

Head and neck cancer transcriptomics

Profiling tumor biology for
chemoradiotherapy response



Martijn van der Heijden

HEAD AND NECK CANCER TRANSCRIPTOMICS

Profiling tumor biology
for chemoradiotherapy response

Martijn van der Heijden

Colofon

Head and neck cancer transcriptomics. Profiling tumor biology for chemoradiotherapy response

Copyright © **2023** Martijn van der Heijden, Amsterdam, the Netherlands.

All rights reserved. No part of this publication may be reproduced, stored in a retrieval system, or transmitted in any form or by any means, electronic, mechanical, by photocopying, recording, or otherwise, without the prior written permission of the author.

Cover: Jaxon Nash by courtesy of Dead End Gallery Amsterdam

Lay-out: Publiss | www.publiss.nl

Printing: Ridderprint | www.ridderprint.nl

The publication of this thesis was financially supported by the Netherlands Cancer Institute/ Antoni van Leeuwenhoek ziekenhuis, Academisch Centrum Tandheelkunde Amsterdam (ACTA), Stichting Michel Keijzer Fonds, Merz medical, Chipsoft, BAP Medical, Daleco Pharma, ABN Amro, Allergy Therapeutics.

Head and neck cancer transcriptomics

Profiling tumor biology for chemoradiotherapy response

ACADEMISCH PROEFSCHRIFT

ter verkrijging van de graad van doctor
aan de Universiteit van Amsterdam
op gezag van de Rector Magnificus
prof. dr. ir. P.P.C.C. Verbeek

ten overstaan van een door het College voor Promoties ingestelde commissie,
in het openbaar te verdedigen in de Agnietenkapel
op dinsdag 24 oktober 2023, te 16.00 uur

door

Martijn van der Heijden
geboren te Nijmegen

Promotiecommissie

<i>Promotoren:</i>	prof. dr. M.W.M. van den Brekel	Universiteit van Amsterdam
	prof. dr. R.H. Brakenhoff	Vrije universiteit Amsterdam
<i>Copromotoren:</i>	dr. C. Vens	NKI-AvL
	dr. P.B.M. Essers	Twentynext BV
<i>Overige leden:</i>	prof. dr. L.E. Smeele	Universiteit van Amsterdam
	prof. dr. M.J. van de Vijver	Universiteit van Amsterdam
	dr. M.F. Bijlsma	Universiteit van Amsterdam
	prof. dr. C.L. Zuur	Universiteit Leiden
	prof. dr. J. Bussink	Radboud Universiteit
	prof. dr. R.P. Takes	Radboud Universiteit

Faculteit der Tandheelkunde

Table of contents

Chapter 1	Introduction	7
Chapter 2	Acute hypoxia profile is a stronger prognostic factor than chronic hypoxia in advanced stage head and neck cancer patients	45
Chapter 3	Epithelial-to-mesenchymal transition is a prognostic marker for patient outcome in advanced stage HNSCC patients treated with chemoradiotherapy	77
Chapter 4	Drug Sensitivity Prediction Models Reveal a Link between DNA Repair Defects and Poor Prognosis in HNSCC Predicting DNA Repair Defects in HNSCC	125
Chapter 5	Biological determinants of chemo-radiotherapy response in HPV-negative head and neck cancer: a multicentric external validation	179
Chapter 6	General discussion	237
Chapter 7	Summary	255
Chapter 8	Nederlandse samenvatting	261
Appendices	Authors and affiliations	268
	Author contributions	270
	Funding	271
	PhD portfolio	272
	List of publications	274
	Dankwoord	276
	Curriculum Vitae	282



CHAPTER 1

Introduction

Introduction

Cancer was the leading cause of death in the Netherlands in 2019. Head and neck cancer is, with annually ~3,000 new cases, the 9th most common type of cancer in the Netherlands. Head and neck cancer is treated with a combination of surgery, radiotherapy or chemotherapy. Although treatment outcome improved over the last decades, the prognosis of head and neck cancer is still disappointing with a 5-year overall survival of around 50%. Treating head and neck cancer is invasive and causes many side effects. A first step in addressing this problem is to identify the patients with high-risk and low-risk tumors upfront and select them for alternative, personalized, treatment options. The research presented in this thesis is intended to identify biological characteristics that have prognostic value and that may also present potential targets for therapy.

1. Head and neck cancer

Cancer is a disease that results from a normal and healthy tissue cell undergoing genetic changes which lead to malignant properties such as uncontrollable growth, indicated as one of the 'Hallmarks of Cancer'. Head and neck cancer is the all-encompassing term for cancers arising from the head and neck area, including oral cavity, oropharynx, nasal cavity, paranasal sinuses, nasopharynx, hypopharynx, larynx, skin, salivary glands and thyroid (Figure 1). Each site has its own prevalence, diagnostic work-up, treatment and prognosis. The mucosa of the subsites mostly consists of squamous epithelial cells and tumors that originated from these cells are therefore called Head and Neck Squamous Cell Carcinoma (HNSCC). The most common subsites that are affected in the Western world are the oral cavity, larynx, oropharynx and hypopharynx. In this thesis we focus on HNSCC arising from the larynx, oropharynx and hypopharynx.

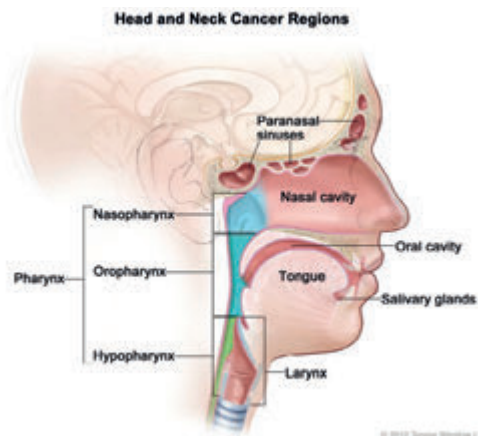


Figure 1. Head and neck cancer sites. Illustration of anatomical sites in the head and neck area. Figure adapted from the NIH National Cancer Institute website¹.

1.1 Epidemiology and etiology

Estimations based on Global cancer statistics, GLOBOCAN, indicate around 185,000 patients with newly diagnosed laryngeal, around 100,000 patients with oropharyngeal and 85,000 patients with hypopharyngeal cancer, in 2020 worldwide. This makes up for 1.9% of all cancers worldwide. The estimated deaths in 2020 were around 100,000, 50,000 and 40,000 for laryngeal, oropharyngeal and hypopharyngeal cancers, respectively². According to the Dutch Cancer Registration (IKNL), there were approximately 3,000 new cases of mucosal head and neck cancer diagnosed in 2020. Of these, around 900 patients presented with oro- or hypopharyngeal carcinoma and around 650 with laryngeal carcinoma. Men are affected more frequently than women and the median age of diagnosis is between 60-70 years old.

Main etiological factors for head and neck cancer are alcohol consumption and smoking. Tobacco is the direct cause for most HNSCC and alcohol consumption has a synergistic effect, increasing the risk for HNSCC. For oropharyngeal and nasopharyngeal carcinoma, viral infections with human papillomavirus (HPV) and Epstein-Bar virus (EBV) respectively, are independent risk factors³. There are geographical differences in prevalence, incidence and etiology in head and neck cancer. For example, nasopharyngeal carcinoma is rare in the Netherlands and is predominantly found in South-East Asia and the Mediterranean countries. Over the last decades, there has been an increase in oropharyngeal carcinoma in the Western world, which attributes to the rising numbers of HPV related oropharyngeal carcinomas. HPV is a sexually transmitted virus which can cause squamous cell carcinomas arising in different areas of the body, including the oropharynx. There are over a 100 different types of HPV and 15 so-called high-risk types, predominantly type 16 and 18, have been associated with cancer so far. The exact mechanism of carcinogenesis will be elaborated on below in this introduction. Recent estimates suggest that near 80% of current newly diagnosed oropharyngeal carcinomas are caused by HPV infection⁴. HPV-positive HNSCC have different demographic characteristics when compared to HPV-negative HNSCC. HPV-positive HNSCC patient tend to be younger, more promiscuous and less likely to smoke.

Besides environmental factors, genetic susceptibility plays a role in the etiology in head and neck cancer. Individuals with certain germline genetic alterations have an increased risk of developing cancer, and some more specific for head and neck cancer. Roughly 5-10% of head and neck cancers have a hereditary origin⁵. Fanconi anemia (FA) patients have a 500 to 700-fold higher chance of developing head and neck cancer^{6,7}, but also have a high chance to develop acute myeloid leukemia and esophageal, gastrointestinal, vulvar and anal cancer. FA is characterized by inherited chromosomal instability, congenital malformations, bone marrow failure and cancer predisposition. This genetic susceptibility arises from alterations

in one of the genes of the FA/BRCA pathway. The FA/BRCA pathway plays a central role in repairing DNA-crosslinks. In contrast, Li-Fraumeni patients have germline mutations in the *TP53* gene and are therefore susceptible to multiple types of cancer, including HNSCC. Similarly, Ataxia-Telangiectasia and Bloom syndrome patients are characterized by a mutation in the *ATM* and *BLM* gene, respectively, and have an increased risk for developing any type of cancer. These hereditary cancer syndromes patients have an increased risk of developing cancer at a young age⁸.

1.2 Clinical presentation and diagnosis

Patients with head and neck cancer usually present themselves in the outpatient clinic with complaints such as hoarseness, globus, pain, a non-healing ulcer or a nodule in the neck. When head and neck cancer is suspected or diagnosed, they are referred to a head and neck cancer center. HNSCC care is centralized in eight centers in the Netherlands. After referral, a detailed endoscopic examination of the total mucosal lining of the head and neck area is performed. Estimation of the extent of the disease requires palpation of the tumor and regional lymph nodes. Imaging with ultrasound, CT and/or MRI of the head and neck is used to determine extension of the primary tumor into other organs/structures of the head and neck area and to determine the number and location of pathological lymph nodes. FDG-PET imaging is increasingly used to help segmenting the tumor for radiotherapy and assess metastatic dissemination. Confirmation of head and neck cancer diagnosis follows after histopathological examination of biopsy material. This is essential to distinguish malignant from benign disease and to determine the histotype and HPV status. As stated before, in this thesis, we focus on HPV-negative HNSCC.

All the information gathered during the diagnostic process is summarized in the TNM categorization system for head and neck cancer. The T-category summarizes the characteristics of the primary tumor, which differs between each tumor site. The N-category translates the extent of lymph node involvement. The M-stage classifies patients with disseminated disease at distant sites or not. Partly based on this prognostic classification system, patients are selected for different treatment protocols. The most recent, 8th edition of the TNM staging system was published in 2017. However, data collection for this thesis was completed prior to 2017 and therefore tumors are staged according to the 7th edition. The most significant difference between the two editions, is recognizing p16^{Ink4A}-positive oropharyngeal cancer, as a surrogate for HPV status, as a different disease entity with a separate TNM staging system. This leads to down-staging of up to 92% of patients with HPV-positive HNSCC, better representing their prognosis in relation to stage⁹. Secondly, depth of

infiltration was added for oral cavity cancers. Third, for all HPV-negative head and neck cancers extra-nodal extension was added for N-staging, resulting in N3 stage when present¹⁰. Based on the TNM-staging systems patients can be categorized into four clinical stages, stage I-IV (Table 1)¹¹. Stage I-II contain small primary tumors without lymph node involvement. Stage III-IV are considered advanced disease stages with either larger primary tumors (T3-T4) and/or more extensive lymph node involvement. Stage IV is further subdivided in A, B and C. A is used for large tumors, B for large tumors for which surgery is no longer feasible and C is for disseminated disease. When first diagnosed with HNSCC, 40% of patients present with stage I-II, 50% with advanced stage III-IVB and 10% with distant metastasis, stage IVC. In this thesis we focus on patients with advanced stage (III-IVB) disease.

1.3 Treatment

Early-stage disease is preferably treated with single modality being surgery or radiotherapy alone. These patients have a good prognosis with 5-year overall survival above 70-80%. However, almost two-thirds of the patients present with advanced stage disease. For advanced stage disease multimodality treatment is the standard of care, meaning a combination of surgery, radiotherapy, chemotherapy or all three. Except for T4b tumors, surgery with or without adjuvant (chemo)radiotherapy is a curable treatment option for all head and neck tumors and is still the first choice of treatment for oral cavity tumors and locally advanced (T4) laryngeal and hypopharyngeal carcinoma. Patients with oral cavity tumors respond better to surgery with adjuvant treatment than to chemoradiotherapy^{12,13}. When a large laryngeal or hypopharyngeal tumor is irradiated, this often affects function of the larynx. However, surgical resection of large tumors in the head and neck area is associated with major morbidities, such as disfigurement and problems with swallowing, speaking and breathing. Due to this morbidity, HNSCC treatment paradigm shifted from surgery towards the organ and function preservation of radiotherapy. In the Netherlands, around 70% of patients with advanced oropharyngeal, hypopharyngeal and laryngeal carcinoma are treated with radiotherapy, often combined with cisplatin-based chemotherapy¹⁴. In the Netherlands, conventional radiotherapy for HNSCC delivers a total dose of 70 Gy in daily fractions of 2 Gy for 7 weeks. Radiotherapy technology drastically improved over the last few decades. Efficacy and safety of radiotherapy depends on the opportunity to irradiate the tumor with the highest dose possible without unacceptable toxicity to the surrounding healthy tissue. The older treatment technique involved three simple fields of radiation, that limited the ability to provide conformal dose distributions and resulted in substantial normal tissue exposure and toxicity. By modulating dose distributions with multiple radiation beams in intensity-modulated radiation therapy (IMRT), the tumor is still subjected to the full radiation dose while more effectively sparing the surrounding healthy

tissue, resulting in less toxicity¹⁵. In 2007, volumetric modulated arc therapy (VMAT) was introduced into routine standard practice. The 360° rotational arc allowed delivering of the radiation beam from a continuous radiation source. Together, the improvement in technology for delivering radiotherapy, and the reduced toxicity associated with it, makes radiotherapy indispensable in the treatment for head and neck cancer. Also the boosting techniques on the gross tumor volume have been improved with the better results for simultaneous integrated boosting¹⁶.

Accelerated schemes, adaptive radiotherapy and proton beam radiotherapy are more recent advances in radiation oncology. For consistent positioning, patients are being irradiated in masks. During treatment, tumor and nodule volumes and patient position change within the 7-week treatment schedule. Adaptive radiotherapy accounts for this change by re-adjusting the radiotherapy plan during radiation, increasing treatment efficacy and reducing toxicity. Clinical benefits are currently investigated^{17,18}. Proton beam radiotherapy uses protons instead of the photons used in conventional radiotherapy. Different from photons, protons release their energy at a defined depth, enabling narrower radiotherapy planning and reduced damage to the tumor surrounding healthy tissue. In head and neck cancer treatment this results in decreased xerostomia and dysphagia¹⁹. Adaptive radiotherapy and proton radiation therapy is not yet widely used. Current clinical studies will further examine the clinical benefit and best indications of these techniques.

The role for systemic therapies has also been extensively investigated. The large meta-analysis by Pignon *et al.* in 2009 showed a benefit of 7% increase in overall survival for advanced stage head and neck cancer when treated with cisplatin-based concurrent chemoradiotherapy compared to radiotherapy alone²⁰. This was recently confirmed by an updated meta-analysis²¹. Since the early 2000s, for large head and neck tumors, cisplatin-based concurrent chemoradiotherapy is the standard of care in the Netherlands. Cisplatin has many side effects, especially renal failure, cardiovascular side-effects, neuropathy and ototoxicity. Due to these side effects, a significant proportion of patients is not eligible for chemotherapy or will not complete the full chemotherapy schedule. As documented by several studies this is associated with decreased survival, where the cutoff point seems to be at 200 mg/m²²². Not all patients are suitable for treatment with cisplatin added to radiotherapy. Currently, however, there are not many alternatives. Cetuximab is the only approved alternative. Bonner *et al.* showed, that when cetuximab is added to radiotherapy, median survival improved to 49.0 months, compared to 29.3 months with radiotherapy alone²³. However several studies showed very little effect of cetuximab in HPV-positive tumors, whereas the value in HPV-negative tumors is probably also less than cisplatin^{24,25}.

TGTGAAGGAGATTAAATAAGATGGTGTGATATAAGTATCTGGAGAAAACTTAGGGTGTGGATATTACGAAAGCCTTCCTAAAAATGAATTAACTGATGAGAAGAAAGGATCCAGCTGAGAGCAAAAGAAAAAGCTTCTTCCTTC

These data from clinical trials were confirmed in real world setting²⁶, resulting in a reduction of concomitant cetuximab-RT treatments in Dutch HNSCC patients. In patients ineligible for cisplatin, there is a choice between cetuximab or carboplatin.

HPV-positive HNSCC is currently treated the same as HPV-negative HNSCC. However HPV-positive HNSCC show a good prognosis under current treatment strategies²⁷. Due to the good prognosis of HPV-positive HNSCC, many de-escalation studies have been conducted. Recently, two studies have shown that in HPV-positive oropharyngeal carcinomas concurrent, chemoradiotherapy is still superior to radiotherapy combined with cetuximab^{24,25}. Other possibilities to de-escalate in T3-4 HPV positive oropharyngeal carcinomas are a decrease in the radiotherapy dose, combining a transoral resection with a reduced radiotherapy dose or diminishing radiotherapy fields, e.g. unilateral radiotherapy²⁸⁻³⁰.

Recurrent local or regional disease is preferably treated by surgery if operable. This so called “salvage” surgery is, due to the earlier irradiation, associated with increased complications and morbidity. Wound healing problems, infections, fistula and bleeding is regularly seen. Reirradiation can be successful, especially if recurrences occur after a long interval. But irradiating the same location twice comes with increased risk of necrosis of the surrounding healthy tissue. Metastatic disease at distant sites is frequently incurable and treated with palliative intent. Monotherapies with cisplatin or cetuximab can sometimes give response, but overall do not influence survival. The EXTREME treatment regimen (cetuximab, 5-fluorouracil (5-FU) and platinum) showed modest increased overall survival³¹. The most recent advancements in the treatment for head and neck cancer are from immunotherapy regimens such as nivolumab and pembrolizumab that demonstrated improved overall survival in the recurrent/metastatic setting and a prolonged survival in a small subset of patients^{32,33}. Larger successes have been reported for neoadjuvant immunotherapy, prior to surgery³⁴, although not yet in randomized controlled trials that are currently underway. However, studies combining immunotherapy with radiotherapy or chemoradiotherapy did not reach their endpoints. These combinations are not successful yet.

