based cervical cancer precursor screening. Gynecologic oncology. 2015;138:676-82.

[137] Holzinger D, Flechtenmacher C, Henfling N, Kaden I, Grabe N, Lahrmann B, et al. Identification of oropharyngeal squamous cell carcinomas with active HPV16 involvement by immunohistochemical analysis of the retinoblastoma protein pathway. International journal of cancer Journal international du cancer. 2013;133:1389-99.

[138] 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. Clinical chemistry. 2005;51:1845-53.

[139] Castro F, Dirks WG, Fahnrich S, Hotz-Wagenblatt A, Pawlita M, Schmitt M. High-throughput SNP-based authentication of human cell lines. International journal of cancer Journal international du cancer. 2013;132:308-14.

[140] Rivlin N, Brosh R, Oren M, Rotter V. Mutations in the p53 Tumor Suppressor Gene: Important Milestones at the Various Steps of Tumorigenesis. Genes & Cancer. 2011;2:466-74.

[141] Schmitt M, Pawlita M. High-throughput detection and multiplex identification of cell contaminations. Nucleic acids research. 2009;37:e119.

[142] Beadle BM, William WN, Jr., McLemore MS, Sturgis EM, Williams MD. p16 expression in cutaneous squamous carcinomas with neck metastases: a potential pitfall in identifying unknown primaries of the head and neck. Head & neck. 2013;35:1527-33.

[143] Rodrigo JP, Heideman DA, Garcia-Pedrero JM, Fresno MF, Brakenhoff RH, Diaz Molina JP, et al. Time trends in the prevalence of HPV in oropharyngeal squamous cell carcinomas in northern Spain (1990-2009). International journal of cancer Journal international du cancer. 2014;134:487-92.

[144] Stein AP, Saha S, Kraninger JL, Swick AD, Yu M, Lambert PF, et al. Prevalence of Human Papillomavirus in Oropharyngeal Cancer: A Systematic Review. Cancer journal (Sudbury, Mass). 2015;21:138-46.

[145] Motz K, Qualliotine JR, Rettig E, Richmon JD, Eisele DW, Fakhry C. Changes in Unknown Primary Squamous Cell Carcinoma of the Head and Neck at Initial Presentation in the Era of Human Papillomavirus. JAMA otolaryngology-- head & neck surgery. 2016;142:223-8.

[146] Bussu F, Sali M, Gallus R, Petrone G, Autorino R, Santangelo R, et al. HPV and EBV Infections in Neck Metastases from Occult Primary Squamous Cell Carcinoma: Another Virus-Related Neoplastic Disease in the Head and Neck Region. Ann Surg Oncol. 2015.

[147] Chaturvedi AK, Engels EA, Pfeiffer RM, Hernandez BY, Xiao W, Kim E, et al. Human papillomavirus and rising oropharyngeal cancer incidence in the United States. Journal of clinical oncology : official journal of the American Society of Clinical Oncology. 2011;29:4294-301.

[148] Bosch FX, Ritter D, Enders C, Flechtenmacher C, Abel U, Dietz A, et al. Head and neck tumor sites differ in prevalence and spectrum of p53 alterations but these have limited prognostic value. International journal of cancer Journal international du cancer. 2004;111:530-8.

[149] Kimple RJ, Smith MA, Blitzer GC, Torres AD, Martin JA, Yang RZ, et al. Enhanced radiation sensitivity in HPV-positive head and neck cancer. Cancer research. 2013;73:4791-800.

[150] Qian X, Wagner S, Ma C, Klussmann JP, Hummel M, Kaufmann AM, et al. ALDH1-positive cancer stem-like cells are enriched in nodal metastases of oropharyngeal squamous cell carcinoma independent of HPV status. Oncology reports. 2013;29:1777-84.

[151] Klein CA. Parallel progression of primary tumours and metastases. Nature reviews Cancer. 2009;9:302-12.

[152] de Villiers EM. Cross-roads in the classification of papillomaviruses. Virology. 2013;445:2-10.

[153] Svajdler M, Jr., Kaspirkova J, Mezencev R, Laco J, Torday T, Dubinsky P, et al. Human papillomavirus and Epstein-Barr virus in nasopharyngeal carcinoma in a non-endemic eastern european population. Neoplasma. 2016;63:107-14.

[154] Jiang W, Chamberlain PD, Garden AS, Kim BY, Ma D, Lo EJ, et al. Prognostic value of p16 expression in Epstein-Barr virus-positive nasopharyngeal carcinomas. Head & neck. 2016;38 Suppl 1:E1459-66.

