The background is a vibrant, abstract watercolor illustration. It features a central profile of a human head and neck, outlined in a thin red line. The interior of the head and neck is filled with various colors and textures, including shades of pink, yellow, green, and blue. The surrounding space is also filled with these colors, creating a complex, layered effect. The overall style is artistic and modern.

# **Novel prognostic biomarkers in head and neck cancer**

Steven W. Mes



# **Novel prognostic biomarkers in head and neck cancer**

**Steven W. Mes**

The research in this thesis was performed at the department of Otolaryngology and Head and Neck Surgery of the Amsterdam University Medical Center, location VUmc.

The research was funded by the European Union's Seventh Framework Project grant 611425: OraMod.

Cover design: Pieter Goderie.

Layout and printing: Nauka.

ISBN: 9789491688157

© S.W. Mes, Haarlem, The Netherlands 2020.

No parts of this thesis may be reproduced or transmitted in any form by any means, electronic or mechanical, including photocopying, recording or any information retrieval system, without permission in writing from the author.

VRIJE UNIVERSITEIT

# **Novel prognostic biomarkers in head and neck cancer**

ACADEMISCH PROEFSCHRIFT

ter verkrijging van de graad Doctor aan  
de Vrije Universiteit Amsterdam,  
op gezag van de rector magnificus  
prof.dr. V. Subramaniam,  
in het openbaar te verdedigen  
ten overstaan van de promotiecommissie  
van de Faculteit der Geneeskunde  
op vrijdag 5 maart 2021 om 11.45 uur  
in de aula van de universiteit,  
De Boelelaan 1105

door  
Steven Willem Mes  
geboren te Amersfoort

promotoren: prof.dr. R.H. Brakenhoff  
prof.dr. C.R. Leemans

copromotor: dr. P. de Graaf

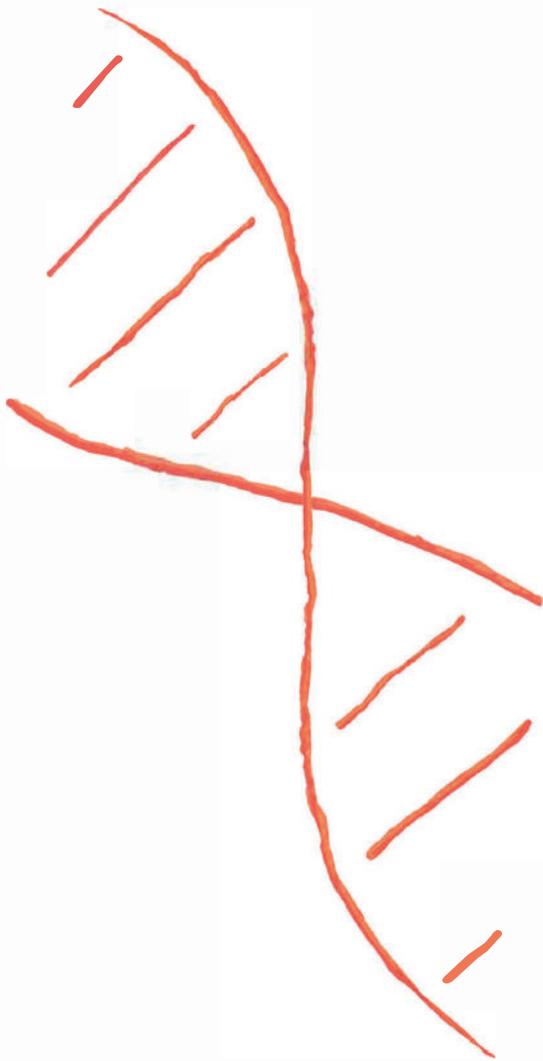
**'Every science begins as philosophy and ends as art.'**

Will Durant, 1885-1981



# Table of contents

<b>Chapter 1</b>	General Introduction	1
<b>Chapter 2</b>	Development and Validation of a Novel and Rapid Molecular Detection Method for High-Risk Human Papillomavirus in Formalin-Fixed, Paraffin-Embedded Tumor Tissue	29
<b>Chapter 3</b>	Prognostic modeling of oral cancer by gene profiles and clinicopathological co-variables	47
<b>Chapter 4</b>	Outcome prediction of head and neck squamous cell carcinoma by MRI radiomic signatures	97
<b>Chapter 5</b>	Comprehensive multiparameter genetic analysis improves circulating tumor DNA detection in head and neck cancer patients	127
<b>Chapter 6</b>	Molecular Characterization of Locally Relapsed Head and Neck Cancer after Concomitant Chemoradiotherapy	157
<b>Chapter 7</b>	General Discussion	183
<b>Chapter 8</b>	Summary	195
<b>Chapter 9</b>	Nederlandse samenvatting	201
<b>Chapter 10</b>	Addendum	207



# 1

## General Introduction

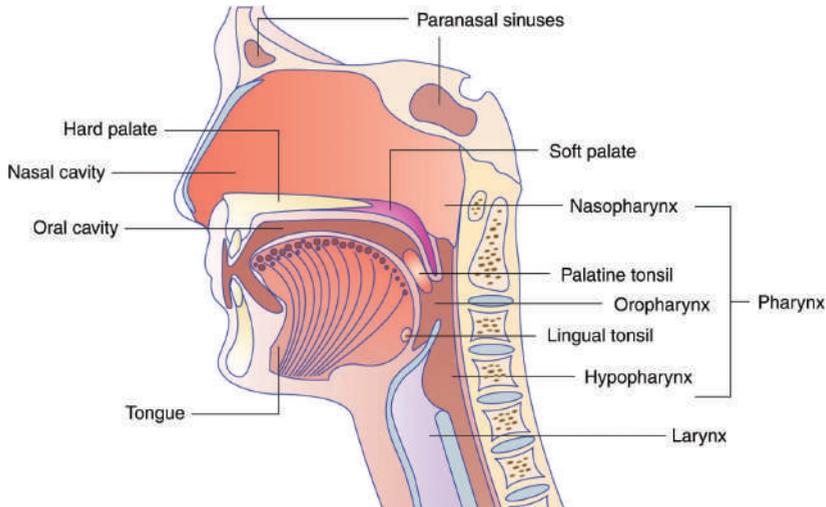
Steven W. Mes, C. René Leemans, Ruud H. Brakenhoff

Partly published in:  
**'Applications of molecular diagnostics for personalized treatment  
of head and neck cancer: state of the art'**

**Expert Review of Molecular Diagnostics 2016; 16(2): 205-21**

## EPIDEMIOLOGY

Head and neck squamous cell carcinoma (HNSCC) arises from the epithelium of the upper aerodigestive tract, i.e. the oral cavity, sinonasal cavity, pharynx and larynx<sup>1</sup>, see Figure 1. Annually, approximately 700,000 patients are diagnosed with HNSCC worldwide, accounting for 3.9% of the total global cancer incidence<sup>2</sup>. In the Netherlands, the reported annual incidence is approximately 2150<sup>3</sup>.



**Figure 1.** Anatomy of the head and neck region  
Subsites of origin of head and neck squamous cell carcinoma are indicated<sup>4</sup>.

Historically, the main risk factors included cigarette use and alcohol consumption. In oral cavity squamous cell carcinoma (OCSCC), smoking increases the risk of developing a tumor threefold. Moreover, concomitant alcohol consumption acts synergistically and further raises the risk 10- to 15-fold<sup>5</sup>. Also the use of smokeless tobacco and especially betel quid, which is commonly used in Southeast Asia and India, is associated with an increased risk for OCSCC<sup>6</sup>. In oropharyngeal squamous cell carcinoma (OPSCC), a growing fraction of tumors is caused by infection with oncogenic human papillomavirus (HPV)<sup>7</sup>. In OPSCC, attributable fractions vary between 10 and 90%, and are particularly high in the United States, Canada and Western Europe<sup>7</sup>. HPV-positive OPSCC is different at the molecular<sup>8</sup> and clinical level, with a more favorable prognosis<sup>9</sup>, and is considered a separate disease entity. The role of HPV in OCSCC is controversial. A meta-analysis showed a worldwide pooled HPV DNA estimate of 24.2% in oral cavity tumors<sup>10</sup>, but more recent studies showed lower attributable fractions, ranging from 4% to 15%<sup>11-13</sup>. This discrepancy is likely caused by different HPV detection strategies, and lower fractions were reported by studies that used more reliable testing algorithms. Finally, genetic syndromes, such as Fanconi anemia and dyskeratosis congenita, are associated with a high risk of developing HNSCC. For instance, Fanconi anemia patients have an increased risk of over 700-fold (95% CI 260–1540)<sup>14</sup>.

## CLINICAL AND HISTOPATHOLOGICAL STAGING

Clinical and histopathological classification is performed according to the TNM-classification system of the Union for International Cancer Control (UICC)<sup>15</sup> and the American Joint Committee on Cancer (AJCC)<sup>16</sup>. This system consists of three components: the dimensions of the primary tumor and invasion in surrounding structures (T), the presence, number, dimensions and extent of cervical lymph node metastasis (N), and the presence of distant metastasis (M). The TNM-classification was updated in 2017 (8<sup>th</sup> edition), and important

changes were the inclusion of a separate staging system for HPV-positive OPSCC, depth of invasion for OCSCC and extranodal extension in the N-stage of HPV-negative HNSCC. Staging of HNSCC is performed by physical examination often including investigation under general anesthesia, imaging, cytology of lymph nodes and histopathology after surgical excision.

## STAGING OF THE CLINICALLY NO NECK

Prominent lymph node metastases are discovered by palpation, but smaller metastases cannot be discriminated from negative nodes. It has been shown that approximately 30–40% of early-stage OCSCC patients have metastatic disease after pathological examination of elective neck dissection specimen, despite being clinically staged as N0 (cN0) and apparently free of lymph node metastasis<sup>17</sup>. Hence, if we rely on the clinical staging only and do not treat a cN0 neck in T1/T2 OCSCC patients, we will undertreat 30–40% of the patients. However, overtreatment will occur in 60–70% of the cN0 patients when the neck is treated electively, that is, treated when there is no clinical diagnosis of nodal disease. This dilemma fueled both imaging and molecular research and led to the development of several methods to select patients at high risk for occult metastasis while preventing overtreatment in the low-risk group.

To assist manual palpation, several radiological techniques (including ultrasound imaging, computed tomography, magnetic resonance imaging and positron emission tomography) have been tested, but for the detection of occult metastasis most lack the required accuracy<sup>18</sup>. However, a combination of ultrasound (US) combined with fine-needle aspiration cytology (FNAC) yields better performance. In this case, the nodal size on ultrasound is used to select lymph nodes at risk of nodal disease. Subsequently, these nodes are assessed by FNAC, the whole procedure abbreviated as USgFNAC. USgFNAC can achieve a sensitivity of 76% and specificity of 100%<sup>19</sup>. However, the diagnostic performance depends on the criteria used to select the lymph nodes at risk, the experience of the radiologist and pathologist, as well as the subgroup of patients studied<sup>19</sup>. In practice, the reported sensitivity in multicenter studies is generally lower<sup>20–22</sup>. The sensitivity is reduced by sampling error (in case of a micrometastasis) and aspiration of the wrong nodes. Furthermore, up to 20% of the aspirates are not evaluable due to insufficient quantity or quality of the material. To increase the sensitivity of USgFNAC and increase the number of evaluable aspirates, a real-time quantitative PCR assay was developed using squamous cell-specific marker hLy6D<sup>23</sup>. This assay was quite successful, but unfortunately false positives were introduced, possibly due to the introduction of skin keratinocytes in the aspirates and the suboptimal specificity of the marker. Other studies demonstrated markers with higher specificity like pemphigus vulgaris antigen<sup>24</sup> and squamous cell carcinoma antigen (SCCA)<sup>25</sup>. Use of these markers could reduce the number of false positives, but large studies are still awaited to prove that molecular analysis of the aspirates really improves USgFNAC.

Sampling error issues when analyzing aspirates can be circumvented by using a sentinel lymph node biopsy (SLNB), whose value in staging of cN0 OCSCC patients became increasingly evident during the last decades. This technique acts on the premise that cancer cells will metastasize to the first draining lymph node, i.e. the sentinel node (SN), in advance of spreading more widely throughout the lymphatic system of the neck. The SN is identified by visualization of a radioactive tracer that is injected in the tumor area. Subsequently, the SN is surgically removed using gamma probe, and optionally blue dye, guidance concurrent with transoral excision of the primary tumor. In some cases more than one SN can be identified<sup>26</sup>. After resection, the SN undergoes extensive histopathological examination that includes step-serial sectioning and immunostaining to detect deposits as small as isolated tumor cells. Eligible patients are diagnosed with primary tumors that can be reliably resected transorally with a clear margin without need for entry of the neck for resection or reconstruction<sup>27</sup>. Moreover, only patients who are staged cN0 by imaging are qualified for a SLNB. Removing the first draining node followed by meticulous histopathological examination should discriminate between N0 and N+ cases. A neck dissection is performed in a second surgical procedure when the SN turns out tumor-positive. Reported sensitivities are between 80% and 91% and negative predictive values (NPVs) between

88% and 94%<sup>28-30</sup>. Recently, a large multicenter study reported a sensitivity of 86% and a NPV of 95%<sup>31</sup>. A caveat is the reduced performance in tumors located at the floor of mouth due to difficult detection of the sentinel lymph node<sup>30</sup>. In addition, the need of two surgical procedures when the sentinel node is positive of tumor hampers introduction of SNLB in common clinical practice. Also, fibrosis of the neck due to SLNB may complicate future surgical procedures that will be necessary if a delayed lymph node metastasis develops in the operated region of the neck.

Reverse transcriptase PCR (RT-qPCR) assays have been developed to assess the presence of cancer cells in the removed lymph node during surgery next to scrutinizing histopathology. The accuracies of these assays are quite acceptable, ranging from 86% to 100% using conventional pathological examination as reference<sup>25,32</sup>. Unfortunately, the actual value of these assays has not been assessed in a prospective clinical setting. An intrinsic problem is that it is almost impossible to prove additional value over scrutinizing histopathology. In most studies, the lymph nodes are halved. Next, one half is analyzed by histopathology and the other half by RT-qPCR. Differences between histology and molecular assays may, therefore, be explained by sampling error instead of different test performances.

Alternative approaches have been evaluated. In studies in breast cancer patients, it was observed that metastasizing and non-metastasizing breast tumors could be distinguished at baseline on basis of their expression profiles<sup>33</sup>. In a number of subsequent papers, the value of gene expression profiling has been determined in HNSCC to predict whether a tumor has metastasized to the neck or not<sup>34-38</sup>. Gene expression signatures showed surprisingly little overlap and reported accuracies vary significantly. Many studies have not been reproduced properly, limiting possible clinical implementation. Most promising in this respect was the initial study published by Roepman et al. The authors identified a set of 852 genes that predicted N-stage<sup>38,39</sup>, and this profile was validated using a dedicated diagnostic microarray platform in a multicenter setting<sup>40</sup>. Furthermore, the authors specifically explored the performance of their signature in the clinically relevant cT1-2N0 HPV-negative oral cancer subgroup. They found a higher NPV (89%) and sensitivity (86%) in these early-stage patients compared to all patients. Nonetheless, despite the validation in multicenter setting, this profile has also not been implemented in clinical practice due to high costs and absence of an application to formalin-fixed paraffin-embedded (FFPE) tissue. In particular for the clinically relevant cT1-T2 oral cancer group, frozen samples are difficult to obtain. Furthermore, to be competitive with other techniques such as SLNB, the NPV should be preferably 95% or higher to make it the method of choice. Ultimately, a combination of both methods could be an interesting approach<sup>41</sup>.

## IMAGING

In HNSCC, imaging is used for determination of the stage of disease and in some cases for follow-up after treatment. For these indications, the commonly used modalities are ultrasound, computed tomography (CT), magnetic resonance imaging (MRI) and positron emission tomography (PET). Ultrasound is mainly used for staging of the neck<sup>42</sup>. More recently, ultrasound has also been applied to assess the depth of invasion of the primary tumor<sup>43</sup>, which might be used for staging of OCSCC.

Computed tomography uses absorption of ionizing radiation to create images of underlying structures. This technique shows very clear bone detail, but lacks soft tissue contrast. Moreover, dental implants may cause severe artifacts, which reserves CT imaging in OCSCC and OPSCC predominantly for detection of bone invasion<sup>44,45</sup>. For staging of laryngeal squamous cell carcinoma (LSCC) it can be very useful, and depending on local preferences may be the modality of choice<sup>45</sup>. Also, CT is commonly used for radiation treatment planning and a chest-CT may be acquired to exclude lung metastases in patients who are at risk for this<sup>46</sup>.

MRI is the preferred modality for visualization of soft tissue, and is particularly appropriate to evaluate the extent of the primary OCSCC, OPSCC and depending on local preferences also LSCC. MRI protocols typically

include the following sequences: T1, T2, contrast-enhanced T1, and a sequence with fat-suppression<sup>47</sup>. More recently, these conventional sequences are complemented by functional MRI such as diffusion-weighted imaging (DWI) and perfusion weighted imaging<sup>48</sup>. MRI has a relatively long examination time as compared to CT, which makes it susceptible to motion artifacts caused by patient movement, swallowing and breathing.

In HNSCC, positron emission tomography, often combined with CT, is used to detect distant metastasis in high-risk patients<sup>49</sup>, and may also be used to determine treatment response when the patient is treated with primary chemoradiotherapy (CRT)<sup>50-54</sup>. The technique measures uptake of a positron emitting tracer. The most commonly used tracer is 18F-Fluoro-deoxyglucose (18F-FDG), a marker of tissue glucose metabolism, that is increased in malignant tissue, but also in inflammation, and intrinsically high in heart and neural tissue.

## TREATMENT

Treatment protocols vary between different tumor locations and stage of disease. Early stage tumors can be treated by surgery or radiotherapy alone. Advanced stage tumors are generally treated by multimodality therapies, explained in more detail below. The mainstay of treatment for OSCC is surgery, and the goals of treatment are complete resection of the primary tumor with adequate margins and staging and treatment of the neck. The surgical approach depends on tumor location and extension. Small tumors can often be resected transorally, but locally advanced tumors and/or tumors that originate from posterior areas may require more extensive surgical approaches like a visor flap with lingual release or a lip-splitting incision with or without mandibulotomy. Moreover, marginal and segmental mandibulectomy are added to the surgery when preoperative evaluation shows of invasion of periosteum (marginal mandibulectomy) or invasion of cortex (segmental mandibulectomy). Management of the neck was traditionally performed by selective neck dissection in cN0 patients, i.e. levels I to III or levels I to IV (oral tongue), or modified radical neck dissection in cN-positive cases. More recently however, cT1N0 or cT2N0 cases may be treated with transoral excision and SLNB (see above). In the past, observation of low-risk cN0 patients was also studied, but current guidelines do not recommend observation of cN0 patients<sup>55,56</sup>. Primary radiotherapy or concomitant CRT is generally not advised for OSCC. The occurrence of osteoradionecrosis is high and acute side-effects are also higher in OSCC compared to other HNSCC subsites<sup>57,58</sup>. One study even reported lower survival rates for CRT<sup>59</sup>. Therefore, the mainstay of OSCC treatment is surgery, and in OSCC primary CRT remains mostly reserved for treatment of functionally non-resectable tumors at the time.

The oral cavity consists of important structures for speech, swallowing and facial appearance. Surgical resection affects all three functions and successful treatment in terms of quality of life therefore often involves reconstruction. Small defects can often be closed primarily or by using regional flaps, e.g. fascial artery musculomucosal (FAMM) flap<sup>60</sup> or buccal fat pad flap<sup>61</sup>. Larger defects require larger regional flaps or preferably microvascular free-tissue transfer, e.g. a radial forearm free flap or a osteocutaneous fibula free flap. Microvascular free flap reconstruction show better functional outcome than regional flap reconstruction<sup>62</sup>.

On the contrary, pharyngeal and more advanced laryngeal cancers are frequently treated with organ preservation protocols, i.e. radiotherapy or CRT<sup>63,64</sup>. Standard protocols include 60-70 Gy delivered in 2 Gy fractions 5 times per week for 7 weeks combined with high-dose cisplatin (100 mg/m<sup>2</sup>, day 1, 22, 43). Especially for HPV-positive OPSCC this protocol is very effective with high survival rates. The successful treatment of these patients even evolved in treatment de-escalation trials to minimize treatment related side-effect, but unfortunately several large trials showed lower survival rates after treatment with adjusted protocols<sup>65,66</sup>.

Other treatment options are transoral laser surgery for early stage LSCC<sup>67</sup>, hypopharyngeal squamous cell carcinoma (HPSCC)<sup>68</sup> and OPSCC<sup>69</sup>, and total laryngectomy with or without partial pharyngectomy for advanced laryngeal and hypopharyngeal cancers (T4a). More recently, transoral robotic surgery became available for early stage OPSCC and unknown primary tumors with adequate tumor control<sup>70,71</sup>. Advantages

are that patients can be treated with a single procedure and that they may experience better functional results compared to primary radiotherapy. However, a recent randomized trial did not confirm this and instead showed slightly better swallowing-related quality of life scores in patients treated with primary radiotherapy<sup>72</sup>.

Depending on tumor-specific risk factors, surgical management of HNSCC may require adjuvant treatment. Criteria for postoperative (chemo)radiotherapy are locally advanced disease, pN2 or pN3 disease, perineural invasion, extracapsular spread and tumor-positive margins. Adjuvant treatment with radiotherapy improves overall survival in these patients with poor prognosis<sup>73</sup>. Side-effects of this treatment include xerostomia and osteoradionecrosis, which are less common with intensity-modulated radiation therapy (IMRT)<sup>74</sup>. This radiotherapy technique is considered the standard nowadays. In the future, patients may also benefit from proton beam therapy to further reduce side effects as demonstrated in normal tissue complication probability models<sup>75-77</sup>. Moreover, an additional survival benefit has been found in patients with tumor-positive margins, multiple positive lymph nodes or extracapsular spread for postoperative, platinum-based CRT<sup>78,79</sup>.

Long-term survival in patients with recurrent or metastatic HNSCC (R/M HNSCC) is possible in the minority of patients when salvage surgery or reirradiation is possible<sup>80</sup>. Several palliative treatment regimens exist, but the cornerstone has been the combination of cisplatin and fluorouracil (FU) for many years. In the last decades, the addition of cetuximab, which is a monoclonal antibody directed against the epidermal growth factor receptor (EGFR), was shown to establish a survival benefit<sup>81</sup> but unfortunately long-term responses are very sporadic. Even more recently, immunotherapy was added to the pallet of systemic options for R/M HNSCC. Several monoclonal antibodies were shown to induce long-term responses in approximately 20% of patients with platinum refractory R/M HNSCC<sup>82-84</sup>. Future studies will establish the exact regimen of immunotherapy and also focus on optimal patient selection<sup>85</sup>.

## PROGNOSIS

The 5-year survival of HNSCC is 58% in the Netherlands in the period of 2007 to 2011<sup>3</sup>, but these survival rates vary from 10% to 98% depending on site, stage, age, comorbidity and subsite<sup>86</sup>. Despite advances in diagnostic and therapeutic management of HNSCC patients, overall survival in the Netherlands improved only 1% in the past 20 years<sup>3</sup>. The most important predictor of overall survival is the presence of HPV, but only in OPSCC<sup>13,87,88</sup>. For other sites, overall survival highly depends on the presence of metastatic LNs and ECS. 5-year overall survival drops from 75% in pN0 patients to 50% in pN-positive and ECS-negative patients and to 30% in ECS-positive patients<sup>89</sup>. This effect is caused by an increase of locoregional recurrence<sup>90</sup> and distant metastases<sup>91</sup>, and possibly treatment-related morbidity<sup>92</sup>.

The most important prognostic factors at present are disease stage including ECS, HPV and histological findings after surgery. To assess stage, accurate diagnostic modalities are critical but meet limitations. Consequently, prognostic models are imprecise and treatments cannot be adequately tailored to the individual patient. For these reasons there is active research into molecular biomarkers to improve diagnosis, to predict treatment responses and to predict prognosis in general more accurately.

## MOLECULAR TOOLBOX

Historically, to detect genetic alterations, specific genes or genomic regions of interest were examined individually, for example, by Sanger sequencing (point mutations) and microsatellite analysis (patterns of allelic loss). These techniques provided recurrent changes, but generally at a low resolution and with labor-intensive methods, which hampered a comprehensive identification of candidate cancer genes and associated genetic changes. Moreover, the cloning of these genes was cumbersome and the molecular tools to characterize these genes were limited to ectopic expression in cell lines.

After its discovery in 1992, comparative genomic hybridization (CGH)<sup>93</sup> made the detection of copy number aberrations (CNAs) accessible on a genome-wide scale albeit with limited resolution. Additionally, after the implementation of array CGH, with DNA microarrays replacing the original metaphase chromosome spreads, it even became possible to analyze CNAs across the entire genome with initially moderate and later ultra-high resolution<sup>94-96</sup>. The same hybridization-based techniques were used to determine gene expression of many genes simultaneously in a tumor or other sample<sup>97,98</sup>. Later, platforms became available for genome-wide expression profiling, supporting integral molecular profiling of the transcriptome. During the last years, next-generation sequencing (NGS) has taken over the descriptive genomics field and is currently used to determine somatic mutations (whole-exome sequencing), gene expression profiles (RNA sequencing) and copy number alterations (low-coverage whole-genome sequencing).

Some limitations came along with this era of high throughput descriptive genomics. First of all, the costs are still very high. To be able to clinically implement results of high-throughput studies, researchers have developed dedicated research platforms that are less expensive; examples are the MammaPrint dedicated microarray platform<sup>99</sup> and Ampliseq Cancer Hotspot Panel for targeted sequencing<sup>100</sup>. Another possibility has been explored by Nielsen et al.<sup>101</sup> They developed an immunohistochemistry assay to distinguish the molecular classes of breast cancer that were previously described by Perou et al.<sup>102</sup> on the basis of expression profiles.

A second limitation involves the interpretation of the data and the validation of results. New statistical techniques had to be developed to handle the statistical challenges of big data research. Interpretation became more difficult for biologists and clinicians, emphasizing the important role of biostatisticians and bioinformaticians. Initial studies were usually underpowered and suffered from a lack of thorough validation in independent series. Finally, the results of microarray analysis are affected by the use of different platforms, different experimental protocols and many other variables. Quite some effort has been made to improve the reproducibility, but it is far from optimal<sup>103</sup>. NGS platforms are expectantly less sensitive to these difficulties, but the available gene expression data sets are still limited.

Besides the descriptive genomics toolbox, also the functional genomics toolbox increased tremendously with the introduction of retroviral and lentiviral vectors, germline modification of the mouse, small interference RNA gene expression knockdown and recently genome editing by, for example, CRISPR/Cas9 technology. The evolving molecular tools have been used to investigate head and neck carcinogenesis, and the molecular landscape of head and neck cancer.

## CARCINOGENESIS OF HPV-NEGATIVE HNSCC

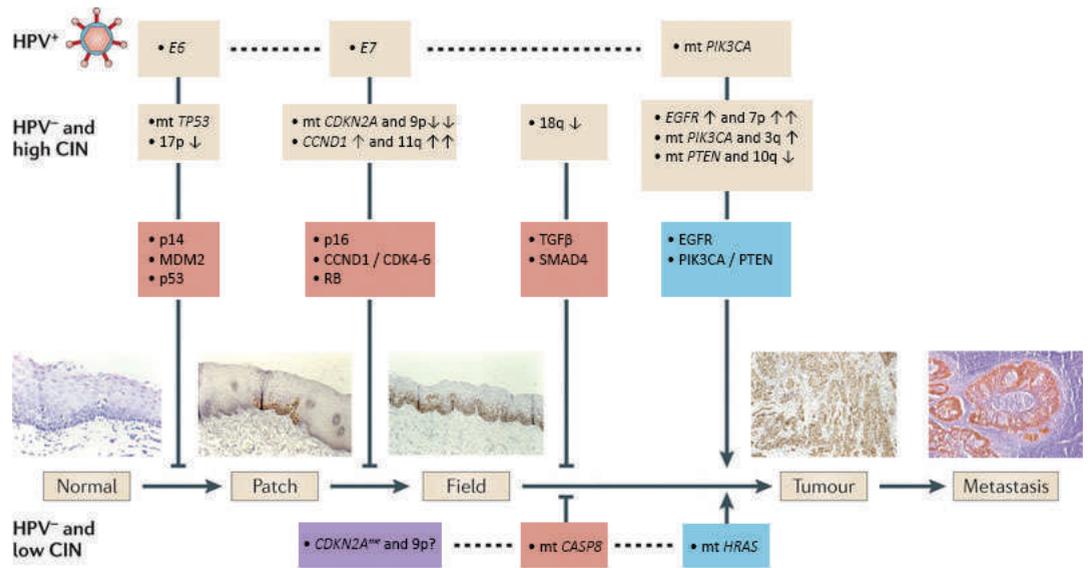
Cancer, including HNSCC, is caused by the accumulation of genetic and epigenetic alterations of genes in various signaling pathways. The most frequently affected pathways in HNSCC involve the cell cycle checkpoints (e.g. p53/ pRb pathways), RTK/RAS/PI(3)K signaling (e.g. epidermal growth factor receptor (EGFR), PIK3CA, PTEN), differentiation (e.g. NOTCH1, FAT1, TP63) and cell death (e.g. FADD, CASP8). Most genes involved can be classified as tumor suppressor genes. An important exception is PIK3CA, which harbors activating mutations or copy number gains in one-tenth to one-third of the HPV-negative tumors<sup>104-106</sup>.

In 1996, Califano et al.<sup>107</sup> presented the first stepwise genetic progression model of HNSCC, based on molecular analysis of histological premalignant mucosal changes and invasive cancers. They identified that the number of chromosomal aberrations (i.e. loss of heterozygosity (LOH) at specific loci) increases with histopathological progression. Furthermore, specific alterations were recurrently found in histological lesions that represent the early steps in the development of invasive cancer, that is, LOH at chromosomes 3p, 9p and 17p. These early genomic alterations were also the most frequent changes in the invasive tumors.

Mutations in TP53 (17p) can be found in 60–84% of the HPV-negative tumors<sup>104,105,108</sup> and mutations and/or focal deletions of CDKN2A (9p) in 57%<sup>105</sup>. Deletions of chromosome 3p (often the entire arm) can be found

in 80% of the tumors<sup>105</sup>. Remarkably the critical cancer gene on 3p is still a mystery. Other genetic alterations were found in more advanced premalignant lesions and invasive cancer and are believed to occur later during carcinogenesis. This initial model was later revised by Braakhuis et al.<sup>109</sup> that is described below (Section 'Field cancerization' and Figure 2).

Recent NGS studies revealed many more candidate cancer genes. Alterations in these genes are commonly shared, but are generally less frequent and sometimes very infrequent. Examples are mutations in NOTCH1 (26%) and FAT1 (32%). An explanation for this finding might be that cell cycle regulation needs to be disturbed by few specific key players as an early step in the carcinogenesis process, whereas the other hallmarks of cancer are likely regulated by more divergent pathways and genes.



**Figure 2.** Progression model of molecular carcinogenesis of oral squamous cell carcinoma.

An adapted version of the genetic progression model of head and neck squamous cell carcinoma (HNSCC)<sup>110</sup>.

A genetically altered stem cell forms a 'patch' of clonally related daughter cells with the same genetic alteration (e.g. mutation in TP53). Such a patch can, for example, be detected by immunostaining for mutated p53. Subsequent genetic changes provide a survival benefit or enhanced proliferation for the altered stem cells and they laterally replace the normal epithelial stem cells, which results in the formation of a larger precancerous field. As the field becomes larger, additional genetic hits give rise to various subclones within the field (clonal divergence) that take over the field. Finally, a subclone acquires sufficient molecular alterations to transform into an invasive tumor cell. Three critical steps can be discriminated in this model: the first mutation causing a genetically altered patch, the outgrowth of a single mutated stem cell into a group of mutated stem cells generating the field and the transformation of a premalignant field into invasive cancer. Predictors of malignant transformation are aneuploidy and the accumulation of cancer-associated genetic changes. Well-known altered pathways in HNSCC are depicted for the three genetic subtypes that are now distinguished, that is, human papillomavirus (HPV)-positive HNSCC, HPV-negative HNSCC with few copy number alterations (low chromosomal instability (CIN)) and HPV-negative HNSCC with many copy number alterations (high CIN). HPV-negative, low CIN tumors are characterized by enrichment of HRAS activating mutations, CASP8 inactivating mutations and absence of TP53 mutations. Future studies are necessary to further characterize these tumors molecularly, particularly to determine the early steps in carcinogenesis.

Moreover, many recurrent genomic alterations cannot be depicted in the progression model to date. For some alterations, the timing during carcinogenesis is known, but the involved genes and pathways are not. Examples are 3p loss (early step), 7q gain (late step) and 8p loss (late step). For other alterations, the involved genes and pathways are known, but their timing is not. Examples are inactivating mutations in FAT1, AJUBA, NOTCH1 and activation of FGFR1 (HPV-negative tumors) and FGFR3 (HPV-positive tumors).

Important genetic and chromosomal alterations are indicated in the upper yellow boxes. A distinction is made between oncogenic pathways (blue boxes) and tumor-suppressive pathways (orange boxes). ↑ indicates overexpression or gain; ↑↑ indicates high-level amplification; ↓ indicates loss; and ↓↓ indicates homozygous loss. CASP8: caspase 8, apoptosis-related cysteine peptidase; CCND1: cyclin D1; CDK: cyclin-dependent kinase; CDKN2A: cyclin-dependent kinase inhibitor 2A (p16); EGFR: epidermal growth factor receptor; HRAS: Harvey rat sarcoma viral oncogene homolog; mt: mutated; PIK3CA: phosphoinositide-3 kinase subunit-α; PTEN: phosphatase and tensin homolog; TGFβ: transforming growth factor-β.

## CARCINOGENESIS OF HPV-POSITIVE HNSCC

The HPV genome consists of six early genes (E1–E7) and two late genes (L1 and L2). Two early genes (i.e. E6 and E7) have a role in oncogenesis<sup>111</sup>, particularly when expressed in the basal epithelial layers. The E6 protein binds to tumor suppressor protein p53 causing its inactivation and degradation<sup>112</sup>. The E7 protein binds to the RB pocket proteins also promoting their degradation<sup>113</sup>. By E6 and E7 gene expression, the virus bypasses the G1/S block, permitting S-phase entry and viral replication<sup>110</sup>. The induction of these processes by the viral oncogenes can be regarded as early changes in the carcinogenesis of HPV-positive tumors, comparable to mutations and/or deletions of TP53 and CDKN2A, the latter acting in the Rb pathway (see Figure 2). Evidence from studies of cervical cancer indicated that while E6 and E7 are necessary to initiate and maintain the transformation of normal epithelial cells to invasive cancer, expression of these viral oncogenes is not sufficient to develop a full-blown tumor. The alteration of additional oncogenic pathways is required for malignant progression<sup>114</sup>. However, the modification of additional pathways is accelerated by the chromosomal instability that is caused by E6 and E7 expression<sup>115</sup>. The deregulated cellular S-phase entry promotes DNA damage, and when p53-mediated DNA damage response is abrogated by E6, this may lead to DNA changes<sup>116</sup>. Frequently affected pathways in HPV-positive HNSCC involve RTK/RAS/PI(3)K signaling (e.g. PIK3CA, FGFR3), differentiation (e.g. NOTCH1, TP63) and cell death (e.g. TRAF3)<sup>105</sup>.

In a recent review a model of HPV infection was proposed with productive virus producing infections in the oral cavity mucosa, and transforming infections in the oropharyngeal crypt epithelial cells<sup>117</sup>. This model explains most research findings on HPV-mediated head and neck carcinogenesis.

## FIELD CANCERIZATION

Besides describing the progression of oral mucosa to invasive cancer, Califano et al. noted that some genetic alterations were shared between the invasive tumors and the associated surrounding abnormal (noninvasive) mucosal cells, which suggests that these cells originated from a common progenitor<sup>107</sup>. This finding supported the concept of ‘field cancerization’, hypothesized by Slaughter et al. in 1953 to explain the frequent development of local recurrences and SPTs in oral squamous cell carcinoma patients<sup>118</sup>. They examined 783 resected oral tumors and concluded that these cancers were often surrounded by histologically abnormal epithelium possibly caused by long-term carcinogen exposure. Moreover, the finding of multiple independent invasive foci suggested a multiclonal origin of squamous cell cancer. The genetic explanation of the field cancerization concept was studied in depth and reported in 2001 by Tabor et al.<sup>119</sup> These authors analyzed biopsies of tumors and macroscopically normal adjacent mucosa and found shared genetic alterations in 50% of the cases. These genetically abnormal cells were coined as fields. In 25% of the cases, these fields extended into the surgical margins and apparently remained behind, causing local relapses and SPTs<sup>120,121</sup>.

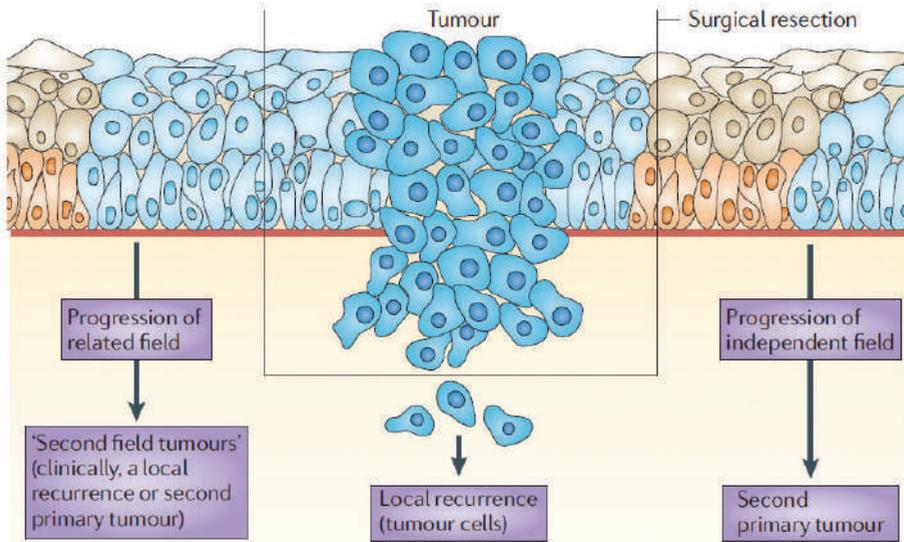
Subsequent research of immunohistochemical staining patterns of mutated p53 allowed microscopic visualization of the likely earliest changes in the mucosal epithelium<sup>122</sup>. These alterations presented as small patches that preceded the fields. This observation refined the progression model for HNSCC<sup>109</sup>, which is depicted schematically in Figure 2. Of note, these models are mainly based on oral cancer and oropharyngeal cancer, but there are no reasons to assume that laryngeal or hypopharyngeal tumors follow other models. It was hypothesized that a mucosal stem cell becomes genetically altered and forms a ‘patch’ of clonally related daughter cells with the same genetic alteration (TP53 mutation). This clonal unit concept has recently been demonstrated by lineage tracing experiments in the skin epithelium in mice<sup>123</sup>. Subsequent alterations would provide a survival benefit or enhanced proliferation to the genetically altered stem cell, thereby laterally

replacing the normal epithelial stem cells. This results in the formation of a larger precancerous field. Some of these fields become visible and present as white (leukoplakia) or red (erythroplakia) lesions. As the field becomes larger, additional genetic hits give rise to various subclones within the field (clonal divergence). These subclones may take over the field and ultimately acquire sufficient genetic alterations to transform into an invasive tumor.

The concept of field cancerization seems at present to be specific for HPV-negative tumors, but has also been studied for HPV-positive tumors under the assumption that HPV infection is the first event in HPV-mediated carcinogenesis and that the presence of transcriptionally active HPV could serve as reliable biomarker of field cancerization. In this study, HPV16-E6-mRNA could not be identified in resection margins of 20 HPV-positive cancers<sup>124</sup>, suggesting that HPV-positive tumors are not surrounded by these large fields. Although the apparent absence of fields surrounding HPV-positive tumors is in line with the favorable prognosis of HPV-positive disease, it strongly relies on the assumption that HPV infection indeed is the starting event. There are also some contrasting observations. There are reports of cases presenting with multiple synchronous HPV-positive HNSCCs (e.g. McGovern et al.<sup>125</sup>). This could be explained by the presence of preneoplastic fields, although multiple transforming infections could explain this as well as suggested in a recent HPV carcinogenesis model by Leemans et al.<sup>117</sup> This model proposed that the oral cavity is the major site of productive infections, while transforming infections occur in specific oropharyngeal cells. Altogether, the concept of field cancerization in HPV-positive tumors remains an interesting topic for future studies.

## **EARLY DETECTION OF RECURRENT DISEASE BY MOLECULAR DIAGNOSIS**

As stated above, the frequent development of local and regional recurrences is a major problem in the treatment of HNSCC patients. Depending on stage and site, local and/or regional recurrence rates have been reported of 20–50%<sup>126,127</sup>. Moreover, distant metastasis occur in 10–40%<sup>128</sup> during follow-up. Remarkably, local recurrences occur even when the surgical margins are histologically tumor-free. Two explanations of this relatively high incidence of locoregional recurrence have been suggested<sup>110</sup>. First, residual tumor cells that are undetectable by standard histological and radiological examination can remain in situ after surgery. These residual cells have been defined as minimal residual cancer (MRC). Furthermore, it has been shown and described above that tumors develop in precancerous fields that can be 5–10 cm in diameter. Part of this precancerous field can extend into the surgical margins and as most are not macroscopically visible by the naked eye, they remain behind unnoticed. Additional genetic alterations can cause outgrowth of a second tumor from the residual premalignant field ('second field tumor'). Second field tumors can occur in close proximity of the primary resected tumor and within 3 years. These tumors are then clinically diagnosed as local recurrences. Second field tumors can also develop at more distant sites or after 3 years and are then diagnosed as SPTs. Hence, local recurrences can be subdivided in two categories, those resulting from MRC and those resulting from fields that remained behind. The different types of recurrent disease are illustrated in Figure 3.



**Figure 3.** Field cancerization and local relapse.

Field cancerization and its role in the development of local recurrences and second primary tumors is shown<sup>110</sup>. Premalignant fields are mucosal areas with epithelial cells that have cancer-associated genetic or epigenetic alterations. Precursor fields, which are shown in light blue, consist of cells that do not show invasive growth despite their cancer-associated genetic changes. An important clinical implication of a field is that it may be the source of local recurrences and second primary tumors after surgical resection of the initial carcinoma. This figure describes the three sources of recurrent disease: true local recurrence that derives from cancer cells that stayed behind; a second tumor from premalignant cells that stayed behind and is clonally related to the index tumor ('second field tumor'); and a second tumor from unrelated premalignant cells ('second primary tumor'). At present, the distinction is made on the basis of clinical criteria. Local recurrences are defined as tumors that arise within 2 cm and within 3 years of the original primary tumor. Second primary tumors are defined by any tumor that arises beyond 2 cm of the primary tumor or after 3 years at the same location<sup>120</sup>. The distinction between local recurrences and second field tumors can only be made with molecular analysis. Index tumor, premalignant field in corresponding surgical margins and recurrent tumor should be of common clonal origin for the recurrent tumor to be considered a second field tumor. Less laborious techniques to distinguish between the two types of local recurrence would be helpful, but have not been developed at the time. The distinction does not have clinical consequences at present, but this will change when targeted chemopreventive treatment for fields becomes available.

## DETECTION OF MINIMAL RESIDUAL HEAD AND NECK CANCER

First attempts to detect residual cancer cells in margins that were histologically free of tumor were published 20 years ago by Brennan et al. These authors used TP53 mutations as molecular marker, which was elegantly detected by 'plaque assays'<sup>129</sup>. The authors showed that mutated TP53 in the surgical margins is associated with a higher risk of locoregional recurrence. In many subsequent studies, the paradigm remained unchanged: detection of TP53 mutations (if present in the primary tumor) in the corresponding margins is a risk factor for locoregional recurrence<sup>130,131</sup>. However, techniques based on detection of mutated DNA lack specificity. Partly, this can be explained by the use of mutated TP53. This is an early change in carcinogenesis and will not detect only residual tumor cells but also residual precancerous fields that do not always progress. What is more, residual tumor cells should be treated by postoperative radiotherapy, but this is not indicated for residual fields. Furthermore, mutated DNA leaks from necrotic tumors. In particular, the latter combined with the high sensitivity of the methods causes specificity problems. Other research groups have suggested to exploit other DNA markers such as methylation markers<sup>132</sup> or microsatellite markers<sup>133</sup> to determine the presence of minimal residual disease, but the performance was similar to mutated TP53 analysis and most studies were small. We assume that next generation sequencing approaches may revive this research field.

DNA is a very stable molecule that easily causes contamination problems. Therefore, RNA-based assays to detect MRC have been evaluated as well. These tests were expected to be more specific than DNA-based

assays<sup>130</sup>. Several genes were considered promising markers since they were highly expressed in HNSCC tumor cells, but not in stromal cells. Examples are hLy6D, SCCA and EGFR. These RT-PCR-based assays usually have a very high sensitivity and are, for example, able to detect a single tumor cell in a background of  $10^7$  nucleated blood cells<sup>134</sup>. However, a prospective study including 105 cases failed to show prognostic relevance of hLy6D expression in deep surgical margins<sup>135</sup>. Even when the tumors with high hLy6D expression were analyzed as subgroup and between 5 and 15 biopsies per patient were analyzed, there was no association with clinical outcome.

One of the major issues that was indicated by the authors and has to be addressed in future research is the sampling of these margins. Even 15 biopsies are subject to sampling-error given the large resections. A more comprehensive way of sampling the margins could overcome the false-negative results. Wrapping the specimen in nitrocellulose and analysis of the DNA by methylation markers has been tried, but thus far only in very small series<sup>136</sup>.

Despite that the concept of detecting residual tumor cells in tissue samples by molecular methods is very attractive, the sampling error problem, which also hampers histological examination, is an issue that is not easily solved. RT-qPCR methods using histogenic markers have the additional disadvantage that false-positive results are easily found due to the presence of normal epithelial cells in the superficial margins. Tumor-specific DNA markers (e.g. TP53 mutations) have the limitation that DNA from necrotic tumor cells contaminate the samples, causing false-positive results, and might detect precancerous changes in superficial margins. Hence, the results of all these studies still leave much to be desired. In later studies, expression profiling studies to predict local recurrences have been tried as well, but the associations were generally not strong and profiles have not yet been adequately validated<sup>137,138</sup>.

Nonetheless, the gains are high: postoperative radiotherapy or nowadays postoperative chemoradiation could be tailored on the basis of the presence of residual disease. An intrinsic problem remains that local recurrences may develop from two different sources: tumor cells and precancerous fields that stayed behind. This has major biological but also clinical consequences. Hence, it will be required to make a separation between residual tumor cells and residual fields, particularly when superficial margins containing mucosal epithelium are analyzed.

The fact that DNA is stable and easily contaminates margin samples has an intriguing other application. The shedding of DNA might well result in tumor-specific alterations in the blood. Already in the 1990s microsatellite alterations in the serum of HNSCC patients were discovered<sup>139</sup>. Since then, new techniques have been developed to detect what is now called circulating tumor DNA (ctDNA). At present, two frequently applied techniques are NGS and digital PCR. Relatively few studies focused on the detection of ctDNA in HNSCC. In 2014, Bettgowda et al.<sup>140</sup> explored the detectability of ctDNA in 15 types of cancer including 10 advanced stage HNSCC patients. CtDNA was detectable in the majority, but results are difficult to interpret given the small number of cases. More recently, Wang et al.<sup>141</sup> published a study that focused specifically on the detection of ctDNA in HNSCC. They performed targeted sequencing of recurrently mutated genes and the HPV16 E6 gene in DNA extracted from plasma (47 patients) and saliva samples (93 samples). Depending on the subsite, ctDNA in plasma was detectable in 80–100% of the cases, whereas tumor DNA was detected in 47–100% in saliva. Moreover, post-treatment saliva samples in a subset of cases ( $n = 9$ ) showed positive tumor DNA in advance of clinical relapse in 4/9 cases. The remaining five cases had no detectable tumor DNA in their saliva during follow-up and remained clinically disease free. These results suggest new approaches for early detection of residual/recurrent disease. However, numbers are small and postoperative saliva sampling was performed at random time points. Longitudinal sampling combined with optimized detection methods will be required to assess the clinical implications.

## DETECTION OF RESIDUAL PRENEOPLASTIC FIELDS

Residual preneoplastic fields are responsible for approximately 1/3 of local relapses. Moreover, SPTs in the upper aerodigestive tract occur in an annual rate of approximately 2–3%. SPTs, too, can derive from residual precancerous fields. Second primary tumors can either be clonally related to the first tumor or derive from an independent field. Early detection of these (residual) preneoplastic fields could provide opportunities to prevent its progression to invasive carcinoma, for example, by chemoprevention. Several methods have been published to identify fields and assess the risk for malignant progression. Importantly, histological changes (e.g. dysplasia) have been shown to be suboptimal to detect premalignant changes or predict malignant potential of the fields. Possibly, the recently recognized morphological change indicated as ‘differentiated dysplasia’ might change the picture<sup>142</sup>.

Other methods have been evaluated that typically focus on the detection of genomic alterations in tissue biopsies. This can be tested in biopsies from both visible lesions (e.g. leukoplakia) and the mucosal margins of surgical specimens. Retrospective studies as early as 1996 have convincingly shown that detection of specific genetic alterations (e.g. LOH at chromosome 9p and 3p) are important for risk assessment of oral precancer<sup>143-145</sup>. These findings were also found to be applicable for analyses of surgical margins<sup>146</sup>. Moreover, prospective evaluation of these markers was successful<sup>147,148</sup>, ultimately demonstrating the added value of detection of these genomic alterations in premalignant fields. Remarkably, this has not become the new clinical standard. Clinical implementation is hampered mainly by the fact that, despite excellent prognostic stratification of patients, for the individual patient it does not have clinical consequences. Although not very effective, it is generally accepted to remove visible preneoplastic lesions (leukoplakia) surgically, followed by active surveillance<sup>149</sup>. This policy is not affected by molecular stratification of patients. This may change when more effective, but also more toxic, treatments become available. When effective therapeutic interventions have been developed and tested to reduce malignant transformation and improve overall survival, a simple LOH-based or NGS-based test might become an important companion diagnostic assay.

Also, screening for and monitoring of fields might become an option by noninvasive diagnostic approaches using brushed cells or saliva. Promising results were shown in a study on the detection of specific genomic alterations in exfoliated cells sampled by brush cytology. In a pilot study using allelic imbalance of cancer-associated chromosome regions assessed by 12 microsatellite markers, an estimated sensitivity of 78% and a specificity of 100% was determined<sup>150</sup>. However, in a larger follow-up study the analytical sensitivity reduced to 45%<sup>148</sup>. Future studies with more sensitive NGS techniques may improve these test characteristics.

Besides noninvasive molecular diagnosis, also visualization of premalignant fields may be very helpful. Poh et al. showed that loss of autofluorescence in tumor adjacent mucosa is associated with histological and molecular premalignant changes<sup>151,152</sup>. Furthermore, these authors developed a handheld device to determine loss of fluorescence visualization in the operating theater and are currently assessing its value in a randomized controlled trial (COOLS trial)<sup>153</sup>. Notwithstanding, detecting and risk assessment of fields is one step, but new and perhaps targeted chemopreventive therapies need to be developed and translated from early experimental phases to clinical trials.

In conclusion, the detection of residual precancerous fields is possible and clinically relevant when using tissue biopsies. The applied methods have been validated thoroughly. Future research should focus on noninvasive detection of these fields. A noninvasive test could also be applied to invisible premalignant fields, which could be specifically interesting for screening of high-risk subjects. However, in parallel to diagnostic approaches, chemopreventive therapies have to be developed, allowing detected precancerous changes to be treated.

## MOLECULAR CLASSIFICATION OF HNSCC: HPV

As indicated above, molecular detection and risk assessment of preneoplastic changes is quite effective. In contrast, detection of residual disease has not been very successful to identify patients at high risk for locoregional recurrence or distant metastases. The intrinsic issue of sampling error hampers applications and other approaches are required. With the upcoming of reliable, high-throughput profiling methods, molecular stratification of HNSCC seems to be the most interesting alternative approach. The most pronounced molecular subgroup in HNSCC is defined by HPV status. Intriguingly, these tumors show very different molecular profiles including a lack of TP53 mutations and fewer chromosomal aberrations than typically present in HPV-negative HNSCC<sup>154-157</sup>. Moreover, HPV-positive and -negative tumors can be distinguished by gene expression profiles<sup>158</sup> and epigenetic signatures<sup>159,160</sup>. Updates of these molecular differences using NGS have recently been published by the Cancer Genome Atlas<sup>105</sup> and Seiwert et al.<sup>106</sup> The majority of previous findings were confirmed. Additionally, new findings provided a more comprehensive overview of the molecular landscape of HPV-positive HNSCC. In contrast to results of previous studies<sup>104,108</sup>, the absolute number of somatic mutations was found to be comparable between HPV-positive and HPV-negative tumors<sup>105</sup>. However, as expected, the landscape of somatic mutations and copy number alterations differed importantly. Altogether, HPV-positive tumors can be considered a very well-validated molecular subgroup of HNSCC with major impact on prognosis. In fact, HPV-positive OPSCC patients have such a favorable prognosis that treatment de-intensification trials were initiated.

The increasing importance of HPV-positive tumors as a separate molecular and clinical subgroup of HNSCC also fueled research for reliable HPV detection methods, particularly in archival FFPE tumor specimen. In theory, this is easily determined by HPV DNA detection, but in practice this turned out to be quite a challenge. Most assays have been borrowed from the cervical cancer screening research field and these PCR-based assays are generally very sensitive, causing false-positive results. The generally accepted 'gold standard' test for tumor tissue is the detection of E6 and E7 mRNA expression detected by RT-PCR. The RT-PCR assay can be reliably performed on mRNA from fresh-frozen tumor tissue. However, these standard assays are less reliable on mRNA that is isolated from FFPE tissue since the RNA is degraded and of poor quality. This is a major disadvantage because often only FFPE material of tumors is available. Hence, other HPV detection methods have been developed with variable performances.

The most frequently applied methods are PCR-based approaches to detect HPV DNA. However, due to the high sensitivity, these methods result in false-positive results<sup>161</sup>. A second, commonly utilized method is fluorescence in situ hybridization. This test is highly specific, but relatively difficult to implement and the sensitivity is suboptimal (83–88%)<sup>162,163</sup>. The third procedure relies on the detection of surrogate markers in combination with HPV detection assays. The most widely applied surrogate marker is the expression of p16 (encoded by CDKN2A). In OPSCC, there is a high correlation between p16 expression and the presence of a transcriptionally active HPV infection. However, this test lacks specificity when used as individual test (false-positive rates between 15% and 20%)<sup>162-164</sup>. Algorithms have been developed to overcome the weaknesses of using individual tests. Smeets et al.<sup>163</sup> assessed the performance of several combinations of individual HPV tests in OPSCC. They found that an algorithm of p16 expression followed by detection of high-risk HPV DNA by GP5+/6+ PCR on the p16-positive samples showed the highest correlation to E6/E7 mRNA expression in the corresponding frozen specimen. This was later confirmed in a follow-up study of independent cases<sup>165</sup>. Thus far, no algorithm has been developed for other subsites besides the oropharynx. The sensitivity of immunostaining for p16 in subsites other than the oropharynx seems to be much lower<sup>12,13</sup>. Hence, a reliable and simple test for HPV applicable to all tumor sites still needs to be developed and, more importantly, thoroughly validated. This would enable us to obtain reliable figures on incidence and prevalence of HPV-positive tumors outside the oropharynx and, determine whether HPV presence is also a major prognostic

marker for non-OPSCC. For OPSCC, test algorithms that include p16-immunostaining followed by some kind of HPV DNA detection can be considered as well validated and reliable.

## MOLECULAR CLASSIFICATION OF HNSCC: OTHER SUBCLASSES

Besides the subclass of tumors defined by HPV-status, a third molecular subgroup is emerging. This subgroup was first discovered by Smeets et al.<sup>166</sup> and is characterized by a very limited number of genetic changes and wild-type TP53 status. Tumors are HPV-negative and arise primarily in the oral cavity. A recent report of the TCGA<sup>105</sup> confirmed that approximately 15% of the cases showed very few chromosomal aberrations (somatic copy number alteration (SCNA) quiet). Moreover, this subset was enriched with activating mutations in HRAS, inactivating mutations in CASP8 and wild-type TP53. Molecular findings are supported by clinical behavior as this SCNA quiet subtype has a markedly favorable prognosis compared to other HPV-negative tumors<sup>105</sup>.

The described stratifications rely on HPV presence, genetic changes and mutational profiles. In the landmark study by Perou et al.<sup>102</sup> subclasses of breast cancer were identified using expression profiling, an observation that changed the breast cancer research field. Along these lines, several groups have studied the molecular classification of HNSCC by gene expression signatures. The first study was performed by Chung et al. using microarray data of 60 tumors<sup>34</sup>. They identified four subtypes that were later referred to as basal (BA), mesenchymal (MS), atypical (AT) and classical (CL)<sup>167</sup>. In retrospect, a limitation of this study was that assessment of HPV status was not standard at that period. More recently, meta-analyses have been performed to detect and validate molecular subtypes in larger cohorts<sup>168,169</sup>, which support the existence of these subclasses. The referred studies show a correlation to the original subtypes defined by Chung et al.<sup>34</sup> The breast cancer classification shows a clear difference in clinical response between the different subgroups<sup>170</sup>. These findings have been carefully validated<sup>171-173</sup>. A comparable clinical relevance of molecular subtypes was demonstrated in HNSCC. However, despite the correlation of the different signatures, these studies showed conflicting clinical associations. For example, while 'basal-like' tumors showed the worst prognosis in the paper of Chung et al.<sup>34</sup>, Walter et al.<sup>167</sup> showed that 'atypical' tumors had the poorest outcome. Another difficulty is caused by the inclusion of HPV-positive tumors. Cluster analyses often show enrichment of HPV-positive cases in certain clusters, which is expected based on their different biological background. However, HPV-positive tumors also cluster together with HPV-negative tumors frequently<sup>105</sup>. Unfortunately, this separation of HPV-positive tumors in different clusters is not explained by a separate clinical behavior<sup>169</sup>. Hence, the subclasses identified by unsupervised cluster analyses of expression profiles do not match with the relevant genetic subtypes, that is, HPV-positive and HPV-negative tumors and within the HPV-negative tumors the SCNA-quiet group. Also interesting are the molecular subtypes based on expression profiling of HPV-positive tumors only<sup>174-176</sup>. These subtypes relate to virus characteristics and prognosis.

In summary, molecular subtypes of HNSCC characterized by either HPV status or few genomic alterations are emerging; independent final validation in large cohorts of various populations is still necessary. Moreover, additional molecular subtypes of HNSCC identified by gene expression profiling seem to be present, but clinical applications need further investigation and validation.

## PROGNOSTIC MODELING USING BIOMARKERS

One of the most important factors in patient management is disease prognosis, as it impacts the chosen treatment and determines the clinical outcome for the patient. Prognostic models are aimed to predict the prognosis of patients with a certain disease using multivariable datasets. In HNSCC, outcome measures are often overall survival, disease-specific survival, locoregional control and recurrence-free survival. HNSCC prognostic models combine clinical variables with histopathology and biomarkers to optimize the model performance, i.e. the accuracy of the model to predict the prognosis of a patient correctly. A frequently

used measure of a model's goodness of fit is the C-index. It gives the probability that a randomly selected patient who experienced an event (e.g. death or recurrence) had a higher risk score than a patient who had not experienced the event. For time-to-event data analysis, the C-index should be integrated over a defined time period, e.g. 5 years. The C-index ranges from 0.5 to 1.0, where a C-index of 0.5 resembles a model with a random chance of correctly stratifying a patient that experiences an event (e.g. death or recurrence) or not, and a value of 1.0 resembles a model with perfect stratification of patients that experience an event or not. As a rule of thumb, a C-index of  $>0.7$  is considered a good model<sup>177</sup>. Features to be included in the model are selected by using a specific method such as stepwise regression (for datasets with features  $\ll$  patients) or LASSO regression (for datasets with features  $\gg$  patients) and trained to estimate the coefficients of the regression model. Selection and training is generally carried out using the same dataset. This procedure, however, is sensitive to overfitting: a term used for a model that describes random fluctuations in the data rather than true relationships between the variables. This results in a model with a very high C-index when applied to the training dataset, but with little external validity when applied to a different dataset. Cross-validation can be used to prevent overfitting, but independent validation using additional datasets that were not used for feature selection or model training is considered mandatory to show external validity of a model.

The best known prognostic model for OCSCC, or cancer in general, is the TNM-staging system as mentioned before. TNM is most accurate when using the pathologic staging system with a C-index of 0.65 (TNM 7) and 0.70 (TNM 8) for overall survival<sup>178</sup>. For OPSCC, p16 status is included in TNM 8 as surrogate for HPV status. This makes TNM 8 superior over TNM 7, also in patients from the Netherlands<sup>88</sup>. Other prognostic models have been described, which usually combine clinical factors with histopathology and/or biomarkers<sup>179,180</sup>. These models show similar or slightly better C-indexes, but caution is required when comparing performances because the C-index also depends on cohort characteristics such as number of events. Next to the TNM-staging model, no prognostic models have been widely implemented in daily clinical routine. An ideal model would outperform TNM-staging significantly. Moreover, TNM-staging is particularly useful as prognostic model by using histopathology. These data are only available after surgery, which precludes accurate pretreatment prognostication. Well-timed availability of prognosis could provide better patient counselling and selection of patients for treatment intensification or deintensification trials. Options for pretreatment prognostic biomarkers are imaging<sup>181</sup> and molecular analysis that can be performed on pretreatment biopsies<sup>39,138,182-184</sup>.

## PROGNOSTIC MODELING OF HNSCC

Besides TNM and some histological features, the HPV status is the most important prognostic marker in HNSCC, and therefore it is included in TNM 8 by p16-immunostaining as surrogate marker. Several studies from a variety of countries demonstrated that patients with HPV-positive OPSCC have a strikingly prognostic advantage compared to HPV-negative OPSCC patients<sup>87,88,185</sup>. A number of multivariate prognostic risk models that include HPV status and other risk factors have been developed for OPSCC<sup>87,185-187</sup>.

Interestingly, patients with HPV-positive OPSCC show a favorable prognosis, but a subgroup of these patients does not. To some extent, the above-mentioned risk models enable stratification of HPV-positive patients in high and low risk by the use of clinical variables (e.g. smoking, comorbidity, stage of disease). However, the identification at baseline of HPV-positive tumors with poor prognosis is still suboptimal. Biomarkers are under investigation and promising results include the quantification of tumor-infiltrating lymphocytes<sup>188,189</sup>, expression of human leukocyte antigen class I<sup>190</sup> and expression of cancer stem cell enrichment markers (e.g. CD98)<sup>191</sup>. Moreover, recent gene expression profiling studies stratify patients, but in 2 studies HPV-positive OPSCC could be divided in 2 groups<sup>174,176</sup>, while in a third study it was divided in three groups<sup>175</sup>. Altogether, results should be considered experimental and validation in larger cohorts is required.

Other biomarkers for prognostic modeling are still awaited, especially to stratify HPV-negative HNSCC patients in high- and low-risk groups. One approach is through gene expression profiling. Several groups have used microarray data to predict metastatic behavior, locoregional recurrence and/or overall survival<sup>139,138,182-184</sup>. However, clinical implementation lags behind. Thinkable reasons are the inadequate reproducibility of the classification algorithms, the high costs and the lack of applications for FFPE specimens. Combining multiple data sets is a possible solution to improve the reproducibility<sup>183</sup>, but very well-annotated clinical data (including HPV-status, pN-stage, follow-up data, etc.) are required to make maximum use of the data sets. Independent validation by other techniques would probably substantiate the reproducibility, but is frequently missing. Moreover, in most studies heterogeneous patient populations are studied (e.g. regarding HPV status, treatment and tumor subsite) and often studies are underpowered. This does not affect generation of research hypotheses, but it hampers clinical implementation in the end.

## RADIOMICS

Traditionally, imaging of HNSCC is primarily used to describe locoregional extension of the tumor and to assist staging<sup>192</sup>. However, from these diagnostic images a variety of quantitative features can be extracted as well: a process that is termed “radiomics”<sup>193</sup>. These radiomic features may subsequently be used for prognostic modelling. Radiomic analyses have been applied in HNSCC patients, mostly focusing on CTs from radiotherapy treatment plans. Aerts et al. were first to describe a prognostic radiomic signature in OPSCC<sup>194</sup>. This model was actually a lung cancer based signature that was applied to OPSCC. Later, they developed a OPSCC-specific model<sup>195</sup>, and their approach was followed by others<sup>196-198</sup>. The preference for CT is explained by: (1) each elemental region of the CT image (voxel) is expressed in terms of Hounsfield units (HU) corresponding to the x-ray attenuation (or tissue density)<sup>193</sup>; and (2) the availability of delineated tumor volumes from radiation treatment plans.

However, as mentioned before, MRI is the preferred imaging modality for OCSCC, because of the superior soft tissue contrast and lower sensitivity to dental artifacts. Yet, radiomic analysis of MRI is more complicated because the signal intensities are influenced by MRI vendor and acquisition protocols<sup>199</sup>. Nonetheless, MRI radiomics has successfully been performed in cancer research, for instance in breast cancer, glioblastoma and prostate cancer<sup>200-202</sup>. In HNSCC, prognostic models based on MRI radiomics were described in several small series of oropharyngeal cancer<sup>203,204</sup> and in a larger mixed HNSCC cohort<sup>205</sup>.

## OUTLINE OF THIS THESIS

As described above, treatment choices for HNSCC patients are largely dependent on TNM-staging and clinical variables such as subsite and age at diagnosis. Prognostic models could help to stratify patients more accurately in groups with a favorable and unfavorable prognosis to optimize personalization of treatment and counselling of patients. In the past, several models have been published, but these models often lack precision or adequate validation. Moreover, most models were trained and validated using heterogeneous HNSCC cohorts. In this thesis, novel prognostic models were developed and validated with homogeneous OCSCC or OPSCC cohorts.

The etiologic role of HPV in OPSCC has been well established in the past. In non-oropharyngeal HNSCC however, much less is clear about the etiologic and prognostic role of HPV. Partly, this is caused by the absence of a reliable HPV-test. Usually HPV DNA test are employed, but these are too sensitive causing false positive results. In OPSCC, a test algorithm is most frequently used to determine HPV status that consists of p16<sup>INK4A</sup>(p16) immunostaining followed by PCR-based detection of high-risk HPV DNA on the p16-immunopositive samples. Unfortunately, this algorithm fails in non-oropharyngeal HNSCC because of a lower sensitivity of p16-immunostaining. In **Chapter 2**, a new molecular HPV-detection algorithm is presented that is suitable for

non-oro-pharyngeal HNSCC. Punch biopsies were obtained from tumor enriched regions of the formalin-fixed paraffin-embedded tissue block, and combined DNA and RNA extraction was performed from the punch biopsy. Next, PCR-based detection of HPV-DNA with partial genotyping was performed for 15 HR-HPV types, and positive results were validated by detection of E6 mRNA.

Tumors may be biologically characterized by gene expression, and this information can contribute to optimal prognostication of OSCCC patients. Comprehensive gene expression of the entire genome can be acquired by using microarray analyses or RNA sequencing. However, these techniques are expensive and require considerable experience in bioinformatics for analyses. Ideally, gene expression is assessed with a platform that is more widely available and results can be interpreted by physicians. In **Chapter 3**, a gene expression prognostic model is presented that uses microarray data to select genes that are subsequently assessed with quantitative polymerase chain reaction (qPCR) in a validation cohort. Genes predict overall and disease-free survival, but also the presence of nodal metastasis. Moreover, the gene expression prognostic model is combined with clinical variables and histopathology to optimize accuracy, and also to assess the additional value of the genes.

OSCC patients often undergo MRI imaging during the diagnostic process, but these scans are primarily used to describe the tumor qualitatively and to assist in staging. However, quantitative features can be obtained by radiomic analyses, and these features may possess important prognostic information. In **Chapter 4**, radiomic analyses have been performed using native T1 MRI scans of OSCCC and OPSCC patients. 545 features are extracted from the scans and used for prognostic modelling. These models are subsequently compared to and combined with clinical prognostic models to obtain the most accurate models that can be used in advance of treatment.

Recurrent HNSCC can be salvaged surgically, but only when detected in an early stage. Regrettably, relapses are frequently discovered in an advanced stage though, and only non-curative treatment may be offered for these patients. The gold standard for disease monitoring remains investigation under general anesthesia with biopsies, but this is obviously very invasive. Imaging is also possible, but lacks clinical sensitivity to detect disease in a very early stage. Recently, detection of circulating tumor DNA (ctDNA) has been explored for disease monitoring. This ctDNA is shed to the circulation by apoptotic tumor cells and can be distinguished from germline DNA by its molecular aberrations. In HNSCC, several methods have been described that focus on specific molecular aberrations (e.g. mutations, copy-number aberrations, HPV). Ideally however, a comprehensive method focuses on all aberrations to detect the majority of tumors. In **Chapter 5**, a ctDNA detection method is presented for HNSCC that detects somatic mutations, copy number aberrations and HPV DNA with a single sequencing library strategy. This method is applied to 40 HNSCC patients and 20 non-cancer controls.

Theoretically, gene expression should be able to describe the biological behavior of a tumor very accurately, but, despite extensive efforts by us and others, C-indexes tend not to exceed 0.80. An explanation might come from molecular heterogeneity. In this concept, a tumor actually consists of many subclones that are genetically linked by clonal evolution, but also show important differences. The use of a single biopsy of such a tumor for gene expression analysis may not capture the most aggressive part of the tumor and hence predict a wrongly favorable prognosis. Moreover, relapses, which are a major source of disease-specific death, may develop from small, but aggressive, subclones of which residual cells remain in the patient after treatment. In **Chapter 6**, we assess genetic heterogeneity by analyzing copy number aberrations in multiple biopsies that are taken from a single tumor. The data is subsequently used to develop an algorithm that predicts genetic similarity and dissimilarity between two biopsies, and this algorithm was combined with targeted sequencing to assess genetic relationships of paired primary tumors and local recurrences.

In **Chapter 7** we discuss the findings in broader context, and present suggestions to proceed.

In **Chapter 8** a summary of the thesis is provided.

## REFERENCES

1. Chow LQM. Head and Neck Cancer. *N Engl J Med* 2020; 382(1): 60-72.
2. Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin* 2018; 68(6): 394-424.
3. Braakhuis BJ, Leemans CR, Visser O. Incidence and survival trends of head and neck squamous cell carcinoma in the Netherlands between 1989 and 2011. *Oral Oncol* 2014; 50(7): 670-5.
4. Sabatini ME, Chiocca S. Human papillomavirus as a driver of head and neck cancers. *Br J Cancer* 2020; 122(3): 306-14.
5. Hashibe M, Brennan P, Chuang SC, 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 Epidemiol Biomarkers Prev* 2009; 18(2): 541-50.
6. Neville BW, Day TA. Oral cancer and precancerous lesions. *CA Cancer J Clin* 2002; 52(4): 195-215.
7. Gillison ML, Chaturvedi AK, Anderson WF, Fakhry C. Epidemiology of Human Papillomavirus-Positive Head and Neck Squamous Cell Carcinoma. *J Clin Oncol* 2015; 33(29): 3235-42.
8. Hayes DN, Van Waes C, Seiwert TY. Genetic Landscape of Human Papillomavirus-Associated Head and Neck Cancer and Comparison to Tobacco-Related Tumors. *J Clin Oncol* 2015; 33(29): 3227-34.
9. Bhatia A, Burtneess B. Human Papillomavirus-Associated Oropharyngeal Cancer: Defining Risk Groups and Clinical Trials. *J Clin Oncol* 2015; 33(29): 3243-50.
10. Ndiaye C, Mena M, Alemany L, et al. HPV DNA, E6/E7 mRNA, and p16INK4a detection in head and neck cancers: a systematic review and meta-analysis. *Lancet Oncol* 2014; 15(12): 1319-31.
11. Castellsague X, Alemany L, Quer M, et al. HPV Involvement in Head and Neck Cancers: Comprehensive Assessment of Biomarkers in 3680 Patients. *J Natl Cancer Inst* 2016; 108(6): djv403.
12. Lingen MW, Xiao W, Schmitt A, et al. Low etiologic fraction for high-risk human papillomavirus in oral cavity squamous cell carcinomas. *Oral Oncol* 2013; 49(1): 1-8.
13. Chung CH, Zhang Q, Kong CS, et al. p16 protein expression and human papillomavirus status as prognostic biomarkers of nonoropharyngeal head and neck squamous cell carcinoma. *J Clin Oncol* 2014; 32(35): 3930-8.
14. Rosenberg PS, Greene MH, Alter BP. Cancer incidence in persons with Fanconi anemia. *Blood* 2003; 101(3): 822-6.
15. Brierley JDG, M. K.; Wittekind, C., eds. *TNM Classification of Malignant Tumours*. 8th Edition ed. Hoboken: Wiley-Blackwell; 2016.
16. Amin MB, Edge SB, American Joint Committee on C. *AJCC cancer staging manual*. Eighth edition. ed. Switzerland: Springer; 2017.
17. van den Brekel MW, van der Waal I, Meijer CJ, Freeman JL, Castelijns JA, Snow GB. The incidence of micrometastases in neck dissection specimens obtained from elective neck dissections. *Laryngoscope* 1996; 106(8): 987-91.
18. de Bree R, Takes RP, Castelijns JA, et al. Advances in diagnostic modalities to detect occult lymph node metastases in head and neck squamous cell carcinoma. *Head Neck* 2015; 37(12): 1829-39.
19. van den Brekel MW, Castelijns JA, Stel HV, et al. Occult metastatic neck disease: detection with US and US-guided fine-needle aspiration cytology. *Radiology* 1991; 180(2): 457-61.
20. Righi PD, Kopecky KK, Caldemeyer KS, Ball VA, Weisberger EC, Radpour S. Comparison of ultrasound-fine needle aspiration and computed tomography in patients undergoing elective neck dissection. *Head Neck* 1997; 19(7): 604-10.
21. Takes RP, Righi P, Meeuwis CA, et al. The value of ultrasound with ultrasound-guided fine-needle aspiration biopsy compared to computed tomography in the detection of regional metastases in the clinically negative neck. *Int J Radiat Oncol Biol Phys* 1998; 40(5): 1027-32.
22. de Bondt RB, Nelemans PJ, Hofman PA, et al. Detection of lymph node metastases in head and neck cancer: a meta-analysis comparing US, USgFNAC, CT and MR imaging. *Eur J Radiol* 2007; 64(2): 266-72.
23. Nieuwenhuis EJ, Jaspars LH, Castelijns JA, et al. Quantitative molecular detection of minimal residual head and neck cancer in lymph node aspirates. *Clin Cancer Res* 2003; 9(2): 755-61.
24. Ferris RL, Xi L, Raja S, et al. Molecular staging of cervical lymph nodes in squamous cell carcinoma of the head and neck. *Cancer Res* 2005; 65(6): 2147-56.
25. Solassol J, Burcia V, Costes V, et al. Pempfigus vulgaris antigen mRNA quantification for the staging of sentinel lymph nodes in head and neck cancer. *Br J Cancer* 2010; 102(1): 181-7.

26. Den Toom IJ, Bloemena E, van Weert S, Karagozoglu KH, Hoekstra OS, de Bree R. Additional non-sentinel lymph node metastases in early oral cancer patients with positive sentinel lymph nodes. *Eur Arch Otorhinolaryngol* 2017; 274(2): 961-8.
27. Schilling C, Stoeckli SJ, Vigili MG, et al. Surgical consensus guidelines on sentinel node biopsy (SNB) in patients with oral cancer. *Head Neck* 2019; 41(8): 2655-64.
28. Broglie MA, Haile SR, Stoeckli SJ. Long-term experience in sentinel node biopsy for early oral and oropharyngeal squamous cell carcinoma. *Ann Surg Oncol* 2011; 18(10): 2732-8.
29. Flach GB, Bloemena E, Klop WM, et al. Sentinel lymph node biopsy in clinically N0 T1-T2 staged oral cancer: the Dutch multicenter trial. *Oral Oncol* 2014; 50(10): 1020-4.
30. Alkureishi LW, Ross GL, Shoaib T, et al. Sentinel node biopsy in head and neck squamous cell cancer: 5-year follow-up of a European multicenter trial. *Ann Surg Oncol* 2010; 17(9): 2459-64.
31. Schilling C, Stoeckli SJ, Haerle SK, et al. Sentinel European Node Trial (SENT): 3-year results of sentinel node biopsy in oral cancer. *Eur J Cancer* 2015; 51(18): 2777-84.
32. Ferris RL, Stefanika P, Xi L, Gooding W, Seethala RR, Godfrey TE. Rapid molecular detection of metastatic head and neck squamous cell carcinoma as an intraoperative adjunct to sentinel lymph node biopsy. *Laryngoscope* 2012; 122(5): 1020-30.
33. van 't Veer LJ, Dai H, van de Vijver MJ, et al. Gene expression profiling predicts clinical outcome of breast cancer. *Nature* 2002; 415(6871): 530-6.
34. Chung CH, Parker JS, Karaca G, et al. Molecular classification of head and neck squamous cell carcinomas using patterns of gene expression. *Cancer Cell* 2004; 5(5): 489-500.
35. Onken MD, Winkler AE, Kanchi KL, et al. A surprising cross-species conservation in the genomic landscape of mouse and human oral cancer identifies a transcriptional signature predicting metastatic disease. *Clin Cancer Res* 2014; 20(11): 2873-84.
36. Warner GC, Reis PP, Jurisica I, et al. Molecular classification of oral cancer by cDNA microarrays identifies overexpressed genes correlated with nodal metastasis. *Int J Cancer* 2004; 110(6): 857-68.
37. O'Donnell RK, Kupferman M, Wei SJ, et al. Gene expression signature predicts lymphatic metastasis in squamous cell carcinoma of the oral cavity. *Oncogene* 2005; 24(7): 1244-51.
38. Roepman P, Wessels LF, Kettelarij N, et al. An expression profile for diagnosis of lymph node metastases from primary head and neck squamous cell carcinomas. *Nat Genet* 2005; 37(2): 182-6.
39. Roepman P, Kemmeren P, Wessels LF, Slootweg PJ, Holstege FC. Multiple robust signatures for detecting lymph node metastasis in head and neck cancer. *Cancer Res* 2006; 66(4): 2361-6.
40. van Hooff SR, Leusink FK, Roepman P, et al. Validation of a gene expression signature for assessment of lymph node metastasis in oral squamous cell carcinoma. *J Clin Oncol* 2012; 30(33): 4104-10.
41. Leusink FK, van Es RJ, de Bree R, et al. Novel diagnostic modalities for assessment of the clinically node-negative neck in oral squamous-cell carcinoma. *Lancet Oncol* 2012; 13(12): e554-61.
42. de Bree R, Castelijns JA, Hoekstra OS, Leemans CR. Advances in imaging in the work-up of head and neck cancer patients. *Oral Oncol* 2009; 45(11): 930-5.
43. Klein Nulent TJW, Noorlag R, Van Cann EM, et al. Intraoral ultrasonography to measure tumor thickness of oral cancer: A systematic review and meta-analysis. *Oral Oncol* 2018; 77: 29-36.
44. Arya S, Rane P, Deshmukh A. Oral cavity squamous cell carcinoma: role of pretreatment imaging and its influence on management. *Clin Radiol* 2014; 69(9): 916-30.
45. Seeburg DP, Baer AH, Ayyun N. Imaging of Patients with Head and Neck Cancer: From Staging to Surveillance. *Oral Maxillofac Surg Clin North Am* 2018; 30(4): 421-33.
46. de Bree R, Deurloo EE, Snow GB, Leemans CR. Screening for distant metastases in patients with head and neck cancer. *Laryngoscope* 2000; 110(3 Pt 1): 397-401.
47. Juliano A, Moonis G. Computed Tomography Versus Magnetic Resonance in Head and Neck Cancer: When to Use What and Image Optimization Strategies. *Magn Reson Imaging Clin N Am* 2018; 26(1): 63-84.
48. Dai YL, King AD. State of the art MRI in head and neck cancer. *Clin Radiol* 2018; 73(1): 45-59.
49. de Bree R, Haigentz M, Jr., Silver CE, et al. Distant metastases from head and neck squamous cell carcinoma. Part II. Diagnosis. *Oral Oncol* 2012; 48(9): 780-6.
50. Schouten CS, de Graaf P, Alberts FM, et al. Response evaluation after chemoradiotherapy for advanced nodal disease in head and neck cancer using diffusion-weighted MRI and 18F-FDG-PET-CT. *Oral Oncol* 2015; 51(5): 541-7.

51. Mehanna H, Wong WL, McConkey CC, et al. PET-CT Surveillance versus Neck Dissection in Advanced Head and Neck Cancer. *N Engl J Med* 2016; 374(15): 1444-54.
52. Helsen N, Van den Wyngaert T, Carp L, Stroobants S. FDG-PET/CT for treatment response assessment in head and neck squamous cell carcinoma: a systematic review and meta-analysis of diagnostic performance. *Eur J Nucl Med Mol Imaging* 2018; 45(6): 1063-71.
53. Anderson CM, Chang T, Graham MM, et al. Change of maximum standardized uptake value slope in dynamic triphasic [18F]-fluorodeoxyglucose positron emission tomography/computed tomography distinguishes malignancy from postradiation inflammation in head-and-neck squamous cell carcinoma: a prospective trial. *Int J Radiat Oncol Biol Phys* 2015; 91(3): 472-9.
54. Gupta T, Jain S, Agarwal JP, et al. Diagnostic performance of response assessment FDG-PET/CT in patients with head and neck squamous cell carcinoma treated with high-precision definitive (chemo)radiation. *Radiother Oncol* 2010; 97(2): 194-9.
55. Network NCC. Head and Neck Cancers (version 1.2018). [https://www.nccn.org/professionals/physician\\_gls/pdf/head-and-neck.pdf](https://www.nccn.org/professionals/physician_gls/pdf/head-and-neck.pdf) (accessed February 11th 2020).
56. Dutch Association of Otorhinolaryngology and Head & Neck Surgery NVvK-N-OeHvH-H. Richtlijn Hoofdhals tumoren. 2014. [https://richtlijnen database.nl/richtlijn/hoofd-halstumoren/hoofd-halstumoren\\_-\\_korte\\_beschrijving.html](https://richtlijnen database.nl/richtlijn/hoofd-halstumoren/hoofd-halstumoren_-_korte_beschrijving.html) (accessed February 11th 2020).
57. Stenson KM, Kunnavakkam R, Cohen EE, et al. Chemoradiation for patients with advanced oral cavity cancer. *Laryngoscope* 2010; 120(1): 93-9.
58. Cohen EE, Baru J, Huo D, et al. Efficacy and safety of treating T4 oral cavity tumors with primary chemoradiotherapy. *Head Neck* 2009; 31(8): 1013-21.
59. Gore SM, Crombie AK, Batstone MD, Clark JR. Concurrent chemoradiotherapy compared with surgery and adjuvant radiotherapy for oral cavity squamous cell carcinoma. *Head Neck* 2015; 37(4): 518-23.
60. Joshi A, Rajendraprasad JS, Shetty K. Reconstruction of intraoral defects using facial artery musculomucosal flap. *Br J Plast Surg* 2005; 58(8): 1061-6.
61. Kim MK, Han W, Kim SG. The use of the buccal fat pad flap for oral reconstruction. *Maxillofac Plast Reconstr Surg* 2017; 39(1): 5.
62. Hanasono MM, Friel MT, Klem C, et al. Impact of reconstructive microsurgery in patients with advanced oral cavity cancers. *Head Neck* 2009; 31(10): 1289-96.
63. Dillon MT, Harrington KJ. Human Papillomavirus-Negative Pharyngeal Cancer. *J Clin Oncol* 2015; 33(29): 3251-61.
64. Forastiere AA, Weber RS, Trotti A. Organ Preservation for Advanced Larynx Cancer: Issues and Outcomes. *J Clin Oncol* 2015; 33(29): 3262-8.
65. Mehanna H, Robinson M, Hartley A, et al. Radiotherapy plus cisplatin or cetuximab in low-risk human papillomavirus-positive oropharyngeal cancer (De-ESCALaTE HPV): an open-label randomised controlled phase 3 trial. *Lancet* 2019; 393(10166): 51-60.
66. Gillison ML, Trotti AM, Harris J, et al. Radiotherapy plus cetuximab or cisplatin in human papillomavirus-positive oropharyngeal cancer (NRG Oncology RTOG 1016): a randomised, multicentre, non-inferiority trial. *Lancet* 2019; 393(10166): 40-50.
67. Baird BJ, Sung CK, Beadle BM, Divi V. Treatment of early-stage laryngeal cancer: A comparison of treatment options. *Oral Oncol* 2018; 87: 8-16.
68. Suarez C, Rodrigo JP, Silver CE, et al. Laser surgery for early to moderately advanced glottic, supraglottic, and hypopharyngeal cancers. *Head Neck* 2012; 34(7): 1028-35.
69. Haughey BH, Hinni ML, Salassa JR, et al. Transoral laser microsurgery as primary treatment for advanced-stage oropharyngeal cancer: a United States multicenter study. *Head Neck* 2011; 33(12): 1683-94.
70. Moore EJ, Hinni ML. Critical review: transoral laser microsurgery and robotic-assisted surgery for oropharynx cancer including human papillomavirus-related cancer. *Int J Radiat Oncol Biol Phys* 2013; 85(5): 1163-7.
71. Holsinger FC, Ferris RL. Transoral Endoscopic Head and Neck Surgery and Its Role Within the Multidisciplinary Treatment Paradigm of Oropharynx Cancer: Robotics, Lasers, and Clinical Trials. *J Clin Oncol* 2015; 33(29): 3285-92.
72. Nichols AC, Theurer J, Prisman E, et al. Radiotherapy versus transoral robotic surgery and neck dissection for oropharyngeal squamous cell carcinoma (ORATOR): an open-label, phase 2, randomised trial. *Lancet Oncol* 2019; 20(10): 1349-59.
73. Lundahl RE, Foote RL, Bonner JA, et al. Combined neck dissection and postoperative radiation therapy in the management of the high-risk neck: a matched-pair analysis. *Int J Radiat Oncol Biol Phys* 1998; 40(3): 529-34.