Although knowledge on cancer biology has greatly progressed and many cancer types gained new therapeutic options based on these new insights, the curative treatment arsenal of HNSCC remained largely unchanged over the last decades except for the recent introduction of immunotherapy. Targeted treatments have not been very successful in HNSCC, mainly as the disease is typically characterized by tumor suppressor gene mutations. In the last decades, efforts have been made to develop novel treatments that harness individual tumor characteristics such as targeting EGFR overexpression or activated PI3K pathways

with inhibitors^{35–37}. Hypoxic tumors are targeted via hypoxia modifiers such as carbogen or nimorazole^{38,39}. To increase radiosensitivity, radiosensitizers such as those targeting DNA-repair (e.g. PARP inhibitors or ATM inhibitors) are tested^{40,41}. As indicated, immune checkpoint inhibition is a new promising therapy. Currently, many efforts are undertaken to discover the benefit in adjuvant and neoadjuvant setting³⁴. Also many studies focus on the tumor immune microenvironment to identify response predicting biomarkers, but the holy grail has not been identified yet. Currently, benefit is closest within the realm of personalized medicine. Biomarkers (see below) are essential to the success of such novel targeted strategies.

1.4 Outcomes

Advanced stage HPV-negative HNSCC patients have a poor prognosis, with 5-year overall survival rates between 40-50%^{42,43}, depending on clinical factors and treatment. Around 25% of patient have a locoregional recurrence¹⁶. TNM-stage is the most important prognostic factor. Tumor-related factors such as primary tumor site, tumor volume and many biological and pathological characteristics play an important role in prognosis^{44,45}. Additionally, patient-related clinical factors such as age, gender, alcohol/tobacco consumption, co-morbidity, blood hemoglobin levels, neutrophil-to-lymphocyte ratio and WHO-index may play an important role in prognosis^{16,46–50}.

2. Head and neck cancer biology

In 2000, Hanahan and Weinberg published their first six hallmarks of cancer⁵¹. Recently in 2022 an updated, third version was released with a total of 14 hallmarks⁵². These hallmarks are characteristics for normal cells to acquire in their transformation to cancer cells and are depicted in Figure 2. Like other cancer types, also head and neck cancer squamous cell carcinomas have frequent genomic alterations resulting in the acquisition of some of these hallmarks (Table 1)⁵³.

Approximately 80-90% of cancers arise from epithelial cells and are therefore called carcinoma. Amongst many types, the two major subtypes of carcinomas are squamous cell carcinoma and adenocarcinomas. Adenocarcinomas arise in epithelial cells with glandular origin. Most common sites for adenocarcinoma are breast, colon, prostate and lung. Squamous cells can be found throughout many parts of the body, including head and neck, esophagus, lungs, the anogenital region and the skin. Squamous cell carcinomas from different sites share similar genomic alterations⁵⁴. Genetic alterations typical tot HNSCC are shown in Table 1. They can be assigned into different biological processes which deregulation is an important hallmark in HNSCC development.

TGTGAAGGAGATTAAATAAGATGGTGTGATATAAGTATCTGGGAGAAAAGCTTAGGGTGTGGATATTAAGAAAGCCTTCCTAAAAAATGACATTTAACTGATGAGAGAAAGGATCCAGCTGAGAGCAAAAGCAAAAGCTTTCTTCCTTC



Figure 2. Depiction of the recently updated hallmarks of cancer as presented in the paper by Hanahan and Weinberg⁵².

Table 1. Genes with frequent and highly significant somatic genetic changes in HPV-negative HNSCC. Adapted from Leemans *et al.*⁵³.

Cellular process	Gene	Protein	Type of gene	Mutation frequency(%)	CNA frequency (%)
Cell cycle	CDKN2A	p16 ^{INK4A}	Tumor suppressor	22	32
	TP53	p53	Tumor suppressor	72	1.4
	CCND1	G1-S-specific cyclin D1	Oncogene	0.6	25
Growth signal	EGFR	EGFR	Oncogene	4	11
Survival	PIK3CA	Catalytic p110α subunit of class 1 PI3Ks	Oncogene	18	21
	PTEN	PTEN	Tumor suppressor	3	4
WNT signaling	FAT1	Protocadherin FAT1	Tumor suppressor	23	8
	AJUBA	LIM domain-containing protein AJUBA	Tumor suppressor	7*	1
Epigenetic regulation	NOTCH1	NOTCH1	Tumor suppressor	18	4
	KMT2D	KMT2D	Tumor suppressor	16	0.4
	NSD1	NSD1	Tumor suppressor	12*	0.8

Data from⁵⁵. Mutation data taken from The Cancer Genome Atlas (TCGA)(n=504) using the cBioPortal. CNA=copy number alterations; EGFR=epidermal growth factor receptor. *Putative passenger mutation that requires further functional studies.

2.1 Cell cycle

The cell cycle is most frequently and consistently affected in head and neck cancer. CDK4/CDK6 and CyclinD1 (*CCND1* gene product) form a complex that drives S-phase entry by phosphorylating the Retinoblastoma protein (pRB) and releasing E2F (Figure 3). *CDKN2A*, a tumor suppressor gene, inhibits cell cycle progression via its product p16^{INK4A}.

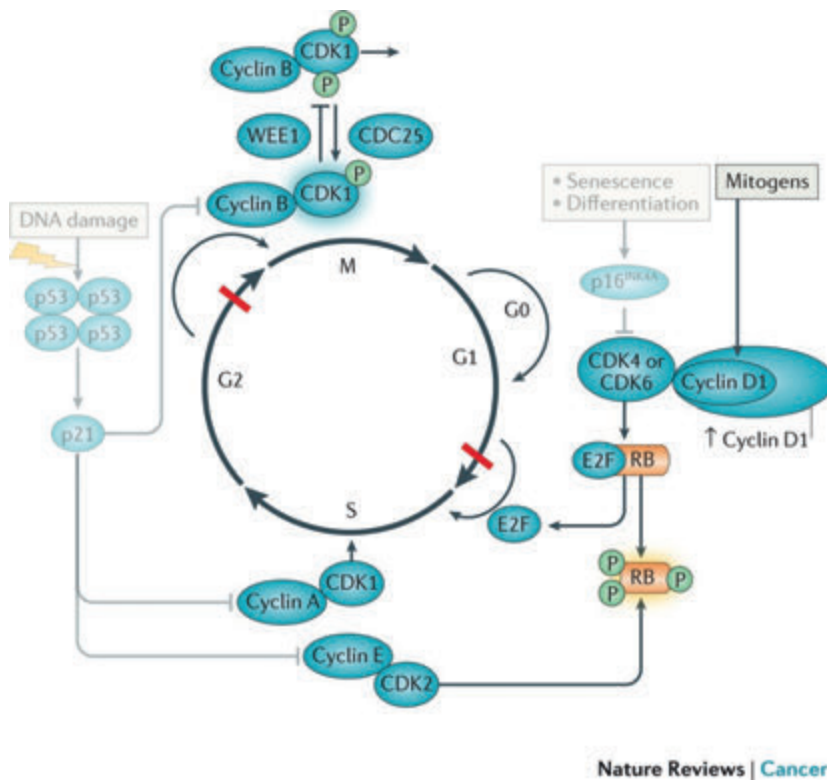


Figure 3. Schematic overview of proteins involved in cancer cell cycle. Figure adapted from Leemans, C. R., Braakhuis, B. J. M. & Brakenhoff, R. H. The molecular biology of head and neck cancer. *Nat. Rev. Cancer* **11**, 9–22 (2011), Macmillan Publishers Limited⁵⁶.

CCND1 amplification and *CDKN2A* loss are frequent findings in HPV-negative HNSCC, resulting in cell-cycle progression (Figure 4)^{55,57,58}. Unscheduled S-phase entry causes replication stress and DNA damage. The accompanying DNA-damage normally activates p53, resulting in cell cycle arrest and apoptosis. However, in most HPV-negative head and neck cancers, inactivating genetic alterations are found in the *TP53* gene, further supporting uncontrolled cell-cycle progression and proliferation. *CDKN2A* loss is associated with poor overall survival in HNSCC patients⁵⁹.

In HPV-positive head and neck cancer cell cycle control is disabled via an alternative route. The HPV genome contains the viral oncogenes E6 and E7. The E6 protein disables TP53 and E7 inhibits the RB protein. This results in cell-cycle progression and proliferation. Because of the expression of the viral oncogenes E6 and E7, genomic alterations in cell cycle genes are not required and thus are rare events in HPV-positive HNSCC. Overall, HPV driven HNSCC are biologically very different from HPV-negative, with different tumor biology, clinical characteristics and treatment response.

2.2 Proliferation

The hallmark ‘sustaining proliferative signaling’ is also frequently affected in head and neck cancer (Figure 4). EGFR (Epidermal Growth Factor Receptor) signaling is activated in 12–15% of HPV-negative HNSCC (Figure 4, Table 2). In a third of HPV-negative HNSCC this is through EGFR gene amplification and overexpression. EGFR activates many downstream pathways including the mTOR/PI3K/AKT pathway, resulting in proliferation and pro-survival signaling. Besides EGFR, the *PIK3CA* gene also shows activating alterations in 13–34% of HPV-negative HNSCC and even in more than 50% of HPV-positive HNSCC (Figure 4).

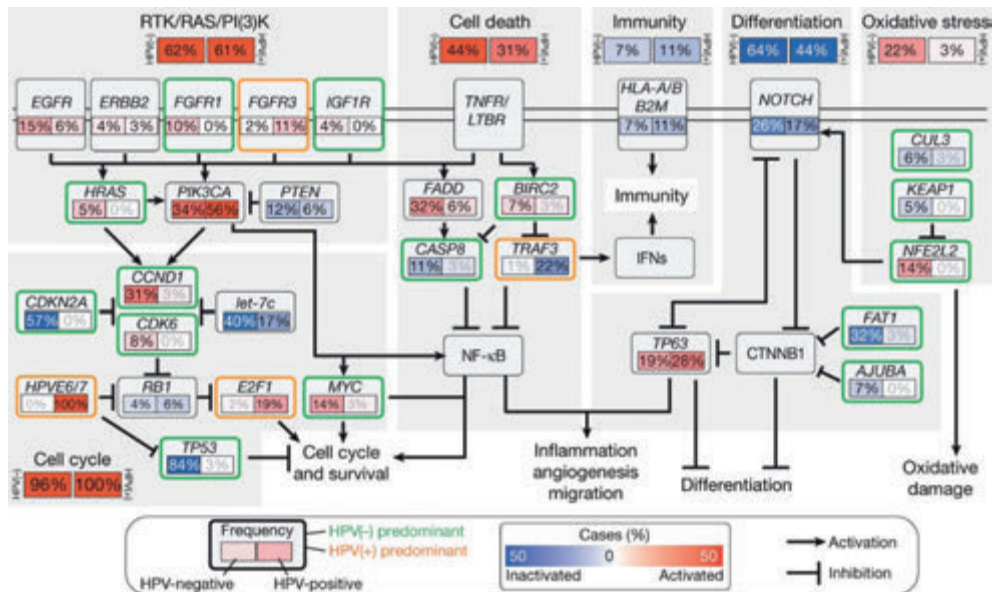


Figure 4. Deregulation of signaling pathways and transcription factors in HNSCC. The frequency (%) of genetic alterations for HPV(-) and HPV(+) tumors are shown separately within sub-panels and highlighted. Pathway alterations include homozygous deletions, focal amplifications and somatic mutations. Activated and inactivated pathways/genes, and activating or inhibitory symbols are based on predicted effects of genome alterations and/or pathway functions. Figure and figure legend copied and adapted from MS Lawrence *et al. Nature*⁶².

PI3K is inhibited by PTEN and its pathway activation is further sustained by PTEN inactivation in around 4-15% of HPV-negative HNSCC^{55,60}. Interestingly, alterations in the PI3K/PTEN pathway are associated with a better locoregional control rate after radiotherapy, although this study contains HPV-positive and –negative patients⁶¹.

2.3 Hypoxia

As stated before, radiotherapy is a corner stone of head and neck cancer treatment. Up to two-third of patients receive radiotherapy. The efficacy of radiotherapy depends on the extent of oxygenation and with that the ability to form DNA damaging free oxygen radicals. Cells residing in an hypoxic environment, or hypoxic cells, are therefore exposed to less radiation damage^{63–65}. Many clinical and preclinical *in vivo* studies have shown that tumor hypoxia is a negative prognostic factor⁶⁶. Hypoxia can be analyzed via various methods. It can be measured directly, by radiological imaging, or indirectly using molecular markers on biopsy material. The Eppendorf pO₂ method uses an oxygen-sensitive needle to directly measure the oxygen concentration in the tumor⁶⁷. In PET-imaging studies, a variety of tracers have been used to identify more and less hypoxic tumors^{68,69}. In research, the most used method to identify hypoxic tumors is immunohistochemistry (IHC). Under hypoxic condition, hypoxia inducible factor 1 alpha (HIF1- α) and 2 alpha (HIF2- α) are stabilized. These hypoxia-activated transcription factors induce and raise transcription of *CAIX*, *GLU1* and *VEGF*. All these molecular markers are used to identify hypoxic tumors^{70–74}. Pimonidazole is a good marker for hypoxia^{75,76}. Pimonidazole visualizes hypoxic areas by binding to cellular molecules at low oxygen levels. More recently, many hypoxia-specific gene expression profiles have been described^{77–81}. No matter which technique was used, they all showed that hypoxia is associated with worse prognosis and therapy resistance in head and neck cancer^{71,73,82–85}. There are several ways for a tumor micro-environment to become hypoxic. Rapidly proliferating tumors require nutrients and oxygen to grow, which is why tumors induce angiogenesis, the formation of new blood vessels and one of the hallmarks of cancer. Diffusing oxygen from the blood vessels result in an oxygen gradient. Cells that are situated furthest from blood vessels, at the end of this oxygen gradient, are continuously and increasingly oxygen deprived and therefore chronically hypoxic. In contrast to chronic hypoxia there is acute hypoxia in tumors⁷⁸. The hypoxia-induced neovasculature is malformed and vulnerable, resulting in fluctuations in blood flow. Acute hypoxic cells are a result of highly fluctuating blood flow, causing quick changes in oxygen concentrations. The occasional occlusion of blood vessels results in diminished blood flow and acute hypoxia⁷⁸.

TGTGAAGGAGATTAAATAAGATGGTGTGATATAAGTATCTGGGAGAAAACTTAGGGTGTGGATATTACGAAAGCCTTCCTAAAAAATGAATTAACTGATGAGAAGAAAGGATCCAGCTGAGAGCAAAAGCAAAAGCTTTCTCCTTC

2.4 DNA-repair

Cells are constantly subjected to DNA damage. This is intrinsic by DNA replication errors, and can be due to endogenously generated metabolites of normal cell metabolism. But also exogenous factors such as smoking, alcohol and UV light cause DNA damage. Therapeutic regimen such as radiation and chemotherapeutic agents inflict DNA damage to provoke apoptosis and cell death. The human genome consists of 3 billion nucleotides, but the exome, the part that can be translated to proteins, is roughly 1% of the whole genome. Replication of the genome is very accurate and errors occur rarely, around once in 10^{108} . Normal cells have a complex apparatus of signaling networks/checkpoints and mechanisms to repair different kinds of DNA damage⁸⁶. For each different type of DNA damage there is a mechanism to repair the damage (Figure 5). Although the chances that mutations affect protein coding regions are small, this chance increases with prolonged exposure to DNA damaging factors. Prolonged exposure will ultimately result in driver mutations such as previously mentioned in Figure 4. Especially when mutations occur in DNA-repair genes or when checkpoints are disabled, the integrity of the genome is challenged. Such alterations result in a rapidly increasing number of mutations and other genome aberrations. This process is called genomic instability, another hallmark of cancer⁸⁷. Head and neck cancer shows a relatively high rate of mutations per mega base of the genome^{88,89}, the so-called tumor mutational burden.

Chemotherapeutics, such as cisplatin, and radiotherapy generate DNA damage. Cisplatin, the most frequently used chemotherapeutic agent in HNSCC, causes DNA interstrand cross-links⁹⁰. DNA crosslinks cause DNA replication blocks in proliferating cells, resulting in replication fork collapse and double strand breaks. In normal cells, such replication blocking crosslinks are resolved and repaired using the Fanconi anemia/homology-directed repair (FA/HR) pathway. Double strand breaks are caused by ionizing radiation either directly or indirectly when DNA with single stranded breaks are replicated. These are primarily repaired by the error-prone non-homologues end-joining pathway and homology-directed repair pathways in replicating cells. Unrepaired double strand breaks, the most lethal type of DNA damage, cause cell cycle arrest and cell death. Tumor cells with impaired DNA repair pathways are genomic unstable because they cannot repair these crosslinks or resolve replication fork stalling which makes them more susceptible for cisplatin. In addition to double strand and single strand breaks, radiotherapy also induces a variety of other DNA alteration including base and nucleotide alterations. However, DNA double strand breaks induced by radiotherapy are thought to be the primary cause of cell death after radiation even though the minority of radiation induced DNA-damage are double strand breaks⁹¹.

Cisplatin together with radiation synergistically increases the number of DNA double strand breaks and increases the therapeutic effect, but the local side effects as well⁹².

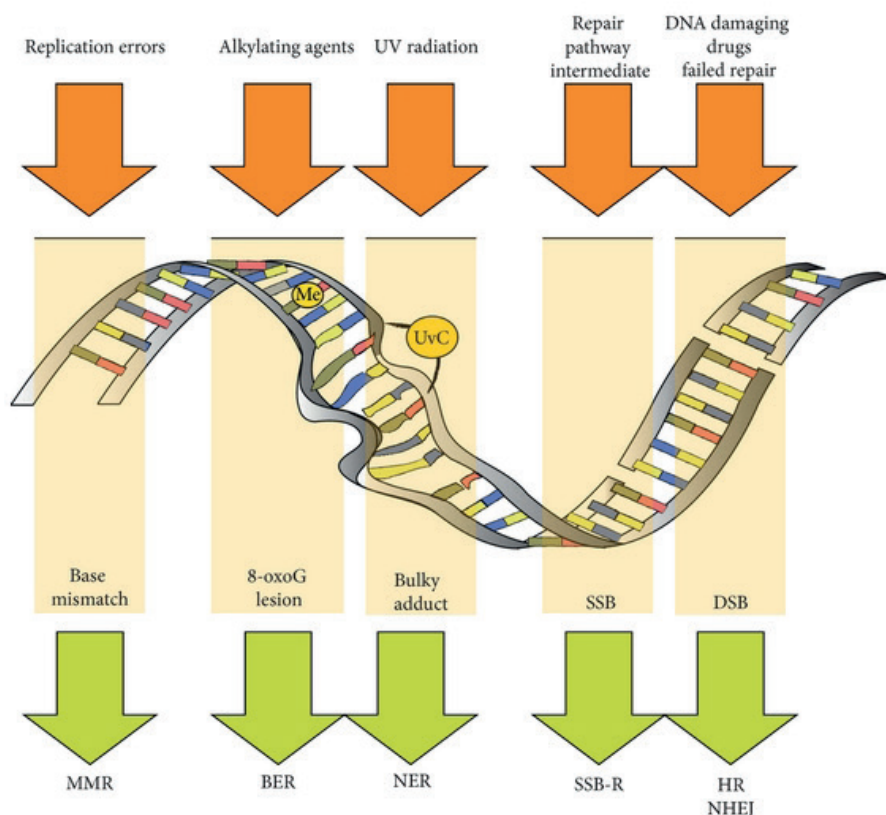


Figure 5. Examples of DNA lesions and repair pathway choice. SSB = single strand break. DSB = Double strand break. MMR = Mismatch repair. BER = Base excision repair. NER = Nucleotide excision repair. SSB-R = Single strand break repair. HR = Homologues recombination. NHEJ = Non-homologues end-joining. Figure and legend adapted from Jenkins *et al.*, "Genome Stability Pathways in Head and Neck Cancers", *International Journal of Genomics*⁹³.