[155] Robinson M, Suh YE, Paleri V, Devlin D, Ayaz B, Pertl L, et al. Oncogenic human papillomavirus-associated nasopharyngeal carcinoma: an observational study of correlation with ethnicity, histological subtype and outcome in a UK population. Infectious agents and cancer. 2013;8:30.

[156] Dogan S, Hedberg ML, Ferris RL, Rath TJ, Assaad AM, Chiosea SI. Human papillomavirus and Epstein-Barr virus in nasopharyngeal carcinoma in a low-incidence population. Head & neck. 2014;36:511-6.

[157] Oliveira CM, Nguyen HT, Ferraz AR, Watters K, Rosman B, Rahbar R. Robotic surgery in otolaryngology and head and neck surgery: a review. Minimally invasive surgery. 2012;2012:286563.

[158] Mehta V, Johnson P, Tassler A, Kim S, Ferris RL, Nance M, et al. A new paradigm for the diagnosis and management of unknown primary tumors of the head and neck: a role for transoral robotic surgery. The Laryngoscope. 2013;123:146-51.

[159] Mirghani H, Amen F, Tao Y, Deutsch E, Levy A. Increased radiosensitivity of HPV-positive head and neck cancers: Molecular basis and therapeutic perspectives. Cancer treatment reviews. 2015;41:844-52.

[160] Chera BS, Amdur RJ, Tepper J, Qaqish B, Green R, Aumer SL, et al. Phase 2 Trial of De-intensified Chemoradiation Therapy for Favorable-Risk Human Papillomavirus-Associated Oropharyngeal Squamous Cell Carcinoma. International journal of radiation oncology, biology, physics. 2015;93:976-85.

[161] Heusinkveld M, Goedemans R, Briet RJ, Gelderblom H, Nortier JW, Gorter A, et al. Systemic and local human papillomavirus 16-specific T-cell immunity in patients with head and neck cancer. International journal of cancer Journal international du cancer. 2012;131:E74-85.

[162] Ward MJ, Thirdborough SM, Mellows T, Riley C, Harris S, Suchak K, et al. Tumour-infiltrating lymphocytes predict for outcome in HPV-positive oropharyngeal cancer. British journal of cancer. 2014;110:489-500.

[163] Albers A, Abe K, Hunt J, Wang J, Lopez-Albaitero A, Schaefer C, et al. Antitumor activity of human papillomavirus type 16 E7-specific T cells against virally infected squamous cell carcinoma of the head and neck. Cancer research. 2005;65:11146-55.

[164] Kim HS, Lee JY, Lim SH, Park K, Sun J-M, Ko YH, et al. Association Between PD-L1 and HPV Status and the Prognostic Value of PD-L1 in Oropharyngeal Squamous Cell Carcinoma. Cancer Research and Treatment : Official Journal of Korean Cancer Association. 2016;48:527-36.

[165] Bauml JM, Cohen RB, Aggarwal C. Immunotherapy for head and neck cancer: latest developments and clinical potential. Therapeutic Advances in Medical Oncology. 2016;8:168-75.

Publications

Schroeder L, Boscolo-Rizzo P, Dal Cin E, Romeo S, Baboci L, Dyckhoff G, Hess J, Lucena-Porcel C, Byl A, Becker N, Alemany L, Castellsagué X, Quer M, León X, Wiesenfarth M, Pawlita M, Holzinger D (2017). Human papillomavirus as prognostic marker with rising prevalence in neck squamous cell carcinoma of unknown primary: A retrospective multicentre study. European Journal of Cancer *74*, 73-81.

Boscolo-Rizzo P, **Schroeder L**, Romeo S, and Pawlita M (2015). The prevalence of human papillomavirus in squamous cell carcinoma of unknown primary site metastatic to neck lymph nodes: a systematic review. Clinical & experimental metastasis.

Holzinger D, Wichmann G, Baboci L, Michel A, Höfler D, Wiesenfarth M, **Schroeder L**, Boscolo-Rizzo P, Herold-Mende C, Dyckhoff G, Boehm A, Del Mistro A, Bosch F, Dietz A, Pawlita M, Waterboer T (2017). Sensitivity and specificity of antibodies against HPV16 E6 and other early proteins for the detection of HPV16-driven oropharyngeal squamous cell carcinoma. International Journal of Cancer.