74. Gupta T, Agarwal J, Jain S, et al. Three-dimensional conformal radiotherapy (3D-CRT) versus intensity modulated radiation therapy (IMRT) in squamous cell carcinoma of the head and neck: a randomized controlled trial. *Radiother Oncol* 2012; 104(3): 343-8.
75. van Dijk LV, Steenbakkers RJ, ten Haken B, et al. Robust Intensity Modulated Proton Therapy (IMPT) Increases Estimated Clinical Benefit in Head and Neck Cancer Patients. *PLoS One* 2016; 11(3): e0152477.
76. Kierkels RGJ, Fredriksson A, Both S, Langendijk JA, Scandurra D, Korevaar EW. Automated Robust Proton Planning Using Dose-Volume Histogram-Based Mimicking of the Photon Reference Dose and Reducing Organ at Risk Dose Optimization. *Int J Radiat Oncol Biol Phys* 2019; 103(1): 251-8.
77. van der Laan HP, van de Water TA, van Herpt HE, et al. The potential of intensity-modulated proton radiotherapy to reduce swallowing dysfunction in the treatment of head and neck cancer: A planning comparative study. *Acta Oncol* 2013; 52(3): 561-9.
78. Cooper JS, Pajak TF, Forastiere AA, et al. Postoperative concurrent radiotherapy and chemotherapy for high-risk squamous-cell carcinoma of the head and neck. *N Engl J Med* 2004; 350(19): 1937-44.
79. Bernier J, Dornic C, Ozsahin M, et al. Postoperative irradiation with or without concomitant chemotherapy for locally advanced head and neck cancer. *N Engl J Med* 2004; 350(19): 1945-52.
80. Sacco AG, Cohen EE. Current Treatment Options for Recurrent or Metastatic Head and Neck Squamous Cell Carcinoma. *J Clin Oncol* 2015; 33(29): 3305-13.
81. Vermorken JB, Mesia R, Rivera F, et al. Platinum-based chemotherapy plus cetuximab in head and neck cancer. *N Engl J Med* 2008; 359(11): 1116-27.
82. Seiwert TY, Burtneis B, Mehra R, et al. Safety and clinical activity of pembrolizumab for treatment of recurrent or metastatic squamous cell carcinoma of the head and neck (KEYNOTE-012): an open-label, multicentre, phase 1b trial. *Lancet Oncol* 2016; 17(7): 956-65.
83. Ferris RL, Blumenschein G, Jr., Fayette J, et al. Nivolumab for Recurrent Squamous-Cell Carcinoma of the Head and Neck. *N Engl J Med* 2016; 375(19): 1856-67.
84. Cohen EEW, Soulieres D, Le Tourneau C, et al. Pembrolizumab versus methotrexate, docetaxel, or cetuximab for recurrent or metastatic head-and-neck squamous cell carcinoma (KEYNOTE-040): a randomised, open-label, phase 3 study. *Lancet* 2019; 393(10167): 156-67.
85. Bauml JM, Aggarwal C, Cohen RB. Immunotherapy for head and neck cancer: where are we now and where are we going? *Ann Transl Med* 2019; 7(Suppl 3): S75.
86. Surveillance E, and End Results (SEER) Program ([www.seer.cancer.gov](http://www.seer.cancer.gov)) Research Data (1975-2016), National Cancer Institute, DCCPS, Surveillance Research Program, released April 2019, based on the November 2018 submission.
87. Ang KK, Harris J, Wheeler R, et al. Human papillomavirus and survival of patients with oropharyngeal cancer. *N Engl J Med* 2010; 363(1): 24-35.
88. Nauta IH, Rietbergen MM, van Bokhoven A, et al. Evaluation of the eighth TNM classification on p16-positive oropharyngeal squamous cell carcinomas in the Netherlands and the importance of additional HPV DNA testing. *Ann Oncol* 2018; 29(5): 1273-9.
89. Myers JN, Greenberg JS, Mo V, Roberts D. Extracapsular spread. A significant predictor of treatment failure in patients with squamous cell carcinoma of the tongue. *Cancer* 2001; 92(12): 3030-6.
90. Leemans CR, Tiwari R, Nauta JJ, van der Waal I, Snow GB. Recurrence at the primary site in head and neck cancer and the significance of neck lymph node metastases as a prognostic factor. *Cancer* 1994; 73(1): 187-90.
91. Leemans CR, Tiwari R, Nauta JJ, van der Waal I, Snow GB. Regional lymph node involvement and its significance in the development of distant metastases in head and neck carcinoma. *Cancer* 1993; 71(2): 452-6.
92. Baxi SS, Pinheiro LC, Patil SM, Pfister DG, Oeffinger KC, Elkin EB. Causes of death in long-term survivors of head and neck cancer. *Cancer* 2014; 120(10): 1507-13.
93. Kallioniemi A, Kallioniemi OP, Sudar D, et al. Comparative genomic hybridization for molecular cytogenetic analysis of solid tumors. *Science* 1992; 258(5083): 818-21.
94. Solinas-Toldo S, Lampel S, Stilgenbauer S, et al. Matrix-based comparative genomic hybridization: biochips to screen for genomic imbalances. *Genes Chromosomes Cancer* 1997; 20(4): 399-407.
95. Pinkel D, Seagraves R, Sudar D, et al. High resolution analysis of DNA copy number variation using comparative genomic hybridization to microarrays. *Nat Genet* 1998; 20(2): 207-11.
96. Pollack JR, Perou CM, Alizadeh AA, et al. Genome-wide analysis of DNA copy-number changes using cDNA microarrays. *Nat Genet* 1999; 23(1): 41-6.

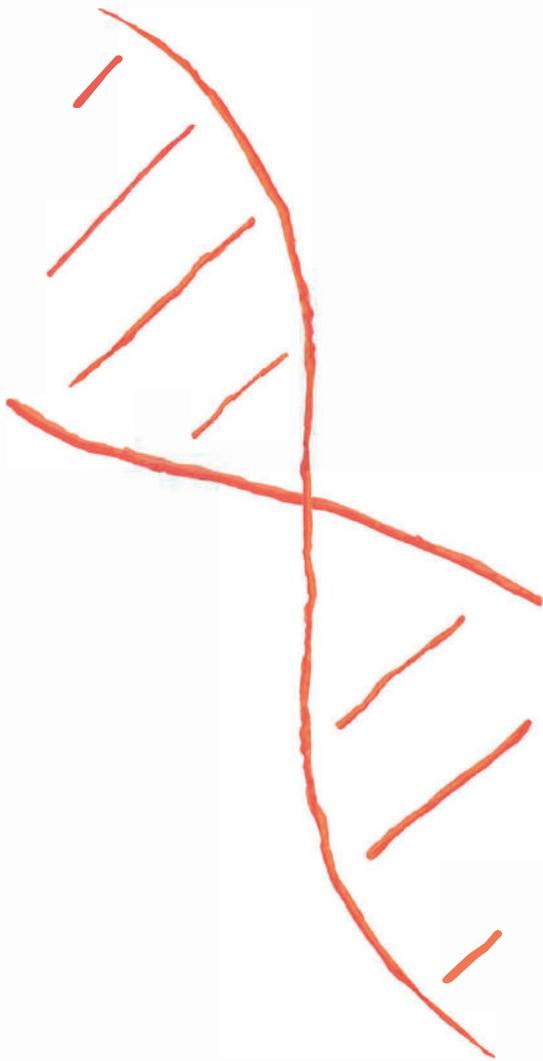
97. Schena M, Shalon D, Davis RW, Brown PO. Quantitative monitoring of gene expression patterns with a complementary DNA microarray. *Science* 1995; 270(5235): 467-70.
98. Lashkari DA, DeRisi JL, McCusker JH, et al. Yeast microarrays for genome wide parallel genetic and gene expression analysis. *Proc Natl Acad Sci U S A* 1997; 94(24): 13057-62.
99. Cardoso F, Van't Veer L, Rutgers E, Loi S, Mook S, Piccart-Gebhart MJ. Clinical application of the 70-gene profile: the MINDACT trial. *J Clin Oncol* 2008; 26(5): 729-35.
100. Singh RR, Patel KP, Routbort MJ, et al. Clinical validation of a next-generation sequencing screen for mutational hotspots in 46 cancer-related genes. *J Mol Diagn* 2013; 15(5): 607-22.
101. Nielsen TO, Hsu FD, Jensen K, et al. Immunohistochemical and clinical characterization of the basal-like subtype of invasive breast carcinoma. *Clin Cancer Res* 2004; 10(16): 5367-74.
102. Perou CM, Sorlie T, Eisen MB, et al. Molecular portraits of human breast tumours. *Nature* 2000; 406(6797): 747-52.
103. Ioannidis JP, Allison DB, Ball CA, et al. Repeatability of published microarray gene expression analyses. *Nat Genet* 2009; 41(2): 149-55.
104. Stransky N, Egloff AM, Tward AD, et al. The mutational landscape of head and neck squamous cell carcinoma. *Science* 2011; 333(6046): 1157-60.
105. Cancer Genome Atlas N. Comprehensive genomic characterization of head and neck squamous cell carcinomas. *Nature* 2015; 517(7536): 576-82.
106. Seiwert TY, Zuo Z, Keck MK, et al. Integrative and comparative genomic analysis of HPV-positive and HPV-negative head and neck squamous cell carcinomas. *Clin Cancer Res* 2015; 21(3): 632-41.
107. Califano J, van der Riet P, Westra W, et al. Genetic progression model for head and neck cancer: implications for field cancerization. *Cancer Res* 1996; 56(11): 2488-92.
108. Agrawal N, Frederick MJ, Pickering CR, et al. Exome sequencing of head and neck squamous cell carcinoma reveals inactivating mutations in NOTCH1. *Science* 2011; 333(6046): 1154-7.
109. Braakhuis BJ, Tabor MP, Kummer JA, Leemans CR, Brakenhoff RH. A genetic explanation of Slaughter's concept of field cancerization: evidence and clinical implications. *Cancer Res* 2003; 63(8): 1727-30.
110. Leemans CR, Braakhuis BJ, Brakenhoff RH. The molecular biology of head and neck cancer. *Nat Rev Cancer* 2011; 11(1): 9-22.
111. zur Hausen H. Papillomaviruses and cancer: from basic studies to clinical application. *Nat Rev Cancer* 2002; 2(5): 342-50.
112. Hubbert NL, Sedman SA, Schiller JT. Human papillomavirus type 16 E6 increases the degradation rate of p53 in human keratinocytes. *J Virol* 1992; 66(10): 6237-41.
113. 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* 1989; 243(4893): 934-7.
114. Steenbergen RD, Snijders PJ, Heideman DA, Meijer CJ. Clinical implications of (epi)genetic changes in HPV-induced cervical precancerous lesions. *Nat Rev Cancer* 2014; 14(6): 395-405.
115. Moody CA, Laimins LA. Human papillomavirus oncoproteins: pathways to transformation. *Nat Rev Cancer* 2010; 10(8): 550-60.
116. Bartkova J, Horejsi Z, Koed K, et al. DNA damage response as a candidate anti-cancer barrier in early human tumorigenesis. *Nature* 2005; 434(7035): 864-70.
117. Leemans CR, Snijders PJF, Brakenhoff RH. The molecular landscape of head and neck cancer. *Nat Rev Cancer* 2018; 18(5): 269-82.
118. Slaughter DP, Southwick HW, Smejkal W. Field cancerization in oral stratified squamous epithelium; clinical implications of multicentric origin. *Cancer* 1953; 6(5): 963-8.
119. Tabor MP, Brakenhoff RH, van Houten VM, et al. Persistence of genetically altered fields in head and neck cancer patients: biological and clinical implications. *Clin Cancer Res* 2001; 7(6): 1523-32.
120. Braakhuis BJ, Tabor MP, Leemans CR, van der Waal I, Snow GB, Brakenhoff RH. Second primary tumors and field cancerization in oral and oropharyngeal cancer: molecular techniques provide new insights and definitions. *Head Neck* 2002; 24(2): 198-206.
121. Tabor MP, Brakenhoff RH, Ruijter-Schippers HJ, et al. Multiple head and neck tumors frequently originate from a single preneoplastic lesion. *Am J Pathol* 2002; 161(3): 1051-60.
122. van Houten VM, Tabor MP, van den Brekel MW, et al. Mutated p53 as a molecular marker for the diagnosis of head and neck cancer. *J Pathol* 2002; 198(4): 476-86.

123. Lim X, Tan SH, Koh WL, et al. Interfollicular epidermal stem cells self-renew via autocrine Wnt signaling. *Science* 2013; 342(6163): 1226-30.
124. Rietbergen MM, Braakhuis BJ, Moukhtari N, et al. No evidence for active human papillomavirus (HPV) in fields surrounding HPV-positive oropharyngeal tumors. *J Oral Pathol Med* 2014; 43(2): 137-42.
125. McGovern SL, Williams MD, Weber RS, et al. Three synchronous HPV-associated squamous cell carcinomas of Waldeyer's ring: case report and comparison with Slaughter's model of field cancerization. *Head Neck* 2010; 32(8): 1118-24.
126. Vermorken JB, Specenier P. Optimal treatment for recurrent/metastatic head and neck cancer. *Ann Oncol* 2010; 21 Suppl 7: vii252-61.
127. Ho AS, Kraus DH, Ganly I, Lee NY, Shah JP, Morris LG. Decision making in the management of recurrent head and neck cancer. *Head Neck* 2014; 36(1): 144-51.
128. Takes RP, Rinaldo A, Silver CE, et al. Distant metastases from head and neck squamous cell carcinoma. Part I. Basic aspects. *Oral Oncol* 2012; 48(9): 775-9.
129. Brennan JA, Mao L, Hruban RH, et al. Molecular assessment of histopathological staging in squamous-cell carcinoma of the head and neck. *N Engl J Med* 1995; 332(7): 429-35.
130. van Houten VM, Leemans CR, Kummer JA, et al. Molecular diagnosis of surgical margins and local recurrence in head and neck cancer patients: a prospective study. *Clin Cancer Res* 2004; 10(11): 3614-20.
131. Poeta ML, Manola J, Goldenberg D, et al. The Ligamp TP53 Assay for Detection of Minimal Residual Disease in Head and Neck Squamous Cell Carcinoma Surgical Margins. *Clin Cancer Res* 2009; 15(24): 7658-65.
132. Shaw RJ, Hall GL, Woolgar JA, et al. Quantitative methylation analysis of resection margins and lymph nodes in oral squamous cell carcinoma. *Br J Oral Maxillofac Surg* 2007; 45(8): 617-22.
133. Temam S, Casiraghi O, Lahaye JB, et al. Tetranucleotide microsatellite instability in surgical margins for prediction of local recurrence of head and neck squamous cell carcinoma. *Clin Cancer Res* 2004; 10(12 Pt 1): 4022-8.
134. Colnot DR, Nieuwenhuis EJ, Kuik DJ, et al. Clinical significance of micrometastatic cells detected by E48 (Ly-6D) reverse transcription-polymerase chain reaction in bone marrow of head and neck cancer patients. *Clin Cancer Res* 2004; 10(23): 7827-33.
135. Graveland AP, Braakhuis BJ, Eerenstein SE, et al. Molecular diagnosis of minimal residual disease in head and neck cancer patients. *Cell Oncol (Dordr)* 2012; 35(5): 367-75.
136. Roh JL, Westra WH, Califano JA, Sidransky D, Koch WM. Tissue imprint for molecular mapping of deep surgical margins in patients with head and neck squamous cell carcinoma. *Head Neck* 2012; 34(11): 1529-36.
137. Salvi D, Picone M, Arredondo MT, et al. Merging person-specific bio-markers for predicting oral cancer recurrence through an ontology. *IEEE Trans Biomed Eng* 2013; 60(1): 216-20.
138. Chung CH, Parker JS, Ely K, et al. Gene expression profiles identify epithelial-to-mesenchymal transition and activation of nuclear factor-kappaB signaling as characteristics of a high-risk head and neck squamous cell carcinoma. *Cancer Res* 2006; 66(16): 8210-8.
139. Nawroz H, Koch W, Anker P, Stroun M, Sidransky D. Microsatellite alterations in serum DNA of head and neck cancer patients. *Nat Med* 1996; 2(9): 1035-7.
140. Bettgowda C, Sausen M, Leary RJ, et al. Detection of circulating tumor DNA in early- and late-stage human malignancies. *Sci Transl Med* 2014; 6(224): 224ra24.
141. Wang Y, Springer S, Mulvey CL, et al. Detection of somatic mutations and HPV in the saliva and plasma of patients with head and neck squamous cell carcinomas. *Sci Transl Med* 2015; 7(293): 293ra104.
142. Wils LJ, Poell JB, Evren I, et al. Incorporation of differentiated dysplasia improves prediction of oral leukoplakia at increased risk of malignant progression. *Mod Pathol* 2020.
143. Mao L, Lee JS, Fan YH, et al. Frequent microsatellite alterations at chromosomes 9p21 and 3p14 in oral premalignant lesions and their value in cancer risk assessment. *Nat Med* 1996; 2(6): 682-5.
144. Partridge M, Pateromichelakis S, Phillips E, Emilion GG, A'Hern RP, Langdon JD. A case-control study confirms that microsatellite assay can identify patients at risk of developing oral squamous cell carcinoma within a field of cancerization. *Cancer Res* 2000; 60(14): 3893-8.
145. Rosin MP, Cheng X, Poh C, et al. Use of allelic loss to predict malignant risk for low-grade oral epithelial dysplasia. *Clin Cancer Res* 2000; 6(2): 357-62.
146. Graveland AP, Golusinski PJ, Buijze M, et al. Loss of heterozygosity at 9p and p53 immunopositivity in surgical margins predict local relapse in head and neck squamous cell carcinoma. *Int J Cancer* 2011; 128(8): 1852-9.

147. Zhang L, Poh CF, Williams M, et al. Loss of heterozygosity (LOH) profiles--validated risk predictors for progression to oral cancer. *Cancer Prev Res (Phila)* 2012; 5(9): 1081-9.
148. Graveland AP, Bremmer JF, de Maaker M, et al. Molecular screening of oral precancer. *Oral Oncol* 2013; 49(12): 1129-35.
149. van der Waal I. Potentially malignant disorders of the oral and oropharyngeal mucosa; present concepts of management. *Oral Oncol* 2010; 46(6): 423-5.
150. Bremmer JF, Graveland AP, Brink A, et al. Screening for oral precancer with noninvasive genetic cytology. *Cancer Prev Res (Phila)* 2009; 2(2): 128-33.
151. Poh CF, MacAulay CE, Zhang L, Rosin MP. Tracing the "at-risk" oral mucosa field with autofluorescence: steps toward clinical impact. *Cancer Prev Res (Phila)* 2009; 2(5): 401-4.
152. Poh CF, Zhang L, Anderson DW, et al. Fluorescence visualization detection of field alterations in tumor margins of oral cancer patients. *Clin Cancer Res* 2006; 12(22): 6716-22.
153. Poh CF, Durham JS, Brasher PM, et al. Canadian Optically-guided approach for Oral Lesions Surgical (COOLS) trial: study protocol for a randomized controlled trial. *BMC Cancer* 2011; 11: 462.
154. Braakhuis BJ, Snijders PJ, Keune WJ, et al. Genetic patterns in head and neck cancers that contain or lack transcriptionally active human papillomavirus. *J Natl Cancer Inst* 2004; 96(13): 998-1006.
155. van Houten VM, Snijders PJ, van den Brekel MW, et al. Biological evidence that human papillomaviruses are etiologically involved in a subgroup of head and neck squamous cell carcinomas. *Int J Cancer* 2001; 93(2): 232-5.
156. 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(10): 1510-7.
157. Smeets SJ, Braakhuis BJ, Abbas S, et al. Genome-wide DNA copy number alterations in head and neck squamous cell carcinomas with or without oncogene-expressing human papillomavirus. *Oncogene* 2006; 25(17): 2558-64.
158. Slebos RJ, Yi Y, Ely K, et al. Gene expression differences associated with human papillomavirus status in head and neck squamous cell carcinoma. *Clin Cancer Res* 2006; 12(3 Pt 1): 701-9.
159. Lleras RA, Smith RV, Adrien LR, et al. Unique DNA methylation loci distinguish anatomic site and HPV status in head and neck squamous cell carcinoma. *Clin Cancer Res* 2013; 19(19): 5444-55.
160. Kostareli E, Holzinger D, Bogatyrova O, et al. HPV-related methylation signature predicts survival in oropharyngeal squamous cell carcinomas. *J Clin Invest* 2013; 123(6): 2488-501.
161. Snijders PJ, van den Brule AJ, Meijer CJ. The clinical relevance of human papillomavirus testing: relationship between analytical and clinical sensitivity. *J Pathol* 2003; 201(1): 1-6.
162. Schache AG, Liloglou T, Risk JM, et al. Evaluation of human papilloma virus diagnostic testing in oropharyngeal squamous cell carcinoma: sensitivity, specificity, and prognostic discrimination. *Clin Cancer Res* 2011; 17(19): 6262-71.
163. Smeets SJ, Hesselink AT, Speel EJ, et al. A novel algorithm for reliable detection of human papillomavirus in paraffin embedded head and neck cancer specimen. *Int J Cancer* 2007; 121(11): 2465-72.
164. Weinberger PM, Yu Z, Haffty BG, et al. Molecular classification identifies a subset of human papillomavirus--associated oropharyngeal cancers with favorable prognosis. *J Clin Oncol* 2006; 24(5): 736-47.
165. Rietbergen MM, Leemans CR, Bloemena E, et al. Increasing prevalence rates of HPV attributable oropharyngeal squamous cell carcinomas in the Netherlands as assessed by a validated test algorithm. *Int J Cancer* 2013; 132(7): 1565-71.
166. Smeets SJ, Brakenhoff RH, Ylstra B, et al. Genetic classification of oral and oropharyngeal carcinomas identifies subgroups with a different prognosis. *Cell Oncol* 2009; 31(4): 291-300.
167. Walter V, Yin X, Wilkerson MD, et al. Molecular subtypes in head and neck cancer exhibit distinct patterns of chromosomal gain and loss of canonical cancer genes. *PLoS One* 2013; 8(2): e56823.
168. De Cecco L, Nicolau M, Giannoccaro M, et al. Head and neck cancer subtypes with biological and clinical relevance: Meta-analysis of gene-expression data. *Oncotarget* 2015; 6(11): 9627-42.
169. Keck MK, Zuo Z, Khattri A, et al. Integrative analysis of head and neck cancer identifies two biologically distinct HPV and three non-HPV subtypes. *Clin Cancer Res* 2015; 21(4): 870-81.
170. Sorlie T, Perou CM, Tibshirani R, et al. Gene expression patterns of breast carcinomas distinguish tumor subclasses with clinical implications. *Proc Natl Acad Sci U S A* 2001; 98(19): 10869-74.
171. Sorlie T, Tibshirani R, Parker J, et al. Repeated observation of breast tumor subtypes in independent gene expression data sets. *Proc Natl Acad Sci U S A* 2003; 100(14): 8418-23.

172. Fan C, Oh DS, Wessels L, et al. Concordance among gene-expression-based predictors for breast cancer. *N Engl J Med* 2006; 355(6): 560-9.
173. Carey LA, Perou CM, Livasy CA, et al. Race, breast cancer subtypes, and survival in the Carolina Breast Cancer Study. *JAMA* 2006; 295(21): 2492-502.
174. Zhang Y, Koneva LA, Virani S, et al. Subtypes of HPV-Positive Head and Neck Cancers Are Associated with HPV Characteristics, Copy Number Alterations, PIK3CA Mutation, and Pathway Signatures. *Clin Cancer Res* 2016; 22(18): 4735-45.
175. Locati LD, Serafini MS, Ianno MF, et al. Mining of Self-Organizing Map Gene-Expression Portraits Reveals Prognostic Stratification of HPV-Positive Head and Neck Squamous Cell Carcinoma. *Cancers (Basel)* 2019; 11(8).
176. Gleber-Netto FO, Rao X, Guo T, et al. Variations in HPV function are associated with survival in squamous cell carcinoma. *JCI Insight* 2019; 4(1).
177. Hosmer DWL, S. *Applied Logistic Regression*. New York, NY: John Wiley & Sons; 2000.
178. Moeckelmann N, Ebrahimi A, Tou YK, et al. Prognostic implications of the 8th edition American Joint Committee on Cancer (AJCC) staging system in oral cavity squamous cell carcinoma. *Oral Oncol* 2018; 85: 82-6.
179. Tham T, Machado R, Herman SW, Kraus D, Costantino P, Roche A. Personalized prognostication in head and neck cancer: A systematic review of nomograms according to the AJCC precision medicine core (PMC) criteria. *Head Neck* 2019; 41(8): 2811-22.
180. Rivera C, Oliveira AK, Costa RAP, De Rossi T, Paes Leme AF. Prognostic biomarkers in oral squamous cell carcinoma: A systematic review. *Oral Oncol* 2017; 72: 38-47.
181. Marcu LG, Reid P, Bezak E. The Promise of Novel Biomarkers for Head and Neck Cancer from an Imaging Perspective. *Int J Mol Sci* 2018; 19(9).
182. Rickman DS, Millon R, De Reynies A, et al. Prediction of future metastasis and molecular characterization of head and neck squamous-cell carcinoma based on transcriptome and genome analysis by microarrays. *Oncogene* 2008; 27(51): 6607-22.
183. De Cecco L, Bossi P, Locati L, Canevari S, Licitra L. Comprehensive gene expression meta-analysis of head and neck squamous cell carcinoma microarray data defines a robust survival predictor. *Ann Oncol* 2014; 25(8): 1628-35.
184. Lohavanichbutr P, Mendez E, Holsinger FC, et al. A 13-gene signature prognostic of HPV-negative OSCC: discovery and external validation. *Clin Cancer Res* 2013; 19(5): 1197-203.
185. Rietbergen MM, Brakenhoff RH, Bloemena E, et al. Human papillomavirus detection and comorbidity: critical issues in selection of patients with oropharyngeal cancer for treatment De-escalation trials. *Ann Oncol* 2013; 24(11): 2740-5.
186. O'Sullivan B, Huang SH, Siu LL, et al. Deintensification candidate subgroups in human papillomavirus-related oropharyngeal cancer according to minimal risk of distant metastasis. *J Clin Oncol* 2013; 31(5): 543-50.
187. Rios Velazquez E, Hoebbers F, Aerts HJ, et al. Externally validated HPV-based prognostic nomogram for oropharyngeal carcinoma patients yields more accurate predictions than TNM staging. *Radiother Oncol* 2014; 113(3): 324-30.
188. Oguejiofor K, Hall J, Slater C, et al. Stromal infiltration of CD8 T cells is associated with improved clinical outcome in HPV-positive oropharyngeal squamous carcinoma. *Br J Cancer* 2015; 113(6): 886-93.
189. Ward MJ, Thirdborough SM, Mellows T, et al. Tumour-infiltrating lymphocytes predict for outcome in HPV-positive oropharyngeal cancer. *Br J Cancer* 2014; 110(2): 489-500.
190. Nasman A, Andersson E, Marklund L, et al. HLA class I and II expression in oropharyngeal squamous cell carcinoma in relation to tumor HPV status and clinical outcome. *PLoS One* 2013; 8(10): e77025.
191. Rietbergen MM, Martens-de Kemp SR, Bloemena E, et al. Cancer stem cell enrichment marker CD98: a prognostic factor for survival in patients with human papillomavirus-positive oropharyngeal cancer. *Eur J Cancer* 2014; 50(4): 765-73.
192. Drage N, Qureshi S, Lingam R. Imaging patients with cancer of the oral cavity. *Br Dent J* 2018; 225(9): 827-32.
193. Kumar V, Gu Y, Basu S, et al. Radiomics: the process and the challenges. *Magn Reson Imaging* 2012; 30(9): 1234-48.
194. Aerts HJ, Velazquez ER, Leijenaar RT, et al. Decoding tumour phenotype by noninvasive imaging using a quantitative radiomics approach. *Nat Commun* 2014; 5: 4006.
195. Parmar C, Leijenaar RT, Grossmann P, et al. Radiomic feature clusters and prognostic signatures specific for Lung and Head & Neck cancer. *Sci Rep* 2015; 5: 11044.
196. Leger S, Zwanenburg A, Pilz K, et al. A comparative study of machine learning methods for time-to-event survival data for radiomics risk modelling. *Sci Rep* 2017; 7(1): 13206.
197. Vallieres M, Kay-Rivest E, Perrin LJ, et al. Radiomics strategies for risk assessment of tumour failure in head-and-neck cancer. *Sci Rep* 2017; 7(1): 10117.

198. Bogowicz M, Riesterer O, Stark LS, et al. Comparison of PET and CT radiomics for prediction of local tumor control in head and neck squamous cell carcinoma. *Acta Oncol* 2017; 56(11): 1531-6.
199. Jethanandani A, Lin TA, Volpe S, et al. Exploring Applications of Radiomics in Magnetic Resonance Imaging of Head and Neck Cancer: A Systematic Review. *Front Oncol* 2018; 8: 131.
200. Li H, Zhu Y, Burnside ES, et al. Quantitative MRI radiomics in the prediction of molecular classifications of breast cancer subtypes in the TCGA/TCIA data set. *NPJ Breast Cancer* 2016; 2.
201. Gevaert O, Mitchell LA, Achrol AS, et al. Glioblastoma multiforme: exploratory radiogenomic analysis by using quantitative image features. *Radiology* 2014; 273(1): 168-74.
202. Gnep K, Fargeas A, Gutierrez-Carvajal RE, et al. Haralick textural features on T2 -weighted MRI are associated with biochemical recurrence following radiotherapy for peripheral zone prostate cancer. *J Magn Reson Imaging* 2017; 45(1): 103-17.
203. Dang M, Lysack JT, Wu T, et al. MRI texture analysis predicts p53 status in head and neck squamous cell carcinoma. *AJNR Am J Neuroradiol* 2015; 36(1): 166-70.
204. Jansen JF, Lu Y, Gupta G, et al. Texture analysis on parametric maps derived from dynamic contrast-enhanced magnetic resonance imaging in head and neck cancer. *World J Radiol* 2016; 8(1): 90-7.
205. Yuan Y, Ren J, Shi Y, Tao X. MRI-based radiomic signature as predictive marker for patients with head and neck squamous cell carcinoma. *Eur J Radiol* 2019; 117: 193-8.



# 2

## **Development and Validation of a Novel and Rapid Molecular Detection Method for High-Risk Human Papillomavirus in Formalin-Fixed, Paraffin-Embedded Tumor Tissue**

Steven W. Mes, Daniëlle A.M. Heideman, Elisabeth Bloemena, Arjen Brink, Martijn Bogaarts, C. René Leemans, Ruud H. Brakenhoff

**The Journal of Molecular Diagnostics 2020; 22(2): 262-71**

## **ABSTRACT**

The most widely applied algorithm for human papillomavirus (HPV) detection in formalin-fixed, paraffin-embedded (FFPE) specimens of oropharyngeal head and neck squamous cell carcinoma (HNSCC) consists of p16INK4A immunostaining followed by PCR-based detection of high-risk HPV DNA on p16INK4A-immunopositive samples. However, in nonoropharyngeal HNSCC this algorithm fails, hampering correct interpretation of the prevalence and prognosis of HPV in these cases. In this study, we developed and validated a molecular HPV detection method for FFPE specimens of oropharyngeal and nonoropharyngeal HNSCC. Sectioning of FFPE blocks was circumvented by using punch biopsies from tumor-enriched regions of FFPE tissue blocks, and combined extraction was applied to obtain high-quality DNA and RNA from the punch biopsy. Next, PCR-based detection of HPV DNA was performed for 15 high-risk HPV types with subsequent detection of E6 mRNA for validation. The combined DNA/RNA FFPE test of tissue cores was assessed in well-characterized cohorts with known HPV status based on earlier work, that is, a cohort of oropharyngeal HNSCC (n = 80) and oral cavity HNSCC (n = 25), and reached an accuracy of 97% and 100%, respectively. In conclusion, our method is rapid, simple, and shows an excellent diagnostic performance for detection of HPV type 16. Ultimately, it can be applied for large cohort studies to determine the etiologic fraction and prognostic implication of HPV in nonoropharyngeal HNSCC.

## INTRODUCTION

Head and neck squamous cell carcinoma (HNSCC) develops in the epithelial lining of the oral cavity, pharynx, and larynx, and accounts for approximately 4% of the global cancer incidence and mortality<sup>1</sup>. Classic etiologic factors are tobacco use and consumption of alcohol-containing beverages<sup>2</sup>. In 2007, the World Health Organization declared that infection with human papillomavirus (HPV), mostly type 16 (HPV16), is an additional causative factor for HNSCC<sup>3</sup>. The virus causes tumors most frequently in the oropharynx [oropharyngeal squamous cell carcinoma (OPSCC)], and particularly in the tonsil. This subgroup of HPV-positive OPSCCs is a separate disease entity when compared with the classic HNSCCs because of its clinical behavior<sup>4</sup>, specific microscopic appearance<sup>5</sup>, and different molecular alterations<sup>6</sup>. The HPV-attributable fraction of OPSCC differs geographically between 9% in Southern Europe<sup>7</sup> and 60% or higher in North America<sup>8</sup>. It convincingly has been shown that patients with HPV-positive OPSCC have a much better prognosis than HPV-negative OPSCC patients<sup>9-11</sup>. This has led to a new TNM staging system for p16INK4A (p16)-positive OPSCC<sup>12</sup>, using p16 immunostaining as a surrogate marker for HPV infection, and even prompted clinical trials focusing on treatment de-intensification<sup>13</sup>.

Accurate HPV detection methods are required to obtain reliable data on the prevalence and prognosis of HPV-positive OPSCCs, and are absolutely crucial in patient selection for treatment de-escalation trials. RNA expression of the E6/E7 oncogenes on fresh-frozen samples generally is considered to be the gold standard for detection of oncogenic HPV infections<sup>14,15</sup>, but various alternative HPV detection methods have been developed and validated for use on formalin-fixed, paraffin-embedded (FFPE) tissue. Single-modality HPV-DNA testing by PCR methods is too sensitive, and may detect futile productive infections<sup>16</sup> or minor carryover contaminations<sup>17</sup>. Consequently, detection of HPV-DNA only is not specific enough. On the other hand, DNA in situ hybridization is less sensitive, and misses HPV-positive tumors<sup>14,18</sup>. Recently, RNA in situ hybridization techniques were developed<sup>19,20</sup> that combine the advantage of in situ hybridization (ie, localizing the virus directly to malignant cells) with higher sensitivity than DNA in situ hybridization<sup>21-23</sup>. However, this method is expensive and requires specific expertise, and at present the most frequently used approach is an algorithm using p16 as a surrogate marker. This algorithm comprises p16 immunostaining followed by PCR-based detection of high-risk HPV DNA on the p16-immunopositive samples<sup>14</sup>, and is accurate for OPSCC with a sensitivity of 98% and a specificity of 96%<sup>24</sup>. However, the algorithm relies on the sensitivity of the p16 immunostaining to select tumors for HPV-DNA testing. In two studies evaluating head and neck tumors outside of the oropharynx, the sensitivity of p16 immunohistochemistry was a mere 74% and 79%<sup>25,26</sup>, compared with almost 100% in OPSCC<sup>14,24</sup>. A lower sensitivity of p16 immunostaining will cause false-negative findings, hampering conclusions of the clinical relevance of HPV status in nonoropharyngeal HNSCC.

Because of these challenges of accurate HPV assessment, much less is known about the HPV-etiological fraction in tumors outside the oropharynx. Moreover, accurate data on the prognostic impact of HPV in nonoropharyngeal HNSCC currently are scarce as well. Older meta-analyses showed a worldwide pooled HPV-DNA estimate of 24.2% in oral cavity tumors<sup>8</sup>, but these estimates relied on PCR-based HPV-DNA detection methods, including even ultrasensitive nested PCR approaches. Recent studies that relied on additional markers for HPV status including RNA detection showed much lower attributable fractions, somewhere between 4% and 15% in the oral cavity<sup>7,25,26</sup>. In these studies, p16 immunostaining was bypassed either by using DNA in situ hybridization<sup>26</sup> or detection of E6/E7 RNA expression in FFPE specimen blocks<sup>7,25</sup>. Nonetheless some concerns remain. One aspect is the lack of clinical validation of the applied detection techniques, hampering interpretation and comparative analysis of the presented results, and obstructing the implementation for clinical decision making. A second concern is the standard practice of tissue sectioning on microtomes. Although many investigators use cleaning protocols to avoid sample-

to-sample contamination, this is a nuisance and prone to errors. Moreover, cutting and mounting of the sections is laborious and requires considerable hands-on time, which hampers the large studies that are necessary. On the other hand, a check of tumor presence and an estimate of tumor percentage is required for proper interpretation of molecular test results.

Based on these considerations and our previous experiences and those of others, we developed and validated a new, simple molecular HPV detection method suitable for FFPE specimens of HNSCC both inside and outside the oropharynx. We circumvent sectioning of the FFPE blocks to reduce contamination and hands-on time, apply combined extraction of high-quality DNA and RNA, and perform HPV-DNA detection of 15 high-risk HPV types with partial genotyping to allow subsequent detection of E6 mRNA for interpretation of the findings. We focused on HPV16 because tumors in the head and neck region are caused primarily by this virus type, but methods can be extended to other types as well. The new test algorithm that we named HPV rapid tissue core (HPV-rTcore) assay was validated clinically on both the DNA level and the RNA level for HPV16 using a cohort of OPSCCs and oral cavity squamous cell carcinomas (OSCCs).

## **MATERIALS AND METHODS**

### **Patients and Tissue Specimens**

To evaluate the presented testing method, two cohorts were used. The first cohort was described before<sup>24</sup>. In short, cohort 1 comprised 86 OPSCCs from patients who presented between 2008 and 2011 at Amsterdam University Medical Center at VUmc, and were included consecutively in a prospective sampling study. A biopsy of the tumor was performed under general anesthesia and two specimens were taken, of which one was fixed in 4% buffered formalin and embedded in paraffin, and the other was snap-frozen and stored in liquid nitrogen at  $-196^{\circ}\text{C}$ . The second cohort consisted of 25 OSCCs from historical sample series<sup>27,28</sup>. In addition, in the second cohort both FFPE and fresh-frozen material were available, either from pretreatment biopsy specimens or the resection specimen.

### **Historical Group Classification**

RNA extraction of the OPSCC and OSCC frozen biopsy specimens was described previously<sup>24,27,28</sup>, and the RNA subsequently was stored at  $-80^{\circ}\text{C}$ . Both DNA and RNA were analyzed. The presence of E6 mRNA in the fresh-frozen biopsy specimens of both the OPSCC and the OSCC cohort was considered the gold standard for etiologic HPV involvement<sup>14</sup>, and was detected by an HPV16 E6\*I RT-PCR–enzyme immunoassay (EIA) as previously described<sup>14</sup>. Among the 86 OPSCC tumors, 24 were HPV positive and 62 were HPV negative<sup>24</sup>. In total, 21 of 24 HPV-positive tumors contained HPV16, two contained HPV33, and one contained HPV35<sup>24</sup>. Among the 25 OSCCs, four were HPV16 positive<sup>27</sup> and 21 were HPV negative<sup>28</sup>.

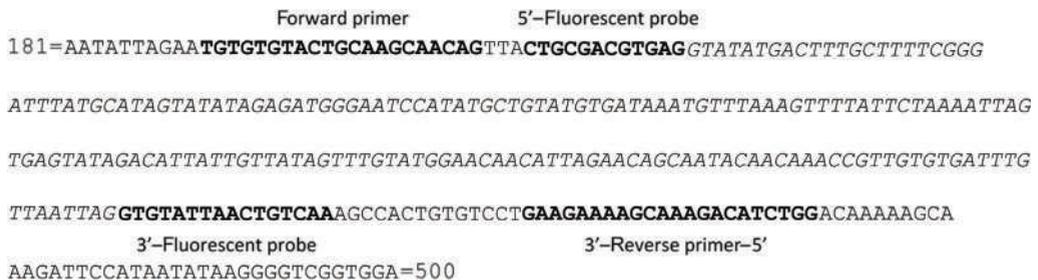
### **Nucleic Acid Extraction on FFPE Tissue Samples**

For sampling of tumor-enriched regions, sterile, disposable 1-mm biopsy punches with a plunger system (a gift from Claudia Küpper, Kai Europe GmbH, Solingen, Germany) were used as described previously<sup>29</sup>. The archived hematoxylin and eosin–stained slides from the tissue block for initial histologic diagnosis were used to select a tumor-enriched region from which to take a biopsy specimen. The hematoxylin and eosin section and tumor block were aligned and a punch was taken only when they matched macroscopically. When they did not match, a novel diagnostic section was prepared as the first step. No surface treatment of the FFPE blocks was applied. Although it has been considered, it was not required. The punch biopsy specimens were stored in microcentrifuge vials (Eppendorf, Hamburg, Germany) with a screw cap at room temperature.

Total nucleic acid extraction from the 1-mm punch biopsies was performed without deparaffinization by overnight lysis at 58°C in 250 µL lysis buffer containing 237.4 µL 10 mmol/L Tris-HCL (Invitrogen, Carlsbad, CA) pH 8.0/0.25% SDS (Bio-Rad Laboratories B.V., Veenendaal, the Netherlands), and 12.6 µL 20 mg/mL proteinase K (Invitrogen). The samples were incubated in an Eppendorf thermomixer with continuous shaking (500 rpm), and proteinase K was added again after four hours. DNA and RNA were isolated by a NucliSense EasyMag isolation procedure (bioMérieux SA, Marcy l'Étoile, France) and dissolved in 30 µL elution buffer (catalog number 280132; bioMérieux). DNA and RNA concentrations were measured using the Qubit DNA HS Assay Kit and Qubit RNA HS Assay Kit (Invitrogen) on a Qubit 4.0 fluorometer (Invitrogen) according to the manufacturer's protocol. Total nucleic acid isolates were stored at -80°C until use.

## HPV-DNA Detection and Genotyping

High-risk HPV-DNA detection and partial genotyping were performed using the HPV-Risk Assay according to the supplier's instructions (Self-screen B.V., Amsterdam, the Netherlands)<sup>30,31</sup>. In short, 5 µL isolated DNA (maximal DNA concentration, 20 ng/µL) was used as input for the HPV-Risk Assay and was run on a VIA 7 Real-Time PCR system (Applied Biosystems, Foster City, CA). This assay targets the E7 region of 15 high-risk HPV types (ie, HPV16, 18, 31, 33, 35, 39, 45, 51, 52, 56, 58, 59, 66, 67, and 68). Four different fluorescent channels are used: HPV16 and HPV18 are detected in separate fluorescent channels, detection of the other 13 HPV types is combined in the third fluorescent channel, and the human β-globin gene (HBB) is detected in a fourth channel. HBB serves as an internal control to confirm that the sample has sufficient amplifiable DNA. A sample was scored as HPV negative when the threshold cycle (CT) value for HBB was ≤36 and CT values for all HPV channels were >36. A sample was considered positive for HPV16, HPV18, and/or other high-risk HPV types when a CT value ≤36 in the channel for HPV16, HPV18, and/or other HPV types was obtained. A sample was recorded as invalid when the CT values for all channels (HPV and HBB) were >36. The cut-off value of 36 was chosen because it was the highest CT value shown by replicate assays to have the same CT value. At higher CT values, replicates became unreliable. See Results for more detail.



**Figure 1.** Primer and probe design of the E6\*I real-time quantitative PCR assay for human papillomavirus type 16. The E6\*I spliced-out sequence is shown in italics. Sequences complementary to the primers and probe are shown in bold.

## HPV16 E6 mRNA Detection

A TaqMan (Applied Biosystems)-based mRNA assay was developed to detect HPV16 E6 mRNA expression. The TaqMan assay was designed to specifically detect RNA and not DNA, using a splice-site spanning probe to detect the most abundant splice variant within the HPV16 E6 open reading frame (ie, E6\*I). Moreover, the amplicon length was kept short to facilitate optimal amplification of RNA from FFPE specimen blocks. A schematic overview of the new assay is shown in Figure 1. First, 8 µL RNA was treated with DNase I, amplification grade

(cat. 18068015; Invitrogen) in a 10- $\mu$ L reaction volume according to the manufacturer's protocol. The DNase-treated mRNA was diluted to a total volume of 26.4  $\mu$ L and used for real-time quantitative RT-PCR (RT-qPCR). RT-qPCR was performed on an ABI7500 Real-Time PCR system (Applied Biosystems) in a one-step reverse transcription and subsequent PCR reaction using the GoTaq Probe 1-Step RT-qPCR System (Promega, Madison, WI) according to the manufacturer's protocol. RT-qPCR analyses were performed in 20  $\mu$ L containing 10  $\mu$ L GoTaq Probe qPCR Master Mix (2 $\times$ ), 0.4  $\mu$ L GoScript RT Mix for 1-Step RT-qPCR (50 $\times$ ), 900 nmol/L forward and reverse primers (Biolegio B.V., Nijmegen, the Netherlands), 250 nmol/L hydrolysis probe (Biolegio B.V.), 4  $\mu$ L RNA, and 2.6  $\mu$ L nuclease-free water. Before use, 30 nmol/L carboxy-X-rhodamine reference dye was added to the GoTaq Probe qPCR Master Mix. Cycling conditions were 15 minutes at 45°C and two minutes at 95°C, followed by 40 cycles of 15 seconds at 95°C of denaturation, and 60 seconds at 60°C of annealing/extension. In parallel, a previously developed assay for the housekeeping gene encoding  $\beta$ -glucuronidase (GUSB)<sup>14</sup> was performed to examine the sample RNA quality. Sequences of primers and probes are listed in Table 1. The TaqMan assay was evaluated using two HPV16-positive HNSCC cell lines and one HPV-negative HNSCC cell line: UD-SCC-2<sup>32</sup> (viral copy number: 23.4<sup>33</sup>), UPCI:SCC-90<sup>34</sup> (viral copy number: 483.0<sup>33</sup>), and VU-SCC-120, which is HPV negative<sup>35</sup>. The detection limit was determined with a serial dilution of RNA from UD-SCC-2 and UPCI:SCC-90 using 20 ng to 2 pg RNA per reaction. The RNA specificity of the assay was determined by evaluation of RNA from UD-SCC-2 in four combinations: with or without pretreatment of the RNA with DNase I, amplification grade (Invitrogen), and with or without addition of reverse transcriptase. In addition, cDNA calibration curves were determined in the number of cDNA copies using constructed and synthesized gBlocks (Integrated DNA Technologies, Coralville, IA) that are fusions of HPV16 and GUSB sequences. The gBlock had the following sequence: 5'-CAAAAACGCAGAAAATACGTGGTTGGAGAGCTCATTTGGAATTTGCCGATTCATGACTGAACAGTCACCGA CGAGAGTGCTGGGGAATAAAAAGGGGATCTTCACTCGGCAGAGACAACCAAAAAGTACATGATATAATATTAGAAT GTGTGACTGCAAGCAACAGTTACTGCGACGTGAGGTGTATTAAGTGTCAAAGCCACTGTGTCCTGAAGAAAAGCAAAGA CATCTGGACAAAAGCAAAGATTCCAT-3' (standard font style corresponds to GUSB; italic font style corresponds to HPV16). Similar to the HPV-Risk Assay for DNA, a sample was scored as HPV16 negative when the CT value for GUSB was  $\leq$ 36 and the CT value for E6\*I was  $>$ 36. When the E6\*I CT value was  $\leq$ 36, the sample was considered HPV16 positive. A sample was recorded as invalid when the CT values for both GUSB and E6\*I were  $>$ 36.

### Assessment of Reproducibility

A series of samples were analyzed in duplicate to assess reproducibility, the effect of retesting, and potential sampling heterogeneity. Both the same RNA/DNA isolate ( $n = 27$ ) was used in an independent reaction at a later time point, and second punch biopsies were taken at a later time point from the same specimen blocks of 18 tumors. These biopsy specimens were processed similarly to the first punch biopsy, but at a later time point to examine the effect of retesting and sampling error. For both reproducibility assays, an interval of approximately three months was used.

### Statistical Analysis

HPV status assessed by HPV16 E6\*I RT-PCR-EIA assay<sup>14</sup> on nucleic acids isolated from the matched frozen specimens was used as the gold standard for the calculation of sensitivity, specificity, and accuracy. A Spearman rank correlation coefficient with corresponding P value was used to compare repeat samples because heteroskedasticity of the data precluded the use of a Pearson product-moment correlation coefficient. All statistical analyses were performed using package stats in R version 3.6.1 (R Project for Statistical Computing, <https://www.r-project.org>, last accessed July 5, 2019).

**Table 1.** Primer and Probe Sequences Used for E6\*I and GUSB Expression Analyses

Target	Sequence	Primer location in genome sequence (length)	Amplicon length, bp
HPV16 E6*I (NC_001526) <sup>a</sup>	F: 5'-TGTGTGTAAGCAAGCAACAG-3'	191-211 (21)	89
	R: 5'-CCAGATGCTCTTTGCTTTCTTC-3'	440-461 (22)	
	TP <sup>b</sup> : 5'-CTGCGACGTGAGGTGATTAAGTCAAAA-3'	215-226^409-425 <sup>c</sup> (29)	
GUSB (NM_000181) <sup>a</sup>	F: 5'-TTTGAATTTGCCGATTCAT-3'	7:65964334-65964355 (22)	84
	R: 5'-GTCTCTGCCGAGTGAAGAT-3'	7:65961013-65961031 (19)	
	TP <sup>b</sup> : 5'-CCAGCACTCTGGTCGGTGACTGTTCA-3'	7: 65961046-65964330 (26)	

Abbreviations: F, forward primer; GUSB,  $\beta$ -glucuronidase; HPV, human papillomavirus; R, reverse primer; TP, TaqMan probe.

<sup>a</sup> GenBank accession number (<https://www.ncbi.nlm.nih.gov/genbank/>).

<sup>b</sup> With 6-Carboxyfluorescein 5'-reporter and Black Hole Quencher 1 3'-quencher.

<sup>c</sup> HPV type 16 E6 is spliced between chromosome positions 226 and 409.

## RESULTS

Validation of the HPV-rTcore workflow presented was performed in subsequent steps. First, the HPV16 E6\*I RT-qPCR test was developed and its analytical performance was evaluated using a serial dilution of RNA from cell lines. Second, the HPV16 E6\*I RT-qPCR test was compared with the previously developed HPV16 E6\*I RT-PCR-EIA14 using archived RNA from fresh-frozen biopsy specimens of OPSCC patients to confirm the diagnostic accuracy. Next, HPV-DNA and mRNA detection was performed on nucleic acid isolates of FFPE punch biopsies of OPSCC and OSCC patients. The HPV status in this cohort was well defined and assessed by HPV16 E6\*I RT-PCR-EIA on RNA from the matched frozen samples. Finally, the accuracy was evaluated by a testing algorithm that combined HPV-DNA and mRNA testing on FFPE biopsy specimens. The results of each step are described in the following sections.

### Analytical Validation of RT-qPCR Assay for Detection of HPV16 E6\*I

Performance of the RT-qPCR TaqMan assay for HPV16 E6\*I was tested using RNA from two HPV16-positive cell lines, and one HPV-negative cell line. The assay was positive in the HPV16-positive cell lines, and negative in the HPV-negative cell line and technical negative controls (with or without DNase digestion and without addition of reverse transcriptase). No effect of DNase digestion was found on E6\*I and GUSB CT values. The detection limit of the assay using a serial dilution of fresh RNA from HPV-positive cell line UD-SCC-2 and UPCI:SCC-90 in sterile water was found to be 20 pg RNA (UD-SCC-2) and 2 pg RNA (UPCI:SCC-90). Moreover, at a CT value of 36, three DNA copies of GUSB and seven copies of HPV16 can be detected according to the gBlock calibration curve (Supplementary Figure 1). The assay was highly specific for the detection of RNA because assays without addition of reverse transcriptase did not detect any product with a CT value <36.

### Clinical Validation of HPV16 E6\*I RT-qPCR Assay on Fresh-Frozen Biopsy Specimens of OPSCC Patients

After technical confirmation of the RT-qPCR TaqMan assay performance, the diagnostic accuracy was benchmarked against our previous HPV16 E6\*I RT-PCR-EIA assay<sup>14</sup> using archived RNA from snap-frozen biopsy specimens of a previously reported cohort of 86 OPSCC patients with well-defined results<sup>24</sup>. The comparison is shown in Table 2. There was good agreement between the results of the previous EIA assay and the new TaqMan-based RT-qPCR assay (96%). The HPV16 E6\*I RT-qPCR assay has a high sensitivity (95%) and specificity (97%).

**Table 2.** HPV E6\*I Detection on RNA of Snap-Frozen Biopsy Specimens from OPSCC Patients by Taqman Assay versus the HPV16 E6\*I RT-PCR-EIA as Gold Standard

Taqman Assay		HPV16 E6*I RT-PCR-EIA gold standard <sup>a</sup>		
		HPV negative	HPV16 positive	other HPV positive
RT-qPCR Taqman <sup>b</sup>	HPV negative	60	1	3
	HPV16 positive	2	20	-

Abbreviations: EIA, enzyme immunoassay; HPV, human papillomavirus; HPV16, human papillomavirus type 16; RT-qPCR, real-time quantitative RT-PCR.

<sup>a</sup> Frozen samples.

<sup>b</sup> Identical frozen samples.

## Nucleic Acid Isolation from FFPE Material Using Punch Biopsies

Sampling of tumor cells for nucleic acid isolation was performed using sterile punch biopsies from tumor-enriched regions of FFPE specimens. The median estimated percentage of tumor of the enriched and sampled region was 50% (range, 10% to 80%; SD, 16%). After sampling, combined extraction of DNA and RNA was performed. The median DNA and RNA yields in 30  $\mu$ L elution volume were 10 ng/ $\mu$ L (range, 5 to 15 ng/ $\mu$ L; SD, 7.1 ng/ $\mu$ L), and 12 ng/ $\mu$ L (range, 4 to 26 ng/ $\mu$ L; SD, 5.8 ng/ $\mu$ L), respectively.

**Table 3.** HPV DNA Detection by the HPV-Risk Assay on FFPE Punch Biopsies from OPSCC and OSCC Patients versus the Gold Standard (HPV16 E6\*I RT-PCR-EIA on RNA From Matched Frozen Material)

HPV-Risk assay		HPV16 E6*I RT-PCR-EIA gold standard <sup>a</sup>		
		HPV negative	HPV16 positive	other HPV positive
OPSCC cohort <sup>b</sup>	HPV negative	55	1	-
	HPV16 positive	1	19	-
	other HPV positive	-	-	3
	Invalid <sup>c</sup>	1	-	-
OSCC cohort <sup>b</sup>	HPV negative	20	-	-
	HPV16 positive	-	4	-
	other HPV positive	-	-	-
	Invalid <sup>c</sup>	1	-	-

Abbreviations: EIA, enzyme immunoassay; FFPE, formalin-fixed paraffin-embedded; HPV, human papillomavirus; HPV16, human papillomavirus type 16; OPSCC, oropharyngeal squamous cell carcinoma; OSCC, oral cavity squamous cell carcinoma; RT-qPCR, real-time quantitative RT-PCR.

<sup>a</sup> RNA-based detection on frozen samples.

<sup>b</sup> DNA-based detection on FFPE punch biopsies.

<sup>c</sup> A sample was recorded as invalid when the threshold cycle values for all HPV channels and *HBB* were  $\geq 36$ .

**Table 4.** HPV16 E6\*I RNA Detection by RT-qPCR TaqMan Assay on FFPE Punch Biopsies from OPSCC and OSCC Patients versus the Gold Standard (HPV16 E6\*I RT-PCR-EIA on RNA from Matched Frozen Material)

RT-qPCR TaqMan assay		HPV16 E6*I RT-PCR-EIA gold standard <sup>a</sup>		
		HPV negative	HPV16 positive	other HPV positive
OPSCC cohort <sup>b</sup>	HPV negative	56	1	3
	HPV16 positive	1	19	-
	Invalid <sup>c</sup>	-	-	-
OSCC cohort <sup>b</sup>	HPV negative	21	-	-
	HPV16 positive	-	4	-
	Invalid <sup>c</sup>	-	-	-

Abbreviations: EIA, enzyme immunoassay; FFPE, formalin-fixed paraffin-embedded; HPV, human papillomavirus; HPV16, human papillomavirus type 16; OPSCC, oropharyngeal squamous cell carcinoma; OSCC, oral cavity squamous cell carcinoma; RT-qPCR, real-time quantitative RT-PCR.

<sup>a</sup> RNA-based detection on frozen samples.

<sup>b</sup> RNA-based detection on FFPE punch biopsies.

<sup>c</sup> A sample was recorded as invalid when the threshold cycle values for E6\*I and GUSB were  $\geq 36$ .

## Detection of HPV DNA of 15 High-Risk HPV Types in Nucleic Acid Isolates of Punch Biopsies

For confident HPV detection, we aimed for molecular HPV testing on FFPE biopsy specimens by combining a DNA test and an RNA test. First, we compared the results of the DNA test method against the results of the previous RT-PCR-EIA assay<sup>14</sup> on matched fresh-frozen tumor material, which we considered the gold standard. Two cohorts were used for validation: the OPSCC cohort (80 of 86 FFPE biopsy specimens could be retrieved), and an OSCC cohort consisting of 25 tumors. Matched fresh-frozen and FFPE samples were available. Results found in both cohorts are shown in Table 3. In total, 103 of 105 samples had valid DNA-based test results, with 1 false-positive and 1 false-negative result as compared with the gold standard. Hence, the HPV-Risk Assay on isolated DNA already was highly accurate in both cohorts (96% in both the OPSCC and OSCC cohorts). These results were stable when retesting the same DNA sample ( $\rho = 0.75$ ;  $P = 8 \times 10^{-6}$ ) (Figure 2A), and when evaluating a second punch biopsy from the same FFPE biopsy ( $\rho = 0.82$ ;  $P = 2 \times 10^{-5}$ ) (Figure 2B).

The reproducibility of results with CT values  $>36$  was low for HBB (Figure 2A) and HPV (Figure 2B). Because the template DNA input was standardized on 100 ng by Qubit DNA quantitation, the signal intensity depended mostly on the level of DNA template degradation, which was unpredictable. Therefore, the cut-off was set at a CT value of 36 to call a sample diagnostic (HBB) and positive (HPV). The limitation of using a CT value of 36 as the cut off is that sensitivity of the assay should be checked regularly, which is performed easily using a large aliquoted batch of DNA of a HPV-positive cell line or using synthetic DNA templates.

**Table 5.** HPV16 Detection by Combination of HPV DNA Detection by the HPV-Risk Assay and RT-qPCR TaqMan Assay on FFPE Punch Biopsies from OPSCC and OSCC Patients versus the Gold Standard (HPV16 E6\*1 RT-PCR-EIA on RNA from Matched Frozen Material)

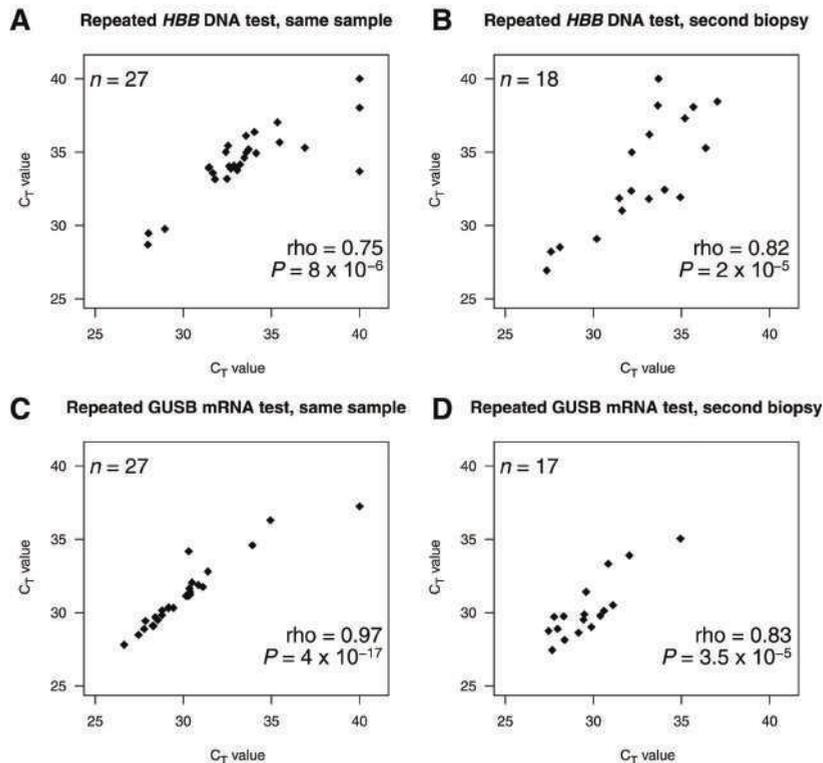
HPV-Risk Assay + RT-qPCR TaqMan assay		HPV16 E6*1 RT-PCR-EIA gold standard <sup>a</sup>		
		HPV negative	HPV16 positive	other HPV positive
OPSCC cohort <sup>b</sup>	HPV negative	56	1	3
	HPV16 positive	1	19	-
	Invalid <sup>c</sup>	-	-	-
OSCC cohort <sup>b</sup>	HPV negative	21	-	-
	HPV16 positive	-	4	-
	Invalid <sup>c</sup>	-	-	-

Abbreviations: EIA, enzyme immunoassay; FFPE, formalin-fixed paraffin-embedded; HPV, human papillomavirus; HPV16, human papillomavirus type 16; OPSCC, oropharyngeal squamous cell carcinoma; OSCC, oral cavity squamous cell carcinoma; RT-qPCR, real-time quantitative RT-PCR.

<sup>a</sup> RNA-based detection on frozen samples.

<sup>b</sup> DNA- and RNA-based detection on FFPE punch biopsies.

<sup>c</sup> A sample was recorded as invalid when both DNA and RNA testing were invalid.



**Figure 2.** Technical reproducibility.

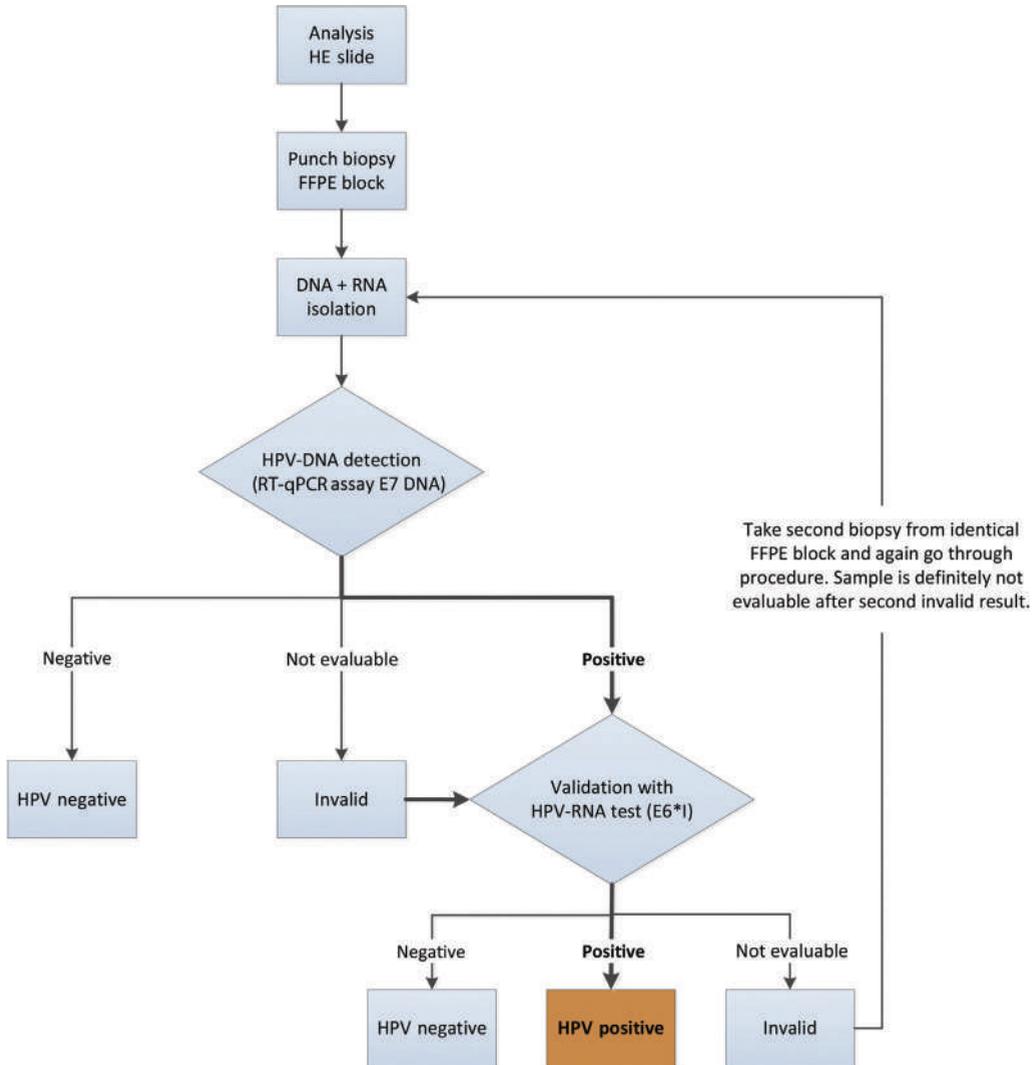
The x axis and y axis represent threshold cycle (CT) values of the housekeeping genes of the two replicates under comparison. Each diamond symbol represents a specific sample. Correlations were assessed with a Spearman rank correlation coefficient (rho) and corresponding P value. A: Technical reproducibility of the human papillomavirus (HPV)–Risk Assay by retesting the same DNA sample. The target is HBB. B: Reproducibility of HPV-Risk Assay on DNA samples from second punch biopsies of the same FFPE specimen. The target is HBB. C: Technical reproducibility of E6\*I real-time quantitative PCR (qPCR) assay by retesting the same RNA sample. The target is  $\beta$ -glucuronidase (GUSB). D: Reproducibility of the E6\*I qPCR assay on RNA samples from second punch biopsies of the same FFPE specimen. The target is GUSB. Note that reproducibility problems between replicates started to occur when one of the samples had a CT value of 36 or higher, which was the reason for setting the cut-off point at 36. n = 27 (A and C); n = 18 (B); n = 17 (D).

## Detection of HPV16 E6\*I in Nucleic Acid Isolates of Punch Biopsies

Next, the HPV16 E6\*I TaqMan RT-qPCR assay was validated using the same FFPE punch biopsy–derived nucleic acid isolates, and the results, similarly as described in the previous paragraph, were compared against the results of the previous HPV16 E6\*I RT-PCR-EIA assay<sup>14</sup> on the matched fresh–frozen tumor material (Table 4). All samples had a valid result in RT-qPCR. Because the assay was designed to detect HPV16, the three non–HPV16-positive OPSCC samples were negative, and these samples were neglected for calculation of the assay accuracy for molecular detection of HPV16, leaving one false-positive and one false-negative result as compared with the gold standard. The accuracy for HPV16 by E6\*I TaqMan RT-qPCR assay was very high (97% in the OPSCC cohort, and 100% in the OSCC cohort). Stability of the assay was ensured using the same nucleic acids of a sample (rho = 0.97; P = 4E-17) (Figure 2C), and a second isolate from the second punch biopsies (rho = 0.83; P = 3.5E-5) (Figure 2D). Again, reproducibility between replicates was low for CT values of 36 or higher. Hence, for RNA analyses, a CT value of 36 also was chosen as the cut-off point to call a sample diagnostic (GUSB) and positive (E6\*I).

## Development of Molecular HPV Testing Algorithm for HPV16 Suitable for FFPE Material

Finally, an algorithm that sequentially combines the DNA and RNA test was evaluated. A schematic overview is presented in Figure 3. The overall accuracy of the algorithm for HPV16 was 97% and 100% in the OPSCC and OSCC cohorts (Table 5), respectively.



**Figure 3.** Human papillomavirus rapid tissue core (HPV-rTcore) workflow for sample processing and determination of HPV in formalin-fixed, paraffin-embedded (FFPE) specimens.

The method is based on contamination-free sampling (punch biopsy with disposables), simultaneous DNA and RNA extraction, HPV-DNA detection with genotyping (HPV-Risk Assay), and HPV-RNA detection on the HPV-DNA-positive samples by real-time quantitative RT-PCR (RT-qPCR). HE, hematoxylin and eosin.

## DISCUSSION

In this study we developed and validated a molecular HPV detection method suitable for FFPE tissue blocks of HNSCC from all subsites. A molecular method was developed based on a contamination-proof sampling method using disposable punches and combined extraction of DNA and RNA that can be automated, followed by HPV-DNA detection with partial genotyping and detection of E6 mRNA in the HPV16 DNA-positive cases using a newly developed RT-qPCR test. This new RT-qPCR test is less laborious and more user-friendly than the previously developed HPV16 E6\*I RT-PCR-EIA assay<sup>14</sup> that was based on an EIA read out. DNA-based testing for 15 high-risk HPV types by HPV-Risk Assay was shown to be reproducible and highly accurate in both cohorts (96% in both the OPSCC and OSCC cohorts), even without combined RNA analysis. The combined method can further prove the presence of transcriptionally active HPV, and performed well with high reproducibility and an overall accuracy of 97% and 100% for the detection of HPV16 in the OPSCC and OSCC cohorts, respectively. Moreover, the RNA analysis added to the performance of the DNA test by increasing the percentage of valid test results from 98% to 100%.

This study compares the new assays with the original results of the HPV16 E6\*I RT-PCR-EIA on RNA from the matched frozen samples. However, whether this latter test is the absolute gold standard is obviously arbitrary. Three samples were found in which the results of the HPV16 E6\*I RT-PCR-EIA did not match the results of the new RT-qPCR test. It remains questionable which result should be considered correct. However, because the new test is being evaluated in this study, we chose to comply with the previous test results. The choice of this gold standard also influenced the calling of false-positive and false-negative test results of the FFPE tests. One sample was negative in the previously used HPV16 E6\*I RT-PCR-EIA, but HPV positive according to the new RT-qPCR test on RNA from frozen material, and also HPV-DNA and HPV-RNA were detected in FFPE. Moreover, this sample also had strongly positive staining for p16. Most likely, this is a true HPV-positive sample that was missed by the original assay. Moreover, a second discordant sample was found. This sample was HPV negative according to the new tests on RNA from frozen material and FFPE and p16 staining was negative. However, the HPV16 E6\*I RT-PCR-EIA was positive and we had to consider this sample to be HPV positive based on the original data. Most likely, this was a true HPV-negative sample with a false-positive EIA. Hence, the new assay may in fact be even more accurate than the presented data indicate.

Between samples, a large range of CT values that indicated a positive test result because of differences in relative tumor content, HPV-DNA copy number, expression level, and particularly nucleic acid quality was expected. Indeed, a relative difference of up to 8 CT values (DNA) and 10 CT values (RNA) was found between the samples with the highest and lowest levels of HPV-positive test results (data not shown), whereas the nucleic acid template input was comparable (50 to 100 ng). Given these differences in the level of HPV-positive test results, an absolute threshold rather than a relative one ( $\Delta$ CT) was used to score a test positive or negative. This threshold for the housekeeping genes was set at a CT value of 36 for both the DNA and RNA tests. This cut-off value was chosen based on the reproducibility of replicate assays. To implement this test in a different laboratory it is important to validate it on local equipment and other reagents before using the test for diagnostic purposes. The data of HPV-positive cell lines or synthetic DNA templates can be used for implementing the assay.

A relatively large number of samples were evaluable compared with other studies that applied combined DNA and RNA HPV detection. In the combined analysis for HPV16 testing all cases were diagnostic. Other studies reported 8.5%<sup>7</sup> and 4.9%<sup>25</sup> of nonevaluable cases. This difference may be explained by different DNA/RNA extraction methods or detection methods, and also differences in specimen quality.

There were three major strengths to this study. First, a thorough clinical validation of the developed testing algorithm for optimal interpretation of the results was included. Second, the use of disposable sampling methods diminishes the major problem of cross-contamination introduced by sectioning. The paraffin

microtomes are a suspected source of cross-sample contamination despite cleaning procedures and knife switches. This was tested before (data not shown), and sectioning in diagnostic testing was omitted. The routinely made hematoxylin and eosin slides can serve well macroscopically to guide the area for punching. Finally, the current workflow resulted in 100% evaluable cases.

There were also limitations to our approach. In this study, nucleic acid isolation was performed using the NucliSens EasyMag, which may not be available in every diagnostic laboratory. However, this method also should be applicable to any column-based nucleic acid isolation kit. Another limitation of this study was the focus on detection of HPV16 only at the RNA level. Future research will overcome this limitation because E6\*I assays for non-16 high-risk HPV types have been developed<sup>36</sup>, and a technical validation with synthetic DNA fragments (gBlocks) can be performed. Moreover, cell lines with HPV33 have been reported<sup>37</sup> and constructed<sup>38</sup>.

This simple and reliable detection method that we termed HPV-rTcore assay can be applied in large cohort studies to determine the attributable fraction of HPV in archival specimens, allowing analysis of prognostic implications.

## **ACKNOWLEDGMENTS**

The authors are grateful to professor Peter Snijders, who unfortunately passed away in 2018, for his support of this project.

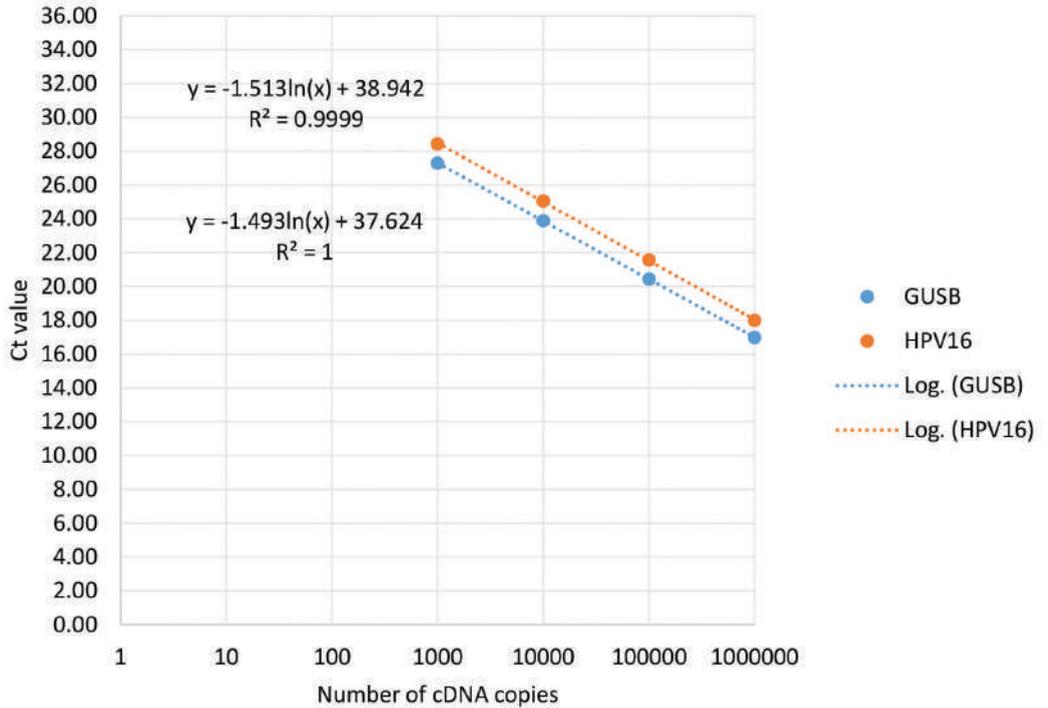
## **CONFLICTS OF INTEREST**

D.A.M.H. has served on the speakers bureau of Qiagen and on the scientific advisory boards of Pfizer and Bristol-Myers Squibb and is a minority shareholder of Self-screen B.V., a spin-off company of VUmc. Self-screen B.V. holds patents related to the work, and has developed and manufactured the high-risk human papillomavirus risk assay, which is licensed to Qiagen (as the QIAscreen human papillomavirus PCR test). Biopsy punches were provided by Kai Europe GmbH.

## **FUNDING**

Supported by the European Union's Seventh Framework Project grant 611425: OraMod (R.H.B.).

## SUPPLEMENTARY FIGURES



**Supplementary Figure S1.** Calibration curves of HPV16 E6\*I cDNA and  $\beta$ -glucuronidase (GUSB) cDNA gBlocks.

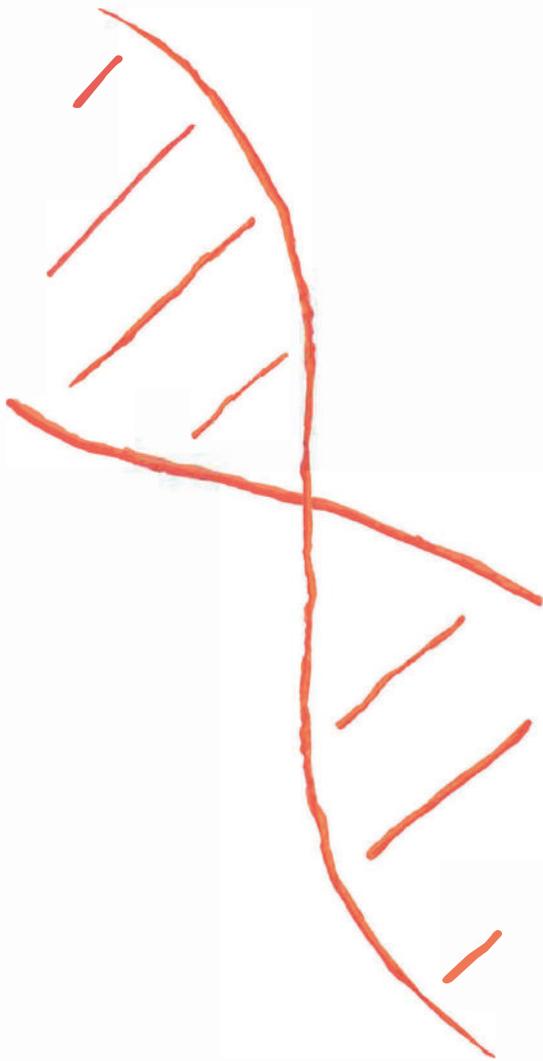
The x axis represents the absolute number of cDNA copies and the y axis represents the corresponding CT value. Solid dots are the average CT values (of three replicates) found when analyzing a given number of cDNA copies. Dashed lines are the actual calibration curves, and its functions are shown in the figure.

## REFERENCES

1. Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin* 2018; 68(6): 394-424.
2. Forastiere A, Koch W, Trotti A, Sidransky D. Head and neck cancer. *N Engl J Med* 2001; 345(26): 1890-900.
3. IARC. IARC Monographs on the evaluation of carcinogenic risks to humans, vol. 90. Human Papillomaviruses. , 2007; Lyon, France; , 2007.
4. Gillison ML, Chaturvedi AK, Anderson WF, Fakhry C. Epidemiology of Human Papillomavirus-Positive Head and Neck Squamous Cell Carcinoma. *J Clin Oncol* 2015; 33(29): 3235-42.
5. Westra WH. The morphologic profile of HPV-related head and neck squamous carcinoma: implications for diagnosis, prognosis, and clinical management. *Head Neck Pathol* 2012; 6 Suppl 1: S48-54.
6. Hayes DN, Van Waes C, Seiwert TY. Genetic Landscape of Human Papillomavirus-Associated Head and Neck Cancer and Comparison to Tobacco-Related Tumors. *J Clin Oncol* 2015; 33(29): 3227-34.
7. Castellsague X, Alemany L, Quer M, et al. HPV Involvement in Head and Neck Cancers: Comprehensive Assessment of Biomarkers in 3680 Patients. *J Natl Cancer Inst* 2016; 108(6): djv403.
8. Ndiaye C, Mena M, Alemany L, et al. HPV DNA, E6/E7 mRNA, and p16INK4a detection in head and neck cancers: a systematic review and meta-analysis. *Lancet Oncol* 2014; 15(12): 1319-31.
9. Ang KK, Harris J, Wheeler R, et al. Human papillomavirus and survival of patients with oropharyngeal cancer. *N Engl J Med* 2010; 363(1): 24-35.
10. O'Sullivan B, Huang SH, Siu LL, et al. Deintensification candidate subgroups in human papillomavirus-related oropharyngeal cancer according to minimal risk of distant metastasis. *J Clin Oncol* 2013; 31(5): 543-50.
11. Rietbergen MM, Brakenhoff RH, Bloemena E, et al. Human papillomavirus detection and comorbidity: critical issues in selection of patients with oropharyngeal cancer for treatment De-escalation trials. *Ann Oncol* 2013; 24(11): 2740-5.
12. Brierley JDG, M. K.; Wittekind, C., eds. TNM Classification of Malignant Tumours. 8th Edition ed. Hoboken: Wiley-Blackwell; 2016.
13. Mirghani H, Amen F, Blanchard P, et al. Treatment de-escalation in HPV-positive oropharyngeal carcinoma: ongoing trials, critical issues and perspectives. *Int J Cancer* 2015; 136(7): 1494-503.
14. Smeets SJ, Hesselink AT, Speel EJ, et al. A novel algorithm for reliable detection of human papillomavirus in paraffin embedded head and neck cancer specimen. *Int J Cancer* 2007; 121(11): 2465-72.
15. 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 Res* 2012; 72(19): 4993-5003.
16. Kreimer AR, Bhatia RK, Messegue AL, Gonzalez P, Herrero R, Giuliano AR. Oral human papillomavirus in healthy individuals: a systematic review of the literature. *Sex Transm Dis* 2010; 37(6): 386-91.
17. Boy S, Van Rensburg EJ, Engelbrecht S, Dreyer L, van Heerden M, van Heerden W. HPV detection in primary intra-oral squamous cell carcinomas--commensal, aetiological agent or contamination? *J Oral Pathol Med* 2006; 35(2): 86-90.
18. Schache AG, Liloglou T, Risk JM, et al. Evaluation of human papilloma virus diagnostic testing in oropharyngeal squamous cell carcinoma: sensitivity, specificity, and prognostic discrimination. *Clin Cancer Res* 2011; 17(19): 6262-71.
19. Wang F, Flanagan J, Su N, et al. RNAscope: a novel in situ RNA analysis platform for formalin-fixed, paraffin-embedded tissues. *J Mol Diagn* 2012; 14(1): 22-9.
20. Ukpo OC, Flanagan JJ, Ma XJ, Luo Y, Thorstad WL, Lewis JS, Jr. High-risk human papillomavirus E6/E7 mRNA detection by a novel in situ hybridization assay strongly correlates with p16 expression and patient outcomes in oropharyngeal squamous cell carcinoma. *Am J Surg Pathol* 2011; 35(9): 1343-50.
21. Bishop JA, Ma XJ, Wang H, et al. Detection of transcriptionally active high-risk HPV in patients with head and neck squamous cell carcinoma as visualized by a novel E6/E7 mRNA in situ hybridization method. *Am J Surg Pathol* 2012; 36(12): 1874-82.
22. Rooper LM, Gandhi M, Bishop JA, Westra WH. RNA in-situ hybridization is a practical and effective method for determining HPV status of oropharyngeal squamous cell carcinoma including discordant cases that are p16 positive by immunohistochemistry but HPV negative by DNA in-situ hybridization. *Oral Oncol* 2016; 55: 11-6.
23. Kerr DA, Arora KS, Mahadevan KK, et al. Performance of a Branch Chain RNA In Situ Hybridization Assay for the Detection of High-risk Human Papillomavirus in Head and Neck Squamous Cell Carcinoma. *Am J Surg Pathol* 2015; 39(12): 1643-52.

24. Rietbergen MM, Leemans CR, Bloemena E, et al. Increasing prevalence rates of HPV attributable oropharyngeal squamous cell carcinomas in the Netherlands as assessed by a validated test algorithm. *Int J Cancer* 2013; 132(7): 1565-71.
25. Lingen MW, Xiao W, Schmitt A, et al. Low etiologic fraction for high-risk human papillomavirus in oral cavity squamous cell carcinomas. *Oral Oncol* 2013; 49(1): 1-8.
26. Chung CH, Zhang Q, Kong CS, et al. p16 protein expression and human papillomavirus status as prognostic biomarkers of nonoropharyngeal head and neck squamous cell carcinoma. *J Clin Oncol* 2014; 32(35): 3930-8.
27. van Houten VM, Leemans CR, Kummer JA, et al. Molecular diagnosis of surgical margins and local recurrence in head and neck cancer patients: a prospective study. *Clin Cancer Res* 2004; 10(11): 3614-20.
28. Mes SW, Te Beest D, Poli T, et al. Prognostic modeling of oral cancer by gene profiles and clinicopathological co-variables. *Oncotarget* 2017; 8(35): 59312-23.
29. Rodrigo JP, Heideman DA, Garcia-Pedrero JM, et al. Time trends in the prevalence of HPV in oropharyngeal squamous cell carcinomas in northern Spain (1990-2009). *Int J Cancer* 2014; 134(2): 487-92.
30. Hesselink AT, Berkhof J, van der Salm ML, et al. Clinical validation of the HPV-risk assay, a novel real-time PCR assay for detection of high-risk human papillomavirus DNA by targeting the E7 region. *J Clin Microbiol* 2014; 52(3): 890-6.
31. Polman NJ, Ostrbenk A, Xu L, et al. Evaluation of the Clinical Performance of the HPV-Risk Assay Using the VALGENT-3 Panel. *J Clin Microbiol* 2017; 55(12): 3544-51.
32. Ballo H, Koldovsky P, Hoffmann T, et al. Establishment and characterization of four cell lines derived from human head and neck squamous cell carcinomas for an autologous tumor-fibroblast in vitro model. *Anticancer Res* 1999; 19(5B): 3827-36.
33. Akagi K, Li J, Broutian TR, et al. Genome-wide analysis of HPV integration in human cancers reveals recurrent, focal genomic instability. *Genome Res* 2014; 24(2): 185-99.
34. White JS, Weissfeld JL, Ragin CC, et al. The influence of clinical and demographic risk factors on the establishment of head and neck squamous cell carcinoma cell lines. *Oral Oncol* 2007; 43(7): 701-12.
35. Hermsen MA, Joenje H, Arwert F, et al. Centromeric breakage as a major cause of cytogenetic abnormalities in oral squamous cell carcinoma. *Genes Chromosomes Cancer* 1996; 15(1): 1-9.
36. Halec G, Schmitt M, Dondog B, et al. Biological activity of probable/possible high-risk human papillomavirus types in cervical cancer. *Int J Cancer* 2013; 132(1): 63-71.
37. Lin CJ, Grandis JR, Carey TE, et al. Head and neck squamous cell carcinoma cell lines: established models and rationale for selection. *Head Neck* 2007; 29(2): 163-88.
38. Schutze DM, Snijders PJ, Bosch L, Kramer D, Meijer CJ, Steenbergen RD. Differential in vitro immortalization capacity of eleven (probable) [corrected] high-risk human papillomavirus types. *J Virol* 2014; 88(3): 1714-24.





# 3

## **Prognostic modeling of oral cancer by gene profiles and clinicopathological co-variables**

Steven W. Mes\*, Dennis te Beest\*, Tito Poli, Silvia Rossi,  
Kathrin Scheckenbach, Wessel N. van Wieringen,  
Arjen Brink, Nicoletta Bertani, Davide Lanfranco,  
Enrico M. Silini, Paul J. van Diest, Elisabeth Bloemena,  
C. René Leemans, Mark A. van de Wiel\*, Ruud H. Brakenhoff\*

*\*These authors have contributed equally to this work.*

**Oncotarget 2017; 8(35): 59312-23**

## **ABSTRACT**

Accurate staging and outcome prediction is a major problem in clinical management of oral cancer patients, hampering high precision treatment and adjuvant therapy planning. Here, we have built and validated multivariable models that integrate gene signatures with clinical and pathological variables to improve staging and survival prediction of patients with oral squamous cell carcinoma (OSCC). Gene expression profiles from 249 human papillomavirus (HPV)-negative OSCCs were explored to identify a 22-gene lymph node metastasis signature (LNMsig) and a 40-gene overall survival signature (OSsig). To facilitate future clinical implementation and increase performance, these signatures were transferred to quantitative polymerase chain reaction (qPCR) assays and validated in an independent cohort of 125 HPV-negative tumors. When applied in the clinically relevant subgroup of early-stage (cT1-2N0) OSCC, the LNMsig could prevent overtreatment in two-third of the patients. Additionally, the integration of RT-qPCR gene signatures with clinical and pathological variables provided accurate prognostic models for oral cancer, strongly outperforming TNM. Finally, the OSsig gene signature identified a subpopulation of patients, currently considered at low-risk for disease-related survival, who showed an unexpected poor prognosis. These well-validated models will assist in personalizing primary treatment with respect to neck dissection and adjuvant therapies.

## INTRODUCTION

Head and neck squamous cell carcinoma (HNSCC) is the 7th most common tumor in the world<sup>1</sup>. HNSCC originates in the mucosal linings of the oral cavity, oropharynx, hypopharynx and larynx. The majority of patients (30-40%) present with oral squamous cell carcinoma (OSCC)<sup>2</sup>. Classical risk factors for HNSCC are tobacco use and alcohol consumption. Additionally, human papillomavirus (HPV) infection became manifest as a cause during the last decade. The HPV-attributable fraction is highest in oropharyngeal squamous cell carcinoma (OPSCC), and varies from 20-90% depending on the geographical region<sup>3</sup>. Also oral cancers may arise from HPV infection, but the attributable fraction is lower, ranging from 0-6%<sup>4</sup>. OPSCCs caused by HPV infection are different at the molecular level<sup>5</sup> and have a highly favorable prognosis<sup>6</sup>. This different clinical behavior led to treatment de-intensifying trials to personalize treatment and a staging adaptation in the 8th edition of the TNM Classification of Malignant Tumors of the Union for International Cancer Control (UICC)<sup>7</sup>.

The 5-years overall survival for OSCC is 60%, but ranges from 10 to 80% depending on the extent of the tumor at diagnosis<sup>8</sup>, as defined by the TNM stage. TNM staging is based on prognosis and employed for treatment planning in patients with OSCC<sup>9</sup>, but is group-based and meets limitations for personalizing treatment of the individual patient.