2.5 Epithelial to mesenchymal transition

In healthy, normal tissue, epithelial cells are polarized and bound to the basement membrane. During epithelial to mesenchymal transition (EMT), normal epithelial cells gain mesenchymal characteristics, i.e. migratory properties, invasiveness and change in shape. This ultimately translates into the degradation of the basement membrane and invasion of tumor cells (Figure 6)⁹⁴. This process plays a role in lymphatic or blood vessel invasion and is associated with increased nodal and distant metastasis. Six core transcription factors

TGTGAAGAGATTAAATAGATGGTGTGATATAAGTATCTGGGAGAAACCTTAGGGTGTGGATATTAAGAAAGCCTTCCTAAAAAATGAATTAACTGATGAGAAGAAAGGATCCAGCTGAGAGCAAAAGCAAAAGCTTTCTCTCTTC

(ZEB1, ZEB2, SNAI1, SNAI2, TWIST1 and TWIST2) regulate EMT. These transcription factors are activated by the TGF β /SMAD and the EGFR/MAPK/ERK pathway⁹⁵. Upregulation of N-cadherin, vimentin and β -catenin and downregulation of E-cadherin are characteristic changes in EMT. In many cancer types, EMT is associated with worse prognosis. Pan-cancer EMT analysis have shown that tumors with high EMT markers show a worse overall survival⁹⁶. The clinical impact of EMT in HNSCC is not well understood. Assessment of individual EMT markers in HNSCC using immunohistochemistry has been linked to impaired overall survival, although others found no impact^{97–102}. Recent meta-analysis of several EMT related transcription factors (TWIST1, SNAI1, SNAI2, ZEB1) show that overexpression is associated with a poor overall survival^{103,104}. Chung *et al.* developed a gene expression signature to identify high-risk disease. This signature included genes involved in EMT¹⁰⁵. EMT activation is associated with an increase in immune activation¹⁰⁶.

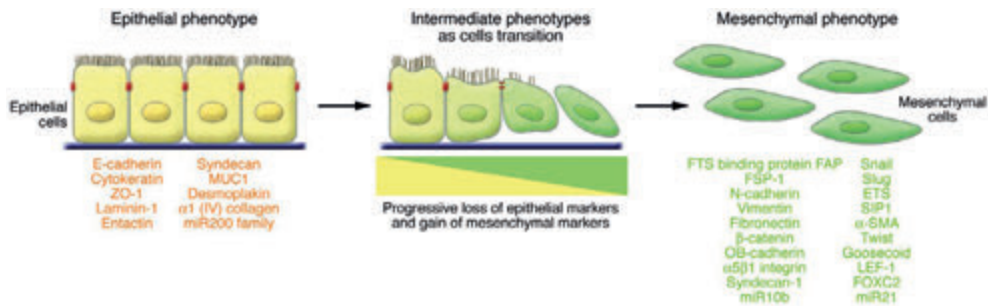


Figure 6. EMT involves a functional transition of polarized epithelial cells into mobile and extracellular matrix (ECM) component-secreting mesenchymal cells. The epithelial and mesenchymal cell markers commonly used EMT research are listed. Colocalization of these two sets of distinct markers define an intermediate phenotype of EMT, indicating cells that have passed only partly through EMT. ZO-1, zona occludens 1; MUC1, mucin 1, cell surface associated; miR200, microRNA 200; SIP1, survival of motor neuron protein interacting protein 1; FOXC2, forkhead box C2. Figure and legend adapted from Kalluri *et al.* J Clin Invest. 2009⁹⁴.

2.6 Immunogenicity

Over the last decade, efforts have revealed the important role of the immune system in the development of cancer. Mutations in the genome result in altered proteins, so called neoantigens, which are normally recognized by the immune system as not-self and initiate an immune response that may lead to cell death. Cancer cells interact with the immune system to mold this immune response. The immune response between HPV-negative and HPV-positive tumors is different¹⁰⁷. Tumor cells can express immune checkpoints of which programmed death-1 (PD-L1) and cytotoxic T-lymphocyte antigen 4 (CTLA4) are the most well-known. By expressing these checkpoints on their cell surface they inhibit the

immune response and prevent cell death¹⁰⁸. A new type of drugs called immune checkpoint inhibitors (ICI) intervene at the tumor cell and immune cell interaction. By blocking this interaction, the anticancer response of the immune system can be activated. Nivolumab and pembrolizumab, both anti-PD-1 drugs, have shown improved survival in patients with platinum refractory recurrent/metastatic HNSCC when compared to standard of care^{109,110}. However, avelumab, an anti-PD-L1 drug, showed no additional benefit when added to the standard of care (chemoradiation) compared to placebo¹¹¹. Around 20 percent of HNSCC patients respond to immunotherapy¹¹². Prediction of immunotherapy response is currently an important focus area^{34,112}.

3. Outcome prognosis and prognostic biomarkers in HNSCC

The prognosis for head and neck cancer patients, particularly with advanced stage of disease, is poor and improved little over the last decades. While approximately half of the patients benefit from current treatment regimens, cancer will recur in the other half. Prognostic clinical factors and prognostic biomarkers can help to identify patients in need for alternative or intensified treatments or can help to lower treatment associated toxicities in good prognosis patients by de-escalation.

3.1 Patient prognosis

As listed above there are currently several clinical factors that are reliably associated with patient outcome to support patient prognosis estimates or influence treatment choice. TNM stage or tumor volumes are, in the case of chemoradiotherapy (CRT) treated advanced HNSCC, known prognostic factors. Using clinical factors only, many models besides TNM are currently available, however not widely used¹¹³. Some factors may be prognostic with regards to some patient outcome endpoints but not to others (see Table 1). The most relevant endpoints in head and neck cancer are overall survival, progression-free survival, locoregional control and distant metastasis-free survival. Locoregional recurrences are generally regarded as treatment failures but this is likely an oversimplified view, and given the generally poor health of HNSCC patients, overall survival is highly influenced by comorbidity as well. As apparent in the representative Kaplan-Meier curves in Roest *et al.*¹⁶, 5 year follow up is considered sufficient for most endpoints for association analyses with potential prognostic factors.

In contrast to prognostic factors, predictive factors are associated with the response to the treatments. Identification of predictive factors requires the comparison between treated and non-treated patients or between different treatment regimens, which is hard in the curative

setting and the current study. In randomized clinical trials (RCTs) comparing differently treated patients, this is more feasible. However, large RCTs, are rare in the curative setting in HNSCC. This further complicates the identification of predictive biomarkers in HNSCC.

3.2 Technological advancements

With the introduction of micro-arrays and SNP arrays, gene-expression or genome changes could be quantified for a large number of genes at the same time. This provided profound insights in cancer biology. This was further accelerated with the introduction of the next-generation sequencing (NGS) machines in the mid-2000s. With NGS, high-throughput sequencing of the DNA of many patients became affordable and feasible. This resulted in large pan-cancer sequencing projects of which The Cancer Genome Atlas is the most well-known. In addition to TCGA many smaller HNSCC-specific sequencing studies were undertaken^{55,57,58,60,105,114–118}. The technological advances currently include single cell sequencing methods, that provided in depth inside in tumor biology and the interaction with the tumor micro-environment¹¹⁹. Together with technological advancements, such as NGS, computational biology also showed great progress. This resulted in machine learning techniques where computational power is used to find patterns in large datasets. Algorithms used in these techniques are a form of artificial intelligence that aims to learn from the data. Complex modelling by these algorithms can reveal biological connections and be useful to identify biomarkers for clinical practice¹²⁰. A comparable development is the clinical application of machine learning on imaging data, termed radiomics. With radiomics, software analysis tools extract many features of radiological imaging. With this technique thousands of data points are extracted of a single image, i.e. a CT or MRI scan. These large datasets can then be analyzed using different machine learning techniques, also creating markers of response with potential for clinical application. Together, these new technological advancements have provided a rapid increase in the understanding of cancer biology.

3.3 Biomarkers

The developments in personalized treatment and the novel treatment opportunities call for improved prognostic markers. In order to identify high-risk patients and to explore novel therapeutic targets, biomarker research has increased rapidly. The underlying assumption is that tumor biology determines aggressiveness and response to treatments. Biomarkers can therefore be both, prognostic and predictive. Prognostic biomarkers can be a valuable tool to identify patients with high-risk tumors and a diminished prognosis even if not causally

linked to known tumor biology processes. With predictive biomarkers, researchers and clinicians aim to identify tumors that predict response to certain therapies. In many cases, predictive biomarkers are related to the targeted treatments, such as the use of gefitinib, which targets EGFR signaling, in EGFR mutated tumors. RCTs with monotherapy drugs are ideal to test such associations as they allow the assessment of response with or without the treatment in question.

Many prognostic biomarkers may implicitly also be predictive, such as HPV positivity. HPV-positive HNSCC have been shown to be more radiosensitive and more sensitive to cisplatin as outcomes are better under current (chemo)radiotherapy treatments. In part this has been formally proven in the clinical setting^{24,25}.

HPV

The most established prognostic biomarker is the presence of HPV, most particularly in oropharyngeal cancers. As mentioned before, HPV is the most important biomarker in HNSCC and is a biology-driven biomarker that identifies tumors caused by HPV and are expressing HPV oncogenes E6 and E7. With a different carcinogenesis mechanism, their biology and behavior is different from HPV-negative HNSCC. This resulted in the latest TNM in a separate p16-positive TNM classification, with p16 immunostaining as surrogate marker for HPV. The prognosis of HPV-positive HNSCC patients is good after conventional treatments with radiotherapy with or without chemotherapy, but also after surgical treatment using TORS²⁷. HPV assessment can be done using a variety of methods with different levels of accuracy. Due to easy access to immunohistochemistry worldwide, p16 positivity as determined by IHC is taken as a marker for HPV-positive HNSCC classification in the 8th TNM classification system¹⁰. However, p16 overexpression is not exclusively occurring in HPV-positive tumors^{121,122}. In the Netherlands, around 10 percent of p16 positive oropharyngeal carcinomas are HPV DNA and E6 RNA negative (as determined by PCR) and show more similarities to HPV-negative HNSCC¹²². Given that HPV-positive largely lack p53 pathway mutations, IHC of p53 can improve accuracy in HPV status classification¹²³. However, the gold standard requires conformation of HPV DNA by PCR¹²². Based on clinical factor, prognostic factors and molecular characteristics, HPV-positive HNSCC are different from HPV-negative HNSCC. In this thesis we focus on HPV-negative HNSCC.

Proliferation

Proliferation, as determined by the proliferation gene expression profile by Starmans *et al.*, has been linked to aggressive disease or disease progression in multiple cancer types;

TGTGAAGGAGATTAAATAAGATGGTGTGATATAAGTATCTGGGAGAAAACCTTAGGGTGTGGATATTAAGAAAGCCTTCCTAAAAATGAATTAACTGATGAGAAGAAAGGATCCAGCTGAGAGCAAAAGCAAAAGCTTTCTTCCTTC

unfortunately this was not assessed in HNSCC¹²⁴. EGFR is an important modulator of proliferation and overexpressed in the majority of HNSCC^{125,126}. In a large review in 2001, EGFR expression was associated with patient prognosis in many cancer types, especially in HNSCC¹²⁷. IHC studies, in clinical and preclinical settings, have shown that EGFR overexpression is associated with poor prognosis in radiotherapy treated tumors, but not for surgery treated tumors, suggesting decreased radiosensitivity^{128–131}. Whether EGFR overexpression is a prognostic factor in HNSCC for patients treated with chemoradiotherapy is inconclusive^{132–135}. More recent research demonstrated that EGFR expression seems a predictive biomarker for accelerated radiotherapy^{131,136–138}. EGFR pathway activation can be targeted by the EGFR binding antibody cetuximab. The combination of radiotherapy with cetuximab has shown efficacy in HNSCC^{23,131,138,139}. However, EGFR expression or increased EGFR gene copy numbers appear not to be predictive for cetuximab response¹³¹. Recently, Bossi *et al.* also developed a gene expression signature that identified patients that respond to cetuximab¹⁴⁰. Responders showed strong EGFR signaling and hypoxia, emphasizing the role of EGFR in radiation response.

Another important modulator of proliferation is PI3K. Activation of ERBB proteins, of which EGFR is one, results in activation of the PI3K/Akt pathway¹⁴¹. However, the PI3K protein can be activated via mechanisms such as hot spot mutations or amplification of PIK3CA or loss of function mutation in inhibitors such as PTEN. Activation of PI3K results in the activation of AKT, a strong cell survival signal. PI3K is frequently affected in HNSCC, mostly in HPV-positive HNSCC^{55,118,142}. The effect of PI3K alterations has not been studied widely. A single study showed that mutations in this pathway have been associated with better prognosis after radiotherapy, although no information on HPV status was given⁶¹. In line with these results, PI3K alterations are also associated with better prognosis in breast and rectal cancers treated with radiotherapy or chemoradiotherapy in cervical cancer^{143–146}. In contrast, another study showed an association with poor overall survival for cervical cancer patients when treated with chemoradiotherapy¹⁴⁷.

Hypoxia

Direct measurements of hypoxia in different areas of tumors, the Eppendorf method, showed that hypoxia is associated with poor survival⁶⁷, however this method is labor intensive and invasive. In HPV-negative tumors, hypoxia is associated with poor prognosis^{85,148,149}. All these previously described methods have been confirmed to be associated with poor survival in HNSCC. Expression of HIF1 α , a transcription factor induced by hypoxia, is a poor prognostic factor in surgically treated HNSCC patients. HIF1 α and CAIX expression are associated with

local recurrences in laryngeal carcinoma⁷¹. PET imaging using different tracers is also used to identify hypoxia in tumors¹⁵⁰. Hypoxia markers and their correlation in advanced stage HNSCC patients treated with chemoradiation is less extensively examined but published results are in line with other treatments. Koukourakis *et al.* showed that increased CAIX, HIF1 α and HIF2 α expression was associated with worse prognosis in chemoradiotherapy treated HNSCC^{70,151}. De Schutter *et al.* showed similar results for the hypoxia markers CAIX, GLUT1 and VEGF⁸⁴.

Transcriptomic approaches and data mining have revealed several gene expression profiles for hypoxia^{77,79,81,85}. Like other methods to assess hypoxia, they show a relation to patient outcome in HNSCC, mostly in surgically treated patients. Eustace *et al.* was developed on patients with laryngeal carcinoma treated with radiotherapy⁷⁷. Winter *et al.* did the same but the patients in their study were primarily treated with surgery⁷⁹. Buffa *et al.* used several datasets across multiple cancer types, including four HNSCC datasets, to develop and validate their hypoxia signature. Almost all patients included were primarily treated with surgery or radiotherapy. None of the patients in these HNSCC datasets were primarily treated with chemoradiotherapy⁸¹. Tawk *et al.* combined all gene expression profiles and tested them on the HNSCC data of the TCGA⁸⁵. There are few patients with hypopharyngeal carcinoma and the group of patients is very heterogenous regarding treatment. The impact of hypoxia on outcome in hypopharyngeal and oropharyngeal HNSCC patients treated with chemoradiotherapy are less studied. Whether the gene expression profiles represent either acute or chronic hypoxia is not well understood either.

EMT

Although the relevance of EMT is shown in other cancer types, it remains ill-defined in HNSCC. A variety of studies have been conducted that assessed markers for EMT using IHC. In all studies it was shown that EMT is associated with worse prognosis. Assessed individually, decreased E-cadherin expression or increased expression of N-cadherin, β -catenin, vimentin, TWIST and ZEB, all factors related to EMT, have been linked to a reduced overall survival in HNSCC^{97,152}. However only Nijkamp *et al.* included patients treated with primary chemoradiotherapy. Recent meta-analysis confirmed the prognostic value of the six key transcription factors (TWIST1, TWIST2, SNAI1, SNAI2, ZEB1, ZEB2, N-cadherin) in several subsites of HNSCC^{99,104}. Hypo- and oropharyngeal carcinomas were underrepresented in most studies. All studies that included hypo-, oropharyngeal and laryngeal tumors, reported results from patients treated with surgery. Hence, associations with chemoradiation cannot be extrapolated. Together, there is very little known about EMT and chemoradiotherapy response. Currently there is no HNSCC-specific EMT gene expression profile. A prognostic gene-expression profile in HNSCC however included genes involved in EMT¹⁰⁵. And a pan-

TGTGAAGGAGATTAAATAAGATGGTGTGATATAAGATATCGGGAGAAAAGCTTAGGGTGTGGATATTAAGAAAGCCTTCCTAAAAAATGAATTAACTGATGAGAGAAGAGGATCCAGCTGAGAGCAAAAGCAAAAGCTTCTTCCTTC

cancer analysis of EMT, using TCGA data, showed that HNSCC show high expression levels of their EMT gene expression profile⁹⁶, but a relation to outcome was not shown for HNSCC.

1

DNA repair

DNA-repair is achieved by a compendium of many different pathways to repair different kinds of DNA-damage. ERCC1 is a gene involved in the Nucleotide Excision Repair pathway. This pathway is involved in removing cisplatin adducts¹⁵³. Older studies reported that high expression of ERCC1 is associated with patient outcome in HNSCC patients treated with cisplatin^{154–156}. Also somatic mutations in DNA repair-related genes have been reported in larger HNSCC sequencing efforts^{125,157–160}. Noteworthy are the data on the FA/HR pathways, involved in the repair of cisplatin induced DNA crosslinks. Promotor methylation of several FA/HR pathway genes have been reported in some HNSCC cases¹⁶¹. Association analyses of FA/HR DNA-repair pathway defects and outcomes in HNSCC are scarce. A small study that compared HR pathway engagement in 13 HNSCC patients found an impairment to be associated with a poorer patient outcome¹⁶² and mutations in DNA-repair genes have been associated with worse locoregional control and overall survival in HNSCC patients treated with chemoradiotherapy in a study conducted by our group¹⁶³.

Recently a new type of drugs, named PARP inhibitors, have been introduced. Polyadenosine diphosphate-ribose polymerase (PARP) plays a key-role in the repair of single strand breaks (SSBs). Inhibition of PARP, leads to unrepaired SSBs, which will lead to replication fork stalling and ultimately to increased double strand breaks (DSBs). These, in turn, are normally repaired using the homology directed repair pathway, which shows overlap with the FA/HR pathway. Tumors with defects in this pathway, such as BRCA1 and BRCA2, may therefore benefit from PARP inhibitors such as olaparib or niraparib¹⁶⁴. In ovarian cancer this led to recommendations to treat BRCA-associated ovarian cancer with PARP inhibitors¹⁶⁵. PARP inhibition in combination with chemo- and/or radiotherapy can work synergistically^{157,166,167}. Currently, several PARP inhibitors are tested in HNSCC¹⁵⁷. FA/HR defects may function as a predictive biomarker for response to PARP inhibition.

Immunology

Although immunotherapy is currently not yet used in the curable setting in routine practice, immunotherapy is at the forefront of clinical research. As only subgroups of patients respond, biomarker research for immunotherapy response and the role of the tumor microenvironment has become a hot topic. Using the TCGA dataset, Mandal *et al.* showed that markers for regulatory T cells (Treg), NK cells and CD8⁺ T cells are prognostic in

HNSCC¹⁰⁷. Immunohistochemically determined high CD8⁺ T cell numbers are associated with good prognosis in HNSCC patients treated with postoperative chemoradiotherapy¹⁶⁸. A good prognosis association with IHC CD8⁺ TIL density was found in patients with oropharyngeal squamous cell carcinoma treated with surgery or (chemo) radiotherapy and in a similarly mixed treatment cohort of hypopharyngeal SCC patients^{169–171}. The consistent effect of CD8⁺ T cell depletion on radiation responses in preclinical studies expose the relevance of certain immune cell populations in radiation response and resistance^{172,173}. Current studies focus on optimizing response to immune response modulators to improve radiotherapy outcomes^{174–177}.

PD-1 and PD-L1 expression is associated with increased response to immune checkpoint inhibition (ICI) in the recurrent/metastatic HNSCC setting^{110,178,179}. The combined positive score (CPS) is a score currently used as a biomarker for prediction of response to pembrolizumab. This method combines the proportion of PD-L1 positive tumor cells and PD-L1 positive immune cells. The CPS is used as the ratio of the PD-L1 positive cells of the total viable tumor cells and is a good predictor for the effect of pembrolizumab¹⁸⁰. Tumor mutational burden (TMB) is another widely investigated biomarker for the effect of immunotherapy. In lung-cancer, TMB was associated with increased efficacy of immunecheckpoint inhibitors¹⁸¹. In a pan-cancer analysis that included HNSCC, Valero *et al.* showed that a high TMB was a predictive marker for the use of checkpoint inhibitors¹⁸².

General clinical biomarkers

The high-throughput data generation of micro-arrays was the dawn of a new era in biomarker discovery research. With the introduction of NGS, biomarker research exploded. Especially, the high-throughput data of transcriptome data has proven an almost inexhaustible source of prognostic gene expression profiles (GEPs) or signatures^{183–187}. For example, GEPs are available for N-stage in HNSCC^{188,189}, for prognosis for different subsites of HNSCC^{190–192} or for HNSCC in general^{193,194} and there are even GEPs available for prognostication of advanced stage HNSCC after treatment with chemoradiation^{183,185}. Many of these have even been validated in independent cohorts. However, none of these prognostic biomarkers are in clinical use. There are several reasons for the lack of implementation. First, most GEPs depend on frozen biopsies that are scarce in routine practice. In addition, the GEP associated with lymph node metastases that is very well validated, suffers from a competing approach: sentinel node biopsy. Other GEPs are insufficiently validated, or give no direction to alternative treatments or new insights in the tumor biology of HNSCC.

4. Thesis outline

Despite invasive treatment protocols, resulting in increased morbidity, and advances in technology and new therapies, the survival of advanced stage HPV-negative head and neck cancer patients is poor and has not or minimally improved over the last few decades. Most advanced stage HNSCC tumors are treated with concomitant chemoradiation either definitive or post-operative. It is unknown as to why some tumors respond well to this treatment regimens while others, seemingly similar tumors, do not. If we can predict upfront which tumors respond well, we can better select patients for alternative treatment and improve their prognosis and decrease their treatment related toxicities. The current perception is that tumor biology plays a central role in treatment response. In order to adequately select the best treatment for each patient, there is an unmet need for prognostic biomarkers and novel treatment options for this patient population. For this purpose, the DESIGN consortium was set up. This is a multicenter consortium, funded by KWF kankerbestrijding, and consisting of Amsterdam UMC, the Netherlands Cancer Institute/Antoni van Leeuwenhoek (NKI/AvL) in Amsterdam, UMC Utrecht, the MAASTRO clinic in Maastricht, and Maastricht University. The consortium focused on identifying clinical, radiological, pathological and genetic biomarkers for prognosis in HPV-negative HNSCC of oropharyngeal, hypopharyngeal and laryngeal subsites, treated with definitive chemoradiotherapy. Poor responding tumors could be treated with upfront surgery as an alternative. As part of the DESIGN consortium, NKI/AvL explored the transcriptome for the identification and/or validation of prognostic biomarkers. The department of head and neck surgery of the Netherlands Cancer institute has a long-lasting history of biomarker research in HNSCC. Collaborations with the radiation oncology department and department of biological stress response, have been indispensable. These studies have been led by the late Prof. Adrian Begg, Prof. Fons Balm, Prof. Michiel van den Brekel, Dr. Conchita Vens and Prof. Marcel Verheij.

Led by the late Prof. A Begg, HNSCC research at the NKI has a long-standing interest into the role of hypoxia for outcome. Dr. de Jong's research into gene expression based hypoxia biomarkers confirmed its potential and relevance for patient outcome^{74,195}. In **Chapter 2** we focused on hypoxia, a known determinant of radiosensitivity and important prognostic factor in HNSCC. However, current gene expression profiles make no distinction between acute and chronic hypoxia. Using published gene expression profiles for acute and chronic hypoxia generated on cell lines, we compare the previous described hypoxia profiles. We investigate whether the clinically used profiles represent both types of hypoxia. We test all profiles on different cohorts, totaling 398 patients, to determine the impact on outcome.

EMT is an important factor in cancer progression and patient outcome in multiple cancer types. *In vitro* studies conducted by Dr. de Jong at our institute also revealed a role in radioresistance¹⁹⁶. However, ways to measure EMT needed to be adapted for HNSCC and little is known about the relation of EMT to chemoradiotherapy outcome in HNSCC. In **Chapter 3** we investigated several gene expression profiles for EMT created in other cancer types and determine their prognostic value in a large HNSCC cohort. Using machine learning we developed a HNSCC-specific EMT profile which is validated on publicly available cell line data and in an independent cohort.

Given the relevance in cisplatin damage repair, previous studies conducted at our institute by Dr. Verhagen, Dr. Vossen and Dr. Wreesman focused on the relevance of the Fanconi Anemia pathway, an important DNA crosslink repair pathway, in HNSCC¹⁶³. Using HNSCC cell lines, these studies revealed that many HNSCC cell lines have mutations in the Fanconi Anemia pathway. It was also confirmed that cell lines with mutations in the FA/HR pathway, are hypersensitive to the DNA crosslinking agent Mitomycin C and that tumors with mutations acquired in the FA pathway are frequently present in HNSCC patients and associated with a worse prognosis. In **Chapter 4** we use their previously acquired cell line data on sensitivity to DNA crosslinking agent Mitomycin C as a proxy for FA/HR DNA repair defects. With NGS data of these cell lines, transcriptomics and machine learning we developed a gene expression profile to identify cell lines with a high risk of these DNA repair defects. This gene expression profile was tested on two patient cohorts with HPV-negative HNSCC.

In **Chapter 5** we performed a large multivariable analysis that compares the importance and correlation between previously identified markers as described in literature and other important novel markers, such as DNA repair and immunology related. These analyses reveal novel contributors and complement studies in a comparable direction from our institute as carried out by Dr. vd Broek and Dr. Pramana, who used gene expression data from various sources to identify important prognostic biomarkers in HNSCC^{197,198}. Together these studies illustrate the potential of transcriptomic based biomarkers for HNSCC outcome prediction but also illuminate the role of the different biological processes in determining outcomes.

Finally, in **Chapter 6** the results of the previous chapters are placed in context with relation to literature and future perspectives.

References

1. Head and Neck Cancers - NCI. <https://www.cancer.gov/types/head-and-neck/head-neck-fact-sheet>. Accessed May 7, 2022.
2. Sung H, Ferlay J, Siegel RL, *et al*. Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. *CA Cancer J Clin*. 2021;71(3):209-249. doi:10.3322/caac.21660
3. Leemans CR, Braakhuis BJM, Brakenhoff RH. The molecular biology of head and neck cancer. *Nat Rev Cancer*. 2011;11(1):9-22. doi:10.1038/nrc2982
4. Rettig EM, D'Souza G. Epidemiology of Head and Neck Cancer. *Surg Oncol Clin N Am*. 2015;24(3):379-396. doi:10.1016/j.soc.2015.03.001
5. van der Kamp MF, Halmos GB, Guryev V, *et al*. Age-specific oncogenic pathways in head and neck squamous cell carcinoma - are elderly a different subcategory? *Cell Oncol*. 2022;45(1). doi:10.1007/S13402-021-00655-4
6. Auerbach AD. Fanconi Anemia and its Diagnosis. *Mutation Res*. 2010;668(1-2):4-10. doi:10.1016/j.mrfmmm.2009.01.013.Fanconi
7. 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):106-112. doi:10.1001/archotol.129.1.106
8. Friedlander PL. *Genomic Instability in Head and Neck Cancer Patients*. Vol 23.; 2001.
9. Cramer JD, Hicks KE, Rademaker AW, Patel UA, Samant S. Validation of the eighth edition American Joint Committee on Cancer staging system for human papillomavirus-associated oropharyngeal cancer. *Head Neck*. 2018;40(3):457-466. doi:10.1002/hed.24974
10. Lydiatt WM, Patel SG, O'Sullivan B, *et al*. Head and neck cancers-major changes in the American Joint Committee on cancer eighth edition cancer staging manual. *CA Cancer J Clin*. 2017;67(2):122-137. doi:10.3322/caac.21389
11. Deschler DG, Moore MG, Smith R V. *TNM Staging of Head and Neck Cancer and Neck Dissection Classification*. Vol Fourth edi.; 2014.
12. Lin CY, Wang HM, Kang CJ, *et al*. Primary tumor site as a predictor of treatment outcome for definitive radiotherapy of advanced-stage oral cavity cancers. *Int J Radiat Oncol Biol Phys*. 2010;78(4):1011-1019. doi:10.1016/j.IJROBP.2009.09.074
13. 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):518-523. doi:10.1002/HED.23626
14. IKNL. Integraal kanker centrum Nederland - Hoofd-hals kanker. <https://iknl.nl/kankersoorten/hoofd-halskanker>. 2022.
15. Cramer JD, Burtneess B, Le QT, Ferris RL. The changing therapeutic landscape of head and neck cancer. *Nat Rev Clin Oncol*. 2019;16(11):669-683. doi:10.1038/s41571-019-0227-z
16. de Roest RH, van der Heijden M, Wesseling FWR, *et al*. Disease outcome and associated factors after definitive platinum based chemoradiotherapy for advanced stage HPV-negative head and neck cancer. *Radiother Oncol*. 2022;175:112-121. doi:10.1016/j.radonc.2022.08.013
17. Schwartz DL. Current progress in adaptive radiation therapy for head and neck cancer. *Curr Oncol Rep*. 2012;14(2):139-147. doi:10.1007/s11912-012-0221-4
18. Heukelom J, Hamming O, Bartelink H, *et al*. Adaptive and innovative Radiation Treatment FOR improving Cancer treatment outcome (ARTFORCE); a randomized controlled phase II trial for individualized treatment of head and neck cancer. *BMC Cancer*. 2013. doi:10.1186/1471-2407-13-84

CTTCAGACCTGTCTCCCTCATTCAAAAAATATTATTATCAGCTCTTACTTGTACCCAGCACTGATATAGGCCTCAGGAATAACAATGAATAAGATAGTAAAAAATCTATATCTCATAAGGTTACGTTCCATGTACTGAAG

19. Kim JK, Leeman JE, Riaz N, McBride S, Tsai CJ, Lee NY. Proton Therapy for Head and Neck Cancer. *Curr Treat Options Oncol*. 2018;19(6). doi:10.1007/s11864-018-0546-9
20. Pignon JP, Maître A le, Maillard E, Bourhis J. 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. doi:10.1016/j.radonc.2009.04.014
21. Lacas B, Carmel A, Landais C, *et al*. Meta-analysis of chemotherapy in head and neck cancer (MACH-NC): An update on 107 randomized trials and 19,805 patients, on behalf of MACH-NC Group. *Radiother Oncol*. 2021;156:281-293. doi:10.1016/j.radonc.2021.01.013
22. 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 Oto-Rhino-Laryngology*. 2017. doi:10.1007/s00405-017-4687-4
23. Bonner JA, Harari PM, Giralt J, *et al*. Radiotherapy plus cetuximab for locoregionally advanced head and neck cancer: *Lancet Oncol*. 2010. doi:10.1016/S1470-2045(09)70311-0
24. 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. doi:10.1016/S0140-6736(18)32752-1
25. 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. doi:10.1016/S0140-6736(18)32779-X
26. Nauta IH, Klausch T, van de Ven PM, *et al*. The important role of cisplatin in the treatment of HPV-positive oropharyngeal cancer assessed by real-world data analysis. *Oral Oncol*. 2021;121:105454. doi:10.1016/J.ORALONCOLOGY.2021.105454
27. Kian Ang K, Harris J, Wheeler R, *et al*. Human Papillomavirus and Survival of Patients with Oropharyngeal Cancer A BS TR AC T. *N Engl J Med*. 2010;363:24-35. doi:10.1056/NEJMoa0912217
28. De Veij Mestdagh PD, Schreuder WH, Vogel W V, *et al*. Mapping of sentinel lymph node drainage using SPECT/CT to tailor elective nodal irradiation in head and neck cancer patients (SUSPECT-2): a single-center prospective trial. doi:10.1186/s12885-019-6331-8
29. Riaz N, Sherman E, Pei X, *et al*. Precision Radiotherapy: Reduction in Radiation for Oropharyngeal Cancer in the 30 ROC Trial. doi:10.1093/jnci/djaa184
30. Solimeno LS, Park YM, Lim JY, Koh YW, Kim SH. Treatment outcomes of neoadjuvant chemotherapy and transoral robotic surgery in locoregionally advanced laryngopharyngeal carcinoma. *Head Neck*. 2021;43(11):3429-3436. doi:10.1002/HED.26838
31. Vermorken JB, Mesia R, Rivera F, *et al*. *Platinum-Based Chemotherapy plus Cetuximab in Head and Neck Cancer A Bs Tr Ac T*. Vol 359.; 2008. www.nejm.org.
32. Chow LQM, Haddad R, Gupta S, *et al*. Antitumor Activity of Pembrolizumab in Biomarker-Unselected Patients With Recurrent and/or Metastatic Head and Neck Squamous Cell Carcinoma: Results From the Phase Ib KEYNOTE-012 Expansion Cohort. *J Clin Oncol*. 2016;34(32):3838-3845. doi:10.1200/JCO.2016.68.1478
33. Harrington KJ, Frer F, Ferris RL, *et al*. Impact of nivolumab vs standard, single-agent therapy of investigator's choice on patient-reported outcomes in recurrent or metastatic squamous cell carcinoma of the head and neck: health-related quality-of-life results from CheckMate 141, a randomized, phase 3 trial HHS Public Access. *Lancet Oncol*. 2017;18(8):1104-1115. doi:10.1016/S1470-2045(17)30421-7

TGTGAAGGAGATTAAATAGATGGTGTGATATAAGTATCTGGGAGAAAAGCTTAGGGTGTGGATATTAGGAAAGCCTTCCTAAAAATGAATTAACTGATGAGAGAAAGGATCCAGCTGAGAGCAAAAGCAAAAGCTTTCTCTCTTC

34. Vos JL, Elbers JBW, Krijgsman O, *et al.* Neoadjuvant immunotherapy with nivolumab and ipilimumab induces major pathological responses in patients with head and neck squamous cell carcinoma. doi:10.1038/s41467-021-26472-9
35. Massacesi C, di Tomaso E, Urban P, *et al.* PI3K inhibitors as new cancer therapeutics: implications for clinical trial design. *Onco Targets Ther.* 2016;9:203-210. doi:10.2147/OTT.S89967
36. Xu MJ, Johnson DE, Grandis JR. EGFR-targeted therapies in the post-genomic era. *Cancer Metastasis Rev.* 2017;36(3):463-473. doi:10.1007/s10555-017-9687-8
37. De Felice F, Guerrero Urbano T. New drug development in head and neck squamous cell carcinoma: The PI3-K inhibitors. *Oral Oncol.* 2017;67:119-123. doi:10.1016/J.ORALONCOLOGY.2017.02.020
38. Janssens GO, Rademakers SE, Terhaard CH, *et al.* Accelerated radiotherapy with carbogen and nicotinamide for laryngeal cancer: Results of a phase III randomized trial. *J Clin Oncol.* 2012;30(15):1777-1783. doi:10.1200/JCO.2011.35.9315
39. Overgaard J, Hansen HS, Overgaard M, *et al.* A randomized double-blind phase III study of nimorazole as a hypoxic radiosensitizer of primary radiotherapy in supraglottic larynx and pharynx carcinoma. Results of the Danish Head and Neck Cancer Study (DAHANCA) Protocol 5-85. *Radiother Oncol.* 1998;46(2):135-146. doi:10.1016/S0167-8140(97)00220-X
40. Dohmen AJC, Qiao X, Duursma A, *et al.* Identification of a novel ATM inhibitor with cancer cell specific radiosensitization activity. *Oncotarget.* 2017;8(43):73925. doi:10.18632/ONCOTARGET.18034
41. De Haan R, Van Werkhoven E, Van Den Heuvel MM, *et al.* Study protocols of three parallel phase 1 trials combining radical radiotherapy with the PARP inhibitor olaparib. *BMC Cancer.* 2019;19(1). doi:10.1186/S12885-019-6121-3
42. Timmermans AJ, van Dijk BAC, Overbeek LIH, *et al.* Trends in treatment and survival for advanced laryngeal cancer: A 20-year population-based study in The Netherlands. *Head Neck.* 2015. doi:10.1002/hed.24200
43. Petersen JF, Timmermans AJ, van Dijk BAC, *et al.* Trends in treatment, incidence and survival of hypopharynx cancer: a 20-year population-based study in the Netherlands. *Eur Arch Oto-Rhino-Laryngology.* 2018;275(1):181-189. doi:10.1007/s00405-017-4766-6
44. Knegjens JL, Hauptmann M, Pameijer FA, *et al.* Tumor volume as prognostic factor in chemoradiation for advanced head and neck cancer. *Head Neck.* 2011;33(3):375-382. doi:10.1002/hed.21459
45. Hoebbers FJP, Pameijer FA, De Bois J, *et al.* Prognostic value of primary tumor volume after concurrent chemoradiation with daily low-dose cisplatin for advanced-stage head and neck carcinoma. *Head Neck.* 2008;30(9):1216-1223. doi:10.1002/HED.20865
46. Bøje CR, Dalton SO, Grønberg TK, *et al.* The impact of comorbidity on outcome in 12 623 Danish Head and Neck Cancer Patients : A population based study from the DAHANCA database The impact of comorbidity on outcome in 12 623 Danish Head and Neck Cancer Patients : A population based study from th. 2013:284-293. doi:10.3109/0284186X.2012.742964
47. Hu K, Harrison LB. Impact of Anemia in Patients with Head and Neck Cancer Treated with Radiation Therapy. *Curr Treat Options Oncol.* 2005;6:31-45.
48. Mascarella MA, Mannard E, Silva SD, Zeitouni A. Neutrophil-to-lymphocyte ratio in head and neck cancer prognosis: A systematic review and meta-analysis. *Head Neck.* 2018;40(5):1091-1100. doi:10.1002/hed.25075
49. Gronhøj C, Jensen JS, Wagner S, *et al.* Impact on survival of tobacco smoking for cases with oropharyngeal squamous cell carcinoma and known human papillomavirus and p16-status: a multicenter retrospective study. *Oncotarget.* 2019;10(45):4655-4663. doi:10.18632/oncotarget.27079