OSCC is mainly treated by surgery with or without postoperative radiotherapy or chemoradiotherapy, and besides TNM stage, additional important prognostic features are derived from histopathological examination of the surgical specimen. For example, tumor-positive surgical margins (R+) and lymph node metastasis (LNM) with extracapsular spread (ECS) are classical treatment-decisive prognostic factors and indicators for postoperative chemoradiotherapy. Of note, histopathological examination of the specimen is only available for postoperative therapy decisions, and not for pre-treatment prediction of prognosis and treatment planning. Particularly for patients with a clinically N0 neck an important choice has to be made between elective treatment of the neck, with associated morbidity, or active surveillance with the risk of occult lymph node metastases that will become manifest during follow-up. Molecular profiling of tumor specimen may provide additional, objective information to improve current prognostication, and can even be performed on pretreatment biopsies to stage the neck.

Several prognostic models based on molecular profiles have been evaluated for HNSCC in general, or for OSCC specifically<sup>10-13</sup>. These models predicted survival of the studied populations, and added independent information to other established prognostic factors. However, none of these models has been introduced in clinical practice. Reasons are (1) insufficient clinical validation of the models, (2) the complexity and lack of reproducibility of the different profiling platforms<sup>14</sup>, (3) heterogeneous study populations regarding HPV status and tumor subsite, (4) the high costs of transcriptomic profiling, and (5) the lack of compatibility with formalin-fixed paraffin-embedded (FFPE) tissue specimen. Translation of expression profiles to quantitative real-time polymerase chain reaction (qPCR) platforms using selected gene panels may overcome most of these disadvantages.

Another argument holds true for expression profiles associated with the clinically N0 neck. Previously, an expression profile has been identified and appropriately validated in a multicenter trial<sup>15-17</sup>. The signature remained accurate with negative predictive values (NPV) of 88% to 90% in the clinically relevant subgroup. However, the sentinel node biopsy is a competing diagnostic modality in this patient group with an even higher NPV of 95%<sup>18</sup>. Notwithstanding, sentinel node biopsy has not been introduced widely, has a poor performance for floor of mouth tumors, and has the obvious disadvantage that it remains a surgical procedure with radioactive tracers, whereas for gene expression analysis only a biopsy is required. Particularly, switching to RT-qPCR analysis of a thoroughly selected gene panel may further enhance the predictive power of the gene signature because of the large dynamic range of RT-qPCR.

We therefore aimed to identify and test gene expression signatures to address these important challenges in head and neck oncology: prediction of lymph node metastasis (LNM) and overall survival (OS). First, signatures of informative genes were selected from gene expression data by regression methods. Next, a limited number of genes were selected for platform transition to RT-qPCR assays, and the prognostic power was validated

using an independent cohort of surgically-treated HPV-negative OSCC patients. The molecular data were further combined with clinical and pathology data to provide the most accurate models for clinical practice to predict nodal metastatic disease and prognosis.

## **MATERIALS AND METHODS**

### **Patients**

Four independent cohorts of human papillomavirus (HPV)-negative OSCC patients were included (1) a cohort of 2 merged tumor gene expression profiles (array cohort 1, AC1) from the University Medical Center Utrecht (UMCU) and VU University Medical Center Amsterdam (VUmc); (2) a cohort of tumor gene expression profiles (array cohort 2, AC2) from the University Hospital Parma Medical Center (UHPMC); (3) an independent cohort of frozen tumor samples from VUmc, UHPMC and University Hospital Düsseldorf (UHD) for RT-qPCR gene expression profiling (qPCR cohort); and (4) an RNAseq dataset of OSCC tumors from The Cancer Genome Atlas (TCGA) Network<sup>19</sup>. Use of tissue from surgical specimen adhered to nation- and institution-specific procedures and guidelines. Informed consent was obtained of enrolled patients, when required. This study followed the Guidelines for the REporting of tumor MARKer Studies (REMARK)<sup>20</sup> (Supplementary Table 6).

### **HPV status**

HPV status was either determined with p16 immunostaining followed by HPV DNA PCR on p16-positive samples (AC1) and/or with HPV16 E6\*I RT-PCR in the AC1 and qPCR cohorts. Both assays have been validated and described before<sup>21</sup>. In AC2, the HPV status was not available. In the other cohorts on the other hand, 1 out of 151 (AC1) and 1 out of 126 (qPCR cohort) tumors were HPV-positive. Hence, the contribution of HPV positive tumors in AC2 was assumed to be low and no samples were excluded.

### **Gene expression datasets**

Similarly preprocessed VUmc (GSE84846) and UMCU (GSE30788) microarray datasets were combined, and comparability of the expression data of both centers was ensured. Data from AC2 (GSE84846) were not combined to the other datasets, because of a different reference design: Universal Human Reference RNA (cat. 740000, Agilent Technologies, Santa Clara, CA, USA) in AC1 and a pool of cell line RNA in AC2 (CAL 27, ATCC CRL-2095, American Type Culture Collection, Manassas, VA, USA). All preprocessing steps of the microarray data were performed using the limma package<sup>22</sup> in R (Supplementary Materials).

### **RT-qPCR**

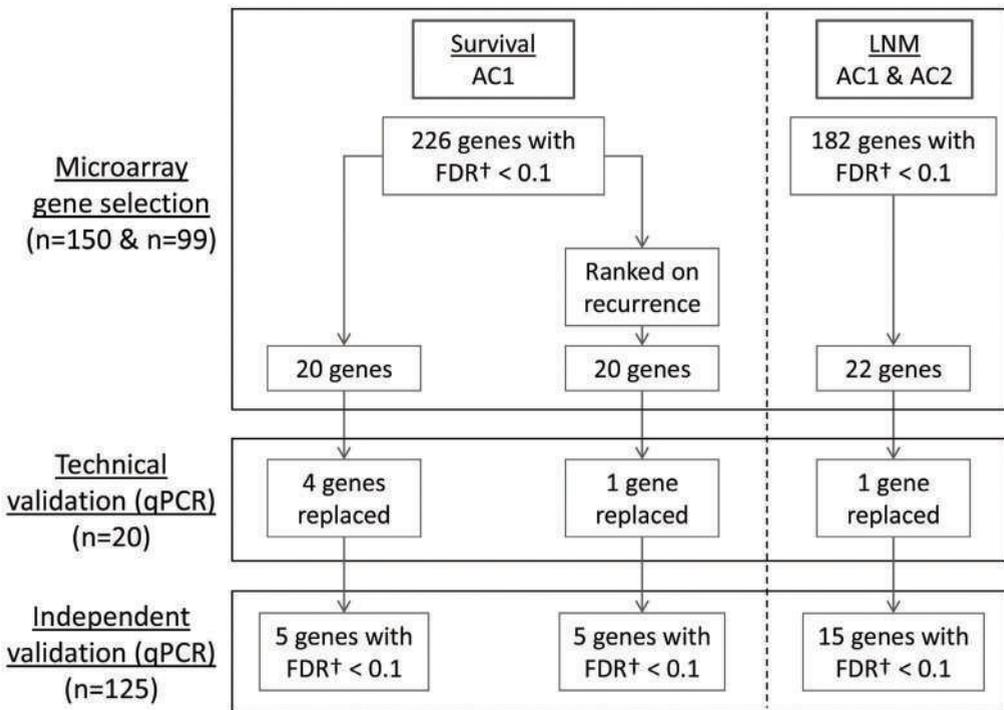
RNA was purified from fresh frozen tumor tissue and synthesis of cDNA was performed from 1 µg of total RNA using the High-Capacity RNA-to-cDNA Kit (cat. 4387406, Applied Biosystems; Foster City, CA). qPCR was performed using Taqman Low-Density Array (TLDA) Cards (cat. 4346800, Applied Biosystems) (Supplementary Table 2). qPCR Ct values were determined with predefined thresholds that were equal per gene for all patients. Relative gene expression was determined by the  $\Delta\Delta\text{Ct}$  method<sup>23</sup> using GUSB Ct-values for normalization. GUSB was selected as the most stable housekeeping gene (see Supplementary Table 7) out of four candidate genes (GAPDH, GUSB, RPLP0, and RPL4).

### **Statistical analyses**

Per dataset, the predictive power for LNM and survival was assessed with the global test<sup>24,25</sup>. Datasets with significant predictive power ( $p < 0.05$ ) were used for gene selection. Genes were selected from the microarray data by using a combination (detailed later) of lasso logistic regression or lasso Cox regression and univariable FDR-based association analysis. The latter was included to enhance reproducibility of individual markers assayed

by qPCR. The gene selection procedure is displayed in Figure 1 and further detailed in the Supplementary Materials. For the LNM genes, the p-values per gene of AC1 and AC2 were combined by Fisher's combined probability test, whereas for the prognostic genes only p-values of AC1 were considered, because the AC2 data did not pass the global test. For technical validation, the correlation between microarray and RT-qPCR data of 20 cases was determined by Pearson's correlation coefficient. For the RT-qPCR data, the univariable association of delta Ct values of the selected genes with either LNM or OS was determined by logistic or Cox regression, respectively. For prediction on independent samples, clinical variables were selected using stepwise regression, followed by adding the selected genes in a logistic (Cox) ridge regression to render multi-type prediction models. Model performance was assessed by bootstrapping. The prediction models for outcome consisted of (1) prognostic genes, (2) significant clinical factors and pathological TNM-stage (pTNM), (3) significant clinical factors and a composite pathological variable (positive if ECS or R+ surgical margins or >1 LNM was present), and the combinations (4) 1+2 and (5) 1+3. The predictive performance was assessed by area-under-the-ROC-curve (AUC) and integrated AUC (iAUC) over 5-year follow-up time for LNM and OS, respectively, complemented for LNM by the negative predictive value (NPV). Additive value of the gene signature was assessed with the global test. All statistical tests performed were two-sided. Univariable p-values were corrected for multiple testing using the Benjamini-Hochberg FDR procedure<sup>26</sup>.

3



**Figure 1.** Schematic representation of the different phases of the study.

Two microarray cohorts (Array Cohort 1 (AC1), n=150; Array Cohort 2 (AC2), n=99) were explored by univariable and multivariable gene selection to identify a 22-gene lymph node metastasis signature (LNMsig) and a 40-gene overall survival signature (OSsig). For the OSsig, 20 genes were selected that were predictive for OS, and 20 additional genes were selected after the genes were ranked on their predictive value for recurrent disease to account for disease-specific death. For LNM prediction, a previously validated multigene microarray signature<sup>(15–17)</sup> was used as preselection. Subsequently, our signatures were transferred to RT-qPCR assays and correlated to the microarray data in 20 cases (technical validation). After this technical validation, 6 genes with poor correlation coefficients were replaced by the second best genes from the initial microarray analyses. Finally, the definitive signatures were validated on an independent cohort of 125 tumors (independent validation). †Univariable p-values were corrected for multiple testing using the Benjamini-Hochberg FDR procedure. AC1, Array Cohort 1; AC2, Array Cohort 2; FDR, false discovery rate; LNM, lymph node metastasis; qPCR, quantitative polymerase chain reaction.

## RESULTS

Microarray data from two cohorts, 150 OSCC patients from The Netherlands (Array Cohort 1, AC1) and 99 OSCC patients from Italy (Array Cohort 2, AC2), were used to identify genes related to LNM and OS (Table 1). LNM was present in 60% of AC1 patients and 49.5% of AC2 patients. In AC1, the median overall follow-up time was 7.2 years (95% CI = 6.7 – 8.1). In AC2, the median overall follow-up time was 3.5 years (95% CI = 3.3 – 4.3).

**Table 1.** Characteristics of Patients in the Four Study Cohorts<sup>a</sup>

Characteristic	Array Cohort 1 (n = 150)	Array Cohort 2 (n = 99)	qPCR Cohort (n = 125)	TCGA Cohort (n = 160)	P <sup>b</sup> Value
Age, mean (SD)	62 (10.7)	66 (10.3)	63 (12.6)	62 (13.6)	P=0.06
Gender					
Male (%)	90 (60.0)	54 (54.5)	72 (57.6)	105 (65.6)	P=0.30
Female (%)	60 (40.0)	45 (45.5)	53 (42.2)	55 (34.4)	
Smoking (PY)					
0-10 (%)	36 (24.0)	51 (51.5)	41 (32.8)	47 (29.4)	P<0.001
11-24 (%)	19 (12.7)	10 (10.1)	13 (10.4)	13 (8.1)	
>24 (%)	95 (63.3)	38 (38.4)	71 (56.8)	60 (37.5)	
Unknown (%)	-	-	-	40 (25.0)	
Subsite					
Oral tongue (%)	53 (35.3)	41 (41.4)	48 (38.4)	-	P=0.62
Other oral cavity (%)	97 (64.7)	58 (58.6)	77 (61.6)	-	
TNM stage					
I (%)	18 (12.0)	22 (22.2)	16 (12.8)	10 (6.3)	P=0.02
II (%)	22 (14.7)	12 (12.1)	27 (21.6)	32 (20.0)	
III (%)	31 (20.7)	21 (21.2)	26 (20.8)	25 (15.6)	
IV (%)	79 (52.7)	44 (44.4)	56 (44.8)	82 (51.3)	
Unknown (%)	-	-	-	11 (6.9)	
N-stage					
Negative (%)	60 (40)	48 (48.5)	61 (48.8)	57 (35.6)	P=0.35
Positive (%)	90 (60)	49 (49.5)	64 (51.2)	76 (47.5)	
Unknown	-	2 (2.0)	-	27 (16.9)	
pCompVar <sup>c</sup>					
Negative (%)	-	-	79 (63.2)	-	
Positive (%)	-	-	38 (30.4)	-	
Unknown (%)	-	-	8 (6.4)	-	

Abbreviation: pCompVar, pathological composite variable; PY, packyears; SD, standard deviation.

<sup>a</sup> Percentages may not total 100 because of rounding.

<sup>b</sup> P values were calculated with the use of One-Way ANOVA for continuous variables and  $\chi^2$  test for categorical variables.

<sup>c</sup> Scored positive if extracapsular spread or positive resection margins or >1 lymph node metastasis was present.

### Identification of genes for prediction of lymph node metastasis and survival in OSCC

The gene selection strategy is summarized in Figure 1 and described in detail in the Supplementary Materials. In short, the previously published LNM gene profile<sup>15,17</sup> was evaluated to predict N-stage in AC1 and AC2. Using the global test with pathological N-stage as outcome, these genes had a p-value of 9.3E-06 and 9.9E-03 in AC1 and AC2, respectively. Combined univariable analysis identified 221 significant genes (FDR<0.1, Supplementary Table 1). From these genes, 22 genes were selected for RT-qPCR validation based on their ranking in univariable and multivariable analysis.

For survival, a similar gene pre-selection strategy was hampered by the lack of thoroughly validated prognostic gene signatures in the public domain. We therefore included other techniques to reduce the dimensions of the data, but also explored all genes to ensure that important prognostic genes were not missed. We only used AC1 for gene selections, as AC2 did not pass the global test due to the shorter follow-up time (global test

p-values AC1: 7.8E-3 and AC2: 0.73). Univariable analysis of all genes identified 226 (out of 37,662) significant genes in AC1 (FDR<0.1, Supplementary Table 1). Next, 20 genes were selected by univariable and multivariable analyses for survival, and 20 additional genes were selected after ranking the genes on their predictive value for recurrent disease to account for disease-specific death (see Figure 1). Two genes overlapped between the 40 survival genes and the 22 LNM genes (Supplementary Figure 1), rendering an overall signature of 60 target genes for technical and independent RT-qPCR validation (Supplementary Table 2).

### Technical RT-qPCR validation of identified genes

First, the 60 target genes were technically validated in a subset of 20 cases from AC2 to evaluate the platform transition. For these 20 cases, correlation coefficients were calculated between microarray and corresponding RT-qPCR data (Supplementary Table 3). In total, 52 of 60 genes validated well, as they showed a good correlation between microarray and RT-qPCR data (mean  $r=0.64$ ,  $SD=0.26$ ). The remaining 8 genes correlated poorly, showing a correlation coefficient  $>1$  SD below the mean. Cox regression nonetheless indicated that two of these eight genes did correlate with survival, i.e. EIF5 ( $P=0.011$ ) and ATP6V0A1 ( $P=0.057$ ), and these were therefore kept in the panel. The remaining 6 genes were replaced by the second best genes from the initial microarray analyses (Supplementary Table 1), and subsequently analyzed.

### Independent RT-qPCR validation of selected genes

The RT-qPCR validation cohort consisted of 125 OSCC cases that were independent from both microarray cohorts. In this validation cohort, nodal metastasis was detected in 51.2% of patients, and the median overall follow-up time was 5.1 years (95% CI = 4.4 – 6.3) (Table 1). The selected genes were run on customized microfluidic RT-qPCR cards, and the results were tested by univariable analyses and corrected for multiple testing. From the LNMsig 15 of 22 genes had an FDR<0.1 (Supplementary Table 4). From the OSSig 10 of 40 genes had an FDR<0.1 for OS, seven of which also significantly associated with disease-free survival (DFS). Thus, after correction for multiple testing, in total 25 of 60 genes selected from microarray datasets could be validated with RT-qPCR assays in an independent patient cohort.

**Table 2.** Performance Metrics of Gene Signature in N-stage prediction

	qPCR validation, all (n = 125)	qPCR validation, cT1-2N0 (n = 54)
NPV (95% CI) <sup>a</sup>	66 (57.1-74.7)	84 (71.7-95.2)
TN	40	26
TN + FN	61	31
PPV (95% CI) <sup>a</sup>	67 (59.1-76.6)	43 (21.5-64.5)
TP	43	10
TP + FP	64	23
Sensitivity (95% CI) <sup>a</sup>	67 (42.3-83.5)	67 (29.6-93.2)
TP	43	10
TP + FN	64	15
Specificity (95% CI) <sup>a</sup>	66 (39.3-83.2)	67 (39.7-86.2)
TN	40	26
TN + FP	61	39
AUC (95% CI) <sup>a</sup>	0.69 (0.63-0.75)	0.66 (0.52-0.78)

Abbreviations: AUC, Area Under the ROC Curve; CI, confidence interval; FN, false negative; FP, false positive; NPV, negative predictive value; PPV, positive predictive value; TN, true negative; TP, true positive.

<sup>a</sup>. Confidence intervals were assessed by bootstrapping.

## **A gene expression-based model to predict lymph node metastasis in OSCC**

The performance of the LNM predictive signature is summarized in Table 2; see Supplementary Table 4 for the estimates per gene. When all clinical stages of disease are considered, the AUC of this model was 0.69 (Table 2), with an NPV of 66% (Table 2). Next, we performed a subgroup analysis on the clinically relevant subset of tumors with clinical stages I and II (n=54), because these tumors qualify for transoral resection without treatment of the neck. In this subgroup, the AUC (0.66, Table 2) and the sensitivity of the LNMsig (67%, Table 2) were comparable with the performance statistics in all stages. The NPV, however, increased from 66% to 84% (Table 2). There were no clinical variables that correlated to LNM (data not shown) and data from histopathology is not available before surgery planning. Moreover, the fraction of occult lymph node metastasis was comparable in cT1 and cT2 tumors (i.e. 25% and 29% respectively). Previously, Van Hooff et al.<sup>17</sup> proposed a clinical decision model that recommends an elective neck dissection when the gene expression signature prediction indicates N+ or active surveillance when the prediction is N0, and estimated the benefit. Following this decision model, the LNMsig shows a similar benefit and could have prevented overtreatment in over 66% of the pN0 cases (72% or 24% overtreatment without or with the clinical decision model, respectively; see Figure 2).

## **A gene expression-based prognostic model for OSCC with independent prognostic value**

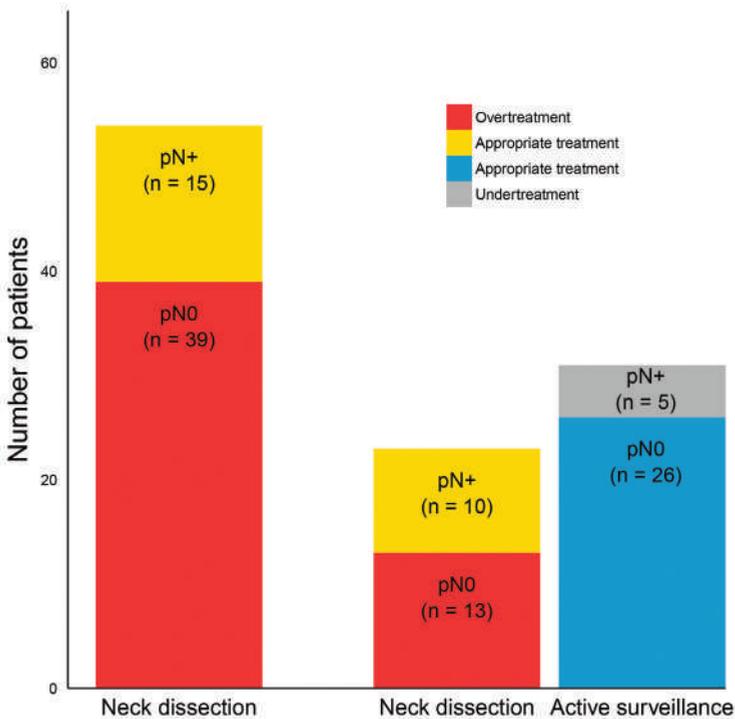
The 40 survival genes significantly discriminated between high and low risk cases (OS: iAUC=0.63, P=1.6E-3 (global test), Table 3 and Figure 3A-left; DFS: iAUC=0.65, P=6.8E-3 (global test), Table 3 and Figure 3A-right; see Supplementary Table 4 for Ridge estimates per gene). In a clinical setting the genes should add prognostic information to established parameters. Hence, the gene signature was analyzed in context of clinical and histopathological data.

Several clinical factors were associated with OS, and none with DFS. A model was trained with the most important clinical factors for this dataset and pathological TNM-stage (pTNM). The clinical factors selected and included in the model were: age at diagnosis and smoking (i.e. packyears, PY), see Supplementary Table 5 for univariable p-values. The model with these two clinical factors and pTNM accurately predicted overall survival (iAUC=0.66, Table 3), but not DFS (iAUC=0.53, Table 3). Adding OSsig to this model improved the accuracy (OS: iAUC=0.68, OSSig: P=0.03 (global test), Table 3 and Supplementary Figure 2A). For DFS, a model based on the two clinical variables + pTNM and the OSsig gave an iAUC of 0.60. Note that this is lower than a model based on the OSsig only (iAUC 0.65).

Besides pTNM, other histopathological variables are important to decide on adjuvant treatment. In the Dutch guidelines, decisive criteria for adjuvant postoperative therapy are extracapsular spread (ECS), tumor-positive margins (R+) and multiple metastatic lymph nodes (>1 LNM). We created a composite variable (pCompVar) that was scored positive if ECS or R+ or >1 LNM was present. This composite variable was combined with clinical factors (i.e. age, PY) in a prognostic model (OS: iAUC=0.73, DFS: iAUC=0.62; see Table 3). The OSsig improved the accuracy of the model (OS: iAUC=0.74, OSSig: P=0.02 (global test), Table 3 and Supplementary Figure 2B; DFS: iAUC=0.68, OSSig: P=0.01 (global test), Table 3). DFS was most accurately predicted by a model that combined the OSsig and pCompVar, not including pTNM (iAUC=0.70; OSSig: P=5.6E-3 (global test)).

A subgroup analysis was performed with patients without criteria for postoperative radiotherapy, i.e. cases that were pCompVar-negative (n=79, Figure 3B-left). For these cases a multi-type prognostic model was built that included clinical factors (age and smoking) and the OSsig. The iAUC increased from 0.70 to 0.73 by adding the prognostic genes (Table 3 and Figure 3B-right). Predictive models for DFS were less accurate in this subgroup, although a predictive model with genes only showed some predictive power (iAUC=0.65, OSSig: P=0.27) (Supplementary Figure 3, Supplementary Table 3).

These findings show that the prognostic value of the OSSig adds to established clinical and pathological prognostic variables.



**Figure 2.** Incorporation of the LNMsig in a clinical decision model that was previously proposed for patients with clinically early stage (cT1-T2N0) oral squamous cell carcinoma (OSCC).

At present, early-stage OSCCs are treated with an elective neck dissection (END, levels I-III or I-IV depending on location) in most centers. This would cause overtreatment in 39 patients (first bar, indicated in red). The clinical decision model recommends performing an END when the gene expression signature prediction is N+ or active surveillance when the prediction is N0. The hypothetical situation when using this decision model is represented in the second and third bar. Following the decision model, only 23 patients are directly treated with an elective neck dissection (second bar), overtreatment is restricted to 13 cases, and 26 patients receive appropriate treatment (third bar). The patients who are pN+ and receive an END are labeled as receiving appropriate treatment (indicated by yellow color).

### External validation of LNMsig and OSSig with TCGA RNAseq data

For additional external validation, we used RNAseq data of HPV-negative OSCC tumors from the TCGA Network publication<sup>19</sup> (n=160, Table 1). The 22-gene LNMsig was significantly associated to pathological N-stage ( $P = 7.6E-06$ , global test). Moreover, the LNMsig could accurately classify the tumors with an AUC of 0.73 (95% CI = 0.67 to 0.78). The performance of the 40-gene OSSig was also significant (iAUC=0.59,  $P=0.02$  (global test), Supplementary Figure 4). The OSSig was less informative since the average follow-up time for the 89 non-deceased cases was very short (2.2 years, SD = 2.35, Supplementary Figure 5A), and the baseline hazard was relatively high when compared to the RT-qPCR validation cohort (Supplementary Figure 5B).

Table 3. Univariable and multivariable analysis of genomic, clinical, pathological and combined models in validation cohort

	<b>Overall survival</b>		<b>Disease free survival</b>	
	<b>iAUC<sup>a</sup> (95% CI<sup>b</sup>)</b>	<b>P<sup>c</sup> value</b>	<b>iAUC<sup>a</sup> (95% CI<sup>b</sup>)</b>	<b>P<sup>c</sup> value</b>
<b>Unitype</b>				
OSsig	0.63 (0.57-0.68)	0.002	0.65 (0.60-0.70)	0.007
Clinical	0.66 (0.59-0.73)		0.54 (0.49-0.61)	
pTNM	0.51 (0.47-0.57)		0.51 (0.47-0.57)	
pCompVar <sup>d</sup>	0.64 (0.56-0.71)		0.63 (0.56-0.71)	
<b>Multitype</b>				
Clinical+pTNM	0.66 (0.60-0.73)		0.53 (0.47-0.60)	
OSsig+clinical+pTNM	0.68 (0.64-0.73)	0.03	0.60 (0.55-0.64)	0.01
Clinical+pCompVar <sup>d</sup>	0.73 (0.67-0.80)		0.62 (0.54-0.70)	
OSsig+clinical+pCompVar <sup>d</sup>	0.74 (0.69-0.79)	0.02	0.68 (0.63-0.73)	0.01
<b>pCompVar<sup>d</sup> negative subgroup</b>				
OSsig	0.71 (0.65-0.76)	0.01	0.65 (0.61-0.68)	0.28
Clinical	0.70 (0.61-0.79)		0.53 (0.43-0.68)	
OSsig+clinical	0.73 (0.68-0.78)	0.02	0.52 (0.46-0.65)	0.47

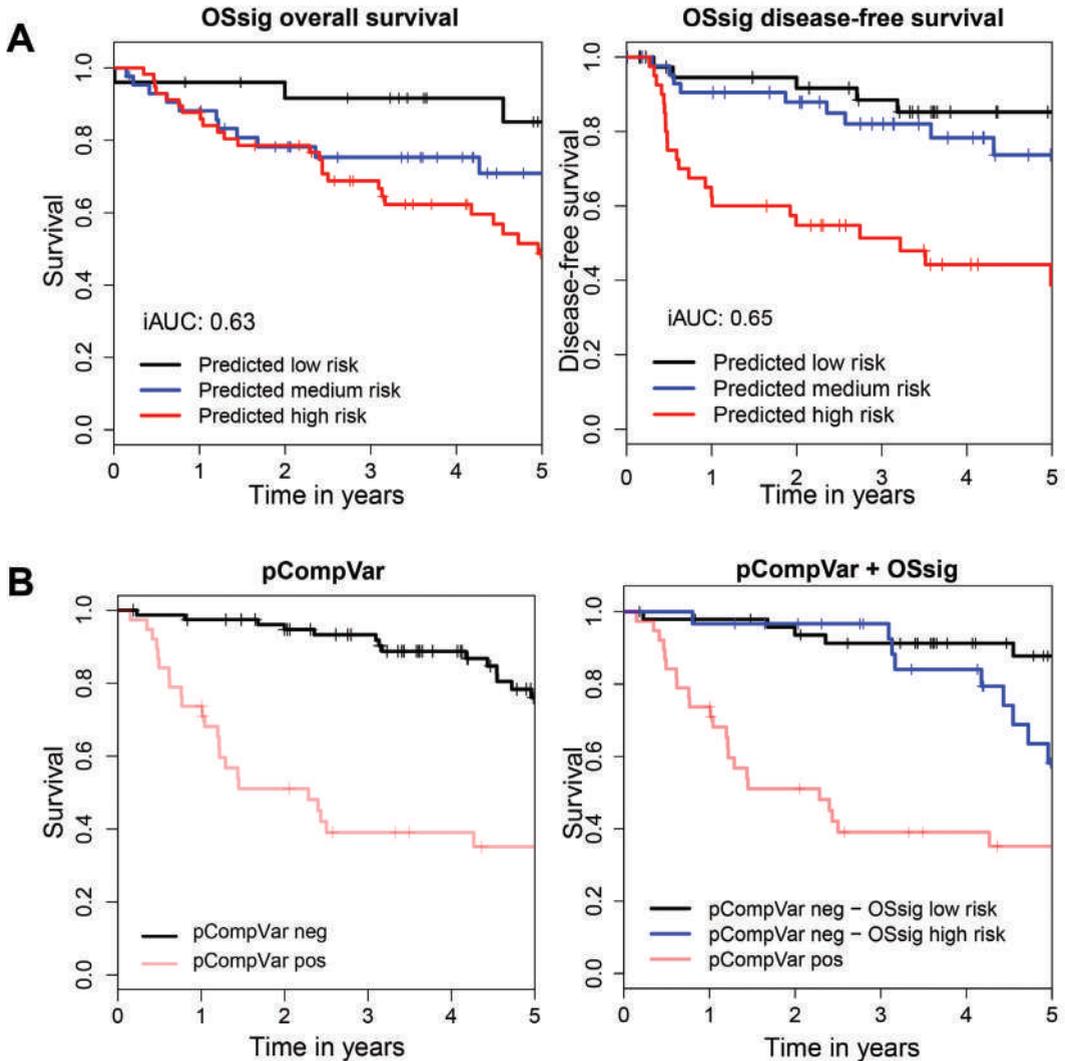
Abbreviations: iAUC, integrated Area Under the Curve; OSsig, Overall Survival signature; pCompVar, pathological composite variable; pTNM, pathological TNM stage

<sup>a</sup> Area under the curve was integrated over 5 year follow-up time.

<sup>b</sup> CIs were assessed by bootstrapping on out-of-bag samples.

<sup>c</sup> Significance of the OSsig was assessed with the global test<sup>22,23</sup>.

<sup>d</sup> Scored positive if extracapsular spread or positive resection margins or >1 lymph node metastasis was present.



3

**Figure 3.** The overall survival signature (OSsig) predicts overall survival and disease-free survival, also in low-risk patients.

(A) Kaplan-Meier analysis of overall survival (left) and disease-free survival (right) with risk groups defined by tertile predicted hazards by the OSsig analyzed with qPCR in the independent validation cohort of 125 OSCC patients. We also considered threshold optimization for creating the three groups; resulting KM curves were very similar and are hence not displayed. (B) On the left, a Kaplan-Meier analysis is shown for overall survival in the independent validation group with risk groups defined by pCompVar, which is scored positive when during histopathological examination either extracapsular spread (ECS) or involved resection margins (R+) or >1 lymph node metastasis was identified. These are routinely used histopathological criteria for adjuvant treatment. On the right the result of a subgroup analysis is shown to improve the stratification of the pCompVar-negative patients (n=79). TNM-staging was not informative to stratify this group (data not shown), but the OSsig was able to identify a subgroup of patients (blue line) with relatively poor prognosis who might have benefited from adjuvant treatment (OS: iAUC=0.71; OSsig: P=0.01 (global test)). The performances of all predicting models are listed in Table 3. Area under the curve was integrated over 5 year follow-up time. Tick marks on curves indicate censoring. iAUC, integrated Area Under the Curve; OSsig, Overall Survival signature; pCompVar, pathological composite variable.

## DISCUSSION

We identified prognostic gene expression signatures that are predictive of LNM and OS in OSCC by rigorous gene selection and validation. First, we selected 60 genes using microarray data, and these genes were validated in an independent cohort of OSCC patients by the use of RT-qPCR assays. Finally, we built 2 multivariable genomic models: a lymph node metastasis model (LNMsig) and overall survival model (OSsig) and confirmed the additive value of the gene signatures to existing and established variables.

The LNMsig with 22 genes predicted nodal metastatic disease with an NPV of 84% in clinical stages I and II. These diagnostic performance statistics are comparable to previous results using a 732-probe microarray signature<sup>17</sup>. However, the RT-qPCR approach facilitates clinical implementation considerably, because a comparable performance was achieved with less genes and a more user-friendly platform. A high NPV is necessary to identify patients who can be spared an elective neck dissection. Recent reports showed that the sentinel node biopsy (SNB), which is an alternative staging technique, is more accurate with an NPV of 95%<sup>18</sup> at comparable prevalence rates of LNM. The SNB, however, is an invasive surgical procedure with associated risks and costs, and with lower sensitivity in floor of mouth tumors<sup>27-29</sup>. Moreover, it has not been introduced widely. It has been suggested that a combination of an expression signature and SNB may be more accurate for staging of the clinically N0 neck<sup>30</sup>.

The OSsig could be used to personalize treatment. By itself, the OSsig predicted overall survival with an iAUC of 0.63, which is already promising compared to the iAUC of 0.51 of standard pTNM. For prediction of DFS, the OSsig was even more valuable, particularly when combined with histopathology, as clinical variables were not informative for DFS. These data confirm the predictive value of the OSsig, but also indicate that integrating clinical, molecular and histopathological variables delivers most accurate predictive models.

The design of this study enabled the identification of robust associations in three ways. First, we used different gene expression platforms to cancel out platform-specific findings. Second, we studied homogeneous patient cohorts: only HPV-negative, surgically treated OSCCs were included. Finally, we considered patients from 3 European countries, thereby excluding the discovery of population-specific gene signatures.

Our findings may be limited by two factors. First, intra-tumor heterogeneity may cause differences in gene expression profiles within a tumor; although previous findings suggest that expression profiles seem stable in HNSCC<sup>31</sup>. Second, all cohorts investigated were retrospective. It should be mentioned, however, that retrospective data obtained in The Netherlands are generally accurate, because treatment and follow-up of HNSCC patients has been centralized to a few clinical centers and clinical management adheres to standardized national guidelines.

Our findings suggest at least two implications. First, the prognostic model may be used for treatment escalation in patients with tumors that do not fulfill the current criteria for postoperative radiotherapy, i.e. margin involvement, >1 metastatic lymph node or ECS. Second, a model that integrates clinical variables and the OSsig accurately predicts prognosis without the addition of histopathology. This model may specifically be important to predict survival in patients who are treated with primary radiotherapy or chemoradiotherapy, since histopathology is not available for these patients. These are important directions for future work. Since frozen material is not always available in these cases, future research should also include applications for FFPE tissue. Ultimately, prospective clinical trials will be required to determine whether the integrated risk models could guide clinical decision making and improve treatment results with respect to outcome and morbidity.

## ACKNOWLEDGEMENTS

We thank V. Okpanyi and L. Colter for sample preparation and assembly of clinical data; F. Rustenburg, P. Eijk, S. Smeets and B. Ylstra for processing of the VUmc microarrays; Agendia for making the UMCU data available; M.

Doeleman and D. Heideman for the E6\*I RT-PCR HPV testing; and A. Visser and C. Oudejans for their assistance with the qPCR microfluidic cards.

## **CONFLICTS OF INTEREST**

The authors have declared that no conflicts of interest exists.

## **FUNDING**

This work was supported by grants from the Dutch Cancer Society, KWF (grant number 2008-4201) and European Union's Seventh Framework Project (grant agreement 224483: NeoMark, and grant agreement 611425: OraMod).

## SUPPLEMENTARY METHODS

### Patients

Four independent cohorts of oral squamous cell carcinoma (OSCC) patients with frozen biopsies were included in this study: (1) a cohort of 150 patients from the University Medical Center Utrecht (UMCU) and VU University Medical Center Amsterdam (VUMC) for microarray gene expression profiling (array cohort 1, AC1); (2) a cohort of 99 patients from the University Hospital Parma Medical Center (UHPMC) for microarray gene expression profiling (array cohort 2, AC2); (3) a cohort of 125 patients from VUMC, UHPMC and University Hospital Düsseldorf (UHD) for qPCR gene expression profiling; and (4) an RNAseq dataset of HPV-negative OSCC tumors from the Cancer Genome Atlas Network publication<sup>19</sup>. Inclusion criteria were: presentation with a squamous cell carcinoma in the oral cavity, date of incidence prior to July 1st 2012, and treated surgically with curative intent. All patients were 18 years or older and had no previous malignancy that impacts outcome. Included ICD-10 codes were: C00.3-4, C02.0-3, C02.8-9, C03, C04, C05.0, and C06. Human papilloma virus (HPV) positive tumors were excluded from further analysis. Informed consent was obtained of enrolled patients when required, and nation- and institution-specific procedures and guidelines were followed in addition. For instance, for the Netherlands use of residual tissue from surgical specimen adhered to the Code of conduct for responsible use by the Federation of Dutch Medical Scientific Societies (FDMSS). This study followed the Guidelines for the REporting of tumor MARKer Studies (REMARK)<sup>20</sup> (Supplementary Table 6).

### Specimen

Biopsies were taken from the surgical specimen at time of surgery, snap frozen and subsequently stored in liquid nitrogen. Five to ten 20 µm sections were used for RNA isolation. Before and after sampling of the 20 µm sections, 5 µm sections were made and stained by haematoxylin and eosin to ensure that at least 50% tumor cells were present in the biopsy. RNA isolation was performed using TRIzol (cat. 15596026, Life Technologies, Breda, The Netherlands; AC1) or with columns using the RNeasy Mini Kit (cat. 74104, Qiagen, Hilden, Germany; AC2 and qPCR cohort), according to the protocol of the suppliers. HPV status was either determined with p16 immunostaining followed by HPV DNA PCR on p16-positive samples (AC1) or with HPV16 E6\*I RT-PCR (qPCR cohort) in the AC1 and AC3 cohorts. Both assays have been validated and described before<sup>21</sup>. In AC2, the HPV status was not determined. In the other cohorts on the other hand, 1 out of 151 (AC1) and 1 out of 126 (qPCR cohort) tumors were HPV-positive. Hence, the contribution of HPV positive tumors in AC2 was assumed low and no further samples were excluded. Quantity and quality of the RNA was tested with the Nanodrop (cat. ND-1000, Thermo Fisher Scientific, Amsterdam, The Netherlands) and the Bioanalyzer 2100 (cat. G2939AA, Agilent Technologies, Amstelveen, The Netherlands) using the RNA Nanokit (cat. 5067-1511, Agilent). RNA Integrity Numbers (RIN-value) were between 6.3 and 10.0.

### Expression arrays

Two independent cohorts of OSCC patients were processed for gene expression microarray analysis: (1) a cohort of 2 merged tumor gene expression profiles (array cohort 1, AC1); and (2) array cohort 2 from the University Hospital Parma Medical Center (UHPMC). In AC1, array hybridization was performed, using 0.5 µg total RNA in the Low RNA Input Linear Amplification Kit (cat. 5184-3523, Agilent) and the 4x44K Whole Human Genome Arrays, according to the manufacturer (Agilent) using dual color labeling. The handling of the UMCU samples and their RNA isolation has been published before<sup>17</sup>. Microarray data of UMCU were retrieved from the Gene Expression Omnibus (GSE30788). Additional information on these samples was obtained from Agendia (Amsterdam, the Netherlands). In AC2, 0.2 µg of total RNA was labeled and simultaneously amplified following the "Two-Color Microarray-Based Gene Expression Analysis (Quick Amp Labeling) Protocol" (Agilent Technologies). Labeled-amplified RNA samples were then hybridized on 4X44K Whole Human Genome

DNA microarray slides (cat. G4112F, Agilent) according to the instructions of the manufacturer. An Agilent Technologies Scanner G2505B US45102976 and the Feature Extraction (FE) Software v 9.5.1.1 with the GE2-v5\_95\_Feb07 Gene Expression protocol were used to scan microarray slides and extract data, respectively.

### Preprocessing microarray data

AC1 gene expression data generated at VUmc was preprocessed in an identical way as those previously preprocessed at UMCU<sup>17</sup>. This comprised: 1) extraction of the median signal from the raw data files without background correction, and 2) median and loess within-array normalization as implemented in the Limma-package (<http://www.bioconductor.org>). Preprocessed VUmc and UMCU data sets were combined by limiting both to the probes that overlapped (using their Agilent probe identifiers). Finally, comparability of the expression data of both centers was ensured by a) joint between-array normalization (Aquantile as implemented in the Limma-package), and b) removal of possible batch (i.e. center) effects using the Combat-package<sup>32</sup>. Raw and processed data are publicly available in the gene expression omnibus (GEO) database (GSE85446).

Data from AC2 were not combined to the other datasets, because of a different reference design: Universal Human Reference RNA (cat. 740000, Agilent Technologies, Santa Clara, CA) in AC1 and a pool of cell line RNA in AC2 (CAL 27, ATCC CRL-2095, American Type Culture Collection, Manassas, VA). AC2 data were preprocessed in the same way as data set AC1. Probes with more than 20% missing values were deleted. AC2 consisted of 106 samples, but seven were excluded because of poor quality MA plots. Remaining missing values were imputed with nearest neighbor averaging with R package *impute*. For the LNM analysis, two additional patients were excluded because information on the LNM was missing. Raw and processed data are registered in the gene expression omnibus (GEO) database (GSE84846), but are not publicly available until July 25<sup>th</sup>, 2018 or after publication of this manuscript.

### Gene selection

Per data set, the predictive significance for LNM and survival was assessed with the global test<sup>24,25</sup>. Data sets with significant results ( $p < 0.05$ ) were used for gene selection. For prediction of LNM, we obtained significant results for both AC1 and AC2 (AC1,  $p=9.3E-06$ ; AC2,  $p=9.9E-03$ ), hence both were used for gene selection. For prediction of survival, only AC1 showed predictive significance (AC1,  $p=7.8E-3$ ; AC2,  $p=0.73$ ) and was used for further analysis. This difference is likely explained by a shorter follow-up time in AC2 compared to AC1 (AC1, mean overall follow-up time: AC1: 4.7 years (SD=3.2), AC2: 3.0 years (SD=1.5)).

The selection of genes was based on univariable and multivariable analysis with an equal contribution to the final signature (50% of genes from univariable analysis, 50% from multivariable analysis). For the multivariable selection our aim was to come to an optimal set of genes that orthogonally contribute on top of each other (see Supplementary Example R code below). Due to selection from a large number of genes using limited sample series, it is to be expected that some genes will not validate. Also, the platform transition we conducted increases the probability that genes cannot be validated, although this is somewhat counterbalanced by the technical validation. For this reason we planned some redundancy in the selected genes. This is best achieved by selecting those genes that have the highest signal to noise ratio (e.g. low  $p$ -value). The total number of selected genes was limited to 60 genes of interest because of the chosen qPCR array card design (60 target genes + 4 housekeeping genes, 3 replicates, and 2 samples / array card). Univariable analysis included  $t$ -tests (LNM) and Cox regression (overall survival). For LNM  $p$ -values of AC1 and AC2 were combined with Fisher's combined  $p$ -value.  $P$ -values were adjusted with the Benjamini-Hochberg procedure to control the false discovery rates (FDR)<sup>26</sup>. Multivariable analysis consisted of lasso logistic regression (LNM) and lasso Cox regression (overall survival), as implemented in package *glmnet* in R. To stabilize the selection, the lasso

was run repeated N times per analysis (as a leave-one-out cross-validation (LOOCV)). For LNM, the selection frequencies of AC1 and AC2 were added up. We next selected those genes that were used most often in the models. For survival, models were tested with and without addition of pathological stage of disease (pTNM) and age as unpenalized covariate, but the additions did not change the list of selected genes considerably.

To reduce dimensionality and enrich for relevant predicting genes, previously published HNSCC gene signatures were used as input for gene selection. For LNM, we used a previously published LNM predicting gene profile<sup>15,16</sup> consisting of 732 probes, which was later validated in a multicenter trial<sup>17</sup>. For survival, thoroughly validated prognostic gene signatures were missing. Therefore, we combined a set of 9 prognostic gene expression profiles<sup>10-12,33-37</sup> (1,426 probes) and an in-house discovered prognostic profile of genes for which copy number alterations and gene expression were best correlated (348 probes). This combined survival profile consisted of 1,762 probes. Twenty genes were selected from the combined survival profile (1,762 probes), and 20 genes were selected from analyses that included all probes (37,622 probes in AC1).

### **Quantitative real-time PCR**

A 384-well Taqman Low-Density Array (TLDA) Card was designed with the selected 60 prognostic genes + 4 housekeeping genes (GAPDH, GUSB, RPL4, RPLP0). Each gene expression assay was represented by 3 replicates. The initial design (TLDA.v1) and optimized design (TLDA.v2) after technical validation (see below) are shown in Supplementary Table 2. One  $\mu\text{g}$  of mRNA was treated with DNase I, Amplification Grade (cat. 18068015, Invitrogen; Carlsbad, CA) in a 10  $\mu\text{l}$  reaction volume. The DNase-treated mRNA was subsequently used for cDNA synthesis with the High-Capacity RNA-to-cDNA Kit (cat. 4387406, Applied Biosystems; Foster City, CA) in a 24.4  $\mu\text{l}$  reaction volume. The qPCR reaction mix consisted of (1) 20  $\mu\text{l}$  cDNA (819  $\mu\text{g}$ ), (2) 190  $\mu\text{l}$  water, and (3) 210  $\mu\text{l}$  2X TaqMan Gene Expression Master Mix (cat. 4369016, Applied Biosystems). Subsequently, the reaction mix was loaded on the TLDA cards according to the protocol of the supplier. Reaction mixes of 2 samples were loaded per TLDA card. Experiments were performed on an ABI Prism 7900HT Fast Real-Time PCR System (cat. 4329001, Applied Biosystems). Thermal cycling conditions were: 50°C for 2 minutes; 94.5°C for 10 minutes; 40 cycles of denaturation at 97°C for 30 seconds and annealing and extension at 59.7°C for 1 minute. The median result of a triplicate assay was used in downstream analysis.

### **Gene expression analysis from qPCR data**

For each qPCR reaction, the Ct-value was determined as the cycle number at which the fluorescence signal reached a fixed threshold using the SDS RQ Manager Version 1.2.2 (Applied Biosystems). Next, the Ct-values were normalized to the GUSB expression level, which was the most stable housekeeping gene in this dataset. To select the most stable housekeeping gene, we determined the standard deviation of the gene expression within all samples and correlation of the housekeeping gene to the average gene expression of all target genes (Supplementary Table 7). GUSB had both the lowest SD (SD=0.94) and the highest correlation to the average expression of all target genes ( $r=0.77$ ).

### **Technical qPCR validation**

The 60 selected target genes were technically validated using a subset of 20 cases from AC2 to evaluate the platform transition. Hence, the qPCR data were correlated to the array data of the same samples. These 20 cases were selected to contain a 1:1 ratio of N0/N+ cases and patients that survived/deceased. For these 20 cases, Pearson's correlation coefficients were calculated between microarray and corresponding qPCR data (Supplementary Table 3). Poor correlation was defined as an  $r$  of  $>1$  SD below the mean and a  $p$ -value  $>0.1$  (because of the small sample size). Logistic and Cox regression analyses were performed to determine the predictive performance of the genes in this cohort.

## Histopathology

Formalin-fixed paraffin embedded slides of the surgical specimens were examined by two specialized pathologists (EB + EMS), according to the guidelines of the Royal College of Pathologists UK (<https://www.rcpath.org/resourceLibrary/dataset-for-histopathology-reporting-of-mucosal-malignancies-of-the-oral-cavity.html>). Tumors were staged according to TNM classification of Malignant Tumors, 7th Edition, published in affiliation of the Union for International Cancer Control (UICC)<sup>38</sup>. The margin status was evaluated and divided into three groups: (1) involved margins when carcinoma was present in or within 1 mm of the margin, (2) negative margins when the excised carcinoma was > 5 mm from the surgical margin, and (3) close margins when the tumor was 1-5 mm from the surgical margin<sup>39</sup>. For model building, the margin status was subsequently subdivided in two groups: involved margins when carcinoma was present in or within 1 mm of the margin (R+); or negative margins, when the excised carcinoma was > 1 mm from the surgical margin (R0). The presence of lymph node metastasis (LNM) was determined by standard histopathological examination of the neck dissection specimen if present. When the neck was left untreated, two scenarios were possible: a patient was diagnosed N+ when a delayed lymph node metastasis developed during follow-up ( $\leq$  three years after treatment) or remained N0 when not. In the different cohorts, 128 of 150 (85.3%, AC1), 89 of 99 (89.9%, AC2) and 103 of 125 (82.4%, qPCR cohort) the neck was treated with a primary neck dissection. Extracapsular spread (ECS) was present if the tumor extended beyond the capsule of the lymph node. When there was doubt, the case was classified as having ECS according to the guidelines of the Royal College of Pathologists UK<sup>40</sup>. We created a pathological composite variable (pCompVar) that was scored positive if ECS or R+ or >1 LNM was present.

## Clinical data

Several clinical variables were used for prognostic model building. These included age at diagnosis, gender, smoking behavior in packyears (1 packyear equals 20 cigarettes a day during 1 year), ECOG Performance Status<sup>41</sup>, and comorbidity. Comorbidity was classified using the Adult Comorbidity Evaluation 27 (ACE-27)<sup>42</sup>, in which an overall comorbid score is graded in four levels: none, mild, moderate or severe. For smoking we only considered the packyears in the model building. Compared to categorical smoking variables, packyears contains the most information about the smoking habits and was also the most significant smoking variable for OS.

## Outcomes

Overall survival (OS) was defined as time from date of incidence to death from any cause. Disease free survival (DFS) was defined as time from date of incidence to development of locoregional recurrence, distant metastasis or second primary HNSCC. Mean survival times for the various data sets were calculated with the reverse Kaplan-Meier as suggested by Schemper et al.<sup>43</sup> Patients who died of other causes or develop second primary tumors outside the head and neck region (SPT), were censored on the date of death or incidence date of the SPT. Local recurrences were scored when these developed within two centimeters of the index tumor and within three years after therapy, whereas a regional recurrence was documented when it developed in a treated neck within three years after treatment.

## Statistical analysis of the RT-qPCR dataset

For the qPCR data, the univariable association of delta Ct values of the selected genes with either LNM or OS/DFS was determined with logistic or Cox regression, respectively. Multivariable models with the selected genes were made with logistic ridge regression (LNM) or Cox ridge regression (OS/DFS). For the clinical variables univariable p-values of clinical variables were determined by Cox proportional hazards regression (OS, DFS)

or logistic regression (LNM). Patients with moderate and severe comorbidity (ACE-27) were considered as one group in the analysis, because the group with severe comorbidity (ACE-27) was very small. TNM-stage was dichotomized as early stage disease (pTNM I+II) and advanced stage disease (pTNM III+IV). Variables with p-value lower than 0.15 were considered as candidates for a multivariable model. For LNM prediction, no clinical variables met this criterion. Next, stepwise regression was performed to identify a multivariable model with clinical variables (using procedure 'step' in R). Stepwise selection with Akaike Information Criterion (AIC) was performed and in each step a variable was added or dropped, which identified the best model. For OS, the stepwise selection procedure selects age and packyears. The prediction models for outcome consisted of (1) prognostic genes only, (2) clinical variables and pathological TNM-stage (pTNM), (3) clinical variables and a composite pathological variable (positive if ECS or R+ or >1 LNM was present), and the combinations (4) 1+2 and (5) 1+3. In combined clinical and genomic models, the clinical variables were not penalized and the genes were incorporated with a ridge penalty to avoid overfitting. The predictive performance was measured by area-under-the-ROC-curve (AUC) and integrated AUC (iAUC)<sup>44</sup> at 5-year follow-up time for LNM and OS/DFS, respectively, complemented for LNM by the NPV, i.e. the proportion of true negatives among all negative tests. Model performance was assessed by bootstrapping, confidence intervals around the AUC, sensitivity, specificity, PPV, and NPV were calculated according to methods described by Jiang et al.<sup>45-47</sup> (see Supplementary Outline of Statistics and Supplementary Example R code). Model performance only takes into account the uncertainty in the genomic, clinical, and pathological coefficients. The variable selection of clinical variables was not bootstrapped, the pathological variables (pTNM and CompVar) were selected based on their known clinical relevance. For the genomics variables we did not perform further selection on the RT-qPCR data. For OS and RFS the subgroup analysis was performed by refitting the model to the cases of that subgroup. For LNM subgroup analysis, refitting was not possible due to the small sample size of the subgroup (n=54) with a low number of cases (15). Therefore, the AUC for the subgroup was computed by first fitting the model to all cases (n=125), and then considering the subgroup. Additive value of the gene signature was assessed with the global test<sup>24,25</sup>. All statistical tests performed were two-sided. In multiple testing settings, univariable p-values were corrected using the Benjamini-Hochberg FDR procedure<sup>26</sup>.

### External validation with TCGA RNAseq dataset

Only the 279 patients that were included in the Cancer Genome Atlas Network publication<sup>19</sup> were used for this analysis, because the RNAseq derived HPV-status, which was considered the most accurate, was not available for the other cases. Of these cases, the normalised RNASeqv2 TCGA data were downloaded with R package TCGA2STAT. Additional clinical information was downloaded directly from the Broad Institute ([http://gdac.broadinstitute.org/runs/awg\\_hnsc\\_\\_2013\\_03\\_30/reports/cancer/HNSC-TP-HPV-positive-36/mutsignozzlerereportscv/nozzle.html](http://gdac.broadinstitute.org/runs/awg_hnsc__2013_03_30/reports/cancer/HNSC-TP-HPV-positive-36/mutsignozzlerereportscv/nozzle.html)). Only HPV-negative tumors of the oral cavity were considered (n=160). For lymph node metastasis, patients with a pathological NX-stage were excluded. This left 133 patients for lymph node metastases (LNM). For OS, the survival time of one patient was missing and 159 patients were available for analysis. The LNM outcome was defined as having a pathological N-stage larger than 0. Of the 60 genes selected on the microarray data, 1 gene could not be matched to the TCGA data (LRCOL1).

Before analysis the data were transformed by taking the square root and scaling the data (e.g. all genes transformed to zero mean and unit variance). For OS and for LNM, the global test was used to assess the association with the genomic signatures. Further predictive performance was assessed by fitting and bootstrapping a logistic (for LNM) and Cox (for OS) regression model with ridge penalization.

### Sample size considerations

Exact sample size calculations are inherently difficult for ridge regression. They also require knowledge of the effect sizes, which are unknown in our multivariate setting. To provide some insights regarding correct

sample size, we performed some ad hoc tests with the qPCR cohort. Random samples of 45-115 patients were repeatedly drawn (50 times) without replacement from the data, supplemented with the actual data for  $n=125$ . For each outcome variable, the global test was used to assess the predictive performance of the genomic signatures. For each same sample size we took the medium p-value across the random samples. Sample sizes of 85 cases and more gave consistent p-values  $<0.05$  for each outcome variable (see Supplementary Figure 6), assuring that the sample size of 125 is sufficient to assess the performance of the signatures.

### **Relation of the genomic signatures to other prognostic markers**

We additionally analyzed the association between the prognostic markers and the OSsig (here taken as the linear predictor of the Cox ridge regression, i.e. the log of the hazard ratios), as recommended in REMARK criterion 14 (Supplementary Table 6). In this analysis we considered the variables used in the prognostic models (i.e. age, smoking, pTNM, pCompVar) and additionally considered gender and comorbidity (ACE27). For the age and packyears, a Pearson's correlation coefficient and an additional p-value were calculated. For sex, pTNM, and pCompVar we performed a t-test and for ACE27 we performed an ANOVA. Significant correlation between OSsig and gender, smoking, ACE27 and pCompVar were found (Supplementary Figure 7). However, as can be seen in this figure, the size of the effects was small. Secondly, we assessed the iAUC of the genomic predictor in different subgroups (Supplementary Table 8). Subgroups included were: (1) pCompVar negative and positive; (2) Age  $<70$  and  $\geq 70$  years at diagnosis; (3) Smoking: packyears  $<$ median and  $\geq$ median value in this dataset; (4) Comorbidity (ACE27): ACE27 0-1 and ACE27 2-3; (5) Male and female gender; and (6) pathological stage I+II and III+IV. This analysis showed that the OSsig has good discriminative power in the various subgroups.

## SUPPLEMENTARY OUTLINE OF STATISTICS

1. Gene selection with microarray data.
  - a. Univariable selection
    - i. p-value per gene with false discovery rate (FDR) control.
      1. Welch t-test for LNM
      2. Cox regression for OS
  - b. Multivariable selection with lasso (see example code)
    - i. Repeatedly fitting the lasso. Genes are ranked based on their selection frequency across the lasso fits (see example code of the lasso selection).
    - ii. Analysis conducted with and without mandatory covariates. Mandatory covariates were not subject to the lasso penalty (e.g. unpenalized).
    - iii. Logistic regression for LNM, Cox regression for OS
2. Technical validation
  - a. Pearson's correlation coefficient (and associated p-value) between qPCR – microarray data per gene.
3. Validate selected genes on independent data (i.e. qPCR data).
  - a. Univariable assessment
    - i. p-value per gene with false discovery correction (FDR) control.
      1. Logistic regression for LNM
      2. Cox regression for OS/RFS
4. Fit with selected genes on the (independent) qPCR data.
  - a. Low dimensional models (i.e. clinical and/or pathological) fitted with standard regression techniques.
    - i. Clinical variable are selected on the qPCR data for OS. First we conducted a univariate screening with p-value of 0.15. Next the clinical model was made by stepwise regression based on AIC.
  - b. Genomics models fitted with ridge regression
  - c. Combined models (e.g. clinical, and/or pathological and/or genomics) for RFS/OS fitted with ridge regression. Clinical and/or pathological are not subject to the ridge penalty.
5. Predictive accuracy of the models is assessed by bootstrapping.
  - a. Bootstrap assessed the parameter uncertainty.
  - b. For the clinical data the variable selection process was not bootstrapped.
  - c. For genomics and pathological variables there was no further variable selection.
  - d. For bootstrap code see <http://github.com/DennisBeest/BootPredError>.
6. Step 4b and step 5 are repeated with the (independent) TCGA data.

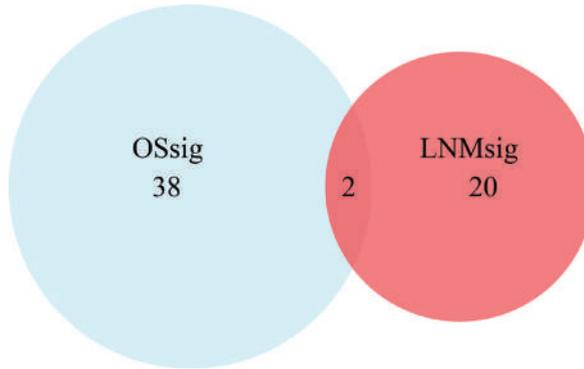
## SUPPLEMENTARY EXAMPLE R CODE OF THE MULTIVARIABLE GENE SELECTION (LASSO)

```
#####
#---Example code
library(glmnet)
#####
#---Simulated some survival data, replace these by real data
#####
P <- 200
N <- 100
Train <- matrix(nrow=N,ncol=P)
Train[] <- runif(P*N)
Survival <- rep(c(0,1),each=N/2)
Train[Survival==1,1:5] <- Train[Survival==1,5]+0.5
Time <- runif(N)
Time[Survival==1] <- Time[Survival==1]/2
colnames(Train) <- 1:P
penfac <- rep(1,times=P)
#####
#---Selection with lasso in the form of a LOOCV.
#---Alternatively the lasso can be run repeatedly on the whole data, or on bootstraps.
#---The lasso is likely to select a different set of variables each time it is run on slightly different data or when
the cross-validation folds are changed. Especially when the the data are stongly correlated.The aim of running
it repeatedly is to stabilise the selection.
#---An unpenalized covariate can be incorporated by setting penfac to 0 for that variable.
#####
N <- length(Survival)      #Number of patients
Selected <- list()         #List for saving selected genes
SavePrediction <- numeric() #Optionally for saving cross-validated predictions
for(i in 1:N)              #Repeat for each patient
{
  #####
  #Leave one patient out
  Y <- cbind(time=Time[-i],status=Survival[-i])
  X <- Train[-i,]
  #####
  #Fit lasso
  model <- cv.glmnet(X,Y,family="cox",standardize = FALSE,alpha = 1,nfolds=5,penalty.factor=penfac)
```

```

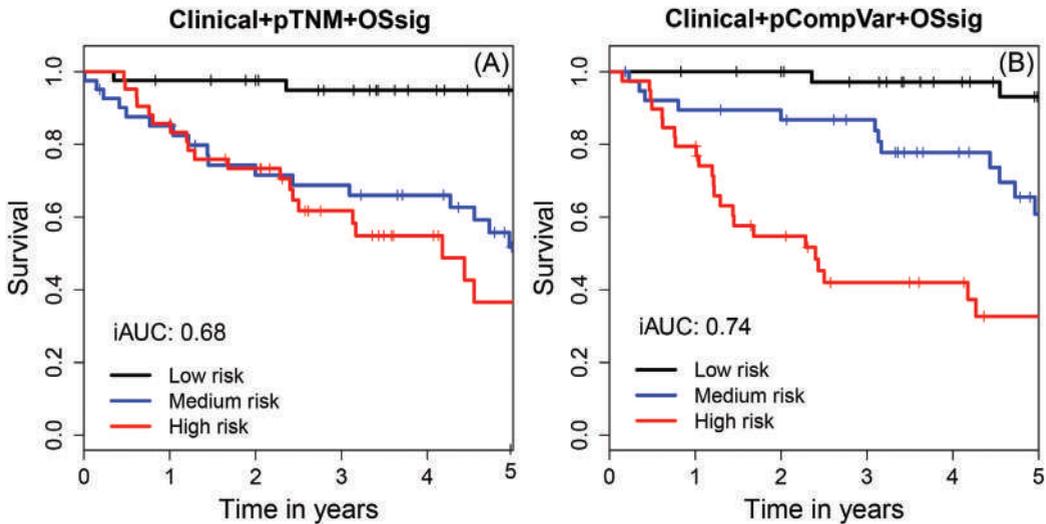
#####
#Keep track of selected variables
betas <- model$glmnet.fit$beta[,model$glmnet.fit$lambda==model$lambda.min]
Selected[[i]] <- names(betas)[which(betas!=0)]
#####
#Optionally get an indication of the predictive value
XTest <- Train[i,,drop=FALSE]
SavePrediction[i] <- predict(model,newx=XTest,s=c("lambda.min"))
}
#####
#----Add selected variables together and sort/rank
tab <- table(unlist(Selected))
o <- order(tab,decreasing=TRUE)
print(cbind(tab[o]))
#####
#----End of code

```



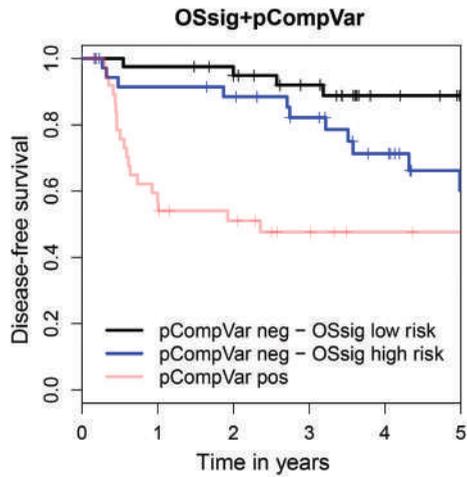
3

**Supplementary Figure 1.** Venn diagram of selected gene signatures shows 2 overlapping genes between OSsig and LNMsig. Venn diagram of selected overall survival gene signature (OSsig, 40 genes) and lymph node metastasis gene signature (LNMsig, 22 genes) shows that 2 genes overlap, limiting the combined signatures to 60 genes. LNMsig, lymph node metastasis signature; OSsig, overall survival signature.



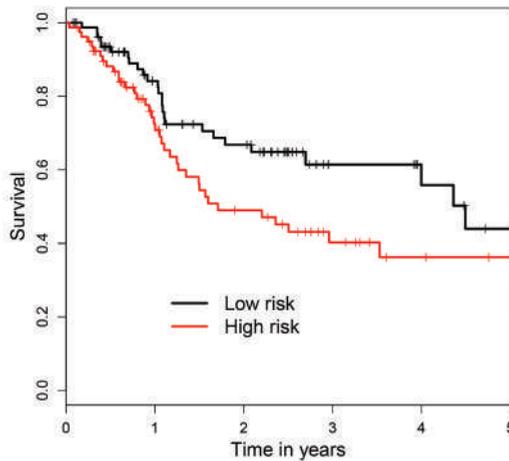
**Supplementary Figure 2.** Best predicting integrated models of clinical variables, histopathological variables and the overall survival signature (OSsig).

(A-B) Results of the Kaplan-Meier analysis are depicted for overall survival in the independent validation cohort using the best predicting models with risk groups defined by tertile predicted hazards of (A) the OSsig combined with significant clinical variables (i.e. age at diagnosis, smoking) and pathological TNM stage (pTNM) (iAUC=0.68, OSsig:  $P=0.03$  (global test)), and (B) the OSsig combined with the same clinical variables and the a composite histopathology variable (pCompVar) that was scored positive if extracapsular spread (ECS) or involved resection margins (R+) or >1 lymph node metastasis was present; all three variables are currently used as indicators for adjuvant treatment (iAUC=0.74, OSsig:  $P=0.02$  (global test)). Area under the curve was integrated over 5 year follow-up time. Tick marks on curves indicate censoring. iAUC, integrated Area Under the Curve; OSsig, overall survival signature; pCompVar, pathological composite variable; pTNM, pathological TNM stage.



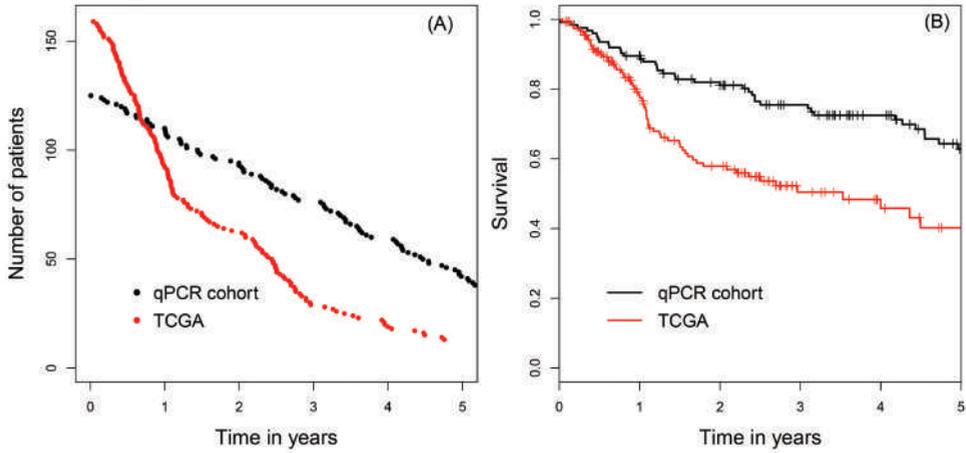
**Supplementary Figure 3.** The Overall Survival genomic signature (OSSig) predicted disease-free survival of oral squamous cell carcinoma (OSCC) in a subgroup of prognostically favorable patients.

A subgroup of 79 prognostically favorable patients was identified based on histopathological variables, i.e. tumor-free margins (R0),  $\leq 1$  lymph node metastasis (LNM), and without extracapsular spread (ECS-neg). Depicted is a Kaplan-Meier analysis for disease-free survival in these pCompVar-negative patients of the independent validation group with risk groups defined by median predicted hazards of the OSSig (black and blue lines; integrated area under the curve (iAUC)=0.65, OSSig:  $P=7E-3$  (global test)). Area under the curve was integrated over 5 year follow-up time. Abbreviations: iAUC, integrated Area Under the Curve; OSSig, Overall Survival signature; pCompVar, pathological composite variable.



**Supplementary Figure 4.** Additional external validation of the overall survival signature (OSSig).

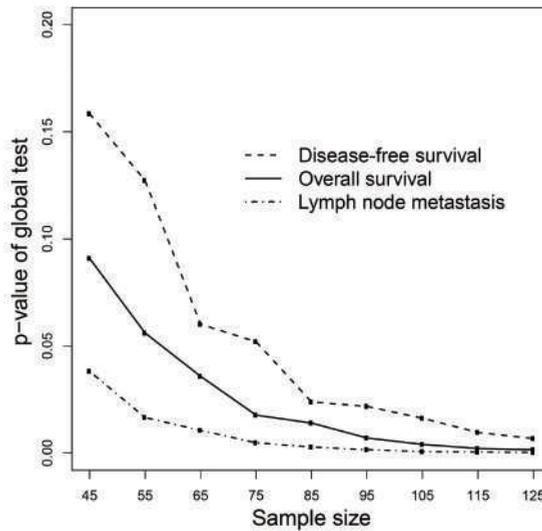
Kaplan-Meier analysis of overall survival with risk groups defined by median predicted hazards by the OSSig. RNAseq data of the TCGA cohort head and neck squamous cell carcinoma cohort were used of HPV-negative, OSCC patients ( $n=159$ ; iAUC=0.59; OSSig:  $P=0.02$  (global test)). Area under the curve was integrated over 5 year follow-up time. Tick marks on curves indicate censoring. Abbreviations: iAUC, integrated Area Under the Curve; OSSig, Overall Survival signature.



3

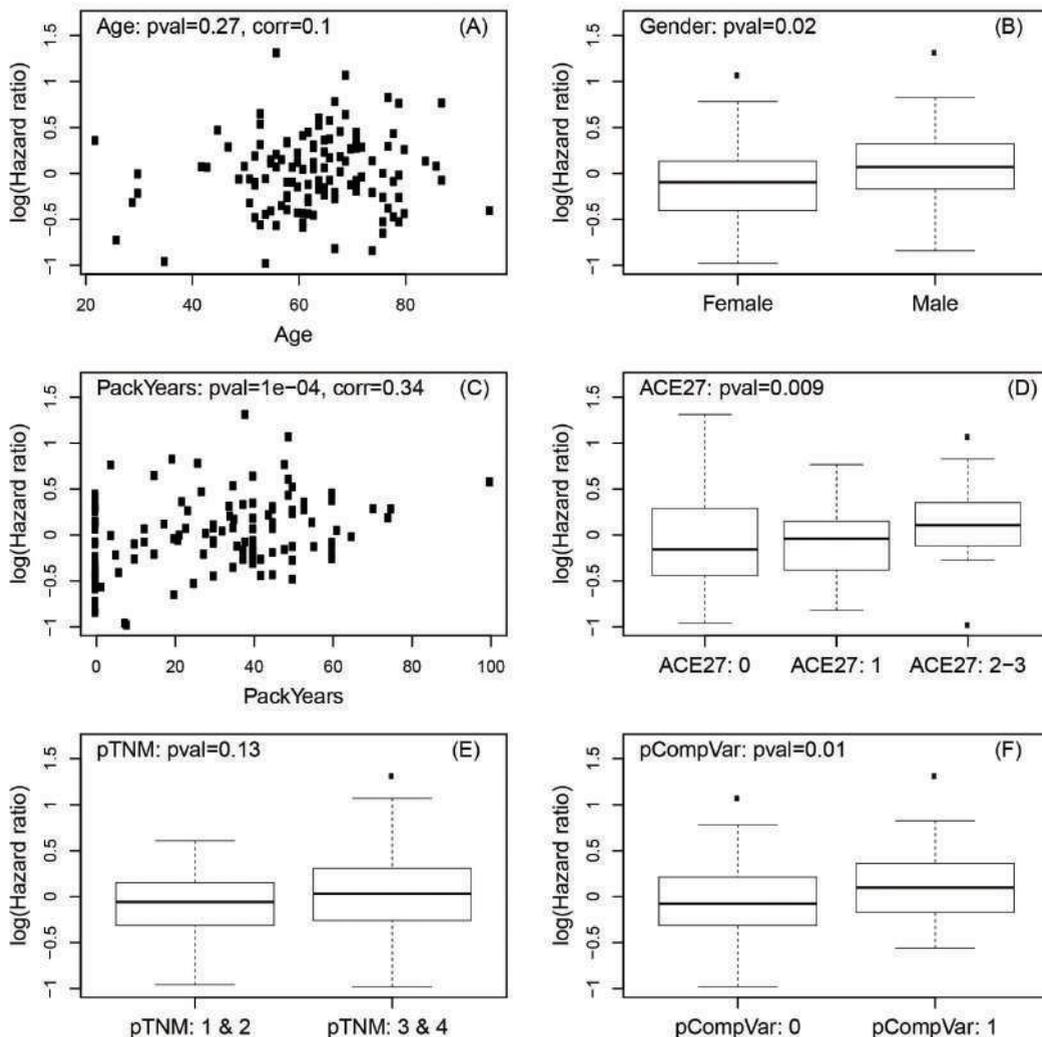
**Supplementary Figure 5.** Comparison of follow-up time and baseline survival curves shows significant differences between our qPCR validation cohort and the TCGA cohort.

**(A)** Number of patients under consideration (y-axis) in relation to follow-up time (x-axis) of qPCR cohort (black) and TCGA cohort (red). **(B)** Baseline Kaplan-Meier analysis of qPCR cohort (black) and TCGA cohort (red) differed significantly (cox regression: HR= 2.0, 95% CI =1.4–2.9),  $P=3E-4$ ).



**Supplementary Figure 6.** Random sampling from independent qPCR cohort shows clear relationship between sample size and global test p-value, and provides rationale for the tested sample size.

Random samples of increasing sample size were repeatedly (50 times) drawn without replacement from the data (y-axis). For each outcome variable, the global test was used to assess the predictive performance of the genomic signatures. P-values were averaged between random samples of the same size (x-axis). Sample sizes of 85 cases and more gave consistent p-values  $<0.05$  for each outcome variable, assuring that the sample size of 125 cases of the independent validation cohort should be sufficient to assess the performance of the signatures.



**Supplementary Figure 7.** Relation of Overall Survival signature (OSsig) to other prognostic markers.

(A-F) Relation of the OSsig linear predictor of the Cox ridge regression, i.e. the log of the hazard ratios (y-axis), and other prognostic variables (x-axis). A Pearson's correlation coefficient and an additional p-value were calculated for numerical variables, for ACE27 we performed an ANOVA, and for the remaining categorical variables, a t-test was performed. Relation between OSsig and (A) age (in years,  $r=0.1$ ,  $P=0.27$ ), (B) gender ( $P=0.02$ ), (C) smoking (in packyears;  $r=0.34$ ,  $P=0.0001$ ), (D) comorbidity (ACE27;  $P=0.009$ ), (E) pathological TNM stage (pTNM;  $P=0.13$ ), and (F) a composite pathology variable (pCompVar) that was positive if extracapsular spread or tumor-positive resection margins or  $>1$  lymph node metastasis was present ( $P=0.01$ ). For each marker the size of the effects was small, even when a significant relation was found. corr, Pearson's correlation coefficient; pCompVar, pathological composite variable; pTNM, pathological TNM stage.