50. Becker A, Stadler P, Lavey RS, *et al.* Severe anemia is associated with poor tumor oxygenation in head and neck squamous cell carcinomas. *Int J Radiat Oncol Biol Phys.* 2000;46(2):459-466. doi:10.1016/S0360-3016(99)00384-3
51. Hanahan D, Weinberg RA. The hallmarks of cancer. *Cell.* 2000;100(1):57-70. doi:10.1007/s00262-010-0968-0
52. Hanahan D. Hallmarks of Cancer: New Dimensions. *Cancer Discov.* 2022;12(1):31-46. doi:10.1158/2159-8290.CD-21-1059
53. Leemans CR, Snijders PJF, Brakenhoff RH. The molecular landscape of head and neck cancer. *Nat Rev Cancer.* 2018;18(5):269-282. doi:10.1038/nrc.2018.11
54. Lorz C, Segrelles C, Errazquin R, Garcia-Escudero R. Comprehensive Molecular Characterization of Squamous Cell Carcinomas. In: *Squamous Cell Carcinoma - Hallmark and Treatment Modalities.* ; 2020:1-12. doi:10.5772/intechopen.85988
55. Lawrence MS, Sougnez C, Lichtenstein L, *et al.* Comprehensive genomic characterization of head and neck squamous cell carcinomas. *Nature.* 2015;517(7536):576-582. doi:10.1038/nature14129
56. Leemans CR, Braakhuis BJM, Brakenhoff RH. The molecular biology of head and neck cancer. *Nat Rev Cancer.* 2011;11(1):9-22. doi:10.1038/nrc2982
57. Stransky N, Egloff AM, Tward AD, *et al.* The mutational landscape of head and neck squamous cell carcinoma. *Science (80-).* 2011;333(6046):1157-1160. doi:10.1126/science.1208130
58. Agrawal N, Frederick MJ, Pickering CR, *et al.* Exome sequencing of head and neck squamous cell carcinoma reveals inactivating mutations in NOTCH1. *Science (80-).* 2011;333(6046):1154-1157. doi:10.1126/science.1206923
59. Chen WS, Bindra RS, Mo A, *et al.* CDKN2A copy number loss is an independent prognostic factor in HPV-negative head and neck squamous cell carcinoma. *Front Oncol.* 2018;8(APR). doi:10.3389/fonc.2018.00095
60. 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):632-641. doi:10.1158/1078-0432.CCR-13-3310
61. Rowinski E, Magné N, Fayette J, *et al.* Radioresistance and genomic alterations in head and neck squamous cell cancer: Sub-analysis of the ProFiLER protocol. *Head Neck.* 2021;43(12):3899-3910. doi:10.1002/hed.26891
62. Comprehensive genomic characterization of head and neck squamous cell carcinomas. *Nature.* 2015;517(7536):576-582. doi:10.1038/nature14129
63. GRAY LH, CONGER AD, EBERT M, HORNSEY S, SCOTT OC. The concentration of oxygen dissolved in tissues at the time of irradiation as a factor in radiotherapy. *Br J Radiol.* 1953. doi:10.1259/0007-1285-26-312-638
64. WRIGHT EA, HOWARD-FLANDERS P. The influence of oxygen on the radiosensitivity of mammalian tissues. *Acta radiol.* 1957. doi:10.1177/028418515704800105
65. Ma NY, Tinganelli W, Maier A, Durante M, Kraft-Weyrather W. Influence of chronic hypoxia and radiation quality on cell survival. *J Radiat Res.* 2013. doi:10.1093/jrr/rrs135
66. Janssen HL, Haustermans KM, Balm AJ, Begg AC. Hypoxia in head and neck cancer: How much, how important? *Head Neck.* 2005. doi:10.1002/hed.20223
67. Nordsmark M, Overgaard J. A confirmatory prognostic study on oxygenation status and loco-regional control in advanced head and neck squamous cell carcinoma treated by radiation therapy. *Radiation Oncol.* 2000. doi:10.1016/S0167-8140(00)00223-1

TGTGAAGGAGATTAAATAGATGGTGTGATATAAGATCTTGGGAGAAAAGCTTAGGGTGTGGATATTAAGAAAGCTTCTCTAAAAATGAATTAACTGATGAGAAGAAAGGATCCAGCTGAGAGCAAAAGCAAAAGCTTTCTCTCTTC

68. Fleming IN, Manavaki R, Blower PJ, *et al.* Imaging tumour hypoxia with positron emission tomography. *Br J Cancer*. 2015. doi:10.1038/bjc.2014.610
69. Huizing J, Garousi J, Lok J, *et al.* cAix-targeting radiotracers for hypoxia imaging in head and neck cancer models. doi:10.1038/s41598-019-54824-5
70. Koukourakis MI, Giatromanolaki A, Sivridis E, *et al.* Hypoxia-regulated carbonic anhydrase-9 (CA9) relates to poor vascularization and resistance of squamous cell head and neck cancer to chemoradiotherapy. *Clin Cancer Res*. 2001. doi:10.1158/1078-0432.ccr-13-0542
71. Schrijvers ML, van der Laan BFAM, de Bock GH, *et al.* Overexpression of Intrinsic Hypoxia Markers HIF1 α and CA-IX Predict for Local Recurrence in Stage T1-T2 Glottic Laryngeal Carcinoma Treated With Radiotherapy. *Int J Radiat Oncol Biol Phys*. 2008. doi:10.1016/j.ijrobp.2008.05.025
72. Aebbersold DM, Burri P, Beer KT, *et al.* Expression of hypoxia-inducible factor-1 α : A novel predictive and prognostic parameter in the radiotherapy of oropharyngeal cancer. *Cancer Res*. 2001.
73. Rademakers SE, Hoogsteen IJ, Rijken PF, *et al.* Pattern of CAIX expression is prognostic for outcome and predicts response to ARCON in patients with laryngeal cancer treated in a phase III randomized trial. *Radiother Oncol*. 2013. doi:10.1016/j.radonc.2013.04.022
74. Janssen HLK, Haustermans KMG, Sprong D, *et al.* HIF-1 α , pimonidazole, and iododeoxyuridine to estimate hypoxia and perfusion in human head-and-neck tumors. *Int J Radiat Oncol Biol Phys*. 2002. doi:10.1016/S0360-3016(02)03935-4
75. Begg AC, Janssen H, Sprong D, *et al.* Hypoxia and Perfusion Measurements in Human Tumors&Initial Experience with Pimonidazole and IUdR. *Acta Oncol (Madr)*. 2001;40(8):924-928. doi:10.1080/02841860152708198
76. Kaanders JH, Wijffels KI, Marres HA, *et al.* Pimonidazole binding and tumor vascularity predict for treatment outcome in head and neck cancer. *Cancer Res*. 2002.
77. Eustace A, Mani N, Span PN, *et al.* A 26-gene hypoxia signature predicts benefit from hypoxia-modifying therapy in laryngeal cancer but not bladder cancer. *Clin Cancer Res*. 2013. doi:10.1158/1078-0432.CCR-13-0542
78. Seigneuric R, Starmans MHW, Fung G, *et al.* Impact of supervised gene signatures of early hypoxia on patient survival. *Radiother Oncol*. 2007. doi:10.1016/j.radonc.2007.05.002
79. 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. doi:10.1158/0008-5472.CAN-06-3322
80. Toustrup K, Sørensen BS, Nordsmark M, *et al.* Development of a hypoxia gene expression classifier with predictive impact for hypoxic modification of radiotherapy in head and neck cancer. *Cancer Res*. 2011. doi:10.1158/0008-5472.CAN-11-1182
81. Buffa FM, Harris AL, West CM, Miller CJ. Large meta-analysis of multiple cancers reveals a common, compact and highly prognostic hypoxia metagene. *Br J Cancer*. 2010. doi:10.1038/sj.bjc.6605450
82. Nordsmark M, Overgaard M, Overgaard J. Pretreatment oxygenation predicts radiation response in advanced squamous cell carcinoma of the head and neck. *Radiother Oncol*. 1996. doi:10.1016/S0167-8140(96)91811-3
83. Nordsmark M, Bentzen SM, Rudat V, *et al.* Prognostic value of tumor oxygenation in 397 head and neck tumors after primary radiation therapy. An international multi-center study. *Radiother Oncol*. 2005. doi:10.1016/j.radonc.2005.06.038
84. De Schutter H, Landuyt W, Verbeken E, Goethals L, Hermans R, Nuyts S. The prognostic value of the hypoxia markers CA IX and GLUT I and the cytokines VEGF and IL 6 in head and neck squamous cell carcinoma treated by radiotherapy \pm chemotherapy. *BMC Cancer*. 2005. doi:10.1186/1471-2407-5-42

CTTCAGACCTGTCCTCCCTCATTTCAAAAAATATTATTATGAGCTCTTACTTGCTACCCAGCACTGATATAGGCACTCAGGAATAACAATGAATAGATAGTAAAAAATCTATATCCTCATAAGGTTACGTTCCATGTACTGAAG

85. Tawk B, Schwager C, Deffaa O, *et al.* Comparative analysis of transcriptomics based hypoxia signatures in head- and neck squamous cell carcinoma. *Radiother Oncol.* 2016;118(2):350-358. doi:10.1016/j.radonc.2015.11.027
86. Roos WP, Thomas AD, Kaina B. DNA damage and the balance between survival and death in cancer biology. 2016. doi:10.1038/nrc.2015.2
87. Negrini S, Gorgoulis VG, Halazonetis TD. Genomic instability an evolving hallmark of cancer. *Nat Rev Mol Cell Biol.* 2010;11(3):220-228. doi:10.1038/nrm2858
88. Kandoth C, McLellan MD, Vandin F, *et al.* Mutational landscape and significance across 12 major cancer types. *Nature.* 2013;502(7471):333-339. doi:10.1038/nature12634
89. Alexandrov LB, Nik-Zainal S, Wedge DC, *et al.* Signatures of mutational processes in human cancer. *Nature.* 2013;500(7463):415-421. doi:10.1038/nature12477
90. Rezaee M, Sanche L, Hunting DJ. Cisplatin enhances the formation of DNA single- and double-strand breaks by hydrated electrons and hydroxyl radicals. *Radiat Res.* 2013;179(3):323-331. doi:10.1667/RR3185.1
91. Seiwert TY, Salama JK, Vokes EE. The concurrent chemoradiation paradigm - General principles. *Nat Clin Pract Oncol.* 2007;4(2):86-100. doi:10.1038/NCPONC0714
92. Chenoufi N, Raoul J-L, Lescoat G, Brissot P, Bourguet P. In Vitro Demonstration of Synergy Between Radionuclide and Chemotherapy. *J Nucl Med.* 1998;39(5).
93. Jenkins G, O'Byrne KJ, Panizza B, Richard DJ. Genome stability pathways in head and neck cancers. *Int J Genomics.* 2013;2013. doi:10.1155/2013/464720
94. Kalluri R, Weinberg R a. Review series The basics of epithelial-mesenchymal transition. *J Clin Invest.* 2009;119(6):1420-1428. doi:10.1172/JCI39104.1420
95. Baumeister P, Zhou J, Canis M, Gires O. Epithelial-to-mesenchymal transition-derived heterogeneity in head and neck squamous cell carcinomas. *Cancers (Basel).* 2021;13(21). doi:10.3390/CANCERS13215355
96. Gibbons DL, Creighton CJ. Pan-cancer survey of epithelial-mesenchymal transition markers across the Cancer Genome Atlas. *Dev Dyn.* 2018;247(3):555-564. doi:10.1002/dvdy.24485
97. Nijkamp MM, Span PN, Hoogsteen IJ, Van Der Kogel AJ, Kaanders JHAM, Bussink J. Expression of E-cadherin and vimentin correlates with metastasis formation in head and neck squamous cell carcinoma patients. *Radiother Oncol.* 2011;99(3):344-348. doi:10.1016/j.radonc.2011.05.066
98. Zhu G, Song P, Zhou H, *et al.* Role of epithelial-mesenchymal transition markers E-cadherin, N-cadherin, β -catenin and ZEB2 in laryngeal squamous cell carcinoma. *Oncol Lett.* January 2018. doi:10.3892/ol.2018.7751
99. Luo Y, Yu T, Zhang Q, *et al.* Upregulated N-cadherin expression is associated with poor prognosis in epithelial-derived solid tumours: A meta-analysis. *Eur J Clin Invest.* 2018;(January):e12903. doi:10.1111/eci.12903
100. Zhang P, Hu P, Shen H, Yu J, Liu Q, Du J. Prognostic role of Twist or Snail in various carcinomas: A systematic review and meta-analysis. *Eur J Clin Invest.* 2014;44(11):1072-1094. doi:10.1111/eci.12343
101. Lefevre M, Rousseau A, Rayon T, *et al.* Epithelial to mesenchymal transition and HPV infection in squamous cell oropharyngeal carcinomas: The papillophar study. *Br J Cancer.* 2017;116(3):362-369. doi:10.1038/bjc.2016.434
102. Göppel J, Möckelmann N, Münscher A, Sauter G, Schumacher U. Expression of Epithelial-Mesenchymal Transition Regulating Transcription Factors in Head and Neck Squamous Cell Carcinomas. *Anticancer Res.* 2017;37(10):5435-5440. doi:10.21873/anticancer.11971

TGTGAAGAGATTAAATAAGATGGTGTGATATAAGTATCTGGAGAAAACTTAGGGTGTGGATATTAAGAAAGCCTTCTTAAAAATGAATTTAACTGATGAGAAGAAAGGATCCAGCTGAGAGCAAAAGCAAAAGCTTTCTCTCTTC

103. Shaw R, Beasley N. Aetiology and risk factors for head and neck cancer: United Kingdom National Multidisciplinary Guidelines. *J Laryngol Otol.* 2016;130(S2):S9-S12. doi:10.1017/s0022215116000360
104. Wan Y, Liu H, Zhang M, *et al.* Prognostic value of epithelial-mesenchymal transition-inducing transcription factors in head and neck squamous cell carcinoma: A meta-analysis. *Head Neck.* 2020;42(5):1067-1076. doi:10.1002/HED.26104
105. Chung CH, Parker JS, Ely K, *et al.* Gene expression profiles identify epithelial-to-mesenchymal transition and activation of nuclear factor- κ B Signaling as characteristics of a high-risk head and neck squamous cell carcinoma. *Cancer Res.* 2006;66(16):8210-8218. doi:10.1158/0008-5472.CAN-06-1213
106. Mak MP, Tong P, Diao L, *et al.* A Patient-Derived, Pan-Cancer EMT Signature Identifies Global Molecular Alterations and Immune Target Enrichment Following Epithelial-to-Mesenchymal Transition. *Clin Cancer Res.* 2016;22(3):609-620. doi:10.1158/1078-0432.CCR-15-0876
107. Mandal R, Şenbabaoğlu Y, Desrichard A, *et al.* The head and neck cancer immune landscape and its immunotherapeutic implications. *JCI Insight.* 2016;1(17):1-18. doi:10.1172/jci.insight.89829
108. Freeman GJ, Long AJ, Iwai Y, *et al.* Engagement of the PD-1 Immunoinhibitory Receptor by a Novel B7 Family Member Leads to Negative Regulation of Lymphocyte Activation. Vol 192.; 2000. <http://www.jem.org/cgi/content/full/192/7/1027>.
109. Ferris RL, Blumenschein GJ, Fayette J, *et al.* Nivolumab for Recurrent Squamous-Cell Carcinoma of the Head and Neck. *N Engl J Med.* 2016;375(19):1856-1867. doi:10.1056/NEJMoa1602252
110. 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 (London, England).* 2019;393(10167):156-167. doi:10.1016/S0140-6736(18)31999-8
111. Lee NY, Ferris RL, Psyrri A, *et al.* Avelumab plus standard-of-care chemoradiotherapy versus chemoradiotherapy alone in patients with locally advanced squamous cell carcinoma of the head and neck: a randomised, double-blind, placebo-controlled, multicentre, phase 3 trial. *Lancet Oncol.* 2021;22(4):450-462. doi:10.1016/S1470-2045(20)30737-3
112. Chowell D, Yoo SK, Valero C, *et al.* Improved prediction of immune checkpoint blockade efficacy across multiple cancer types. *Nat Biotechnol.* 2021. doi:10.1038/s41587-021-01070-8
113. Hoesseini A, Van Leeuwen N, Marinella J, *et al.* Predicting survival in head and neck cancer: External validation and update of the prognostic model OncologIQ in 2189 patients. 2021. doi:10.1002/hed.26716
114. Cancer Genome Atlas Research Network, Weinstein JN, Collisson EA, *et al.* The Cancer Genome Atlas Pan-Cancer analysis project. *Nat Genet.* 2013;45(10):1113-1120. doi:10.1038/ng.2764
115. Lechner M, Frampton GM, Fenton T, *et al.* Targeted next-generation sequencing of head and neck squamous cell carcinoma identifies novel genetic alterations in HPV+ and HPV- tumors. *Genome Med.* 2013;5(5):49. doi:10.1186/gm453
116. Pickering CR, Zhang J, Yoo SY, *et al.* Integrative genomic characterization of oral squamous cell carcinoma identifies frequent somatic drivers. *Cancer Discov.* 2013;3(7):770-781. doi:10.1158/2159-8290.CD-12-0537
117. Hedberg ML, Goh G, Chiosea SI, *et al.* Genetic landscape of metastatic and recurrent head and neck squamous cell carcinoma. *J Clin Invest.* 2015;126(6):169-180. doi:10.1172/JCI82066DS1
118. Lui VWY, Hedberg ML, Li H, *et al.* Frequent mutation of the PI3K pathway in head and neck cancer defines predictive biomarkers. *Cancer Discov.* 2013;3(7):761-769. doi:10.1158/2159-8290.CD-13-0103