**Supplementary Table 1.** Genes with false discovery rate <0.1 on microarray data

FDR < 0.10 for survival of AC1					FDR based on Fisher combined p-value < 0.10 for N stage of AC1 + AC2					
Gene	Probe	FDR survival	p-value Survival	p-value Recurrence	Gene	Probe	Fisher's combined test FDR	Fisher's combined test p-value	p-value AC1	p-value AC2
TANC2	A_23_P431404	5.76E-03	1.53E-07	5.97E-02	FN1	A_24_P334130	2.53E-09	2.50E-06	8.3E-08	1.3E-03
SELP	A_23_P137697	2.55E-02	1.69E-06	2.65E-02	P4HA1	A_24_P406693	5.20E-08	2.57E-05	4.0E-05	6.3E-05
C9orf116	A_23_P422115	2.55E-02	2.06E-06	4.13E-02	ADAM12	A_23_P350512	1.03E-07	3.39E-05	8.4E-08	6.1E-02
C9orf50	A_23_P406785	2.55E-02	2.71E-06	1.40E-01	FN1	A_24_P85539	5.35E-07	1.32E-04	3.5E-06	8.4E-03
CDK9	A_23_P158083	2.89E-02	4.99E-06	5.30E-03	SCG5	A_23_P62081	1.03E-06	2.04E-04	2.4E-07	2.4E-01
ENST00000362067	A_24_P919789	2.89E-02	6.13E-06	5.54E-03	SPOCK1	A_24_P354689	1.40E-06	2.31E-04	4.8E-06	1.7E-02
FLJ25410	A_23_P89101	2.89E-02	6.08E-06	2.12E-01	TIMM8B	A_23_P98382	1.64E-06	2.31E-04	3.4E-04	2.8E-04
SPINK4	A_23_P71880	2.89E-02	4.37E-06	4.20E-01	AL050204	A_24_P937691	3.06E-06	3.77E-04	5.0E-04	3.7E-04
R78584	A_24_P225862	3.41E-02	9.05E-06	7.70E-03	COL5A1	A_23_P158593	3.96E-06	4.35E-04	1.2E-06	2.1E-01
ZNF366	A_23_P407096	3.41E-02	9.50E-06	3.93E-02	ADAM12	A_23_P202327	6.02E-06	5.95E-04	3.0E-06	1.3E-01
TCEB3C	A_23_P315910	3.41E-02	9.96E-06	1.76E-01	COL5A1	A_23_P83818	1.28E-05	1.15E-03	5.8E-06	1.5E-01
CCDC88	A_23_P24384	4.39E-02	1.40E-05	2.16E-03	COL6A1	A_24_P331918	1.54E-05	1.18E-03	2.2E-05	4.8E-02
ADCY4	A_23_P381261	5.13E-02	3.01E-05	1.02E-03	SERPINH1	A_23_P76006	1.55E-05	1.18E-03	6.9E-05	1.5E-02
LOC642730	A_24_P683013	5.13E-02	2.59E-05	3.38E-03	EVA1	A_23_P150379	1.71E-05	1.21E-03	3.8E-05	3.1E-02
BC015370	A_24_P205154	5.13E-02	3.56E-05	3.76E-03	MGC11257	A_23_P134477	2.22E-05	1.46E-03	1.4E-03	1.1E-03
CCDC88	A_23_P24389	5.13E-02	2.92E-05	5.84E-03	SDC2	A_24_P380734	2.61E-05	1.61E-03	4.3E-04	4.3E-03
RASA4	A_24_P943263	5.13E-02	3.31E-05	1.17E-02	CALD1	A_23_P42575	3.18E-05	1.85E-03	1.1E-05	2.1E-01
C21orf125	A_32_P214178	5.13E-02	2.47E-05	1.20E-02	COL11A1	A_23_P11806	3.53E-05	1.87E-03	8.3E-05	3.1E-02
ENST00000380632	A_32_P6646	5.13E-02	3.36E-05	1.42E-02	LGALS1	A_23_P166459	3.59E-05	1.87E-03	1.9E-05	1.4E-01
CBFA2T3	A_23_P500741	5.13E-02	2.89E-05	3.34E-02	FAP	A_23_P56746	4.97E-05	2.46E-03	1.6E-05	2.3E-01
IL27RA	A_24_P348326	5.13E-02	3.41E-05	4.92E-02	MLL	A_24_P281913	5.54E-05	2.61E-03	4.4E-05	9.4E-02
NRIP3	A_23_P47682	5.13E-02	2.54E-05	7.10E-02	PLA2G4B	A_23_P403424	6.41E-05	2.88E-03	2.6E-04	1.9E-02
TMEM31	A_23_P352717	5.13E-02	2.99E-05	1.09E-01	CSTB	A_23_P154894	6.80E-05	2.92E-03	4.5E-04	1.2E-02
PTPN14	A_23_P149111	5.13E-02	3.68E-05	1.63E-01	CCND1	A_24_P193011	7.71E-05	3.18E-03	1.2E-03	5.0E-03
SPATA17	A_23_P346912	5.13E-02	3.18E-05	2.06E-01	COL5A2	A_23_P10391	8.54E-05	3.38E-03	2.2E-05	3.0E-01
CCND1	A_24_P193011	5.13E-02	2.43E-05	2.54E-01	TRIO	A_24_P42603	9.49E-05	3.61E-03	7.5E-05	9.9E-02
ESM1	A_23_P144843	5.13E-02	2.67E-05	7.80E-01	POSTN	A_24_P347411	9.94E-05	3.64E-03	2.0E-04	3.9E-02
RASGRP2	A_23_P64058	5.18E-02	3.85E-05	2.70E-01	LARP6	A_23_P117782	1.09E-04	3.83E-03	1.5E-04	5.8E-02
ZNF406	A_32_P109922	5.21E-02	4.01E-05	2.72E-01	KDELRL2	A_24_P42517	1.15E-04	3.87E-03	1.3E-04	7.1E-02
LOC645277	A_32_P49867	5.66E-02	5.56E-05	6.30E-03	SEC11L1	A_23_P380917	1.20E-04	3.87E-03	2.6E-05	3.7E-01
PVRL3	A_23_P80763	5.66E-02	5.51E-05	7.92E-03	TPM2	A_23_P216501	1.21E-04	3.87E-03	6.6E-05	1.5E-01
C6orf189	A_23_P145054	5.66E-02	4.84E-05	2.59E-02	IGF1R	A_23_P305680	1.32E-04	4.01E-03	2.7E-03	3.9E-03
LOC645733	A_32_P3572	5.66E-02	5.25E-05	3.86E-02	CLEC11A	A_23_P153489	1.34E-04	4.01E-03	2.3E-04	4.7E-02
LOC92689	A_23_P132915	5.66E-02	5.49E-05	1.06E-01	DENND2D	A_23_P85952	1.56E-04	4.54E-03	1.2E-03	1.1E-02
THC2278340	A_24_P479364	5.66E-02	4.78E-05	2.16E-01	TGM3	A_23_P57118	1.63E-04	4.60E-03	1.9E-03	6.8E-03
TMEFF1	A_24_P274987	5.66E-02	5.54E-05	2.44E-01	TPM1	A_23_P391586	1.77E-04	4.80E-03	6.2E-05	2.4E-01

Supplementary Table 1. (continued)

FDR < 0.10 for survival of AC1					FDR based on Fisher combined p-value < 0.10 for N stage of AC1 + AC2					
Gene	Probe	FDR survival	p-value Survival	p-value Recurrence	Gene	Probe	Fisher's combined test FDR	Fisher's combined test p-value	p-value AC1	p-value AC2
RESP18	A_24_P59236	5.66E-02	4.68E-05	4.42E-01	PPL	A_23_P106906	1.80E-04	4.80E-03	8.1E-04	1.8E-02
CPNE5	A_23_P360804	5.78E-02	5.83E-05	1.57E-02	TPM1	A_32_P89709	1.89E-04	4.89E-03	9.5E-05	1.6E-01
TNFRSF10B	A_23_P169030	5.84E-02	6.51E-05	1.26E-01	TPM1	A_23_P206018	1.96E-04	4.89E-03	1.2E-04	1.4E-01
RAI1	A_23_P77807	5.84E-02	6.21E-05	1.37E-01	FBXO32	A_23_P82814	2.00E-04	4.89E-03	3.5E-03	4.8E-03
OPRK1	A_32_P33576	5.84E-02	6.39E-05	2.52E-01	CCND1	A_24_P124550	2.03E-04	4.89E-03	1.0E-03	1.6E-02
EFNA1	A_23_P113005	5.84E-02	6.10E-05	4.14E-01	C1QTNF6	A_24_P211565	2.12E-04	4.89E-03	4.9E-05	3.6E-01
SESTD1	A_23_P367610	5.86E-02	6.81E-05	4.26E-02	NDUFV3	A_23_P211285	2.20E-04	4.89E-03	3.1E-03	6.0E-03
FLT3	A_23_P99442	5.86E-02	6.85E-05	1.79E-01	COL1A2	A_24_P265274	2.21E-04	4.89E-03	2.1E-04	8.9E-02
GCET2	A_24_P182947	6.03E-02	7.21E-05	1.22E-01	SPARC	A_23_P7642	2.23E-04	4.89E-03	6.3E-05	3.0E-01
AK097371	A_24_P661612	6.21E-02	8.62E-05	4.76E-04	NID2	A_23_P163087	2.28E-04	4.90E-03	2.5E-05	7.8E-01
NARG2	A_32_P129419	6.21E-02	9.78E-05	1.08E-02	SRPX2	A_23_P136978	2.33E-04	4.90E-03	2.1E-04	9.2E-02
FBXO36	A_23_P432554	6.21E-02	9.64E-05	6.28E-02	ALOX12B	A_23_P83634	2.60E-04	5.35E-03	6.8E-04	3.2E-02
RP11-138L2.1	A_23_P20532	6.21E-02	8.32E-05	6.56E-02	RAB3D	A_24_P236956	2.85E-04	5.74E-03	1.1E-03	2.3E-02
WDR31	A_24_P6674	6.21E-02	8.11E-05	7.44E-02	TGM1	A_23_P65618	2.91E-04	5.75E-03	6.1E-03	4.1E-03
IL18	A_23_P104798	6.21E-02	9.75E-05	1.43E-01	AK022065	A_24_P478423	3.02E-04	5.86E-03	4.7E-03	5.5E-03
COL4A5	A_24_P290153	6.21E-02	8.43E-05	1.69E-01	RAB10	A_23_P165879	3.26E-04	6.20E-03	5.4E-02	5.3E-04
STC2	A_23_P110686	6.21E-02	9.89E-05	1.78E-01	SERPINB2	A_24_P245379	3.73E-04	6.96E-03	7.4E-04	4.4E-02
BANK1	A_23_P10232	6.21E-02	9.29E-05	1.88E-01	DBI	A_23_P79199	3.87E-04	7.08E-03	3.0E-02	1.2E-03
NAPSA	A_23_P90130	6.21E-02	9.36E-05	2.05E-01	TRIM29	A_23_P203267	4.16E-04	7.47E-03	6.0E-04	6.2E-02
P2RY14	A_24_P165864	6.21E-02	8.48E-05	3.05E-01	CALD1	A_24_P921366	4.28E-04	7.55E-03	4.3E-05	9.0E-01
CTTN	A_23_P202823	6.21E-02	8.44E-05	3.06E-01	P4HA2	A_23_P30363	4.50E-04	7.80E-03	7.2E-05	5.6E-01
TNFRSF19	A_24_P56310	6.21E-02	8.20E-05	3.75E-01	RGS5	A_23_P46045	4.62E-04	7.81E-03	4.7E-03	8.9E-03
CXCL13	A_23_P121695	6.21E-02	9.00E-05	5.54E-01	MLL	A_24_P127812	4.66E-04	7.81E-03	3.2E-04	1.3E-01
GLCE	A_23_P151870	6.21E-02	8.78E-05	6.54E-01	MGC4677	A_24_P273143	5.07E-04	8.34E-03	2.6E-04	1.8E-01
FBXO36	A_24_P254702	6.69E-02	1.08E-04	1.51E-01	RAB11FIP1	A_23_P31873	5.28E-04	8.56E-03	2.8E-04	1.8E-01
FBXO9	A_23_P254120	6.74E-02	1.31E-04	7.98E-03	P4HA2	A_23_P18966	6.04E-04	9.54E-03	1.5E-04	3.6E-01
SELE	A_23_P97112	6.74E-02	1.34E-04	1.45E-02	HSPC159	A_23_P430818	6.08E-04	9.54E-03	4.6E-03	1.2E-02
CCDC48	A_23_P166566	6.74E-02	1.29E-04	1.89E-02	COL5A2	A_23_P33196	6.23E-04	9.61E-03	6.4E-05	9.0E-01
CCDC113	A_24_P73730	6.74E-02	1.34E-04	3.53E-02	INVS	A_23_P157970	6.32E-04	9.61E-03	8.7E-05	6.8E-01
AFF1	A_23_P169619	6.74E-02	1.27E-04	4.44E-02	TGFB1	A_23_P156327	6.79E-04	1.01E-02	1.0E-04	6.1E-01
GFRA2	A_24_P96505	6.74E-02	1.29E-04	6.18E-02	COL5A3	A_23_P55749	6.83E-04	1.01E-02	2.6E-04	2.5E-01
CLEC3B	A_23_P69497	6.74E-02	1.33E-04	8.92E-02	UBTD1	A_23_P161501	6.98E-04	1.01E-02	5.2E-04	1.3E-01
MTL5	A_24_P25234	6.74E-02	1.32E-04	9.66E-02	EIF4A2	A_32_P134402	7.09E-04	1.01E-02	2.3E-01	2.9E-04
KIAA0746	A_23_P426021	6.74E-02	1.34E-04	1.15E-01	TAGLN	A_23_P87013	7.13E-04	1.01E-02	7.3E-03	9.2E-03
PPP1R16B	A_23_P352535	6.74E-02	1.24E-04	1.28E-01	GDPD3	A_23_P26511	8.32E-04	1.16E-02	5.2E-03	1.5E-02
CD19	A_23_P113572	6.74E-02	1.33E-04	2.01E-01	COL6A1	A_32_P32254	8.43E-04	1.16E-02	1.7E-04	4.7E-01

Supplementary Table 1. (continued)

FDR < 0.10 for survival of AC1					FDR based on Fisher combined p-value < 0.10 for N stage of AC1 + AC2					
Gene	Probe	FDR survival	p-value Survival	p-value Recurrence	Gene	Probe	Fisher's combined test FDR	Fisher's combined test p-value	p-value AC1	p-value AC2
THC2375612	A_32_P144999	6.74E-02	1.26E-04	2.29E-01	SEC11L1	A_24_P79413	8.68E-04	1.17E-02	1.5E-04	5.5E-01
BTN3A2	A_24_P252078	6.74E-02	1.16E-04	5.02E-01	DEFB103A	A_23_P169017	8.84E-04	1.18E-02	1.6E-02	5.3E-03
BTN3A2	A_23_P391264	6.74E-02	1.30E-04	5.16E-01	SNRP70	A_23_P4902	9.38E-04	1.24E-02	2.3E-03	4.0E-02
LOC149351	A_24_P520767	7.05E-02	1.42E-04	1.77E-02	GALNT2	A_24_P353794	9.96E-04	1.29E-02	1.7E-03	5.7E-02
AK000901	A_32_P126410	7.35E-02	1.50E-04	4.16E-01	SPINK5	A_23_P356494	1.01E-03	1.29E-02	1.1E-02	9.1E-03
C14orf81	A_24_P153558	7.43E-02	1.56E-04	2.52E-02	OPN3	A_23_P74391	1.15E-03	1.41E-02	2.1E-03	5.5E-02
THC2319172	A_32_P875465	7.43E-02	1.54E-04	5.18E-01	CDH2	A_23_P38732	1.15E-03	1.41E-02	1.8E-04	6.5E-01
C12orf35	A_24_P273561	7.56E-02	1.66E-04	7.35E-02	PRSS23	A_23_P150789	1.19E-03	1.41E-02	3.1E-04	3.9E-01
IL3RA	A_32_P217750	7.56E-02	1.67E-04	1.23E-01	COL6A3	A_32_P156322	1.20E-03	1.41E-02	2.9E-03	4.1E-02
BLNK	A_24_P64344	7.56E-02	1.68E-04	1.69E-01	EVA1	A_24_P278552	1.20E-03	1.41E-02	5.5E-03	2.2E-02
MS4A1	A_23_P116371	7.56E-02	1.66E-04	2.37E-01	IGHG1	A_23_P158817	1.20E-03	1.41E-02	1.5E-03	8.0E-02
FCRL3	A_23_P103803	7.56E-02	1.69E-04	2.56E-01	PLA2G4B	A_23_P218203	1.20E-03	1.41E-02	8.4E-04	1.4E-01
TCL1A	A_23_P357717	7.57E-02	1.71E-04	3.35E-01	SLC2A4RG	A_23_P102575	1.26E-03	1.47E-02	2.3E-03	5.6E-02
THC2444078	A_23_P88988	7.69E-02	1.76E-04	7.49E-02	MAN1B1	A_23_P94857	1.35E-03	1.55E-02	2.2E-04	6.1E-01
STEAP3	A_24_P200000	7.98E-02	1.91E-04	6.49E-03	COL5A2	A_32_P218734	1.37E-03	1.55E-02	1.4E-03	9.8E-02
CCDC93	A_23_P91062	7.98E-02	1.88E-04	1.60E-01	IGHG1	A_23_P218126	1.42E-03	1.60E-02	1.1E-03	1.3E-01
C6orf32	A_24_P941359	7.98E-02	1.86E-04	7.19E-01	ADC	A_23_P103371	1.45E-03	1.61E-02	1.8E-03	8.1E-02
ERGIC1	A_23_P333227	7.98E-02	1.90E-04	7.63E-01	WAC	A_23_P201996	1.54E-03	1.69E-02	2.5E-03	6.2E-02
AY358804	A_23_P76136	8.02E-02	2.02E-04	6.60E-03	PDGFC	A_23_P58396	1.58E-03	1.70E-02	8.3E-04	2.0E-01
KIAA1772	A_23_P119040	8.02E-02	2.04E-04	9.69E-03	RAB11FIP1	A_24_P945029	1.58E-03	1.70E-02	2.1E-04	7.9E-01
GJA4	A_23_P1083	8.02E-02	2.08E-04	2.17E-02	PLEC1	A_24_P913056	1.62E-03	1.72E-02	3.7E-04	4.5E-01
GRK5	A_23_P12884	8.02E-02	1.96E-04	3.31E-02	TPM1	A_24_P44462	1.69E-03	1.78E-02	4.8E-04	3.7E-01
PPP1R3B	A_23_P216199	8.02E-02	1.96E-04	7.57E-02	GPX7	A_24_P418816	1.86E-03	1.94E-02	1.7E-03	1.2E-01
GARNL3	A_24_P136522	8.02E-02	2.05E-04	1.02E-01	SPRR3	A_23_P62709	1.90E-03	1.95E-02	1.4E-03	1.4E-01
AL514561	A_32_P123176	8.02E-02	2.00E-04	1.55E-01	PPP2R2A	A_23_P123539	1.91E-03	1.95E-02	4.7E-03	4.3E-02
CCR7	A_23_P343398	8.02E-02	2.09E-04	2.93E-01	KDELRL2	A_23_P19938	2.02E-03	2.04E-02	4.0E-03	5.4E-02
ICAM3	A_23_P164691	8.08E-02	2.12E-04	3.84E-01	PCOLCE	A_23_P251499	2.16E-03	2.16E-02	9.3E-04	2.5E-01
RNF32	A_23_P19816	8.31E-02	2.22E-04	6.80E-02	THBS2	A_24_P605612	2.23E-03	2.20E-02	3.2E-04	7.4E-01
DCX	A_23_P500328	8.31E-02	2.24E-04	9.73E-02	PDGFC	A_24_P163168	2.25E-03	2.21E-02	2.6E-02	9.3E-03
NAPSB	A_23_P90125	8.31E-02	2.27E-04	1.75E-01	NDUFV3	A_24_P416951	2.36E-03	2.29E-02	4.7E-03	5.5E-02
TPBG	A_23_P59261	8.31E-02	2.26E-04	5.61E-01	BASP1	A_23_P213385	2.40E-03	2.30E-02	2.9E-04	9.0E-01
RAMP3	A_23_P111737	8.54E-02	2.38E-04	1.64E-01	TRIM29	A_23_P340123	2.42E-03	2.30E-02	8.3E-04	3.2E-01
PGM5	A_24_P120907	8.54E-02	2.36E-04	5.04E-01	ADC	A_24_P11462	2.46E-03	2.32E-02	2.6E-03	1.0E-01
PECAM1	A_23_P252471	8.57E-02	2.46E-04	6.51E-02	LAMB3	A_23_P86012	2.51E-03	2.34E-02	7.5E-04	3.6E-01
PARP11	A_24_P188056	8.57E-02	2.42E-04	1.13E-01	COL4A1	A_24_P68342	2.53E-03	2.34E-02	9.9E-04	2.8E-01
LRAT	A_32_P113066	8.57E-02	2.46E-04	5.70E-01	SERPINB13	A_23_P432978	2.62E-03	2.39E-02	5.9E-03	4.9E-02

Supplementary Table 1. (continued)

FDR < 0.10 for survival of AC1					FDR based on Fisher combined p-value < 0.10 for N stage of AC1 + AC2					
Gene	Probe	FDR survival	p-value Survival	p-value Recurrence	Gene	Probe	Fisher's combined test FDR	Fisher's combined test p-value	p-value AC1	p-value AC2
C6orf182	A_24_P189112	8.59E-02	2.81E-04	9.95E-03	SLC2A4RG	A_24_P365442	2.67E-03	2.40E-02	3.8E-03	7.6E-02
POLR2A	A_24_P308128	8.59E-02	2.77E-04	1.64E-02	CNFN	A_23_P27473	2.69E-03	2.40E-02	7.4E-03	4.0E-02
C20orf160	A_23_P91414	8.59E-02	2.51E-04	1.76E-02	PRSS23	A_24_P937405	2.70E-03	2.40E-02	3.8E-04	7.8E-01
PLCG2	A_23_P106675	8.59E-02	2.80E-04	3.60E-02	COL3A1	A_24_P402242	2.76E-03	2.43E-02	3.4E-03	9.0E-02
KIAA2002	A_23_P65851	8.59E-02	2.89E-04	4.11E-02	MICAL2	A_23_P24843	2.77E-03	2.43E-02	3.4E-04	9.0E-01
SPARCL1	A_23_P113351	8.59E-02	2.79E-04	7.19E-02	UTP15	A_23_P213441	3.03E-03	2.63E-02	4.8E-04	7.1E-01
LRRTM4	A_24_P60268	8.59E-02	2.69E-04	7.39E-02	ADAMTS2	A_23_P213615	3.07E-03	2.64E-02	1.2E-03	2.8E-01
SDF2L1	A_23_P6344	8.59E-02	2.81E-04	8.47E-02	SLPI	A_23_P91230	3.22E-03	2.75E-02	1.3E-03	2.8E-01
CNR2	A_23_P310931	8.59E-02	2.94E-04	8.90E-02	PI3	A_23_P210465	3.41E-03	2.85E-02	5.9E-03	6.6E-02
PRSS12	A_23_P121637	8.59E-02	2.60E-04	9.17E-02	TRIO	A_23_P425880	3.41E-03	2.85E-02	4.9E-03	7.9E-02
IFT57	A_23_P121386	8.59E-02	2.62E-04	1.32E-01	AP2M1	A_23_P155624	3.50E-03	2.91E-02	9.6E-04	4.1E-01
PCDHA5	A_23_P334045	8.59E-02	2.83E-04	1.37E-01	PPP2CB	A_23_P134693	3.71E-03	3.05E-02	1.0E-02	4.1E-02
KIAA1909	A_23_P81640	8.59E-02	2.94E-04	1.75E-01	TRIO	A_24_P913431	3.73E-03	3.05E-02	8.7E-04	4.9E-01
TDRKH	A_24_P41975	8.59E-02	2.59E-04	1.88E-01	PCNT	A_23_P57347	3.83E-03	3.10E-02	6.8E-03	6.4E-02
TBCC	A_23_P251248	8.59E-02	2.69E-04	2.37E-01	DENND2D	A_23_P311346	4.14E-03	3.33E-02	5.2E-02	9.1E-03
FCRL3	A_23_P358438	8.59E-02	2.53E-04	2.43E-01	EMP1	A_24_P921446	4.20E-03	3.33E-02	3.8E-03	1.3E-01
CR620293	A_24_P854492	8.59E-02	2.92E-04	2.84E-01	TIMP2	A_23_P107401	4.21E-03	3.33E-02	2.6E-03	1.9E-01
CCND1	A_24_P124550	8.59E-02	2.85E-04	3.47E-01	C10orf116	A_23_P161439	4.29E-03	3.36E-02	3.4E-02	1.5E-02
THC2340670	A_32_P154121	8.59E-02	2.87E-04	4.21E-01	LOC147645	A_23_P101246	4.40E-03	3.43E-02	2.7E-02	1.9E-02
PUS3	A_23_P13073	8.59E-02	2.91E-04	5.10E-01	HOP	A_24_P913146	4.45E-03	3.44E-02	3.0E-02	1.7E-02
C6orf32	A_23_P215009	8.59E-02	2.90E-04	6.01E-01	FTH1	A_24_P58337	4.79E-03	3.67E-02	9.8E-04	5.8E-01
USHBP1	A_24_P366859	8.60E-02	2.99E-04	2.17E-02	SPINK7	A_23_P213832	4.93E-03	3.73E-02	1.8E-01	3.3E-03
CPEB3	A_23_P46813	8.60E-02	2.98E-04	7.99E-02	MYO10	A_24_P46357	4.95E-03	3.73E-02	1.6E-03	3.7E-01
NCF1	A_23_P42746	8.64E-02	3.03E-04	3.50E-01	PLCXD1	A_23_P61180	5.03E-03	3.75E-02	2.2E-02	2.7E-02
EIF5	A_24_P398810	8.69E-02	3.39E-04	1.15E-03	CREB3	A_23_P423389	5.05E-03	3.75E-02	6.8E-04	8.8E-01
TMEM87A	A_24_P65098	8.69E-02	3.38E-04	1.71E-02	CXorf56	A_23_P171223	5.25E-03	3.87E-02	8.4E-03	7.5E-02
A_32_P33434	A_32_P33434	8.69E-02	3.32E-04	6.24E-02	KLK12	A_23_P500010	5.41E-03	3.96E-02	3.5E-03	1.9E-01
KLRF1	A_32_P158966	8.69E-02	3.38E-04	7.73E-02	THBS2	A_23_P62021	5.88E-03	4.27E-02	7.8E-04	9.1E-01
TMEM24	A_23_P353056	8.69E-02	3.16E-04	8.33E-02	FBXL10	A_23_P87919	5.98E-03	4.31E-02	7.0E-03	1.0E-01
IRF4	A_23_P214360	8.69E-02	3.26E-04	1.15E-01	EIF4A2	A_32_P110751	6.23E-03	4.46E-02	2.1E-01	3.6E-03
AX775927	A_32_P148627	8.69E-02	3.31E-04	1.89E-01	FGFBP1	A_23_P30126	6.48E-03	4.61E-02	8.4E-02	9.4E-03
PAEP	A_23_P257129	8.69E-02	3.32E-04	2.41E-01	RGS5	A_23_P51518	6.84E-03	4.78E-02	1.5E-01	5.6E-03
VEGF	A_23_P81805	8.69E-02	3.22E-04	3.31E-01	C1orf42	A_23_P12155	6.88E-03	4.78E-02	2.1E-02	4.1E-02
DMXL1	A_24_P101128	8.69E-02	3.25E-04	4.05E-01	PSMD2	A_24_P42681	6.97E-03	4.78E-02	3.8E-03	2.3E-01
BCNP1	A_24_P940348	8.69E-02	3.39E-04	4.85E-01	ENY2	A_23_P82748	6.99E-03	4.78E-02	2.4E-03	3.6E-01
AK024456	A_24_P928281	8.69E-02	3.31E-04	5.05E-01	TPM2	A_23_P20566	7.11E-03	4.78E-02	3.1E-03	2.9E-01

Supplementary Table 1. (continued)

FDR < 0.10 for survival of AC1					FDR based on Fisher combined p-value < 0.10 for N stage of AC1 + AC2					
Gene	Probe	FDR survival	p-value Survival	p-value Recurrence	Gene	Probe	Fisher's combined test FDR	Fisher's combined test p-value	p-value AC1	p-value AC2
MLKL	A_24_P185044	8.69E-02	3.17E-04	5.55E-01	MYO10	A_23_P7596	7.12E-03	4.78E-02	3.3E-03	2.7E-01
DVL3	A_23_P218884	8.69E-02	3.17E-04	8.66E-01	PAM	A_24_P933704	7.14E-03	4.78E-02	1.1E-01	7.8E-03
PACIN1	A_23_P258088	8.69E-02	3.16E-04	9.36E-01	HPS3	A_23_P40817	7.18E-03	4.78E-02	1.2E-02	7.5E-02
ATP2A3	A_23_P207632	8.85E-02	3.48E-04	3.24E-01	EPOR	A_23_P367899	7.20E-03	4.78E-02	1.6E-03	5.8E-01
GIMAP6	A_23_P145631	8.85E-02	3.50E-04	4.49E-01	SLC22A16	A_24_P919084	7.21E-03	4.78E-02	3.9E-01	2.3E-03
ABL2	A_23_P138099	8.88E-02	3.56E-04	2.51E-01	FTH1	A_32_P111565	7.77E-03	5.10E-02	1.3E-03	7.8E-01
CCND1	A_23_P202837	8.88E-02	3.54E-04	4.23E-01	POF1B	A_23_P159764	7.80E-03	5.10E-02	1.1E-02	8.8E-02
AVPR2	A_23_P346798	8.91E-02	3.68E-04	8.51E-02	PCNT	A_24_P8350	8.08E-03	5.23E-02	7.7E-03	1.3E-01
ABHD7	A_23_P43898	8.91E-02	3.71E-04	1.17E-01	TBRG1	A_24_P21752	8.14E-03	5.23E-02	1.3E-01	7.7E-03
NCR3	A_23_P251881	8.91E-02	3.71E-04	1.81E-01	SLC7A1	A_24_P253251	8.29E-03	5.23E-02	2.0E-02	5.4E-02
SPIB	A_23_P39067	8.91E-02	3.63E-04	3.28E-01	SERPINB13	A_23_P119015	8.29E-03	5.23E-02	7.7E-03	1.4E-01
CES4	A_23_P374892	8.91E-02	3.70E-04	3.83E-01	DEFB4	A_23_P157628	8.31E-03	5.23E-02	2.1E-02	5.0E-02
SYNPO2	A_23_P310094	8.91E-02	3.70E-04	8.26E-01	KIFAP3	A_23_P62920	8.31E-03	5.23E-02	1.2E-03	8.8E-01
NAPSA	A_32_P107029	8.95E-02	3.75E-04	1.69E-01	S100A8	A_23_P434809	8.42E-03	5.27E-02	1.9E-02	5.8E-02
LAX1	A_24_P291278	9.06E-02	3.83E-04	1.55E-01	PDZK1IP1	A_23_P160920	8.56E-03	5.32E-02	4.2E-03	2.6E-01
SOX9	A_23_P26847	9.08E-02	3.86E-04	6.83E-01	MMP1	A_32_P429083	8.83E-03	5.45E-02	6.2E-01	1.8E-03
MMRN2	A_23_P150057	9.23E-02	3.95E-04	1.51E-02	IL8	A_32_P87013	8.99E-03	5.52E-02	2.0E-02	5.9E-02
ECSM2	A_23_P72651	9.26E-02	4.04E-04	1.88E-01	TAGLN	A_23_P87011	9.58E-03	5.84E-02	4.6E-02	2.7E-02
WDR31	A_23_P251324	9.26E-02	4.02E-04	2.04E-01	STXBP1	A_23_P135310	9.89E-03	5.94E-02	1.2E-02	1.1E-01
SP100	A_24_P916816	9.26E-02	4.01E-04	2.45E-01	C1orf42	A_24_P191047	9.90E-03	5.94E-02	1.4E-01	9.0E-03
A_32_P180185	A_32_P180185	9.26E-02	4.06E-04	5.83E-01	ADAMTS2	A_23_P321307	9.93E-03	5.94E-02	2.2E-03	5.9E-01
KCNA5	A_23_P417173	9.37E-02	4.15E-04	5.39E-02	GPSM3	A_24_P230521	1.02E-02	6.06E-02	3.4E-02	4.0E-02
THC2280976	A_32_P58437	9.37E-02	4.19E-04	7.52E-02	CLDN7	A_23_P164284	1.03E-02	6.06E-02	5.3E-03	2.6E-01
LOC646686	A_24_P937649	9.37E-02	4.18E-04	8.54E-02	SLC2A4RG	A_23_P102571	1.03E-02	6.06E-02	1.7E-02	8.0E-02
TMEM105	A_32_P919718	9.37E-02	4.23E-04	2.12E-01	SULT2B1	A_23_P107981	1.05E-02	6.10E-02	2.3E-02	6.0E-02
ENST00000371030	A_32_P206479	9.37E-02	4.21E-04	3.02E-01	SMS	A_24_P305764	1.05E-02	6.10E-02	4.9E-02	2.8E-02
AL525862	A_32_P97046	9.38E-02	4.32E-04	1.48E-02	SAPS1	A_23_P119448	1.06E-02	6.10E-02	3.0E-01	4.6E-03
LOC441208	A_24_P145009	9.38E-02	4.43E-04	2.18E-02	RPL37A	A_23_P142724	1.07E-02	6.17E-02	4.0E-02	3.6E-02
KCNA3	A_23_P201138	9.38E-02	4.39E-04	2.41E-02	KDELRL2	A_23_P19936	1.08E-02	6.18E-02	6.0E-03	2.4E-01
CFP	A_23_P22444	9.38E-02	4.31E-04	3.37E-01	COL4A1	A_23_P65240	1.09E-02	6.18E-02	1.7E-03	8.5E-01
PRKCB1	A_23_P206585	9.38E-02	4.40E-04	4.04E-01	SEMA3C	A_23_P256473	1.10E-02	6.18E-02	5.7E-03	2.6E-01
CLEC10A	A_23_P141505	9.38E-02	4.36E-04	5.96E-01	FSTL1	A_23_P212696	1.11E-02	6.18E-02	2.9E-03	5.0E-01
RNF36	A_24_P50543	9.38E-02	4.30E-04	7.60E-01	DSG3	A_23_P153120	1.11E-02	6.18E-02	2.1E-02	7.1E-02
SMARCA3	A_24_P277155	9.38E-02	4.43E-04	9.02E-01	IVL	A_23_P353524	1.15E-02	6.37E-02	7.2E-02	2.1E-02
ENST00000377515	A_23_P361679	9.40E-02	4.59E-04	5.83E-03	ARL14	A_23_P92161	1.18E-02	6.49E-02	2.4E-03	6.6E-01
DDIT4	A_23_P104318	9.40E-02	4.61E-04	3.16E-02	MLLT7	A_23_P217487	1.22E-02	6.68E-02	7.0E-01	2.3E-03

**Supplementary Table 1.** (continued)

FDR < 0.10 for survival of AC1					FDR based on Fisher combined p-value < 0.10 for N stage of AC1 + AC2					
Gene	Probe	FDR survival	p-value Survival	p-value Recurrence	Gene	Probe	Fisher's combined test FDR	Fisher's combined test p-value	p-value AC1	p-value AC2
C1orf67	A_32_P87531	9.40E-02	4.67E-04	4.94E-02	SEMA4D	A_24_P261169	1.28E-02	7.00E-02	1.0E-02	1.7E-01
ENST00000311695	A_32_P188825	9.40E-02	4.57E-04	8.48E-02	SLC9A3R1	A_23_P308519	1.29E-02	7.01E-02	3.9E-03	4.6E-01
ZNF573	A_23_P339079	9.40E-02	4.66E-04	1.12E-01	DMKN	A_23_P320261	1.30E-02	7.03E-02	2.8E-02	6.3E-02
BC032907	A_24_P870362	9.40E-02	4.57E-04	1.19E-01	MGC3207	A_24_P279797	1.36E-02	7.26E-02	7.0E-03	2.7E-01
THC2407434	A_24_P565390	9.40E-02	4.51E-04	1.76E-01	AK021531	A_24_P503729	1.36E-02	7.26E-02	7.0E-03	2.7E-01
ENST00000374472	A_32_P48054	9.40E-02	4.64E-04	1.83E-01	HOP	A_23_P254507	1.37E-02	7.29E-02	2.5E-02	7.6E-02
GIMAP8	A_24_P132383	9.40E-02	4.58E-04	4.53E-01	PDE4DIP	A_24_P253100	1.39E-02	7.33E-02	4.7E-03	4.0E-01
AK056689	A_24_P548060	9.63E-02	4.81E-04	1.56E-01	SLPI	A_24_P190472	1.40E-02	7.33E-02	5.9E-03	3.2E-01
FAM27E2	A_24_P458479	9.64E-02	4.89E-04	3.67E-02	POF1B	A_24_P250815	1.41E-02	7.38E-02	2.6E-02	7.5E-02
ZC3H12D	A_24_P187826	9.64E-02	4.85E-04	1.11E-01	AEBP1	A_23_P157299	1.42E-02	7.40E-02	1.2E-02	1.6E-01
NCF1	A_32_P116203	9.64E-02	4.88E-04	3.85E-01	RAB25	A_23_P115091	1.43E-02	7.40E-02	6.8E-03	2.9E-01
HOXB9	A_23_P27013	9.65E-02	4.98E-04	1.26E-02	GPX7	A_23_P73972	1.46E-02	7.53E-02	1.2E-02	1.8E-01
TSPAN19	A_23_P2322	9.65E-02	4.92E-04	2.27E-02	YWHAH	A_23_P103070	1.49E-02	7.61E-02	1.3E-02	1.6E-01
M69012	A_24_P127159	9.65E-02	4.97E-04	9.75E-02	DKK3	A_24_P261417	1.49E-02	7.61E-02	3.8E-03	5.5E-01
TANC2	A_23_P218346	9.65E-02	4.99E-04	2.15E-01	DKK3	A_23_P162047	1.51E-02	7.64E-02	2.7E-03	7.7E-01
THC2308938	A_32_P99032	9.67E-02	5.06E-04	2.11E-01	EPPK1	A_23_P83388	1.59E-02	7.96E-02	6.0E-03	3.7E-01
CCL21	A_23_P112470	9.67E-02	5.04E-04	2.43E-01	TPM1	A_23_P363344	1.60E-02	7.96E-02	1.6E-01	1.4E-02
WDR68	A_23_P422268	9.84E-02	5.31E-04	4.83E-02	AKR1B10	A_23_P93641	1.60E-02	7.96E-02	5.0E-02	4.5E-02
ENST00000377492	A_23_P348979	9.84E-02	5.31E-04	7.30E-02	KRT2	A_23_P116850	1.60E-02	7.96E-02	4.2E-02	5.3E-02
RG55	A_23_P46045	9.84E-02	5.31E-04	1.23E-01	AGPAT2	A_32_P26103	1.62E-02	7.99E-02	7.6E-03	3.0E-01
CCL15	A_23_P218369	9.84E-02	5.27E-04	1.79E-01	TMEM40	A_23_P29551	1.62E-02	7.99E-02	1.7E-01	1.4E-02
CCR6	A_24_P234921	9.84E-02	5.34E-04	2.00E-01	LLGL2	A_23_P129738	1.68E-02	8.18E-02	1.4E-02	1.7E-01
DARC	A_23_P115161	9.84E-02	5.41E-04	2.78E-01	RPL37A	A_32_P783	1.69E-02	8.18E-02	2.6E-02	9.1E-02
CD79B	A_23_P207201	9.84E-02	5.40E-04	2.83E-01	CALD1	A_24_P255524	1.69E-02	8.18E-02	6.4E-03	3.7E-01
GPCR5C	A_23_P346670	9.84E-02	5.41E-04	3.66E-01	DUOX1	A_24_P316586	1.72E-02	8.26E-02	9.8E-02	2.5E-02
LY9	A_24_P324674	9.84E-02	5.18E-04	5.99E-01	LEPREL2	A_23_P87752	1.72E-02	8.26E-02	4.3E-01	5.8E-03
HSH2D	A_23_P153372	9.84E-02	5.34E-04	7.05E-01	FADS3	A_23_P64404	1.85E-02	8.83E-02	6.6E-03	4.0E-01
BC044628	A_32_P117453	9.89E-02	5.51E-04	4.05E-03	IMPDH2	A_24_P166042	1.88E-02	8.90E-02	2.0E-02	1.4E-01
PRAME	A_24_P216361	9.89E-02	5.53E-04	1.09E-02	FTH1	A_32_P342064	1.88E-02	8.90E-02	2.9E-03	9.5E-01
MMRN1	A_23_P18539	9.89E-02	5.54E-04	2.12E-01	RBP7	A_24_P165423	1.94E-02	9.15E-02	5.3E-02	5.4E-02
ZBP1	A_23_P259141	9.89E-02	5.51E-04	3.55E-01	HTRA1	A_23_P97990	1.97E-02	9.21E-02	4.3E-03	6.7E-01
ARHGEF15	A_24_P359007	9.89E-02	5.64E-04	6.66E-03	PRKAB2	A_24_P917711	2.01E-02	9.35E-02	2.2E-02	1.3E-01
ERGIC1	A_24_P97770	9.89E-02	5.67E-04	7.28E-03	FTH1	A_32_P820503	2.02E-02	9.36E-02	3.9E-03	7.6E-01
PCDHA1	A_24_P146138	9.89E-02	5.67E-04	4.35E-02	PTK6	A_23_P56978	2.04E-02	9.40E-02	3.1E-01	9.7E-03
ZNF540	A_23_P90542	9.89E-02	5.70E-04	5.78E-02	FLJ23447	A_23_P433798	2.05E-02	9.40E-02	3.5E-03	8.5E-01
ENST00000381158	A_32_P77831	9.89E-02	5.71E-04	1.17E-01	C9orf5	A_24_P12904	2.09E-02	9.56E-02	1.1E-02	2.8E-01

Supplementary Table 1. (continued)

FDR < 0.10 for survival of AC1					FDR based on Fisher combined p-value < 0.10 for N stage of AC1 + AC2					
Gene	Probe	FDR survival	p-value Survival	p-value Recurrence	Gene	Probe	Fisher's combined test FDR	Fisher's combined test p-value	p-value AC1	p-value AC2
CAMK2N1	A_24_P117620	9.89E-02	5.73E-04	1.58E-01	S100A9	A_23_P23048	2.15E-02	9.78E-02	2.4E-02	1.3E-01
CCDC69	A_24_P97825	9.89E-02	5.59E-04	5.16E-01	RHO	A_23_P57950	2.16E-02	9.78E-02	7.3E-03	4.4E-01
IL27RA	A_23_P27606	9.89E-02	5.75E-04	5.40E-01	KLK7	A_23_P39056	2.20E-02	9.89E-02	9.9E-02	3.3E-02
BLR1	A_24_P252945	9.92E-02	5.79E-04	2.30E-01	EPOR	A_23_P381954	2.20E-02	9.89E-02	4.4E-03	7.4E-01
FAM112A	A_23_P57020	9.95E-02	5.84E-04	2.56E-01	COL7A1	A_23_P144071	2.23E-02	9.98E-02	2.6E-02	1.3E-01
ZBTB34	A_24_P238365	9.99E-02	5.96E-04	2.12E-01	-	-	-	-	-	-
STC2	A_23_P416395	9.99E-02	5.97E-04	2.36E-01	-	-	-	-	-	-
INSIG2	A_23_P102454	9.99E-02	5.94E-04	4.59E-01	-	-	-	-	-	-
U09197	A_24_P937240	9.99E-02	5.99E-04	5.16E-01	-	-	-	-	-	-
EXOC6	A_23_P169576	9.99E-02	5.90E-04	9.45E-01	-	-	-	-	-	-

**Supplementary Table 2.** Selected gene signature and qPCR assays

Type of gene	Ensembl gene ID	Chromosome name	Band	Associated Gene name	Description	Taqman assay technical validation (TLDA.v1)	Taqman assay final signature (after technical validation) (TLDA.v2)
LNMsig	ENSG00000148848	10	q26.2	ADAM12	ADAM metallopeptidase domain 12 [Source:HGNC Symbol;Acc:HGNC:190]	Hs01106101_m1	Hs01106101_m1
Ossig	ENSG00000129467	14	q12	ADCY4	adenylate cyclase 4 [Source:HGNC Symbol;Acc:HGNC:235]	Hs00934099_m1	Hs00934099_m1
Ossig	ENSG00000116748	1	p13.2	AMPD1	adenosine monophosphate deaminase 1 [Source:HGNC Symbol;Acc:HGNC:468]	Hs00921502_m1	Hs00921502_m1
Ossig	ENSG00000109321	4	q13.3	AREG	amphiregulin [Source:HGNC Symbol;Acc:HGNC:651]	Hs00950669_m1	Hs00950669_m1
Ossig	ENSG00000118520	6	q23.2	ARG1	arginase 1 [Source:HGNC Symbol;Acc:HGNC:663]	NA	Hs00968979_m1
Ossig	ENSG00000033627	17	q21.2	ATP6V0A1	ATPase, H+ transporting, lysosomal V0 subunit a1 [Source:HGNC Symbol;Acc:HGNC:865]	Hs00193110_m1	Hs00193110_m1
Ossig	ENSG00000160345	9	q34.3	C9orf116	chromosome 9 open reading frame 116 [Source:HGNC Symbol;Acc:HGNC:28435]	Hs01077891_m1	Hs01077891_m1
Ossig	ENSG00000179058	9	q34.11	C9orf50	chromosome 9 open reading frame 50 [Source:HGNC Symbol;Acc:HGNC:23677]	Hs01368756_m1	NA
Ossig/ LNMsig	ENSG00000122786	7	q33	CALD1	caldesmon 1 [Source:HGNC Symbol;Acc:HGNC:1441]	Hs00921982_m1	Hs00921982_m1
Ossig	ENSG00000129993	16	q24.3	CBFA2T3	core-binding factor, runt domain, alpha subunit 2; translocated to, 3 [Source:HGNC Symbol;Acc:HGNC:1537]	NA	Hs00602520_m1
Ossig	ENSG00000168071	11	q13.1	CCDC88B	coiled-coil domain containing 88B [Source:HGNC Symbol;Acc:HGNC:26757]	Hs00989954_mH	Hs00989954_mH
Ossig/ LNMsig	ENSG00000110092	11	q13.3	CCND1	cyclin D1 [Source:HGNC Symbol;Acc:HGNC:1582]	Hs00765553_m1	Hs00765553_m1
Ossig	ENSG00000136807	9	q34.11	CDK9	cyclin-dependent kinase 9 [Source:HGNC Symbol;Acc:HGNC:1780]	Hs00977896_g1	NA
Ossig	ENSG00000147889	9	p21.3	CDKN2A	cyclin-dependent kinase inhibitor 2A [Source:HGNC Symbol;Acc:HGNC:1787]	NA	Hs00923894_m1

Supplementary Table 2. (continued)

Type of gene	Ensembl gene ID	Chromosome name	Band	Associated Gene name	Description	Taqman assay technical validation (TLDA.v1)	Taqman assay final signature (after technical validation) (TLDA.v2)
Ossig	ENSG00000163815	3	p21.31	CLEC3B	C-type lectin domain family 3, member B [Source:HGNC Symbol;Acc:HGNC:11891]	Hs00162844_m1	Hs00162844_m1
LNMsig	ENSG0000060718	1	p21.1	COL11A1	collagen, type XI, alpha 1 [Source:HGNC Symbol;Acc:HGNC:2186]	Hs01097664_m1	Hs01097664_m1
Ossig	ENSG00000188153	X	q22.3	COL4A5	collagen, type IV, alpha 5 [Source:HGNC Symbol;Acc:HGNC:2207]	Hs00166712_m1	Hs00166712_m1
LNMsig	ENSG00000130635	9	q34.3	COL5A1	collagen, type V, alpha 1 [Source:HGNC Symbol;Acc:HGNC:2209]	Hs00609133_m1	Hs00609133_m1
LNMsig	ENSG00000142156	21	q22.3	COL6A1	collagen, type VI, alpha 1 [Source:HGNC Symbol;Acc:HGNC:2211]	Hs01095585_m1	Hs01095585_m1
Ossig	ENSG00000085733	11	q13.3	CTTN	cortactin [Source:HGNC Symbol;Acc:HGNC:3338]	Hs01124225_m1	Hs01124225_m1
Ossig	ENSG00000156234	4	q21.1	CXCL13	chemokine (C-X-C motif) ligand 13 [Source:HGNC Symbol;Acc:HGNC:10639]	Hs00757930_m1	Hs00757930_m1
LNMsig	ENSG00000176797	8	p23.1	DEFB103A	defensin, beta 103A [Source:HGNC Symbol;Acc:HGNC:15967]	Hs00218678_m1	Hs00218678_m1
Ossig	ENSG00000100664	14	q32.32	EIF5	eukaryotic translation initiation factor 5 [Source:HGNC Symbol;Acc:HGNC:3299]	Hs01028813_g1	Hs01028813_g1
LNMsig	ENSG00000110723	11	q22.3	EXPH5	exophilin 5 [Source:HGNC Symbol;Acc:HGNC:30578]	Hs00323579_m1	Hs00323579_m1
LNMsig	ENSG00000115414	2	q35	FN1	fibronectin 1 [Source:HGNC Symbol;Acc:HGNC:3778]	Hs00415008_m1	Hs00415008_m1
HKG	ENSG00000111640	12	p13.31	GAPDH	glyceraldehyde-3-phosphate dehydrogenase [Source:HGNC Symbol;Acc:HGNC:4141]	Hs99999905_m1	Hs99999905_m1
HKG	ENSG00000169919	7	q11.21	GUSB	glucuronidase, beta [Source:HGNC Symbol;Acc:HGNC:4696]	Hs00939627_m1	Hs00939627_m1
LNMsig	ENSG00000169429	4	q13.3	IL8	interleukin 8 [Source:HGNC Symbol;Acc:HGNC:6025]	Hs00174103_m1	Hs00174103_m1
Ossig	ENSG00000163083	2	q14.2	INHBB	inhibin, beta B [Source:HGNC Symbol;Acc:HGNC:6067]	Hs00173582_m1	Hs00173582_m1

**Supplementary Table 2.** (continued)

Type of gene	Ensembl gene ID	Chromosome name	Band	Associated Gene name	Description	Taqman assay technical validation (TLDA.v1)	Taqman assay final signature (after technical validation) (TLDA.v2)
Ossig	ENSG00000176842	16	q12.2	IRX5	iroquois homeobox 5 [Source:HGNC Symbol;Acc:HGNC:14361]	Hs04334749_m1	Hs04334749_m1
Ossig	ENSG00000174718	12	p11.21	KIAA1551	KIAA1551 [Source:HGNC Symbol;Acc:HGNC:25559]	Hs01028589_m1	Hs01028589_m1
Ossig	ENSG00000134545	12	p13.2	KLRC1	killer cell lectin-like receptor subfamily C, member 1 [Source:HGNC Symbol;Acc:HGNC:6374]	Hs00970274_m1	Hs00970274_m1
LNMsig	ENSG00000118058	11	q23.3	KMT2A	lysine (K)-specific methyltransferase 2A [Source:HGNC Symbol;Acc:HGNC:7132]	Hs00610538_m1	Hs00610538_m1
LNMsig	ENSG00000108244	17	q21.2	KRT23	keratin 23, type I [Source:HGNC Symbol;Acc:HGNC:6438]	Hs00210096_m1	Hs00210096_m1
Ossig	ENSG00000204583	12	q24.33	LRCOL1	leucine rich colipase-like 1 [Source:HGNC Symbol;Acc:HGNC:44160]	Hs01113075_m1	Hs01113075_m1
LNMsig	ENSG00000139329	12	q21.33	LUM	lumican [Source:HGNC Symbol;Acc:HGNC:6724]	Hs00929860_m1	Hs00929860_m1
LNMsig	ENSG00000213401	X	q28	MAGEA12	melanoma antigen family A, 12 [Source:HGNC Symbol;Acc:HGNC:6799]	Hs04176236_m1	NA
LNMsig	ENSG00000149573	11	q23.3	MPZL2	myelin protein zero-like 2 [Source:HGNC Symbol;Acc:HGNC:3496]	Hs01083647_m1	Hs01083647_m1
Ossig	ENSG00000133055	1	q32.1	MYBPH	myosin binding protein H [Source:HGNC Symbol;Acc:HGNC:7552]	Hs00192226_m1	Hs00192226_m1
Ossig	ENSG00000133026	17	p13.1	MYH10	myosin, heavy chain 10, non-muscle [Source:HGNC Symbol;Acc:HGNC:7568]	Hs00992055_m1	NA
Ossig	ENSG00000104419	8	q24.22	NDRG1	N-myc downstream regulated 1 [Source:HGNC Symbol;Acc:HGNC:7679]	Hs00608387_m1	Hs00608387_m1
LNMsig	ENSG00000122884	10	q22.1	P4HA1	prolyl 4-hydroxylase, alpha polypeptide I [Source:HGNC Symbol;Acc:HGNC:8546]	Hs00914594_m1	Hs00914594_m1
Ossig	ENSG00000110435	11	p13	PDHX	pyruvate dehydrogenase complex, component X [Source:HGNC Symbol;Acc:HGNC:21350]	Hs00185790_m1	Hs00185790_m1
Ossig	ENSG00000154330	9	q21.11	PGM5	phosphoglucomutase 5 [Source:HGNC Symbol;Acc:HGNC:8908]	Hs00222671_m1	Hs00222671_m1

Supplementary Table 2. (continued)

Type of gene	Ensembl gene ID	Chromosome name	Band	Associated Gene name	Description	Taqman assay technical validation (TLDA.v1)	Taqman assay final signature (after technical validation) (TLDA.v2)
Ossig	ENSG00000110777	11	q23.1	POU2AF1	POU class 2 associating factor 1 [Source:HGNC Symbol;Acc:HGNC:9211]	NA	Hs01573371_m1
Ossig	ENSG00000185686	22	q11.22	PRAME	preferentially expressed antigen in melanoma [Source:HGNC Symbol;Acc:HGNC:9336]	Hs01022301_m1	Hs01022301_m1
Ossig	ENSG00000127329	12	q15	PTPRB	protein tyrosine phosphatase, receptor type, B [Source:HGNC Symbol;Acc:HGNC:9665]	Hs01549049_m1	Hs01549049_m1
LNMsig	ENSG00000143248	1	q23.3	RGS5	regulator of G-protein signaling 5 [Source:HGNC Symbol;Acc:HGNC:10001]	Hs01555176_m1	Hs01555176_m1
HKG	ENSG00000174444	15	q22.31	RPL4	ribosomal protein L4 [Source:HGNC Symbol;Acc:HGNC:10353]	Hs03044647_g1	Hs03044647_g1
HKG	ENSG00000089157	12	q24.23	RPLP0	ribosomal protein, large, P0 [Source:HGNC Symbol;Acc:HGNC:10371]	Hs99999902_m1	Hs99999902_m1
LNMsig	ENSG00000166922	15	q13.3	SCG5	secretogranin V [Source:HGNC Symbol;Acc:HGNC:10816]	Hs00161638_m1	Hs00161638_m1
Ossig	ENSG00000007908	1	q24.2	SELE	selectin E [Source:HGNC Symbol;Acc:HGNC:10718]	Hs00950401_m1	Hs00950401_m1
Ossig	ENSG00000174175	1	q24.2	SELP	selectin P (granule membrane protein 140kDa, antigen CD62) [Source:HGNC Symbol;Acc:HGNC:10721]	Hs00927900_m1	Hs00927900_m1
LNMsig	ENSG00000197632	18	q21.33	SERPINB2	serpin peptidase inhibitor, clade B (ovalbumin), member 2 [Source:HGNC Symbol;Acc:HGNC:8584]	Hs01010736_m1	Hs01010736_m1
LNMsig	ENSG00000149257	11	q13.5	SERPINH1	serpin peptidase inhibitor, clade H (heat shock protein 47), member 1, (collagen binding protein 1) [Source:HGNC Symbol;Acc:HGNC:1546]	NA	Hs01060397_g1
Ossig	ENSG00000148942	11	p14.2	SLC5A12	solute carrier family 5 (sodium/monocarboxylate cotransporter), member 12 [Source:HGNC Symbol;Acc:HGNC:28750]	Hs01054637_m1	Hs01054637_m1
Ossig	ENSG00000198021	X	q27.2	SPANXA1	sperm protein associated with the nucleus, X-linked, family member A1 [Source:HGNC Symbol;Acc:HGNC:11218]	Hs03007483_gH	Hs03007483_gH

Supplementary Table 2. (continued)

Type of gene	Ensembl gene ID	Chromosome name	Band	Associated Gene name	Description	Taqman assay technical validation (TLDA.v1)	Taqman assay final signature (after technical validation) (TLDA.v2)
Ossig	ENSG00000122711	9	p13.3	SPINK4	serine peptidase inhibitor, Kazal type 4 [Source:HGNC Symbol;Acc:HGNC:16646]	Hs00205508_m1	NA
LNMsig	ENSG00000152377	5	q31.2	SPOCK1	sparc/osteonectin, cwcv and kazal-like domains proteoglycan (testican) 1 [Source:HGNC Symbol;Acc:HGNC:11251]	Hs00928769_m1	Hs00928769_m1
Ossig	ENSG00000159516	1	q21.3	SPRR2G	small proline-rich protein 2G [Source:HGNC Symbol;Acc:HGNC:11267]	Hs00972901_s1	Hs00972901_s1
Ossig	ENSG00000172403	4	q26	SYNPO2	synaptopodin 2 [Source:HGNC Symbol;Acc:HGNC:17732]	Hs00326493_m1	Hs00326493_m1
Ossig	ENSG00000170921	17	q23.2	TANC2	tetratricopeptide repeat, ankyrin repeat and coiled-coil containing 2 [Source:HGNC Symbol;Acc:HGNC:30212]	Hs00229073_m1	Hs00229073_m1
Ossig	ENSG00000100721	14	q32.13	TCL1A	T-cell leukemia/lymphoma 1A [Source:HGNC Symbol;Acc:HGNC:11648]	NA	Hs00951350_m1
LNMsig	ENSG00000150779	11	q23.1	TIMM8B	translocase of inner mitochondrial membrane 8 homolog B (yeast) [Source:HGNC Symbol;Acc:HGNC:11818]	Hs02339636_g1	Hs02339636_g1
Ossig	ENSG00000114854	3	p21.1	TNNC1	troponin C type 1 (slow) [Source:HGNC Symbol;Acc:HGNC:11943]	Hs00896999_g1	Hs00896999_g1
Ossig	ENSG00000168477	6	p21.32	TNXB	tenascin XB [Source:HGNC Symbol;Acc:HGNC:11976]	Hs00954865_m1	Hs00954865_m1
Ossig	ENSG00000146242	6	q14.1	TPBG	trophoblast glycoprotein [Source:HGNC Symbol;Acc:HGNC:12004]	Hs00907219_m1	NA
LNMsig	ENSG00000140416	15	q22.2	TPM1	tropomyosin 1 (alpha) [Source:HGNC Symbol;Acc:HGNC:12010]	Hs00165966_m1	Hs00165966_m1
Ossig	ENSG00000110900	12	p11.21	TSPAN11	tetraspanin 11 [Source:HGNC Symbol;Acc:HGNC:30795]	Hs01391666_m1	Hs01391666_m1
Ossig	ENSG00000112715	6	p21.1	VEGFA	vascular endothelial growth factor A [Source:HGNC Symbol;Acc:HGNC:12680]	Hs00900055_m1	Hs00900055_m1

**Supplementary Table 3.** Pearson's correlation coefficients and corresponding p-values of technical validation

Gene	Correlation coefficient PCR vs Array	P-value correlation
SPRR2G	-0.98	2.30E-13
SERPINB2	-0.95	1.30E-10
MYBPH	-0.94	4.90E-10
TNNC1	-0.94	8.10E-10
KRT23	-0.93	1.90E-09
CXCL13	-0.93	4.20E-09
IL8	-0.93	1.50E-08
COL11A1	-0.9	9.10E-08
DEFB103	-0.89	1.30E-07
SYNPO2	-0.88	4.00E-07
PRAME	-0.87	5.10E-07
AREG	-0.87	7.20E-07
TIMM8B	-0.86	1.10E-06
PDHX	-0.86	1.50E-06
AMPD1	-0.86	1.50E-06
SPANXA	-0.85	1.90E-06
FN1	-0.85	2.40E-06
SPOCK1	-0.83	5.50E-06
EXPH5	-0.81	1.30E-05
KLRC1	-0.81	1.30E-05
SLC5A12	-0.8	2.00E-05
COL5A1	-0.8	2.70E-05
LUM	-0.79	2.80E-05
TSPAN11	-0.76	8.50E-05
SELE	-0.76	9.60E-05
TPM1	-0.73	2.70E-04
INHBB	-0.7	5.80E-04
MPZL2	-0.7	6.40E-04
CLEC3B	-0.69	7.40E-04
CCND1	-0.68	9.80E-04
TNXB	-0.67	1.10E-03
NDRG1	-0.67	1.30E-03
CTTN	-0.67	1.40E-03
ADAM12	-0.64	2.30E-03
SCG5	-0.64	2.40E-03
KIAA1551	-0.64	2.40E-03
P4HA1	-0.61	4.10E-03
CALD1	-0.57	8.40E-03
ADCY4	-0.56	9.50E-03
SELP	-0.56	1.00E-02
VEGFA	-0.56	1.00E-02
COL4A5	-0.54	1.40E-02
KMT2A	-0.53	1.70E-02
RG55	-0.52	1.90E-02

**Supplementary Table 3.** (continued)

<b>Gene</b>	<b>Correlation coefficient PCR vs Array</b>	<b>P-value correlation</b>
C9orf116	-0.51	2.20E-02
TANC2	-0.5	2.30E-02
PGM5	-0.48	3.30E-02
LRCOL1	-0.45	4.50E-02
COL6A1	-0.45	4.90E-02
PTPRB	-0.44	5.00E-02
CCDC88B	-0.42	6.30E-02
IRX5	-0.4	8.30E-02
MAGEA12 <sup>a</sup>	-0.37	1.10E-01
TPBG <sup>a</sup>	-0.22	3.50E-01
ATP6V0A1 <sup>a</sup>	-0.21	3.80E-01
C9orf50 <sup>a</sup>	-0.21	4.00E-01
CDK9 <sup>a</sup>	-0.17	4.70E-01
SPINK4 <sup>a</sup>	0.15	5.30E-01
MYH10 <sup>a</sup>	-0.07	7.50E-01
EIF5 <sup>a</sup>	0.03	8.90E-01

<sup>a</sup>. Genes with correlation coefficients  $\leq 1$  SD (mean  $r=0.64$ ,  $SD=0.26$ )

**Supplementary Table 4.** Coefficient estimates and false discovery rates of qPCR results of the gene signatures

Gene	Overall survival					Lymph node metastasis					Disease free survival							
	Univariate estimate	Estimate 95%CI	p-value	FDR	Ridge estimate <sup>a</sup>	Univariate estimate	Estimate 95%CI	p-value	FDR	Ridge estimate <sup>a</sup>	Univariate estimate	Estimate 95%CI	p-value	FDR	Ridge estimate <sup>a</sup>			
ADAM12	-0.05	-0.27	0.17	0.66		-0.25	-0.52	0.01	0.07	0.10	2.1E-02	-0.05	-0.28	0.18	0.67			
ADCY4	0.45	0.14	0.76	4.6E-03	0.03	0.04	0.38	-0.02	0.82	0.07		0.53	0.20	0.86	1.8E-03	0.02	0.07	
AMPD1	0.08	-0.02	0.17	0.12	0.26	4.4E-03	-0.04	-0.16	0.08	0.53		0.07	-0.04	0.17	0.20	0.37	-1.9E-02	
AREG	-0.01	-0.18	0.16	0.93	0.95	4.4E-03	-0.22	-0.45	-0.01	0.05		-0.14	-0.33	0.06	0.16	0.33	-2.6E-02	
ARG1	0.01	-0.07	0.09	0.82	0.87	-5.6E-04	0.12	0.03	0.23	0.02		-0.03	-0.12	0.05	0.40	0.50	-4.9E-02	
ATP6V0A1	0.10	-0.28	0.48	0.61	0.82	0.01	0.32	-0.12	0.80	0.17		0.09	-0.31	0.50	0.64	0.70	2.5E-02	
C9orf116	-0.30	-0.60	0.00	0.05	0.12	-0.05	-0.21	-0.57	0.14	0.25		-0.29	-0.61	0.03	0.08	0.21	-0.1	
CALD1	-0.01	-0.35	0.34	0.98	0.98	-0.01	-0.51	-0.96	-0.09	0.02	0.04	-0.1	-0.16	-0.52	0.20	0.39	0.50	-4.3E-02
CBFA2T3	0.47	0.20	0.73	5.4E-04	0.01	0.04	0.36	0.03	0.70	0.03		0.50	0.23	0.76	2.5E-04	0.00	0.07	
CCDC88B	0.31	0.02	0.59	0.03	0.11	0.03	0.25	-0.07	0.59	0.13		0.44	0.15	0.73	2.9E-03	0.02	0.10	
CCND1	-0.08	-0.29	0.14	0.48	0.67	-1.5E-03	-0.25	-0.54	0.02	0.08	0.11	-0.1	-0.13	-0.37	0.10	0.28	0.42	-1.3E-02
CDKN2A	0.02	-0.08	0.11	0.71	0.84	0.01	-0.06	-0.17	0.06	0.32		0.01	-0.09	0.11	0.81	0.81	1.5E-02	
CLEC3B	0.25	0.01	0.50	0.04	0.12	0.02	0.44	0.14	0.77	0.01		0.38	0.12	0.65	4.2E-03	0.02	4.4E-02	
COL11A1	-0.10	-0.22	0.02	0.11			-0.10	-0.23	0.03	0.15	0.17	2.8E-02	-0.07	-0.19	0.06	0.30		
COL4A5	-0.11	-0.32	0.11	0.34	0.54	-0.02	-0.20	-0.49	0.07	0.16		-0.10	-0.34	0.13	0.38	0.50	-6.8E-03	
COL5A1	-0.08	-0.31	0.14	0.45			-0.30	-0.58	-0.03	0.03	0.06	-1.2E-02	-0.17	-0.41	0.06	0.15		
COL6A1	-0.02	-0.29	0.25	0.88			-0.24	-0.56	0.07	0.14	0.17	3.3E-02	-0.03	-0.31	0.25	0.84		
CTTN	-0.11	-0.33	0.10	0.30	0.52	-0.02	0.09	-0.18	0.36	0.51		-0.25	-0.47	-0.03	0.03	0.10	-0.1	
CXCL13	0.16	0.02	0.29	0.02	0.09	0.02	0.09	-0.08	0.26	0.30		0.12	-0.02	0.26	0.09	0.23	6.9E-03	
DEFB103	0.06	-0.02	0.14	0.12			0.13	0.03	0.24	0.02	0.04	0.1	-0.03	-0.12	0.06	0.50		
EIF5	0.15	-0.26	0.56	0.48	0.67	0.02	-0.01	-0.47	0.46	0.98		0.13	-0.30	0.57	0.55	0.62	3.6E-02	
EXPH5	0.26	-0.02	0.53	0.07			0.54	0.23	0.89	1.1E-03	0.01	0.1	0.18	-0.11	0.46	0.22		
FN1	-0.13	-0.30	0.04	0.14			-0.40	-0.65	-0.17	9.3E-04	0.01	-0.1	-0.21	-0.40	-0.03	0.02		
IL8	-0.13	-0.30	0.04	0.15			-0.09	-0.28	0.10	0.34	0.38	-3E-02	-0.16	-0.34	0.03	0.10		
INHBB	-0.05	-0.26	0.17	0.66	0.83	-0.03	-0.04	-0.31	0.24	0.80		0.03	-0.19	0.25	0.78	0.80	-2.3E-02	
IRX5	0.10	-0.16	0.37	0.45	0.66	0.02	0.16	-0.15	0.48	0.31		0.18	-0.11	0.47	0.23	0.38	4.5E-02	
KIAA1551	0.33	0.00	0.66	0.05	0.12	0.04	0.22	-0.14	0.61	0.24		0.24	-0.10	0.58	0.17	0.33	2.3E-02	
KLRC1	0.14	0.01	0.28	0.03	0.11	0.03	-0.17	-0.37	0.02	0.09		0.09	-0.06	0.25	0.24	0.39	-1.6E-03	
KMT2A	0.32	-0.17	0.81	0.20			0.15	-0.40	0.70	0.59	0.59	4.2E-02	0.27	-0.23	0.76	0.29		
KRT23	-0.01	-0.14	0.12	0.92			0.27	0.10	0.47	2.8E-03	0.02	0.1	-0.03	-0.16	0.11	0.72		
LRCOL1	0.19	0.08	0.30	5.3E-04	0.01	0.08	-0.05	-0.19	0.08	0.43		0.13	0.02	0.25	0.03	0.10	0.07	
LUM	0.02	-0.24	0.28	0.89			-0.14	-0.46	0.16	0.37	0.39	-8.9E-03	0.11	-0.14	0.37	0.38		
MPZL2	0.26	0.02	0.49	0.03			0.38	0.09	0.68	0.01	0.03	3.1E-02	0.20	-0.05	0.44	0.11		
MYBPH	0.03	-0.03	0.09	0.34	0.54	-2.0E-04	-0.08	-0.15	0.00	0.04		0.01	-0.05	0.08	0.64	0.70	-2.2E-02	
NDRG1	-0.04	-0.28	0.20	0.74	0.85	0.01	-0.12	-0.40	0.15	0.39		-0.09	-0.34	0.16	0.48	0.58	1.1E-02	
P4HA1	-0.03	-0.41	0.35	0.88			-0.68	-1.18	-0.23	4.4E-03	0.02	-0.1	-0.25	-0.64	0.15	0.22		
PDHX	-0.23	-0.66	0.20	0.29	0.52	-0.03	-0.07	-0.57	0.43	0.78		-0.24	-0.69	0.22	0.31	0.42	-0.1	
PGM5	0.36	0.20	0.53	1.9E-05	7.4E-04	0.06	0.04	-0.17	0.26	0.71		0.28	0.10	0.45	2.1E-03	0.02	3.6E-02	
POU2AF1	0.15	3.0E-04	0.29	0.05	0.12	0.01	0.12	-0.06	0.31	0.20		0.11	-0.04	0.26	0.16	0.33	-2.6E-02	
PRAME	-0.01	-0.08	0.06	0.79	0.87	-0.01	0.07	-0.02	0.15	0.13		0.08	0.00	0.15	0.05	0.15	0.1	

Supplementary Table 4. (continued)

Gene	Overall survival				Lymph node metastasis						Disease free survival					
	Univariate estimate	Estimate 95%CI	p-value	FDR	Ridge estimate <sup>a</sup>	Univariate estimate	Estimate 95%CI	p-value	FDR	Ridge estimate <sup>a</sup>	Univariate estimate	Estimate 95%CI	p-value	FDR	Ridge estimate <sup>a</sup>	
PTPRB	0.27	-0.07 0.61	0.12	0.26	0.01	0.10 -0.27 0.49	0.59				0.19 -0.17 0.56	0.29	0.42	-1.9E-02		
RG55	0.33	0.09 0.58	0.01			0.37 0.02 0.74	0.05	0.07	0.1	0.39	0.13 0.66	0.00				
SCG5	0.02	-0.22 0.26	0.87			-0.31 -0.62 -0.03	0.04	0.07	-3.3E-02	-0.02	-0.27 0.22	0.85				
SELE	0.22	0.05 0.39	0.01	0.05	0.03	0.21 0.00 0.45	0.06				0.29 0.13 0.46	5.9E-04	0.01	0.06		
SELP	0.28	0.12 0.43	4.9E-04	0.01	0.05	0.24 0.03 0.48	0.03				0.33 0.17 0.48	3.3E-05	1.3E-03	0.08		
SERPINB2	-0.02	-0.14 0.10	0.77			0.21 0.05 0.39	0.02	0.04	0.1	-0.08	-0.22 0.05	0.23				
SERPINH1	0.07	-0.22 0.36	0.64			-0.56 -0.96 -0.19	4.1E-03	0.02	-0.1	-0.16	-0.43 0.12	0.26				
SLC5A12	-0.04	-0.13 0.05	0.36	0.55	-0.01	0.04 -0.07 0.16	0.43				0.05 -0.04 0.14	0.28	0.42	0.06		
SPANXA	-0.01	-0.08 0.05	0.69	0.84	0.00	0.04 -0.05 0.12	0.40				0.01 -0.06 0.09	0.70	0.74	4.2E-02		
SPOCK1	-0.05	-0.24 0.14	0.58			-0.32 -0.57 -0.09	0.01	0.03	-4.5E-02	-0.20	-0.40 0.00	0.05				
SPRR2G	0.02	-0.06 0.10	0.66	0.83	-3.7E-03	0.18 0.07 0.31	2.9E-03				-0.03 -0.12 0.06	0.51	0.61	-4.3E-02		
SYNPO2	0.10	-0.05 0.25	0.20	0.39	1.5E-03	-0.04 -0.22 0.13	0.63				0.13 -0.03 0.29	0.11	0.24	7.0E-03		
TANC2	-0.04	-0.39 0.31	0.83	0.87	-4.8E-03	-0.43 -0.87 -0.01	0.05				-0.37 -0.74 0.00	0.05	0.15	-0.07		
TCL1A	0.20	0.08 0.31	1.0E-03	0.01	0.06	0.14 -0.01 0.29	0.08				0.17 0.05 0.30	0.01	0.03	0.06		
TIMM8B	0.17	-0.18 0.52	0.33			0.32 -0.07 0.73	0.11	0.14	2.6E-02	0.20	-0.17 0.57	0.29				
TNNC1	0.05	-0.03 0.12	0.23	0.45	4.0E-03	-0.06 -0.15 0.03	0.22				0.05 -0.03 0.13	0.22	0.38	8.3E-03		
TNXB	0.24	0.04 0.43	0.02	0.09	0.03	0.22 0.00 0.46	0.06				0.31 0.11 0.51	2.5E-03	0.02	0.06		
TPM1	0.04	-0.15 0.24	0.67			-0.45 -0.73 -0.20	7.9E-04	0.01	-0.1	-0.07	-0.28 0.13	0.50				
TSPAN11	0.28	0.04 0.51	0.02	0.09	0.03	0.15 -0.17 0.47	0.36				0.26 -0.01 0.52	0.06	0.16	3.3E-02		
VEGFA	-0.26	-0.51 -0.02	0.04	0.12	-0.04	-0.08 -0.39 0.22	0.59				-0.22 -0.48 0.05	0.11	0.24	-3.5E-02		

<sup>a</sup> Confidence intervals are not provided, because methodology to construct these is lacking for logistic ridge regression. In addition, they are likely uninformative (wide) due to collinearity in the variables.

**Supplementary Table 5.** Coefficients and p-values of clinical and pathological variables for integration with gene signatures

Variable name	Univariate overall survival		Multivariate overall survival		Univariate disease-free survival		
	p-value	Coefficient	p-value	Coefficient	p-value	Coefficient	
ECOG (0 <sup>a</sup> /1)	0.14	0.59			0.92	-0.05	
Age	0.001	0.04	9.50E-04	0.05	0.69	0.005	
ACE27 (none <sup>a</sup> /mild/moderate+severe)	0.02	0.93	1.16		0.28	0.46	-0.06
Sex (male <sup>a</sup> /female)	0.1	0.51			0.64	0.15	
PackYears	0.005	0.02	0.003	0.02	0.32	0.007	
pTNM (1+2 <sup>a</sup> vs. 3+4)	0.44	0.24	0.29	0.33	0.39	0.29	
pCompVar (negative <sup>a</sup> /positive)	7.00E-06	1.4			7.00E-05	1.3	

<sup>a</sup> Reference category

**Supplementary Table 6.** Remark criteria checklist

<b>Item to be reported</b>		
Introduction		
1	State the marker examined, the study objectives, and any pre-specified hypotheses.	X
Material and methods		
Patients		
2	Describe the characteristics (e.g., disease stage or co-morbidities) of the study patients, including their source and inclusion and exclusion criteria.	X
3	Describe treatments received and how chosen (e.g., randomized or rule-based).	X
Specimen characteristics		
4	Describe type of biological material used (including control samples) and methods of preservation and storage.	X
Assay methods		
5	Specify the assay method used and provide (or reference) a detailed protocol, including specific reagents or kits used, quality control procedures, reproducibility assessments, quantitation methods, and scoring and reporting protocols. Specify whether and how assays were performed blinded to the study endpoint.	X
Study design		
6	State the method of case selection, including whether prospective or retrospective and whether stratification or matching (e.g., by stage of disease or age) was used. Specify the time period from which cases were taken, the end of the follow-up period, and the median follow-up time.	X
7	Precisely define all clinical endpoints examined.	X
8	List all candidate variables initially examined or considered for inclusion in models.	X
9	Give rationale for sample size; if the study was designed to detect a specified effect size, give the target power and effect size.	X
Statistical analysis methods		
10	Specify all statistical methods, including details of any variable selection procedures and other model-building issues, how model assumptions were verified, and how missing data were handled.	X
11	Clarify how marker values were handled in the analyses; if relevant, describe methods used for cutpoint determination.	X
Results		
Data		
12	Describe the flow of patients through the study, including the number of patients included in each stage of the analysis (a diagram may be helpful) and reasons for dropout. Specifically, both overall and for each subgroup extensively examined report the numbers of patients and the number of events.	X
13	Report distributions of basic demographic characteristics (at least age and sex), standard (disease-specific) prognostic variables, and tumor marker, including numbers of missing values.	X
Analysis and presentation		
14	Show the relation of the marker to standard prognostic variables.	X
15	Present univariable analyses showing the relation between the marker and outcome, with the estimated effect (e.g., hazard ratio and survival probability). Preferably provide similar analyses for all other variables being analyzed. For the effect of a tumor marker on a time-to-event outcome, a Kaplan-Meier plot is recommended.	X
16	For key multivariable analyses, report estimated effects (e.g., hazard ratio) with confidence intervals for the marker and, at least for the final model, all other variables in the model.	X
17	Among reported results, provide estimated effects with confidence intervals from an analysis in which the marker and standard prognostic variables are included, regardless of their statistical significance.	X
18	If done, report results of further investigations, such as checking assumptions, sensitivity analyses, and internal validation.	X
Discussion		
19	Interpret the results in the context of the pre-specified hypotheses and other relevant studies; include a discussion of limitations of the study.	X
20	Discuss implications for future research and clinical value.	X

**Supplementary Table 7.** Housekeeping gene performance

Housekeeping gene	Correlation coefficient					SD <sup>a</sup>
	GAPDH	GUSB	RPL4	RPLP0	mean-Ct target genes <sup>b</sup>	
GAPDH	1	0.77	0.82	0.83	0.69	1.17
GUSB	0.77	1	0.72	0.71	0.77	0.94
RPL4	0.82	0.72	1	0.87	0.67	1.05
RPLP0	0.83	0.71	0.87	1	0.65	0.98

<sup>a</sup>. Standard deviation of housekeeping gene across all samples.

<sup>b</sup>. Correlation of Ct-value of housekeeping gene and average Ct-value of all target genes (n=60).

**Supplementary Table 8.** Assessment of performance of the gene signature in relevant subgroups by integrated area-under-the-curve (iAUC)

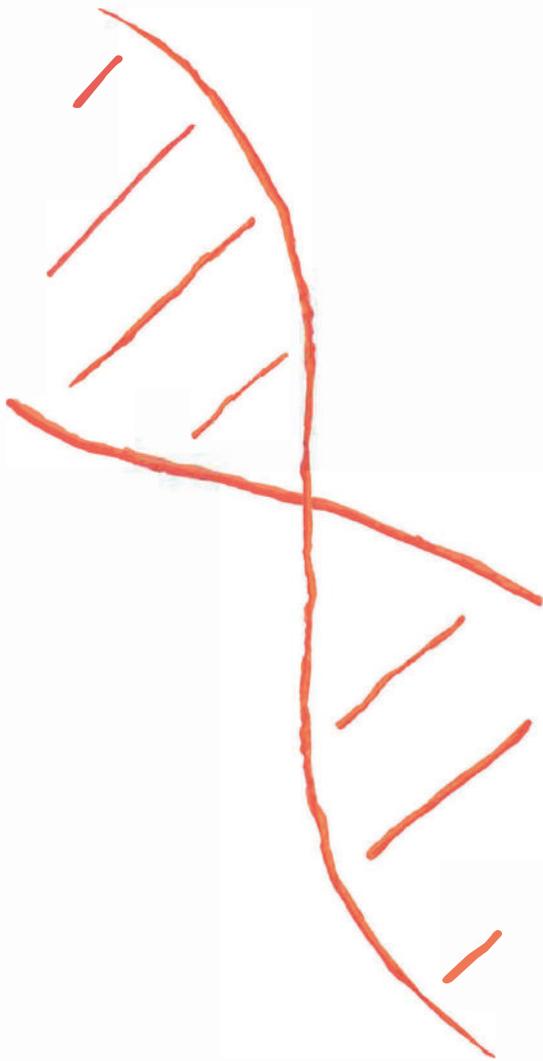
Subgroup	OS			DFS			n
	iAUC	95% CI		iAUC	95% CI		
CompVar: 0	0.71	0.65	0.76	0.65	0.61	0.68	79
CompVar: 1	0.58	0.51	0.65	0.62	0.55	0.7	38
Age<70	0.64	0.59	0.7	0.62	0.57	0.68	88
Age>=70	0.56	0.48	0.66	0.76	0.7	0.82	37
PackYears < median	0.61	0.54	0.69	0.7	0.65	0.75	59
PackYears >= median	0.54	0.48	0.62	0.54	0.47	0.63	66
Female	0.58	0.51	0.67	0.56	0.49	0.63	53
Male	0.63	0.56	0.69	0.64	0.58	0.72	72
pTNM: 1 or 2	0.6	0.52	0.68	0.53	0.45	0.64	43
pTNM: 3 or 4	0.63	0.57	0.69	0.63	0.57	0.7	82
ACE27: 0-1	0.66	0.6	0.73	0.68	0.63	0.75	77
ACE27: 2-3	0.65	0.57	0.74	0.56	0.49	0.63	48
Treatment surgery only	0.51	0.43	0.61	0.69	0.65	0.71	60
Treatment surgery + adjuvant	0.67	0.61	0.73	0.65	0.58	0.73	65

## REFERENCES

1. Ferlay J, Soerjomataram I, Dikshit R, et al. Cancer incidence and mortality worldwide: sources, methods and major patterns in GLOBOCAN 2012. *Int J Cancer* 2015; 136(5): E359-86.
2. Braakhuis BJ, Leemans CR, Visser O. Incidence and survival trends of head and neck squamous cell carcinoma in the Netherlands between 1989 and 2011. *Oral Oncol* 2014; 50(7): 670-5.
3. Gillison ML, Chaturvedi AK, Anderson WF, Fakhry C. Epidemiology of Human Papillomavirus-Positive Head and Neck Squamous Cell Carcinoma. *J Clin Oncol* 2015; 33(29): 3235-42.
4. Castellsague X, Alemany L, Quer M, et al. HPV Involvement in Head and Neck Cancers: Comprehensive Assessment of Biomarkers in 3680 Patients. *J Natl Cancer Inst* 2016; 108(6): djv403.
5. Braakhuis BJ, Snijders PJ, Keune WJ, et al. Genetic patterns in head and neck cancers that contain or lack transcriptionally active human papillomavirus. *J Natl Cancer Inst* 2004; 96(13): 998-1006.
6. O'Sullivan B, Huang SH, Su J, et al. Development and validation of a staging system for HPV-related oropharyngeal cancer by the International Collaboration on Oropharyngeal cancer Network for Staging (ICON-5): a multicentre cohort study. *Lancet Oncol* 2016; 17(4): 440-51.
7. Brierley JDG, M. K.; Wittekind, C., eds. *TNM Classification of Malignant Tumours. 8th Edition ed.* Hoboken: Wiley-Blackwell; 2016.
8. Chinn SB, Myers JN. Oral Cavity Carcinoma: Current Management, Controversies, and Future Directions. *J Clin Oncol* 2015; 33(29): 3269-76.
9. Leemans CR, Braakhuis BJ, Brakenhoff RH. The molecular biology of head and neck cancer. *Nat Rev Cancer* 2011; 11(1): 9-22.
10. Chung CH, Parker JS, Ely K, et al. Gene expression profiles identify epithelial-to-mesenchymal transition and activation of nuclear factor-kappaB signaling as characteristics of a high-risk head and neck squamous cell carcinoma. *Cancer Res* 2006; 66(16): 8210-8.
11. De Cecco L, Bossi P, Locati L, Canevari S, Licitra L. Comprehensive gene expression meta-analysis of head and neck squamous cell carcinoma microarray data defines a robust survival predictor. *Ann Oncol* 2014; 25(8): 1628-35.
12. Lohavanichbutr P, Mendez E, Holsinger FC, et al. A 13-gene signature prognostic of HPV-negative OSCC: discovery and external validation. *Clin Cancer Res* 2013; 19(5): 1197-203.
13. Jung AC, Job S, Ledrappier S, et al. A poor prognosis subtype of HNSCC is consistently observed across methylome, transcriptome, and miRNome analysis. *Clin Cancer Res* 2013; 19(15): 4174-84.
14. Tan PK, Downey TJ, Spitznagel EL, Jr., et al. Evaluation of gene expression measurements from commercial microarray platforms. *Nucleic Acids Res* 2003; 31(19): 5676-84.
15. Roepman P, Kemmeren P, Wessels LF, Slootweg PJ, Holstege FC. Multiple robust signatures for detecting lymph node metastasis in head and neck cancer. *Cancer Res* 2006; 66(4): 2361-6.
16. Roepman P, Wessels LF, Kettelarij N, et al. An expression profile for diagnosis of lymph node metastases from primary head and neck squamous cell carcinomas. *Nat Genet* 2005; 37(2): 182-6.
17. van Hooff SR, Leusink FK, Roepman P, et al. Validation of a gene expression signature for assessment of lymph node metastasis in oral squamous cell carcinoma. *J Clin Oncol* 2012; 30(33): 4104-10.
18. Schilling C, Stoeckli SJ, Haerle SK, et al. Sentinel European Node Trial (SENT): 3-year results of sentinel node biopsy in oral cancer. *Eur J Cancer* 2015; 51(18): 2777-84.
19. Cancer Genome Atlas N. Comprehensive genomic characterization of head and neck squamous cell carcinomas. *Nature* 2015; 517(7536): 576-82.
20. McShane LM, Altman DG, Sauerbrei W, et al. Reporting recommendations for tumor marker prognostic studies. *J Clin Oncol* 2005; 23(36): 9067-72.
21. Smeets SJ, Hesselink AT, Speel EJ, et al. A novel algorithm for reliable detection of human papillomavirus in paraffin embedded head and neck cancer specimen. *Int J Cancer* 2007; 121(11): 2465-72.
22. Ritchie ME, Phipson B, Wu D, et al. limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res* 2015; 43(7): e47.
23. Livak KJ, Schmittgen TD. Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. *Methods* 2001; 25(4): 402-8.
24. Goeman JJ, Oosting J, Cleton-Jansen AM, Anninga JK, van Houwelingen HC. Testing association of a pathway with survival using gene expression data. *Bioinformatics* 2005; 21(9): 1950-7.

25. Goeman JJ, van de Geer SA, de Kort F, van Houwelingen HC. A global test for groups of genes: testing association with a clinical outcome. *Bioinformatics* 2004; 20(1): 93-9.
26. Benjamini Y, Hochberg Y. Controlling the False Discovery Rate - a Practical and Powerful Approach to Multiple Testing. *J R Stat Soc B* 1995; 57(1): 289-300.
27. Alkureishi LW, Ross GL, Shoaib T, et al. Sentinel node biopsy in head and neck squamous cell cancer: 5-year follow-up of a European multicenter trial. *Ann Surg Oncol* 2010; 17(9): 2459-64.
28. Den Toom IJ, Heuveling DA, Flach GB, et al. Sentinel node biopsy for early-stage oral cavity cancer: the VU University Medical Center experience. *Head Neck* 2015; 37(4): 573-8.
29. Civantos FJ, Zitsch RP, Schuller DE, et al. Sentinel lymph node biopsy accurately stages the regional lymph nodes for T1-T2 oral squamous cell carcinomas: results of a prospective multi-institutional trial. *J Clin Oncol* 2010; 28(8): 1395-400.
30. Leusink FK, van Es RJ, de Bree R, et al. Novel diagnostic modalities for assessment of the clinically node-negative neck in oral squamous-cell carcinoma. *Lancet Oncol* 2012; 13(12): e554-61.
31. Pramana J, Pimentel N, Hofland I, et al. Heterogeneity of gene expression profiles in head and neck cancer. *Head Neck* 2007; 29(12): 1083-9.
32. Johnson WE, Li C, Rabinovic A. Adjusting batch effects in microarray expression data using empirical Bayes methods. *Biostatistics* 2007; 8(1): 118-27.
33. Rickman DS, Millon R, De Reynies A, et al. Prediction of future metastasis and molecular characterization of head and neck squamous-cell carcinoma based on transcriptome and genome analysis by microarrays. *Oncogene* 2008; 27(51): 6607-22.
34. Onken MD, Winkler AE, Kanchi KL, et al. A surprising cross-species conservation in the genomic landscape of mouse and human oral cancer identifies a transcriptional signature predicting metastatic disease. *Clin Cancer Res* 2014; 20(11): 2873-84.
35. Thurlow JK, Pena Murillo CL, Hunter KD, et al. Spectral clustering of microarray data elucidates the roles of microenvironment remodeling and immune responses in survival of head and neck squamous cell carcinoma. *J Clin Oncol* 2010; 28(17): 2881-8.
36. Winter SC, Buffa FM, Silva P, et al. Relation of a hypoxia metagene derived from head and neck cancer to prognosis of multiple cancers. *Cancer Res* 2007; 67(7): 3441-9.
37. Chen C, Mendez E, Houck J, et al. Gene expression profiling identifies genes predictive of oral squamous cell carcinoma. *Cancer Epidemiol Biomarkers Prev* 2008; 17(8): 2152-62.
38. Sobin LH, M.K.; Wittekind, C.; eds. *TNM Classification of Malignant Tumours*. 7th Edition ed. Hoboken: Wiley-Blackwell; 2009.
39. Batsakis JG. Surgical excision margins: a pathologist's perspective. *Adv Anat Pathol* 1999; 6(3): 140-8.
40. van den Brekel MW, Lodder WL, Stel HV, Bloemena E, Leemans CR, van der Waal I. Observer variation in the histopathologic assessment of extranodal tumor spread in lymph node metastases in the neck. *Head Neck* 2012; 34(6): 840-5.
41. Oken MM, Creech RH, Tormey DC, et al. Toxicity and response criteria of the Eastern Cooperative Oncology Group. *Am J Clin Oncol* 1982; 5(6): 649-55.
42. Piccirillo JF, Tierney RM, Costas I, Grove L, Spitznagel EL, Jr. Prognostic importance of comorbidity in a hospital-based cancer registry. *JAMA* 2004; 291(20): 2441-7.
43. Schemper M, Smith TL. A note on quantifying follow-up in studies of failure time. *Control Clin Trials* 1996; 17(4): 343-6.
44. Heagerty PJ, Lumley T, Pepe MS. Time-dependent ROC curves for censored survival data and a diagnostic marker. *Biometrics* 2000; 56(2): 337-44.
45. Jiang B, Zhang XG, Cai TX. Estimating the confidence interval for prediction errors of support vector machine classifiers. *J Mach Learn Res* 2008; 9: 521-40.
46. Jiang W, Varma S, Simon R. Calculating confidence intervals for prediction error in microarray classification using resampling. *Stat Appl Genet Mol Biol* 2008; 7(1): Article8.
47. Wahl S, Boulesteix AL, Zierer A, Thorand B, van de Wiel MA. Assessment of predictive performance in incomplete data by combining internal validation and multiple imputation. *BMC Med Res Methodol* 2016; 16(1): 144.





# 4

## **Outcome prediction of head and neck squamous cell carcinoma by MRI radiomic signatures**

Steven W. Mes, Floris H.P. van Velden, Boris Peltenburg, Carel F.W. Peeters, Dennis E. te Beest, Mark A. van de Wiel, Joost Mekke, Doriene C. Mulder, Roland M. Martens, Jonas A. Castelijns, Frank A. Pameijer, Remco de Bree, Ronald Boellaard, C. René Leemans, Ruud H. Brakenhoff, Pim de Graaf

**European Radiology 2020; 30(11):6311-6321**

## **ABSTRACT**

### **Objectives**

Head and neck squamous cell carcinoma (HNSCC) shows a remarkable heterogeneity between tumors, which may be captured by a variety of quantitative features extracted from diagnostic images, termed radiomics. The aim of this study was to develop and validate MRI-based radiomic prognostic models in oral and oropharyngeal cancer.

### **Materials and Methods**

Native T1-weighted images of four independent, retrospective (2005–2013), patient cohorts ( $n = 102$ ,  $n = 76$ ,  $n = 89$ , and  $n = 56$ ) were used to delineate primary tumors, and to extract 545 quantitative features from. Subsequently, redundancy filtering and factor analysis were performed to handle collinearity in the data. Next, radiomic prognostic models were trained and validated to predict overall survival (OS) and relapse-free survival (RFS). Radiomic features were compared to and combined with prognostic models based on standard clinical parameters. Performance was assessed by integrated area under the curve (iAUC).

### **Results**

In oral cancer, the radiomic model showed an iAUC of 0.69 (OS) and 0.70 (RFS) in the validation cohort, whereas the iAUC in the oropharyngeal cancer validation cohort was 0.71 (OS) and 0.74 (RFS). By integration of radiomic and clinical variables, the most accurate models were defined (iAUC oral cavity, 0.72 (OS) and 0.74 (RFS); iAUC oropharynx, 0.81 (OS) and 0.78 (RFS)), and these combined models outperformed prognostic models based on standard clinical variables only ( $p < 0.001$ ).

### **Conclusions**

MRI radiomics is feasible in HNSCC despite the known variability in MRI vendors and acquisition protocols, and radiomic features added information to prognostic models based on clinical parameters.

## INTRODUCTION

Head and neck squamous cell carcinoma (HNSCC) is a malignancy arising in the mucosal lining of the oral cavity, oropharynx, larynx, and hypopharynx<sup>1</sup>. Unfortunately, mortality rates are high<sup>2</sup>, and long-term functional deficits often remain after therapy<sup>3</sup>. Ideally, treatment is personalized to maximize treatment efficacy and minimize side effects. However, treatment personalization is currently only based on stage, site, and histological parameters after surgery, with suboptimal performance<sup>4</sup>.

Despite that HNSCC arise in one tissue type, they are remarkably heterogeneous hampering accurate prediction of clinical behavior<sup>5</sup>. This heterogeneous tumor biology may be captured by imaging<sup>6,7</sup>. In the past, images were mostly described by qualitative features such as dimension and invasion in neighboring structures, but currently images are also being analyzed by extraction of a variety of quantitative features, also termed radiomics<sup>8</sup>.

Radiomic analyses have previously been applied in HNSCC patients, but most studies focused on computed tomography (CT), most particularly for radiotherapy planning. Aerts et al described a prognostic radiomic signature based on CT scans of lung cancer and applied this signature successfully in oropharyngeal cancer<sup>9</sup>. Others followed with comparable approaches<sup>10-14</sup>. The preference for CT is explained by (i) intuitive interpretation of signal intensities that correspond to tissue radiodensity<sup>8</sup>, (ii) standardization of imaging performance across vendors and scanners<sup>8</sup>, and (iii) availability of delineated tumor volumes from radiation treatment plans.

Nonetheless, in clinical practice, magnetic resonance imaging (MRI) is often the modality of choice for imaging of head and neck tumors, because of the superior soft tissue contrast. However, the acquired MRI signal intensities are influenced by scanner parameters and many image acquisition-related factors<sup>15</sup>. Still, MRI can identify physical properties of the tumor by application of separate sequence acquisition protocols (e.g., diffusion-weighted MRI (DWI), dynamic contrast-enhanced (DCE) MRI<sup>16</sup>), and therefore, MRI might better capture overall tumor biology than CT. As such, MRI radiomics was able to categorize breast cancer, glioblastoma, and prostate cancer in different molecular subtypes<sup>17-19</sup>. In HNSCC, prognostic models based on MRI radiomics were only described for small series of less than 20 cases of oropharyngeal cancer<sup>20,21</sup> or heterogeneous cohorts<sup>22,23</sup>.

In this study, we present an MRI radiomics workflow based on T1-weighted images that is applied in two independent patient cohorts of oral cancer (n = 102 and n = 76) and two cohorts of HPV-negative oropharyngeal cancer (n = 89 and n = 56) for prediction of overall survival (OS) and relapse-free survival (RFS).

## MATERIAL AND METHODS

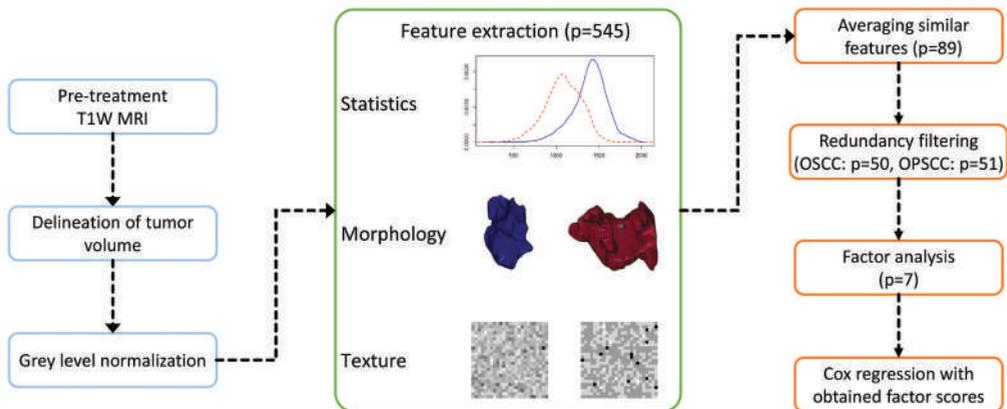
### Patients

Four independent, retrospective cohorts of HNSCC patients included (i) a cohort of oral squamous cell carcinoma (OSCC) patients from Amsterdam UMC, location VUmc (VUMC), treated from 2005 to 2013; (ii) a cohort of OSCC patients from University Medical Center Utrecht (UMCU) treated from 2010 to 2013; (iii) a cohort of HPV-negative oropharyngeal squamous cell carcinoma (OPSCC) patients from VUMC, treated from 2008 to 2012; and (iv) a cohort of HPV-negative OPSCC patients from UMCU treated from 2010 to 2013. All patients were treated with curative intent. HPV status was assessed with p16 immunohistochemistry and subsequent PCR-based HPV DNA detection on p16-immunopositive cases. HPV-positive tumors were excluded because this group is considered to be a separate disease entity within HNSCC<sup>24</sup>, which would interfere with radiomic findings<sup>25</sup> and clinical outcome<sup>26</sup>. The Dutch Medical Research Involving Human Subjects Act (WMO) does not apply to this study and therefore informed consent was waived by the Medical Ethics Review Committee at Amsterdam UMC. Medical records were reviewed to obtain clinical characteristics, including age at diagnosis,

gender, comorbidity, and clinical TNM-stage (7th edition)<sup>27</sup>. Comorbidity was classified using the Adult Comorbidity Evaluation 27 (ACE-27)<sup>28</sup>. Two outcome measures were used: (a) OS, which was defined as time from date of incidence to death from any cause; and (b) RFS, which was defined as time from date of incidence to development of locoregional recurrence, distant metastasis, or second primary HNSCC. For RFS, patients who died of other causes or developed other tumors outside the head and neck region were censored at the date of death or incidence date of the other tumor.

## MRI

The schematic workflow of this study is depicted in Figure 1. Axial 2D T1W images without gadolinium enhancement and short T1 inversion recovery (STIR) (OSCC VUMC, OSCC UMCU, OPSCC VUMC) or T2-weighted (OPSCC UMCU) images were available for all patients. These scans were obtained using scanners of different vendors and protocols (Supplementary Table 1). Native T1W images were used for feature extraction because this sequence was available for all tumors. The STIR sequence was used to facilitate tumor segmentation, and for feature extraction in the OSCC cohorts to assess a possible additional prognostic value. Our protocols of contrast-enhanced T1W imaging changed in time (e.g., slice thickness, 2D versus 3D, with or without fat saturation), and therefore this sequence was not considered in this study.