CTTCAGACCTGTCCTCCCTCATTTCAAAAAATATTATTATCTAGCTCTTACTTGCTACCCAGCACTGATATAGGCATCTAGGAATAACAATGAATAGATAGTAAAAAATCTATATCTCTATAAGCTTACGTTCTCATGTACTGAAG

119. Puram S V, Tirosh I, Parikh AS, *et al.* Single-Cell Transcriptomic Analysis of Primary and Metastatic Tumor Ecosystems in Head and Neck Cancer. *Cell*. 2017;172:1-14. doi:10.1016/j.cell.2017.10.044
120. 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):530-536. doi:10.1038/415530a
121. Hoffmann M, Tribius S. HPV and Oropharyngeal Cancer in the Eighth Edition of the TNM Classification: Pitfalls in Practice. *Transl Oncol*. 2019;12(8):1108-1112. doi:10.1016/J.TRANON.2019.05.009
122. Nauta IH, Rietbergen MM, van Bokhoven AAJD, *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):1273-1279. doi:10.1093/ANNONC/MDY060
123. Van Monsjou HS, Van Velthuysen MLF, Van Den Brekel MWM, Jordanova ES, Melief CJM, Balm AJM. Human papillomavirus status in young patients with head and neck squamous cell carcinoma. doi:10.1002/ijc.26195
124. Starmans MHW, Krishnapuram B, Steck H, *et al.* Robust prognostic value of a knowledge-based proliferation signature across large patient microarray studies spanning different cancer types. *Br J Cancer*. 2008. doi:10.1038/sj.bjc.6604746
125. Romick-Rosendale LE, Lui VWY, Grandis JR, Wells SI. The Fanconi anemia pathway: Repairing the link between DNA damage and squamous cell carcinoma. *Mutat Res - Fundam Mol Mech Mutagen*. 2013;743-744:78-88. doi:10.1016/j.mrfmmm.2013.01.001
126. Tweardy DJ. Elevated Levels of Transforming Growth Factor α and Epidermal Growth Factor Receptor Messenger RNA Are Early Markers of Carcinogenesis in Head and Neck Cancer. *Cancer Res*. 1993;53(15):3579-3584. <http://aacrjournals.org/cancerres/article-pdf/53/15/3579/2451248/cr0530153579.pdf>. Accessed April 17, 2022.
127. Nicholson R., Gee JM., Harper M. EGFR and cancer prognosis. *Eur J Cancer*. 2001. doi:10.1016/s0959-8049(01)00231-3
128. Demiral AN, Sarioglu S, Birlik B, Sen M, Kinay M. Prognostic significance of EGF receptor expression in early glottic cancer. *Auris Nasus Larynx*. 2004;31(4):417-424. doi:10.1016/J.ANL.2004.05.003
129. Maurizi M, Almadori G, Ferrandina G, *et al.* Prognostic significance of epidermal growth factor receptor in laryngeal squamous cell carcinoma. *Br J Cancer* 1996 748. 1996;74(8):1253-1257. doi:10.1038/bjc.1996.525
130. Ang KK, Berkey BA, Tu X, *et al.* Impact of epidermal growth factor receptor expression on survival and pattern of relapse in patients with advanced head and neck carcinoma. *Cancer Res*. 2002.
131. Bossi P, Resteghini C, Paielli N, Licitra L, Pilotti S, Perrone F. Prognostic and predictive value of EGFR in head and neck squamous cell carcinoma. *Oncotarget*. 2016. doi:10.18632/oncotarget.11413
132. Numico G, Russi EG, Colantonio I, *et al.* EGFR status and prognosis of patients with locally advanced head and neck cancer treated with chemoradiotherapy. *Anticancer Res*. 2010;30(2):671-676.
133. Rahimi AS, Wilson DD, Saylor DK, *et al.* p16, Cyclin D1, and HIF-1 α Predict Outcomes of Patients with Oropharyngeal Squamous Cell Carcinoma Treated with Definitive Intensity-Modulated Radiation Therapy. *Int J Otolaryngol*. 2012;2012. doi:10.1155/2012/685951
134. Bossi P, Resteghini C, Paielli N, Licitra L, Pilotti S, Perrone F. *Prognostic and Predictive Value of EGFR in Head and Neck Squamous Cell Carcinoma*. Vol 7.; 2016. www.impactjournals.com/oncotarget.
135. Young RJ, Rischin D, Fisher R, *et al.* Relationship between Epidermal Growth Factor Receptor Status, p16 INK4A, and Outcome in Head and Neck Squamous Cell Carcinoma. doi:10.1158/1055-9965.EPI-10-1262

TGTGAAGAGATTAAATAAGATGGTGTGATATAAGTATCGGAGAAAACTTAGGGTGTGGATATTAAGAAAGCCTTCCTAAAAATGAATTAACTGATGAGAAGAAAGGATCCAGCTGAGAGCAAACTAAAGCTTTCTCCTTC

136. Eriksen JG, Steiniche T, Askaa J, Alsner J, Overgaard J. The prognostic value of epidermal growth factor receptor is related to tumor differentiation and the overall treatment time of radiotherapy in squamous cell carcinomas of the head and neck. *Int J Radiat Oncol Biol Phys*. 2004;58(2):561-566. doi:10.1016/J.IJROBP.2003.09.043
137. Pedicini P, Nappi A, Strigari L, *et al*. Correlation between egfr expression and accelerated proliferation during radiotherapy of head and neck squamous cell carcinoma. 2012. doi:10.1186/1748-717X-7-143
138. Bentzen SM, Atasoy BM, Daley FM, *et al*. Epidermal growth factor receptor expression in pretreatment biopsies from head and neck squamous cell carcinoma as a predictive factor for a benefit from accelerated radiation therapy in a randomized controlled trial. *J Clin Oncol*. 2005. doi:10.1200/JCO.2005.06.411
139. Ang KK. Multidisciplinary Management of Locally Advanced SCCN: Optimizing Treatment Outcomes. *Oncologist*. 2008. doi:10.1634/theoncologist.2007-0157
140. Bossi P, Bergamini C, Siano M, *et al*. Functional genomics uncover the biology behind the responsiveness of head and neck squamous cell cancer patients to cetuximab. *Clin Cancer Res*. 2016;22(15):3961-3970. doi:10.1158/1078-0432.CCR-15-2547
141. Bussink J, van der Kogel AJ, Kaanders JH. Activation of the PI3-K/AKT pathway and implications for radioresistance mechanisms in head and neck cancer. *Lancet Oncol*. 2008;9(3):288-296. doi:10.1016/S1470-2045(08)70073-1
142. Qiu W, Schönleben F, Li X, *et al*. PIK3CA Mutations in Head and Neck Squamous Cell Carcinoma. doi:10.1158/1078-0432.CCR-05-2173
143. Bernichon E, Vallard A, Wang Q, *et al*. Genomic alterations and radioresistance in breast cancer: An analysis of the ProfILER protocol. *Ann Oncol*. 2017;28(11):2773-2779. doi:10.1093/ANNONC/MDX488
144. He Y, Van't Veer LJ, Lopez-Yurda M, Van De Velde CJH, Marijnen CAM. Do Rectal Cancer Patients with PIK3CA Mutations Benefit from Preoperative Radiotherapy with Regard to Local Recurrences? doi:10.1158/1078-0432.CCR-10-1437
145. Hsieh JC-H, Wang H-M, Wu M-H, *et al*. Review of emerging biomarkers in head and neck squamous cell carcinoma in the era of immunotherapy and targeted therapy. *Head Neck*. 2019;41 Suppl 1:19-45. doi:10.1002/hed.25932
146. Wang J, Chai Y-L, Wang T, Liu J-H, Dai P-G, Liu Z. Genetic alterations of PIK3CA and tumor response in patients with locally advanced cervical squamous cell carcinoma treated with cisplatin-based concurrent chemoradiotherapy. 2015. doi:10.1016/j.yexmp.2015.03.014
147. McIntyre JB, Wu JS, Craighead PS, *et al*. PIK3CA mutational status and overall survival in patients with cervical cancer treated with radical chemoradiotherapy. 2012. doi:10.1016/j.ygyno.2012.12.019
148. Linge A, Lock S, Gudziol V, *et al*. Low cancer stem cell marker expression and low hypoxia identify good prognosis subgroups in HPV(-) HNSCC after postoperative radiochemotherapy: A multicenter study of the DTKK-ROG. *Clin Cancer Res*. 2016;22(11):2639-2649. doi:10.1158/1078-0432.CCR-15-1990
149. Toustrup K, Sørensen BS, Metwally MAH, *et al*. Validation of a 15-gene hypoxia classifier in head and neck cancer for prospective use in clinical trials. *Acta Oncol (Madr)*. 2016. doi:10.3109/0284186X.2016.1167959
150. Huang Y, Fan J, Li Y, Fu S, Chen Y, Wu J. Imaging of Tumor Hypoxia With Radionuclide-Labeled Tracers for PET. doi:10.3389/fonc.2021.731503

CTTCAGACCTGTCTCCCTCATTCAAAAAATATTATTATCAGCTCTTACTTGCTACCCAGCACTGATATAGGCACCTCAGGAATAACAATGAATAAGATAGTAGAAAAATCTATATCTCATAAGGTTACGTTCCATGTACTGAAG

151. Koukourakis MI, Giatromanolaki A, Sivridis E, *et al.* Hypoxia-inducible factor (HIF1A and HIF2A), angiogenesis, and chemoradiotherapy outcome of squamous cell head-and-neck cancer. *Int J Radiat Oncol Biol Phys.* 2002. doi:10.1016/S0360-3016(02)02848-1
152. Lefevre M, Rousseau A, Rayon T, *et al.* Epithelial to mesenchymal transition and HPV infection in squamous cell oropharyngeal carcinomas: The papillophar study. *Br J Cancer.* 2017. doi:10.1038/bjc.2016.434
153. Reed E. Platinum-DNA adduct, nucleotide excision repair and platinum based anti-cancer chemotherapy. *Cancer Treat Rev.* 1998;24(5):331-344. doi:10.1016/S0305-7372(98)90056-1
154. Jun HJ, Ahn MJ, Kim HS, *et al.* ERCC1 expression as a predictive marker of squamous cell carcinoma of the head and neck treated with cisplatin-based concurrent chemoradiation. *Br J Cancer.* 2008;99:167-172. doi:10.1038/sj.bjc.6604464
155. Handra-Luca A, Hernandez J, Mountzios G, *et al.* Excision Repair Cross Complementation Group1Immunohistochemical Expression Predicts Objective Response and Cancer-Specific Survival in Patients Treated by Cisplatin-Based Induction Chemotherapy for Locally Advanced Head and Neck Squamous Cell Carcinoma. doi:10.1158/1078-0432.CCR-07-0252
156. Chiu T-J, Chen C-H, Chien C-Y, Li S-H, Tsai H-T, Chen Y-J. *High ERCC1 Expression Predicts Cisplatin-Based Chemotherapy Resistance and Poor Outcome in Unresectable Squamous Cell Carcinoma of Head and Neck in a Betel-Chewing Area.* Vol 9.; 2011. doi:10.1186/1479-5876-9-31
157. Moutafi M, Economopoulou P, Rimm D, Psyrri A. PARP inhibitors in head and neck cancer: Molecular mechanisms, preclinical and clinical data. *Oral Oncol.* 2021;117:105292. doi:10.1016/j.oraloncology.2021.105292
158. de Jong MC, Pramana J, Kneijens JL, *et al.* HPV and high-risk gene expression profiles predict response to chemoradiotherapy in head and neck cancer, independent of clinical factors. *Radiother Oncol.* 2010;95(3):365-370. doi:10.1016/j.radonc.2010.02.001
159. Chung CH, Guthrie VB, Masica DL, *et al.* Genomic alterations in head and neck squamous cell carcinoma determined by cancer gene-targeted sequencing. *Ann Oncol.* 2015;26(6):1216-1223. doi:10.1093/ANNONC/MDV109
160. Chandrasekharappa SC, Chinn SB, Donovan FX, *et al.* Assessing the spectrum of germline variation in Fanconi anemia genes among patients with head and neck carcinoma before age 50. *Cancer.* 2017;123(20):3943-3954. doi:10.1002/cncr.30802
161. Marsit CJ, Liu M, Nelson HH, Posner M, Suzuki M, Kelsey KT. Inactivation of the Fanconi anemia/BRCA pathway in lung and oral cancers: implications for treatment and survival. *Oncogene.* 2004;23:1000-1004. doi:10.1038/sj.onc.1207256
162. Bhide SA, Thway K, Lee J, *et al.* Delayed DNA double-strand break repair following platin-based chemotherapy predicts treatment response in head and neck squamous cell carcinoma. *Br J Cancer.* 2016;115(7):825-830. doi:10.1038/bjc.2016.266
163. Verhagen CVM, Vossen DM, Borgmann K, *et al.* Fanconi anemia and homologous recombination gene variants are associated with functional DNA repair defects in vitro and poor outcome in patients with advanced head and neck squamous cell carcinoma. *Oncotarget.* 2018;9(26):18198-18213. doi:10.18632/oncotarget.24797
164. Bryant HE, Schultz N, Thomas HD, *et al.* *Specific Killing of BRCA2-Deficient Tumours with Inhibitors of Poly(ADP-Ribose) Polymerase.*; 2005. www.nature.com/nature.
165. Konstantinopoulos PA, Lheureux S, Moore KN. PARP Inhibitors for Ovarian Cancer: Current Indications, Future Combinations, and Novel Assets in Development to Target DNA Damage Repair. *Am Soc Clin Oncol Educ book Am Soc Clin Oncol Annu Meet.* 2020;40(40):1-16. doi:10.1200/EDBK_288015

TGTGAAGGAGATTAAATAAGATGGTGTGATATAAGTATTGGGAGAAAAGCTTAGGGTGTGGATATTAAGAAAGCCTTCCTAAAAATGAATTAACTGATGAGAAGAAAGGATCCAGCTGAGAGAAAGGAAAAAGCTTTCTTCCTTC

166. Sakogawa K, Aoki Y, Misumi K, *et al.* Involvement of homologous recombination in the synergism between cisplatin and poly (ADP-ribose) polymerase inhibition. *Cancer Sci.* 2013;104(12):1593-1599. doi:10.1111/cas.12281
167. Molkenkintine JM, Molkenkintine DP, Bridges KA, *et al.* Targeting DNA damage response in head and neck cancers through abrogation of cell cycle checkpoints. *Int J Radiat Biol.* 2021;97(8):1121-1128. doi:10.1080/09553002.2020.1730014
168. Balermipas P, Rodel F, Rodel C, *et al.* CD8+ tumour-infiltrating lymphocytes in relation to HPV status and clinical outcome in patients with head and neck cancer after postoperative chemoradiotherapy: A multicentre study of the German cancer consortium radiation oncology group (DKTK-ROG). *Int J cancer.* 2016;138(1):171-181. doi:10.1002/ijc.29683
169. Ono T, Azuma K, Kawahara A, *et al.* Pre-treatment CD8(+) tumour-infiltrating lymphocyte density predicts distant metastasis after definitive treatment in patients with stage III/IV hypopharyngeal squamous cell carcinoma. *Clin Otolaryngol.* 2018;43(5):1312-1320. doi:10.1111/coa.13171
170. De Meulenaere A, Vermassen T, Aspeslagh S, *et al.* Tumor PD-L1 status and CD8(+) tumor-infiltrating T cells: markers of improved prognosis in oropharyngeal cancer. *Oncotarget.* 2017;8(46):80443-80452. doi:10.18632/oncotarget.19045
171. De Meulenaere A, Vermassen T, Aspeslagh S, *et al.* Prognostic markers in oropharyngeal squamous cell carcinoma: focus on CD70 and tumour infiltrating lymphocytes. *Pathology.* 2017;49(4):397-404. doi:10.1016/j.pathol.2017.02.002
172. Chen H yan, Xu L, Li L feng, Liu X xing, Gao J xin, Bai Y rui. Inhibiting the CD8 + T cell infiltration in the tumor microenvironment after radiotherapy is an important mechanism of radioresistance. *Sci Rep.* 2018. doi:10.1038/s41598-018-30417-6
173. Gupta A, Probst HC, Vuong V, *et al.* Radiotherapy Promotes Tumor-Specific Effector CD8 + T Cells via Dendritic Cell Activation . *J Immunol.* 2012. doi:10.4049/jimmunol.1200563
174. Van Limbergen EJ, De Ruyscher DK, Pimentel VO, *et al.* Combining radiotherapy with immunotherapy: The past, the present and the future. *Br J Radiol.* 2017. doi:10.1259/bjr.20170157
175. Swart M, Verbrugge I, Beltman JB. Combination Approaches with Immune-Checkpoint Blockade in Cancer Therapy. *Front Oncol.* 2016;6:233. doi:10.3389/fonc.2016.00233
176. Ngwa W, Irabor OC, Schoenfeld JD, Hesser J, Demaria S, Formenti SC. Using immunotherapy to boost the abscopal effect. *Nat Rev Cancer.* 2018. doi:10.1038/nrc.2018.6
177. Bristow RG, Alexander B, Baumann M, *et al.* Combining precision radiotherapy with molecular targeting and immunomodulatory agents: a guideline by the American Society for Radiation Oncology. *Lancet Oncol.* 2018;19(5):e240-e251. doi:10.1016/S1470-2045(18)30096-2
178. Zandberg DP, Algazi AP, Jimeno A, *et al.* Durvalumab for recurrent or metastatic head and neck squamous cell carcinoma: Results from a single-arm, phase II study in patients with $\geq 25\%$ tumour cell PD-L1 expression who have progressed on platinum-based chemotherapy. *Eur J Cancer.* 2019;107:142-152. doi:10.1016/j.ejca.2018.11.015
179. Burtneess B, Harrington KJ, Greil R, *et al.* Pembrolizumab alone or with chemotherapy versus cetuximab with chemotherapy for recurrent or metastatic squamous cell carcinoma of the head and neck(KEYNOTE-048): a randomised, open-label, phase 3 study. *Lancet.* 2019;394(10212):1915-1928. doi:10.1016/S0140-6736(19)32591-7
180. Emancipator K, Linggang Huang •, Aurora-Garg D, *et al.* Comparing programmed death ligand 1 scores for predicting pembrolizumab efficacy in head and neck cancer. *Mod Pathol.* 2021;34:532-541. doi:10.1038/s41379-020-00710-9