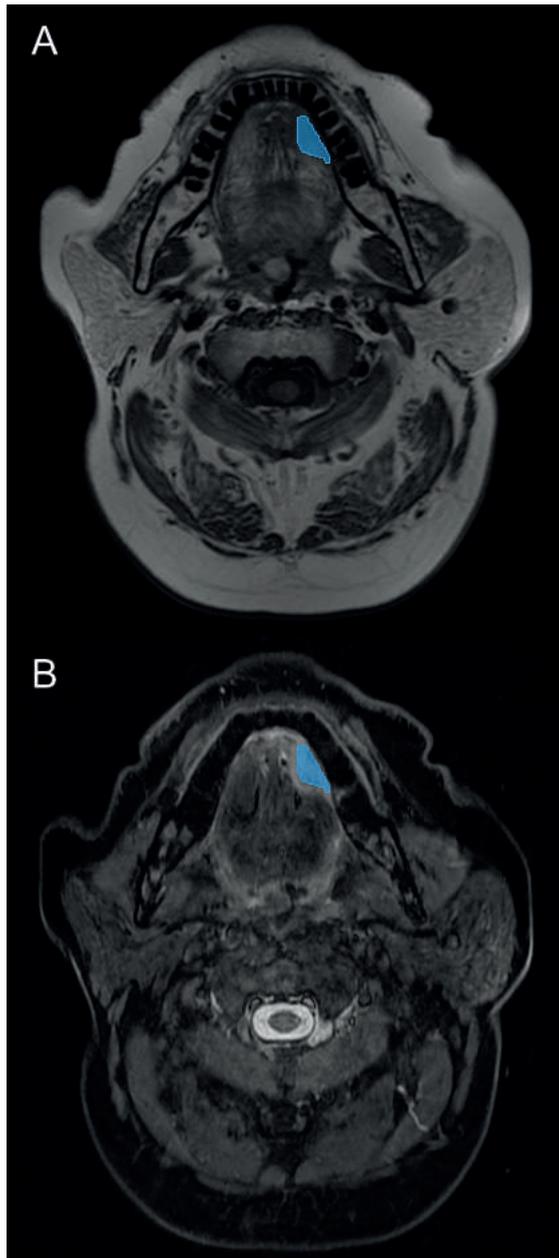


**Figure 1.** Illustration of radiomics pipeline

Abbreviations: MRI, magnetic resonance imaging; OPSCC, oropharyngeal squamous cell carcinoma; OSCC, oral cavity squamous cell carcinoma; T1W, T1-weighted.

## Segmentation

MR images of VUMC patients were transferred to VelocityAI 3.1 (Varian Medical Systems), whereas UMCU MRI scans were transferred to an in-house developed target volume delineation tool<sup>29</sup>. Subsequently, STIR images were automatically co-registered to the T1W images and registration was visually checked. Supervised manual delineation of all primary tumors was performed by S.M. and B.P. (both 3 years of experience) with visual inspection of delineation by senior head and neck radiologists (P.G. or F.P. with 11 and 25 years of experience). In Figure 2, an example of a delineated tumor is shown on T1W MRI and STIR.



4

**Figure 2.** Illustration of tumor segmentation on T1 MRI and STIR  
Exemplary segmentation of a T2N2b tongue tumor on the left side on T1W MRI (A) and STIR (B).

### Feature extraction and processing

The feature extraction and processing can be found in detail in the Supplementary Methods. The extracted features are described in Table 1.

**Table 1.** Radiomic raw features (p = 545)

Group	Number	Name
First order statistics	35	From entire image (before normalization): Maximum gray level, minimum gray level, range, mean, median, standard deviation, maximum gray level of all values over 0.5, median of all values over 0.5, mean of all values over 0.5 From tumor VOI (after normalization): Maximum gray level, minimum gray level, range, mean, median, standard deviation, interquartile range, coefficient of variation (COV, in percentage), skewness, kurtosis, excess kurtosis, median absolute deviation of the median, mean absolute deviation of the median, mean absolute deviation of the mean, mean Laplacian, total energy, variance, root-mean-square (RMS), mean of the maximum voxel and the six adjacent voxels ( $Max_{star}$ ), integrated intensity, entropy <sup>a</sup> , uniformity <sup>a</sup>
Spatial autocorrelation	2	Moran's I, Geary's C
Intensity-volume histogram features	1	Area under a cumulative intensity-volume histogram curve (AUC)
Morphological features	11	Tumor volume, surface area, surface to volume ratio, surface area to surface of an equivolumetric sphere to volume ratio, radius of an equivolumetric sphere, compactness 1, compactness 2, spherical disproportion, sphericity, asphericity, maximum 3D diameter
Fractal features	4	Fractal dimension (calculated), fractal dimension (fitted), fractal abundance, fractal lacunarity
Texture features based on gray level co-occurrence matrix <sup>a,b</sup>	300	Joint maximum, joint average, joint variance, joint entropy, difference average, difference variance, difference entropy, sum average, sum variance, sum entropy, angular second moment, contrast, dissimilarity, inverse difference, inverse difference normalized, inverse difference moment, inverse difference moment normalized, inverse variance, correlation, autocorrelation, cluster tendency, cluster shade, cluster prominence, first measure of information correlation, second measure of information correlation
Texture features based on gray level run length <sup>a,b</sup>	192	Short runs emphasis, long runs emphasis, low gray level run emphasis, high gray level run emphasis, short run low gray level emphasis, short run high gray level emphasis, long run low gray level emphasis, long run high gray level emphasis, gray level non-uniformity, gray level non-uniformity normalized, run length non-uniformity, run length non-uniformity normalized, run percentage, gray level variance, run length variance, run entropy

<sup>a</sup> Obtained using a discretization of 32, 64 or 128 gray level bins

<sup>b</sup> Calculated from matrices per direction and then averaged (average), or from merged matrix created using all matrices over all directions (combined). The matrices were calculated either per x-y plane (2D, but all planes were used in the calculation) or volumetrically (3D).

## Interobserver feature stability

MRI scans of 30 OPSCCs were re-segmented by an independent senior head and neck radiologist (J.C., with 35 years of experience) according to the pipeline described before. Subsequently, feature extraction was performed and the mean value of similar features was determined, leaving n = 89 unique features. The Kendall's coefficient of concordance was determined and a coefficient of  $\geq 0.7$  was considered high concordance.

## Factor analysis and model training

The subsequent steps of predictive modelling that were applied in this study have been described before<sup>30</sup>, and can be found in detail in the Supplementary Methods.

## Influence of vendor and magnetic field strength

As described above, a variety of MRI acquisition protocols and equipment of different vendors were used. Although this may impact the radiomics analyses, it reflects current clinical routine. Ideally, a correlation analyses would be performed of test-retest data from different vendors and magnetic field strengths to standardize the data, but such datasets are not available. Instead, multivariate analysis of variance (MANOVA) was performed to compare the mean factor scores between vendors and magnetic field strength in VUMC patient cohorts. In the UMCU cohorts (OSCC and OPSCC), only the mean factor scores between magnetic field strengths were compared, because most scans were obtained using one MR vendor (Table 2).

## RESULTS

### Patient characteristics

Patient cohorts consisted of 102 patients (VUMC OSCC), 76 patients (UMCU OSCC), 89 patients (VUMC OPSCC), and 56 patients (UMCU OPSCC). Patient characteristics for each cohort are presented in Table 2. VUMC OSCC and UMCU OSCC cohorts had similar distributions of age and gender, but VUMC patients presented with higher comorbidity scores ( $p < 0.001$ ), more advanced T-stage ( $p < 0.01$ ), and consequently a poorer overall survival ( $p = 0.01$ ). In contrast, VUMC OPSCC and UMCU OPSCC cohorts only differed significantly from each other in terms of ACE-27 score ( $p = 0.01$ ). Moreover, the scans were obtained using scanners of different vendors and protocols (see also Supplementary Table 1).

**Table 2.** Patient characteristics

		VUMC OSCC	UMCU OSCC	VUMC OPSCC	UMCU OPSCC	P-value*	P-value±
Number of cases		102	76	89	56		
Median age	years (MAD)	63 (11.9)	66.3 (11.1)	60 (7.4)	64 (11.9)	0.23	0.24
Gender	male	64 (62.7)	46 (60.5)	49 (55.1)	35 (62.5)		
	female	38 (37.3)	30 (39.5)	40 (44.9)	21 (37.5)	0.77	0.48
Smoking	current	51 (50.0)	34 (44.7)	54 (60.7)	34 (60.7)		
	former	35 (34.3)	24 (31.6)	26 (29.2)	13 (23.2)		
	never	16 (15.7)	15 (19.7)	9 (10.1)	6 (10.7)		
	unknown	0 (0)	3 (3.9)	0 (0)	3 (5.4)	0.23	0.16
Alcohol	current	68 (66.7)	49 (64.5)	66 (74.2)	40 (71.4)		
	former	13 (12.7)	6 (7.9)	12 (13.5)	10 (17.9)		
	never	21 (20.6)	17 (22.4)	11 (12.4)	3 (5.4)		
	unknown	0 (0)	4 (5.3)	0 (0)	3 (5.4)	0.11	0.07
ACE27	0	28 (27.5)	27 (35.5)	26 (29.2)	17 (30.4)		
	1	34 (33.3)	40 (52.6)	33 (37.1)	27 (48.2)		
	2	28 (27.5)	4 (5.3)	27 (30.3)	7 (12.5)		
	3	12 (11.8)	5 (6.6)	3 (3.4)	1 (1.8)		
	unknown	0 (0)	0 (0)	0 (0)	4 (7.1)	<0.001	0.01
T-stage	1	12 (11.8)	20 (26.3)	7 (7.9)	6 (10.7)		
	2	36 (35.3)	28 (36.8)	35 (39.3)	17 (30.4)		
	3	21 (20.6)	4 (5.3)	16 (18.0)	13 (23.2)		
	4	33 (32.4)	24 (31.6)	31 (34.8)	20 (35.7)	<0.01	0.67
N-stage	0	62 (60.8)	51 (67.1)	40 (44.9)	18 (32.1)		
	1	20 (19.6)	6 (7.9)	14 (15.7)	7 (12.5)		
	2	20 (19.6)	19 (25.0)	35 (39.3)	30 (53.6)		
	3	0 (0)	0 (0)	0 (0)	1 (1.8)	0.13	0.19
Stage	I	10 (9.8)	18 (23.7)	4 (4.5)	4 (7.1)		
	II	23 (22.5)	17 (22.4)	17 (19.1)	6 (10.7)		
	III	25 (24.5)	6 (7.9)	15 (16.9)	7 (12.5)		
	IV	44 (43.1)	35 (46.1)	53 (59.6)	39 (69.6)	0.01	0.4
Vendor	GE	49 (48.0)	0 (0)	70 (78.7)	0 (0)	<0.001	<0.001
	Philips	4 (3.9)	76 (100)	1 (1.1)	55 (98.2)		
	Siemens	48 (47.1)	0 (0)	18 (20.2)	1 (1.8)		

**Table 2.** (continued)

		VUMC OSCC	UMCU OSCC	VUMC OPSCC	UMCU OPSCC	P-value*	P-value±
	Toshiba	1 (1.0)	0 (0)	0 (0)	0 (0)		
Magnetic field strength	1.0T	12 (11.8)	0 (0)	1 (1.1)	0 (0)	<0.001	<0.001
	1.5 T	83 (81.4)	58 (76.3)	71 (79.8)	21 (37.5)		
	3.0T	7 (6.9)	18 (23.7)	17 (19.1)	35 (62.5)		
Survival	deceased	49 (48.0)	24 (31.6)	47 (52.8)	28 (50.0)		
	alive	53 (52.0)	52 (68.4)	42 (47.2)	28 (50.0)	0.01	0.87
Median time to death	years (MAD)	1.4 (1.2)	1.3 (1.0)	2.1 (1.9)	2.0 (1.5)	0.03	0.17
Median follow-up time (alive patients)	years (MAD)	4.5 (2.0)	3.7 (0.9)	5.9 (1.7)	5.0 (0.5)	<0.001	<0.001

Abbreviations: MAD, median absolute deviation; OPSCC, oropharyngeal squamous cell carcinoma; OSCC, oral cavity squamous cell carcinoma; T, Tesla.

P-value\* = VUMC OSCC compared to UMCU OSCC, and calculated with the use of Student's t-test for continuous variables and  $\chi^2$  test for categorical variables.

P-value± = VUMC OPSCC compared to UMCU OPSCC, and calculated with the use of Student's t-test for continuous variables and  $\chi^2$  test for categorical variables.

## Normalization

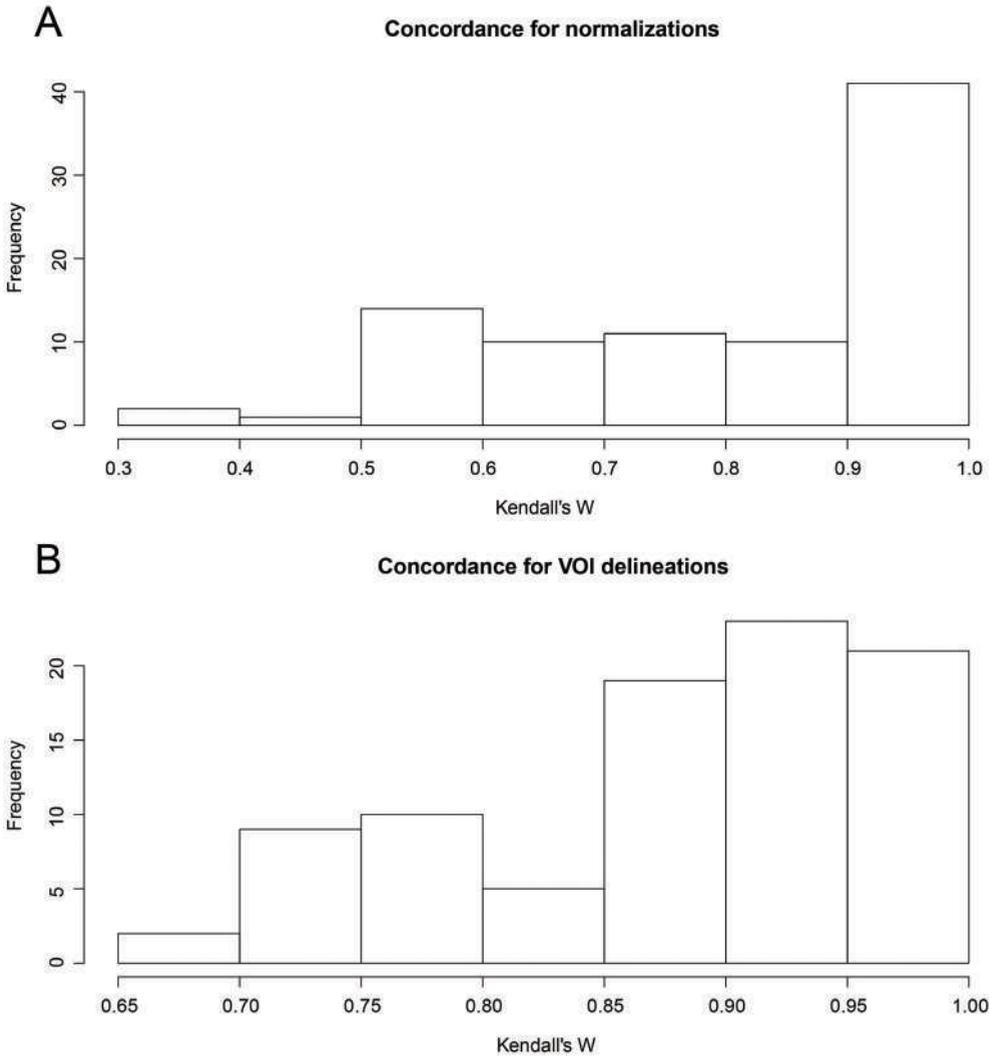
Since different MRI parameters were used on MRI systems supplied by different vendors, we assessed the influence of signal intensities on radiomic analysis<sup>31,32</sup> using five gray level normalization methods that are described in the Supplementary Methods. A high concordance was found for the 89 radiomics features before and after normalization (mean = 0.82, sd = 0.19). Figure 3 a shows a histogram of the concordances of the core 89 radiomics features. Given the minor influence of gray level normalization on these features, it was decided to proceed with unnormalized data.

## Interobserver stability

Another putative important variable in radiomics feature extraction is definition of the tumor contours by manual delineation, which may introduce variability in the data by inconsistency of segmentation<sup>33</sup>. Therefore, the stability of the radiomics features of a random subgroup of 30 VUMC OPSCCs was assessed when the tumors were delineated by two independent radiologists. A high concordance was found of the 89 radiomics features (mean = 0.88, sd = 0.09) suggesting that delineation by experienced radiologists is consistent or minor changes in delineation do not impact radiomic features. Figure 3b displays the concordances of the 89 core radiomics features with multiple delineations.

## Dimension reduction and factor analysis

Redundancy filtering was applied to the 89 core radiomic features to remove highly correlated features which resulted in 50 features (VUMC OSCC dataset) and 51 features (VUMC OPSCC dataset). A regularized estimator of the correlation matrix between the features was obtained, and factor analysis was performed on this matrix, which showed that both VUMC OSCC features and VUMC OPSCC features were described by 7 latent factors. The factors accounted for 78% (VUMC OSCC) and 77% (VUMC OPSCC) of the variation in the data. The 7 factors can be roughly interpreted as representing (i) 3D geometrics, (ii) meta-gray level co-occurrence, (iii) meta-first order, (iv) gray level mix, (v) meta-gray level run length, (vi) geometrics, and (vii) entropy. The exact content of each factor is shown in Supplementary Table 2 (OSCC) and Supplementary Table 3 (OPSCC). The highest variation in both datasets is explained by factors 1 (3D geometrics) and 2 (meta-gray level co-occurrence).



**Figure 3.** Radiomic features showed high concordance before and after gray level normalization and high interobserver stability (A) Five methods of gray level normalization were performed before feature extraction and the concordance was calculated of the 89 averaged radiomics features before and after normalization. The figure shows an histogram of the Kendall's coefficients of concordance (mean = 0.82, sd = 0.19).

(B) For 30 VUMC OPSCCs, interobserver stability was assessed by delineation of the tumors by two independent radiologists. The figure shows an histogram of the Kendall's coefficients of concordance (mean = 0.88, sd = 0.09).

### OSCC prognostic models

The 7 extracted latent factors were used to train a model to predict OS and RFS of OSCC patients. For OS, an iAUC was found of 0.69 in both the VUMC OSCC cohort and the UMCU OSCC cohort (Table 3). For RFS, iAUCs of 0.63 and 0.70 were found in the VUMC OSCC cohort and the UMCU OSCC cohort, respectively (Table 3). These radiomics models were compared to models using (i) tumor volume, and (ii) clinical variables (N-stage, age at diagnosis and gender). Tumor volume only had a limited prognostic value (iAUC 0.50–0.60). Compared to the radiomics only model, the clinical models performed equally or worse (Table 3). Subsequently, the radiomics

and clinical models were combined to assess whether this could further improve the performance. Indeed, the most accurate models were found when radiomics and clinical data were combined (Table 3), and the iAUC improvement was also statistically significant (Supplementary Table 4). Figure 4a and b show Kaplan-Meier curves of the UMCU OSCC cohort with group stratification based on the median predicted risk.

**Table 3.** Performance of radiomic, clinical, and combined models in OSCC and OPSCC cohorts

	Overall survival		Relapse-free survival	
	iAUC (95% CI <sup>a</sup> )	P <sup>b</sup> value	iAUC (95% CI <sup>a</sup> )	P <sup>b</sup> value
OSCC VUMC - training				
Radiomic	0.69 (0.59-0.73)		0.63 (0.50-0.68)	
Clinical <sup>c</sup>	0.69 (0.61-0.75)		0.60 (0.49-0.66)	
Radiomic + clinical <sup>c</sup>	0.75 (0.65-0.77)		0.65 (0.51-0.67)	
OSCC UMCU - validation				
Radiomic	0.69 (0.52-0.75)	0.009	0.70 (0.54-0.75)	0.003
Clinical <sup>c,d</sup>	0.65 (0.51-0.72)	0.02	0.64 (0.51-0.70)	0.08
Radiomic + clinical <sup>c,d</sup>	0.72 (0.55-0.74)	0.01	0.74 (0.58-0.78)	<0.001
OPSCC VUMC - training				
Radiomic	0.71 (0.62-0.76)		0.70 (0.58-0.77)	
Clinical <sup>c</sup>	0.57 (0.46-0.61)		0.56 (0.42-0.61)	
Radiomic + clinical <sup>c</sup>	0.73 (0.62-0.76)		0.70 (0.56-0.75)	
OPSCC UMCU - validation				
Radiomic	0.71 (0.58-0.77)	0.02	0.74 (0.60-0.83)	0.08
Clinical <sup>c,d</sup>	0.74 (0.64-0.83)	<0.001	0.71 (0.58-0.82)	0.01
Radiomic + clinical <sup>c,d</sup>	0.81 (0.68-0.91)	<0.001	0.78 (0.62-0.83)	0.04

Abbreviations: CI, confidence interval; iAUC, integrated Area Under the Curve; OPSCC, oropharyngeal squamous cell carcinoma; OSCC, oral cavity squamous cell carcinoma.

<sup>a</sup> CIs were assessed by bootstrapping.

<sup>b</sup> Assessed by log-rank testing in validation cohorts with group stratification based on the median predicted risk.

<sup>c</sup> Clinical models consisted of N-stage, age at diagnosis and gender.

<sup>d</sup> Recalibration of coefficients of clinical variables was allowed to optimize comparability with radiomic models.

For delineation, STIR imaging was also used since the tumors are more clearly discriminated from normal tissue on this sequence. Radiomic features extracted from this sequence may also further improve the prognostic model, and therefore additional prognostic models based on the combination of STIR and T1W MRI radiomic features were trained and validated. In the training cohort, the iAUC did not improve by using the combination of T1W MRI and STIR (Table 4), whereas in the validation cohort the iAUC did improve, but the precision of the estimated iAUC is low given the wide confidence intervals. The difference between the cohorts might also be explained by the shorter follow-up time in the OSCC UMCU cohort or the smaller cohort size (Table 2).

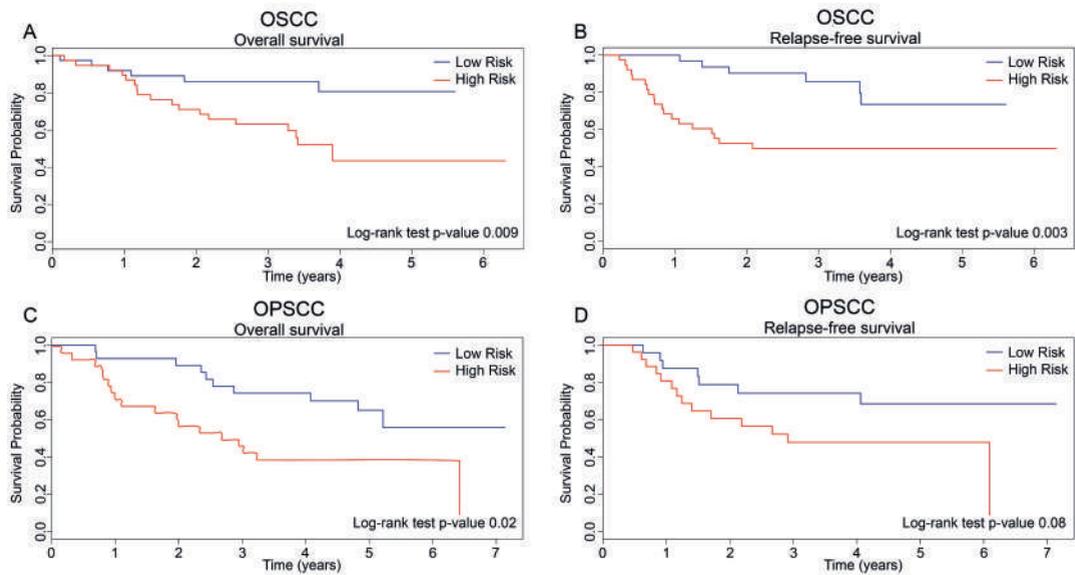
### OPSCC prognostic models

Following the strategy of assessing the relevance of radiomics models in OSCC patients, OPSCC models were trained using radiomics, clinical data, tumor volume, and a combination of both. Note that the study encompassed only HPV-negative cases. Similarly to the OSCC cohorts, radiomics-only models predicted the outcome of OPSCC patients (Table 3). The clinical models, however, were less informative in the VUMC cohort (Table 3). The better performance of the clinical models in the OPSCC UMCU cohort may relate to the shorter follow-up time or the smaller cohort size (Table 2). The combined models showed the highest iAUCs (Table 3), and were significantly better than radiomic and clinical models (Supplementary Table 4). Tumor volume

only had a limited prognostic value (iAUC 0.53–0.64). Figure 4c and d show Kaplan-Meier curves of the UMCU OPSCC cohort with group stratification based on the median predicted risk.

### Influence of vendor and magnetic field strength

Radiomic features were extracted from scans with three different magnetic field strengths (Table 2). The VUMC cohorts also consisted of data extracted from scanners of various MR vendors (Table 2). MANOVA analysis implied that there might be an effect of the field strength on factor 3 (meta-first order), factor 4 (gray level mix), and factor 5 (meta-gray level run length) (Supplementary Table 5). Second, MANOVA analysis presented a possible effect of MR vendor on factor 3 (meta-first order), factor 4 (gray level mix), factor 5 (meta-gray level run length), factor 6 (geometrics), and factor 7 (entropy) (Supplementary Table 6). However, the indicated effects were not consistent across datasets, except for factor 3 (meta-first order).



**Figure 4.** The radiomic signature predicts overall and relapse-free survival in oral cavity squamous cell carcinoma and oropharyngeal squamous cell carcinoma

(A, B) Kaplan-Meier analysis of overall survival (A) and relapse-free survival (B) with risk groups defined by median predicted hazards of the radiomic signature in the UMCU validation cohort of 76 OSCC patients.

(C, D) Kaplan-Meier analysis of overall survival (C) and relapse-free survival (D) of different risk groups defined by median predicted hazards of the radiomic signature in the UMCU validation cohort of 56 OPSCC patients.

All P-values are calculated using a log-rank test. Tick marks on curves indicate censoring.

Abbreviations: OPSCC, oropharyngeal squamous cell carcinoma; OSCC, oral cavity squamous cell carcinoma.

**Table 4.** Performance of radiomic (T1W + STIR), clinical, and combined models in OSCC cohort

	Overall survival		Relapse-free survival	
	iAUC (95% CI <sup>a</sup> )	P <sup>b</sup> value	iAUC (95% CI <sup>a</sup> )	P <sup>b</sup> value
OSCC VUMC - training				
Radiomic	0.67 (0.57-0.71)		0.62 (0.47-0.65)	
Clinical <sup>c</sup>	0.69 (0.61-0.75)		0.60 (0.49-0.66)	
Radiomic + clinical <sup>c</sup>	0.74 (0.64-0.76)		0.65 (0.49-0.66)	
OSCC UMCU - validation				
Radiomic	0.80 (0.68-0.84)	<0.001	0.72 (0.57-0.77)	0.01
Clinical <sup>c,d</sup>	0.65 (0.51-0.72)	0.02	0.64 (0.51-0.70)	0.08
Radiomic + clinical <sup>c,d</sup>	0.82 (0.67-0.83)	<0.001	0.76 (0.61-0.80)	0.001

Abbreviations: CI, confidence interval; iAUC, integrated Area Under the Curve; OPSCC, oropharyngeal squamous cell carcinoma; OSCC, oral cavity squamous cell carcinoma; STIR, Short TI Inversion Recovery; T1W, T1-weighted.

<sup>a</sup> CIs were assessed by bootstrapping.

<sup>b</sup> Assessed by log-rank testing in validation cohorts with group stratification based on the median predicted risk.

<sup>c</sup> Clinical models consisted of N-stage, age at diagnosis and gender.

<sup>d</sup> Recalibration of coefficients of clinical variables was allowed to optimize comparability with radiomic models.

## DISCUSSION

This study was set out to develop prognostic models based on MRI radiomics in oral cavity and oropharyngeal cancer patients. Although MRI is most commonly used in head and neck cancer imaging, clinical routine shows a large variety of MRI vendors and MRI acquisition protocols, which might hamper radiomic analyses. Here we show that despite this potential problem, relevant information can be extracted.

In four patient cohorts, 545 quantitative features were extracted from native T1W MRI, and a four-step method was applied to reduce dimensions while preserving the data's covariation<sup>30</sup>. This method includes redundancy filtering and factor analysis, and provided models based on 7 latent factors both in OSCC and in OPSCC. These factors roughly describe tumor intensity (i.e., "graylevel-mix" and "meta-firstorder"), shape (i.e., "3D geometrics" and "geometrics"), and texture (i.e., "meta-graylevelco-occurrence," "meta-graylevelrunlength," and "entropy"). In validation setting, the prognostic performance of these models was accurate, and the combined models outperformed clinical characteristics alone in predicting both OS and RFS. These results are very promising and indicate that MRI radiomic analysis may have additional value to current prognostic variables.

Furthermore, as with all prognostic models, it is important that it applies in settings outside the reference hospitals involved in the development. Partly, this was overcome by using independent validation cohorts provided by a second institution that uses imaging equipment from different vendors. Moreover, feature stability remained high with and without gray level normalization, and did not depend on interobserver variability. Together this suggests that the external validity of the signature described is expected to be high.

To date, only few prognostic MRI radiomic signatures for HNSCC have been published<sup>20-23</sup>. Most previous studies applied radiomic analyses to CT scans of HNSCC patients<sup>10-13</sup>, and comparable performance of the prognostic models was reported. However, in these studies, delineated CT scans from radiotherapy treatment plans were used, which are often not available in surgically treated patients and thereby not available for many HNSCC patients. Nonetheless, MRI radiomics has been applied to nasopharyngeal carcinoma<sup>34-39</sup>, which is a separate disease entity<sup>40</sup>.

Next to radiomic signatures, there is a myriad of other prognostic biomarkers for HNSCC available that, for instance, are based on imaging<sup>41</sup>, immunohistochemistry<sup>42</sup>, and microarray data<sup>4</sup>. The advantages of our radiomic profile is that it is available before treatment and based on standard diagnostic images, thereby avoiding additional costs and discomfort for the patient. Moreover, radiomic analyses may better capture tumor heterogeneity than biomarkers<sup>43</sup>.

Our study has several strengths. First, standard-of-care native T1W MR images were used to extract the radiomic features. This sequence is used in almost all clinical HNSCC protocols and makes the results broadly applicable. Second, multiple adequately sized patient cohorts were imaged on scanners of different vendors to develop and validate the models, which further contributes to the generalizability of the approach. In addition, features were not very sensitive to delineation. Finally, the prognostic signature is interpretable for clinicians: the latent factors represented different tumor characteristics and were subsequently used in Cox regression. Cox regression is familiar to most clinicians as opposed to machine learning algorithms<sup>44</sup>, alleviating the “black box” effect of many high-throughput prognostic models.

However, there are also limitations to be identified in this study. Foremost, the MRI scans of the tumors in this study stem from scanners of different vendors and were attained with different acquisition settings, causing data variability. Indeed, our analyses indicate that some factor scores might be influenced by the variety of scanning protocols and used MR equipment. This is especially true for factor 3, which is made up of features describing first-order statistics that would be expected to be influenced by acquisition settings and magnetic field strength. However, the largest variability in the data was explained by factors 1 and 2, which appeared not to be influenced by vendor and field strength variability. Nonetheless, more uniform data will likely improve model performance and validity<sup>8</sup>. Finally, the radiomic signatures were combined with several important clinical variables (e.g., N-stage, age at diagnosis), but combination with other important clinical factors such as smoking (packyears) and alcohol consumption (unityears) might improve prediction accuracy further<sup>45</sup>. Of note, the retrospective nature of this study precluded the use of the 8th edition of the UICC TNM Classification because important information was not available (i.e., clinical depth of invasion and clinical extranodal extension). However, it has been shown that the 8th edition outperforms the previous edition<sup>46</sup>, and including the new system in future studies may improve prediction of the clinical and combined clinical-radiomic models.

In conclusion, we developed and validated a prognostic signature based on radiomic features extracted from standard-of-care MRI. This finding suggests that important prognostic information is present in MRI databases of HNSCC patients across the world. It also implicates that MRI acquisition protocols should be further standardized to optimize exchangeability of data and models. Future research could focus on analysis of feature stability by scanning patients on scanners of different vendors, and on the same scanner at multiple time points (test-retest analysis). Moreover, we already show that combining multiple sequences may improve the prognostic performance of the model, while future studies should incorporate functional MRI sequences and multiple imaging modalities (i.e., CT and PET) to capture more aspects of tumor biology.

## ACKNOWLEDGMENTS

The authors thank Furkan Yaz for his help in data acquisition and extraction, and Patricia Doornaert and Tezontl Rosario for the assistance in tumor delineation using VelocityAI 3.1.

## CONFLICTS OF INTEREST

The authors of this manuscript declare no relationships with any companies whose products or services may be related to the subject matter of the article.

## FUNDING

This study has received funding by the European Union’s Seventh Framework Project (grant agreement 611425: OraMod) and by the Netherlands Organisation for Health Research and Development (ZonMw) (grant 10-10400-98-14002).

## SUPPLEMENTARY METHODS

### Feature extraction

In total, 545 radiomic features were extracted from the tumor VOIs projected onto the T1W images using ACCURATE<sup>47,48</sup>, which is available at <https://petralymphoma.org/accurate-tool/>. This version was modified to allow DICOM input of MRI. No voxel interpolation was applied before feature extraction. The features were either based on first order statistics (n=35), intensity-volume histograms (n=1), morphology (n=11), fractals (n=4), spatial autocorrelation (n=2) or texture (n=492). Texture features were derived from gray-level co-occurrence matrices and from gray-level run-length matrices. Each texture feature was calculated in four ways. For the first approach, matrices were created per direction and per x-y plane (2D) for all planes and all four possible directions, and subsequently combined into a single matrix that was used to calculate the feature from (2D combined). For the second approach, the feature was calculated from each matrix, created per direction and per x-y plane (2D) for all planes and all four possible directions, and then averaged (2D averaged). For the third approach, matrices were created per direction for the entire VOI (3D) for all thirteen possible directions, and subsequently combined into a single matrix that was used to calculate the feature from (3D combined). For the fourth approach, the feature was calculated from each matrix, created per direction for all thirteen possible directions, and then averaged (3D averaged). In addition, a discretization of 32, 64 or 128 gray level bins was applied (Table 1). All features were implemented according to the recommendations set by the image biomarker standardization initiative<sup>49</sup>.

### Normalization

Because various MRI parameters were used including intensity features, a gray level normalization was applied to the MRI scans in advance of feature extraction. Four schemes for normalization were considered:

- N-I: no normalization was applied.
- N-II: a multiplicative transformation was used to fix the range of gray levels for all the images<sup>31</sup>. Each gray level was multiplied by the ratio MEDIAN/(median of the reference volume of interest (VOI)) where MEDIAN was a constant, i.e. the median value for all analyzed reference VOIs.
- N-III: the gray levels within the tumor VOI that were located outside the range of three standard deviations of the average voxel intensity were not considered in further analysis<sup>31</sup>.
- N-IV: the median value of the reference VOI was subtracted from gray levels within the tumor VOI and subsequently divided by the median absolute deviation (MAD) within the reference VOI according to the following formula:

$$I_N = \frac{I - \text{median}(I_{ref})}{MAD_{ref}}$$

Where  $I_N$  is the normalized gray level within the tumor VOI,  $I$  is the gray level within the tumor VOI in advance of normalization,  $I_{ref}$  is the gray level within the reference VOI and  $MAD_{ref}$  is the MAD within the reference VOI.

For N-II and N-IV, two reference structures were used for normalization, i.e. the obliquus capitis inferior muscle (N-IIa and N-IVa) and the myelum (N-IIb and N-IVb).

### Feature processing

As described before, texture features were calculated using different settings for coding of gray level intensity, different numbers of considered directions and different methods of including the obtained values in the

considered directions. The average value of similar features with different calculation settings was determined reducing the initial 545 raw radiomic features to 89 core features (see also Table 1). These core features were subsequently used in downstream analysis (see *Factor analysis*).

### Factor analysis

After obtaining the core 89 features, redundancy filtering was performed by removing the minimal number of features under marginal correlation threshold  $\tau$ , which was set  $\tau=0.95$ . Next, features were scaled (centered around 0 and variance 1) to avoid a situation where the features with the largest scale dominate the analysis. A regularized estimator of the correlation matrix between the scaled features was obtained, and a maximum likelihood factor analysis was performed on the matrix. The number of latent features was determined using the Guttman-Kaiser rule<sup>50</sup> on the regularized correlation matrix. Factor scores were obtained by regressing the latent features on the observed data by way of the obtained factor solution. All steps were carried out using the R package “FMradio: Factor modeling for radiomic data”, version 1.1.1 (<https://CRAN.R-project.org/package=FMradio>)<sup>30</sup>.

### Prognostic model training

The resulting factor scores were used as predictors for a Cox model built on the training sets: (1) VUMC OSCC, and (2) VUMC OPSCC. Subsequently, the parameter matrices of the factor analysis performed on the training sets were used to construct factor scores for the samples in the validation sets: (1) UMCU OSCC, and (2) UMCU OPSCC, and these factor scores were used as predictors in a Cox model for the validation sets. Models were built to predict overall and disease-free survival. Performance of all models was assessed in terms of the (integrated) area-under-the-curve (iAUC) with a confidence interval that was assessed by bootstrapping (1000 times). Moreover, the patients of the validation sets were divided into low-risk and high-risk groups based on the median of the model scores and a log-rank test was performed. The radiomics only models were compared to a clinical-variables only model and a model holding both the radiomic features and clinical-variables. Clinical models were based on those features that were uniformly available and for which, *a priori*, some predictive power is expected: i.e. N-stage, age at diagnosis, and gender. T-stage and ACE-27 score held no predictive power in this dataset. Additionally, the radiomics only models were compared to tumor volume alone as proposed by Welch et al<sup>51</sup>. The iAUC of the combined model (radiomics + clinical) was compared to the iAUC of the individual prognostic models (radiomics only + clinical only) with the Wilcoxon rank sum test for dependent samples<sup>52</sup>, and a multiplicity correction was performed using the Holm method. P-values of less than 0.05 were considered statistically significant.

**Supplementary Table 1.** MRI protocols of included patients

Patient ID	Manufacturer	Magnetic field strength	T1 Repetition time	T1 Echo time	T1 x.size mm	T1 y.size mm	T1 z.size mm	STIR Repetition time	STIR Echo time	STIR Inversion time	STIR x.size mm	STIR y.size mm	STIR z.size mm
OSCC_VUMC_1	GE	1.5	440	14	0.5	0.5	3.3	4740	27	150	0.5	0.5	3.3
OSCC_VUMC_2	GE	1.5	540	14	0.5	0.5	4.4	6600	32	160	0.5	0.5	4.4
OSCC_VUMC_3	GE	1.5	520	14	0.5	0.5	4.4	6600	32	160	0.5	0.5	4.4
OSCC_VUMC_4	SIEMENS	1.5	593	15	0.4	0.4	3.3	5540	134	150	0.4	0.4	3.3
OSCC_VUMC_5	SIEMENS	1.0	588	15	0.8	0.8	3.0	5565	30	150	0.8	0.8	3.0
OSCC_VUMC_6	GE	1.5	500	14	0.5	0.5	4.4	6600	32	160	0.5	0.5	4.4
OSCC_VUMC_7	GE	1.5	500	14	0.5	0.5	4.4	6600	33	160	0.5	0.5	4.4
OSCC_VUMC_8	SIEMENS	1.5	456	12	0.8	0.8	4.4	5000	70	150	0.4	0.4	6.5
OSCC_VUMC_9	GE	1.5	500	14	0.5	0.5	4.4	6600	32	160	0.5	0.5	4.4
OSCC_VUMC_10	SIEMENS	1.5	645	11	0.6	0.6	4.4	6050	41	160	0.4	0.4	4.8
OSCC_VUMC_11	SIEMENS	1.5	658	10	0.7	0.7	4.4	6050	41	160	0.4	0.4	4.8
OSCC_VUMC_12	TOSHIBA	3.0	753	10	0.7	0.7	3.3	9930	48	225	0.4	0.4	3.3
OSCC_VUMC_13	GE	1.5	500	14	0.5	0.5	4.4	6600	32	160	0.5	0.5	4.4
OSCC_VUMC_14	SIEMENS	1.5	488	14	0.5	0.5	3.3	7630	62	160	0.5	0.5	3.3
OSCC_VUMC_15	SIEMENS	1.0	540	15	0.8	0.8	3.0	5100	30	150	0.8	0.8	3.0
OSCC_VUMC_16	GE	1.5	460	13	0.5	0.5	3.3	5000	30	150	0.5	0.5	3.3
OSCC_VUMC_17	SIEMENS	1.5	645	11	0.6	0.6	4.4	6810	44	160	0.4	0.4	4.8
OSCC_VUMC_18	GE	1.5	520	14	0.5	0.5	4.4	6840	32	160	0.5	0.5	4.4
OSCC_VUMC_19	GE	1.5	500	14	0.5	0.5	4.4	6600	32	160	0.5	0.5	4.4
OSCC_VUMC_20	SIEMENS	1.5	645	11	0.6	0.6	4.4	6050	41	160	0.4	0.4	4.8
OSCC_VUMC_21	SIEMENS	1.5	580	11	0.6	0.6	4.4	6050	44	160	0.4	0.4	4.8
OSCC_VUMC_22	GE	1.5	540	14	0.5	0.5	4.4	6600	32	160	0.5	0.5	4.4
OSCC_VUMC_23	SIEMENS	1.0	644	15	0.8	0.8	3.0	6095	30	150	0.8	0.8	3.0
OSCC_VUMC_24	SIEMENS	1.0	540	15	0.8	0.8	3.0	5100	30	150	0.8	0.8	3.0
OSCC_VUMC_25	GE	1.5	480	14	0.5	0.5	4.4	6600	32	160	0.5	0.5	4.4
OSCC_VUMC_26	SIEMENS	1.5	522	14	0.4	0.4	3.3	5290	27	150	0.4	0.4	3.3
OSCC_VUMC_27	GE	3.0	620	17	0.5	0.5	3.3	6360	8	160	0.5	0.5	3.3
OSCC_VUMC_28	GE	1.5	500	14	0.5	0.5	4.4	6400	63	150	0.5	0.5	4.4
OSCC_VUMC_29	GE	1.5	500	14	0.5	0.5	4.4	6600	32	160	0.5	0.5	4.4
OSCC_VUMC_30	GE	1.5	500	14	0.5	0.5	4.4	6600	32	160	0.5	0.5	4.4
OSCC_VUMC_31	GE	1.5	500	14	0.5	0.5	4.4	6600	32	160	0.5	0.5	4.4
OSCC_VUMC_32	Philips	3.0	750	16	0.4	0.4	4.4	4534	50	180	0.4	0.4	4.4
OSCC_VUMC_33	GE	3.0	620	17	0.5	0.5	3.3	6360	8	160	0.5	0.5	3.3
OSCC_VUMC_34	GE	1.5	500	14	0.5	0.5	4.4	6600	32	160	0.5	0.5	4.4
OSCC_VUMC_35	SIEMENS	1.5	553	12	0.8	0.8	4.4	5000	70	150	0.4	0.4	6.5
OSCC_VUMC_36	GE	1.5	500	14	0.5	0.5	4.4	6600	32	160	0.5	0.5	4.4
OSCC_VUMC_37	Philips	3.0	666	16	0.4	0.4	4.4	5591	50	180	0.4	0.4	7.7
OSCC_VUMC_38	GE	1.5	500	14	0.5	0.5	4.4	6600	32	160	0.5	0.5	4.4
OSCC_VUMC_39	SIEMENS	1.5	600	15	0.9	0.9	3.0	6095	30	150	1.1	1.1	3.0
OSCC_VUMC_40	GE	1.5	500	14	0.5	0.5	4.4	6600	32	160	0.5	0.5	4.4
OSCC_VUMC_41	SIEMENS	1.5	521	14	0.4	0.4	3.3	5290	27	150	0.4	0.4	3.3

Supplementary Table 1. (continued)

Patient ID	Manufacturer	Magnetic field strength	T1 Repetition time	T1 Echo time	T1 x.size mm	T1 y.size mm	T1 z.size mm	STIR Repetition time	STIR Echo time	STIR Inversion time	STIR x.size mm	STIR y.size mm	STIR z.size mm
OSCC_VUMC_42	SIEMENS	1.5	522	14	0.4	0.4	3.3	5270	27	150	0.4	0.4	3.3
OSCC_VUMC_43	SIEMENS	1.5	598	14	0.5	0.5	3.5	7630	62	160	0.5	0.5	3.5
OSCC_VUMC_44	GE	1.5	640	14	0.5	0.5	3.3	6540	26	150	0.5	0.5	3.3
OSCC_VUMC_45	GE	1.5	500	14	0.5	0.5	4.4	6600	63	160	0.5	0.5	4.4
OSCC_VUMC_46	GE	1.5	500	14	0.5	0.5	4.4	6600	32	160	0.5	0.5	4.4
OSCC_VUMC_47	GE	1.5	500	14	0.6	0.6	4.4	6600	32	160	0.6	0.6	4.4
OSCC_VUMC_48	GE	1.5	500	14	0.5	0.5	4.4	6600	32	160	0.5	0.5	4.4
OSCC_VUMC_49	GE	1.5	540	14	0.5	0.5	4.4	6600	32	160	0.5	0.5	4.4
OSCC_VUMC_50	SIEMENS	1.5	645	11	0.6	0.6	4.4	6050	44	160	0.4	0.4	4.8
OSCC_VUMC_51	SIEMENS	1.5	645	11	0.6	0.6	4.4	7460	41	160	0.4	0.4	4.8
OSCC_VUMC_52	SIEMENS	1.5	645	11	0.6	0.6	4.4	4260	40	160	0.4	0.4	4.8
OSCC_VUMC_53	GE	1.5	460	13	0.5	0.5	3.3	6540	26	150	0.5	0.5	3.3
OSCC_VUMC_54	GE	1.5	500	14	0.5	0.5	4.4	6600	32	160	0.5	0.5	4.4
OSCC_VUMC_55	SIEMENS	1.5	671	14	0.6	0.6	3.6	6770	27	150	0.5	0.5	3.6
OSCC_VUMC_56	GE	3.0	620	17	0.5	0.5	3.3	6760	37	160	0.5	0.5	3.3
OSCC_VUMC_57	SIEMENS	1.0	750	15	0.8	0.8	3.0	7000	30	150	0.8	0.8	3.0
OSCC_VUMC_58	GE	1.5	500	14	0.5	0.5	4.4	6600	32	160	0.5	0.5	4.4
OSCC_VUMC_59	GE	1.5	460	13	0.5	0.5	3.3	6540	26	150	0.5	0.5	3.3
OSCC_VUMC_60	SIEMENS	1.5	600	12	0.8	0.8	4.4	7860	16	150	0.4	0.4	4.4
OSCC_VUMC_61	GE	1.5	500	14	0.5	0.5	4.4	6600	32	160	0.5	0.5	4.4
OSCC_VUMC_62	GE	1.5	500	14	0.5	0.5	4.4	6600	32	160	0.5	0.5	4.4
OSCC_VUMC_63	SIEMENS	1.5	645	10	0.7	0.7	4.4	6660	41	160	0.4	0.4	4.8
OSCC_VUMC_64	SIEMENS	1.5	645	11	0.6	0.6	4.4	6050	41	160	0.4	0.4	4.8
OSCC_VUMC_65	GE	1.5	460	13	0.5	0.5	3.3	6540	26	150	0.5	0.5	3.3
OSCC_VUMC_66	SIEMENS	1.0	600	15	0.9	0.9	3.0	5700	30	150	0.9	0.9	3.0
OSCC_VUMC_67	SIEMENS	1.5	465	14	0.5	0.5	3.3	7630	62	160	0.5	0.5	3.3
OSCC_VUMC_68	GE	1.5	520	14	0.5	0.5	4.4	6600	63	160	0.5	0.5	4.4
OSCC_VUMC_69	SIEMENS	1.5	488	14	0.7	0.7	4.4	6670	62	160	0.5	0.5	4.4
OSCC_VUMC_70	SIEMENS	1.5	546	14	0.4	0.4	4.4	5520	27	150	0.4	0.4	4.4
OSCC_VUMC_71	Philips	3.0	666	16	0.4	0.4	4.4	9415	50	180	0.4	0.4	4.4
OSCC_VUMC_72	GE	1.5	520	14	0.5	0.5	4.4	6600	63	160	0.5	0.5	4.4
OSCC_VUMC_73	SIEMENS	1.5	595	14	0.4	0.4	3.3	6300	27	150	0.4	0.4	3.3
OSCC_VUMC_74	GE	1.5	500	14	0.5	0.5	4.4	6600	32	160	0.5	0.5	4.4
OSCC_VUMC_75	SIEMENS	1.5	600	15	0.8	0.8	3.0	5600	30	150	1.1	1.1	3.0
OSCC_VUMC_76	SIEMENS	1.0	540	15	0.9	0.9	3.0	5100	30	150	0.9	0.9	3.0
OSCC_VUMC_77	SIEMENS	1.5	465	14	0.5	0.5	3.3	7630	62	160	0.5	0.5	3.3
OSCC_VUMC_78	SIEMENS	1.5	621	14	0.4	0.4	3.3	6270	27	150	0.4	0.4	3.3
OSCC_VUMC_79	SIEMENS	1.5	553	14	0.5	0.5	3.3	7630	62	160	0.5	0.5	3.3
OSCC_VUMC_80	SIEMENS	1.5	488	14	0.5	0.5	3.3	7630	62	160	0.5	0.5	3.3
OSCC_VUMC_81	GE	1.5	480	14	0.5	0.5	3.3	4940	33	150	1.0	1.0	3.3
OSCC_VUMC_82	GE	1.5	500	14	0.5	0.5	4.4	6600	32	160	0.5	0.5	4.4

**Supplementary Table 1.** (continued)

Patient ID	Manufacturer	Magnetic field strength	T1 Repetition time	T1 Echo time	T1 x.size mm	T1 y.size mm	T1 z.size mm	STIR Repetition time	STIR Echo time	STIR Inversion time	STIR x.size mm	STIR y.size mm	STIR z.size mm
OSCC_VUMC_83	SIEMENS	1.0	540	15	0.8	0.8	3.0	5100	30	150	0.8	0.8	3.0
OSCC_VUMC_84	SIEMENS	1.5	603	15	0.4	0.4	3.3	5540	134	150	0.4	0.4	3.3
OSCC_VUMC_85	Philips	1.0	561	14	0.4	0.4	4.4	4198	90	-1	0.4	0.4	4.4
OSCC_VUMC_86	SIEMENS	1.5	522	14	0.4	0.4	3.3	5290	27	150	0.4	0.4	3.3
OSCC_VUMC_87	SIEMENS	1.5	522	14	0.4	0.4	3.3	5290	27	150	0.4	0.4	3.3
OSCC_VUMC_88	GE	1.5	540	14	0.5	0.5	4.4	6600	32	160	0.5	0.5	4.4
OSCC_VUMC_89	SIEMENS	1.5	571	14	0.4	0.4	3.5	5770	27	150	0.4	0.4	3.5
OSCC_VUMC_90	GE	1.5	540	14	0.5	0.5	4.4	6600	32	160	0.5	0.5	4.4
OSCC_VUMC_91	SIEMENS	1.0	540	15	0.8	0.8	4.0	5100	30	150	0.8	0.8	4.0
OSCC_VUMC_92	SIEMENS	1.0	650	15	0.8	0.8	3.0	6100	30	150	0.8	0.8	3.0
OSCC_VUMC_93	SIEMENS	1.0	540	15	0.8	0.8	3.0	5100	30	150	0.8	0.8	3.0
OSCC_VUMC_94	SIEMENS	1.5	645	11	0.6	0.6	4.4	4240	40	160	0.4	0.4	4.8
OSCC_VUMC_95	GE	1.5	500	14	0.5	0.5	4.4	6600	32	160	0.5	0.5	4.4
OSCC_VUMC_96	GE	1.5	520	14	0.5	0.5	4.4	5500	33	150	0.5	0.5	4.4
OSCC_VUMC_97	GE	1.5	500	14	0.5	0.5	4.4	6600	63	160	0.5	0.5	4.4
OSCC_VUMC_98	SIEMENS	1.5	521	14	0.4	0.4	3.3	5290	27	150	0.4	0.4	3.3
OSCC_VUMC_99	SIEMENS	1.5	456	12	0.8	0.8	4.4	5000	70	150	0.4	0.4	6.5
OSCC_VUMC_100	GE	1.5	540	14	0.5	0.5	4.4	6600	32	160	0.5	0.5	4.4
OSCC_VUMC_101	GE	1.5	460	13	0.5	0.5	3.3	6540	32	150	1.0	1.0	3.3
OSCC_VUMC_102	GE	1.5	540	14	0.5	0.5	4.4	6600	32	160	0.5	0.5	4.4
UMCU_OSCC_1	Philips	3.0	653	16	0.6	0.6	4.0	4868	25	220	0.6	0.6	4.0
UMCU_OSCC_2	Philips	1.5	583	15	0.6	0.6	4.0	2134	24	165	0.5	0.5	4.0
UMCU_OSCC_3	Philips	1.5	589	15	0.6	0.6	4.0	1733	24	165	0.5	0.5	4.0
UMCU_OSCC_4	Philips	1.5	515	15	0.6	0.6	4.0	1867	24	165	0.5	0.5	4.0
UMCU_OSCC_5	Philips	3.0	653	16	0.6	0.6	4.0	4868	25	220	0.6	0.6	4.0
UMCU_OSCC_6	Philips	1.5	593	15	0.5	0.5	4.4	4200	130	-1	0.5	0.5	5.5
UMCU_OSCC_7	Philips	1.5	596	15	0.6	0.6	4.0	1730	23	165	0.5	0.5	4.0
UMCU_OSCC_8	Philips	1.5	591	15	0.6	0.6	4.0	1733	24	165	0.5	0.5	4.0
UMCU_OSCC_9	Philips	1.5	593	15	0.6	0.6	4.0	2131	24	165	0.5	0.5	4.0
UMCU_OSCC_10	Philips	1.5	652	14	0.4	0.4	4.4	3097	9	165	0.8	0.8	4.4
UMCU_OSCC_11	Philips	1.5	517	15	0.6	0.6	4.0	1998	24	165	0.5	0.5	4.0
UMCU_OSCC_12	Philips	1.5	649	14	0.4	0.4	4.4	3099	9	165	0.8	0.8	4.4
UMCU_OSCC_13	Philips	3.0	653	16	0.6	0.6	4.0	4868	25	220	0.6	0.6	4.0
UMCU_OSCC_14	Philips	1.5	588	15	0.6	0.6	5.0	1734	24	165	0.5	0.5	5.0
UMCU_OSCC_15	Philips	1.5	592	15	0.6	0.6	4.0	2132	24	165	0.5	0.5	4.0
UMCU_OSCC_16	Philips	1.5	652	14	0.4	0.4	4.4	3094	9	165	0.8	0.8	4.4
UMCU_OSCC_17	Philips	1.5	590	15	0.6	0.6	4.0	1734	24	165	0.5	0.5	4.0
UMCU_OSCC_18	Philips	1.5	515	15	0.6	0.6	4.0	1867	24	165	0.5	0.5	4.0
UMCU_OSCC_19	Philips	1.5	591	15	0.6	0.6	4.0	1733	24	165	0.5	0.5	4.0
UMCU_OSCC_20	Philips	3.0	653	16	0.6	0.6	4.0	4868	25	220	0.6	0.6	4.0
UMCU_OSCC_21	Philips	3.0	653	16	0.6	0.6	4.0	4868	25	220	0.6	0.6	4.0

Supplementary Table 1. (continued)

Patient ID	Manufacturer	Magnetic field strength	T1 Repetition time	T1 Echo time	T1 x.size mm	T1 y.size mm	T1 z.size mm	STIR Repetition time	STIR Echo time	STIR Inversion time	STIR x.size mm	STIR y.size mm	STIR z.size mm
UMCU_OSCC_22	Philips	3.0	653	16	0.6	0.6	4.0	4868	25	220	0.6	0.6	4.0
UMCU_OSCC_23	Philips	1.5	652	14	0.4	0.4	4.4	3098	9	165	0.8	0.8	4.4
UMCU_OSCC_24	Philips	3.0	653	16	0.6	0.6	4.0	4868	25	220	0.6	0.6	4.0
UMCU_OSCC_25	Philips	1.5	652	14	0.4	0.4	4.4	3098	9	165	0.8	0.8	4.4
UMCU_OSCC_26	Philips	3.0	653	16	0.6	0.6	4.0	4868	25	220	0.6	0.6	4.0
UMCU_OSCC_27	Philips	3.0	653	16	0.6	0.6	4.0	4868	25	220	0.6	0.6	4.0
UMCU_OSCC_28	Philips	3.0	653	16	0.6	0.6	4.0	4868	25	220	0.6	0.6	4.0
UMCU_OSCC_29	Philips	3.0	653	16	0.6	0.6	4.0	4868	25	220	0.6	0.6	4.0
UMCU_OSCC_31	Philips	1.5	652	14	0.4	0.4	4.4	3098	9	165	0.8	0.8	4.4
UMCU_OSCC_33	Philips	1.5	590	15	0.6	0.6	4.0	1734	24	165	0.5	0.5	4.0
UMCU_OSCC_34	Philips	1.5	590	15	0.6	0.6	4.0	1734	24	165	0.5	0.5	4.0
UMCU_OSCC_35	Philips	1.5	590	15	0.6	0.6	4.0	1734	24	165	0.5	0.5	4.0
UMCU_OSCC_36	Philips	1.5	591	15	0.6	0.6	4.0	1733	24	165	0.5	0.5	4.0
UMCU_OSCC_37	Philips	1.5	590	15	0.6	0.6	4.0	1734	24	165	0.5	0.5	4.0
UMCU_OSCC_38	Philips	1.5	652	14	0.4	0.4	4.4	3098	9	165	0.8	0.8	4.4
UMCU_OSCC_39	Philips	1.5	652	14	0.4	0.4	4.4	3098	9	165	0.8	0.8	4.4
UMCU_OSCC_40	Philips	1.5	652	14	0.4	0.4	4.4	3098	9	165	0.8	0.8	4.4
UMCU_OSCC_41	Philips	3.0	666	16	0.6	0.6	4.0	4868	25	220	0.6	0.6	4.0
UMCU_OSCC_42	Philips	3.0	666	16	0.6	0.6	4.0	4868	25	220	0.6	0.6	4.0
UMCU_OSCC_43	Philips	1.5	587	15	0.6	0.6	4.0	1660	25	165	0.5	0.5	4.0
UMCU_OSCC_44	Philips	1.5	584	15	0.6	0.6	4.0	1660	25	165	0.5	0.5	4.0
UMCU_OSCC_45	Philips	1.5	587	15	0.6	0.6	4.0	1660	25	165	0.5	0.5	4.0
UMCU_OSCC_46	Philips	3.0	666	16	0.6	0.6	4.0	4868	25	220	0.6	0.6	4.0
UMCU_OSCC_47	Philips	1.5	650	14	0.4	0.4	6.2	6047	9	165	0.7	0.7	6.2
UMCU_OSCC_48	Philips	3.0	666	16	0.6	0.6	4.0	4868	25	220	0.6	0.6	4.0
UMCU_OSCC_49	Philips	1.5	591	15	0.6	0.6	4.0	1656	26	165	0.5	0.5	4.0
UMCU_OSCC_50	Philips	1.5	652	14	0.4	0.4	4.4	3098	9	165	0.8	0.8	4.4
UMCU_OSCC_51	Philips	1.5	652	14	0.4	0.4	4.4	3098	9	165	0.8	0.8	4.4
UMCU_OSCC_52	Philips	1.5	589	15	0.6	0.6	4.0	1658	25	165	0.5	0.5	4.0
UMCU_OSCC_53	Philips	1.5	595	15	0.6	0.6	4.2	2130	24	165	0.5	0.5	4.2
UMCU_OSCC_54	Philips	1.5	652	14	0.4	0.4	4.4	3097	9	165	0.8	0.8	4.4
UMCU_OSCC_55	Philips	1.5	652	14	0.4	0.4	4.4	3098	9	165	0.8	0.8	4.4
UMCU_OSCC_56	Philips	1.5	587	15	0.6	0.6	4.0	1657	26	165	0.5	0.5	4.0
UMCU_OSCC_57	Philips	1.5	652	14	0.4	0.4	4.4	3097	9	165	0.8	0.8	4.4
UMCU_OSCC_58	Philips	1.5	590	15	0.6	0.6	4.0	1734	24	165	0.5	0.5	4.0
UMCU_OSCC_59	Philips	1.5	588	15	0.6	0.6	4.0	1734	24	165	0.5	0.5	4.0
UMCU_OSCC_60	Philips	1.5	652	14	0.4	0.4	4.4	3096	9	165	0.8	0.8	4.4
UMCU_OSCC_61	Philips	1.5	594	15	0.6	0.6	4.0	1779	27	165	0.5	0.5	4.0
UMCU_OSCC_62	Philips	3.0	653	16	0.6	0.6	4.0	4868	25	220	0.6	0.6	4.0
UMCU_OSCC_63	Philips	3.0	653	16	0.6	0.6	4.0	4868	25	220	0.6	0.6	4.0
UMCU_OSCC_65	Philips	1.5	589	15	0.6	0.6	4.0	1733	24	165	0.5	0.5	4.0

Supplementary Table 1. (continued)

Patient ID	Manufacturer	Magnetic field strength	T1 Repetition time	T1 Echo time	T1 x.size mm	T1 y.size mm	T1 z.size mm	STIR Repetition time	STIR Echo time	STIR Inversion time	STIR x.size mm	STIR y.size mm	STIR z.size mm
UMCU_OSCC_67	Philips	1.5	652	14	0.4	0.4	4.4	3098	9	165	0.8	0.8	4.4
UMCU_OSCC_68	Philips	1.5	590	15	0.6	0.6	4.0	1734	24	165	0.5	0.5	4.0
UMCU_OSCC_71	Philips	1.5	652	14	0.4	0.4	4.4	3094	9	165	0.7	0.7	4.4
UMCU_OSCC_72	Philips	1.5	590	15	0.6	0.6	4.0	1734	24	165	0.5	0.5	4.0
UMCU_OSCC_73	Philips	1.5	590	15	0.6	0.6	4.0	1734	24	165	0.5	0.5	4.0
UMCU_OSCC_74	Philips	1.5	652	14	0.4	0.4	4.4	3098	9	165	0.7	0.7	4.4
UMCU_OSCC_75	Philips	1.5	590	15	0.6	0.6	4.0	1734	24	165	0.5	0.5	4.0
UMCU_OSCC_77	Philips	1.5	590	15	0.6	0.6	4.0	1734	24	165	0.5	0.5	4.0
UMCU_OSCC_78	Philips	1.5	652	14	0.4	0.4	4.4	3098	9	165	0.8	0.8	4.4
UMCU_OSCC_79	Philips	1.5	590	15	0.6	0.6	5.0	2134	24	165	0.5	0.5	5.0
UMCU_OSCC_80	Philips	1.5	590	15	0.6	0.6	4.0	1734	24	165	0.5	0.5	4.0
UMCU_OSCC_82	Philips	1.5	517	15	0.6	0.6	4.0	1867	24	165	0.5	0.5	4.0
UMCU_OSCC_83	Philips	3.0	653	16	0.4	0.4	3.3	4425	25	220	0.4	0.4	3.3
UMCU_OSCC_84	Philips	1.5	588	15	0.6	0.6	4.0	1734	24	165	0.5	0.5	4.0
OPSCC_VUMC_1	GE	3.0	580	16	0.5	0.5	3.3	NA	NA	NA	NA	NA	NA
OPSCC_VUMC_2	GE	1.5	500	14	0.5	0.5	4.4	NA	NA	NA	NA	NA	NA
OPSCC_VUMC_3	GE	1.5	460	13	0.5	0.5	3.3	NA	NA	NA	NA	NA	NA
OPSCC_VUMC_4	GE	1.5	500	14	0.5	0.5	4.4	NA	NA	NA	NA	NA	NA
OPSCC_VUMC_5	GE	1.5	560	14	0.5	0.5	4.4	NA	NA	NA	NA	NA	NA
OPSCC_VUMC_6	GE	1.5	620	14	0.5	0.5	4.4	NA	NA	NA	NA	NA	NA
OPSCC_VUMC_7	Philips	3.0	666	16	0.4	0.4	4.4	NA	NA	NA	NA	NA	NA
OPSCC_VUMC_8	GE	1.5	620	14	0.5	0.5	4.4	NA	NA	NA	NA	NA	NA
OPSCC_VUMC_9	GE	1.5	500	14	0.5	0.5	4.4	NA	NA	NA	NA	NA	NA
OPSCC_VUMC_10	GE	1.5	720	14	0.5	0.5	4.4	NA	NA	NA	NA	NA	NA
OPSCC_VUMC_11	SIEMENS	1.5	522	14	0.4	0.4	3.3	NA	NA	NA	NA	NA	NA
OPSCC_VUMC_12	SIEMENS	1.5	390	14	0.9	0.9	4.4	NA	NA	NA	NA	NA	NA
OPSCC_VUMC_13	GE	1.5	480	14	0.5	0.5	3.3	NA	NA	NA	NA	NA	NA
OPSCC_VUMC_14	GE	1.5	480	14	0.5	0.5	4.4	NA	NA	NA	NA	NA	NA
OPSCC_VUMC_15	GE	1.5	500	14	0.5	0.5	4.4	NA	NA	NA	NA	NA	NA
OPSCC_VUMC_16	GE	1.5	500	14	0.5	0.5	4.4	NA	NA	NA	NA	NA	NA
OPSCC_VUMC_17	GE	1.5	460	13	0.5	0.5	3.3	NA	NA	NA	NA	NA	NA
OPSCC_VUMC_18	GE	1.5	600	14	0.5	0.5	4.4	NA	NA	NA	NA	NA	NA
OPSCC_VUMC_19	GE	1.5	620	14	0.5	0.5	6.5	NA	NA	NA	NA	NA	NA
OPSCC_VUMC_20	SIEMENS	1.5	546	14	0.4	0.4	4.4	NA	NA	NA	NA	NA	NA
OPSCC_VUMC_21	GE	3.0	600	16	0.5	0.5	3.3	NA	NA	NA	NA	NA	NA
OPSCC_VUMC_22	GE	1.5	440	14	0.5	0.5	3.3	NA	NA	NA	NA	NA	NA
OPSCC_VUMC_23	GE	1.5	420	13	0.5	0.5	3.3	NA	NA	NA	NA	NA	NA
OPSCC_VUMC_24	GE	3.0	720	16	0.5	0.5	3.3	NA	NA	NA	NA	NA	NA
OPSCC_VUMC_25	SIEMENS	1.5	492	14	0.4	0.4	4.4	NA	NA	NA	NA	NA	NA
OPSCC_VUMC_26	GE	1.5	620	14	0.5	0.5	6.5	NA	NA	NA	NA	NA	NA
OPSCC_VUMC_27	GE	1.5	640	14	0.5	0.5	6.5	NA	NA	NA	NA	NA	NA

Supplementary Table 1. (continued)

Patient ID	Manufacturer	Magnetic field strength	T1 Repetition time	T1 Echo time	T1 x.size mm	T1 y.size mm	T1 z.size mm	STIR Repetition time	STIR Echo time	STIR Inversion time	STIR x.size mm	STIR y.size mm	STIR z.size mm
OPSCC_VUMC_28	SIEMENS	1.5	390	14	0.9	0.9	4.4	NA	NA	NA	NA	NA	NA
OPSCC_VUMC_29	GE	1.5	500	14	0.5	0.5	4.4	NA	NA	NA	NA	NA	NA
OPSCC_VUMC_30	SIEMENS	1.5	546	14	0.4	0.4	4.4	NA	NA	NA	NA	NA	NA
OPSCC_VUMC_31	GE	3.0	600	16	0.5	0.5	3.3	NA	NA	NA	NA	NA	NA
OPSCC_VUMC_32	GE	3.0	640	18	0.5	0.5	3.3	NA	NA	NA	NA	NA	NA
OPSCC_VUMC_33	GE	1.5	640	14	0.5	0.5	4.4	NA	NA	NA	NA	NA	NA
OPSCC_VUMC_34	GE	1.5	540	13	0.5	0.5	3.3	NA	NA	NA	NA	NA	NA
OPSCC_VUMC_35	GE	1.5	500	14	0.5	0.5	4.4	NA	NA	NA	NA	NA	NA
OPSCC_VUMC_36	SIEMENS	1.5	571	14	0.4	0.4	3.3	NA	NA	NA	NA	NA	NA
OPSCC_VUMC_37	SIEMENS	1.5	522	14	0.4	0.4	3.3	NA	NA	NA	NA	NA	NA
OPSCC_VUMC_38	GE	1.5	500	14	0.5	0.5	4.4	NA	NA	NA	NA	NA	NA
OPSCC_VUMC_39	GE	1.5	500	14	0.5	0.5	4.4	NA	NA	NA	NA	NA	NA
OPSCC_VUMC_40	GE	1.5	500	14	0.5	0.5	4.4	NA	NA	NA	NA	NA	NA
OPSCC_VUMC_41	GE	1.5	620	16	0.5	0.5	3.3	NA	NA	NA	NA	NA	NA
OPSCC_VUMC_42	SIEMENS	1.5	465	14	0.6	0.6	3.3	NA	NA	NA	NA	NA	NA
OPSCC_VUMC_43	SIEMENS	1.5	671	14	0.5	0.5	3.3	NA	NA	NA	NA	NA	NA
OPSCC_VUMC_44	GE	3.0	560	17	0.5	0.5	3.3	NA	NA	NA	NA	NA	NA
OPSCC_VUMC_45	GE	1.5	460	13	0.5	0.5	3.3	NA	NA	NA	NA	NA	NA
OPSCC_VUMC_46	GE	1.5	500	14	0.5	0.5	4.4	NA	NA	NA	NA	NA	NA
OPSCC_VUMC_47	GE	1.5	500	14	0.5	0.5	4.4	NA	NA	NA	NA	NA	NA
OPSCC_VUMC_48	SIEMENS	1.5	350	14	0.5	0.5	4.4	NA	NA	NA	NA	NA	NA
OPSCC_VUMC_49	GE	1.5	500	14	0.5	0.5	4.4	NA	NA	NA	NA	NA	NA
OPSCC_VUMC_50	GE	3.0	600	17	0.5	0.5	3.3	NA	NA	NA	NA	NA	NA
OPSCC_VUMC_51	SIEMENS	1.0	532	15	0.8	0.8	5.0	NA	NA	NA	NA	NA	NA
OPSCC_VUMC_52	GE	1.5	740	16	0.5	0.5	3.8	NA	NA	NA	NA	NA	NA
OPSCC_VUMC_53	GE	1.5	620	14	0.5	0.5	4.4	NA	NA	NA	NA	NA	NA
OPSCC_VUMC_54	GE	1.5	500	14	0.5	0.5	4.4	NA	NA	NA	NA	NA	NA
OPSCC_VUMC_55	GE	1.5	480	14	0.5	0.5	4.4	NA	NA	NA	NA	NA	NA
OPSCC_VUMC_56	GE	1.5	500	14	0.5	0.5	4.4	NA	NA	NA	NA	NA	NA
OPSCC_VUMC_57	GE	1.5	500	14	0.5	0.5	4.4	NA	NA	NA	NA	NA	NA
OPSCC_VUMC_58	GE	1.5	500	14	0.5	0.5	4.4	NA	NA	NA	NA	NA	NA
OPSCC_VUMC_59	GE	3.0	620	17	0.5	0.5	3.3	NA	NA	NA	NA	NA	NA
OPSCC_VUMC_60	GE	1.5	500	14	0.5	0.5	4.4	NA	NA	NA	NA	NA	NA
OPSCC_VUMC_61	GE	3.0	560	15	0.5	0.5	6.0	NA	NA	NA	NA	NA	NA
OPSCC_VUMC_62	GE	3.0	680	17	0.5	0.5	3.3	NA	NA	NA	NA	NA	NA
OPSCC_VUMC_63	SIEMENS	1.5	522	14	0.4	0.4	3.3	NA	NA	NA	NA	NA	NA
OPSCC_VUMC_64	GE	3.0	600	16	0.5	0.5	3.3	NA	NA	NA	NA	NA	NA
OPSCC_VUMC_65	GE	1.5	460	13	0.5	0.5	3.3	NA	NA	NA	NA	NA	NA
OPSCC_VUMC_66	SIEMENS	1.5	488	14	0.7	0.7	3.3	NA	NA	NA	NA	NA	NA
OPSCC_VUMC_67	GE	1.5	460	13	0.5	0.5	3.3	NA	NA	NA	NA	NA	NA
OPSCC_VUMC_68	GE	1.5	500	14	0.5	0.5	4.4	NA	NA	NA	NA	NA	NA

Supplementary Table 1. (continued)

Patient ID	Manufacturer	Magnetic field strength	T1 Repetition time	T1 Echo time	T1 x.size mm	T1 y.size mm	T1 z.size mm	STIR Repetition time	STIR Echo time	STIR Inversion time	STIR x.size mm	STIR y.size mm	STIR z.size mm
OPSCC_VUMC_69	SIEMENS	1.5	488	14	0.5	0.5	3.3	NA	NA	NA	NA	NA	NA
OPSCC_VUMC_70	GE	1.5	620	14	0.5	0.5	4.4	NA	NA	NA	NA	NA	NA
OPSCC_VUMC_71	GE	3.0	540	17	0.5	0.5	6.0	NA	NA	NA	NA	NA	NA
OPSCC_VUMC_72	GE	1.5	500	14	0.5	0.5	4.4	NA	NA	NA	NA	NA	NA
OPSCC_VUMC_73	SIEMENS	1.5	522	14	0.4	0.4	3.3	NA	NA	NA	NA	NA	NA
OPSCC_VUMC_74	GE	1.5	460	13	0.5	0.5	3.3	NA	NA	NA	NA	NA	NA
OPSCC_VUMC_75	GE	1.5	480	14	0.5	0.5	3.3	NA	NA	NA	NA	NA	NA
OPSCC_VUMC_76	GE	3.0	800	16	0.5	0.5	4.4	NA	NA	NA	NA	NA	NA
OPSCC_VUMC_77	GE	3.0	780	16	0.5	0.5	3.3	NA	NA	NA	NA	NA	NA
OPSCC_VUMC_78	GE	3.0	600	16	0.5	0.5	3.3	NA	NA	NA	NA	NA	NA
OPSCC_VUMC_79	GE	3.0	600	17	0.5	0.5	3.3	NA	NA	NA	NA	NA	NA
OPSCC_VUMC_80	SIEMENS	1.5	488	14	0.7	0.7	3.3	NA	NA	NA	NA	NA	NA
OPSCC_VUMC_81	GE	1.5	500	14	0.5	0.5	4.4	NA	NA	NA	NA	NA	NA
OPSCC_VUMC_82	GE	1.5	500	14	0.5	0.5	4.4	NA	NA	NA	NA	NA	NA
OPSCC_VUMC_83	GE	1.5	500	14	0.5	0.5	4.4	NA	NA	NA	NA	NA	NA
OPSCC_VUMC_84	GE	1.5	500	14	0.5	0.5	4.4	NA	NA	NA	NA	NA	NA
OPSCC_VUMC_85	GE	1.5	460	13	0.5	0.5	3.3	NA	NA	NA	NA	NA	NA
OPSCC_VUMC_86	GE	1.5	720	14	0.5	0.5	4.4	NA	NA	NA	NA	NA	NA
OPSCC_VUMC_87	GE	1.5	500	14	0.5	0.5	4.4	NA	NA	NA	NA	NA	NA
OPSCC_VUMC_88	GE	1.5	500	14	0.5	0.5	4.4	NA	NA	NA	NA	NA	NA
OPSCC_VUMC_89	SIEMENS	1.5	547	14	0.6	0.6	4.4	NA	NA	NA	NA	NA	NA
OPSCC_UMCU_1	Philips	3.0	594	16	0.5	0.5	4.4	NA	NA	NA	NA	NA	NA
OPSCC_UMCU_2	Philips	3.0	594	16	0.5	0.5	4.4	NA	NA	NA	NA	NA	NA
OPSCC_UMCU_3	Philips	1.5	517	15	0.6	0.6	5.4	NA	NA	NA	NA	NA	NA
OPSCC_UMCU_4	Philips	3.0	594	16	0.5	0.5	4.4	NA	NA	NA	NA	NA	NA
OPSCC_UMCU_6	Philips	3.0	594	16	0.5	0.5	4.4	NA	NA	NA	NA	NA	NA
OPSCC_UMCU_7	Philips	3.0	594	16	0.5	0.5	4.4	NA	NA	NA	NA	NA	NA
OPSCC_UMCU_8	Philips	3.0	594	16	0.5	0.5	4.4	NA	NA	NA	NA	NA	NA
OPSCC_UMCU_9	Philips	3.0	594	16	0.5	0.5	4.4	NA	NA	NA	NA	NA	NA
OPSCC_UMCU_10	Philips	3.0	594	16	0.5	0.5	4.4	NA	NA	NA	NA	NA	NA
OPSCC_UMCU_11	Philips	3.0	594	16	0.5	0.5	4.4	NA	NA	NA	NA	NA	NA
OPSCC_UMCU_12	Philips	3.0	594	16	0.5	0.5	4.4	NA	NA	NA	NA	NA	NA
OPSCC_UMCU_13	Philips	3.0	697	15	0.5	0.5	4.4	NA	NA	NA	NA	NA	NA
OPSCC_UMCU_14	Philips	3.0	594	16	0.5	0.5	4.4	NA	NA	NA	NA	NA	NA
OPSCC_UMCU_17	Philips	3.0	594	16	0.5	0.5	4.4	NA	NA	NA	NA	NA	NA
OPSCC_UMCU_18	Philips	3.0	594	16	0.5	0.5	4.4	NA	NA	NA	NA	NA	NA
OPSCC_UMCU_19	Philips	3.0	594	16	0.5	0.5	4.4	NA	NA	NA	NA	NA	NA
OPSCC_UMCU_20	Philips	3.0	594	16	0.5	0.5	4.4	NA	NA	NA	NA	NA	NA
OPSCC_UMCU_21	Philips	3.0	594	16	0.5	0.5	4.4	NA	NA	NA	NA	NA	NA
OPSCC_UMCU_22	Philips	3.0	697	15	0.5	0.5	4.4	NA	NA	NA	NA	NA	NA
OPSCC_UMCU_23	Philips	3.0	594	16	0.5	0.5	4.4	NA	NA	NA	NA	NA	NA

Supplementary Table 1. (continued)

Patient ID	Manufacturer	Magnetic field strength	T1 Repetition time	T1 Echo time	T1 x.size mm	T1 y.size mm	T1 z.size mm	STIR Repetition time	STIR Echo time	STIR Inversion time	STIR x.size mm	STIR y.size mm	STIR z.size mm
OPSCC_UMCU_24	Philips	3.0	594	16	0.5	0.5	4.4	NA	NA	NA	NA	NA	NA
OPSCC_UMCU_25	Philips	3.0	594	16	0.5	0.5	4.4	NA	NA	NA	NA	NA	NA
OPSCC_UMCU_26	Philips	3.0	594	16	0.5	0.5	4.4	NA	NA	NA	NA	NA	NA
OPSCC_UMCU_29	Philips	1.5	593	15	0.5	0.5	4.4	NA	NA	NA	NA	NA	NA
OPSCC_UMCU_30	Philips	3.0	697	15	0.5	0.5	4.4	NA	NA	NA	NA	NA	NA
OPSCC_UMCU_31	Philips	3.0	697	15	0.5	0.5	4.4	NA	NA	NA	NA	NA	NA
OPSCC_UMCU_32	Philips	3.0	697	15	0.5	0.5	4.4	NA	NA	NA	NA	NA	NA
OPSCC_UMCU_33	Philips	1.5	593	15	0.5	0.5	4.4	NA	NA	NA	NA	NA	NA
OPSCC_UMCU_34	Philips	3.0	697	15	0.5	0.5	4.4	NA	NA	NA	NA	NA	NA
OPSCC_UMCU_35	Philips	1.5	593	15	0.5	0.5	4.4	NA	NA	NA	NA	NA	NA
OPSCC_UMCU_36	Philips	1.5	593	15	0.5	0.5	4.4	NA	NA	NA	NA	NA	NA
OPSCC_UMCU_37	Philips	1.5	593	15	0.5	0.5	4.4	NA	NA	NA	NA	NA	NA
OPSCC_UMCU_38	Philips	3.0	697	15	0.5	0.5	4.4	NA	NA	NA	NA	NA	NA
OPSCC_UMCU_39	Philips	3.0	413	15	0.5	0.5	4.4	NA	NA	NA	NA	NA	NA
OPSCC_UMCU_40	Philips	3.0	697	15	0.5	0.5	4.4	NA	NA	NA	NA	NA	NA
OPSCC_UMCU_42	Philips	3.0	697	15	0.5	0.5	4.4	NA	NA	NA	NA	NA	NA
OPSCC_UMCU_44	Philips	1.5	593	15	0.5	0.5	4.4	NA	NA	NA	NA	NA	NA
OPSCC_UMCU_45	Philips	1.5	593	15	0.5	0.5	4.4	NA	NA	NA	NA	NA	NA
OPSCC_UMCU_46	Philips	1.5	593	15	0.5	0.5	4.4	NA	NA	NA	NA	NA	NA
OPSCC_UMCU_47	Philips	1.5	593	15	0.5	0.5	4.4	NA	NA	NA	NA	NA	NA
OPSCC_UMCU_48	Philips	1.5	593	15	0.5	0.5	4.4	NA	NA	NA	NA	NA	NA
OPSCC_UMCU_50	Philips	1.5	593	15	0.5	0.5	4.4	NA	NA	NA	NA	NA	NA
OPSCC_UMCU_51	Philips	1.5	593	15	0.5	0.5	4.4	NA	NA	NA	NA	NA	NA
OPSCC_UMCU_52	Philips	3.0	697	15	0.5	0.5	4.4	NA	NA	NA	NA	NA	NA
OPSCC_UMCU_54	Philips	1.5	593	15	0.5	0.5	4.4	NA	NA	NA	NA	NA	NA
OPSCC_UMCU_56	Philips	1.5	593	15	0.5	0.5	4.4	NA	NA	NA	NA	NA	NA
OPSCC_UMCU_57	Philips	1.5	590	13	0.6	0.6	4.0	NA	NA	NA	NA	NA	NA
OPSCC_UMCU_58	Philips	1.5	517	15	0.6	0.6	4.0	NA	NA	NA	NA	NA	NA
OPSCC_UMCU_59	Philips	3.0	594	16	0.5	0.5	4.4	NA	NA	NA	NA	NA	NA
OPSCC_UMCU_61	Philips	1.5	588	15	0.6	0.6	4.0	NA	NA	NA	NA	NA	NA
OPSCC_UMCU_63	Philips	1.5	593	15	0.5	0.5	4.4	NA	NA	NA	NA	NA	NA
OPSCC_UMCU_64	SIEMENS	1.5	650	18	0.9	0.9	4.8	NA	NA	NA	NA	NA	NA
OPSCC_UMCU_65	Philips	1.5	593	15	0.5	0.5	4.4	NA	NA	NA	NA	NA	NA
OPSCC_UMCU_66	Philips	3.0	594	16	0.5	0.5	4.4	NA	NA	NA	NA	NA	NA
OPSCC_UMCU_67	Philips	3.0	594	16	0.5	0.5	4.4	NA	NA	NA	NA	NA	NA
OPSCC_UMCU_68	Philips	3.0	594	16	0.5	0.5	4.4	NA	NA	NA	NA	NA	NA

**Supplementary Table 2.** Factor loadings of OSCC model

Feature	Factor1	Factor2	Factor3	Factor4	Factor5	Factor6	Factor7
integrated intensity	0.61		0.42				
maximum 3D diameter	0.87						
surface area	0.86						
surface to volume ratio	-0.86					0.32	
radius of an equivolumetric sphere	0.94						
fractal dimension (calculated)	0.51						
fractal dimension (fitted)	0.66						
fractal abundance	0.94						
fractal lacunarity	-0.71						
Moran's I	-0.56			0.54			
Geary's C	0.55			-0.54			
entropy	-0.61				-0.36		
COOC first measure of information correlation	0.64			-0.54			
GLRL run length non-uniformity	0.8						
coefficient of variation		0.8			0.5		
COOC joint variance		0.94					
COOC joint entropy	0.46	0.64		-0.44			-0.33
COOC sum entropy		0.87					-0.32
COOC angular second moment		-0.64		0.47			0.39
COOC contrast		0.71		-0.46			0.34
COOC dissimilarity		0.7		-0.65			
COOC cluster tendency		0.93					
COOC cluster prominence		0.72					0.31
GLRL gray level non-uniformity normalized		-0.88					
GLRL run entropy	0.36	0.82					
gray level median from entire image			0.65				
gray level maximum of all values over 0.5 from entire image			0.87				
gray level mean of all values over 0.5 from entire image			0.91				
gray level median of all values over 0.5 from entire image			0.68				
gray level minimum from tumor ROI		-0.37	0.73				
gray level range from tumor ROI			0.89				
median absolute deviation of the median			0.9				
mean Laplacian			0.87				
energy	0.38		0.59				
variance			0.88				
Max <sub>star</sub>			0.94				
COOC inverse variance		-0.59		0.75			
COOC correlation				0.94			
COOC second measure of information correlation	-0.55			0.71			
GLRL gray level variance		-0.41		0.66			
skewness					0.51		
AUC					-0.73		
GLRL short run low gray level emphasis		0.39			0.78		
GLRL short run high gray level emphasis	-0.38	-0.35			-0.8		
GLRL long run low gray level emphasis				0.37	0.73		
GLRL long run high gray level emphasis		-0.49			-0.74		
compactness 2							-0.91
asphericity							0.94
excess kurtosis					0.39		
COOC cluster shade					0.41		

Abbreviations: COOC, co-occurrence; GLRL, grey-level run-length

**Supplementary Table 3.** Factor loadings of OPSCC model

Feature	Factor1	Factor2	Factor3	Factor4	Factor5	Factor6	Factor7
coefficient of variation	-0.69			0.57			
excess kurtosis	0.58			0.4			
entropy	-0.61	-0.48					
COOC joint entropy	-0.91				-0.33		
COOC sum entropy	-0.96						
COOC angular second moment	0.85				0.32		
COOC contrast	-0.65	-0.32					-0.57
COOC inverse difference normalized	0.69						0.53
COOC inverse variance	0.69	0.32					0.44
COOC cluster prominence	-0.73				0.41		
GLRL gray level non-uniformity normalized	0.9						
GLRL run length non-uniformity normalized	-0.68	-0.51					-0.32
GLRL gray level variance	-0.83						
GLRL run length variance	0.66	0.5			0.35		
GLRL run entropy	-0.88						
maximum 3D diameter		0.82					
surface area		0.89					
surface to volume ratio		-0.7			0.31	0.42	
radius of an equivolumetric sphere		0.91					
energy		0.72	0.31				
fractal dimension (fitted)		0.55			-0.32		
fractal abundance		0.85					
fractal lacunarity		-0.67					
GLRL gray level non-uniformity	0.31	0.87					
GLRL run length non-uniformity		0.9					
gray level median from entire image			0.54				
gray level maximum of all values over 0.5 from entire image			0.77				
gray level mean of all values over 0.5 from entire image			0.85				
gray level minimum from tumor ROI		-0.32	0.62				
gray level range from tumor ROI			0.82				
mean Laplacian			0.83				
variance			0.81				
root-mean-square			0.92				
Maxstar			0.89				
skewness				0.59			
AUC				-0.74			
GLRL short run low gray level emphasis				0.91			
GLRL short run high gray level emphasis				-0.92			
GLRL long run low gray level emphasis	0.4	0.42		0.69			
GLRL long run high gray level emphasis				-0.88			
Moran's I					0.7		
Geary's C					-0.68		
COOC first measure of information correlation		0.37			-0.81		
COOC second measure of information correlation		-0.3			0.82		
compactness 2						-0.86	
asphericity						0.91	
COOC correlation		0.32			0.54		0.61
gray level median of all values over 0.5 from entire image			0.49				
fractal dimension (calculated)		0.39					
COOC inverse difference moment normalized	0.44						0.48
COOC cluster shade				0.45			

Abbreviations: COOC, co-occurrence; GLRL, grey-level run-length

**Supplementary Table 4.** Comparison of iAUC of combined models to single modality models with Wilcoxon rank sum test for dependent samples and multiplicity correction using Holm method

		adjusted p-value
OSCC overall survival	Radiomics + Clinical > Radiomics	P<0.001
	Radiomics + Clinical > Clinical	P<0.001
OSCC relapse-free survival	Radiomics + Clinical > Radiomics	P<0.001
	Radiomics + Clinical > Clinical	P<0.001
OPSCC overall survival	Radiomics + Clinical > Radiomics	P<0.001
	Radiomics + Clinical > Clinical	P<0.001
OPSCC relapse-free survival	Radiomics + Clinical > Radiomics	P<0.001
	Radiomics + Clinical > Clinical	P<0.001

**Supplementary Table 5.** Influence of magnetic field strength on radiomics features as assessed by MANOVA and multiplicity correction using Holm method

Factor	VUMC OSCC	UMCU OSCC	VUMC OPSCC	UMCU OPSCC
Factor 1 - 3D geometrics	0.35	1.00	1.00	1.00
Factor 2 - meta-gray level co-occurrence	1.00	1.00	1.00	1.00
Factor 3 - meta-first order	0.36	1.00	<0.001	<0.001
Factor 4 - gray level-mix	<0.001	0.84	1.00	1.00
Factor 5 - meta-gray level run length	1.00	1.00	0.39	<0.001
Factor 6 - geometrics	1.00	1.00	0.58	1.00
Factor 7 - entropy	1.00	1.00	0.87	1.00

**Supplementary table 6.** Influence of MR vendor on radiomics features as assessed by MANOVA and multiplicity correction using Holm method

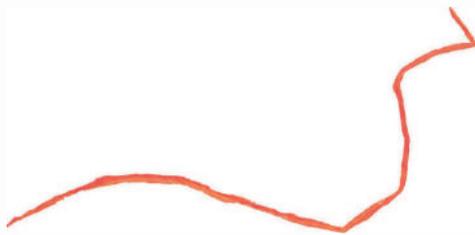
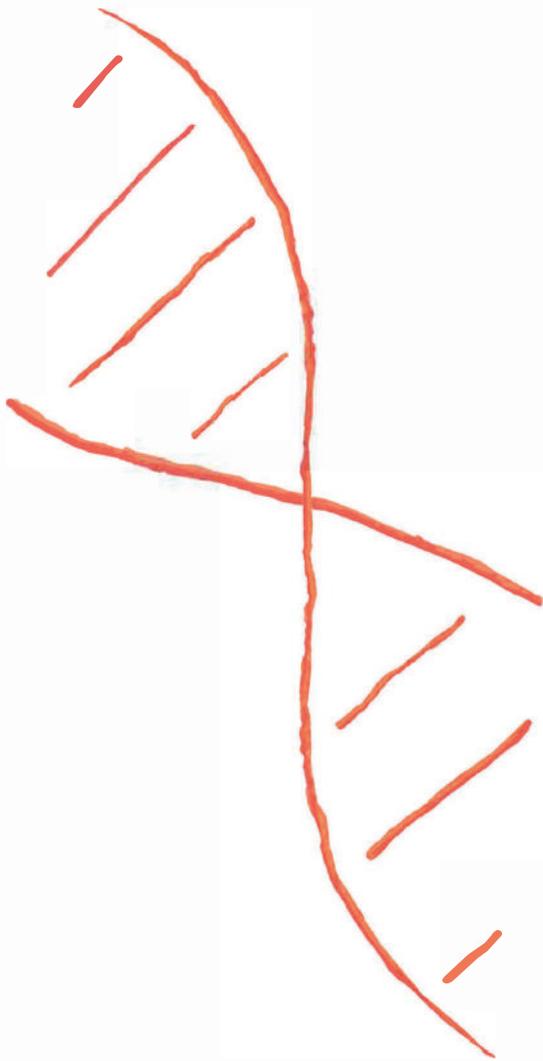
Factor	VUMC OSCC	VUMC OPSCC
Factor 1 - 3D geometrics	1.00	1.00
Factor 2 - meta-gray level co-occurrence	0.96	1.00
Factor 3 - meta-first order	<0.001	<0.001
Factor 4 - gray level-mix	<0.001	1.00
Factor 5 - meta-gray level run length	1.00	<0.01
Factor 6 - geometrics	1.00	1.00
Factor 7 - entropy	1.00	<0.01

## REFERENCES

1. Forastiere A, Koch W, Trotti A, Sidransky D. Head and neck cancer. *N Engl J Med* 2001; 345(26): 1890-900.
2. Surveillance E, and End Results (SEER) Program ([www.seer.cancer.gov](http://www.seer.cancer.gov)) Research Data (1975–2016), National Cancer Institute, DCCPS, Surveillance Research Program, released April 2019, based on the November 2018 submission.
3. Murphy BA, Deng J. Advances in Supportive Care for Late Effects of Head and Neck Cancer. *J Clin Oncol* 2015; 33(29): 3314-21.
4. Mes SW, Leemans CR, Brakenhoff RH. Applications of molecular diagnostics for personalized treatment of head and neck cancer: state of the art. *Expert Rev Mol Diagn* 2016; 16(2): 205-21.
5. Leemans CR, Snijders PJF, Brakenhoff RH. The molecular landscape of head and neck cancer. *Nat Rev Cancer* 2018; 18(5): 269-82.
6. Gevaert O, Xu J, Hoang CD, et al. Non-small cell lung cancer: identifying prognostic imaging biomarkers by leveraging public gene expression microarray data—methods and preliminary results. *Radiology* 2012; 264(2): 387-96.
7. O'Connor JP, Rose CJ, Waterton JC, Carano RA, Parker GJ, Jackson A. Imaging intratumor heterogeneity: role in therapy response, resistance, and clinical outcome. *Clin Cancer Res* 2015; 21(2): 249-57.
8. Kumar V, Gu Y, Basu S, et al. Radiomics: the process and the challenges. *Magn Reson Imaging* 2012; 30(9): 1234-48.
9. Aerts HJ, Velazquez ER, Leijenaar RT, et al. Decoding tumour phenotype by noninvasive imaging using a quantitative radiomics approach. *Nat Commun* 2014; 5: 4006.
10. Parmar C, Leijenaar RT, Grossmann P, et al. Radiomic feature clusters and prognostic signatures specific for Lung and Head & Neck cancer. *Sci Rep* 2015; 5: 11044.
11. Leger S, Zwanenburg A, Pilz K, et al. A comparative study of machine learning methods for time-to-event survival data for radiomics risk modelling. *Sci Rep* 2017; 7(1): 13206.
12. Vallieres M, Kay-Rivest E, Perrin LJ, et al. Radiomics strategies for risk assessment of tumour failure in head-and-neck cancer. *Sci Rep* 2017; 7(1): 10117.
13. Bogowicz M, Riesterer O, Stark LS, et al. Comparison of PET and CT radiomics for prediction of local tumor control in head and neck squamous cell carcinoma. *Acta Oncol* 2017; 56(11): 1531-6.
14. Forghani R, Chatterjee A, Reinhold C, et al. Head and neck squamous cell carcinoma: prediction of cervical lymph node metastasis by dual-energy CT texture analysis with machine learning. *Eur Radiol* 2019; 29(11): 6172-81.
15. Jethanandani A, Lin TA, Volpe S, et al. Exploring Applications of Radiomics in Magnetic Resonance Imaging of Head and Neck Cancer: A Systematic Review. *Front Oncol* 2018; 8: 131.
16. Nooij RP, Hof JJ, van Laar PJ, van der Hoorn A. Functional MRI for Treatment Evaluation in Patients with Head and Neck Squamous Cell Carcinoma: A Review of the Literature from a Radiologist Perspective. *Curr Radiol Rep* 2018; 6(1): 2.
17. Li H, Zhu Y, Burnside ES, et al. Quantitative MRI radiomics in the prediction of molecular classifications of breast cancer subtypes in the TCGA/TCIA data set. *NPJ Breast Cancer* 2016; 2.
18. Gevaert O, Mitchell LA, Achrol AS, et al. Glioblastoma multiforme: exploratory radiogenomic analysis by using quantitative image features. *Radiology* 2014; 273(1): 168-74.
19. Gnep K, Fargeas A, Gutierrez-Carvajal RE, et al. Haralick textural features on T2-weighted MRI are associated with biochemical recurrence following radiotherapy for peripheral zone prostate cancer. *J Magn Reson Imaging* 2017; 45(1): 103-17.
20. Dang M, Lysack JT, Wu T, et al. MRI texture analysis predicts p53 status in head and neck squamous cell carcinoma. *AJNR Am J Neuroradiol* 2015; 36(1): 166-70.
21. Jansen JF, Lu Y, Gupta G, et al. Texture analysis on parametric maps derived from dynamic contrast-enhanced magnetic resonance imaging in head and neck cancer. *World J Radiol* 2016; 8(1): 90-7.
22. Yuan Y, Ren J, Shi Y, Tao X. MRI-based radiomic signature as predictive marker for patients with head and neck squamous cell carcinoma. *Eur J Radiol* 2019; 117: 193-8.
23. Ren J, Tian J, Yuan Y, et al. Magnetic resonance imaging based radiomics signature for the preoperative discrimination of stage I-II and III-IV head and neck squamous cell carcinoma. *Eur J Radiol* 2018; 106: 1-6.
24. Hayes DN, Van Waes C, Seiwert TY. Genetic Landscape of Human Papillomavirus-Associated Head and Neck Cancer and Comparison to Tobacco-Related Tumors. *J Clin Oncol* 2015; 33(29): 3227-34.
25. Leijenaar RT, Bogowicz M, Jochems A, et al. Development and validation of a radiomic signature to predict HPV (p16) status from standard CT imaging: a multicenter study. *Br J Radiol* 2018; 91(1086): 20170498.

26. Nauta IH, Rietbergen MM, van Bokhoven A, et al. Evaluation of the eighth TNM classification on p16-positive oropharyngeal squamous cell carcinomas in the Netherlands and the importance of additional HPV DNA testing. *Ann Oncol* 2018; 29(5): 1273-9.
27. Sobin LH, M.K.; Wittekind, C.; eds. *TNM Classification of Malignant Tumours*. 7th Edition ed. Hoboken: Wiley-Blackwell; 2009.
28. Piccirillo JF, Tierney RM, Costas I, Grove L, Spitznagel EL, Jr. Prognostic importance of comorbidity in a hospital-based cancer registry. *JAMA* 2004; 291(20): 2441-7.
29. Bol GH, Kotte AN, van der Heide UA, Legendijk JJ. Simultaneous multi-modality ROI delineation in clinical practice. *Comput Methods Programs Biomed* 2009; 96(2): 133-40.
30. Peeters CFW, Übelhör C, Mes SW, et al. Stable prediction with radiomics data. arXiv:1903.11696 [stat.ML]
31. Collewet G, Strzelecki M, Mariette F. Influence of MRI acquisition protocols and image intensity normalization methods on texture classification. *Magn Reson Imaging* 2004; 22(1): 81-91.
32. Zhao B, Tan Y, Tsai WY, et al. Reproducibility of radiomics for deciphering tumor phenotype with imaging. *Sci Rep* 2016; 6: 23428.
33. Balagurunathan Y, Gu Y, Wang H, et al. Reproducibility and Prognosis of Quantitative Features Extracted from CT Images. *Transl Oncol* 2014; 7(1): 72-87.
34. Mao J, Fang J, Duan X, et al. Predictive value of pretreatment MRI texture analysis in patients with primary nasopharyngeal carcinoma. *Eur Radiol* 2019; 29(8): 4105-13.
35. Zhang B, He X, Ouyang F, et al. Radiomic machine-learning classifiers for prognostic biomarkers of advanced nasopharyngeal carcinoma. *Cancer Lett* 2017; 403: 21-7.
36. Zhang B, Ouyang F, Gu D, et al. Advanced nasopharyngeal carcinoma: pre-treatment prediction of progression based on multi-parametric MRI radiomics. *Oncotarget* 2017; 8(42): 72457-65.
37. Zhang B, Tian J, Dong D, et al. Radiomics Features of Multiparametric MRI as Novel Prognostic Factors in Advanced Nasopharyngeal Carcinoma. *Clin Cancer Res* 2017; 23(15): 4259-69.
38. Zhuo EH, Zhang WJ, Li HJ, et al. Radiomics on multi-modalities MR sequences can subtype patients with non-metastatic nasopharyngeal carcinoma (NPC) into distinct survival subgroups. *Eur Radiol* 2019; 29(10): 5590-9.
39. Zhao L, Gong J, Xi Y, et al. MRI-based radiomics nomogram may predict the response to induction chemotherapy and survival in locally advanced nasopharyngeal carcinoma. *Eur Radiol* 2020; 30(1): 537-46.
40. Bruce JP, Yip K, Bratman SV, Ito E, Liu FF. Nasopharyngeal Cancer: Molecular Landscape. *J Clin Oncol* 2015; 33(29): 3346-55.
41. Marcu LG, Reid P, Bezak E. The Promise of Novel Biomarkers for Head and Neck Cancer from an Imaging Perspective. *Int J Mol Sci* 2018; 19(9).
42. Rivera C, Oliveira AK, Costa RAP, De Rossi T, Paes Leme AF. Prognostic biomarkers in oral squamous cell carcinoma: A systematic review. *Oral Oncol* 2017; 72: 38-47.
43. Sala E, Mema E, Himoto Y, et al. Unravelling tumour heterogeneity using next-generation imaging: radiomics, radiogenomics, and habitat imaging. *Clin Radiol* 2017; 72(1): 3-10.
44. Parmar C, Grossmann P, Rietveld D, Rietbergen MM, Lambin P, Aerts HJ. Radiomic Machine-Learning Classifiers for Prognostic Biomarkers of Head and Neck Cancer. *Front Oncol* 2015; 5: 272.
45. Mes SW, Te Beest D, Poli T, et al. Prognostic modeling of oral cancer by gene profiles and clinicopathological co-variables. *Oncotarget* 2017; 8(35): 59312-23.
46. Moeckelmann N, Ebrahimi A, Tou YK, et al. Prognostic implications of the 8th edition American Joint Committee on Cancer (AJCC) staging system in oral cavity squamous cell carcinoma. *Oral Oncol* 2018; 85: 82-6.
47. van Velden FH, Kramer GM, Frings V, et al. Repeatability of Radiomic Features in Non-Small-Cell Lung Cancer [(18)F] FDG-PET/CT Studies: Impact of Reconstruction and Delineation. *Mol Imaging Biol* 2016; 18(5): 788-95.
48. Boellaard R. Quantitative oncology molecularanalysis suite: ACCURATE. *J Nucl Med* 2018; 59.
49. Zwanenburg A, Leger S, Vallières M, Löck S. Image biomarker standardisation initiative. arXiv:1612.07003 [cs.CV]
50. Kaiser HF. A second generation little jiffy. *Psychometrika* 1970; 35(4): 401-15.
51. Welch ML, McIntosh C, Haibe-Kains B, et al. Vulnerabilities of radiomic signature development: The need for safeguards. *Radiother Oncol* 2019; 130: 2-9.
52. Haibe-Kains B, Desmedt C, Sotiriou C, Bontempi G. A comparative study of survival models for breast cancer prognostication based on microarray data: does a single gene beat them all? *Bioinformatics* 2008; 24(19): 2200-8.





# 5

## **Comprehensive multiparameter genetic analysis improves circulating tumor DNA detection in head and neck cancer patients**

Steven W. Mes, Arjen Brink, Erik A. Sistermans, Roy Straver, Cees B.M. Oudejans, Jos B. Poell, C. René Leemans, Ruud H. Brakenhoff

**Oral Oncology 2020; 109: 104852**

## **ABSTRACT**

### **Introduction**

Tumor-specific genetic aberrations in cell-free DNA (cfDNA) from plasma are promising biomarkers for diagnosis of recurrent head and neck squamous cell carcinoma (HNSCC). However, the sensitivity when using somatic mutations only in cfDNA is suboptimal. Here, we combined detection of copy number aberrations (CNAs), human papillomavirus (HPV) DNA and somatic mutations in a single sequencing workflow.

### **Methods**

Pretreatment plasmas of 40 patients and 20 non-cancer controls were used for analysis. Plasma DNA underwent low-coverage whole genome sequencing (lcWGS) to detect both CNAs and HPV-DNA, and deep sequencing to detect mutations in 12 frequently altered cancer driver genes in HNSCC using the same sequencing library. A specific analysis pipeline line was developed for data mining. The corresponding tumors were analyzed using slightly adapted protocols.

### **Results**

Using the developed method, somatic mutations and CNAs were detected in plasma DNA of HNSCC patients in 67% and 52%, respectively. HPV-DNA in plasma was detected in 100% of patients with HPV-positive tumors, and not in plasma of patients with HPV-negative tumors or non-cancer controls. Combined analysis increased the detection rate of tumor DNA in plasma to 78%. The detection rate was significantly associated with the stage of disease of the tumor. Neither HPV status nor location of the primary tumor influenced detection of CNAs or somatic mutations in plasma.

### **Conclusions**

This study demonstrates that the combined analysis of CNAs, HPV and somatic mutations in plasma of HNSCC patients is feasible and contributes to a higher sensitivity of the assay compared to single modality analyses.

## INTRODUCTION

Head and neck squamous cell carcinoma (HNSCC) shows a high rate of recurrences, distant metastasis and second primary tumors<sup>1</sup>, which can often only be treated curatively when detected at an early stage. Consequently, there is an imperative need for the development of minimally invasive methods that allow early detection of relapsed disease, but such a biomarker is currently not available. Therefore, diagnosis of recurrent HNSCC remains dependent on conventional imaging and clinical examination, resulting in many patients being diagnosed with recurrent disease at an advanced stage. The introduction of next-generation sequencing (NGS) created new opportunities for post-treatment surveillance. A promising example of NGS applications is the detection of circulating tumor DNA (ctDNA) in plasma, that was first described in colorectal cancer<sup>2</sup> and later also in HNSCC<sup>3</sup>.

In a previous study, Wang et al. studied the detection of somatic mutations and HPV in cell-free DNA (cfDNA) from plasma and saliva of 93 patients<sup>4</sup>. The authors focused on somatic mutations and HPV using a PCR-amplification based sequencing protocol. Limitations of this approach are: (i) it requires selection of mutations from the index tumor, followed by custom primer design for each unique mutation while even cancer driver mutations in the recurrent tumor may differ from those in the index tumor<sup>5</sup>; and (ii) the protocol is complex including a gel-purification step, which may introduce cross-contamination<sup>6</sup>. A more general limitation of PCR-based approaches lies in the fact that ctDNA fragments are generally small, on average 160 bp, and to allow PCR amplification both primer sequences have to be present on the same fragment, limiting detection rates as not all template molecules will be amplified. Others exploited alternative techniques for detection of ctDNA in HNSCC, for instance digital droplet PCR<sup>7</sup>, methylation assays<sup>8</sup> and low-coverage whole genome sequencing (lcWGS) for copy number aberrations (CNA)<sup>9</sup>.

All these methods focused on detection of a specific type of genetic alteration, while there is considerable heterogeneity of molecular alterations in HNSCC<sup>10</sup>. As an example, oncogenesis in approximately two-thirds of HNSCCs is associated with a high number of CNAs, whereas tumor development in the remaining one-third is predominantly characterized by somatic mutations<sup>11</sup>. Furthermore, a subset of oropharynx tumors is caused by oncogenic HPV infections and very few somatic mutations. Ideally, ctDNA detection would cover the variety of genetic alterations including HPV-DNA. Moreover, the majority of plasma DNA template molecules should be represented in the sequencing library to obtain the highest sensitivity. Therefore, we developed and tested a workflow for the combined detection of somatic mutations, HPV-DNA and CNAs using the same sequencing library, and assessed its feasibility in pretreatment plasma samples of 40 HNSCC patients and 20 non-cancer controls.

## METHODS

### Samples

We studied samples from 40 HNSCC patients and 20 anonymous individuals without cancer. The study was approved by the institutional review board and from all subjects signed informed consent was obtained. Patients were treated at Amsterdam UMC, location VUmc. Anonymous blood donors at Sanquin in Amsterdam served as non-cancer controls. Patient characteristics are depicted in Table 1. The control plasma was obtained from anonymized blood donors, and no clinical data was available. Samples of 4 × 6 ml whole blood were taken from patients before treatment of the primary tumor (N = 38) or before salvage treatment (N = 2) using EDTA vacutainers. Plasma was collected within 24 h from whole blood by centrifugation, and further purified with an additional centrifugation step at 20,162 g using a Hettich EBA 12 R microcentrifuge (Hettich, Tuttlingen, Germany) to remove contaminating cells. The schematic workflow of plasma DNA isolation, library preparation and sequencing is depicted in Figure 1. The cfDNA from 4 ml plasma was isolated using a Qiasymphony automated platform (Qiagen, Hilden, Germany). Extra tumor biopsies could be obtained from 27 patients

during examination under general anesthesia and were directly snap-frozen and stored in liquid nitrogen. Tumor DNA was isolated using the Purelink™ Genomic DNA Mini Kit (cat. K182001, Invitrogen, Carlsbad, CA, USA) after macrodissection to ensure that neoplastic cellularity was over 20%. The actual tumor-derived fraction was later precisely quantified from lcWGS using the R package ACE<sup>12</sup>.

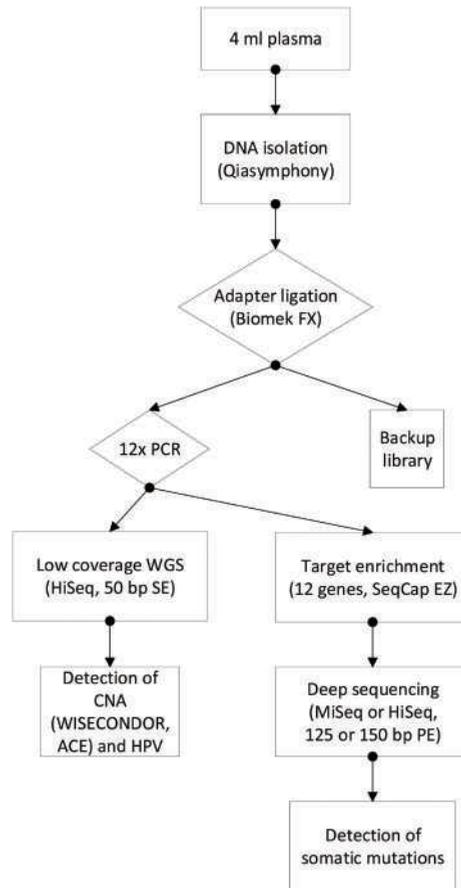
**Table 1.** Patient characteristics

Characteristic	N = 40
Age, mean (SD)	65.6 (9)
Gender	
Male (%)	27 (67.5)
Female (%)	13 (32.5)
Smoking (PY)	
0-10 (%)	11 (27.5)
11-24 (%)	8 (20)
>24 (%)	21 (52.5)
Subsite	
Oral cavity (%)	5 (12.5)
Oropharynx (%)	18 (45)
Hypopharynx (%)	10 (25)
Larynx (%)	5 (12.5)
Unknown primary (%)	2 (5)
TNM stage	
I (%)	2 (5)
II (%)	4 (10)
III (%)	6 (15)
IV (%)	28 (70)
T-stage	
0 (%)	2 (5)
1 (%)	3 (7.5)
2 (%)	13 (32.5)
3 (%)	4 (10)
4 (%)	18 (45)
N-stage	
0 (%)	17 (42.5)
1 (%)	7 (17.5)
2 (%)	15 (37.5)
3 (%)	1 (2.5)
HPV <sup>a</sup>	
Negative (%)	9 (45)
Positive (%)	10 (50)
Unknown <sup>b</sup> (%)	1 (5)

Abbreviations: HPV, human papilloma virus; PY, packyears; SD, standard deviation.

<sup>a</sup> Only determined when primary tumor was located in oropharynx or unknown.

<sup>b</sup> No material available from primary tumor for HPV analysis.



**Figure 1.** Illustration of workflow for detection of ctDNA.

Four ml of pretreatment plasma was used for automated DNA isolation and adapter ligation. The resulting library was divided in two to have a diagnostic backup, amplified, and subsequently used for low-coverage whole-genome sequencing to detect copy number aberrations and HPV, and target enrichment for deep sequencing to detect somatic mutations. Abbreviations: CNA, copy number aberrations; HPV, human papillomavirus; PCR, polymerase chain reaction; PE, paired-end; SE, single-end; WGS, whole genome sequencing.

## NGS library preparation and lcWGS

Plasma NGS libraries were generated with a 5500 SOLiD™ Fragment Library kit (cat. 4464412, LifeTech, Carlsbad, CA, USA) and TruSeq adapters (cat. FC-121-4001 and FC-121-4002, Illumina, San Diego, CA, USA) using a Biomek FX robot (Beckman Coulter, Pasadena, CA, USA). Prior to downstream steps, the plasma sequencing library was divided in a diagnostic library (20  $\mu$ l) and a backup library (25  $\mu$ l), and the diagnostic library underwent 12 cycles of PCR duplication (Figure 1) to generate the fragments for sequencing and allow equimolar pooling. Tumor NGS libraries were manually generated using the KAPA Hyper Prep Kit (cat. 07962347001, KAPA Biosystems, Cape Town, South Africa). Next, lcWGS was performed on the plasma and tumor samples by sequencing the non-enriched NGS libraries using an Illumina HiSeq 2000 instrument (Illumina) with 50 bp single-end reads as described before<sup>13,14</sup> with 8 to 21 samples per lane.

## Target enrichment and deep sequencing

Amplified diagnostic NGS libraries were enriched using a SeqCap EZ Choice kit (cat. 06740251001, Roche Nimblegen, Madison, WI, USA) with a 66 kb custom library of the exons of 12 genes (AJUBA, CASP8, CDKN2A,

FAT1, FBXW7, HRAS, KMT2D, NOTCH1, NSD1, PIK3CA, PTEN, TP53) that were selected based on TCGA data as the most frequently altered cancer driver genes in HNSCC<sup>11</sup>. The enriched libraries were subsequently sequenced on an Illumina HiSeq or Illumina MiSeq instrument with 125 bp or 150 bp paired-end reads.

### **Bioinformatics and statistical analysis**

Specific pipelines to detect CNAs, HPV and somatic mutations in both tumor and plasma DNA are described in the Supplementary Methods and Supplementary Figure 1. Clinical characteristics of patients with ctDNA detected in plasma were compared to those of patients without ctDNA detected in plasma using Fisher's exact test. All statistical analyses were performed using package 'stats' version 3.1.2 in R.

## **RESULTS**

### **Detection of copy number aberrations in tumor and corresponding plasma**

First, 27 tumors, of which biopsies were available, were analyzed for presence of CNAs. On average, each tumor had 34 altered segments (range 0–72, Table 2), including well known losses of 3p, 9p, and 17p, and gains of 3q, 7p and 8q (Supplementary Figure 2). Next, the corresponding cfDNA was sequenced. The CNA profiles of tumor and plasma were compared for corresponding aberrations. An example is shown in Figure 2. Alterations were found in 14 out of 27 (52%) plasma samples (Figure 3A and Table 2), whereas in 20 controls, a deletion was detected in only one case (5%) ( $P < 0.01$  by Fisher's Exact test). This deletion was found in an area with known copy number variations (4q31.21, database of genomic variants (<http://dgv.tcag.ca>)), and most likely depicts a copy number variation (CNV) and not a somatic genetic change.

### **Detection of somatic mutations in tumor and corresponding plasma**

For detection of somatic mutations, the same sequencing library used for copy number analysis was enriched for the 12-gene targeted sequencing panel. The mean coverage for plasma was 478 (range 145–1464). The distribution of the coverage of each gene in the plasma DNA is depicted in Supplementary Figure 3 for the samples with the highest coverage (A, mean 1464) and lowest coverage (B, mean 145). On average, the tumors contained three somatic mutations in the selected genes (range 0–7) with a mean variant allele frequency (VAF) of 29.3 (range 1.1–87.3) (Supplementary Table 1). Next, these mutations were searched for in the sequencing data of the plasma DNA using the bioinformatics pipelines depicted in Supplementary Figure 1C. One or more somatic mutations of the index tumor were detected in the plasma of 18 of 27 (67%) patients (Table 2 and Supplementary Table 1).

### **HPV detection**

An algorithm was developed to map reads from lcWGS to HPV genomes. This algorithm was first evaluated by comparing the results from tumor DNA lcWGS to the results of a validated HPV test to confirm that it could indeed be used for HPV detection<sup>15</sup>. We found 100% agreement in the HPV-positive and HPV-negative tumors analyzed. Next, the developed algorithm was applied to plasma DNA lcWGS, and HPV was detected in plasma of all evaluable patients with HPV-positive tumors (Table 2, Supplementary Figure 4). One sample (592) was not evaluable for this analysis as the backup FASTQ file was corrupted that was required for HPV genome mapping. No false-positive results were detected in plasma of patients with HPV-negative tumors or non-cancer controls (Table 2, Supplementary Figure 4) at the chosen cut-off value.

**Table 2.** Summary of detected tumor associated aberrations in cell-free DNA of plasma

	Patient	Tumor location	Stage of disease	HPV	Number of CNAs in tumor	CNAs found in Plasma	Number of somatic mutations in Tumor	Number of somatic mutations in Plasma	HPV found in Plasma	Combined detection conclusion
	590	hp	IVA		43	Yes	2	2	No	Yes
	595	hp	IVC		31	Yes	4	4	No	Yes
	598	hp	IVB		51	Yes	2	2	No	Yes
	603	op	IVA	HPV16	18	Yes	4	0	HPV16	Yes
	604	la	III		5	No	7	0	No	No
	606	hp	III		36	No	4	2	No	Yes
	607	la	III		40	No	2	1	No	Yes
	610	oc	II		12	No	6	0	No	No
	616	op	IVC	HPV16	23	Yes	5	2	HPV16	Yes
	617	op	IVA	HPV16	3	Yes	0	NA <sup>a</sup>	HPV16	Yes
	618	la	IVA		53	Yes	1	1	No	Yes
	619	op	III	Negative	72	No	1	1	No	Yes
Tumor status known	623	hp	IVA		50	Yes	1	1	No	Yes
	625	op	IVA	HPV16	14	No	1	1	HPV16	Yes
	626	oc	IVA		23	Yes	4	3	No	Yes
	627	op	IVA	HPV16	16	No	2	1 <sup>b</sup>	NA <sup>c</sup>	Yes
	628	op	IVB	Negative	30	No	6	1	No	Yes
	629	hp	IVA		58	No	4	0	No	No
	632	la	III		39	No	2	0	No	No
	633	op	IVA	Negative	48	Yes	6	4	No	Yes
	635	op	IVA	Negative	62	Yes	2	2	No	Yes
	636	oc	I		0	NA <sup>d</sup>	6	0	No	No
	639	op	IVA	Negative	43	Yes	2	2	No	Yes
	640	op	IVA	HPV16	13	No	2	0	HPV16	Yes
	642	hp	II		61	No	1	0	No	No
	647	la	IVA		23	Yes	2	2	No	Yes
	649	hp	IVA		43	Yes	3	2	No	Yes
Tumor status unknown	591	op	II	HPV16	NA	No	NA	0	HPV16	Yes
	592	op	IVA	Unknown <sup>e</sup>	NA	No	NA	1	No	Yes
	601	oc	I		NA	No	NA	0	No	No
	608	oc	IVA		NA	Yes	NA	1	No	Yes
	611	up	IVB	Negative	NA	Yes	NA	1	No	Yes
	612	up	IVA	Negative	NA	Yes	NA	0	No	Yes
	615	hp	IVA		NA	No	NA	2	No	Yes
	622	op	IVA	Negative	NA	No	NA	0	No	No
	630	op	III	HPV16	NA	No	NA	0	HPV16	Yes
	637	op	II	Negative	NA	No	NA	0	No	No
	643	op	IVA	HPV16	NA	Yes	NA	0	HPV16	Yes
	644	hp	IVA		NA	Yes	NA	0	No	Yes
	648	op	IVA	HPV16	NA	No	NA	0	HPV16	Yes

Abbreviations: CNA, copy number aberration; hp, hypopharynx; HPV, human papillomavirus; la, larynx; NA, not applicable; oc, oral cavity; op, oropharynx; up, unknown primary

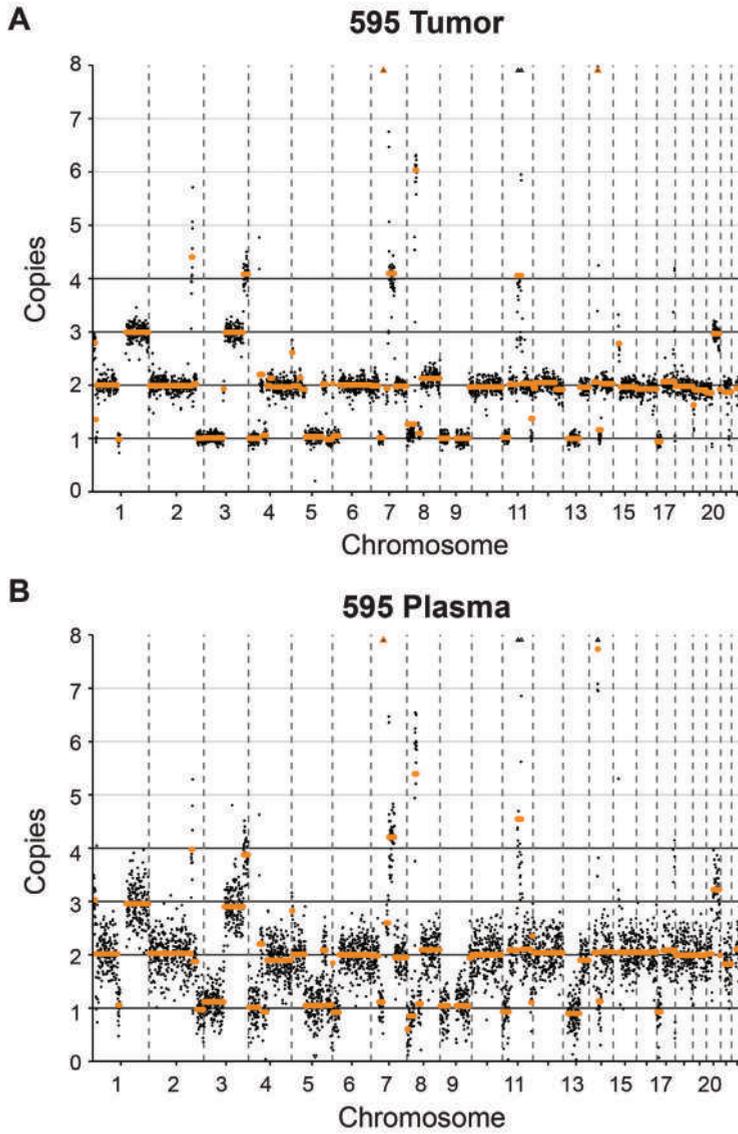
<sup>a</sup> no somatic mutations present in primary tumor

<sup>b</sup> found only in backup half of pre-PCR ligation product

<sup>c</sup> corrupted backup FASTQ file

<sup>d</sup> no CNAs present in primary tumor

<sup>e</sup> no material available from primary tumor for HPV analysis



**Figure 2.** Example of tumor and corresponding plasma copy number profile.

Absolute copy number profile of a cT4aN3M1 hypopharynx tumor (A) and corresponding plasma sample (B) showing, amongst others, well known losses of 3p, 9p and 17p, gains of 1q and 3q, and high-level amplifications of 7p and 11q. 2n ploidy was assumed for both samples. The graphs depict the number of copies on the Y-axis and the bins ordered by chromosomal location on the X-axis. Plot was derived from ACE.

## Detection of ctDNA without prior knowledge of tumor DNA aberrations

Since the algorithms to detect CNAs in plasma are not dependent on knowledge of tumor DNA aberrations, the same methods were applied to the 13 cases of which genetic data of the tumor was not available. CNAs were detected in plasma of 5 of 13 patients (38%, Table 2 and Figure 3). For detection of somatic mutations in plasma without prior knowledge of the somatic mutations in the primary tumor we took additional precautions. More stringent criteria for mutation calling were applied to ensure a low number of false positive mutation calls. Using these criteria, mutations were called in plasma of 4 of 13 patients (31%, Table 2). It should be noted that when we applied these criteria to the 27 patients with known tumor mutation status (Supplementary Table 2), some mutations were identified in cfDNA that were not identified in the primary tumor. These calls may be false positive, but could also relate to another (patho)biological process.

## Combined analysis of CNA, HPV-DNA and somatic mutations

The combined analysis of CNAs, somatic mutations and HPV in patients with known tumor DNA aberrations revealed tumor-associated aberrations in plasma DNA of 21 out of 27 (78%) patients (Table 2). When prior knowledge of tumor DNA aberrations was not available, ctDNA was detected in plasma of 10 of 13 patients (77%). Combining the results of cfDNA analyses in patients with and without prior knowledge on tumor variants, resulted in ctDNA detection in 31 out of 40 patients (78%, Table 2).

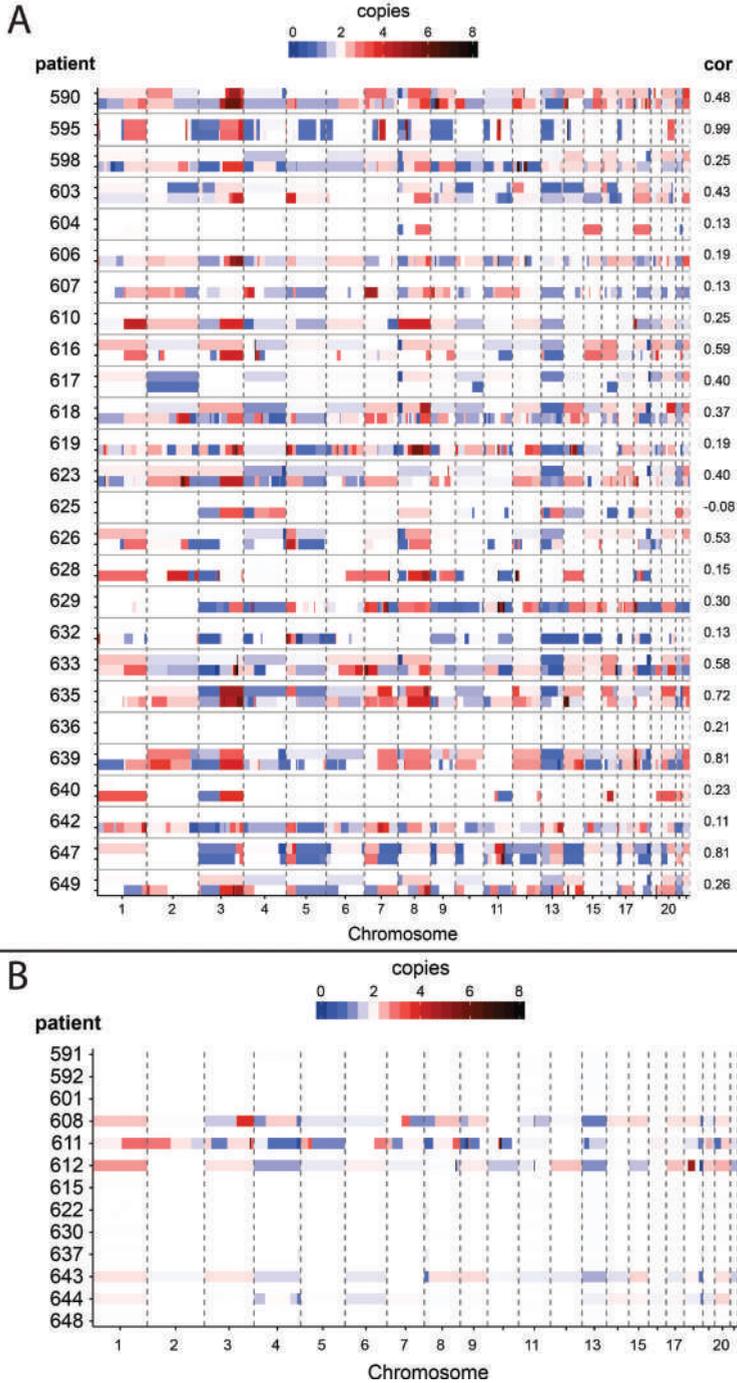
## Influence of clinical factors

In this cohort, the detection of CNAs and somatic mutations in plasma was not influenced by the location of the primary tumor or the HPV status (oropharynx tumors only). However, the TNM-stage was positively correlated to detection of CNAs in plasma ( $P < 0.05$ ), and to mutations in plasma ( $P < 0.05$ ) (Supplementary Figure 5).

## Analyses to improve sensitivity

To determine whether sequencing coverage constituted a rate-limiting factor in our experiments, we resequenced the index library of plasma of three patients using similar sequencing parameters and combined the sequencing results of the first and second run. Despite the increase in the total number of reads, the combined mean coverage was very comparable to the mean coverage of the two separate runs after duplicate removal, and no additional variants were discovered (Supplementary Table 3), indicating that the libraries were sequenced to completion with sufficient coverage.

Next, we prepared target-enriched sequencing libraries from the backup plasma DNA library (Figure 1) of seven patients, which may contain additional DNA templates that were not sequenced before. In one patient without mutations in the primary plasma DNA library, ctDNA was detected in the backup library (Supplementary Table 3). These results demonstrate that input DNA can be the limiting factor for detection of tumor variants in plasma.



**Fig. 3.** Heatmap of copy number aberrations in tumor and plasma samples.

(A) For each of 27 patients, copy number aberrations found in plasma, first row, and in tumor, second row, are shown. On the X-axis are the segments ordered by chromosomal location. Blue indicates an allelic loss (less than two copies), whereas red indicates an allelic gain or amplification (more than two copies). On the right, the correlation is shown between segments of plasma and tumor. (B) For the 13 patients with missing tumor data, only the plasma copy number aberrations are shown.

## DISCUSSION

This study was designed to develop a workflow for detection of CNAs, HPV-DNA and somatic mutations in cfDNA from plasma of HNSCC patients using the same sequencing library, and to determine its performance in plasma samples of 40 HNSCC patients. This method was based on lcWGS for detection of CNAs and HPV-DNA, and targeted deep sequencing of a panel of 12 frequently mutated genes in HNSCC.

In our view this approach combines the best of all available approaches, and we detected CNAs in plasma of HNSCC patients in 52%, HPV-DNA in plasma of 100% of patients with HPV-positive tumors, and somatic mutations in plasma of 67% of HNSCC patients. It was hypothesized that the combined analysis of CNA, HPV and mutations would increase the overall sensitivity of the assay, and this was indeed the case. To reduce costs, one could argue for a stepwise approach, in which first lcWGS is performed to detect CNAs and HPV-DNA, and deep sequencing only when CNAs or HPV-DNA are not detected. In fact, for HPV-positive tumors, our data suggests that the deep sequencing step is redundant since all HPV-positive tumors could already be detected by lcWGS. Other authors demonstrated different HPV-DNA detection methods<sup>16-20</sup>, but the major advantage of our combined approach is that it is comprehensive: a single assay that serves all tumor subtypes.

As genetic alterations in the index and recurrent tumor may differ<sup>5</sup>, ideally post-treatment surveillance does not rely only on detection of genetic changes that were present in the index tumor. In this regard, we show an additional advantage of our comprehensive method. Confident detection of somatic mutations without prior knowledge of tumor DNA mutations may be hampered by false positive calls, but the detection of CNAs and HPV-DNA in plasma did not depend on prior knowledge and can therefore be applied directly to plasma DNA with high specificity (>95%).

To our best knowledge relatively few studies focus on the detection of ctDNA in HNSCC patients. Wang et al. detected ctDNA in 87% of 47 HNSCC patients<sup>4</sup>. Importantly, their assay also included detection of circulating HPV-DNA in patients with HPV-positive tumors, whereby the sensitivity increased from 64% to 87%. Schwaederle et al. found alterations in plasma of 88% of 25 HNSCC patients<sup>21</sup>. However, these alterations were not matched to the mutational profiles of the primary tumors, which in our study was shown to be important to exclude false positive calls. Perdomo et al. detected mutations in 42% of 36 HNSCC cases<sup>22</sup>, Schirmer et al. showed a sensitivity of 74%<sup>9</sup>, and Galot et al. showed a sensitivity of 51%<sup>23</sup>. In other studies<sup>3,7</sup> only very small HNSCC patient cohorts were analyzed, which makes a relevant comparison difficult.

There are several options to further enhance the detection rate, which might be required for clinical implementation. First, it was shown by others that methods of blood processing have a strong influence on cfDNA levels<sup>24</sup>, and optimizing the sampling of blood by using dedicated cell-free DNA tubes may improve detection. Second, increasing the DNA input may overcome undersampling and improve detection of low fractions of ctDNA<sup>25</sup>. This effect was already shown in our study as we detected additional variants when sequencing the backup DNA library, and it further implies that a high conversion rate, i.e. the representation of the cfDNA in the sequencing library, is of large importance. Third, ctDNA has shorter fragment sizes than non-tumor derived plasma DNA<sup>26</sup>, and size selection of fragments between 90 and 150 bp was shown to improve the detection of ctDNA in advanced cancers<sup>27</sup>. Finally, Wang et al. showed that the combination of saliva and plasma analysis also increased sensitivity<sup>4</sup>, which implies that future cohort studies should also include samples of saliva or oral rinses from HNSCC patients.

In conclusion, we developed a method for combined analysis of somatic mutations, HPV-DNA and CNAs in cfDNA from plasma of HNSCC patients. Although prior knowledge of the genetic changes in the tumor is helpful in the analyses, we show that it is not required. Future research should focus on implementing the suggested improvements of the assay and application in a large longitudinal cohort for early detection of recurrent disease to proof the clinical utility.

## **CONFLICTS OF INTEREST**

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

## **FUNDING**

This work was supported by the European Union's Seventh Framework Project (grant agreement 611425: OraMod). Funders had no role in the design and conduct of the study; collection, management, analysis, and interpretation of the data; preparation, review, or approval of the manuscript; and decision to submit the manuscript for publication.

## SUPPLEMENTARY METHODS

### BIOINFORMATICS

#### Detection of CNAs by lcWGS

Detection of CNAs in primary tumors was performed using the R package QDNAseq<sup>13</sup>. For detection of CNAs in plasma DNA, the generated data was analyzed using two approaches. Firstly, aberrations were called based on signal differences compared to a reference set of plasma samples using WISECONDOR<sup>14</sup>. This pipeline searches for aberrations using an individual bin method, sliding window method and chromosomal wide aneuploidy test. Only aberrations detected by the individual bin method and sliding window method were considered. Secondly, the mapped reads were binned (bin size 500 kbp), normalized, and corrected for GC content and mappability<sup>13</sup>. Samples were normalized using the 20 control plasma samples and segmented using the build-in DNAcopy call<sup>28</sup> within QDNAseq. Using the ACE<sup>12</sup> function twosamplecompare, which equalizes segments between two samples, all patient plasma samples were tested against all control plasma samples. Samples were called positive if at least 1 segment had a significantly different mean (multiple testing adjusted P-value < 0.001) from all control samples. CNA plasma profiles were visually compared to the corresponding tumor profiles.

#### Detection of HPV by lcWGS

HPV status of oropharyngeal tumors was determined in the formalin-fixed, paraffin-embedded (FFPE) tumor specimen by using p16 immunostaining followed by HPV-DNA PCR on p16-positive samples, as described before<sup>15</sup>. HPV status of unknown primary tumors was determined by HPV-DNA PCR of the cytology specimen<sup>29</sup>. Subsequently, an algorithm was developed that exploits the non-human reads of lcWGS and maps these reads to the known HPV-genomes. HPV-specific reads were normalized to within-sample human-specific reads. This algorithm was first applied to compare the sequencing results of the primary tumors to the validated HPV test. Next, the developed algorithm was applied to the non-human reads of WGS of the plasma samples of all 40 patients and 20 controls. Samples were deemed positive when the fraction of HPV-specific reads exceeded the highest fraction observed in controls by a factor 2.

#### Identification of somatic mutations in tumor DNA

For detection of somatic mutations in the primary tumors, two pipelines for analysis were used (Supplementary Figure 1A): i. FASTQ files were trimmed using Trimmomatic<sup>30</sup>, mapped to Hg19 with BWA-MEM<sup>31</sup>, and duplicates were removed using Picard tools' MarkDuplicates (<http://broadinstitute.github.io/picard/>). Calling was performed using SAMtools Mpileup<sup>32</sup> and VarScan<sup>33</sup>. Settings for VarScan (SNP and indel) were: min-var-freq 0.01, p-value 0.1, strand-filter 1; ii. Trimming, mapping, and duplicate removal were performed using the same tools, however base quality score recalibration was performed using GATK BaseRecalibrator<sup>34</sup> and calling was performed using GATK MuTect2 with default settings<sup>34</sup>. Identified variants were filtered using the 1000 Genomes Project dataset, and variants with EUR AF >0.01 were removed<sup>35</sup>. Functional annotation of sequence variants was performed by Oncotator<sup>36</sup> and snpEff<sup>37</sup>.

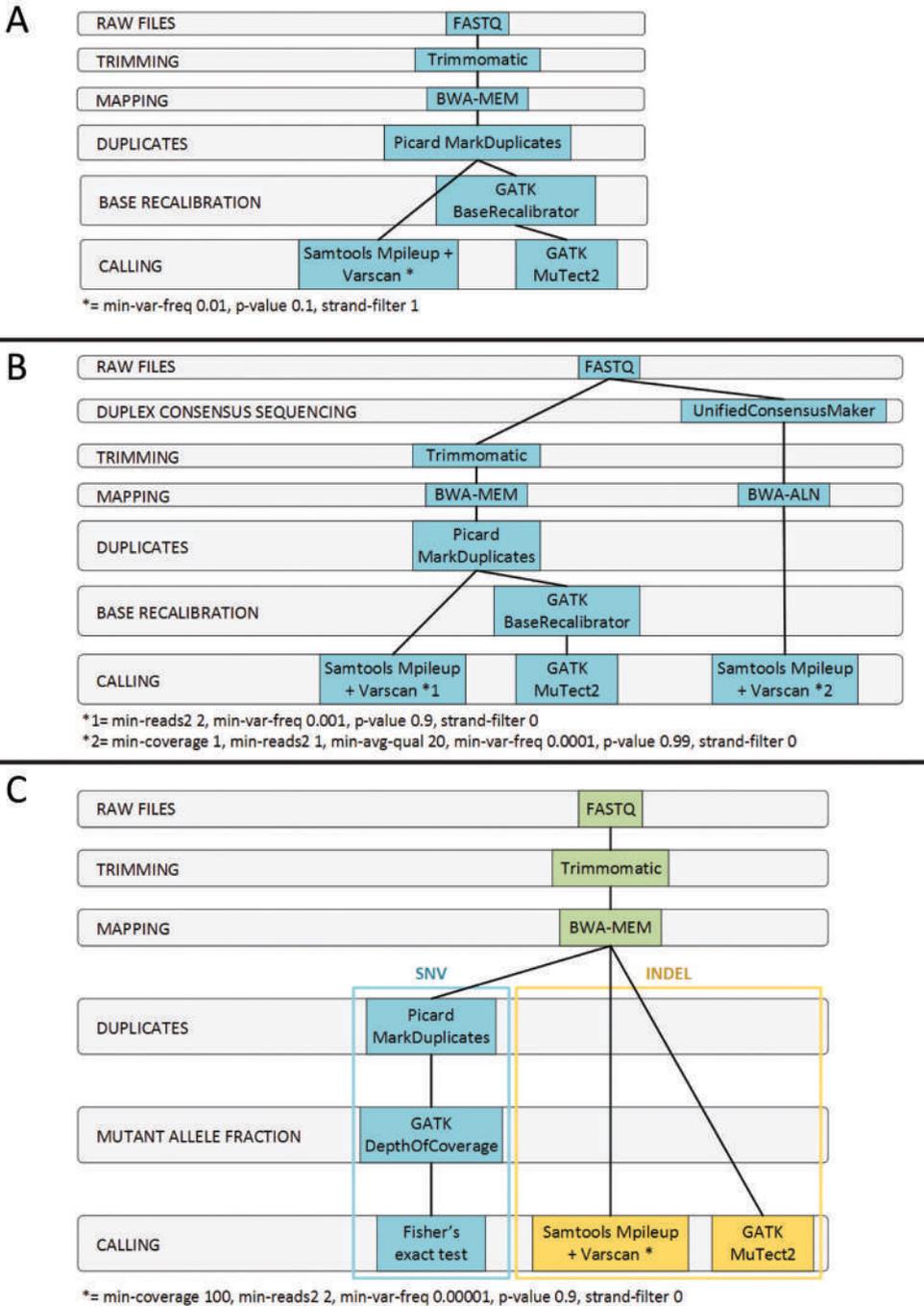
#### Identification of somatic mutations in cfDNA (without knowledge of predetermined tumor variants)

Detection of de novo variants in plasma was performed by combining the same two pipelines as described for tumors but with different settings of VarScan (SNP and indel): min-reads2 2, min-var-freq 0.001, p-value 0.9, strand-filter 0. Additionally, a third pipeline was used to reduce the number of false positive calls

(Supplementary Figure 1B): iii. FASTQ files were collapsed into Duplex Consensus Sequences (DCS) using UnifiedConsensusMaker<sup>38</sup> and mapped to Hg19 with BWA-ALN<sup>31</sup>. Calling was performed using Samtools Mpileup<sup>32</sup> and VarScan<sup>33</sup> using the following settings of VarScan: min-coverage 1, min-reads2 1, min-avg-qual 20, min-var-freq 0.0001, p-value 0.99, strand-filter 0. Detected variants in plasma DNA were considered valid when called by all three pipelines and VAF  $\leq$  25% to remove potentially remaining germline variants. Further filtering of sequence variants and functional annotation were performed as described before.

### **Detection of predetermined tumor variants in cfDNA**

To increase sensitivity and specificity, plasma DNA was examined for the predetermined tumor variants only. The schematic workflow is depicted in Supplementary Figure 1C. FASTQ files were mapped to Hg19 with BWA-MEM<sup>31</sup>. For detection of single nucleotide variations (SNVs), the data was trimmed using Trimmomatic<sup>30</sup>, duplicates were removed using Picard tools' MarkDuplicates, and mutant allele fractions of the tumor variants were determined in the corresponding plasma sample using GATK DepthOfCoverage<sup>34</sup>. These fractions were compared to mutant allele fractions of the same variants in the other samples. Samples in which the mutant allele fractions significantly exceeded their frequencies in the other samples ( $P < 0.05$ ) were scored as positive according to Wang et al.<sup>4</sup> For detection of small insertions and deletions duplicates were not removed. Calling was achieved using: i. Samtools Mpileup<sup>32</sup> and VarScan<sup>33</sup> with the following settings of VarScan (SNP and indel): min-coverage 100, min-reads2 2, min-var-freq 0.00001, p-value 0.9, strand-filter 0; and ii. GATK MuTect2 with default settings<sup>34</sup>.



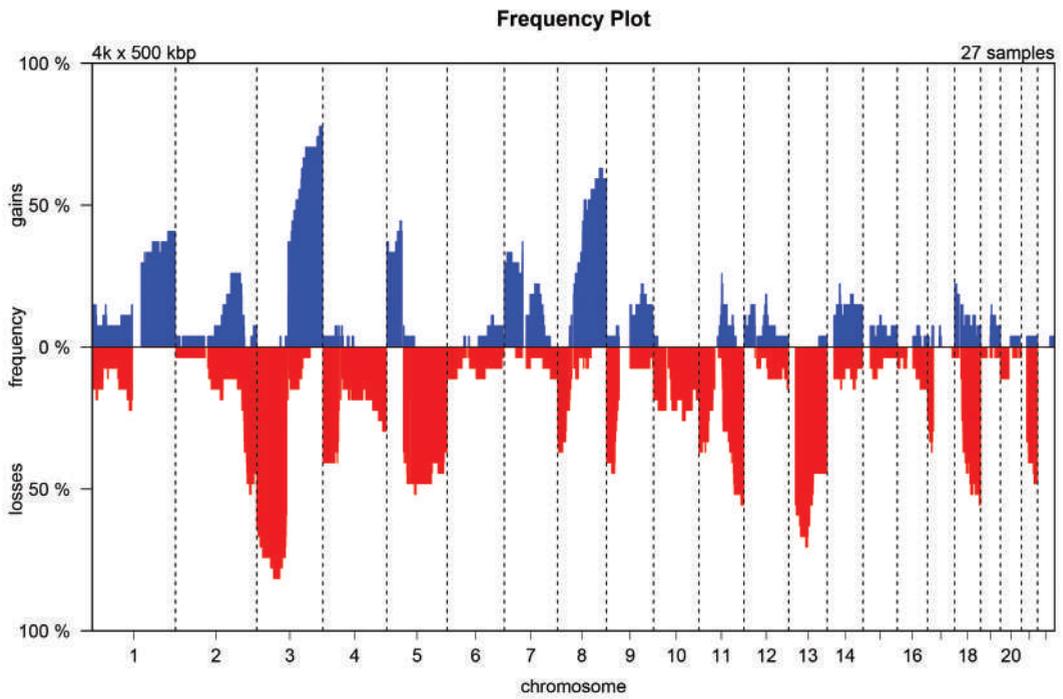
**Supplementary Figure 1.** Illustration of bioinformatics pipelines for detection of somatic mutations

(A) For detection of somatic mutations in the primary tumors two pipelines for analysis were used: i. FASTQ files were trimmed using Trimmomatic<sup>30</sup>, mapped to Hg19 with BWA-MEM<sup>31</sup>, and duplicates were removed using Picard tools' MarkDuplicates (<http://broadinstitute.github.io/picard/>). Calling was performed using SAMtools Mpileup<sup>32</sup> and VarScan<sup>33</sup>. Settings for VarScan (SNP and indel) were: min-var-freq 0.01, p-value 0.1, strand-filter 1; ii. Trimming, mapping, and duplicate removal were performed using the same tools, however base quality score recalibration was performed using GATK BaseRecalibrator<sup>34</sup> and calling was performed using GATK MuTect2 with default settings<sup>34</sup>.

(B) For detection of somatic mutations in cfDNA in plasma without knowledge of predetermined tumor variants we applied the following: the same two pipelines as described in (A) were used but with different settings of VarScan (SNP and indel): min-reads2 2, min-var-freq 0.001, p-value 0.9, strand-filter 0. In addition, FASTQ files were collapsed into Duplex Consensus Sequences (DCS) using UnifiedConsensusMaker<sup>38</sup> and mapped to Hg19 with BWA-ALN<sup>31</sup>. Calling was performed using Samtools Mpileup<sup>32</sup> and VarScan<sup>33</sup> using the following settings of VarScan: min-coverage 1, min-reads2 1, min-avg-qual 20, min-var-freq 0.0001, p-value 0.99, strand-filter 0. Detected variants in plasma DNA were considered valid when called by all three pipelines and VAF  $\leq$  25% to remove potentially remaining germ line variants.

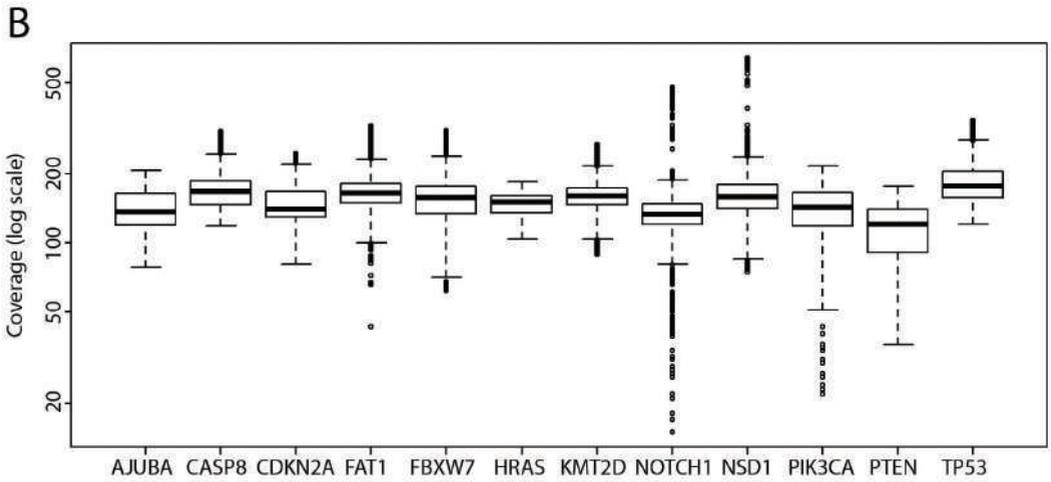
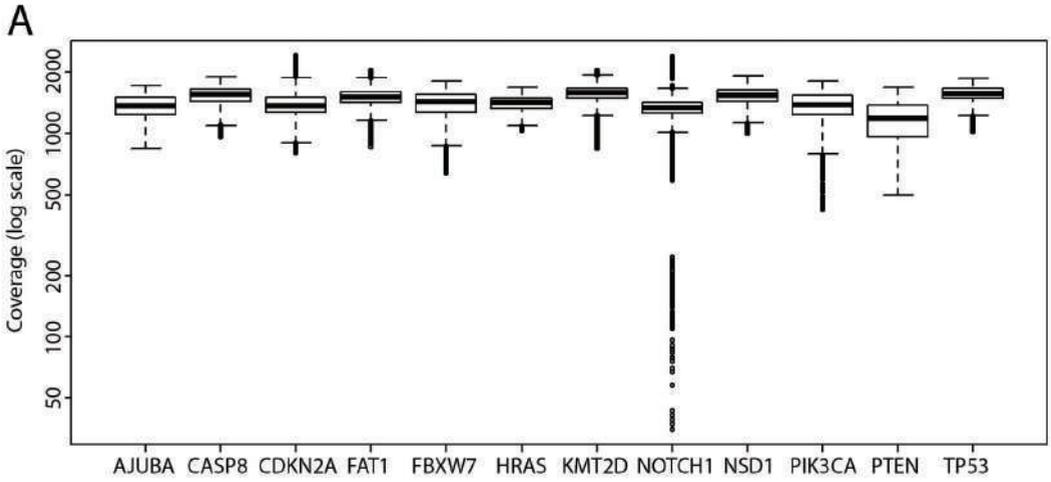
(C) For detection of predetermined tumor variants in cfDNA we used the following: FASTQ files were mapped to Hg19 with BWA-MEM<sup>31</sup>. For detection of single nucleotide variations (SNVs), the data was trimmed using Trimmomatic<sup>30</sup>, duplicates were removed using Picard tools' MarkDuplicates, and mutant allele fractions of the tumor variants were determined in the corresponding plasma sample using GATK DepthOfCoverage<sup>34</sup>. These fractions were compared to mutant allele fractions of the same variants in the other samples. Samples in which the mutant allele fractions significantly exceeded their frequencies in the other samples ( $P < 0.05$ ) were scored as positive according to Wang et al.<sup>4</sup> For detection of small insertions and deletions duplicates were not removed. Calling was achieved using: i. Samtools Mpileup<sup>32</sup> and VarScan<sup>33</sup> with the following settings of VarScan (SNP and indel): min-coverage 100, min-reads2 2, min-var-freq 0.00001, p-value 0.9, strand-filter 0; and ii. GATK MuTect2 with default settings<sup>34</sup>.

Abbreviations: INDEL, insertion or deletion; SNV, single nucleotide variation.



5

**Supplementary Figure 2.** Frequency plot of copy number aberrations in tumor cohort  
The graph represents the relative frequency of gains (in blue) and losses (in red) assessed by low-coverage whole genome sequencing in 27 tumors for each chromosome position.



**Supplementary Figure 3.** Coverage per gene in plasma DNA of sample with highest and lowest mean coverage  
 A uniform coverage was found of the targeted gene panel in plasma samples with highest (A) and lowest (B) mean coverage, except for a few bases in NOTCH1.



**Supplementary Table 1.** Detected somatic mutations in primary tumors and corresponding plasma samples

Tumor								
Sample	HPV <sup>a</sup>	Chromosome	Position	Gene	REF	ALT	Variant_Type	VAF
590	NA	chr4	187542108	FAT1	CAG	C	DEL	29.9
590	NA	chr17	7577579	TP53	G	GTA	INS	36
595	NA	chr4	153249440	FBXW7	C	A	SNP	25.4
595	NA	chr4	187584469	FAT1	TGAAAAGAATCCTTGTC	T	DEL	30
595	NA	chr9	139400061	NOTCH1	G	C	SNP	25.9
595	NA	chr17	7577106	TP53	G	A	SNP	49.7
598	NA	chr4	187628119	FAT1	CAGGATCGTGGGCTTCT	C	DEL	55.5
598	NA	chr17	7578474	TP53	C	CG	INS	70
603	HPV16	chr3	178936091	PIK3CA	G	A	SNP	13.4
603	HPV16	chr4	153247289	FBXW7	G	A	SNP	20.3
603	HPV16	chr9	139395305	NOTCH1	G	A	SNP	6.4
603	HPV16	chr14	23442712	AJUBA	G	C	SNP	4.6
604	NA	chr3	178936082	PIK3CA	G	A	SNP	39.3
604	NA	chr4	187535374	FAT1	G	GC	INS	36.4
604	NA	chr4	187538941	FAT1	T	TA	INS	36.9
604	NA	chr5	176562223	NSD1	C	T	SNP	36.6
604	NA	chr9	139399991	NOTCH1	C	T	SNP	42.1
604	NA	chr11	534289	HRAS	C	T	SNP	39.1
604	NA	chr12	49431773	KMT2D	C	T	SNP	36.5
606	NA	chr3	178921472	PIK3CA	G	A	SNP	8.9
606	NA	chr9	139418427	NOTCH1	CACAGCTGTTG	C	DEL	22.1
606	NA	chr9	139418428	NOTCH1	A	G	SNP	1.39
606	NA	chr17	7578188	TP53	C	A	SNP	37.6
607	NA	chr9	21971120	CDKN2A	G	A	SNP	64.1
607	NA	chr9	139400000	NOTCH1	C	A	SNP	59
610	NA	chr4	153332759	FBXW7	G	A	SNP	8.2
610	NA	chr4	187524477	FAT1	G	A	SNP	4
610	NA	chr4	187541808	FAT1	T	A	SNP	3.4
610	NA	chr5	176638938	NSD1	A	T	SNP	4
610	NA	chr12	49433015	KMT2D	CA	C	DEL	1.6
610	NA	chr17	7578555	TP53	C	T	SNP	5.9
616	HPV16	chr3	178928079	PIK3CA	G	A	SNP	16.7
616	HPV16	chr9	139390581	NOTCH1	G	C	SNP	1.2
616	HPV16	chr9	139391328	NOTCH1	G	C	SNP	1.5
616	HPV16	chr9	139391547	NOTCH1	G	A	SNP	1.1

Plasma									
Duplex consensus sequencing <sup>b</sup>				Mutect2 <sup>c</sup>		VarScan <sup>d</sup>		Summary plasma <sup>e</sup>	
ALT_ reads	REF_ reads	ALT_ reads_ total	REF_ reads_ total	P-value	VOF	TLOD	VOF	P-value	
NA	NA	NA	NA	NA	3.3	216.62	4.18	6.66E-52	Confirmed
NA	NA	NA	NA	NA	2.1	222.39	1.52	1.26E-15	Confirmed
96	1268	98	15176	2.4E-223	NA	NA	NA	NA	Confirmed
NA	NA	NA	NA	NA	6.9	1072.56	7.78	2.14E-102	Confirmed
64	656	64	12006	0	NA	NA	NA	NA	Confirmed
150	1220	155	14603	5.44193546268299e-319	NA	NA	NA	NA	Confirmed
NA	NA	NA	NA	NA	1.9	338.48	2.04	5.27E-34	Confirmed
NA	NA	NA	NA	NA	1.8	60.46	0.31	0.004643	Confirmed
0	422	2	13385	1	NA	NA	NA	NA	Not confirmed
0	464	4	14274	1	NA	NA	NA	NA	Not confirmed
0	433	1	12344	1	NA	NA	NA	NA	Not confirmed
0	324	1	10557	1	NA	NA	NA	NA	Not confirmed
0	387	7	13540	1	NA	NA	NA	NA	Not confirmed
NA	NA	NA	NA	NA	ND	ND	ND	ND	Not confirmed
NA	NA	NA	NA	NA	ND	ND	0.09	0.60479	Not confirmed
0	347	5	12030	1	NA	NA	NA	NA	Not confirmed
0	481	12	15006	1	NA	NA	NA	NA	Not confirmed
2	322	19	11195	0.091133	NA	NA	NA	NA	Not confirmed
0	482	7	17297	1	NA	NA	NA	NA	Not confirmed
1	216	11	13565	0.149449	NA	NA	NA	NA	Not confirmed
NA	NA	NA	NA	NA	1.1	132.58	0.94	4.67E-12	Confirmed
0	210	2	13574	1	NA	NA	NA	NA	Not confirmed
2	195	6	15380	0.001269	NA	NA	NA	NA	Confirmed
0	304	5	13271	1	NA	NA	NA	NA	Not confirmed
2	371	4	14985	0.001243	NA	NA	NA	NA	Confirmed
0	126	0	14058	1	NA	NA	NA	NA	Not confirmed
0	140	2	15240	1	NA	NA	NA	NA	Not confirmed
0	170	1	14532	1	NA	NA	NA	NA	Not confirmed
0	144	1	15889	1	NA	NA	NA	NA	Not confirmed
NA	NA	NA	NA	NA	ND	ND	ND	ND	Not confirmed
0	163	3	14407	1	NA	NA	NA	NA	Not confirmed
2	113	6	11090	0.000818	NA	NA	NA	NA	Confirmed
0	183	1	14531	1	NA	NA	NA	NA	Not confirmed
0	165	0	13568	1	NA	NA	NA	NA	Not confirmed
0	255	6	17133	1	NA	NA	NA	NA	Not confirmed

Supplementary Table 1. (continued)

Tumor								
Sample	HPV <sup>a</sup>	Chromosome	Position	Gene	REF	ALT	Variant_Type	VAF
616	HPV16	chr12	49427912	KMT2D	C	T	SNP	11.2
617 <sup>f</sup>	HPV16	-	-	-	-	-	-	-
618	NA	chr17	7578550	TP53	G	T	SNP	71.2
619	Negative	chr17	7577568	TP53	C	A	SNP	43.2
623	NA	chr17	7578457	TP53	C	A	SNP	73
625	HPV16	chr2	202151270	CASP8	C	T	SNP	16.5
626	NA	chr4	187557993	FAT1	CA	C	DEL	20.9
626	NA	chr9	139409854	NOTCH1	T	A	SNP	32.4
626	NA	chr17	7577141	TP53	C	T	SNP	17.1
626	NA	chr17	7578406	TP53	C	T	SNP	17.9
627	HPV16	chr12	49421581	KMT2D	C	T	SNP	9.3
627	HPV16	chr12	49421705	KMT2D	C	A	SNP	11.8
628	Negative	chr3	178936091	PIK3CA	G	A	SNP	39.9
628	Negative	chr4	187522492	FAT1	C	T	SNP	11.7
628	Negative	chr4	187584620	FAT1	T	A	SNP	41.4
628	Negative	chr4	187584621	FAT1	GT	G	DEL	41.5
628	Negative	chr9	139402593	NOTCH1	T	A	SNP	31.1
628	Negative	chr17	7578448	TP53	G	GC	INS	87.3
629	NA	chr4	187524787	FAT1	G	GA	INS	1.4
629	NA	chr9	21971155	CDKN2A	G	A	SNP	50.5
629	NA	chr17	7577144	TP53	A	G	SNP	35.1
629	NA	chr17	7577518	TP53	T	A	SNP	29.9
632	NA	chr9	21971111	CDKN2A	G	A	SNP	40.2
632	NA	chr17	7578275	TP53	G	A	SNP	33.1
633	Negative	chr4	187584558	FAT1	C	CA	INS	31
633	Negative	chr4	187584563	FAT1	G	T	SNP	30.5
633	Negative	chr9	21974506	CDKN2A	CA	C	DEL	33.5
633	Negative	chr9	21974508	CDKN2A	G	T	SNP	33.6
633	Negative	chr9	21974712	CDKN2A	TG	T	DEL	19.3
633	Negative	chr17	7578474	TP53	CG	C	DEL	28.9
635	Negative	chr4	153332667	FBXW7	C	A	SNP	32.9
635	Negative	chr17	7579415	TP53	C	T	SNP	59.6
636	NA	chr3	178952085	PIK3CA	A	G	SNP	18.6
636	NA	chr4	153247376	FBXW7	T	C	SNP	18.5
636	NA	chr4	187530963	FAT1	CA	C	DEL	16.9

Plasma									
Duplex consensus sequencing <sup>b</sup>					Mutect2 <sup>c</sup>		VarScan <sup>d</sup>		Summary plasma <sup>e</sup>
ALT_ reads	REF_ reads	ALT_ reads_ total	REF_ reads_ total	P-value	VAF	TLOD	VAF	P-value	
4	216	10	15945	1.75E-06	NA	NA	NA	NA	Confirmed
-	-	-	-	-	-	-	-	-	Not confirmed
6	612	9	14838	5.61E-09	NA	NA	NA	NA	Confirmed
2	1533	3	16462	0.004922	NA	NA	NA	NA	Confirmed
12	367	14	15080	3.96E-25	NA	NA	NA	NA	Confirmed
4	731	10	13830	0.000399	NA	NA	NA	NA	Confirmed
NA	NA	NA	NA	NA	4.7	453.32	8.91	5.96E-179	Confirmed
4	170	6	14201	1.36E-08	NA	NA	NA	NA	Confirmed
0	121	4	13023	1	NA	NA	NA	NA	Not confirmed
3	166	17	14653	0.000601	NA	NA	NA	NA	Confirmed
0	1236	11	15745	1	NA	NA	NA	NA	Not confirmed
0	1168	2	16701	1	NA	NA	NA	NA	Not confirmed
0	433	2	13385	1	NA	NA	NA	NA	Not confirmed
0	471	1	14549	1	NA	NA	NA	NA	Not confirmed
0	451	4	14477	1	NA	NA	NA	NA	Not confirmed
NA	NA	NA	NA	NA	ND	ND	ND	ND	Not confirmed
0	444	0	13747	1	NA	NA	NA	NA	Not confirmed
NA	NA	NA	NA	NA	0.7626	31.98	0.87	8.019E-11	Confirmed
NA	NA	NA	NA	NA	ND	ND	ND	ND	Not confirmed
0	159	10	10428	1	NA	NA	NA	NA	Not confirmed
1	163	20	12931	0.215525	NA	NA	NA	NA	Not confirmed
0	171	7	17345	1	NA	NA	NA	NA	Not confirmed
1	380	13	13836	0.287541	NA	NA	NA	NA	Not confirmed
0	352	7	14671	1	NA	NA	NA	NA	Not confirmed
NA	NA	NA	NA	NA	0.9012	58.27	1.9	3.008E-33	Confirmed
3	397	4	16803	2.3E-06	NA	NA	NA	NA	Confirmed
NA	NA	NA	NA	NA	1.9	323.12	3.31	6.245E-65	Confirmed
7	401	12	19367	2.54E-11	NA	NA	NA	NA	Confirmed
NA	NA	NA	NA	NA	ND	ND	1.41	5.878E-23	Not confirmed
NA	NA	NA	NA	NA	2.7	137.49	1.74	5.085E-28	Confirmed
14	432	17	15659	8.99E-27	NA	NA	NA	NA	Confirmed
20	459	24	17860	7.17E-39	NA	NA	NA	NA	Confirmed
1	303	8	12762	0.156577	NA	NA	NA	NA	Not confirmed
0	300	1	13574	1	NA	NA	NA	NA	Not confirmed
NA	NA	NA	NA	NA	ND	ND	ND	ND	Not confirmed

**Supplementary Table 1.** (continued)

<b>Tumor</b>								
Sample	HPV <sup>a</sup>	Chromosome	Position	Gene	REF	ALT	Variant_Type	VAF
636	NA	chr9	21970971	CDKN2A	G	T	SNP	36.9
636	NA	chr9	139412297	NOTCH1	C	T	SNP	38.1
636	NA	chr17	7578517	TP53	G	A	SNP	9.7
639	Negative	chr4	187525704	FAT1	C	T	SNP	12.6
639	Negative	chr17	7577538	TP53	C	A	SNP	27.6
640	HPV16	chr9	139413166	NOTCH1	C	A	SNP	2.2
640	HPV16	chr17	7578555	TP53	C	G	SNP	4.7
642	NA	chr17	7577545	TP53	T	C	SNP	18.2
647	NA	chr14	23450536	AJUBA	C	CT	INS	59.7
647	NA	chr17	7579380	TP53	AGG	A	DEL	60.7
649	NA	chr3	178952085	PIK3CA	A	G	SNP	61.1
649	NA	chr14	23451340	AJUBA	T	TC	INS	57.7
649	NA	chr17	7578474	TP53	C	CG	INS	50.5

Abbreviations: ALT, alternative sequence; DEL, deletion; HPV, human papillomavirus; INS, insertion; NA, not applicable; ND, not detected; REF, reference sequence; SNP, single-nucleotide polymorphism; VAF, variant allele frequency

<sup>a</sup> HPV status only determined in oropharynx tumors.

<sup>b</sup> Detection of SNPs: comparison of mutant fractions in patient under investigation (first two columns) versus control (all other samples; third and fourth column) after duplicate removal. P-value was calculated using a two-sided Fisher's exact test. See also Supplemental Methods.

<sup>c</sup> Detection of small insertions or deletions: default settings of GATK Mutect2 were used. TLOD represents the log odds that an altered allele exists.

<sup>d</sup> Detection of small insertions or deletions: used settings of VarScan (SNP and indel): min-coverage 100, min-reads2 2, min-var-freq 0.00001, p-value 0.9, strand-filter 0.

<sup>e</sup> Somatic mutation is considered confirmed when confirmed by DCS (SNP) or Mutect2 AND VarScan (insertion or deletion).

<sup>f</sup> No somatic mutations present in primary tumor

Plasma									
Duplex consensus sequencing <sup>b</sup>					Mutect2 <sup>c</sup>		VarScan <sup>d</sup>		Summary plasma <sup>e</sup>
ALT_ reads	REF_ reads	ALT_ reads_ total	REF_ reads_ total	P-value	VOF	TLOD	VOF	P-value	
0	301	0	13523	1	NA	NA	NA	NA	Not confirmed
0	274	10	11901	1	NA	NA	NA	NA	Not confirmed
0	348	22	15437	1	NA	NA	NA	NA	Not confirmed
11	335	15	13802	1.84E-19	NA	NA	NA	NA	Confirmed
34	448	35	17506	1.66E-93	NA	NA	NA	NA	Confirmed
0	207	0	12577	1	NA	NA	NA	NA	Not confirmed
0	241	0	14407	1	NA	NA	NA	NA	Not confirmed
0	377	7	17283	1	NA	NA	NA	NA	Not confirmed
NA	NA	NA	NA	NA	4.4	220.42	2.82	3.509E-44	Confirmed
NA	NA	NA	NA	NA	6	796.58	4.31	2.147E-68	Confirmed
48	1491	55	13950	3.45E-66	NA	NA	NA	NA	Confirmed
NA	NA	NA	NA	NA	3.3	278.34	3.3	1.473E-58	Confirmed
NA	NA	NA	NA	NA	1.3	27.6	0.88	2.446E-11	Confirmed

**Supplementary Table 2.** Detected somatic mutations in plasma of all samples without tumor knowledge criteria

Sample	HPV <sup>a</sup>	Chromosome	Position	Gene	REF	ALT	Variante Type	VAF_DCS <sup>b</sup>	VAF_Mutect2 <sup>b</sup>	VAF_VarScan <sup>b</sup>	Present in tumor? <sup>c</sup>
590		chr2	202141692	CASP8	G	A	SNP	2.04	1.3	1.06	No
595		chr4	153249440	FBXW7	C	A	SNP	2.56	5.4	7.21	Yes
595		chr5	176637946	NSD1	C	G	SNP	0.85	6.4	8.12	No
595		chr9	139400061	NOTCH1	G	C	SNP	11.11	4.5	9.76	Yes
595		chr17	7577106	TP53	G	A	SNP	7.41	9.1	12.16	Yes
598		chr12	49427219	KMT2D	G	A	SNP	2.22	2	2.19	No
616	HPV16	chr2	202136245	CASP8	GCT	G	DEL	2.44	2.8	3.25	No
619	Negative	chr9	21974792	CDKN2A	G	A	SNP	0.51	1.2	1.22	No
626		chr4	187557993	FAT1	CA	C	DEL	4.76	2.9	2.83	Yes
633	Negative	chr9	21974506	CDKN2A	CA	C	DEL	2.38	1.5	1.78	Yes
635	Negative	chr4	153332667	FBXW7	C	A	SNP	0.88	3.1	3.45	Yes
635	Negative	chr9	139410512	NOTCH1	G	A	SNP	1.35	1.7	1.06	No
635	Negative	chr17	7577559	TP53	G	A	SNP	2.67	2.5	3.53	No
635	Negative	chr17	7578457	TP53	C	T	SNP	4.82	6.7	7.18	No
635	Negative	chr17	7579415	TP53	C	T	SNP	3.23	4.2	4.62	Yes
639	Negative	chr4	187525704	FAT1	C	T	SNP	1.12	2.7	2.54	Yes
639	Negative	chr17	7577538	TP53	C	A	SNP	1.22	5.4	6.34	Yes
642		chr5	176721936	NSD1	TC	T	DEL	1.18	1.4	0.8	No
647		chr14	23450536	AJUBA	C	CT	INS	0.67	3.2	2.38	Yes
647		chr17	7579380	TP53	AGG	A	DEL	3.23	3.6	4.02	Yes
649		chr3	178952085	PIK3CA	A	G	SNP	1.13	1.3	3.16	Yes
649		chr14	23451340	AJUBA	T	TC	INS	1.06	1.7	2.76	Yes
592	Unknown <sup>d</sup>	chr12	49438644	KMT2D	C	A	SNP	0.49	0.8342	0.74	NA
608		chr9	21971036	CDKN2A	C	A	SNP	1.54	2.8	4.02	NA
611	Negative	chr10	89717594	PTEN	G	C	SNP	0.82	0.9901	1.16	NA
615		chr4	187524181	FAT1	C	T	SNP	0.89	0.9843	0.9	NA
615		chr5	176707589	NSD1	A	T	SNP	2.03	1.6	1.53	NA

Abbreviations: ALT, alternative sequence; DCS, duplex consensus sequencing; DEL, deletion; INS, insertion; NA, not applicable; REF, reference sequence; SNP, single-nucleotide polymorphism; VAF, variant allele frequency

<sup>a</sup> HPV status only determined in oropharynx tumors.

<sup>b</sup> Variants were considered valid when called by three pipelines and VAF  $\leq$  25% to remove potentially remaining germline variants. See also Supplemental Methods.

<sup>c</sup> Confirmation of in plasma detected somatic mutation in primary tumor. Only reported when tumor DNA was available.

<sup>d</sup> No material available from primary tumor for HPV analysis.

**Supplementary Table 3.** Optimization of circulating tumor DNA detection by additional sequencing of primary and backup sequencing library

Patient	CHROM	POS	REF	ALT	ALT reads Plasma 1st capture	ALT reads Plasma resequencing 1st capture	ALT reads Plasma 2nd capture on pre-PCR backup	P-value Plasma 1st capture	P-value Plasma all sequenced captures
606	chr17	7578188	C	A	2	NA <sup>a</sup>	1	0.001	0
606	chr3	178921472	G	A	1	NA <sup>a</sup>	0	0.115	0.246
607	chr9	139400000	C	A	2	NA <sup>a</sup>	0	0.001	0.004
618	chr17	7578550	G	T	4	2	2	0	0
627	chr12	49421581	C	T	0	0	2	1	0.345
627	chr12	49421705	C	A	0	0	2	1	0.018
629	chr17	7577144	A	G	1	NA <sup>a</sup>	0	0.23	0.348
636	chr3	178952085	A	G	1	NA <sup>a</sup>	0	0.124	0.366

a. resequencing of 1st capture was not performed

Abbreviations: ALT, alternative sequence; CHROM, chromosome; NA, not applicable; POS, position; REF, reference sequence

## REFERENCES

1. Leemans CR, Braakhuis BJ, Brakenhoff RH. The molecular biology of head and neck cancer. *Nat Rev Cancer* 2011; 11(1): 9-22.
2. Diehl F, Schmidt K, Choti MA, et al. Circulating mutant DNA to assess tumor dynamics. *Nat Med* 2008; 14(9): 985-90.
3. Bettgowda C, Sausen M, Leary RJ, et al. Detection of circulating tumor DNA in early- and late-stage human malignancies. *Sci Transl Med* 2014; 6(224): 224ra24.
4. Wang Y, Springer S, Mulvey CL, et al. Detection of somatic mutations and HPV in the saliva and plasma of patients with head and neck squamous cell carcinomas. *Sci Transl Med* 2015; 7(293): 293ra104.
5. de Roest RH, Mes S, Brink A, et al. Molecular characterization of locally relapsed head and neck cancer after concomitant chemoradiotherapy. *Clin Cancer Res* 2019; 25(23): 7256-65.
6. Stahlberg A, Krzyzanowski PM, Jackson JB, Egyud M, Stein L, Godfrey TE. Simple, multiplexed, PCR-based barcoding of DNA enables sensitive mutation detection in liquid biopsies using sequencing. *Nucleic Acids Res* 2016; 44(11): e105.
7. van Ginkel JH, Huibers MMH, van Es RJJ, de Bree R, Willems SM. Droplet digital PCR for detection and quantification of circulating tumor DNA in plasma of head and neck cancer patients. *BMC Cancer* 2017; 17(1): 428.
8. Schrock A, Leisse A, de Vos L, et al. Free-Circulating Methylated DNA in Blood for Diagnosis, Staging, Prognosis, and Monitoring of Head and Neck Squamous Cell Carcinoma Patients: An Observational Prospective Cohort Study. *Clin Chem* 2017; 63(7): 1288-96.
9. Schirmer MA, Beck J, Leu M, et al. Cell-Free Plasma DNA for Disease Stratification and Prognosis in Head and Neck Cancer. *Clin Chem* 2018; 64(6): 959-70.
10. Leemans CR, Snijders PJF, Brakenhoff RH. The molecular landscape of head and neck cancer. *Nat Rev Cancer* 2018; 18(5): 269-82.
11. Cancer Genome Atlas N. Comprehensive genomic characterization of head and neck squamous cell carcinomas. *Nature* 2015; 517(7536): 576-82.
12. Poell JB, Mendeville M, Sie D, Brink A, Brakenhoff RH, Ylstra B. ACE: Absolute Copy number Estimation from low-coverage whole-genome sequencing data. *Bioinformatics* 2018.
13. Scheinin I, Sie D, Bengtsson H, et al. DNA copy number analysis of fresh and formalin-fixed specimens by shallow whole-genome sequencing with identification and exclusion of problematic regions in the genome assembly. *Genome Res* 2014; 24(12): 2022-32.
14. Straver R, Sistermans EA, Holstege H, Visser A, Oudejans CB, Reinders MJ. WISECONDOR: detection of fetal aberrations from shallow sequencing maternal plasma based on a within-sample comparison scheme. *Nucleic Acids Res* 2014; 42(5): e31.
15. Smeets SJ, Hesselink AT, Speel EJ, et al. A novel algorithm for reliable detection of human papillomavirus in paraffin embedded head and neck cancer specimen. *Int J Cancer* 2007; 121(11): 2465-72.
16. Lee JY, Garcia-Murillas I, Cutts RJ, et al. Predicting response to radical (chemo)radiotherapy with circulating HPV DNA in locally advanced head and neck squamous carcinoma. *Br J Cancer* 2017; 117(6): 876-83.
17. Chera BS, Kumar S, Beaty BT, et al. Rapid Clearance Profile of Plasma Circulating Tumor HPV Type 16 DNA during Chemoradiotherapy Correlates with Disease Control in HPV-Associated Oropharyngeal Cancer. *Clin Cancer Res* 2019; 25(15): 4682-90.
18. Chera BS, Kumar S, Shen C, et al. Plasma Circulating Tumor HPV DNA for the Surveillance of Cancer Recurrence in HPV-Associated Oropharyngeal Cancer. *J Clin Oncol* 2020; 38(10): 1050-8.
19. Nguyen B, Meehan K, Pereira MR, et al. A comparative study of extracellular vesicle-associated and cell-free DNA and RNA for HPV detection in oropharyngeal squamous cell carcinoma. *Sci Rep* 2020; 10(1): 6083.
20. Hanna GJ, Supplee JG, Kuang Y, et al. Plasma HPV cell-free DNA monitoring in advanced HPV-associated oropharyngeal cancer. *Ann Oncol* 2018; 29(9): 1980-6.
21. Schwaederle M, Chattopadhyay R, Kato S, et al. Genomic Alterations in Circulating Tumor DNA from Diverse Cancer Patients Identified by Next-Generation Sequencing. *Cancer Res* 2017; 77(19): 5419-27.
22. Perdomo S, Avogbe PH, Foll M, et al. Circulating tumor DNA detection in head and neck cancer: evaluation of two different detection approaches. *Oncotarget* 2017; 8(42): 72621-32.
23. Galot R, van Marcke C, Helaers R, et al. Liquid biopsy for mutational profiling of locoregional recurrent and/or metastatic head and neck squamous cell carcinoma. *Oral Oncol* 2020; 104: 104631.