- CTTCAGACCTGTCCTCCCTATTCAAAAAATATTATTATGAGCTCTTACTTGCTACCCAGCACTGATATAGGCACTCAGGAATAACAATGAATAAGATAGTAAAAATCTATATCTCTAAGGTTACGTTCCATGTCTGAAG
181. Rizvi NA, Hellmann MD, Snyder A, *et al.* CANCER IMMUNOLOGY Mutational landscape determines sensitivity to PD-1 blockade in non-small cell lung cancer. <https://www.science.org>. Accessed March 22, 2022.
 182. Valero C, Lee M, Hoen D, *et al.* The association between tumor mutational burden and prognosis is dependent on treatment context. *Nat Genet.* 2021;53(1):11-15. doi:10.1038/s41588-020-00752-4
 183. Pramana J, Van den Brekel MWM, van Velthuysen M-LF, *et al.* Gene expression profiling to predict outcome after chemoradiation in head and neck cancer. *Int J Radiat Oncol Biol Phys.* 2007. doi:10.1016/j.ijrobp.2007.08.032
 184. Liu X, Liu P, Chernock RD, *et al.* A prognostic gene expression signature for oropharyngeal squamous cell carcinoma. *EBioMedicine.* 2020;61:102805. doi:10.1016/j.ebiom.2020.102805
 185. Schmidt S, Linge A, Zwanenburg A, *et al.* Development and validation of a gene signature for patients with head and neck squamous cell carcinomas treated by postoperative radio(chemo)therapy. *Clin Cancer Res.* 2018;24(6):clincanres.2345.2017. doi:10.1158/1078-0432.CCR-17-2345
 186. Shi H, Chen J, Li Y, *et al.* Identification of a six microRNA signature as a novel potential prognostic biomarker in patients with head and neck squamous cell carcinoma. *Oncotarget.* 2016;7(16):21579-21590. doi:10.18632/oncotarget.7781
 187. Chen J, Fu G, Chen Y, Zhu G, Wang Z. Gene-expression signature predicts survival benefit from postoperative chemoradiotherapy in head and neck squamous cell carcinoma. *Oncol Lett.* 2018;16(2):2565-2578. doi:10.3892/ol.2018.8964
 188. Van Hooff SR, Leusink FJK, 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):4104-4110. doi:10.1200/JCO.2011.40.4509
 189. Watanabe H, Mogushi K, Miura M, *et al.* Prediction of lymphatic metastasis based on gene expression profile analysis after brachytherapy for early-stage oral tongue carcinoma. *Radiother Oncol.* 2008;87(2):237-242. doi:10.1016/j.radonc.2007.12.027
 190. MesSW, teBeest D, Poli T, *et al.* Prognostic modeling of oral cancer by gene profiles and clinicopathological co-variables. *Oncotarget.* 2017;8(35):59312-59323. doi:10.18632/oncotarget.19576
 191. Zhong Q, Fang J, Huang Z, *et al.* A response prediction model for taxane, cisplatin, and 5-fluorouracil chemotherapy in hypopharyngeal carcinoma. *Sci Rep.* 2018;8(1):1-8. doi:10.1038/s41598-018-31027-y
 192. Schomberg J, Ziogas A, Anton-Culver H, Norden-Krichmar T. Identification of a gene expression signature predicting survival in oral cavity squamous cell carcinoma using Monte Carlo cross validation. *Oral Oncol.* 2018;78(February):72-79. doi:10.1016/j.oraloncology.2018.01.012
 193. Pavón MA, Parreño M, Téllez-Gabriel M, *et al.* Gene expression signatures and molecular markers associated with clinical outcome in locally advanced head and neck carcinoma. *Carcinogenesis.* 2012;33(9):1707-1716. doi:10.1093/carcin/bgs207
 194. 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):1628-1635. doi:10.1093/annonc/mdu173
 195. De Jong MC. PREDICTING RADIORESISTANCE IN HEAD AND NECK CANCER. 2017. <http://www.moniquedejong-research.eu/phd>. Accessed April 18, 2022.
 196. De Jong MC, Ten Hoeve JJ, Grénman R, *et al.* Pretreatment microRNA expression impacting on epithelial-to-mesenchymal transition predicts intrinsic radiosensitivity in head and neck cancer cell lines and patients. *Clin Cancer Res.* 2015;21(24):5630-5638. doi:10.1158/1078-0432.CCR-15-0454

TGTGAAGGAGATTAAATAAGATGGTGTGATATAAGTATCTGGGAGAAAACTTAGGGTGTGGATATTACGGAAAGCCTTCCTAAAAAATGAATTAACTGATGAGAAGAAAGGATCCAGCTGAGAGCAAAAGCAAAAGCTTTCTTCCTTC

197. Pramana J, Van den Brekel MWM, van Velthuysen MLF, *et al.* Gene Expression Profiling to Predict Outcome After Chemoradiation in Head and Neck Cancer. *Int J Radiat Oncol Biol Phys.* 2007;69(5):1544-1552. doi:10.1016/j.ijrobp.2007.08.032
198. Van Den Broek GB, Wildeman M, Rasch CRN, *et al.* Molecular markers predict outcome in squamous cell carcinoma of the head and neck after concomitant cisplatin-based chemoradiation. doi:10.1002/ijc.24254



CHAPTER 2

TTCCCATCAACCCCTAGG99CTCCTCCT99CTGCTGGAGTTGTAGTCTGAACTTCTATCTT99GAGAGAGCCCTAC9CTCCCCCTACCGAGTCCC99GTAAATCTTAAAGCACTGCTACCC9CCCCCCC999999CTGCTG

Acute Hypoxia Profile is a Stronger Prognostic Factor than Chronic Hypoxia in Advanced Stage Head and Neck Cancer Patients

Martijn van der Heijden
Monique C. de Jong
Caroline V.M. Verhagen
Reinout H. de Roest
Sebastian Sanduleanu
Frank Hoebbers
C. René Leemans
Ruud H. Brakenhoff
Conchita Vens
Marcel Verheij
Michiel W.M. van den Brekel

Abstract

Hypoxic head and neck tumors respond poorly to radiotherapy and can be identified using gene expression profiles. However, it is unknown whether treatment outcome is driven by acute or chronic hypoxia. Gene expression data of 398 head and neck cancers was collected. Four clinical hypoxia profiles were compared to *in vitro* acute and chronic hypoxia profiles. Chronic and acute hypoxia profiles were tested for their association to outcome using Cox proportional hazard analyses. In an initial set of 224 patients, scores of the four clinical hypoxia profiles correlated with each other and with chronic hypoxia. However, the acute hypoxia profile showed a stronger association with local recurrence after chemoradiotherapy ($p=0.02$; HR=3.1) than the four clinical (chronic hypoxia) profiles ($p=0.2$; HR=0.9). An independent set of 174 patients confirmed that acute hypoxia is a stronger prognostic factor than chronic hypoxia for overall survival, progression-free survival, local and locoregional control. Multivariable analyses accounting for known prognostic factors substantiate this finding ($p=0.045$; $p=0.042$; $p=0.018$ and $p=0.003$, respectively). In conclusion, the four clinical hypoxia profiles are related to chronic hypoxia and not acute hypoxia. The acute hypoxia profile shows a stronger association with patient outcome and should be incorporated into existing prediction models.

Introduction

The average overall survival for advanced stage head and neck cancer patients is around 50%¹, but this varies greatly among different groups of patients. Clinical (TNM) staging explains survival variation only partially for these patients²⁻⁴. Human papillomavirus (HPV) positive oropharyngeal tumors represent a distinct subgroup of head and neck squamous cell carcinoma (HNSCC) that is associated with a good prognosis⁵. We have previously shown that the addition of HPV status and a prognostic gene expression profile can improve outcome prediction, suggesting that a substantial part of the survival variation is explained by tumor biology⁵.

Hypoxia is one of the most studied biological factors affecting prognosis in HNSCC⁶. Tumor cells can become hypoxic by chronic (diffusion limited) and acute (perfusion limited) mechanisms⁷. Perfusion or diffusion limited, intermittent or cycling hypoxia are alternative terms that have been used to better reflect the mechanisms that result in acute hypoxia as referred to in this study. Both acute and chronic hypoxia can have different effects on tumor cells and their microenvironment. It is unclear whether prognosis is mostly impacted by chronic or acute hypoxia⁷. Because oxygen is essential for DNA-damage upon irradiation, hypoxic tumors respond poorly to radiation^{8,9}. Since two third of all HNSCC patients receive radiotherapy as part of their treatment, hypoxia can reduce efficacy of the treatment and impact prognosis¹⁰. A meta-analysis of clinical trials showed that in vivo modification of the hypoxia status during radiotherapy improves survival of HNSCC patients, demonstrating that hypoxia is an important factor in radioresistance¹¹. Likewise, hypoxia-inducible factor is a poor prognostic factor in surgically treated HNSCC patients¹². Unfortunately, the benefit from hypoxia modification therapy was modest and comes with added toxicity¹¹. This led to the hypothesis that only patients with hypoxic tumors profit from such a therapeutic intervention, which was further confirmed in two subsequent studies^{13,14}. Initiated in 2014, the NIMRAD study aims to 'prospectively validate a hypoxia gene signature that can be used in clinical practice to personalize treatment and select appropriate patients for hypoxia modifying treatment'¹⁵.

A robust method to quantify hypoxia in tumors is required to select patients for hypoxia modifying treatment. Different techniques have been applied to evaluate the level of hypoxia in a tumor and its impact on radiotherapy response⁶, including an oxygen-sensitive needle probe inserted into the tumor¹⁶⁻¹⁹, exogenous immunohistochemical markers (e.g., pimonidazole²⁰), endogenous biomarkers (e.g., HIF1- α ^{21,22} or carbonic anhydrase IX^{14,21}) and imaging techniques like MRI²³ and PET²⁴. None of these techniques are currently used in routine clinical practice because they are too complex or insufficiently accurate. Tumor

TTCCATCAAGCCCTAGGGTCTCTGTGGTGCTGGGAGTTGTAGTCTGAACGTTCTATCTTGGGAGAAAGGCTCAAGCTCCCCCTACCGAGTCCCGGGTAATCTTAAAGACCTGACCGCCCCCGCGCTGAGAGGGGAG

hypoxia can also be deduced from gene expression profiling and many gene expression profiles for hypoxia have been proposed^{13,21,25–29}.

In HNSCC, four hypoxia profiles have been validated to be prognostic or predictive^{13,25–27}. Winter *et al.*²⁵ obtained a hypoxia 99-gene expression profile that was validated to be associated with recurrence free survival. Buffa *et al.*²⁶ used hypoxia-regulated genes to select co-expressed genes in three HNSCC and five breast cancer studies. The resulting 51-gene profile was validated for its prognostic relevance in four independent datasets. Toustrup *et al.*¹³ generated a profile from *in vitro* experiments together with gene expression data from 58 head and neck cancer biopsies. The resulting 15-gene expression profile proved to be predictive for the response to nimorazole during radiotherapy in 323 HNSCC patients. Eustace *et al.*²⁷ reduced the profile from Buffa *et al.* to a 26-gene profile that predicted regional control in 157 laryngeal cancer patients treated with radiotherapy. Some authors suggest that the use of multiple hypoxia profiles improves the prognostic value³⁰. However, none of these studies made a distinction between acute and chronic hypoxia. Gene expression profiles for acute and chronic hypoxia were generated *in vitro*. Seigneuric *et al.* used temporal changes in human epithelial mammary cell lines in response to hypoxia to generate gene expression profiles for acute and chronic hypoxia and showed that acute hypoxia is a prognostic factor in breast cancer³¹.

In this study we aimed to improve hypoxia-based HNSCC patient prognostication by evaluating the contribution of chronic and acute hypoxia in gene expression profiles. We selected the four clinical HNSCC hypoxia gene expression profiles which are associated with patient outcome in HNSCC^{13,25–27}. In an initial set of 224 patients, these gene expression profiles were compared amongst each other and to the *in vitro* chronic and acute hypoxia gene expression profiles of Seigneuric *et al.*³¹. The prognostic value of these acute and chronic hypoxia expression profiles was then tested in a new cohort of 91 HNSCC patients and validated in an independent cohort of 174 HPV-negative HNSCC patients, all treated with definitive cisplatin based chemoradiotherapy. Lastly, the prognostic value of acute and chronic hypoxia expression profiles was tested in a multivariable analysis with known prognostic factors in HNSCC.

Results

Few Overlapping Genes in the Four Clinical HNSCC Hypoxia Gene Expression Profiles

We compared the four clinical gene expression profiles of Winter *et al.*, Buffa *et al.*, Toustrup *et al.* and Eustace *et al.* to acute and chronic hypoxia profiles of Seigneuric *et al.* (Figure

1). A more extensive overview of the profiles is presented in Supplementary Materials, Table S1 and S2. The four clinical profiles consisted of 147 unique genes. Of these, 82% was present in only one of the four profiles. Three genes (2%) were present in all four signatures: Aldolase A (*ALDOA*), Prolyl 4-Hydroxylase Subunit Alpha 1 (*P4HA1*) and Solute Carrier Family 2 Member 1 (*SLC2A1* a.k.a. *GLUT-1*). Aldolase A is a glycolytic enzyme, the *P4HA1* gene encodes a component of a key enzyme in collagen synthesis and the *SLC2A1* (a.k.a. *GLUT-1*) gene encodes a glucose transporter. None of these three genes were present in the Seigneure chronic *in vitro* profile, however 9 genes in this profile were present in at least one of the four clinical profiles. The Seigneure acute profile had no overlapping genes with any of the clinical profiles or the Seigneure chronic *in vitro* profile.

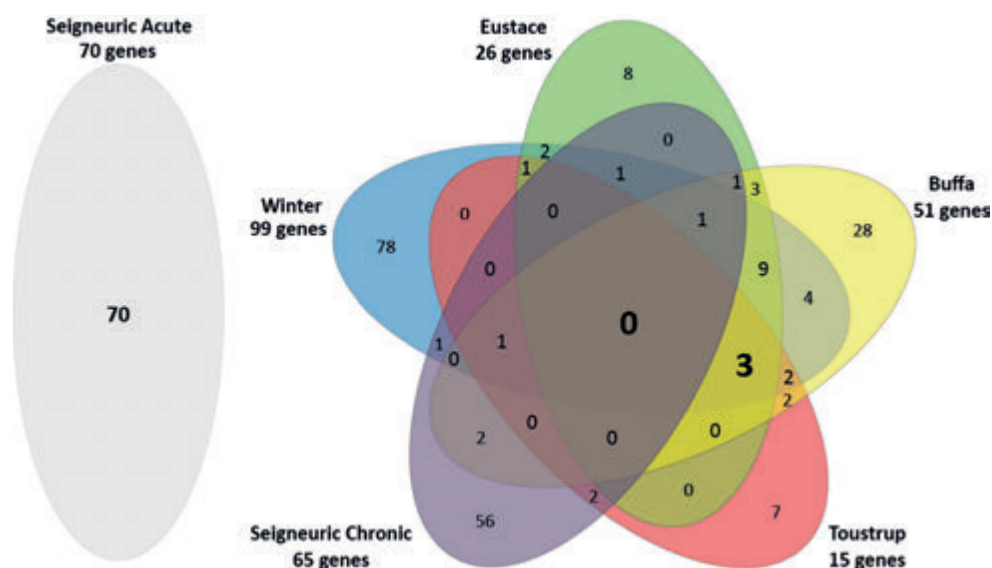


Figure 1. Venn diagram showing the number of overlapping genes in the four clinical hypoxia profiles and the Seigneure acute and chronic *in vitro* profiles.

The Four Clinical HNSCC Hypoxia Gene Expression Profiles are Correlated and Resemble Chronic Hypoxia Response

We tested the conformity of the four clinical hypoxia profiles in 224 HNSCC patients. To this end, we combined gene expression data generated from three HNSCC cohorts (Table 1): Pramana (published), de Jong 1 (expanded expression data set with additional samples that could not be used in the original study) and unpublished gene expression data from a patient cohort as described in de Jong 2 *et al.*^{32–36}. Patient characteristics of the cohorts are available in Supplementary Materials, Tables 3–5. Scores for the four clinical profiles and the

Seigneuric acute and chronic hypoxia profiles were generated for all 224 HNSCC patients. The average Spearman correlation between the scores of the four hypoxia profiles was 0.82 and highly significant (range 0.71–0.90, *p*-values < 0.0001, Figure 2A), demonstrating that the four gene expression profiles rank patients similarly. All four clinical hypoxia profiles were significantly correlated to the Seigneuric chronic hypoxia profile (correlation 0.60, *p*< 0.0001, Figure 2A,). The average correlation with Seigneuric acute hypoxia was -0.09 (*p*=0.2). A clustering based on all scores showed that the four clinical HNSCC hypoxia profiles cluster together with the Seigneuric chronic profile, whereas no correlation was observed with the Seigneuric acute profile (Figure 2B). Together this strongly suggests that the four clinical profiles represent chronic hypoxia and not acute hypoxia.

Table 1. Summary of characteristics of the five HNSCC patient cohorts.

Use	Study	Cohorts	N	Sites	Treatment	Material	Assay
Combined for classification analyses	Pramana ³² (test cohort)	Stage III-IV HNSCCs	91	All head and neck	Chemo-radiotherapy	FF	Dual channel Operon microarray
	De Jong 1 ³³	Larynx / oropharynx	99	Larynx/ oropharynx	Radiotherapy	FF	Illumina beads microarray
	De Jong 2 ³⁴	T2-3 larynx	34	Larynx	Radiotherapy	FFPE	RNAseq
Combined for validation cohort	Van der Heijden - NKI-CRAD ^{35,36}	Stage III-IV HNSCCs	98	Larynx, hypopharynx, HPV-neg oropharynx	Chemo-radiotherapy	FF	RNAseq
	Van der Heijden - DESIGN ^{35,36}	Stage III-IV HNSCCs	76	Larynx, hypopharynx, HPV-neg oropharynx	Chemo-radiotherapy	FF	RNAseq

FF: Fresh-Frozen; FFPE: Formalin Fixed Paraffin Embedded

The Acute Hypoxia Profile is Associated with Local Control

As the inclusion criteria of the three cohorts were not comparable, we did not attempt to combine the three cohorts for outcome analyses. The largest cohort with Local Control (LC) data available (Pramana cohort, *n* = 91) was used to assess the prognostic value of acute and chronic hypoxia³². Unfortunately, no clinically validated acute hypoxia profiles are available for HNSCC. We therefore used the Seigneuric acute hypoxia profile as a surrogate marker for acute hypoxia in HNSCC. Since the four clinically validated HNSCC hypoxia profiles were highly correlated to each other and to the Seigneuric Chronic profile (Figure 2A), for each patient the scores of the four clinical profiles were averaged to obtain a joint chronic hypoxia score. The median was used to define “High” and “Low” hypoxia groups.