24. Parpart-Li S, Bartlett B, Popoli M, et al. The Effect of Preservative and Temperature on the Analysis of Circulating Tumor DNA. *Clin Cancer Res* 2017; 23(10): 2471-7.
25. Newman AM, Bratman SV, To J, et al. An ultrasensitive method for quantitating circulating tumor DNA with broad patient coverage. *Nat Med* 2014; 20(5): 548-54.
26. Jiang P, Chan CW, Chan KC, et al. Lengthening and shortening of plasma DNA in hepatocellular carcinoma patients. *Proc Natl Acad Sci U S A* 2015; 112(11): E1317-25.
27. Mouliere F, Chandrananda D, Piskorz AM, et al. Enhanced detection of circulating tumor DNA by fragment size analysis. *Sci Transl Med* 2018; 10(466).
28. Venkatraman ES, Olshen AB. A faster circular binary segmentation algorithm for the analysis of array CGH data. *Bioinformatics* 2007; 23(6): 657-63.
29. Hesselink AT, Berkhof J, van der Salm ML, et al. Clinical validation of the HPV-risk assay, a novel real-time PCR assay for detection of high-risk human papillomavirus DNA by targeting the E7 region. *J Clin Microbiol* 2014; 52(3): 890-6.
30. Bolger AM, Lohse M, Usadel B. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* 2014; 30(15): 2114-20.
31. Li H, Durbin R. Fast and accurate long-read alignment with Burrows-Wheeler transform. *Bioinformatics* 2010; 26(5): 589-95.
32. Li H, Handsaker B, Wysoker A, et al. The Sequence Alignment/Map format and SAMtools. *Bioinformatics* 2009; 25(16): 2078-9.
33. Koboldt DC, Zhang Q, Larson DE, et al. VarScan 2: somatic mutation and copy number alteration discovery in cancer by exome sequencing. *Genome Res* 2012; 22(3): 568-76.
34. McKenna A, Hanna M, Banks E, et al. The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res* 2010; 20(9): 1297-303.
35. Genomes Project C, Auton A, Brooks LD, et al. A global reference for human genetic variation. *Nature* 2015; 526(7571): 68-74.
36. Ramos AH, Lichtenstein L, Gupta M, et al. Oncotator: cancer variant annotation tool. *Hum Mutat* 2015; 36(4): E2423-9.
37. Cingolani P, Platts A, Wang le L, et al. A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of *Drosophila melanogaster* strain w1118; iso-2; iso-3. *Fly (Austin)* 2012; 6(2): 80-92.
38. Kennedy SR, Schmitt MW, Fox EJ, et al. Detecting ultralow-frequency mutations by Duplex Sequencing. *Nat Protoc* 2014; 9(11): 2586-606.



# 6

## **Molecular Characterization of Locally Relapsed Head and Neck Cancer after Concomitant Chemoradiotherapy**

Reinout H. de Roest, Steven W. Mes, Jos B. Poell, Arjen Brink,  
Mark A. van de Wiel, Elisabeth Bloemena, Elena Thai, Tito Poli,  
C. René Leemans, Ruud H. Brakenhoff

**Clinical Cancer Research 2019; 25(23): 7256-65**

## **ABSTRACT**

### **Purpose**

To investigate the pathobiological origin of local relapse after chemoradiotherapy, we studied genetic relationships of primary tumors (PT) and local relapses (LR) of patients treated with chemoradiotherapy.

### **Experimental design**

First, low-coverage whole genome sequencing was performed on DNA from 44 biopsies of resected head and neck squamous cell carcinoma (HNSCC) specimens (median 3 biopsies/tumor) to assess suitability of copy number alterations (CNAs) as biomarker for genetic relationships. CNAs were compared within and between tumors and an algorithm was developed to assess genetic relationships with consideration of intratumor heterogeneity. Next, this CNA-based algorithm was combined with target enrichment sequencing of genes frequently mutated in HNSCC to assess the genetic relationships of paired tumors and LRs of patients treated with chemoradiotherapy.

### **Results**

Genetic relationship analysis using CNAs could accurately (96%) predict tumor biopsy pairs as patient-matched or independent. However, subsequent CNA analysis of PTs and LRs after chemoradiotherapy suggested genetic relationships in only 20% of cases, and absence in 80%. Target enrichment sequencing for mutations confirmed absence of any genetic relationship in half of the paired PTs and LRs.

### **Conclusions**

There are minor variations in CNA profiles within different areas of HNSCC tumors and many between independent tumors, suggesting that CNA profiles could be exploited as a marker of genetic relationship. Using CNA profiling and mutational analysis of cancer driver genes, relapses after chemoradiotherapy appear to be partially genetically related to the corresponding PTs, but seem often genetically unrelated. This remarkable observation warrants further studies and will impact therapeutic innovations and prognostic modeling when using index tumor characteristics.

## INTRODUCTION

Head and neck squamous cell carcinomas (HNSCC) are among the most common incident cancers worldwide, with more than 600,000 newly diagnosed cases each year<sup>1,2</sup>. HNSCC arises in the mucosal lining of the upper aerodigestive tract encompassing the oral cavity, oropharynx, hypopharynx, and larynx. Most tumors arise in the oral cavity (30%–40%), followed by the larynx (30%–35%)<sup>3</sup>. Classic risk factors for HNSCC are smoking and excessive alcohol consumption. Also, patients with inherited genetic predispositions such as Fanconi anemia have an increased risk for HNSCC<sup>4</sup>. For tumors arising in the oropharynx (OP) infection with high-risk human papillomavirus (HPV), in particular HPV16, is a more recently emerged risk factor<sup>5,6</sup>.

Unfortunately, the majority (~60%) of patients with HNSCC present with advanced stage disease, meaning that the malignancies have invaded neighboring anatomic structures and/or have spread to the lymph nodes in the neck. Clinical management of these patients consists of surgery followed by postoperative (chemo) radiotherapy, locoregional radiotherapy, or concurrent chemoradiotherapy with surgical salvage for residual or relapsed cancers. Upfront chemoradiotherapy is applied when surgical resection is considered too invasive, the tumor is deemed very sensitive to nonsurgical modalities or when severe problems with swallowing and speech are expected after surgery. This multimodality approach has led to improved outcomes in terms of quality of life in patients with advanced stage disease, but prognosis still leaves much to be desired<sup>2,7</sup>. The most critical event in the course of the disease is the development of local relapses (LR)<sup>7–10</sup>. After definitive treatment with concurrent chemoradiotherapy, the local failure rate at 3 years of follow-up is between 15% and 50%<sup>7,11,12</sup>.

There are several pathobiological origins for LRs<sup>13,14</sup>. A local relapse is clinically defined as a lesion that develops within 3 years after and within 2 cm from the treated primary tumor, relapses developing spatially or temporally more distinct are considered to be second primary tumors (SPT)<sup>15</sup>. Our current knowledge on the pathobiological mechanisms of the development of relapses is mainly based on molecular studies of resected oral tumors and their surgical margins<sup>13,14,16</sup>. Two mechanisms seem to play a role. First, even when the surgical margins are histologically diagnosed as tumor-free, tumor cells may have remained behind, also called minimal residual cancer (MRC), and these could give rise to an LR. A genetic relationship between tumor and LR is then expected. Second, fields of preneoplastic cells that preceded and surrounded the primary tumor might stay behind unnoticed after tumor excision, and develop in LRs. These LRs have also been depicted as “second field tumors” (SFT)<sup>13,14,16</sup>. When studied for genetic alterations, these second field tumors share the early changes with the index tumor, such as TP53 mutations, CDKN2A mutations and 9p losses, but not the late changes. Finally, patients with HNSCC are also at risk for second primary tumors, and as these often originate independently from the index tumor, no genetic relationship is expected.

To date, the genetic origin of local relapse after chemoradiotherapy has not been studied very well, and we wondered which mechanisms might play a role. As surgical margins are not available from these patients, the studies have to rely on genetic characterization of the index tumors and local relapses to investigate genetic relationships. Caveats in these analyses are both intratumor genetic heterogeneity, which can be determined by multiple biopsies or single-cell analyses, and treatment-induced genetic changes.

Recent research has demonstrated a role for intratumoral genetic heterogeneity in HNSCC development and progression<sup>17–19</sup>, which may indeed contribute to LR<sup>20</sup>. Therefore, genetic assays to study genetic relationships that are validated by analyzing multiple biopsies of primary tumors, would be of great value to study genetic relationships between index tumors and LRs. Both chromosomal changes and point mutations in bona fide cancer driver genes can be used as markers of genetic relationship. Chromosomal changes are easiest determined, but their discriminating power to assess genetic relationships is more limited. Algorithms that consider chromosome breakpoints and other genetic changes in HNSCC often relate to losses of complete chromosome arms, and are consequently comparable in many different tumors, losing discriminative power.

Somatic mutations in cancer driver genes are therefore considered to be most accurate markers to assess genetic relationships as they are very unique<sup>21,22</sup>. However, low-quality DNA from formalin-fixed paraffin-embedded (FFPE) specimens and PCR errors may cause false positive findings, sometimes with a high variant allele frequency (VAF), which could hamper the interpretation of the data. Therefore, a combination of both approaches might be the preferred strategy to assess genetic relationships.

The aims of this study are to evaluate copy number alterations (CNAs) as biomarker of genetic relationship by studying intratumor genetic heterogeneity using multiple spatially distinct biopsies obtained from individual tumors. Second, genetic relationships between LRs and index tumors in patients treated with chemoradiotherapy were investigated with this approach, and findings extended with mutation analysis of head and neck cancer driver genes.

## **MATERIALS AND METHODS**

### **Tumors analyzed to develop a CNA-based algorithm tumor**

To study the suitability of copy number alterations to assess genetic relationships we had to take into account intratumor genetic heterogeneity within distinct samples from one tumor (intratumor) and comparable alterations (e.g., frequent 3p losses) between different tumors (intertumor). We therefore collected and analyzed 44 biopsies from 13 resected primary oral tumors (3–5 biopsies/tumor). The samples were obtained from the resection specimen available at the department of Pathology at Amsterdam UMC (Amsterdam, the Netherlands), location VUmc, or at the University Hospital Parma (Parma, Italy), and were snap-frozen. All primary tumors were large enough to allow for widely spaced biopsies; care was taken to avoid areas of necrosis and ulceration; all sampled areas were documented histologically. None of the patients had received treatment prior to definitive surgery.

### **Tumors and relapses of chemoradiotherapy-treated patients for assessment of genetic relationships**

We included all patients who developed an LR after cisplatin-based chemoradiotherapy with curative intent for an advanced stage HPV-negative oropharyngeal, hypopharyngeal, or laryngeal squamous cell carcinoma at Amsterdam UMC, location VUmc, between 2009 and 2014. Criteria for LR were residual or recurrent tumor within 2 cm from and within 3 years after the index tumor. All patients were evaluated by medical history, physical examination, examination under general anesthesia/panendoscopy, and imaging (e.g., CT and/or MRI), staging was according to the 7th edition of the American Joint Committee on Cancer staging manual. Only HPV-negative OPSCC tumors were included, determined by p16 immunohistochemical (IHC) staining as described by Smeets and colleagues<sup>23</sup>. Patients received intensity modulated radiotherapy (IMRT) concurrent with cisplatin (100 mg/m<sup>2</sup> every 3 weeks), with a curative intent.

In total, 113 patients with an HPV-negative oropharyngeal, hypopharyngeal, or laryngeal squamous cell carcinoma were treated within the given timespan with curative-intent chemoradiation at our clinic. In total, 20 (17.7%) of these 113 patients developed a local relapse (12) according to the clinical criteria described above or a regional relapse in the neck (8). From the 10 of 12 patients with an LR, we could retrieve the paired FFPE biopsies from the pathology archive. Clinical characteristics and follow-up period of the patients of whom the tumor and LRs were analyzed did not significantly differ from patients who developed a regional relapse or of whom tumor–LR pairs could not be analyzed (Supplementary Table 1). FFPE material with sufficient tumor purity, assessed by inspection of the hematoxylin/eosin–stained slide, was used for DNA isolation. When required, tumor percentage was enriched by macrodissection and was between 45% and 90%. Definitive tumor purity was estimated by ACE (see below).

## Ethical approval

Studies were carried out in accordance to the Declaration of Helsinki. According to the decisions of the Institutional Review Board and when patients consented, these studies were performed following the guidelines of the Code of Conduct for Human Tissue and Medical Research (<https://www.federa.org/codes-conduct>) and the EU General Data Protection Regulation.

## Whole-genome low-coverage sequencing for copy number alterations

Genomic DNA was obtained from the snap-frozen samples using the Life PureLink Kit (Thermo Fisher Scientific) and from the FFPE samples with either the Life PureLink Kit or the QIAamp DNA Micro Kit (Qiagen). DNA yield was analyzed on a Qubit 2.0 (Thermo Fisher Scientific). In total, a minimum of 200 or 250 ng DNA from snap-frozen or FFPE material, respectively, was used as input for library preparation. First, DNA was sheared on a Covaris S2 (Covaris). Library preparation for DNA isolated from FFPE was conducted using the TruSeq Nano Kit (Illumina), following the manufacturer's instructions. In short, sheared DNA was end-repaired, the 3' ends were adenylated, indexed adapters were ligated, and the library was amplified with 10 PCR cycles. Libraries of DNA from frozen specimen were prepared with SOLiD reagents (Applied Biosystems) and amplification with 12 PCR cycles. The quality of the libraries was verified on a Bioanalyzer DNA 7500 chip (Agilent Technologies). Whole-genome sequencing libraries of 20 (FFPE) or 24 (fresh frozen) samples were pooled equimolarly and sequenced on a single lane of a HiSeq 2500 (Illumina) in a single-read 50-cycle run mode (SR50). Raw sequencing reads were mapped to the human reference genome (build GRCh37/hg19) with BWA<sup>24</sup>. Data analysis was performed using the "QDNAseq" R package (version 1.16.0)<sup>25</sup>. Data were filtered against a blacklist of regions known to be germline copy number variants<sup>26</sup>, and log<sub>2</sub> ratios were median-normalized and segmented with the Circular Binary Segmentation algorithm<sup>27</sup>. Estimates of tumor purity and absolute copy numbers were obtained through the "ACE" R package (version 0.99.6). Calling of the segments was done using CGHcall (version 2.42.2)<sup>28</sup>.

6

## Algorithm development to assess genetic relationships by low-coverage NGS

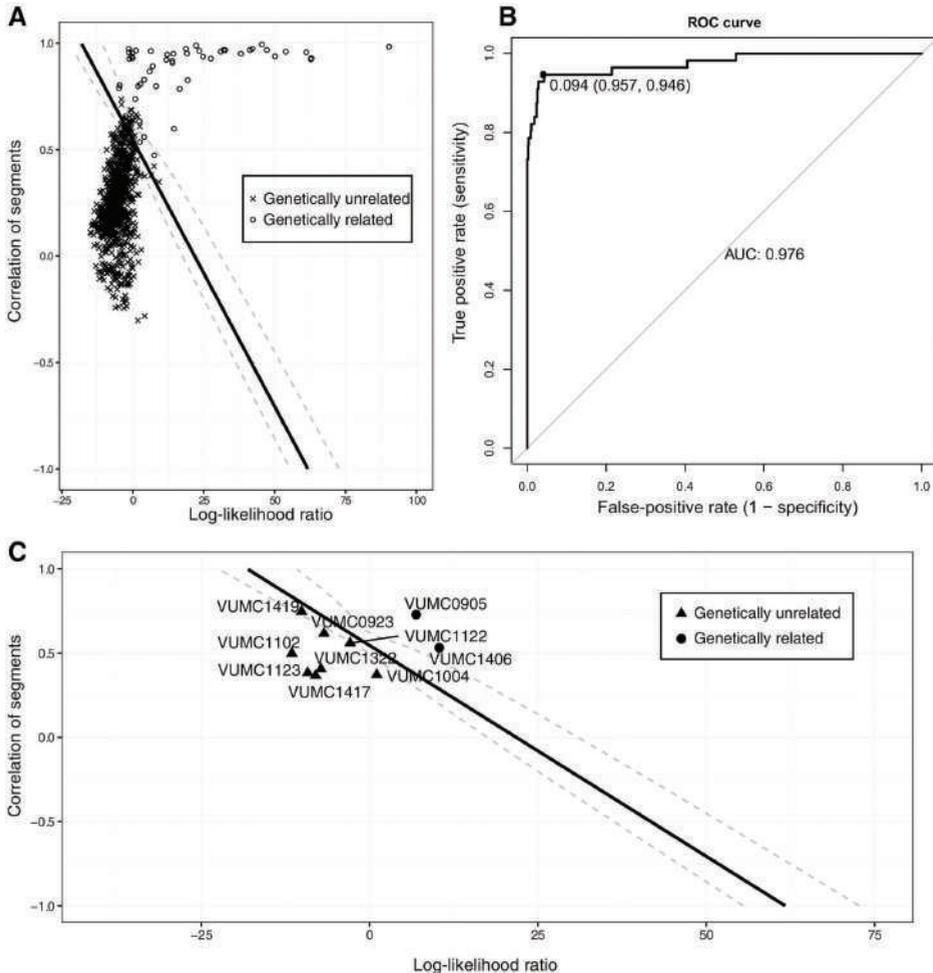
The segmented data of CGHcall were used to compare copy number profiles. To compare copy number profiles and take the specific breakpoints into consideration, we calculated the log-likelihood ratio (LLR) using the "Clonality" R package (version 1.28.0), which quantifies the likeliness of two tumors being clonally related<sup>29</sup>. In addition, we did a correlation analysis of segments as described by Sie and colleagues<sup>30</sup>. Significance of both the LLR and correlation score for discriminating pairs and nonpairs was assessed by a permutation test. Under the null-hypothesis, the scores (LLR or correlation) of pairs and nonpairs are exchangeable. Therefore, we created a permutation null-distribution for the difference in mean score between pairs and nonpairs, where the pairs were created by permutation of the sample profiles. The P value for the observed difference in mean scores was then computed using this null-distribution.

Calculation of both assays were plotted in 2D and the optimal cutoff determined to define genetic relationships. We used the multi-biopsied OSCC samples to develop a classification model for genetic relatedness. For evaluating the performance of the predictive model, leave-2-out cross-validation (CV) was used as other analyses such as bootstrapping demand larger sample sizes. For evaluation by CV, it is essential that the left-out samples are completely independent of the training samples. Because biopsies from the same patient are related, we did not perform CV on the pairs, but in fact on the patients. In short, all samples of 2 patients were left out for testing. The profiles of all remaining patients were used for training the logistic regression model with the pairing indicator as response variable and LLR and correlation score as covariates. The model was then applied to all paired and nonpaired samples of the two left-out patients, and this was repeated for all possible sample couples. The predicted probability scores and true pairing indicators were then used to

produce the ROC-curve<sup>31</sup>, and the corresponding AUC. The optimal cutoff was determined by the highest sensitivity and specificity combination using the Youden index<sup>32</sup>. This threshold was applied to the testing cohort to test the performance of the model.

### **Target enrichment sequencing for mutations**

In addition, mutation analysis of the index tumors and relapses of chemoradiotherapy-treated patients was performed. We selected a panel of 12 cancer driver genes which are frequently altered in HNSCC according to TCGA data (Supplementary Table 2)<sup>33</sup>. The sequencing libraries as described above were used for target-enrichment using the SeqCap EZ (Roche) protocol. We aimed for equimolar pooling of 50 ng of all twenty samples to a combined DNA amount of 1 µg. Because of low yields in some samples, input of individual samples ranged from 6.6 to 83.1 ng, which obviously impacted coverage. The library was hybridized to the SeqCap Oligo Pool. Streptavidin beads were used to capture the complex of oligos and DNA fragments, and unbound fragments were washed away. The captured fragment pool was amplified by 14 PCR cycles, and sequenced on a single lane of a HiSeq 2500 (Illumina) in a paired-end 150-cycle run mode (PE150). Raw sequencing data were mapped to the human genome (build GRCh37/hg19) with BWA and duplicate reads were marked for removal using Picard (<http://broadinstitute.github.io/picard>). Variants were called using multiple tools, including GATK MuTect2<sup>34</sup> and Samtools mpileup<sup>35</sup> with VarScan 2<sup>36</sup>. To assess mutations, we required a minimum coverage of 70 unique reads. We filtered all variants with a frequency higher than 0.01% in the European population of the 1000 Genome Project cohort<sup>37</sup>. In addition, we manually filtered variants that were apparent in both the primary and the relapse with a VAF of around 50%, independent of the tumor purity (germline SNPs). Estimated tumor purity, absolute copy numbers (computed with ACE), and VAF obtained with MuTect2 was used for calculation of the 95% confidence interval (CI) of the VAF and the absolute number of variant (mutant) alleles (aVA) in the tumor genome. To determine the genetic relation of primary and relapse, we looked at the cooccurrence of variants found in the primary, which were present in at least 33.3% of the tumor-derived alleles as determined by ACE, implying that the variant occurs in the larger part of the tumor cell population. Furthermore, we reviewed each variant manually using the IGV viewer to ensure that no false positives were called by the algorithms. Finally, variants that appeared to be specific for either the primary tumor or the relapse sample were cross-checked for low coverage reads to ensure that minor subclones in either the index tumor or relapse were not missed.



**Figure 1.** Classification of sample pairs based on comparison of CNAs.

A, Validation data of the model: x-axis reflects the log likelihood ratio of a tumor pair, y-axis the correlation of segments. Black solid line indicates the classification cutoff with the 95 CI in dotted lines. The circles represent sample pairs of the same tumor (genetically related), and the X signs are independent sample pairs. B, AUC curve of our genetic relationship algorithm, with our optimal cut-off point. C, The CNA-based classification model was applied on primary tumor and LR samples of patients treated for cure with chemoradiotherapy. The pairs indicated as black dots are classified by our algorithm as genetically related, those indicated as black triangles as genetically unrelated.

## RESULTS

### Genetic relationship analysis by copy number profiles

As a first step, we developed and tested an algorithm for assessing genetic relationships on basis of copy number alterations, taking intratumoral genetic heterogeneity into account. We collected multiple biopsies of surgical specimen and isolated DNA. We performed whole-genome low-coverage sequencing to generate copy number profiles for 44 samples of 13 patients. Similarity of profiles was assessed by (i) the calculation of a LLR, which is based on the concordance of gains and losses in two samples and (ii) a correlation score of segments of all intra- and interpatient combinations of tumor pairs. Calculations of the two methods were plotted in a 2D graph. We made a total number of 946 comparisons, of which 56 were intratumor comparisons and 890 intertumor comparisons. The patient-matched intratumor comparisons had a higher mean correlation

coefficient ( $0.83 \pm 0.18$  vs.  $0.25 \pm 0.18$ ) and a higher mean log-likelihood ratio ( $17.3 \pm 21.7$  vs.  $-5.5 \pm 3.5$ ) than the comparisons between different tumor pairs. Both means differed significantly ( $P < 0.001$ ) analyzed by permutation test (10,000 iterations). To classify sample pairs as genetically related or genetically unrelated, we used a logistic regression model analysis on the basis of the LLR and the correlation score together. The probability of genetic relation by CNA profiles could be predicted by the equation:  $-13.11 + 19.87$  (correlation score)  $+ 0.50$  (LLR) in our training cohort. Across patients, the AUCs ranged from 93% to 100%, as assessed by comparing biopsy couples within a given patient with couples when biopsies were from different patients. An optimal cutoff at  $P(Y) > 0.094$  was determined by the highest combination of sensitivity and specificity (Youden index) of the classification. Both the correlation score and LLR were significant predictors of genetic relationship ( $P < 0.001$ ). Our model could identify sample pairs in our testing cohort as being biopsies from the same tumor with a sensitivity of 95%, a specificity of 96%, and an overall accuracy of 96% (Figure 1A and B).

### **Copy number profiles of primary tumors and local relapses after chemoradiotherapy**

Because the developed algorithm using CNAs appeared to be sufficiently accurate to investigate genetic relationships within tumors and between independent tumors, we studied paired primary tumors and local relapses of chemoradiation-treated HNSCC patients likewise. From a consecutive cohort of 113 patients with HPV-negative tumors treated with chemoradiotherapy between 2009 and 2014, 20 cases developed a local or regional relapse within a median follow-up time of 35.7 months (1.05–85.3; Supplementary Table 3). In 10 of 12 cases with a local relapse, the paired biopsies could be retrieved from the pathology archive. The studied cohort of these 10 cases with local relapse consisted of predominantly male patients (80%) all with locally advanced disease (T3–T4), 60% of whom had a primary tumor in the oropharynx (all HPV-negative). On the basis of imaging during follow-up, two patients were diagnosed with residual disease after treatment, which was confirmed by histopathologic assessment of a biopsy or the specimen after salvage surgery (Table 1). In the remaining 8 patients, locoregional control seemed achieved but relapses occurred nonetheless. The median time to diagnosis of residual or recurrent disease, both considered as local relapses, was 8.2 months (5.8–35.1). As expected for patients with relapsed HNSCC, they had a poor overall survival.

From both primary tumors and local relapses, copy number profiles were established by low-coverage whole-genome sequencing. The copy number profiles showed the characteristic alterations for HNSCC with losses and gains at 3p, 8p, 9p, and 11q. Furthermore, a high frequency of amplifications and gains was observed in regions containing PIK3CA (85%), EGFR (65%), and CCND1 (60%; Figure 2).

We applied the genetic relationship algorithm described above to the corresponding pairs of primary tumors and relapses. Primary tumor and relapse of two patients were designated by the algorithm as genetically related, VUMC0905 and VUMC1406, while the eight other pairs were identified as genetically unrelated, although two were borderline and within the 0.95 CI of the cut-off value (Figure 1C).

The result that six or maybe even eight of the local relapses in chemoradiotherapy-treated patients seemed not genetically related to the index tumors by copy number profile analysis was unexpected. This could imply that despite the results reported above, CNA profiling is less suited for assessment of genetic relationships of index tumors and LRs after chemoradiotherapy, also because the biological context is very different. Therefore, we decided to complement the data by mutation analysis of genes commonly mutated in HNSCC and that are assumed cancer driver genes (Supplementary Table 2). All 10 paired tumor and relapse samples were sequenced at a median sequencing coverage of 527 $\times$ , but with a large range (19 – 3,790) depending on the sample quality and nucleotide position. In one sample, DNA quality was too poor and coverage too low to pass our quality checks, and we lost that pair for mutation analysis, but it was convincingly genetically related by copy number analysis. In the nine remaining tumor pairs, sequence variants were found in 10 of 12 tested genes, and the median number of variants were 4 (0–8) and 3 (2–8), in respective primary tumors and relapses.

**Table 1.** Patient characteristics

General characteristics				Primary tumor	Treatment primary tumor			Relapse		Cohort	
Patient ID	Age	Sex	Site <sup>a</sup>	TNM stage	Disease stage	RT dose (Gy) <sup>b</sup>	Cisplatin dose (mg/m <sup>2</sup> )	Type <sup>d</sup>	Interval (months)		
VUMC0905	63	M	LA	T3N2cM0	IVA	70	300	Rec	8.2	<i>CRT cohort</i>	
VUMC0923	64	M	OP	T4bN1M0	IVB	70	300	Rec	11.5		
VUMC1004	62	M	HP	T3N2bM0	IVA	70	300	Rec	35.1		
VUMC1102	76	F	HP	T4bN0M0	IVB	70	300	Rec	6		
VUMC1122	43	F	OP	T4aN2cM0	IVA	70	300	Res	6.3		
VUMC1123	62	M	OP	T4aN2bM0	IVA	70	100	Rec	16.6		
VUMC1322	64	M	OP	T3N2bM0	IVA	70	100 <sup>c</sup>	Res	5.8		
VUMC1406	56	M	HP	T3N2cM0	IVA	70	300	Rec	6.3		
VUMC1417	66	M	OP	T3N1M0	III	70	280	Rec	26.4		
VUMC1419	71	M	OP	T4aN0M0	IVA	70	280	Rec	17.2		
ITGH_3	68	M	OC	T1N0M0	I	NA	NA	NA	NA		<i>Cohort for training and validation of CNA algorithm</i>
ITGH_4	80	M	OC	T4aN2bM0	IVA	NA	NA	NA	NA		
ITGH_5	64	M	OC	T4aN0M0	IVA	NA	NA	NA	NA		
ITGH_6	69	M	OC	T4aN2bM0	IVA	NA	NA	NA	NA		
ITGH_7	67	M	OC	T4aN1M0	IVA	NA	NA	NA	NA		
ITGH_10	53	F	OC	T4aN2cM0	IVA	NA	NA	NA	NA		
ITGH_14	62	M	OC	T4aN0M0	IVA	NA	NA	NA	NA		
ITGH_15	72	F	OC	T4aN2bM0	IVA	NA	NA	NA	NA		
NPR_201	73	M	OC	T4aN0Mx	IVA	NA	NA	NA	NA		
NPR_202	66	M	OC	T4aN2bMx	IVA	NA	NA	NA	NA		
NPR_203	82	F	OC	T4aNxMx	IVA	NA	NA	NA	NA		
NPR_205	89	F	OC	T4aN2cMx	IVA	NA	NA	NA	NA		
NPR_206	86	M	OC	T2N2bMx	IVA	NA	NA	NA	NA		

<sup>a</sup> LA, larynx; OP, oropharynx; HP, hypopharynx; OC, oral cavity.

<sup>b</sup> Cumulative dose on primary tumor.

<sup>c</sup> Switched to carboplatin after development of adverse event.

<sup>d</sup> Rec: clinically local recurrence; Res: clinically residual disease.

As examples, we displayed two cases in Table 2, one that we consider as genetically related and one apparently not. We applied the ACE algorithm on the data that corrects for stromal contamination and ploidy, and provides an estimate of the absolute number of alleles in the tumor cells, and the number of variant alleles. When we analyzed the variants with high number of absolute variant alleles (aVA) and that thus were present in the very large majority of the tumor/relapse cell population, we found that these were shared in case VUMC1406 between primary and relapse, in accordance with the result of the CNA profiling. Case VUMC1406 had a TP53 (c.716A>G) variant that occurred in all of its tumor alleles of the primary tumor (absolute tumor alleles/variant alleles; 0.98/1.02) and in an equal genetic composition in the relapse. This suggests that in this particular case, the original population of cells in the index tumor had recurred. However, we also identified somatic variants in apparently a smaller fraction of tumor cells (range: 0.05%–27.9%) in either the primary tumor or the relapse, which likely reflects slight variations in tumor cell populations. The same was observed in all other cases (VUMC1122, –1419, –1102, and –1322) that showed a genetic relationship within tumor and relapse pairs based on one or more cancer driver mutations with a high aVA in the index tumor (Figure 2; Supplementary Table 4).

In five of the nine cases, there seemed to be no relation at all between the tumor and relapse. As an example, in case VUMC1417 the primary tumor had two mutations in TP53, one with a VAF of 83.4% (c.1022\_1023delTC) and one with 7.0% (c.1013A>G). The mutation with a frequency of 83.4% should certainly be considered as a cancer driver mutation but it lacks completely in the relapse with 2,779 reads on that position. Likewise, the relapse of VUMC1417 had double mutations in FAT1 and TP53, but these variants all lacked in the primary tumor, although it should be mentioned that the coverage in the tumor was somewhat lower on the respective positions but still 96, 68, 181, and 151 reads. All four remaining cases followed this pattern of mutually exclusive mutations: VUMC1004, -0923, -1417, and -1123. Hence, these relapses originated from very minor subclones in the tumor, or they had an independent genetic origin. A striking example is case VUMC1123 with a high VAF mutation in TP53 (c.833C>T) in the tumor, and no mutant reads in 1,495 reads on that position in the relapse, while the relapse had another high VAF TP53 (c.574C>T) mutation with no mutant reads in 984 reads on that position in the index tumor.

**Table 2.** Example data of variants found in two tumor pairs

Patient	Gene	Type	Mutation	
			Mutation	Protein change
<b>VUMC1406 genetically related</b>	AJUBA	Missense	c.733G>A	p.G245R
	AJUBA	Missense	c.865G>A	p.G289S
	KMT2D	Missense	c.2033C>T	p.S678F
	KMT2D	Missense	c.8896C>T	p.R2966W
	NSD1	Missense	c.1267G>A	p.A423T
	NSD1	Missense	c.3034C>G	p.R1012G
	PIK3CA	Missense	c.3140A>G	p.H1047R
	TP53	Missense	c.716A>G	p.N239S
<b>VUMC1417 genetically unrelated</b>	FAT1	Frameshift_del	c.8013_8016delCCTT	p.FF2671fs
	FAT1	Frameshift_ins	c.3445_3446insA	p.M1149fs
	TP53	Missense	c.1013A>G	p.H338R
	TP53	Frameshift_del	c.1022_1023delTC	p.F341fs
	TP53	Missense	c.761T>A	p.I254N
	TP53	Frameshift_del	c.1019_1028delTGTTCCGAGA	p.MFRE340fs

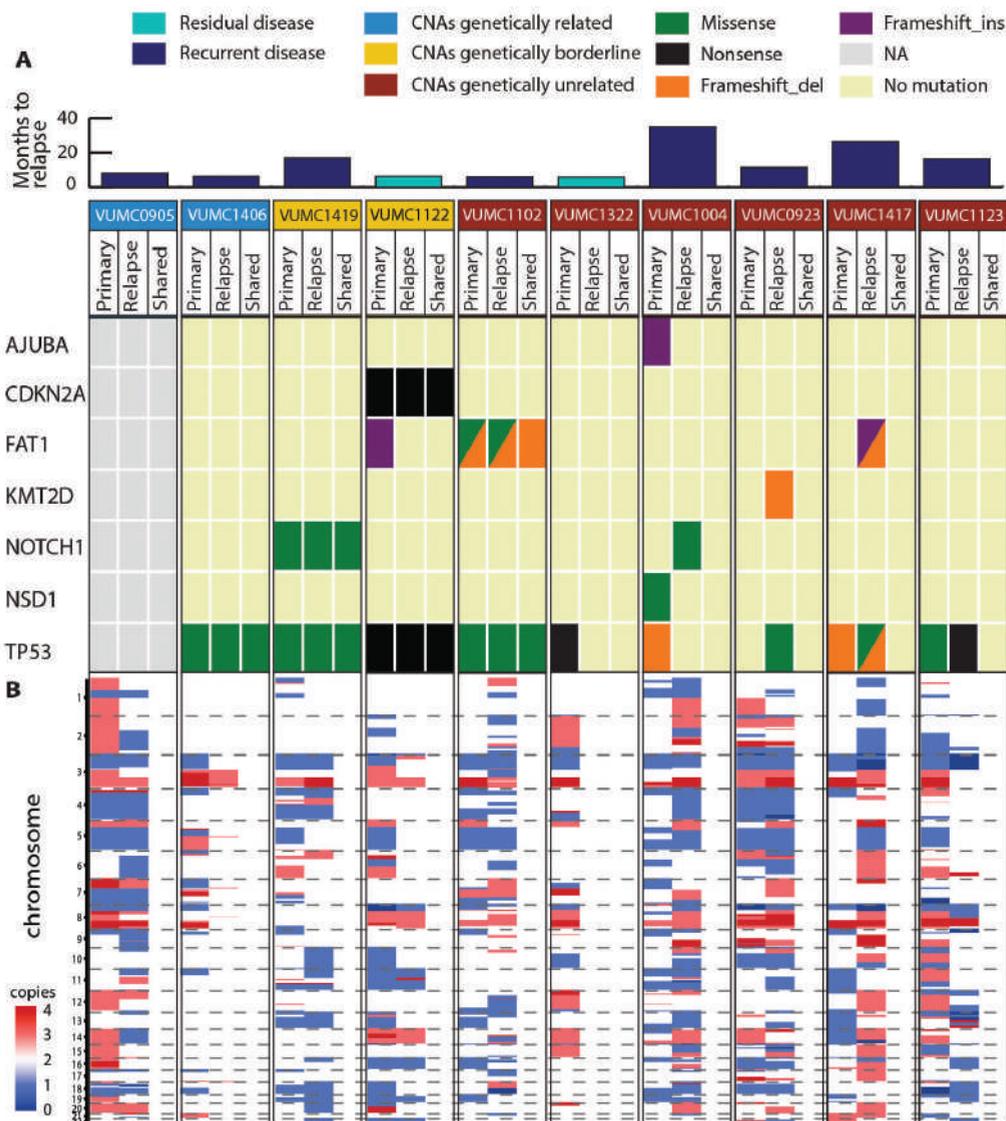
NOTE: Mutations found in two pairs of index tumor and relapse. Case VUMC1406 shares all its dominant variants (highlighted in bold), an example of genetic relatedness. All variants of case VUMC1417 are private to either the primary tumor or the relapse, suggesting that these are genetically unrelated.

<sup>a</sup>VAF, variant allele frequency. Computed by MuTect2.

<sup>b</sup>Number of copies of the region of interest.

<sup>c</sup>Number of mutant alleles calculated with the VAF and the copy numbers of the region.

Primary						Relapse					
Cellularity	Coverage	VAF <sup>a</sup>	CN <sup>b</sup>	aVA <sup>c</sup>	aVA/CN(%)	Cellularity	Coverage	VAF	CN	aVA	aVA/CN (%)
0.84	587	2.3	1.99	0.05	2.74	0.2	747	NA	—	—	—
0.84	546	2.3	1.99	0.05	2.74	0.2	806	NA	—	—	—
0.84	511	4.6	2.06	0.11	5.45	0.2	671	NA	—	—	—
0.84	324	NA	—	—	—	0.2	737	2.3	1.69	0.22	13.2
0.84	293	7.8	2.88	0.25	8.83	0.2	807	NA	—	—	—
0.84	339	NA	—	—	—	0.2	733	5	1.74	0.49	27.93
0.84	267	NA	—	—	—	0.2	861	7.4	4.25	0.91	21.32
0.84	<b>274</b>	<b>75</b>	<b>0.98</b>	<b>1.02</b>	<b>104.13</b>	0.2	<b>675</b>	<b>10.4</b>	<b>0.93</b>	<b>0.93</b>	<b>99.43</b>
0.5	96	NA	—	—	—	0.18	<b>3382</b>	<b>32.1</b>	<b>3.73</b>	<b>4.12</b>	<b>110.44</b>
0.5	68	NA	—	—	—	0.18	<b>3290</b>	<b>22.8</b>	<b>3.73</b>	<b>2.93</b>	<b>78.45</b>
0.5	233	3.6	2.14	0.15	6.96	0.18	2812	NA	—	—	—
0.5	<b>145</b>	<b>43.1</b>	<b>2.14</b>	<b>1.78</b>	<b>83.38</b>	0.18	2779	NA	—	—	—
0.5	181	NA	—	—	—	0.18	<b>3929</b>	<b>27.9</b>	<b>4.75</b>	3.87	81.4
0.5	151	NA	—	—	—	0.18	<b>3619</b>	<b>25.4</b>	<b>4.75</b>	<b>3.52</b>	<b>74.1</b>



**Figure 2.** Summary of CNA and target enrichment sequencing analysis.

A, Mutations found in index tumor and metachronous local relapse. Pairs are ordered from left to right on basis of genetic dependence (blue, genetically related by CNA analysis; yellow, borderline genetically unrelated; dark red, genetically unrelated). For every patient both the cancer gene variants in the index tumor and the relapse are shown. When the variant is shared, it is depicted in the “shared” column. Variants are shown when present in at least 33.3% of the tumor alleles. B, Detailed CNA profiles of index tumors and relapses. Red, copy number gains; blue, copy number losses.

## DISCUSSION

Assessment of the genetic landscape of tumors, relapses, and metastases, is technically challenging when frozen specimens are not available. Often only archived FFPE tissue material is available, which gives poor DNA yields and quality. This negatively impacts genetic analyses as the prepared sequencing libraries are generated on only few DNA strands, hampering coverage and introducing more PCR errors. For CNA analysis this is not too problematic; only the noise will increase, which can be reduced by increasing the bin size. For

mutations, it is more problematic as a basic coverage of at least 30 deep is required to identify a mutation in 10% of the tumor cells in 3 reads. Moreover, formalin fixation may cause nucleotide deamination, causing high sequencing error rates. These limitations can in part be overcome if mutation analysis is restricted to established cancer driver genes that usually show high VAFs and are biologically relevant. However, clonal evolution during tumor progression may impact also bona fide cancer drivers.

Although many HPV-negative HNSCC tumors share specific genomic alterations with frequent patterns of losses in chromosomal regions 3p and 8p, and gains in 3q, 5p, and 8q<sup>33</sup>, most tumors exhibit a unique combination of CNAs, particularly when breakpoints are considered. The combination of the log-likelihood ratio<sup>29</sup> and the correlation of segments<sup>30</sup> on the 44 biopsies of 13 oral cavity samples could make an accurate classification between different biopsies from one tumor and independent tumors using copy number profiles. The high accuracy (96%) of our classification model reflects the much higher degree of inter- than intratumoral heterogeneity, when considering copy number alterations. This result is in line with a recent study by Tabatabaieifar and colleagues<sup>38</sup>. Unfortunately, analysis of genetic relationships by CNA profiling seemed less conclusive for tumors and local relapses of chemoradiotherapy-treated patients. This might relate to the assay itself, but more likely it reflects the complex pathobiological origin of the relapses. Therefore, we extended the biomarker panel with somatic mutation analysis of the known and established HNSCC driver genes. When we consider mutations as gold standard, accuracy of CNA profiling is indeed less accurate to assess genetic relationships as it was in 6 of 9 cases correct (67%). As indicated above, this likely relates to the nature of the relapse after chemoradiotherapy treatment, and comparable studies to lymph node metastases may shed more light on this issue. Moreover, the analysis of 3-5 biopsies might not capture the genetic heterogeneity of a tumor completely, and we may have underestimated intratumor heterogeneity. Nonetheless, CNA profiles are extremely helpful to correct tumor percentages and ploidy using algorithms such as ACE, and provides quantitative estimates of the mutant allele numbers<sup>39</sup>.

We showed by mutation analysis that 5 of 10 relapses of chemoradiotherapy-treated patients analyzed seemed genetically related to the tumor based on one or more cancer driver mutations with a high frequency of mutant alleles. This is not unexpected as tumor cell populations may vary between tumor and relapse, which may relate to the extra cell divisions in relapses but which is also in line with the view that some populations are more resistant to particular treatments than others<sup>40</sup>.

Most remarkable are the five cases of which the relapses seemed not to be genetically related to the index tumor. Whether this truly refers an independent origin remains a challenging question. The availability of only FFPE specimen hampered DNA sequencing, and in some pairs the coverage was somewhat low. Moreover, chemoradiotherapy treatment might have selected resistant clones that were present in the index tumor but with frequencies of less than 1% or even 0.1%, and that are not picked up with the current sequencing coverage of a single randomly taken biopsy. Moreover, treatment may induce genetic changes, although we would expect additional mutations and copy number alterations, and not a disappearance. Hence, this cannot explain why cancer driver mutations in the primary tumor are not present in the relapse, and it does not really support that treatment induced changes underlie the observed phenomenon.

LRs remain a major obstacle in the success rate of HNSCC treatment. Most of our knowledge on relapses is based on studies performed on surgically treated OSCC<sup>13,14,16,17</sup>. Despite the high frequency of local failure, studies to investigate the local relapses after treatment with cisplatin-based chemoradiotherapy are scarce, due to the lack of available biological material in the absence of surgery. Incidence of locoregional failure after concurrent chemotherapy varies between 15% and 50%<sup>7,11,12</sup>. In our cohort of 113 patients, only 20 (17.7%) patients developed a local and/or regional relapse, of whom 12 (10.6%) had a local relapse. These numbers are low but still comparable to other studies on cisplatin-based concomitant chemoradiotherapy in the IMRT era<sup>12</sup>. Nonetheless, the relatively effective treatment in this cohort might relate to the remarkably high number of genetically unrelated relapses.

We evaluated whether the applied treatment protocols played a role. Development of local recurrences from residual cells can be explained by insufficient treatment or an inadequate response to treatment. Patients receiving less than a cumulative dose of 200 mg/m<sup>2</sup> cisplatin have a significant higher risk on developing a locoregional relapse<sup>41</sup>. Also, a lower total irradiation dose or greater tumor volume can potentially lead to the outgrowth of residual cells<sup>42</sup>. All our patients received the planned dose of 70 Gy on their primary tumor. However, two patients (VUMC1123 and VUMC1322) did not reach a cumulative dose of >200 mg/m<sup>2</sup> cisplatin due to unacceptable nephrotoxicity, although patient VUMC1322 switched to a weekly carboplatin regimen. Interestingly, this somewhat less optimal primary tumor management did apparently not lead to outgrowth of the index tumor as the origin of the relapses of both VUMC1123 and VUMC1322 seemed genetically unrelated.

Several molecular mechanisms of action and resistance to chemoradiotherapy are proposed to explain inadequate response to treatment, and preclinical studies show many different mechanisms for resistance<sup>43</sup>. In HNSCC the fraction of DNA-bound platinum and genes in the DNA repair pathway (especially Fanconi anemia/BRCA pathway) are associated with the response to treatment<sup>44,45</sup>. In clinical studies specific mutations (e.g. TP53, NOTCH1)<sup>42,46</sup> or expression of specific genes (e.g., hypoxia-related genes or SDF-1 and CXCR4<sup>42,47</sup>) have been associated with local treatment failure. These studies generally assume that the development of local relapses relates primarily to the outgrowth of residual tumor cells. The idea that so many relapses seem to be genetically unrelated to the bulk of the index tumor may impact the interpretation of these studies and could explain why we failed to identify clinically meaningful predictive biomarkers. A previous study by Hedberg and colleagues also described a case where there was no genetic relationship between the index tumor and the recurrence, and the authors suggested that this recurrence might have originated from a second independent field or developed as a true second primary tumor<sup>17</sup>. Irrespective of the precise origin of these genetically unrelated relapses, this observation warrants further studies. The impact is tremendous for future studies, as it will imply that gene profiling or radiomics studies on the bulk index tumor will not always be able to indicate the accurate risk for relapse. Associations may improve when the genetically unrelated relapses are filtered out. Also, the evaluation of novel experimental agents in clinical trials might benefit from the precise analysis of the relapse type that occurred. In addition, relapses may be treated as new tumors. Given these potential impacts, similar studies should be performed but in larger cohorts, and preferably making use of multiple frozen biopsies to allow ultradeep sequencing and better estimates of intratumoral heterogeneity.

Our study provides insight in the complex biology of relapsed HNSCC after chemoradiotherapy, and might have large consequences for prognostic modeling, the use of predictive and prognostic biomarkers and therapeutic innovations.

## **ACKNOWLEDGMENTS**

We would like to thank D. Sie for providing his “clonality” tool to assess the correlation and log-likelihood analysis of CNAs and Prof Enrico Silini for providing patient material.

## **CONFLICTS OF INTEREST**

C.R. Leemans is a consultant/advisory board member for MSD Global Advisory Board HNSCC. No potential conflicts of interest were disclosed by the other authors.

## **FUNDING**

This work was supported by the Design Project (KWF-A6C7072).

**Supplementary Table 1.** Patient characteristics of local recurrences after treatment with CRT

Patient characteristics		FFPE for analysis available (n=10)	No FFPE for analysis available (n=10)	p value <sup>a</sup>
Gender				
	Male (%)	8(80)	8(80)	
	Female (%)	2(20)	2(20)	1.00
Age, mean (SD)		62.7(8.8)	58.5(8.3)	0.29
Comorbidity score (ACE-27)				
	None (%)	3(30)	4(40)	
	Mild (%)	4(40)	5(50)	
	Moderate (%)	3(30)	1(1)	
	Severe (%)	0(0)	0(0)	0.71
Site				
	Oropharynx (%)	6(60)	6(60)	
	Hypopharynx (%)	3(30)	3(30)	
	Larynx (%)	1(10)	1(1)	1.00
T stage				
	T2 (%)	0(0)	3(30)	
	T3 (%)	5(50)	2(20)	
	T4a (%)	3(30)	2(20)	
	T4b (%)	2(20)	3(30)	0.25
N stage				
	N0 (%)	2(20)	1(10)	
	N1 (%)	2(20)	1(10)	
	N2a (%)	0(0)	1(10)	
	N2b (%)	3(30)	3(30)	
	N2c (%)	3(30)	3(30)	
	N3 (%)	0	1(10)	1.00
Stage				
	III (%)	1(10)	1(10)	
	IVA (%)	7(70)	6(60)	
	IVB (%)	2(20)	3(30)	1.00
Chemotherapy completed				
	Yes (%)	8(80)	8(80)	
	No (%)	2(20)	2(20)	1.00
Mean cumulative dose cisplatin (mg/m <sup>2</sup> )		256(82.6)	278(45.7)	0.47
Radiotherapy completed				
	Yes (%)	10(100)	10(100)	
	No (%)	0(0)	0(0)	1.00
Mean cumulative dose radiotherapy (Gy)		70(0)	70(0)	
Median survival time (95% C.I.)		1.54(0.0 - 3.59)	1.99(0.0 - 4.79)	0.15

<sup>a</sup> P values were calculated with the use of t-test for continuous variables, Fisher exact test for categorical variables and log-rank-test for survival data test for categorical variables

**Supplementary Table 2.** Genepanel of the target enrichment sequencing

Gene	Mutated in HNSCC (%) <sup>a</sup>
TP53	72%
FAT1	23%
CDKN2A	22%
PIK3CA	21%
NOTCH1	19%
KMT2D	18%
NSD1	10%
CASP8	9%
AJUBA	6%
FBXW7	5%
HRAS	4%
PTEN	2%

<sup>a</sup> According to TCGA<sup>33</sup>**Supplementary Table 3.** Follow-up data on the chemoradiation cohort (n=113)

PatientID	Year of diagnosis	Tumor site <sup>a</sup>	Stage	Local/regional relapse	First relapse <sup>b</sup>	Time to relapse (m)	Status	Cause of death	Overall Survival time (m)
VUMC0905	2009	LA	IVA	Yes	Local	8.2	Death	Death with disease	18.5
VUMC0909	2009	OP	III	Yes	Regional	15.5	Death	Other	67.0
VUMC0918	2009	OP	IVB	Yes	Local	9.5	Death	Death with disease	13.6
VUMC0921	2009	OP	IVB	Yes	Regional	11	Death	Death with disease	13.5
VUMC0923	2009	OP	IVB	Yes	Local	11.5	Death	Death with disease	26.7
VUMC1004	2010	HP	IVA	Yes	Local	35.0	Death	Death with disease	45.0
VUMC1009	2010	LA	IVA	Yes	Regional	9.7	Alive	NA	67.9
VUMC1102	2011	HP	IVB	Yes	Local	6.0	Death	Death with disease	8.5
VUMC1109	2011	HP	IVA	Yes	Regional	6.6	Death	Other	48.1
VUMC1122	2011	OP	IVA	Yes	Local	6.2	Death	Death with disease	8.8
VUMC1123	2011	OP	IVA	Yes	Local	16.6	Death	Death with disease	35.7
VUMC1226	2012	OP	IVA	Yes	Regional	9.6	Death	Death with disease	16.6
VUMC1315	2013	OP	IVA	Yes	Local	9.5	Death	Death with disease	19.2
VUMC1322	2013	OP	IVA	Yes	Local	5.8	Death	Death with disease	8.3
VUMC1323	2013	OP	IVA	Yes	Regional	5.3	Alive	NA	34.9

**Supplementary Table 3.** (continued)

PatientID	Year of diagnosis	Tumor site <sup>a</sup>	Stage	Local/regional relapse	First relapse <sup>b</sup>	Time to relapse (m)	Status	Cause of death	Overall Survival time (m)
VUMC1402	2014	HP	IVB	Yes	Regional + Distant metastasis	17.8	Death	Death with disease	23.9
VUMC1406	2014	HP	IVA	Yes	Locoregional	6.4	Death	Death with disease	8.3
VUMC1417	2014	OP	III	Yes	Local	26.5	Alive	NA	30.3
VUMC1419	2014	OP	IVA	Yes	Local	17.3	Alive	NA	24.8
VUMC1421	2014	HP	IVA	Yes	Regional	5.5	Alive	NA	32.1
VUMC0901	2009	HP	III	No	NA	NA	Alive	NA	73.7
VUMC0903	2009	LA	IVB	No	Distant metastasis	6.7	Death	Death with disease	10.3
VUMC0904	2009	HP	IVA	No	NA	NA	Death	Lost to follow up	23.5
VUMC0906	2009	HP	III	No	NA	NA	Death	Other	42.2
VUMC0907	2009	HP	IVA	No	NA	NA	Alive	NA	80.9
VUMC0910	2009	LA	III	No	NA	NA	Death	Other	22.4
VUMC0911	2009	HP	IVA	No	NA	NA	Death	Other	14.9
VUMC0913	2009	HP	IVA	No	NA	NA	Death	Lost to follow up	34.9
VUMC0914	2009	HP	III	No	NA	NA	Alive	NA	81.1
VUMC0915	2009	LA	III	No	NA	NA	Alive	NA	84.7
VUMC0917	2009	OP	III	No	NA	NA	Alive	NA	76.8
VUMC0919	2009	OP	IVB	No	NA	NA	Death	Other	51.2
VUMC0920	2009	OP	IVA	No	NA	NA	Death	Death with disease	1.1
VUMC0925	2009	OP	IVB	No	NA	NA	Death	Other	57.7
VUMC0926	2009	HP	IVA	No	SPT	67.4	Alive	NA	85.3
VUMC1001	2010	HP	IVA	No	Distant metastasis	12.7	Death	Death with disease	35.9
VUMC1002	2010	HP	IVA	No	NA	NA	Death	Lost to follow up	45.2
VUMC1003	2010	HP	IVA	No	NA	NA	Alive	NA	77.7
VUMC1005	2010	HP	III	No	SPT	19.8	Alive	NA	76.2
VUMC1006	2010	HP	III	No	NA	NA	Alive	NA	45.1
VUMC1007	2010	OP	IVA	No	NA	NA	Alive	NA	65.0
VUMC1008	2010	OP	IVA	No	NA	NA	Death	Other	7.5
VUMC1010	2010	HP	IVA	No	Distant metastasis	5.5	Death	Death with disease	16.1
VUMC1101	2011	HP	IVA	No	NA	NA	Alive	NA	63.7
VUMC1103	2011	HP	III	No	NA	NA	Alive	NA	68.6
VUMC1104	2011	LA	III	No	NA	NA	Alive	NA	60.6

**Supplementary Table 3.** (continued)

PatientID	Year of diagnosis	Tumor site <sup>a</sup>	Stage	Local/regional relapse	First relapse <sup>b</sup>	Time to relapse (m)	Status	Cause of death	Overall Survival time (m)
VUMC1105	2011	OP	IVA	No	NA	NA	Alive	NA	63.4
VUMC1106	2011	LA	III	No	NA	NA	Alive	NA	35.6
VUMC1107	2011	HP	IVA	No	NA	NA	Alive	NA	58.0
VUMC1108	2011	HP	III	No	NA	NA	Death	Other	58.5
VUMC1111	2011	HP	IVA	No	NA	NA	Alive	NA	59.4
VUMC1112	2011	HP	IVA	No	SPT	19.6	Alive	NA	59.1
VUMC1114	2011	OP	IVA	No	NA	NA	Alive	NA	64.6
VUMC1115	2011	OP	IVA	No	NA	NA	Alive	NA	63.0
VUMC1116	2011	HP	IVA	No	SPT	11.6	Death	Other	33.6
VUMC1117	2011	OP	IVA	No	Distant metastasis	6.0	Death	Death with disease	7.1
VUMC1119	2011	HP	III	No	NA	NA	Alive	NA	68.8
VUMC1121	2011	OP	IVB	No	NA	NA	Death	Other	16.5
VUMC1124	2011	OP	IVA	No	NA	NA	Alive	NA	58.9
VUMC1125	2011	LA	IVA	No	NA	NA	Alive	NA	58.2
VUMC1201	2012	OP	IVA	No	NA	NA	Alive	NA	58.0
VUMC1202	2012	LA	III	No	NA	NA	Alive	NA	57.3
VUMC1203	2012	OP	IVB	No	SPT	53.9	Alive	NA	57.6
VUMC1204	2012	OP	IVA	No	Distant metastasis	12.1	Death	Death with disease	29.8
VUMC1206	2012	LA	IVA	No	NA	NA	Alive	NA	50.7
VUMC1208	2012	LA	III	No	NA	NA	Alive	NA	50.8
VUMC1209	2012	LA	III	No	NA	NA	Death	NA	46.9
VUMC1211	2012	OP	III	No	NA	NA	Alive	NA	15.4
VUMC1212	2012	OP	IVB	No	NA	NA	Alive	NA	53.0
VUMC1213	2012	HP	IVA	No	NA	NA	Alive	NA	52.0
VUMC1214	2012	HP	IVA	No	NA	NA	Alive	NA	49.0
VUMC1215	2012	OP	IVA	No	NA	NA	Alive	NA	52.0
VUMC1216	2012	HP	IVA	No	NA	NA	Alive	NA	49.9
VUMC1218	2012	OP	IVA	No	NA	NA	Alive	NA	46.2
VUMC1219	2012	OP	IVA	No	NA	NA	Alive	NA	59.1
VUMC1223	2012	OP	IVA	No	NA	NA	Alive	NA	51.8
VUMC1225	2012	OP	IVB	No	NA	NA	Alive	NA	44.6
VUMC1301	2013	OP	IVB	No	NA	NA	Alive	NA	44.2
VUMC1302	2013	OP	IVA	No	Distant metastasis	5.9	Death	Death with disease	16.6
VUMC1303	2013	LA	III	No	NA	NA	Alive	NA	43.9
VUMC1304	2013	OP	IVA	No	NA	NA	Alive	NA	37.8

**Supplementary Table 3.** (continued)

PatientID	Year of diagnosis	Tumor site <sup>a</sup>	Stage	Local/regional relapse	First relapse <sup>b</sup>	Time to relapse (m)	Status	Cause of death	Overall Survival time (m)
VUMC1305	2013	LA	IVA	No	NA	NA	Alive	NA	33.3
VUMC1306	2013	OP	IVA	No	NA	NA	Alive	NA	34.5
VUMC1307	2013	OP	IVA	No	NA	NA	Death	Other	28.9
VUMC1308	2013	OP	IVB	No	NA	NA	Death	Other	43.6
VUMC1309	2013	OP	IVA	No	NA	NA	Alive	NA	42.6
VUMC1311	2013	OP	IVB	No	Distant metastasis	4.9	Death	Death with disease	5.0
VUMC1312	2013	HP	IVA	No	NA	NA	Alive	NA	44.7
VUMC1313	2013	OP	IVA	No	NA	NA	Alive	NA	41.4
VUMC1314	2013	OP	IVA	No	NA	NA	Alive	NA	39.7
VUMC1316	2013	LA	III	No	NA	NA	Death	Other	39.4
VUMC1317	2013	HP	IVB	No	NA	NA	Death	Other	22.2
VUMC1318	2013	LA	IVA	No	Distant metastasis	5.4	Death	Death with disease	26.5
VUMC1319	2013	HP	IVB	No	NA	NA	Alive	NA	39.5
VUMC1325	2013	LA	III	No	NA	NA	Alive	NA	44.0
VUMC1326	2013	HP	IVA	No	Distant metastasis	5.8	Death	Death with disease	15.0
VUMC1327	2013	LA	IVA	No	NA	NA	Alive	NA	41.8
VUMC1328	2013	HP	IVA	No	Distant metastasis	18.8	Alive	NA	34.9
VUMC1329	2013	HP	IVA	No	NA	NA	Alive	NA	34.2
VUMC1403	2014	LA	III	No	NA	NA	Alive	NA	30.9
VUMC1404	2014	OP	IVA	No	NA	NA	Alive	NA	30.9
VUMC1405	2014	OP	III	No	NA	NA	Alive	NA	30.9
VUMC1407	2014	HP	III	No	NA	NA	Alive	NA	25.3
VUMC1408	2014	OP	IVA	No	NA	NA	Alive	NA	26.3
VUMC1409	2014	LA	III	No	NA	NA	Alive	NA	25.8
VUMC1410	2014	OP	IVA	No	NA	NA	Death	Other	14.8
VUMC1411	2014	OP	IVB	No	Distant metastasis	8.2	Death	Death with disease	8.4
VUMC1412	2014	OP	IVA	No	NA	NA	Alive	NA	34.7
VUMC1415	2014	OP	IVA	No	NA	NA	Alive	NA	30.6
VUMC1416	2014	OP	IVA	No	NA	NA	Alive	NA	30.7
VUMC1418	2014	OP	III	No	NA	NA	Alive	NA	24.9
VUMC1420	2014	OP	III	No	NA	NA	Death	Other	10.9
VUMC1422	2014	LA	IVA	No	NA	NA	Alive	NA	32.9

<sup>a</sup> LA means larynx, OP means oropharynx and HP means hypopharynx<sup>b</sup> SPT means Second Primary Tumor

**Supplementary Table 4.** Detailed results of target enrichment sequencing

Patient	Gene	Mutation			
		Type	Location	Protein change	Sample
VUMC1406	AJUBA	Missense	c.733G>A	p.G245R	Primary
VUMC1406	AJUBA	Missense	c.865G>A	p.G289S	Primary
VUMC1406	KMT2D	Missense	c.2033C>T	p.S678F	Primary
VUMC1406	KMT2D	Missense	c.8896C>T	p.R2966W	Relapse
VUMC1406	NSD1	Missense	c.1267G>A	p.A423T	Primary
VUMC1406	NSD1	Missense	c.3034C>G	p.R1012G	Relapse
VUMC1406	PIK3CA	Missense	c.3140A>G	p.H1047R	Relapse
VUMC1406	TP53	Missense	c.716A>G	p.N239S	Both
VUMC1122	CDKN2A	Nonsense	c.329G>A	p.W110*	Both
VUMC1122	FAT1	Frameshift_ins	c.819_820insT	p.A274fs	Primary
VUMC1122	FBXW7	Missense	c.925C>T	p.R309C	Primary
VUMC1122	KMT2D	Missense	c.15091C>T	p.R5031C	Primary
VUMC1122	KMT2D	Missense	c.9730G>A	p.E3244K	Primary
VUMC1122	NOTCH1	Nonsense	c.1408C>T	p.Q470*	Primary
VUMC1122	PTEN	Missense	c.235G>A	p.A79T	Primary
VUMC1122	TP53	Nonsense	c.916C>T	p.R306*	Both
VUMC1419	NOTCH1	Missense	c.995G>A	p.C332Y	Both
VUMC1419	NOTCH1	Missense	c.1154C>A	p.S385Y	Relapse
VUMC1419	NSD1	Missense	c.2903A>G	p.K968R	Relapse
VUMC1419	NSD1	Missense	c.5741G>A	p.R1914H	Relapse
VUMC1419	TP53	Missense	c.707A>G	p.Y236C	Both
VUMC1419	TP53	Missense	c.482C>T	p.A161V	Relapse
VUMC1419	TP53	Missense	c.481G>T	p.A161S	Relapse
VUMC1419	TP53	Missense	c.745A>G	p.R249G	Relapse
VUMC1102	FAT1	Frameshift_del	c.9331delA	p.I3111fs	Both
VUMC1102	FAT1	Missense	c.4841C>T	p.P1614L	Primary
VUMC1102	FAT1	Missense	c.115T>A	p.Y39N	Relapse
VUMC1102	FBXW7	Nonsense	c.1099C>T	p.R367*	Primary
VUMC1102	NSD1	Missense	c.6596G>A	p.R2199H	Primary
VUMC1102	TP53	Missense	c.638G>T	p.R213L	Both
VUMC1322	KMT2D	In_Frame_Del	c.2088_2114delCCCCACATCCCCACC ACCTGAGGACTC	p.696_705SPTSPPPEDS>S	Primary
VUMC1322	KMT2D	In_Frame_Del	c.11235_11237delGCA	p.Q3745del	Relapse
VUMC1322	NOTCH1	Missense	c.6284G>A	p.R2095H	Relapse
VUMC1322	TP53	Nonsense	c.949C>T	p.Q317*	Primary
VUMC1004	AJUBA	Frameshift_ins	c.34_35insC	p.L12fs	Primary
VUMC1004	KMT2D	Nonsense	c.16342C>T	p.R5448*	Primary
VUMC1004	NOTCH1	Missense	c.7210C>A	p.Q2404K	Relapse

Cellularity	Coverage	Primary							Relapse										
		VAF <sup>a</sup>	VAF_low <sup>b</sup>	VAF_high <sup>c</sup>	Copynumbers	aVA	aVA_low <sup>b</sup>	aVA_high <sup>c</sup>	aVA/CN (%)	Cellularity	Coverage	VAF <sup>a</sup>	VAF_low <sup>b</sup>	VAF_high <sup>c</sup>	Copynumbers	aVA	aVA_low <sup>b</sup>	aVA_high <sup>c</sup>	aVA/CN (%)
0.84	587	2.3	0.85	4.94	1.99	0.05	0.02	0.12	2.74	0.20	747	NA							
0.84	546	2.3	1.00	4.48	1.99	0.05	0.02	0.11	2.74	0.20	806	NA							
0.84	511	4.6	1.27	11.36	2.06	0.11	0.03	0.28	5.45	0.20	671	NA							
0.84	324	NA								0.20	737	2.3	0.85	4.94	1.69	0.22	0.08	0.48	13.20
0.84	293	7.8	3.41	14.73	2.88	0.25	0.11	0.48	8.83	0.20	807	NA							
0.84	339	NA								0.20	733	5	2.88	7.99	1.74	0.49	0.28	0.78	27.93
0.84	267	NA								0.20	861	7.4	5.01	10.45	4.25	0.91	0.61	1.28	21.32
0.84	274	75	65.86	83.14	0.98	1.02	0.90	1.13	104.13	0.20	675	10.4	7.14	14.54	0.93	0.93	0.64	1.30	99.43
0.37	568	63.4	57.14	69.44	3.34	4.27	3.85	4.68	128.10	0.41	146	23.1	13.15	35.50	2.89	1.33	0.76	2.05	46.08
0.37	542	41.1	34.66	47.62	2.53	2.44	2.06	2.83	96.48	0.41	181	NA							
0.37	777	2.4	0.97	4.88	2.53	0.14	0.06	0.29	5.63	0.41	172	NA							
0.37	973	33.3								0.41	329	NA							
0.37	644	2.3	0.93	4.69	3.67	0.16	0.07	0.33	4.44	0.41	267	NA							
0.37	674	15	11.25	19.46	3.34	1.01	0.76	1.31	30.31	0.41	238	NA							
0.37	781	1.5	0.55	3.24	2.34	0.09	0.03	0.19	3.68	0.41	267	NA							
0.37	838	42.1	37.57	46.76	2.81	2.62	2.34	2.91	93.08	0.41	302	11	6.00	18.10	3.17	0.67	0.36	1.09	20.99
0.63	1761	73.6	70.50	76.55	4.07	3.86	3.70	4.02	94.84	0.22	1596	25.9	22.79	29.24	2.91	2.59	2.28	2.93	88.91
0.63	1459	NA								0.22	1442	2.2	1.29	3.50	2.91	0.22	0.13	0.35	7.55
0.63	1048	NA								0.22	1189	1.4	0.64	2.64	2.71	0.14	0.06	0.26	5.06
0.63	1077	NA								0.22	1168	1.6	0.77	2.92	2.71	0.16	0.08	0.29	5.78
0.63	1155	61.5	57.47	65.41	2.13	2.03	1.90	2.16	95.37	0.22	1243	13	10.56	15.77	3.38	1.36	1.11	1.65	40.27
0.63	877	NA								0.22	1203	7	5.07	9.37	3.38	0.73	0.53	0.98	21.68
0.63	883	NA								0.22	1218	7	5.07	9.37	3.38	0.73	0.53	0.98	21.68
0.63	1059	NA								0.22	1221	8.3	6.25	10.78	3.38	0.87	0.65	1.13	25.71
0.37	1277	29.4	26.08	32.93	2.13	1.63	1.44	1.82	76.35	0.41	144	47.1	25.71	70.22	2.98	2.76	1.51	4.11	92.61
0.37	1399	31	27.85	34.27	2.13	1.72	1.54	1.90	80.50	0.41	117	NA							
0.37	1381	NA								0.41	98	23.1	7.82	45.37	2.98	1.35	0.46	2.66	45.42
0.37	1375	6.2	4.64	8.10	2.13	0.34	0.26	0.45	16.10	0.41	121	NA							
0.37	1585	1	0.52	1.74	2.17	0.06	0.03	0.10	2.57	0.41	112	NA							
0.37	1643	8.5	6.77	10.51	2.77	0.52	0.42	0.65	18.96	0.41	140	29.2	15.42	45.90	1.95	1.41	0.75	2.22	72.19
0.38	152	8.6	1.80	23.06	2.80	0.52	0.11	1.40	18.63	0.27	409	NA							
0.38	96	NA								0.27	450	2.7	0.88	6.19	2.21	0.21	0.07	0.47	9.31
0.38	119	NA								0.27	431	3.8	1.41	8.08	2.02	0.28	0.10	0.60	13.99
0.38	83	26.7	9.15	51.20	2.28	1.48	0.51	2.84	64.86	0.27	360	NA							
0.53	679	25	20.14	30.38	1.08	0.71	0.57	0.87	66.09	0.34	126	NA							
0.53	760	6.8	4.53	9.74	2.17	0.27	0.18	0.38	12.35	0.34	49	NA							
0.53	752	NA								0.34	107	17.6	6.06	36.89	4.31	1.44	0.50	3.02	33.45

**Supplementary Table 4.** (continued)

Patient	Gene	Mutation			
		Type	Location	Protein change	Sample
VUMC1004	NOTCH1	In_Frame_Del	c.7162_7197delCAAACCTTACAGATG CAGCAGCAGAACCTGCAGCCA	p.QNLQMQQQNLQP2388del	Relapse
VUMC1004	NSD1	Missense	c.3106G>C	p.A1036P	Primary
VUMC1004	TP53	Frameshift_del	c.352delA	p.T118fs	Primary
VUMC0923	KMT2D	Frameshift_ins	c.6476_6477insT	p.L2159fs	Relapse
VUMC0923	KMT2D	Frameshift_del	c.1143delC	p.P381fs	Relapse
VUMC0923	NOTCH1	Frameshift_del	c.823_863delGGCGTGAACACCTACA ACTGCCGCTGCCGCCAGAGTGGAC	p.GVNTYNCRCPEWT275fs	Relapse
VUMC0923	NOTCH1	Missense	c.812C>T	p.A271V	Relapse
VUMC0923	TP53	Missense	c.734G>T	p.G245V	Relapse
VUMC1417	FAT1	Frameshift_del	c.8013_8016delCTTT	p.FF2671fs	Relapse
VUMC1417	FAT1	Frameshift_ins	c.3445_3446insA	p.M1149fs	Relapse
VUMC1417	TP53	Missense	c.1013A>G	p.H338R	Primary
VUMC1417	TP53	Frameshift_del	c.1022_1023delTC	p.F341fs	Primary
VUMC1417	TP53	Missense	c.761T>A	p.I254N	Relapse
VUMC1417	TP53	Frameshift_del	c.1019_1028delTGTTCCGAGA	p.MFRE340fs	Relapse
VUMC1123	FAT1	Missense	c.1507G>A	p.A503T	Relapse
VUMC1123	FBXW7	Missense	c.1691G>A	p.R564H	Primary
VUMC1123	KMT2D	Missense	c.16214G>A	p.R5405H	Primary
VUMC1123	KMT2D	Missense	c.15481G>A	p.E5161K	Primary
VUMC1123	NSD1	Missense	c.5789G>A	p.R1930H	Primary
VUMC1123	TP53	Missense	c.833C>T	p.P278L	Primary
VUMC1123	TP53	Nonsense	c.574C>T	p.Q192*	Relapse
VUMC1123	TP53	Missense	c.997C>T	p.R333C	Relapse

<sup>a</sup> Calculated using MuTect

<sup>b</sup> Lower limit 95% CI

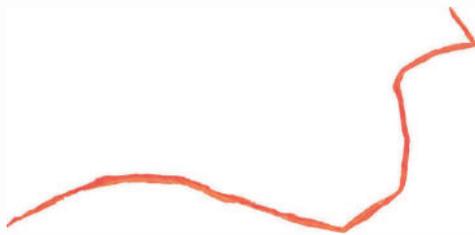
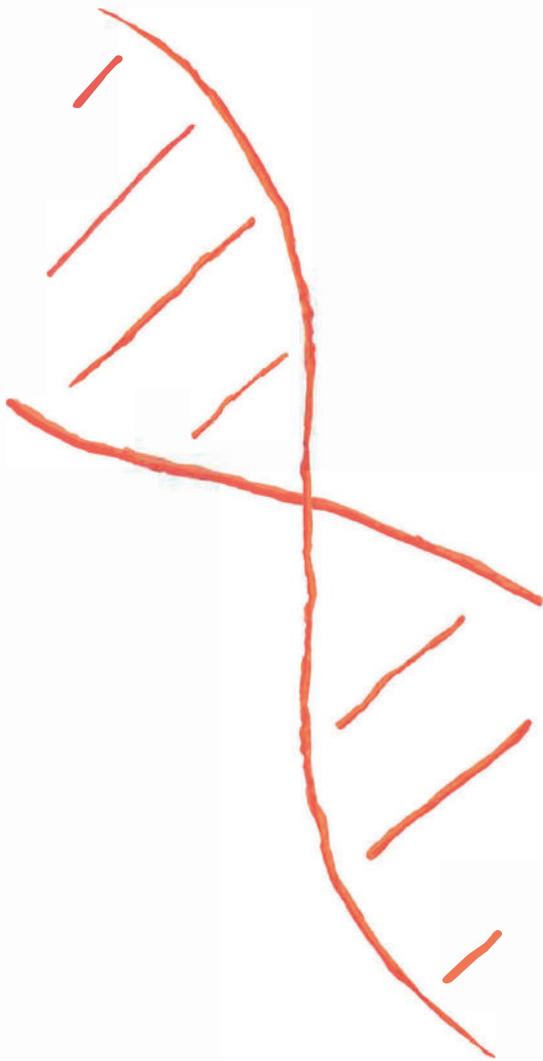
<sup>c</sup> Upper limit 95% CI

Cellularity	Coverage	Primary								Cellularity	Coverage	Relapse							
		VAF <sup>a</sup>	VAF_low <sup>b</sup>	VAF_high <sup>c</sup>	Copynumbers	aVA	aVA_low <sup>b</sup>	aVA_high <sup>c</sup>	aVA/CN (%)			VAF <sup>a</sup>	VAF_low <sup>b</sup>	VAF_high <sup>c</sup>	Copynumbers	aVA	aVA_low <sup>b</sup>	aVA_high <sup>c</sup>	aVA/CN (%)
0.53	729	NA								0.34	101	13	4.41	28.09	4.31	1.07	0.36	2.30	24.71
0.53	726	38.7	32.59	44.97	1.16	1.13	0.95	1.32	98.05	0.34	47	NA							
0.53	754	26.3	21.13	32.04	1.91	0.97	0.78	1.18	50.71	0.34	43	NA							
0.37	109	NA								0.60	147	14.7	5.57	29.17	3.08	0.65	0.25	1.29	21.06
0.37	87	NA								0.60	143	52.4	35.82	69.02	3.08	2.31	1.58	3.05	75.08
0.37	146	NA								0.60	208	3.9	1.07	9.65	3.08	0.17	0.05	0.43	5.59
0.37	191	NA								0.60	198	5.9	1.63	14.38	3.08	0.26	0.07	0.64	8.45
0.37	112	NA								0.60	136	35.7	20.82	53.78	2.76	1.46	0.85	2.20	52.94
0.50	96	NA								0.18	3382	32.1	30.33	33.91	3.73	4.12	3.90	4.35	110.44
0.50	68	NA								0.18	3290	22.8	21.17	24.49	3.73	2.93	2.72	3.15	78.45
0.50	233	3.6	0.75	10.20	2.14	0.15	0.03	0.42	6.96	0.18	2812	NA							
0.50	145	43.1	30.85	55.96	2.14	1.78	1.28	2.32	83.38	0.18	2779	NA							
0.50	181	NA								0.18	3929	27.9	26.37	29.47	4.75	3.87	3.66	4.09	81.40
0.50	151	NA								0.18	3619	25.4	23.87	26.98	4.75	3.52	3.31	3.74	74.10
0.59	654	NA								0.43	1101	1.4	0.67	2.56	2.03	0.07	0.03	0.12	3.23
0.59	1113	2.3	1.19	3.98	2.90	0.10	0.05	0.17	3.40	0.43	1187	NA							
0.59	1356	2.1	1.23	3.34	3.90	0.11	0.06	0.18	2.85	0.43	1767	NA							
0.59	1276	1	0.48	1.83	3.90	0.05	0.03	0.10	1.36	0.43	1553	NA							
0.59	1180	2	1.07	3.40	2.78	0.08	0.04	0.14	3.00	0.43	1390	NA							
0.59	1195	64.7	59.20	66.75	2.89	2.70	2.53	2.86	93.28	0.43	1495	NA							
0.59	948	NA								0.43	1191	43.8	39.70	47.98	2.10	2.08	1.89	2.28	98.97
0.59	1006	NA								0.43	1197	1.4	0.73	2.43	2.10	0.07	0.03	0.12	3.16

## REFERENCES

1. Ferlay J, Soerjomataram I, Dikshit R, et al. Cancer incidence and mortality worldwide: sources, methods and major patterns in GLOBOCAN 2012. *Int J Cancer* 2015; 136(5): E359-86.
2. Carvalho AL, Nishimoto IN, Califano JA, Kowalski LP. Trends in incidence and prognosis for head and neck cancer in the United States: a site-specific analysis of the SEER database. *Int J Cancer* 2005; 114(5): 806-16.
3. Braakhuis BJ, Leemans CR, Visser O. Incidence and survival trends of head and neck squamous cell carcinoma in the Netherlands between 1989 and 2011. *Oral Oncol* 2014; 50(7): 670-5.
4. Kutler DI, Auerbach AD, Satagopan J, et al. High incidence of head and neck squamous cell carcinoma in patients with Fanconi anemia. *Arch Otolaryngol Head Neck Surg* 2003; 129(1): 106-12.
5. Gillison ML, Koch WM, Capone RB, et al. Evidence for a causal association between human papillomavirus and a subset of head and neck cancers. *J Natl Cancer Inst* 2000; 92(9): 709-20.
6. Rietbergen MM, Leemans CR, Bloemena E, et al. Increasing prevalence rates of HPV attributable oropharyngeal squamous cell carcinomas in the Netherlands as assessed by a validated test algorithm. *Int J Cancer* 2013; 132(7): 1565-71.
7. Pignon JP, le Maitre A, Maillard E, Bourhis J, Group M-NC. Meta-analysis of chemotherapy in head and neck cancer (MACH-NC): an update on 93 randomised trials and 17,346 patients. *Radiother Oncol* 2009; 92(1): 4-14.
8. Braakhuis BJ, Brakenhoff RH, Leemans CR. Treatment choice for locally advanced head and neck cancers on the basis of risk factors: biological risk factors. *Ann Oncol* 2012; 23 Suppl 10: x173-7.
9. Argiris A, Karamouzis MV, Raben D, Ferris RL. Head and neck cancer. *Lancet* 2008; 371(9625): 1695-709.
10. Leemans CR, Braakhuis BJ, Brakenhoff RH. The molecular biology of head and neck cancer. *Nat Rev Cancer* 2011; 11(1): 9-22.
11. Oksuz DC, Prestwich RJ, Carey B, et al. Recurrence patterns of locally advanced head and neck squamous cell carcinoma after 3D conformal (chemo)-radiotherapy. *Radiat Oncol* 2011; 6: 54.
12. Leeman JE, Li JG, Pei X, et al. Patterns of Treatment Failure and Postrecurrence Outcomes Among Patients With Locally Advanced Head and Neck Squamous Cell Carcinoma After Chemoradiotherapy Using Modern Radiation Techniques. *JAMA Oncol* 2017; 3(11): 1487-94.
13. Tabor MP, Brakenhoff RH, Ruijter-Schippers HJ, Kummer JA, Leemans CR, Braakhuis BJ. Genetically altered fields as origin of locally recurrent head and neck cancer: a retrospective study. *Clin Cancer Res* 2004; 10(11): 3607-13.
14. Graveland AP, Golusinski PJ, Buijze M, et al. Loss of heterozygosity at 9p and p53 immunopositivity in surgical margins predict local relapse in head and neck squamous cell carcinoma. *Int J Cancer* 2011; 128(8): 1852-9.
15. Braakhuis BJ, Tabor MP, Leemans CR, van der Waal I, Snow GB, Brakenhoff RH. Second primary tumors and field cancerization in oral and oropharyngeal cancer: molecular techniques provide new insights and definitions. *Head Neck* 2002; 24(2): 198-206.
16. van Houten VM, Leemans CR, Kummer JA, et al. Molecular diagnosis of surgical margins and local recurrence in head and neck cancer patients: a prospective study. *Clin Cancer Res* 2004; 10(11): 3614-20.
17. Hedberg ML, Goh G, Chiosea SI, et al. Genetic landscape of metastatic and recurrent head and neck squamous cell carcinoma. *J Clin Invest* 2016; 126(1): 169-80.
18. Mroz EA, Rocco JW. MATH, a novel measure of intratumor genetic heterogeneity, is high in poor-outcome classes of head and neck squamous cell carcinoma. *Oral Oncol* 2013; 49(3): 211-5.
19. Puram SV, Tirosh I, Parikh AS, et al. Single-Cell Transcriptomic Analysis of Primary and Metastatic Tumor Ecosystems in Head and Neck Cancer. *Cell* 2017; 171(7): 1611-24 e24.
20. Giefing M, Wierzbička M, Szyfter K, et al. Moving towards personalised therapy in head and neck squamous cell carcinoma through analysis of next generation sequencing data. *Eur J Cancer* 2016; 55: 147-57.
21. Geurts TW, van Velthuysen ML, Broekman F, et al. Differential diagnosis of pulmonary carcinoma following head and neck cancer by genetic analysis. *Clin Cancer Res* 2009; 15(3): 980-5.
22. van Ginkel JH, de Leng WW, de Bree R, van Es RJ, Willems SM. Targeted sequencing reveals TP53 as a potential diagnostic biomarker in the post-treatment surveillance of head and neck cancer. *Oncotarget* 2016; 7(38): 61575-86.
23. Smeets SJ, Hesselink AT, Speel EJ, et al. A novel algorithm for reliable detection of human papillomavirus in paraffin embedded head and neck cancer specimen. *Int J Cancer* 2007; 121(11): 2465-72.
24. Li H, Durbin R. Fast and accurate long-read alignment with Burrows-Wheeler transform. *Bioinformatics* 2010; 26(5): 589-95.

25. Scheinin I, Sie D, Bengtsson H, et al. DNA copy number analysis of fresh and formalin-fixed specimens by shallow whole-genome sequencing with identification and exclusion of problematic regions in the genome assembly. *Genome Res* 2014; 24(12): 2022-32.
26. MacDonald JR, Ziman R, Yuen RK, Feuk L, Scherer SW. The Database of Genomic Variants: a curated collection of structural variation in the human genome. *Nucleic Acids Res* 2014; 42(Database issue): D986-92.
27. Venkatraman ES, Olshen AB. A faster circular binary segmentation algorithm for the analysis of array CGH data. *Bioinformatics* 2007; 23(6): 657-63.
28. van de Wiel MA, Kim KI, Vosse SJ, van Wieringen WN, Wilting SM, Ylstra B. CGHcall: calling aberrations for array CGH tumor profiles. *Bioinformatics* 2007; 23(7): 892-4.
29. Ostrovskaya I, Seshan VE, Olshen AB, Begg CB. Clonality: an R package for testing clonal relatedness of two tumors from the same patient based on their genomic profiles. *Bioinformatics* 2011; 27(12): 1698-9.
30. Sie DB, A.; Veeramachaneni, R.; Bengtsson, H.; Schmidt, B.L.; Pinkel, D.; Ylstra, B.; Albertson, D.G. <https://research.vu.nl/ws/portalfiles/portal/41546932>: VU University medical center. 2017: 93-110.
31. Robin X, Turck N, Hainard A, et al. pROC: an open-source package for R and S+ to analyze and compare ROC curves. *BMC Bioinformatics* 2011; 12: 77.
32. Youden WJ. Index for rating diagnostic tests. *Cancer* 1950; 3(1): 32-5.
33. Cancer Genome Atlas N. Comprehensive genomic characterization of head and neck squamous cell carcinomas. *Nature* 2015; 517(7536): 576-82.
34. Cibulskis K, Lawrence MS, Carter SL, et al. Sensitive detection of somatic point mutations in impure and heterogeneous cancer samples. *Nat Biotechnol* 2013; 31(3): 213-9.
35. Li H, Handsaker B, Wysoker A, et al. The Sequence Alignment/Map format and SAMtools. *Bioinformatics* 2009; 25(16): 2078-9.
36. Koboldt DC, Zhang Q, Larson DE, et al. VarScan 2: somatic mutation and copy number alteration discovery in cancer by exome sequencing. *Genome Res* 2012; 22(3): 568-76.
37. Genomes Project C, Auton A, Brooks LD, et al. A global reference for human genetic variation. *Nature* 2015; 526(7571): 68-74.
38. Tabatabaieifar S, Thomassen M, Larsen MJ, Larsen SR, Kruse TA, Sorensen JA. The subclonal structure and genomic evolution of oral squamous cell carcinoma revealed by ultra-deep sequencing. *Oncotarget* 2017; 8(10): 16571-80.
39. Poell JB, Mendeville M, Sie D, Brink A, Brakenhoff RH, Ylstra B. ACE: Absolute Copy number Estimation from low-coverage whole-genome sequencing data. *Bioinformatics* 2018.
40. Niehr F, Eder T, Pilz T, et al. Multilayered Omics-Based Analysis of a Head and Neck Cancer Model of Cisplatin Resistance Reveals Intratumoral Heterogeneity and Treatment-Induced Clonal Selection. *Clin Cancer Res* 2018; 24(1): 158-68.
41. Al-Mamgani A, de Ridder M, Navran A, Klop WM, de Boer JP, Tesselaar ME. The impact of cumulative dose of cisplatin on outcome of patients with head and neck squamous cell carcinoma. *Eur Arch Otorhinolaryngol* 2017; 274(10): 3757-65.
42. Linge A, Lohaus F, Lock S, et al. HPV status, cancer stem cell marker expression, hypoxia gene signatures and tumour volume identify good prognosis subgroups in patients with HNSCC after primary radiochemotherapy: A multicentre retrospective study of the German Cancer Consortium Radiation Oncology Group (DKTK-ROG). *Radiother Oncol* 2016; 121(3): 364-73.
43. Galluzzi L, Senovilla L, Vitale I, et al. Molecular mechanisms of cisplatin resistance. *Oncogene* 2012; 31(15): 1869-83.
44. Martens-de Kemp SR, Brink A, van der Meulen IH, et al. The FA/BRCA Pathway Identified as the Major Predictor of Cisplatin Response in Head and Neck Cancer by Functional Genomics. *Mol Cancer Ther* 2017; 16(3): 540-50.
45. Martens-de Kemp SR, Dalm SU, Wijnolts FM, et al. DNA-bound platinum is the major determinant of cisplatin sensitivity in head and neck squamous carcinoma cells. *PLoS One* 2013; 8(4): e61555.
46. Tinhofer I, Stenzinger A, Eder T, et al. Targeted next-generation sequencing identifies molecular subgroups in squamous cell carcinoma of the head and neck with distinct outcome after concurrent chemoradiation. *Ann Oncol* 2016; 27(12): 2262-8.
47. De-Colle C, Menegakis A, Monnich D, et al. SDF-1/CXCR4 expression is an independent negative prognostic biomarker in patients with head and neck cancer after primary radiochemotherapy. *Radiother Oncol* 2018; 126(1): 125-31.



# 7

## General Discussion

Current treatment guidelines of HNSCC are largely based on disease stage according to TNM-classification, clinical variables such as subsite, age at diagnosis, comorbidities, and histological findings<sup>1,2</sup>. However clinical decision making is not precise, and tumor-related biological factors are not often considered except for the presence of HPV, while it is often a major determinant for treatment response and prognosis. Studies in this thesis and by others<sup>3-9</sup> show that prognostic models can be developed that are able to stratify patients more accurately in groups with a favorable and unfavorable prognosis than TNM-staging. A more accurate stratification of HNSCC patients will be of value to optimize personalization of treatment and counseling of patients for shared decision making, as has already been shown in other types of cancer<sup>10,11</sup>.

Exemplary is the progress in OPSCC where the discovery of an etiologic role of HPV infections and the improved prognosis of patients with these tumors led to the development of very accurate prognostic models<sup>12-14</sup> and a subsequent adaptation of the TNM-classification<sup>15</sup>. However, thus far this did not translate into a change in treatment policy, since chemoradiotherapy, which is typically the preferred treatment for OPSCC, is particularly effective for HPV-positive tumors<sup>16,17</sup>, and adaptation of this treatment regimen had a negative effect on survival and no lower morbidity<sup>18,19</sup>. Extrapolation of our knowledge of HPV as a causative and prognostic factor to other subsites of HNSCC remains uncertain. Although, it seems likely that the oral cavity is frequently exposed to HPV infections, the available data on HPV is limited and sometimes conflicting. For a large part, this may be caused by the HPV detection methods that have been applied to these tumors and that may be unreliable in non-oropharyngeal HNSCC. In OPSCC, the most frequently used test algorithm consists of p16<sup>ink4A</sup>(p16)-immunostaining followed by PCR-based detection of high-risk HPV DNA on the p16-immunopositive samples. This algorithm has been rigorously validated in OPSCC<sup>20,21</sup>. Unfortunately, this algorithm failed in non-oropharyngeal HNSCC because of a lower sensitivity of p16-immunostaining<sup>22,23</sup>. Consequently, researchers relied on sensitive DNA PCR based methods, causing false positive results and hampering the collection of accurate data on the role of HPV in non-oropharyngeal HNSCC. Therefore, we developed a new test method in this thesis that combines contamination-free sample handling with sequential detection of HPV DNA and mRNA by PCR.

Other methods for HPV detection are on the market, most notably the RNAscope Assay<sup>24-28</sup>. This assay is based on RNA *in situ* hybridization technology to detect E6/E7 mRNA expression of seven high-risk HPV genotypes (HPV16, 18, 31, 33, 35, 52, and 58). This technology is directly applied to diagnostic histological sections which enables direct visualization of the virus in the tumor cells. Theoretically, this improves specificity of the assay because the results are not influenced by non-carcinogenic, mucosal infections, whereas PCR-based techniques require nucleic acid extraction which destroys the tumor tissue hampering morphological correlation. Disadvantages are the costs of the assay itself and dependence on the interpretation by a pathologist. And although additional histological sections can be easily obtained in a prospective setting, sectioning of large retrospective series is time consuming and requires significant manpower. We have circumvented the sectioning by taking tumor core biopsies with a sterile, disposable biopsy punch using the H&E slide for guidance. Core biopsies are taken from tumor enriched areas which makes the risk of false negative results by sampling error and false positive results from a non-oncogenic infection much less likely. Results from studies with stringent detection techniques indicate that the prevalence of HPV in non-oropharyngeal HNSCC is very low<sup>22,23,29</sup>. Similarly, using the newly described detection algorithm, we found a prevalence of approximately 2% in a multicenter study of 940 oral cavity tumors, and a comparable prognosis of patients with HPV-positive and HPV-negative OSCC<sup>30</sup>. Therefore, although important for research, it is unlikely that the developed testing algorithm will be implemented for non-oropharyngeal tumors in future clinical practice.

It has been observed that at the molecular level HPV-positive tumors are very different from HPV-negative tumors<sup>31</sup>, although these results were obtained mostly from studies with OPSCCs. This implies that prognostic models based on DNA aberrations or gene expression profiles of HPV-negative tumor biopsies are not

automatically applicable to HPV-positive tumors. This may not only be true for molecular biomarkers, but also for radiomic prognostic models, since radiomic signatures appear to differ between HPV-positive and HPV-negative tumors as well<sup>32</sup>. Therefore, if molecular or radiomic prognostic models are used in the future, our newly developed HPV testing algorithm may be valuable to exclude the small group of tumors that are caused by HPV and consequently differ at the molecular level, since this may impact correct classification by the applied models. Of note, the newly developed HPV-rTcore workflow could be well applied for OPSCC in addition.

In the past, multiple prognostic models based on gene expression profiles have been evaluated for HNSCC, but none of these models are currently being used in clinical practice. This lack of implementation may be caused by the heterogeneity in patient cohorts under investigation and the lack of rigorous validation in the context of clinical and histopathological prognostic features. Moreover, the continuous global use of formalin fixation and paraffin embedding of tissue, demands the use of biomolecules isolated from an FFPE specimen to allow rapid and worldwide clinical implementation. Another complication is that previously published prognostic signatures are mostly based on microarray data that are often platform-dependent. Moreover, microarray technology became obsolete for expression profiling by the introduction of RNA sequencing (RNAseq). In this thesis, we have also developed a prognostic model based on microarray gene expression data, but in contrast to many other studies, we verified the findings using RNAseq datasets, transferred the platform to RT-qPCR and validated the gene signature by using the RT-qPCR platform. In our study, a homogeneous series of HPV-negative oral cavity tumors was used, and detailed clinical and histopathological information was available. It was indeed possible to identify patients with a good and poor prognosis on top of variables that are currently being used in clinical decision making. This result awaits prospective validation, and could be used to select patients for personalized treatment in the future, but the added value was somewhat limited. The limited added value relates to the fact that oral cancer is typically treated by surgery, which allows investigation of the specimen by histopathology. The microscopic findings of margin involvement, lymph node metastasis, extracapsular spread, depth of invasion, and growth pattern are already strong prognostic biomarkers, leaving only small room for improvements. Obviously, this situation is completely different for tumors from other subsites that are treated by radiotherapy, chemoradiotherapy and increasingly more by immunotherapy, and expression profiles might have much more added value for prognosis or therapy prediction. The same holds true for adjuvant treatment of oral cancers by postoperative radiotherapy or chemoradiotherapy, or future (neo)adjuvant immunotherapy.

In our studies we also attempted to solve another important issue in OSCC: the selection of patients with occult lymph node metastasis, also using gene expression profiles from the primary tumor. Previously, a gene signature had been identified and validated using a dedicated microarray platform, which was quite successful with a negative predictive value (NPV) of 89%<sup>33-35</sup>. We hypothesized that a platform transition to RT-qPCR could increase this performance because of the better standardization and larger dynamic range of qPCR and it would certainly enhance clinical implementation since qPCR is a diagnostic tool that is routinely used in molecular pathology. Remarkably, platform transition to qPCR did not further enhance the performance and we found a similar NPV in our cohort. Hence, the gene profiling approach seems to meet intrinsic limitations for staging of the clinically N0 neck, and likely additional relevant biological factors such as intratumor heterogeneity and stromal contamination, need to be explored to improve predictions. As an alternative, sentinel lymph node biopsy (SLNB) is used for detection of occult lymph node metastasis and its performance is superior as was demonstrated in a large multicenter trial when compared to palpation with USgFNAC, the most widely used methods to stage the clinically N0 neck<sup>36</sup>. Hence, SLNB remains the standard next to elective neck dissection. In floor-of-mouth tumors however, the performance of SLNB is less convincing because the signal from the primary tumor tends to mask the signal of the first echelon lymph node. Moreover, SLNB demands a second surgical procedure when the SN is tumor-positive. Although it is also

possible that technological improvements of the SLNB will overcome current limitations<sup>37-39</sup>, there may be a role for gene expression profiling<sup>40</sup> when the performance could be improved.

Although prediction of prognosis and presence of lymph node metastases by gene expression profiling was generally accurate in our study, misclassifications did occur. These misclassifications seemed to relate at least in part to the percentage of tumor cells in the biopsy under investigation. Traditionally, we estimated the surface area of tumor cells on H&E slides using a microscope. However, this method remained an estimate. To overcome this problem, our lab developed a new method that can more accurately determine the tumor percentage from DNA copy number data<sup>41</sup>. Using this method, we indeed found that the percentage of tumor cells was often lower than estimated. How strong this impacts prognostic accuracy remains to be determined: not the tumor cells but other cells in the biopsy may be important for prediction of prognosis. For instance, lymphocyte infiltrates may impact prognosis, and several immune profiles have been published that may be present in expression signatures derived from tumor biopsies<sup>42,43</sup>. An option to diminish the influence of tumor cell fraction is to use single cell RNA sequencing<sup>44</sup>. However, at present this technique is very technically challenging and expensive, and more suited to identify cells than to profile them.

Further, we need to consider that oral cancers, as a subgroup of HNSCCs, may contain subclasses of tumors that differ at the molecular level which may have impact on prognosis and performance of biomarkers such as gene expression signatures. In this respect, it has already been shown that a separate genetic subgroup of HPV-negative HNSCC exists with few DNA copy number alterations and a different mutational profile<sup>45</sup> (most notably many are p53 wild type). This subgroup was referred to as 'SCNA quiet' by the TCGA consortium and shown to have a favorable prognosis<sup>46</sup>. Recently we showed that also at least one cell line has been generated from this particular subgroup of tumors<sup>47</sup>. We have initiated a research project to more precisely define this group of copy number silent tumors, develop faster and easier methods for detection, as well as characterize them at the molecular and immunological level.

A final explanation for the misclassifications when applying gene expression profiling may be caused by intratumor heterogeneity, which we analyzed by studying DNA alterations in multiple biopsies of separate tumors. Although DNA profiles of most biopsies of a single tumor were very similar, some genetic differences were apparent indeed. These genetic differences may cause variance in gene expression profiles which could also explain incorrect predictions. Indeed, when we performed gene expression profiling of these multiple biopsies, we found that gene expression profiles in separate biopsies of one tumor varied significantly causing opposite predictions of prognosis and lymph node metastasis (unpublished data).

Besides using molecular methods, intratumor heterogeneity may be captured by advanced analysis of diagnostic imaging. Traditionally, prognostic biomarkers in imaging were qualitative features scored by an experienced radiologist such as invasion in neighboring structures, extra-nodal extension, etc., but more recently biomarkers have been identified by extraction of quantitative features, an approach termed "radiomics"<sup>48</sup>. Often, a large part of these features focus on tumor texture, which may indeed reflect intratumor heterogeneity. In the past, radiomic prognostic models were published based on extracted features from CT<sup>49-54</sup>, but since MRI is generally the diagnostic imaging modality of choice in HNSCC<sup>55</sup>, a study in this thesis was focused on MRI radiomics. Since radiomic analyses use images from the entire tumor as an input for feature extraction, this approach may indeed capture the complete tumor phenotype and is, in theory, able to determine and quantify heterogeneity. Furthermore, functional imaging such as diffusion-weighted imaging (DWI), perfusion weighted imaging and positron emission tomography (PET) are increasingly being used in HNSCC<sup>56</sup>, and may be of interest for future integrated radiomic analyses by itself or combined (e.g. PET-MRI<sup>57</sup>). Finally, imaging is acquired non-invasively, and can be performed at multiple time points during treatment, which could identify minor subclones that do not respond to therapy.

An important issue that we came across in our radiomics study is the lack of standardization between MRI scanners between centers and vendors. Moreover, this technique has evolved dramatically, for instance regarding acquisition protocols and magnetic field strength. These factors can influence radiomic features to some extent, which obstructs optimal usage of large historical databases and standardization in studies. Standardization should be a major future objective, and although challenging, it was already performed in PET<sup>58</sup> imaging.

Furthermore, manual delineation of the volume of interest is time consuming and may impact radiomic analyses<sup>59</sup>. In our study we did not find a major impact of delineation by different observers, but automatic and semi-automatic delineation may decrease interobserver variability and increase the speed of analysis. Algorithms that show these benefits exist for tumor delineation of CT images, and have been tested in lung cancer<sup>60,61</sup>. However, lung cancer tissue is more easy to distinguish from the air-containing surrounding normal lung tissue than HNSCC from its surrounding soft tissue. Nonetheless, future studies will hopefully reveal similar algorithms that can be successfully applied to MRI in HNSCC. For instance, deep learning based segmentation has been explored in HNSCC recently<sup>62</sup>.

Besides prediction of prognosis at baseline, post-treatment monitoring of patients could also identify high-risk patients for relapse and detect recurrent cancer. Patients are eligible for salvage therapy when the relapse is detected at an early stage, but currently recurrent disease is often diagnosed at an advanced stage and beyond curative salvage treatment options. Currently, an intensified follow-up regimen would imply more frequent visits and taking biopsies from any suspicious lesion. However, this lays a burden on the patient and health care system especially when a biopsy is obtained under general anesthesia. Preferably, a less invasive screening method would be available that detects recurrent tumors at an earlier stage, and detection of circulating tumor DNA (ctDNA) in plasma could meet with this demand.

In this thesis, a method for ctDNA detection was developed and tested in 40 newly diagnosed patients with different stages of HNSCC. The results are promising since tumor DNA was detected in 78% of the patients. However, our results are not applicable to a follow-up situation as the patients under investigation often had advanced stage disease, which is likely different from occult recurrent tumor. Ideally, the sensitivity should be higher and the assay should at least be able to detect (almost) all advanced stage tumors. At present, this is clearly not the case and future alterations to the assay should be aimed to increase sensitivity. However, it is possible that certain tumors do not induce any ctDNA as different mechanisms of leakage of tumor DNA may contribute to its presence<sup>63</sup>. Small tumors may not produce apoptotic cells or circulating tumor cells<sup>64</sup>. Nonetheless, tumors, including small tumors, might induce up- or downregulation of other processes, and this activity could be detected using other tools. For instance, it is known that expression profiles of platelets are changed by tumor presence and these changes may already be detectable in early stage recurrent tumors<sup>65,66</sup>. Therefore, this technique may be even more promising.

Next to applications in follow-up, detection of ctDNA may be of interest to select patients for targeted therapy treatment, which is most relevant in the setting of metastatic disease. Recently, a study was published in metastatic lung cancer in which 30% of patients had actionable mutations in plasma. Moreover, 85.7% of patients who were treated based on plasma testing showed a complete or a partial response or stable disease<sup>67</sup>. Similar results were later described by Sabari et al.<sup>68</sup> In general, less actionable mutations are known in HNSCC, but targeted approaches have been applied in recurrent/metastatic HNSCC<sup>69</sup>. For this application, the gene panel should be adjusted, or a commercially available kit could be used.

However, the ctDNA approach might be impacted by the observations in Chapter 6. Although fascinating, the molecular differences that were found between primary tumor and local recurrence after chemoradiotherapy, indicate that mutations or copy numbers of the primary are not suited as early biomarkers of recurrence. In this study, it appeared that copy numbers and mutations of paired tumors and recurrences differed

completely in approximately half of the cases. It should be noted, however, that the data were obtained by sequencing formalin-fixed paraffin-embedded tumor biopsies, which causes a lower coverage, sequencing artefacts and possibly induced mutations due to nucleotide deamination<sup>70</sup>. However, technical issues cannot explain the complete lack of identical molecular alterations that were found in the patients with genetically unrelated recurrences. In other types of cancer, it has been shown that therapy can select resistant cells<sup>71</sup>. Also in our study, it is possible that CRT caused selection of a minor subclone of cells that was not detected and consequently overlooked in the primary tumor. Either the fraction of these cells may have been below the detection rate of target-enrichment DNA sequencing (<0.1-1%) or sampling-error occurred, and the analyzed biopsy did not contain this subclone. More advanced sequencing techniques on multiple biopsies may detect such minor subclones in the tumor that may be predictive for treatment response in the future. Moreover, imaging may guide sampling for biomarker analyses by determination of the most aggressive area of the tumor<sup>72</sup>. Even more remarkable was that the observed heterogeneity was detected in driver genes, whereas one would expect that early mutations in carcinogenesis (e.g. in *TP53*) are still shared. Obviously, therapy may also induce mutations<sup>73-75</sup>, but in this situation these mutations are expected to be additive to earlier mutations and hence in that situation most driver mutations should have remained identical. Finally, the genetically unrelated relapses may evolve from the surrounding mucosa. Most likely, the surrounding contains preneoplastic fields which may further evolve into invasive tumors during follow-up. Often these preneoplastic fields are genetically related to the primary tumor, but genetically unrelated fields also exist<sup>76</sup>. In support of this hypothesis, Sherborne et al. found that germline *TP53* mutations in pediatric patients may sensitize these patients to develop second malignant neoplasms after genotoxic treatment<sup>77</sup>. Preneoplastic fields often harbor *TP53* mutations<sup>78</sup>, and may therefore be at greater risk for progression after chemoradiotherapy. Hence, a hypothesis of surrounding, genetically unrelated, preneoplastic fields that progress in part because of genotoxic treatment, is not unlikely and supported by some clinical data. This remarkable observation will certainly fuel future research as it may be a major contributor to treatment failure. Besides the underlying biological process, it will also be important to determine whether this effect can also be found after radiotherapy alone, while we expect that it will not play a role after surgery.

Finally, an important development in the treatment of HNSCC should be mentioned in this discussion. In the past years, immunotherapy gained attention with very promising results even as second-line treatment for recurrent/metastatic HNSCC. In a fraction of patients, treatment with checkpoint inhibitors of programmed cell death protein 1 (PD-1) showed durable responses<sup>79-81</sup>. However, immunotherapy is expensive and needs to be administered repeatedly for several months to years. Moreover, serious side effects may occur, which occasionally include life-threatening immune-mediated pneumonitis. Therefore, selection of patients that are likely to respond is critical to prevent unwanted side-effects and reduce costs. Expression levels of PD-L1 are known to predict treatment response to some degree, but a significant number of patients with high expression of PD-L1 do not respond to this treatment<sup>82</sup>. Other biomarkers include microsatellite instability (MSI) and high tumor mutation burden in general<sup>83,84</sup>, and for these biomarkers a liquid biopsy approach has been published recently<sup>85</sup>. Future research will likely determine its usability in HNSCC and also focus on HNSCC specific biomarkers.

Major challenges in the clinical management of HNSCC are to improve locoregional control of the tumor, to develop better treatments for recurrent/metastatic disease, to find biomarkers for early detection of recurrent disease, to improve prediction of prognosis, and to identify markers of personalized treatment. In Chapters 3 and 4 different approaches to discover novel biomarkers in HNSCC to improve prediction of prognosis are presented. In Chapter 3, gene expression analyses were explored, whereas in Chapter 4, MRI radiomics was investigated. These methods might be complementary and both can be applied in advance of surgery or organ-preserving treatment. Additionally, HPV could be incorporated to the prognostic models presented in Chapters 3 and 4 using the simple and accurate method for HPV detection in non-oropharyngeal cancer, which

is the subject of Chapter 2. Future research will show if HPV is of interest in these tumors. Clinical outcome may also improve with better follow up methods to detect recurrent disease at an early stage. One possibility is to screen patients for detection of ctDNA in plasma, and in Chapter 5 we developed and tested a method for this. However, further optimization steps are required as well as prospective, clinical validation of early detection of recurrent disease. An overarching complication seems intratumor genetic heterogeneity. This may be of major influence to the accuracy of the prognostic models in Chapters 3 and 4, and impact ctDNA as biomarker for early detection of recurrent disease. It might further be a key factor in treatment success on its own. This phenomenon was explored in Chapter 6 and if not tackled in future studies this may indeed get in the way of future solutions to address the clinical challenges.

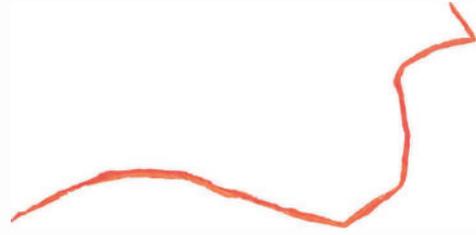
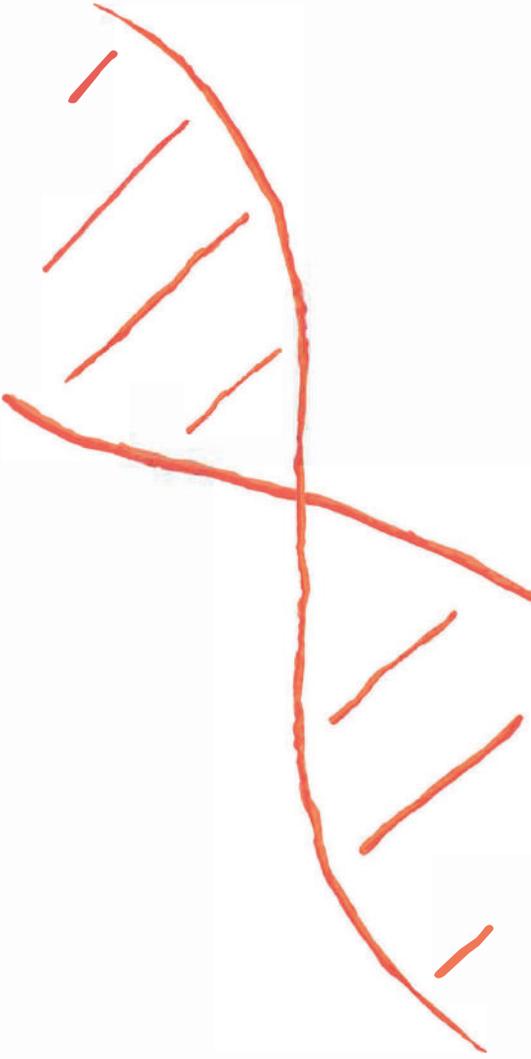
## REFERENCES

1. Network NCC. Head and Neck Cancers (version 1.2018). [https://www.nccn.org/professionals/physician\\_gls/pdf/head-and-neck.pdf](https://www.nccn.org/professionals/physician_gls/pdf/head-and-neck.pdf) (accessed February 11th 2020).
2. Dutch Association of Otorhinolaryngology and Head & Neck Surgery NVvK-N-OeHvhH-H. Richtlijn Hoofdhals tumoren. 2014. [https://richtlijndatabase.nl/richtlijn/hoofd-halstumoren/hoofd-halstumoren\\_-\\_korte\\_beschrijving.html](https://richtlijndatabase.nl/richtlijn/hoofd-halstumoren/hoofd-halstumoren_-_korte_beschrijving.html) (accessed February 11th 2020).
3. Mes SW, Te Beest D, Poli T, et al. Prognostic modeling of oral cancer by gene profiles and clinicopathological co-variables. *Oncotarget* 2017; 8(35): 59312-23.
4. Mes SW, van Velden FHP, Peltenburg B, et al. Outcome prediction of head and neck squamous cell carcinoma by MRI radiomic signatures. *Eur Radiol* 2020.
5. Zhang Y, Koneva LA, Virani S, et al. Subtypes of HPV-Positive Head and Neck Cancers Are Associated with HPV Characteristics, Copy Number Alterations, PIK3CA Mutation, and Pathway Signatures. *Clin Cancer Res* 2016; 22(18): 4735-45.
6. Locati LD, Serafini MS, Ianno MF, et al. Mining of Self-Organizing Map Gene-Expression Portraits Reveals Prognostic Stratification of HPV-Positive Head and Neck Squamous Cell Carcinoma. *Cancers (Basel)* 2019; 11(8).
7. Gleber-Netto FO, Rao X, Guo T, et al. Variations in HPV function are associated with survival in squamous cell carcinoma. *JCI Insight* 2019; 4(1).
8. De Cecco L, Bossi P, Locati L, Canevari S, Licitra L. Comprehensive gene expression meta-analysis of head and neck squamous cell carcinoma microarray data defines a robust survival predictor. *Ann Oncol* 2014; 25(8): 1628-35.
9. Lohavanichbutr P, Mendez E, Holsinger FC, et al. A 13-gene signature prognostic of HPV-negative OSCC: discovery and external validation. *Clin Cancer Res* 2013; 19(5): 1197-203.
10. van 't Veer LJ, Dai H, van de Vijver MJ, et al. Gene expression profiling predicts clinical outcome of breast cancer. *Nature* 2002; 415(6871): 530-6.
11. Cardoso F, van't Veer LJ, Bogaerts J, et al. 70-Gene Signature as an Aid to Treatment Decisions in Early-Stage Breast Cancer. *N Engl J Med* 2016; 375(8): 717-29.
12. Ang KK, Harris J, Wheeler R, et al. Human papillomavirus and survival of patients with oropharyngeal cancer. *N Engl J Med* 2010; 363(1): 24-35.
13. O'Sullivan B, Huang SH, Siu LL, et al. Deintensification candidate subgroups in human papillomavirus-related oropharyngeal cancer according to minimal risk of distant metastasis. *J Clin Oncol* 2013; 31(5): 543-50.
14. Rietbergen MM, Brakenhoff RH, Bloemena E, et al. Human papillomavirus detection and comorbidity: critical issues in selection of patients with oropharyngeal cancer for treatment De-escalation trials. *Ann Oncol* 2013; 24(11): 2740-5.
15. Brierley JDG, M. K.; Wittekind, C., eds. TNM Classification of Malignant Tumours. 8th Edition ed. Hoboken: Wiley-Blackwell; 2016.
16. Cheraghlou S, Yu PK, Otremba MD, et al. Treatment deintensification in human papillomavirus-positive oropharynx cancer: Outcomes from the National Cancer Data Base. *Cancer* 2018; 124(4): 717-26.
17. Vokes EE, Agrawal N, Seiwert TY. HPV-Associated Head and Neck Cancer. *J Natl Cancer Inst* 2015; 107(12): djv344.
18. Mehanna H, Robinson M, Hartley A, et al. Radiotherapy plus cisplatin or cetuximab in low-risk human papillomavirus-positive oropharyngeal cancer (De-ESCALaTE HPV): an open-label randomised controlled phase 3 trial. *Lancet* 2019; 393(10166): 51-60.
19. Gillison ML, Trotti AM, Harris J, et al. Radiotherapy plus cetuximab or cisplatin in human papillomavirus-positive oropharyngeal cancer (NRG Oncology RTOG 1016): a randomised, multicentre, non-inferiority trial. *Lancet* 2019; 393(10166): 40-50.
20. Smeets SJ, Hesselink AT, Speel EJ, et al. A novel algorithm for reliable detection of human papillomavirus in paraffin embedded head and neck cancer specimen. *Int J Cancer* 2007; 121(11): 2465-72.
21. Rietbergen MM, Leemans CR, Bloemena E, et al. Increasing prevalence rates of HPV attributable oropharyngeal squamous cell carcinomas in the Netherlands as assessed by a validated test algorithm. *Int J Cancer* 2013; 132(7): 1565-71.
22. Lingen MW, Xiao W, Schmitt A, et al. Low etiologic fraction for high-risk human papillomavirus in oral cavity squamous cell carcinomas. *Oral Oncol* 2013; 49(1): 1-8.
23. Chung CH, Zhang Q, Kong CS, et al. p16 protein expression and human papillomavirus status as prognostic biomarkers of nonoropharyngeal head and neck squamous cell carcinoma. *J Clin Oncol* 2014; 32(35): 3930-8.

24. Wang F, Flanagan J, Su N, et al. RNAscope: a novel in situ RNA analysis platform for formalin-fixed, paraffin-embedded tissues. *J Mol Diagn* 2012; 14(1): 22-9.
25. Ukpo OC, Flanagan JJ, Ma XJ, Luo Y, Thorstad WL, Lewis JS, Jr. High-risk human papillomavirus E6/E7 mRNA detection by a novel in situ hybridization assay strongly correlates with p16 expression and patient outcomes in oropharyngeal squamous cell carcinoma. *Am J Surg Pathol* 2011; 35(9): 1343-50.
26. Bishop JA, Ma XJ, Wang H, et al. Detection of transcriptionally active high-risk HPV in patients with head and neck squamous cell carcinoma as visualized by a novel E6/E7 mRNA in situ hybridization method. *Am J Surg Pathol* 2012; 36(12): 1874-82.
27. Rooper LM, Gandhi M, Bishop JA, Westra WH. RNA in-situ hybridization is a practical and effective method for determining HPV status of oropharyngeal squamous cell carcinoma including discordant cases that are p16 positive by immunohistochemistry but HPV negative by DNA in-situ hybridization. *Oral Oncol* 2016; 55: 11-6.
28. Kerr DA, Arora KS, Mahadevan KK, et al. Performance of a Branch Chain RNA In Situ Hybridization Assay for the Detection of High-risk Human Papillomavirus in Head and Neck Squamous Cell Carcinoma. *Am J Surg Pathol* 2015; 39(12): 1643-52.
29. Castellsague X, Alemany L, Quer M, et al. HPV Involvement in Head and Neck Cancers: Comprehensive Assessment of Biomarkers in 3680 Patients. *J Natl Cancer Inst* 2016; 108(6): djv403.
30. Nauta IH, Heideman DAM, Brink A, et al. The unveiled reality of human papillomavirus as risk factor for oral cavity squamous cell carcinoma. *Submitted*.
31. Hayes DN, Van Waes C, Seiwert TY. Genetic Landscape of Human Papillomavirus-Associated Head and Neck Cancer and Comparison to Tobacco-Related Tumors. *J Clin Oncol* 2015; 33(29): 3227-34.
32. Leijenaar RT, Bogowicz M, Jochems A, et al. Development and validation of a radiomic signature to predict HPV (p16) status from standard CT imaging: a multicenter study. *Br J Radiol* 2018; 91(1086): 20170498.
33. Roepman P, Kemmeren P, Wessels LF, Slootweg PJ, Holstege FC. Multiple robust signatures for detecting lymph node metastasis in head and neck cancer. *Cancer Res* 2006; 66(4): 2361-6.
34. Roepman P, Wessels LF, Kettelarij N, et al. An expression profile for diagnosis of lymph node metastases from primary head and neck squamous cell carcinomas. *Nat Genet* 2005; 37(2): 182-6.
35. van Hooff SR, Leusink FK, Roepman P, et al. Validation of a gene expression signature for assessment of lymph node metastasis in oral squamous cell carcinoma. *J Clin Oncol* 2012; 30(33): 4104-10.
36. Schilling C, Stoeckli SJ, Haerle SK, et al. Sentinel European Node Trial (SENT): 3-year results of sentinel node biopsy in oral cancer. *Eur J Cancer* 2015; 51(18): 2777-84.
37. Heuveling DA, van Schie A, Vugts DJ, et al. Pilot study on the feasibility of PET/CT lymphoscintigraphy with 89Zr-nanocolloidal albumin for sentinel node identification in oral cancer patients. *J Nucl Med* 2013; 54(4): 585-9.
38. Heuveling DA, Visser GW, de Groot M, et al. Nanocolloidal albumin-IRDye 800CW: a near-infrared fluorescent tracer with optimal retention in the sentinel lymph node. *Eur J Nucl Med Mol Imaging* 2012; 39(7): 1161-8.
39. van den Berg NS, Brouwer OR, Klop WM, et al. Concomitant radio- and fluorescence-guided sentinel lymph node biopsy in squamous cell carcinoma of the oral cavity using ICG-(99m)Tc-nanocolloid. *Eur J Nucl Med Mol Imaging* 2012; 39(7): 1128-36.
40. Leusink FK, van Es RJ, de Bree R, et al. Novel diagnostic modalities for assessment of the clinically node-negative neck in oral squamous-cell carcinoma. *Lancet Oncol* 2012; 13(12): e554-61.
41. Poell JB, Mendeville M, Sie D, Brink A, Brakenhoff RH, Ylstra B. ACE: Absolute Copy number Estimation from low-coverage whole-genome sequencing data. *Bioinformatics* 2018.
42. Wondergem NE, Nauta IH, Muijlwijk T, Leemans CR, van de Ven R. The Immune Microenvironment in Head and Neck Squamous Cell Carcinoma: on Subsets and Subsites. *Curr Oncol Rep* 2020; 22(8): 81.
43. de Ruiter EJ, Ooft ML, Devriese LA, Willems SM. The prognostic role of tumor infiltrating T-lymphocytes in squamous cell carcinoma of the head and neck: A systematic review and meta-analysis. *Oncoimmunology* 2017; 6(11): e1356148.
44. Zhang AW, O'Flanagan C, Chavez EA, et al. Probabilistic cell-type assignment of single-cell RNA-seq for tumor microenvironment profiling. *Nat Methods* 2019; 16(10): 1007-15.
45. Smeets SJ, Brakenhoff RH, Ylstra B, et al. Genetic classification of oral and oropharyngeal carcinomas identifies subgroups with a different prognosis. *Cell Oncol* 2009; 31(4): 291-300.
46. Cancer Genome Atlas N. Comprehensive genomic characterization of head and neck squamous cell carcinomas. *Nature* 2015; 517(7536): 576-82.
47. van Harten AM, Poell JB, Buijze M, et al. Characterization of a head and neck cancer-derived cell line panel confirms the distinct TP53-proficient copy number-silent subclass. *Oral Oncol* 2019; 98: 53-61.

48. Kumar V, Gu Y, Basu S, et al. Radiomics: the process and the challenges. *Magn Reson Imaging* 2012; 30(9): 1234-48.
49. Aerts HJ, Velazquez ER, Leijenaar RT, et al. Decoding tumour phenotype by noninvasive imaging using a quantitative radiomics approach. *Nat Commun* 2014; 5: 4006.
50. Parmar C, Leijenaar RT, Grossmann P, et al. Radiomic feature clusters and prognostic signatures specific for Lung and Head & Neck cancer. *Sci Rep* 2015; 5: 11044.
51. Leger S, Zwanenburg A, Pilz K, et al. A comparative study of machine learning methods for time-to-event survival data for radiomics risk modelling. *Sci Rep* 2017; 7(1): 13206.
52. Vallieres M, Kay-Rivest E, Perrin LJ, et al. Radiomics strategies for risk assessment of tumour failure in head-and-neck cancer. *Sci Rep* 2017; 7(1): 10117.
53. Bogowicz M, Riesterer O, Stark LS, et al. Comparison of PET and CT radiomics for prediction of local tumor control in head and neck squamous cell carcinoma. *Acta Oncol* 2017; 56(11): 1531-6.
54. Forghani R, Chatterjee A, Reinhold C, et al. Head and neck squamous cell carcinoma: prediction of cervical lymph node metastasis by dual-energy CT texture analysis with machine learning. *Eur Radiol* 2019; 29(11): 6172-81.
55. Juliano A, Moonis G. Computed Tomography Versus Magnetic Resonance in Head and Neck Cancer: When to Use What and Image Optimization Strategies. *Magn Reson Imaging Clin N Am* 2018; 26(1): 63-84.
56. Dai YL, King AD. State of the art MRI in head and neck cancer. *Clin Radiol* 2018; 73(1): 45-59.
57. Martens RM, Koopman T, Noij DP, et al. Adherence to pretreatment and intratreatment imaging of head and neck squamous cell carcinoma patients undergoing (chemo) radiotherapy in a research setting. *Clin Imaging* 2020; 69: 82-90.
58. Boellaard R, Delgado-Bolton R, Oyen WJ, et al. FDG PET/CT: EANM procedure guidelines for tumour imaging: version 2.0. *Eur J Nucl Med Mol Imaging* 2015; 42(2): 328-54.
59. Balagurunathan Y, Gu Y, Wang H, et al. Reproducibility and Prognosis of Quantitative Features Extracted from CT Images. *Transl Oncol* 2014; 7(1): 72-87.
60. Parmar C, Rios Velazquez E, Leijenaar R, et al. Robust Radiomics feature quantification using semiautomatic volumetric segmentation. *PLoS One* 2014; 9(7): e102107.
61. Rios Velazquez E, Aerts HJ, Gu Y, et al. A semiautomatic CT-based ensemble segmentation of lung tumors: comparison with oncologists' delineations and with the surgical specimen. *Radiother Oncol* 2012; 105(2): 167-73.
62. Schouten JPE, Noteboom S, Martens RM, et al. Automatic segmentation of head and neck primary tumors on MRI using a multi-view CNN. *Submitted*.
63. Thierry AR, El Messaoudi S, Gahan PB, Anker P, Stroun M. Origins, structures, and functions of circulating DNA in oncology. *Cancer Metastasis Rev* 2016; 35(3): 347-76.
64. Cescon DWB, S.V.; Chan, S.M.; Siu, L.L. Circulating tumor DNA and liquid biopsy in oncology. *Nature Cancer* 2020; 1(3): 276-90.
65. Best MG, Sol N, In 't Veld S, et al. Swarm Intelligence-Enhanced Detection of Non-Small-Cell Lung Cancer Using Tumor-Educated Platelets. *Cancer Cell* 2017; 32(2): 238-52 e9.
66. Best MG, Sol N, Kooi I, et al. RNA-Seq of Tumor-Educated Platelets Enables Blood-Based Pan-Cancer, Multiclass, and Molecular Pathway Cancer Diagnostics. *Cancer Cell* 2015; 28(5): 666-76.
67. Aggarwal C, Thompson JC, Black TA, et al. Clinical Implications of Plasma-Based Genotyping With the Delivery of Personalized Therapy in Metastatic Non-Small Cell Lung Cancer. *JAMA Oncol* 2019; 5(2): 173-80.
68. Sabari JK, Offin M, Stephens D, et al. A Prospective Study of Circulating Tumor DNA to Guide Matched Targeted Therapy in Lung Cancers. *J Natl Cancer Inst* 2019; 111(6): 575-83.
69. Sacco AG, Cohen EE. Current Treatment Options for Recurrent or Metastatic Head and Neck Squamous Cell Carcinoma. *J Clin Oncol* 2015; 33(29): 3305-13.
70. Do H, Dobrovic A. Sequence artifacts in DNA from formalin-fixed tissues: causes and strategies for minimization. *Clin Chem* 2015; 61(1): 64-71.
71. Holohan C, Van Schaeybroeck S, Longley DB, Johnston PG. Cancer drug resistance: an evolving paradigm. *Nat Rev Cancer* 2013; 13(10): 714-26.
72. Gates EDH, Lin JS, Weinberg JS, et al. Guiding the first biopsy in glioma patients using estimated Ki-67 maps derived from MRI: conventional versus advanced imaging. *Neuro Oncol* 2019; 21(4): 527-36.
73. Newhauser WD, Durante M. Assessing the risk of second malignancies after modern radiotherapy. *Nat Rev Cancer* 2011; 11(6): 438-48.
74. El Ghissassi F, Baan R, Straif K, et al. A review of human carcinogens--part D: radiation. *Lancet Oncol* 2009; 10(8): 751-2.

75. Sherborne AL, Davidson PR, Yu K, Nakamura AO, Rashid M, Nakamura JL. Mutational Analysis of Ionizing Radiation Induced Neoplasms. *Cell Rep* 2015; 12(11): 1915-26.
76. Leemans CR, Braakhuis BJ, Brakenhoff RH. The molecular biology of head and neck cancer. *Nat Rev Cancer* 2011; 11(1): 9-22.
77. Sherborne AL, Lavergne V, Yu K, et al. Somatic and Germline TP53 Alterations in Second Malignant Neoplasms from Pediatric Cancer Survivors. *Clin Cancer Res* 2017; 23(7): 1852-61.
78. Braakhuis BJ, Tabor MP, Kummer JA, Leemans CR, Brakenhoff RH. A genetic explanation of Slaughter's concept of field cancerization: evidence and clinical implications. *Cancer Res* 2003; 63(8): 1727-30.
79. Seiwert TY, Burtneess B, Mehra R, et al. Safety and clinical activity of pembrolizumab for treatment of recurrent or metastatic squamous cell carcinoma of the head and neck (KEYNOTE-012): an open-label, multicentre, phase 1b trial. *Lancet Oncol* 2016; 17(7): 956-65.
80. Ferris RL, Blumenschein G, Jr., Fayette J, et al. Nivolumab for Recurrent Squamous-Cell Carcinoma of the Head and Neck. *N Engl J Med* 2016; 375(19): 1856-67.
81. Cohen EEW, Soulieres D, Le Tourneau C, et al. Pembrolizumab versus methotrexate, docetaxel, or cetuximab for recurrent or metastatic head-and-neck squamous cell carcinoma (KEYNOTE-040): a randomised, open-label, phase 3 study. *Lancet* 2019; 393(10167): 156-67.
82. Cramer JD, Burtneess B, Ferris RL. Immunotherapy for head and neck cancer: Recent advances and future directions. *Oral Oncol* 2019; 99: 104460.
83. Le DT, Uram JN, Wang H, et al. PD-1 Blockade in Tumors with Mismatch-Repair Deficiency. *N Engl J Med* 2015; 372(26): 2509-20.
84. Le DT, Durham JN, Smith KN, et al. Mismatch repair deficiency predicts response of solid tumors to PD-1 blockade. *Science* 2017; 357(6349): 409-13.
85. Georgiadis A, Durham JN, Keefer LA, et al. Noninvasive Detection of Microsatellite Instability and High Tumor Mutation Burden in Cancer Patients Treated with PD-1 Blockade. *Clin Cancer Res* 2019; 25(23): 7024-34.



8

**Summary**

## SUMMARY

In the introduction (**Chapter 1**) general information on head and neck cancer is provided and it is outlined that treatment of HNSCC patients is primarily based on TNM-staging, imaging and clinical variables such as age at diagnosis and tumor subsite, and for surgically treated patients also on histological examination of the specimen<sup>1</sup>. In general, patients with a good prognosis (early stage disease) are treated with single modality therapy (surgery or radiotherapy), whereas patients with a poor prognosis (advanced stage disease) receive multimodality treatment (surgery with adjuvant (chemo)radiotherapy or upfront concomitant chemoradiotherapy). Also immunotherapy is being introduced, and for specific indications bioradiotherapy is applied using cetuximab instead of cisplatin. Although prognostic stratification of patients has improved with the introduction of TNM 8<sup>2</sup>, refined prognostic models will be required to stratify patients more accurately in groups with a favorable and unfavorable prognosis to optimize personalization of treatment and counselling of patients. One approach to improve prediction of prognosis is to include biomarkers in these models, which may better represent the differences in tumor biology. In HNSCC, the most prominent biomarker is HPV in oropharyngeal cancer, and this has been introduced in TNM 8<sup>2</sup> by the surrogate marker p16 immunostaining. For non-oropharyngeal HNSCC such a strong biomarker has not been identified and introduced in clinical practice. HPV may also be relevant for non-oropharyngeal HNSCC, but the standard testing algorithms fail in these tumors<sup>3,4</sup> and an accurate alternative is lacking, hampering studies on the role of HPV in non-oropharyngeal HNSCC.

The aim of this thesis was to develop and evaluate novel prognostic biomarkers in HNSCC that can be applied in general or to specific tumor subsites. Moreover, a liquid biopsy test was developed that may be used in the future for more comprehensive genetic analysis of the tumor and for follow-up after therapy. Finally, we studied intratumor heterogeneity and molecular differences between primary and recurrent tumors.

In **Chapter 2** a molecular HPV testing algorithm (HPV-rTcore assay) was developed that can be applied to archival specimen of oropharyngeal and non-oropharyngeal HNSCC. Tumor sampling was performed with punch biopsies from tumor enriched regions, and DNA and RNA was extracted from the cores subsequently. Next, PCR-based detection of HPV-DNA was performed for 15 HR-HPV types, and positive results were validated by detection of E6 mRNA. This testing algorithm was validated in OPSCC and OSCC samples, and reached an overall accuracy of 97% in OPSCC and 100% in OSCC. The HPV-rTcore assay is currently used to determine the prevalence of HPV in non-oropharyngeal HNSCC and its prognostic impact, and might become a new standard for OPSCC in the diagnostic workup.

In **Chapter 3** a prognostic model is presented based on gene expression profiling. Microarray data of OSCC was used to select a 22-gene signature to predict N-stage and a 40-gene signature to predict prognosis. These signatures were transferred to a quantitative PCR platform and validated in an independent OSCC cohort. The signature predicted the presence of occult lymph node metastasis in cT1-2N0 patients with an NPV of 84%. Additionally, the integration of the 40-gene signature with clinical and pathological variables provided accurate prognostic models that outperformed TNM. Finally, the 40-gene signature identified a subpopulation of patients, currently considered at low-risk for disease-related death, who showed an unexpected poor prognosis. These results may be used for counselling and to select patients for active surveillance instead of neck dissection or to select patients that may benefit from adjuvant therapy.

Another approach to develop prognostic models for HNSCC was presented in **Chapter 4** using imaging biomarkers: native T1-weighted MRI scans of OSCC and HPV-negative OPSCC patients were used to extract quantitative features from manually delineated tumors. This approach is also known as 'radiomics'. In the research reported in this chapter, 545 features were extracted that describe tumor signal intensity, shape and texture. To these features, redundancy filtering and factor analysis was applied and the acquired factors were used for prognostic modelling. These models based on MRI radiomics provided additional prognostic

information to known clinical variables, with the best performance of models with a combination of clinical and radiomic variables. Interestingly, although variation in MRI vendors and acquisition protocols was large, this did not preclude the generation of radiomic prognostic models in this study. These models can be integrated with standard diagnostic work-up, as native T1-weighted images are used in most diagnostic protocols.

Recurrent HNSCC can be salvaged surgically, but only when detected at an early stage, which is currently often not the case. Recurrent cancer is difficult to detect and the gold standard, examination under general anesthesia with biopsy, is very invasive and unsuited for routine screening. Possibly, early discovery can be achieved by detection of circulating tumor DNA (ctDNA) which is shedded in plasma by apoptotic tumor cells. In **Chapter 5** a ctDNA detection method was developed for HNSCC that focuses on somatic mutations, copy number aberrations and HPV-DNA detection. It was hypothesized in this chapter that a comprehensive approach that identifies different molecular alterations would increase sensitivity of ctDNA detection given the considerable genetic heterogeneity between HNSCCs. Somatic mutations and CNAs were detected in plasma DNA in 67% and 52% of HNSCC patients, whereas HPV-DNA in plasma was detected in 100% of patients with HPV-positive tumors. The combined analysis increased the detection rate to 78%. Calling somatic mutations in plasma DNA improved with prior knowledge of mutations in the tumor, but prior knowledge was not required to detect CNAs and HPV-DNA in plasma. This study established an encouraging opening for early detection of recurrent disease by longitudinal screening of a large cohort.

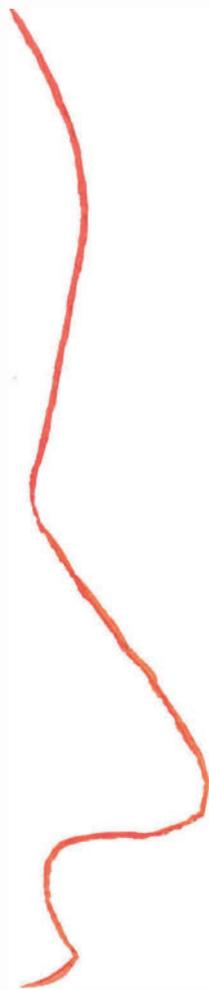
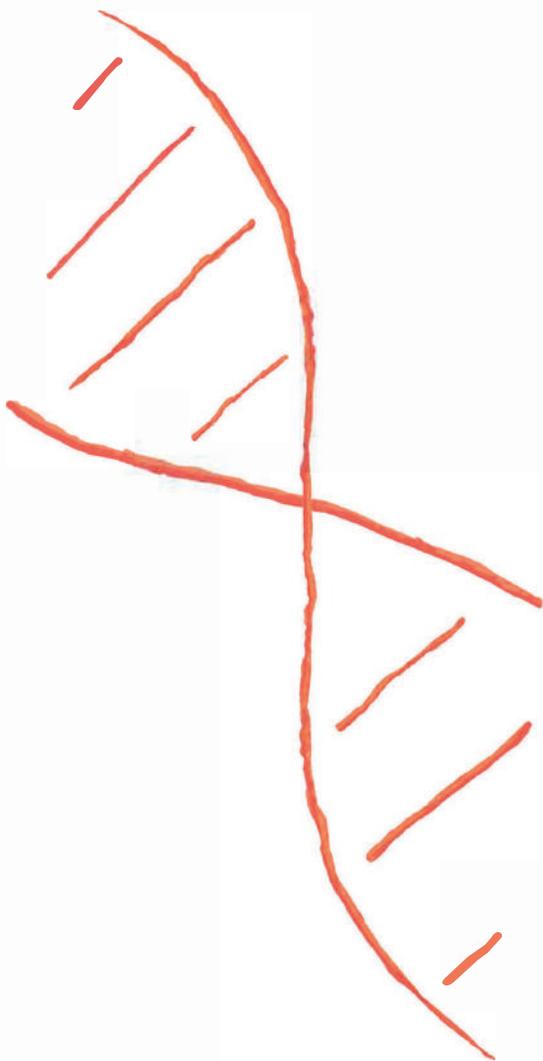
Although biomarker-based prognostic modelling appears to be possible, all model accuracies appears to plateau at some point, suggesting that a particular biological factor limits the application of tumor biomarkers analyzed at baseline. In **Chapter 6** it was hypothesized that this may be related to intratumor genetic heterogeneity either detectable at the time of diagnosis, or exposed by treatment and reflected by genetic differences between primary and recurrent tumor. Indeed differences were found in multiregion biopsies of surgical specimen, but CNA profiles appeared to be largely similar. Heterogeneity might be more profound when sequencing for mutations, which is currently under investigation. However, heterogeneity became most prominent when tumors and recurrences were genetically compared as 50% of recurrent tumors after chemoradiotherapy appeared to be genetically unrelated to the primary tumor. This finding indicates that recurrent tumors likely evolve from very small subclones that are not detected by molecular strategies on bulk tumor DNA. Given this observation, the performance of prognostic models based on bulk tumor characteristic may indeed reach limitations as minor subclones may determine recurrent disease and consequently prognosis, and future studies should include this apparent key factor.

In **Chapter 7** the data presented in this thesis are discussed and put in a larger perspective.

## REFERENCES

1. Chow LQM. Head and Neck Cancer. *N Engl J Med* 2020; 382(1): 60-72.
2. Brierley JDG, M. K.; Wittekind, C., eds. *TNM Classification of Malignant Tumours*. 8th Edition ed. Hoboken: Wiley-Blackwell; 2016.
3. Lingen MW, Xiao W, Schmitt A, et al. Low etiologic fraction for high-risk human papillomavirus in oral cavity squamous cell carcinomas. *Oral Oncol* 2013; 49(1): 1-8.
4. Chung CH, Zhang Q, Kong CS, et al. p16 protein expression and human papillomavirus status as prognostic biomarkers of nonoropharyngeal head and neck squamous cell carcinoma. *J Clin Oncol* 2014; 32(35): 3930-8.





# 9

**Nederlandse samenvatting**

**Hoofdstuk 1** is een algemene introductie met informatie over hoofd-halskanker. Dit hoofdstuk beschrijft onder meer de behandeling van patiënten met een hoofd-halsplaveiselcelcarcinoom (HNSCC). De behandeling is met name gebaseerd op de TNM-stadiëring, beeldvorming en klinische variabelen zoals leeftijd op moment van diagnose en locatie van de tumor. Daarnaast is de behandeling afhankelijk van histologisch onderzoek, maar dit geldt met name voor tumoren die primair chirurgisch worden behandeld<sup>1</sup>. Door de recente introductie van TNM 8<sup>2</sup> kan de prognose van patiënten al beter voorspeld worden, maar aanvullende prognostische modellen zijn nodig om patiënten nauwkeuriger in groepen met goede en slechte prognose in te delen zodat het mogelijk is de behandeling hierop aan te passen en patiënten beter te kunnen informeren. Het toevoegen van biomarkers aan prognostische modellen is een mogelijkheid om deze modellen nauwkeuriger te maken. Biomarkers kunnen belangrijke maar subtiele verschillen in tumorbiologie beter weergeven dan conventionele prognostische variabelen zoals tumorgrootte. Voor oropharynx tumoren (OPSCC) is aan- of afwezigheid van HPV de belangrijkste biomarker. Derhalve werd een aparte stadiëring voor deze tumoren opgenomen in TNM 8<sup>2</sup>, waarbij gebruik gemaakt werd van een surrogaatmarker voor HPV (p16 immunohistochemie). HPV zou ook relevant kunnen zijn voor tumoren buiten de oropharynx, maar de standaardtest die gebruikt wordt voor oropharynx tumoren is onvoldoende betrouwbaar voor deze subsites<sup>3,4</sup> en een nauwkeurig alternatief ontbreekt. Hierdoor worden studies met betrekking tot de rol van HPV buiten de oropharynx bemoeilijkt.

Het doel van dit onderzoek was de identificatie en evaluatie van nieuwe prognostische biomarkers voor HNSCC in het algemeen of voor HNSCC op specifieke tumorlocaties. Er is een betrouwbaar HPV-teststelsel voor tumoren in de oropharynx, maar helaas niet voor tumoren in ander subsites terwijl het wel van groot prognostisch belang zou kunnen zijn. In het onderzoek dat wordt beschreven in **Hoofdstuk 2** werd daarom een moleculair HPV-teststelsel (HPV-rTcore assay) ontwikkeld dat toegepast kan worden op gearchiveerd materiaal van niet-oropharyngeale HNSCC. Tumormateriaal werd verzameld door stansbipten te nemen uit tumorrijke regio's, en vervolgens werd DNA en RNA uit deze bipten geëxtraheerd. Voorts werden PCR-technieken gebruikt voor detectie van HPV-DNA van 15 hoog-risico HPV typen. Een positief resultaat werd gevalideerd door middel van E6- mRNA-detectie. Dit testalgoritme werd gevalideerd in oropharynx tumoren (OPSCC) en mondholte tumoren (OSCC) met een nauwkeurigheid van 97% in OPSCC en 100% in OSCC. De HPV-rTcore assay wordt momenteel gebruikt om de prevalentie te bepalen van HPV in niet-oropharyngeale HNSCC, en zou de nieuwe standaard kunnen worden voor HPV diagnostiek in OPSCC. In dit proefschrift is de test gebruikt om HPV te detecteren daar waar nodig.

In **Hoofdstuk 3** wordt een prognostisch model gepresenteerd wat is gebaseerd op genexpressieprofielen. Allereerst wordt gebruik gemaakt van microarraydata van OSCC om een genexpressieprofiel te selecteren van 22 genen die voorspellend zijn voor lymfekliermetastasering (N-stadium) en een profiel van 40 genen die voorspellend zijn voor de prognose van patiënt. Vervolgens werden deze profielen gevalideerd in een onafhankelijk OSCC-cohort, waarbij gebruik gemaakt werd van kwantitatieve PCR om de mate van expressie te bepalen. Het N-stadium profiel kon de aanwezigheid van occulte lymfekliermetastasen voorspellen met een negatieve voorspellende waarde van 84% in patiënten met een cT1-2N0 tumor. Daarnaast werd een nauwkeurig prognostisch model gevonden door het prognostische genexpressieprofiel te combineren met klinische en pathologische variabelen. Dit model presteerde beter dan het TNM-stadium. Tot slot werd een subpopulatie geïdentificeerd met een slechte prognose, ondanks ontbreken van bekende negatieve prognostische kenmerken. Deze resultaten kunnen gebruikt worden voor het beter informeren van patiënten over hun prognose en voor behandeling op maat: selectie van patiënten waarbij geen halsklierdissectie vereist is of waarbij adjuvante therapie meerwaarde kan hebben. Hierbij moeten we wel opmerken dat de gouden standaard, histopathologisch onderzoek van het chirurgisch preparaat, eveneens een zeer sterke voorspellende waarde bleek te hebben. Biomarkeranalyses zijn dus belangrijker voor tumoren die niet-chirurgisch behandeld worden.

In **Hoofdstuk 4** is voor een andere benadering gekozen voor de ontwikkeling van prognostische modellen voor HNSCC: biomarkers op basis van beeldvorming. Er wordt gebruik gemaakt van blanco T1-gewogen

MRI-scans van patiënten met OSCC en HPV-negatieve OPSCC om kwantitatieve eigenschappen af te leiden van de handmatig ingetekende tumoren. Deze methode wordt ook wel 'radiomics' genoemd. In deze studie werden 545 eigenschappen afgeleid die de signaalintensiteit beschrijven van de tumor, alsmede de vorm en textuur. Overbodige eigenschappen werden verwijderd, gevolgd door een factoranalyse met de overgebleven kerneigenschappen. Deze werden vervolgens gebruikt voor trainen en valideren van prognostische modellen. De modellen gebaseerd op MRI-radiomics bevatten additionele prognostische informatie. De meest nauwkeurige modellen werden gevonden door combinaties te maken van klinische variabelen en MRI-radiomics. Er was een grote variatie in MRI-fabrikanten en acquisitieprotocollen in deze studie, maar dit hinderde het vervaardigen van de prognostische modellen niet. Dat laat onverlet dat standaardisatie van apparatuur en protocollen, en het gebruik van meer data uit de MRI-beelden, de prognostische modellen kan verbeteren. De vervaardigde modellen kunnen makkelijk geïntegreerd worden met de standaard diagnostiek aangezien blanco T1-gewogen MRI al gebruikt wordt in de meeste protocollen.

Recidief HNSCC kan vaak chirurgisch nog behandeld worden, mits de recidief-tumor ontdekt wordt in een vroeg stadium. Dit is helaas vaak niet het geval. Recidief-tumoren zijn lastig te ontdekken en de gouden standaard, onderzoek in narcose met bioptafname van verdachte afwijkingen, is invasief en daarmee niet geschikt voor routinematige screening. Vroegdiagnostiek zou wellicht mogelijk gemaakt kunnen worden door detectie van circulerend tumor DNA (ctDNA) wat wordt verspreid door apoptotische tumorcellen. In **Hoofdstuk 5** behandelt een detectiemethode voor ctDNA in HNSCC, gericht op het opsporen van somatische mutaties, chromosomale afwijkingen (kopie nummer aberratie (CNA)) en HPV-DNA. De hypothese was dat de sensitiviteit van ctDNA-detectie toe zou nemen door gebruik te maken van een gecombineerde analyse van verschillende moleculaire veranderingen vanwege de genetische heterogeniteit tussen hoofd-halstumoren. Somatische mutaties werden gedetecteerd in plasma van 67% van de patiënten, terwijl CNAs aanwezig waren in het plasma van 52%. HPV-DNA werd gedetecteerd in 100% van de patiënten met een HPV-positieve tumor. Een gecombineerde analyse verhoogde inderdaad het percentage van gedetecteerd tumor-DNA naar 78%. Kennis van de moleculaire veranderingen in de primaire tumor verbeterde de detectie van somatische mutaties in plasma, maar detectie van CNAs en HPV-DNA werd hierdoor niet beïnvloed. De positieve resultaten uit deze studie bieden ruimte voor verdere validatie door longitudinale screening in een prospectief cohort.

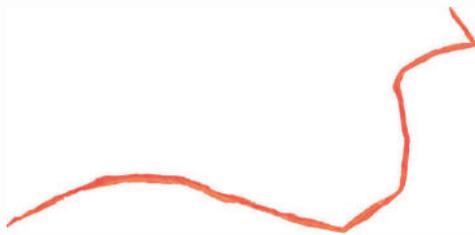
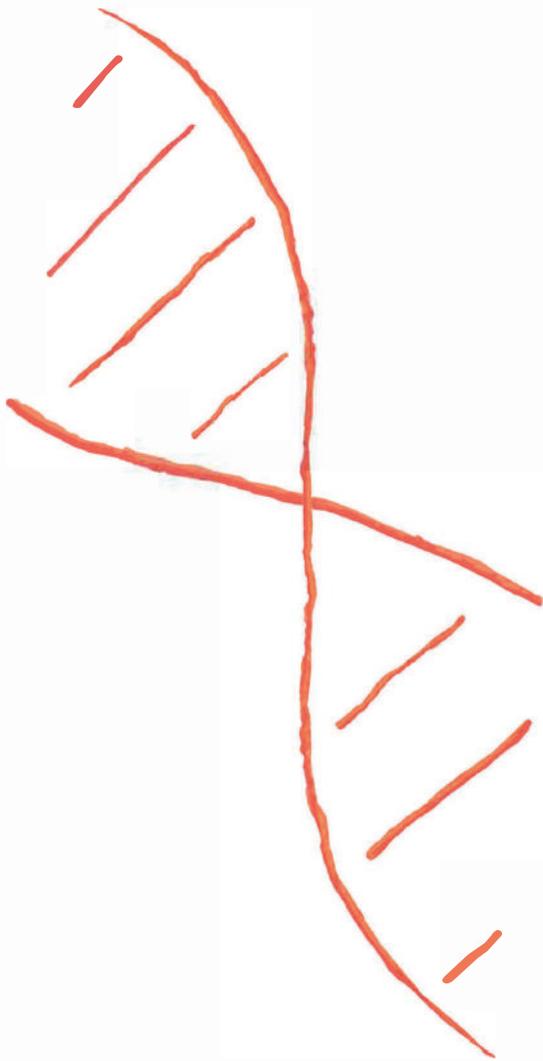
Hoewel het mogelijk lijkt om prognostische modellen te ontwikkelen op basis van biomarkers, lijken de prestaties van de verschillende modellen een suboptimaal maximum te bereiken. Dit impliceert dat er wellicht een biologische factor is waardoor de applicatie van tumorbiomarkers gelimiteerd wordt. In **Hoofdstuk 6** wordt gehypothetiseerd dat dit gerelateerd kan zijn aan intratumorale genetische heterogeniteit die gedetecteerd kan worden op moment van diagnose of naar voren komt door de behandeling en daardoor gereflecteerd wordt door genetische verschillen tussen de primaire tumor en het lokaal recidief. In deze studie werden inderdaad genetische verschillen gevonden tussen bipten uit verschillende regio's van de tumor, maar de CNA-profielen leken sterk op elkaar. Analyses van somatische mutaties zijn wellicht superieur om intratumorale heterogeniteit aantonen, en dit wordt momenteel onderzocht in een vervolgstudie. Daarnaast werd in deze studie heterogeniteit meer uitgesproken door het vergelijken van de primaire tumor en het lokaal recidief na chemoradiatie. De helft van de recidief-tumoren na chemoradiotherapie leken genetisch niet gerelateerd aan de primaire tumor, wat impliceert dat recidief-tumoren waarschijnlijk ontstaan uit zeer kleine subklonen in de primaire tumor die niet worden gedetecteerd door analyse van bulk-DNA, maar die chemoradiatieresistent zijn. Deze bevinding impliceert ook dat de prestaties van prognostische modellen met biomarkers gebaseerd op bulk-DNA/mRNA tegen beperkingen aanlopen, aangezien kleine subklonen bepalend kunnen zijn voor het optreden van recidief en derhalve de prognose. Toekomstige studies dienen rekening te houden met deze bepalende factor.

In **Hoofdstuk 7** worden gepresenteerde data bediscussieerd en in een breder perspectief gezet.

## REFERENCES

1. Chow LQM. Head and Neck Cancer. *N Engl J Med* 2020; 382(1): 60-72.
2. Brierley JDG, M. K.; Wittekind, C., eds. *TNM Classification of Malignant Tumours*. 8th Edition ed. Hoboken: Wiley-Blackwell; 2016.
3. Lingen MW, Xiao W, Schmitt A, et al. Low etiologic fraction for high-risk human papillomavirus in oral cavity squamous cell carcinomas. *Oral Oncol* 2013; 49(1): 1-8.
4. Chung CH, Zhang Q, Kong CS, et al. p16 protein expression and human papillomavirus status as prognostic biomarkers of nonoropharyngeal head and neck squamous cell carcinoma. *J Clin Oncol* 2014; 32(35): 3930-8.





**10**

**Addendum**

## DANKWOORD

Dit onderzoek wordt gedragen door de buitengewone inzet van een enorm gevarieerd team. Allen die betrokken zijn geweest bij de totstandkoming van dit proefschrift wil ik daarom hartelijk bedanken! Mijn dank is bijzonder groot voor de patiënten die tijdens een moeilijke periode ervoor gekozen hebben om deel te nemen aan wetenschappelijk onderzoek en zo bij te dragen aan verbetering van de kwaliteit van zorg in de toekomst. Tevens wil ik alle co-auteurs danken voor hun bijdrage aan de verschillende hoofdstukken.

Daarnaast wil ik een aantal mensen persoonlijk bedanken:

Prof. dr. R.H. Brakenhoff, beste Ruud, dank voor het vertrouwen in mij om dit onderzoek tot een goed einde te brengen. Ik ben je eeuwig dankbaar dat ik via de achterdeur alsnog mocht solliciteren op mijn droomfunctie als onderzoeker en toekomstig arts-assistent KNO. Aan het begin van mijn promotietraject hadden we een aantal hoofdstukken voor ogen, maar je hebt me ook de kans gegeven aanvullende onderzoeken op te starten waardoor ik met veel plezier mijn interesse kon ontwikkelen in onder meer niet-invasieve tumordetectie en intratumorale heterogeniteit. In het kader van het Europees onderzoek zijn we frequent aanwezig geweest op vergaderingen in het buitenland waarbij we stevige discussies gevoerd hebben met onze collega-onderzoekers, maar waarna ook altijd ruimte was voor een plaatselijke speciaalbierproeverij. Je kennis en drive zijn altijd inspirerend geweest, alsmede je vermogen om vruchtbare samenwerkingen op te zetten en te onderhouden.

Prof. dr. C.R. Leemans, beste René, ik wil u danken voor de mogelijkheid dit promotietraject te doorlopen en de specialisatie tot KNO-arts te volgen. Uw overtuigende woorden aan de telefoon zal ik nooit vergeten en hebben mij uiteindelijk over de streep getrokken om in Amsterdam aan de slag te gaan in plaats van de stad van uw grote voetballiefde. Tevens dank voor uw kritische, klinische blik; zowel op de artikelen als mijn klinische werkzaamheden.

Dr. P. de Graaf, beste Pim, dank voor alle begeleiding in het radiologische werk binnen dit proefschrift en daarbuiten. De vele uren die jij vrijgemaakt hebt ondanks je drukke klinische activiteiten waardeer ik enorm. Jouw enthousiasme was aanstekelijk en hielp mij altijd weer verder, ook als ik even pixelmoe was van alle scans. Daarnaast hebben we ook altijd gezellige tijden gehad tijdens de OraMod-vergaderingen en met name daarbuiten. Je bent voor mij echt een rolmodel als arts die hoogkwalitatief klinisch werk weet te combineren met gevarieerd wetenschappelijk onderzoek.

Prof. dr. E. Bloemena, beste Elisabeth, ik ben je erg dankbaar voor de (her)beoordeling van alle histopathologie en de vele nuttige aanvullingen op de manuscripten. Je maakte altijd tijd vrij voor mij en mijn stapels met coupes en hebt me bovendien erg veel geleerd over de pathologie van het hoofd-hals plaveiselcelcarcinoom.

Dr. E.A. Siermans, beste Erik, ik wil je bedanken voor de hulp in het opzetten van de ctDNA-studie en de verdere begeleiding hierin. Wat begon met een afwijking in de NIPT-data van één van jullie patiënten, is uiteindelijk geëindigd in een mooie studie en een vervolgbeurs samen met Ruud. Ik houd alle vertrouwen in de verdere toekomst hiervan en hoop dat de hoofd-halskankerpatiënten er veel baat bij mogen hebben.

De overige leden van mijn leescommissie, prof. dr. R.J. Baatenburg de Jong, prof. dr. M.W.M. van den Brekel, prof. dr. P. Lambin, prof. dr. E.M.D. Schuurin, wil ik hartelijk danken voor de bereidheid om dit manuscript te beoordelen en voor het plaatsnemen in de promotiecommissie.

Dan mijn collega's van de Tumorbiologie. Om te beginnen wil ik de rol van Arjen uitlichten. Jij hebt mij vanaf het begin enorm geholpen met het praktische werk en later ook steeds meer met de bio-informatica, een rol die jij jezelf echt eigen hebt gemaakt de laatste jaren. In bijna ieder hoofdstuk was voor jou een belangrijke rol weggelegd. Daarnaast heb ik je leren kennen als een fijn mens met een onnavolgbare precisie. Bij eventueel toekomstig onderzoek hoop ik echt op een nieuwe Arjen.

Marijke B., jou heb ik leren kennen als zeer kundig in het lab wat ik erg waardeerde, maar meer nog ben ik erg gesteld op jouw sprankelende persoonlijkheid. Het was altijd even lachen om bij je binnen te vallen op je kamer of in het (kweek)lab.

Marijke S., van jouw jarenlange ervaring heb ik veel geleerd met name met betrekking tot het histologische werk. Dank hiervoor. Daarnaast kon ik altijd genieten van je verhalen over het begin van de Tumorbiolegroep en je escapades met de proefdieren.

Anne, ik ben altijd enorm onder de indruk geweest van je kennis en je vermogen je ergens helemaal in vast te bijten. Jij bezit alle eigenschappen van een goede onderzoeker en ik verwacht dan ook dat je hierin een mooie carrière tegemoet ziet.

Vicky, jouw onderzoek was voor mij altijd erg tastbaar en ik ben blij voor je dat je het hebt kunnen afronden en je promotie binnenkort gepland staat. Daarnaast vind ik je een warm persoon met een fijn gevoel voor humor. Samen met Marijke was je altijd een perfect team!

Jos, jouw komst heeft een nieuwe boost gegeven aan het lab en ik ben erg dankbaar voor je bijdrage aan hoofdstuk 5 en 6. Daarnaast heb ik erg genoten van je droge humor. Veel succes in je verdere carrière die hopelijk nog lang vervolgd mag worden bij de KNO-afdeling.

Boudewijn, dank voor jouw grote hulp bij het mij wegwijs maken binnen het hoofd-hals oncologisch onderzoek. Jij had altijd tijd voor uitleg en om samen te sparren over nieuwe ideeën. Daarnaast vond ik onze microdissectiesessies fantastisch. Hopelijk heb je gevonden wat je zocht na je pensioen!

Sanne, toen ik begon, mocht ik vrijwel direct getuige zijn van jouw promotie. De toon was gezet! Daarna als postdoc heb ik je leren kennen als een eigenzinnige onderzoeker en gevat persoon.

Costa, ik heb van je intelligente opmerkingen genoten tijdens de werkbijeenkomsten en daarbuiten. Je hebt je enorm moeten ontwikkelen om jouw project tot een goed einde te brengen en hier heb ik veel respect voor!

Sonja, Richard, Ana en Jantine, jullie immunologisch onderzoek heb ik altijd fascinerend gevonden. Dank voor het bijspijkeren van mijn kennis en voor alle gezelligheid!

Een speciale vermelding gaat uit naar de collega's van de biostatistiek: dr. Dennis te Beest, dr. Wessel van Wieringen, dr. Carel Peeters en prof. dr. Mark van de Wiel. Zonder jullie hulp had de totstandkoming van dit proefschrift geen schijn van kans. Dennis, jij bent tijdens de OraMod-tijd degene geweest met wie ik het meeste heb samengewerkt en het voelde echt als een team als we samen (als botte Nederlanders) in discussie gingen met de overige leden van het consortium. Carel, dank voor jouw frisse blik op de radiomicsdata. Jij hebt deze studie echt gered. Mark, ik heb veel bewondering voor jouw vermogen om je in te leven in klinische vraagstukken en je creatieve oplossingen. Daarnaast zal ik je relativeringsvermogen binnen OraMod ook nooit vergeten.

My dear OraMod friends, working in an international consortium with people from different scientific backgrounds has been very special to me. Thank you all for the opportunity to learn from you and for the introduction to your cultures. Hope to meet you again in science or elsewhere.

Beste mensen van de moleculaire pathologie, dr. Daniëlle Heideman, Martijn Bogaarts en Marije Doeleman, ik wil jullie van harte danken voor jullie hulp bij de ontwikkeling van de HPV test en de analyses hiervan. Jullie jarenlange ervaring op dit gebied heeft mij veel gebracht!

De radiomicsstudie zou nooit in deze vorm tot stand zijn gekomen zonder de uiterst deskundige hulp van dr. Floris van Velden en prof. dr. Ronald Boellaard. Dank voor al jullie werk. Specifiek wil ik Floris uitlichten die een berg werk verzet heeft om de radiomicssoftware geschikt te maken voor onze MRI's en VOI's. Daarnaast ben je ook altijd erg betrokken geweest bij de vooruitgang van het project wat mij erg gemotiveerd heeft. Hopelijk kunnen we in de toekomst nog eens een nieuw project starten, ik heb wel een aantal ideeën hiervoor!

Daarnaast wil ik de heren danken van het UMC Utrecht voor hun bijdrage aan deze studie. Uit ervaring weet ik hoeveel tijd gaat zitten in het intekenen en controleren van iedere MRI-scan... Dus dr. Boris Peltenburg, Furkan Yaz, dr. Frank Pameier en prof. dr. Remco de Bree veel dank voor jullie werk!

Ditzelfde geldt uiteraard voor Roland Martens en prof. dr. Jonas Castelijns. Ook jullie ben ik erg dankbaar voor het intekenen van de MRI-scans die we gebruikten ter bevestiging van de interobserver agreement. Roland, succes met de laatste loodjes!

Naast Erik wil ik nog een aantal andere mensen noemen die meegewerkt hebben aan de ctDNA-studie en wiens bijdrage onmisbaar was. Dr. Roy Straver dank voor je analyses en aanpassingen aan het script. Prof. dr. Cees Oudejans, ik ben u erg dankbaar voor de conceptuele bijdrages. Daarnaast zou het OraMod-project zonder de TLDA-faciliteiten in uw lab een enorme uitdaging zijn geweest. Praktisch heb ik erg veel gehad aan de hulp van Allerdien Visser, hartelijk dank hiervoor.

Naast mijn werk als tutor, heb ik verschillende studenten begeleid met onderzoek en bachelorscripties. Cynthia, Joost, Joran, Leonard en Pieter, bedankt voor jullie input en de leermomenten die jullie mij geboden hebben.

Mijn kamer heb ik tijdens mijn onderzoeksperiode gedeeld met Inge Braspenning, Femke Jansen, Laura Korsten Maarten van Loon, Irene Nauta, Annette van Nieuwenhuizen, Reinout de Roest, en Charlotte Schouten. Veel dank voor jullie gezelligheid, de fijne koffie- en lunchmomenten, en de wetenschappelijke inzichten die jullie deelden. Laura, waarmee ik gezamenlijk ben gestart, wat heb ik altijd respect gehad voor de manier waarop je een enorm multicenter project coördineerde. Reinout, jij kwam het artsenteam van de Tumorbioogie gelukkig weer versterken. Ik heb het altijd een mooie tijd met je gevonden en ben erg trots op onze gezamenlijke studie. Irene, fantastisch hoe jij verdergegaan bent met de HPV-assay en hoe jij je hebt ontwikkeld tot internationale HPV-goeroe. Voor de mensen die nog bezig zijn met hun PhD: succes allemaal met de laatste loodjes! Sandra Biemans, dank voor al je hulp en voor de relativerende koffiemomenten waardoor ik altijd weer even landde.

De zaadjes voor mijn interesse in wetenschappelijk onderzoek zijn geplant in Rotterdam op twee afdelingen. Prof. dr. Jon Laman heeft mij destijds overtuigd om mijn wetenschappelijke stage in te vullen met basaal onderzoek naar multipele sclerose. Jouw aanstekelijk enthousiasme, oprechte interesse en scherpe analyses heb ik altijd bewonderd! Ook wil ik graag de overige groepsleden danken: dr. Marvin van Luijn, dr. Karim Kreft, Annet Wierenga-Wolf, Marie-José Melief en Marjan van Meurs.

Toen ik vervolgens mijn keuze gemaakt heb voor de KNO-Heelkunde ben ik klinisch onderzoek gaan doen onder begeleiding van dr. Emilie Dronkers, dr. Marc van der Schroeff en prof. dr. Rob Baatenburg de Jong. Deze periode, gecombineerd met mijn tijd als oudste co-assistent op jullie afdeling, heeft me overtuigd om te gaan voor een promotietraject binnen de hoofd-halschirurgie en de opleiding tot KNO-arts. Ik heb me altijd erg welkom gevoeld en onderdeel van het team. Bedankt!

Graag wil ik alle hoofd-halschirurgen danken voor hun interesse en patiëntinclusie. Dr. Simone Eerenstein, dr. Jan-Jaap Hendrickx, Hakkı Karagozoglu, Stijn van Weert en de verschillende fellows, zonder jullie was dit onderzoek nooit mogelijk geweest.

Graag wil ik de staf danken van de afdeling KNO van het Amsterdam UMC, locatie VUmc en AMC, voor hun vertrouwen in mij en de fantastische opleiding die ik mag genieten. Uiteraard mogen ook de KNO-artsen van het Diaconessenhuis Utrecht en Dijklander ziekenhuis, locatie Hoorn, niet ontbreken. Bij jullie heb ik een grote stap gezet in mijn chirurgische ontwikkeling waarvoor ik jullie erg dankbaar ben, de opleiders dr. Jasper Quak en dr. Loet Bauwens in het bijzonder. Daarnaast heb ik een hele fijne samenwerking gekend met alle medewerkers van de poliklinieken en operatiecomplexen op de verschillende locaties, de medewerkers van het audiologisch centrum, de verpleegkundigen van 1C, Trudi Limpens, Hanneke Tielens, Ton Houffelaar,

Jacqueline Geskus, Esther de Koster en de dames van het secretariaat op de VU: Boukje, Gerrie, Marjon en Vanessa.

Collegae AIOS! De opleiding tot KNO-arts is echt een prachtige tijd en dit komt voor een belangrijk deel doordat je alles kunt delen met een hechte groep collega's. De ski- en assistentenweekenden waren onvergetelijk. Het is een beetje alsof je een tweede studententijd krijgt soms, fantastisch! Het wordt hoog tijd om weer eens wat te plannen, en ik hoop dat dit allemaal snel weer kan.

Een groot deel van mijn studententijd heb ik doorgebracht in Delft, een briljante periode van mijn leven. VS42-ers, CGS-ers, Dopie M-ers en commissiegenoten dank voor de vele uurtjes aan de bar, weekendjes en ontelbare levenslessen. Nu dit werk af is hoog tijd om weer wat meer tijd in jullie te investeren, dus tot snel!

Dan een woord aan mijn paranimfen, Jeroen Kraak en Thadé Goderie.

Beste Jeroen, vanaf het eerste uur gezamenlijk op de onderzoekskamer heb ik het enorm goed met je kunnen vinden. Later toen ik begon met de opleiding heb ik me altijd aan je opgetrokken en gezien als voorbeeld. Ik bewonder je enthousiasme en enorme inzet heel erg. De cursussen (met name de tijd daarbuiten) en skiweekenden waren fantastisch en de zondagavond in Groningen onvergetelijk. Dat jij je droom kunt naleven als toekomstig hoofd-halschirurg in Amsterdam UMC gun ik je van harte! Heel knap dat je naast je fulltimebaan een gezin hebt weten op te bouwen en ook nog een promotieonderzoek hebt opgepakt wat zich inmiddels alweer in een vergevorderd stadium bevindt.

Beste Thadé, wij hebben een andere gezamenlijke achtergrond waarin ik je heb leren kennen als zwager ruim voordat de KNO als toekomstig vakgebied voor mij überhaupt speelde. Jij bent zeker één van de mensen die mij heeft geïnspireerd om de KNO te verkennen en ik ben je veel dank verschuldigd voor jouw ingang tot de sollicitaties van de VU. Daarnaast ben ik je erg dankbaar voor je coachende rol als opleider waarin je ook aanmoedigt om verder te kijken richting de toekomst. Ook naast het werk heb ik fijne tijden met je beleefd. Het skiën voorafgaand aan het KNO-skiweekend vond ik briljant en ben pas in de week daarna weer hersteld van de schnapsroute geloof ik. Ik hoop nog vaak te mogen genieten van dergelijke momenten met jou, Michelle en de kinderen.

Lieve Pieter en Tanja, dank voor alle support en het liefdevol opnemen van mij in jullie familie. Ik vind het altijd enorm fijn om bij jullie langs te komen en geniet van de fijne gesprekken. Pieter, veel dank voor al je werk aan de kaft van dit manuscript. Het is prachtig geworden! Steven en Irena, ik kan me geen fijnere schoonfamilie wensen en ben blij dat de jaarlijkse samenkomst met oud en nieuw een traditie geworden is. De volgende mijlpaal is hopelijk voor jullie in 2022!

Lieve Jolijn en Niek, jullie zijn een geweldig paar en ik heb het altijd erg naar mijn zin met jullie. Jolijn, jij bent de eerste van de familie met eigen drukwerk in de kast waardoor ik natuurlijk wel moest volgen. Dank voor je hulp met de Nederlandse teksten! Ik ben trots op jullie en hoop dat er nog vele avondjes mogen volgen.

Lieve pap en mam, jullie zijn altijd heel erg ondersteunend geweest in alles wat ik heb gedaan waarvoor ik jullie oneindig dankbaar ben. Vooral aan het einde van mijn middelbare school en de eerste jaren in Delft ben ik niet de makkelijkste geweest, maar desondanks bleven jullie altijd liefdevol en stuurden jullie me ongemerkt in de goede richting. Ik heb altijd geluk gehad dat ik me mocht ontwikkelen in elke gewenste richting en ik besef dat dit een bijzonder voorrecht is. Ondanks dat ik sinds mijn 17<sup>e</sup> niet meer thuis woon, voelt jullie huis altijd als thuishomen. Dit geldt ook voor Sophia, jullie hebben haar echt deel laten voelen van de familie. Pap, je bent een sportief en sociaal persoon en ik bewonder je oprechte interesse in iedereen om je heen enorm. Mam, je bent bijzonder liefdevol en attent en jouw vermogen om iedereen bij elkaar te brengen en zich welkom te laten voelen vind ik erg bijzonder.

Tot slot, lieve Sophia. Ik ben enorm dankbaar om jou al vele jaren in mijn leven te hebben. Vanaf het begin voelt het echt alsof het altijd zo had moeten zijn. Ik kan volledig mijzelf zijn met je en voel me samen met

jou precies op mijn plek. Jij bent degene die mij heeft aangemoedigd om te kiezen voor de KNO en voor dit promotietraject en jouw discipline en doorzettingsvermogen inspireerde mij om me ook zo op te stellen ten aanzien van het onderzoek. Ik kan oprecht zeggen dat zonder jou dit werk nooit afgemaakt was in deze vorm en waarschijnlijk überhaupt nooit door mij geïnitieerd. Naast het werk moet je af en toe je stoom kwijt kunnen en ook die uitlaatklep ben jij voor mij. Zowel voor de haard in ons heerlijke huis in Haarlem als samen op het strand met onze hond. Heerlijk ook hoe ik met jou heb mogen ontdekken dat de bergen ook in de zomer prachtig zijn. Ik ben trots op de stappen die jij hebt gezet in je eigen carrière en kijk enorm uit naar de toekomst met jou!

## CURRICULUM VITAE

Steven Mes was born on December 8, 1987 in Amersfoort, the Netherlands. After primary school, he attended the Odulphus Lyceum in Tilburg, where he obtained his gymnasium diploma, cum laude, in 2005. In the same year he started at the Delft University of Technology to study Marine Technology. The next year Steven went to Rotterdam to study medicine at the Erasmus Medical Center. During his medical training, he developed a specific interest in translational research while studying immunology and cell biology of multiple sclerosis in the lab of prof. dr. J.D. Laman. Moreover, enthused by the otolaryngologists at the Erasmus MC, Steven also became passionate about otolaryngology and studied noncompliance to guidelines in head and neck cancer treatment under supervision of dr. Marc van der Schroeff and dr. Emilie Dronkers.



After obtaining his medical degree, cum laude, in 2014, Steven got the opportunity to combine his passions for translational research and otolaryngology at the department of Otolaryngology and Head and Neck Surgery of the Amsterdam UMC, location VUmc, and studied prognostic markers in head and neck cancer, which resulted in this thesis. His project was part of a European study (OraMod), which enabled him to collaborate with and learn from researchers with diverse backgrounds. Also during this period, he presented his work at national and international conferences on several occasions.

In 2017 he commenced the otorhinolaryngology specialty training program at the Amsterdam UMC, location VUmc, under supervision of prof. dr. C.R. Leemans and prof. dr. P. Merkus consecutively. Part of his training program took place at the Diaconessenhuis in Utrecht under supervision of dr. J.J. Quak and at the Dijklander Ziekenhuis in Hoorn under supervision of dr. L.J.J.M. Bauwens. During his residency, Steven chose to differentiate in otology and he became a member of the Otology Workgroup of Dutch Society of Otolaryngology/Head-Neck Surgery. Steven lives in Haarlem together with his girlfriend Sophia Goderie.

## LIST OF PUBLICATIONS

### **Comprehensive multiparameter genetic analysis improves circulating tumor DNA detection in head and neck cancer patients**

**Mes SW**, Brink A, Sisternans EA, Straver R, Oudejans CBM, Poell JB, Leemans CR, Brakenhoff RH

Oral Oncology 2020 Jun 23;109:104852

### **Outcome prediction of head and neck squamous cell carcinoma by MRI radiomic signatures**

**Mes SW**, van Velden FHP, Peltenburg B, Peeters CFW, te Beest DE, van de Wiel MA, Mekke J, Mulder DC, Martens RM, Castelijns JA, Pameijer FA, de Bree R, Boellaard R, Leemans CR, Brakenhoff RH, de Graaf P

European Radiology 2020 Nov;30(11):6311-6321

### **Stable prediction with radiomics data**

Peeters CFW, Übelhör C, **Mes SW**, Martens RM, Koopman T, de Graaf P, van Velden FHP, Boellaard R, Castelijns JA, te Beest D, Heymans MW, van de Wiel MA

arXiv:1903.11696 [stat.ML], 2019

### **Development and Validation of a Novel and Rapid Molecular Detection Method for High-Risk Human Papillomavirus in Formalin-Fixed, Paraffin-Embedded Tumor Tissue**

**Mes SW**, Heideman DAM, Bloemena E, Brink A, Bogaarts M, Leemans CR, Brakenhoff RH

The Journal of Molecular Diagnostics 2020 Feb;22(2):262-271

### **Molecular Characterization of Locally Relapsed Head and Neck Cancer after Concomitant Chemoradiotherapy**

De Roest RH, **Mes SW**, Poell JB, Brink A, van de Wiel MA, Bloemena E, Thai E, Poli T, Leemans CR, Brakenhoff RH

Clinical Cancer Research 2019 Dec 1;25(23):7256-7265

### **Improved high-dimensional prediction with Random Forests by the use of co-data**

Te Beest DE, **Mes SW**, Wilting SM, Brakenhoff RH, van de Wiel MA

BMC Bioinformatics 2017 Dec 28;18(1):584

### **Prognostic modeling of oral cancer by gene profiles and clinicopathological co-variables**

**Mes SW**, Te Beest D, Poli T, Rossi S, Scheckenbach K, van Wieringen WN, Brink A, Bertani N, Lanfranco D, Silini EM, van Diest PJ, Bloemena E, Leemans CR, van de Wiel MA, Brakenhoff RH

Oncotarget 2017 Jul 26;8(35):59312-59323

### **Applications of molecular diagnostics for personalized treatment of head and neck cancer: state of the art**

**Mes SW**, Leemans CR, Brakenhoff RH.

Expert Review of Molecular Diagnostics 2016;16(2):205-21

**Noncompliance to guidelines in head and neck cancer treatment; associated factors for both patient and physician**

Dronkers EA, **Mes SW**, Wieringa MH, van der Schroeff MP, Baatenburg de Jong RJ  
BMC Cancer 2015 Jul 11;15:515

**Multiple sclerosis-associated CLEC16A controls HLA class II expression via late endosome biogenesis**

Van Luijn MM, Kreft KL, Jongsma ML, **Mes SW**, Wierenga-Wolf AF, van Meurs M, Melief MJ, van der Kant R, Janssen L, Janssen H, Tan R, Priatel JJ, Neeffjes J, Laman JD, Hintzen RQ  
Brain 2015 Jun;138(Pt 6):1531-47

**Severe intracranial complications during FESS; endoscopic reconstruction with cartilage from the cavum conchae**

**Mes SW**, Kropveld A

Het Nederlands Tijdschrift voor Keel- Neus- Oorheelkunde 2014 Apr; 20 (2): 54-59