TTCCATCAAGCCCTAGGGCTCCTCTGTGGTGTCTGGGAGTTGTAGTCTGAACGCTTCTATCTTGGGAGAGAGGCTACGCTCCGCCCTACCGAGTCCCGGGTAACTCTTAAAGACCTGACCGCCCCCGCGCGCTGTAGAGGGGAG

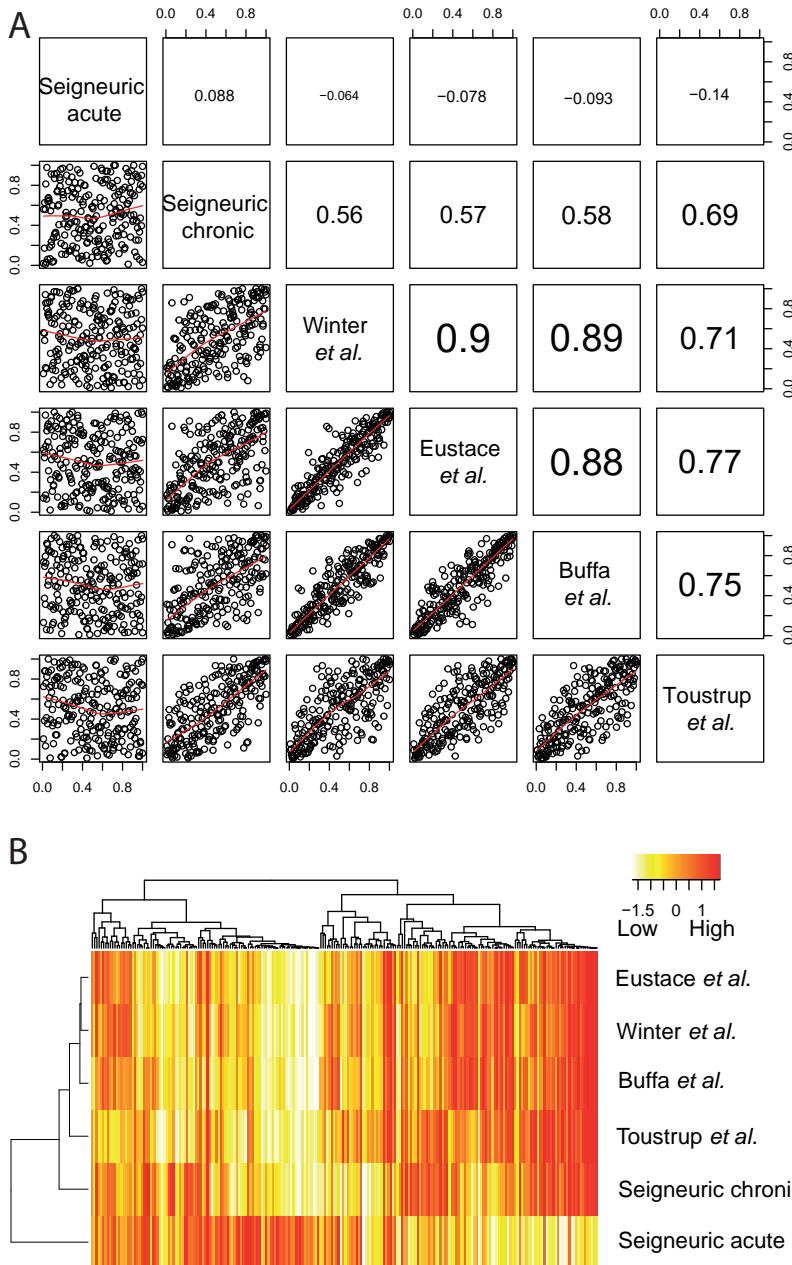


Figure 2. Correlations and clustering of the hypoxia scores from the clinical hypoxia profiles and the Seigneureic acute and chronic *in vitro* hypoxia profiles in 224 HNSCC patients. **(A)** Spearman correlations (upper right panels) and scatter plots (lower left panels) of all possible pairs of hypoxia profiles for 224 patients. All Spearman correlations were significant at the $p < 0.0001$ level. **(B)** Heatmap showing the scores for the four hypoxia profiles and *in vitro* profiles as indicated in 224 patients.

TTCCATCAGCCCTAGGGTCTCTGTGGTGCTGGGATTGTAGTCTGAACGTTCTATCTTGGGAGAGAGCCTACGTCCTCCCTACCGAGTCCCGGGTAATCTTAAAGACCTGACCGCCCCCGCGCTGTAGAGGGGAG

The acute hypoxia profile was significantly associated with local control ($p=0.02$). Patients in the acute hypoxia “High” group showed a higher local recurrence rate (Hazard Ratio (HR): 3.1, 95%CI: 1.1–8.6), compared to patients in the acute hypoxia “Low” group. Patients in chronic hypoxia “High” group tended to have a worse local control rate ($p=0.2$, HR=1.9; Figure 3). Kaplan Meier curves and hazard ratios for the individual profiles are provided in Supplementary Figure 1. To study the combined effect of the acute hypoxia profile and the chronic hypoxia profile we generated three groups: (1) acute and chronic hypoxia profile both low, (2) acute or chronic hypoxia profile high and (3) acute and chronic hypoxia profile both high (Figure 3C). The comparison in Figure 3C shows that, local control rates are worst when both acute and chronic hypoxia profiles are high ($p=0.04$).

Validation of the Acute Hypoxia Profile as a Prognostic Marker for Outcome in HNSCC

HPV is a major prognostic factor in HNSCC and the Pramana cohort contained HPV-positive and HPV-negative tumors. So, to test whether the prognostic value of acute hypoxia is independent of HPV, we aimed to validate the prognostic relevance of acute hypoxia in a HPV-negative HNSCC cohort. To this end, we collected clinical and gene expression data from a study by Van der Heijden *et al.*^{35,36}. This data derived from pre-treatment patient tumor material of two cohorts of HPV-negative advanced stage HNSCC patients, NKI-CRAD and DESIGN, comprising 174 patients in total. All patients were treated with chemo-radiotherapy and an overview of patient characteristics is available in Supplementary Materials, Table S6. Overall Survival (OS), Progression Free Survival (PFS), Local Control (LC), Locoregional Control (LRC) and Distant Metastasis free survival (DM) data were available for this dataset. As before, scores for all 174 patients were calculated. The Seigneure acute hypoxia profile was used as a marker for acute hypoxia and the scores of the four clinical HNSCC hypoxia profiles (Winter *et al.*, Buffa *et al.*, Toustrup *et al.* and Eustace *et al.*) were averaged to obtain a joint chronic hypoxia score. The distribution of the profiles is shown in Supplemental Figure 2A. Median splits were used to divide patients in “High” and “Low” groups.

Confirming the findings in the Pramana cohort, the acute hypoxia “High” group had a significantly worse LC ($p=0.006$, HR=3.3; Figure 4A), OS (HR=1.58, $p=0.023$), PFS (HR=1.66, $p=0.009$) and LRC; (HR=2.4, $p=0.008$; Supplementary Figure 2B). The chronic hypoxia “High” group showed a significant difference in PFS only (HR=1.52, $p=0.03$) and a trend for OS (HR=1.43, $p=0.075$; supplementary Figure 2C), but no significant difference in LC ($p=0.34$; HR=1.43; Figure 4B) and LRC (HR=1.59, $p=0.141$). We next combined the chronic hypoxia and acute hypoxia groups as described earlier to obtain three groups. When both acute and

chronic hypoxia expression profiles are “High”, OS, PFS, LC and LRC rates are significantly worse than when both are low (Figure 4C-E and Supplementary Figure 3).

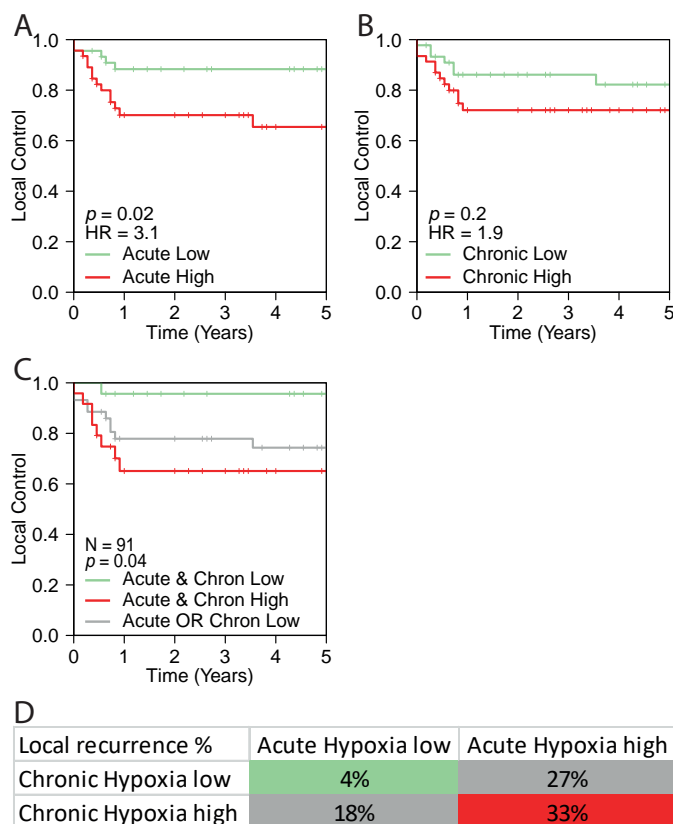


Figure 3. Combined acute and chronic hypoxia scores. (A) Kaplan–Meier curve showing local control of the “High” versus “Low” groups of the acute hypoxia profile. (B) Kaplan–Meier curve showing local control of the “High” versus “Low” groups of the chronic hypoxia profile. (C) Kaplan–Meier curve showing local control for 3 groups: (1) acute and chronic hypoxia both low, (2) acute or chronic hypoxia high, or (3) acute and chronic hypoxia both high. The crosstab shows the percent of local recurrences per subgroup in 91 chemoradiotherapy patients. Samples were divided into two groups using the median. Cells are colored in a color corresponding with the line color in the Kaplan–Meier curve. P-value represent the log-rank p -value. (D) Crosstab showing the percentage local recurrences per subgroup.

Multivariable Analysis Confirms the Prognostic Value of Acute Hypoxia

To determine whether acute hypoxia has additional prognostic value to known prognostic markers, other than HPV, we performed a multivariable in the Van der Heijden cohort. For 149 patients, tumor volume data were available. Within these 149 patients all available variables were tested in a univariable analysis for their association with OS, PFS, LC, LRC and DM (Supplementary Materials, Table S7). To avoid arbitrary cutoffs, acute hypoxia, chronic

hypoxia (the average of the four clinical hypoxia scores) and tumor volume were tested as continuous variables.

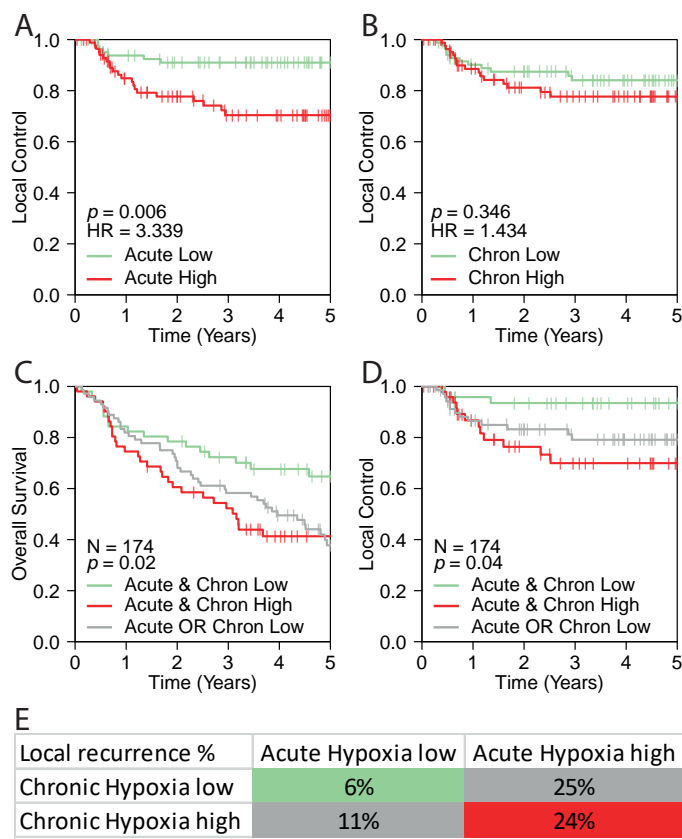


Figure 4. Validation of prognostic value of combined acute and chronic hypoxia scores in the van der Heijden cohort. (A) Kaplan–Meier curve showing local control of the “High” versus “Low” acute hypoxia. (B) Kaplan–Meier curve showing local control of the “High” versus “Low” chronic hypoxia. (C) Kaplan–Meier curve overall survival (C) and local control (D) for 3 groups: acute and chronic hypoxia both low, acute or chronic hypoxia high or acute and chronic hypoxia both high. (E) Crosstab showing the percentage of local recurrences per subgroup in the validation cohort consisting of 174 patients treated with chemoradiotherapy. Cells are colored in a color corresponding with the line color in the Kaplan–Meier curves in panel C and D.

Acute hypoxia, gender, tumor site, disease stage, cisplatin dose, and tumor volume showed significant associations with multiple patient outcome measures in univariable analysis (Supplementary Materials, Table S7). Chronic hypoxia only showed associations with locoregional control. Variables with significant associations to patient outcome were combined in a multivariable Cox proportional hazard model (Table 2). Previous research has shown an

interaction between hypoxia and tumor volume³⁷, this interaction was therefore incorporated in the multivariable analysis. Results of the multivariable analysis show that acute hypoxia is significantly associated with OS (HR=3.5, $p=0.045$), PFS (HR=3.3, $p=0.042$), LC (HR=38.2, $p=0.018$) and LRC (HR=27, $p=0.003$), whereas chronic hypoxia is not associated with any of the outcomes. Tumor volume (per cm³ increase) is also associated with OS (HR=1.01, $p=4.12E-05$), PFS (HR=1.008, $p=0.0007$) and LC (HR=1.014, $p=0.027$). The interaction of acute hypoxia with tumor volume was significantly associated with OS (HR=0.98, $p=0.004$), PFS (HR=0.98, $p=0.02$), and LRC (HR=0.95, $p=0.016$). The interaction of chronic hypoxia with tumor volume was not significantly associated with any of the reported outcomes. These analyses confirm the prognostic value of acute hypoxia independent from clinical factors, tumor volume and chronic hypoxia.

Table 2. Multivariable Cox proportional hazard analysis of parameters with patient outcome in the Van der Heijden cohort.

OS: Overall Survival; PFS: Progression Free Survival; LC: Local Control; LRC: Locoregional Control;

DM: Distant Metastasis Free Survival; HR: Hazard Ratio. *: Interaction between the two variables.

	OS		PFS		LC		LRC		DM	
Variable	HR	p-value	HR	p-value	HR	p-value	HR	p-value	HR	p-value
Gender										
Female	0,83	0,04	0,56	0,04	0,16	0,013	0,19	0,004	0,28	0,04
Male	REF		REF		REF		REF		REF	
Tumor site										
Oropharynx	REF		REF		REF		REF		REF	
Hypopharynx	0,64	0,09	0,69	0,15	0,96	0,12	0,97	0,94	0,52	0,16
Larynx	0,58	0,9	0,79	0,45	1,27	0,56	1,08	0,86	1,08	0,87
Disease stage										
II-III	0,83	0,57	0,78	0,45	3,74	0,037	1,68	0,3	0,38	0,2
IVA-IVB	REF		REF		REF		REF		REF	
Cumulative cisplatin < 200 mg/m²										
No	REF		REF		REF		REF		REF	
Yes	2,03	0,003	1,93	0,004	6,4	0,0003	2,7	0,012	0,81	0,64
Tumor Volume	1,01	4,40E-05	1,008	0,0006	1,014	0,034	1,007	0,16	1,006	0,19
Acute Hypoxia	4,14	0,021	4,02	0,017	46,4	0,014	30	0,0022	5,5	0,13
Chronic Hypoxia	0,83	0,65	1	0,99	0,51	0,48	1,23	0,75	1,26	0,73
Acute Hypoxia *	0,98	0,007	0,98	0,03	0,96	0,12	0,96	0,017	0,98	0,26
Tumor Volume										
Chronic Hypoxia *	1	0,99	0,99	0,7	1	0,98	0,99	0,52	0,99	0,57
Tumor Volume										

OS: Overall Survival; PFS: Progression Free Survival; LC: Local Control; LRC: Locoregional Control; DM: Distant Metastasis Free Survival; HR: Hazard Ratio. *: Interaction between the two variables.

Discussion

The four clinical gene expression profiles for hypoxia, which have been validated to predict outcome in HNSCC, have few overlapping genes. Nevertheless, they classified patients similarly, indicating that they reflect a similar underlying biological process: chronic and not acute hypoxia. Since the clinical hypoxia profiles were correlated, we combined them in a joint chronic hypoxia score. In contrast to the chronic hypoxia profile, the acute profile was associated with local control in the Pramana cohort. The poor association of the acute hypoxia profile with patient outcome was validated in an independent validation cohort ($n=174$) and multivariable analysis showed that acute hypoxia is a significant prognostic factor independent of clinical factors, tumor volume and chronic hypoxia.

The phenomenon that different gene expression profiles, with different genes, can describe the same process, has been reported before³⁸. Given the fact that over 4,000 genes are hypoxia-influenced, it seems reasonable to assume that multiple robust, but different, hypoxia gene expression profiles can be assembled³⁹. It should be noted that all gene expression profiles have been applied with the same method in our study. Due to lack of access to the reference cohort, it was not always possible to apply gene expression profiles as in the original publication. To be able to compare the gene expression profiles, we decided to apply the same method to all gene expression profiles.

Acute and Chronic Hypoxia

The terms acute and chronic hypoxia are simplified terms to describe a complex spectrum of hypoxic micro environmental alterations in a tumor⁴⁰. While an absolute distinction between the two cannot be made, many suggestions for the separate origin, measurement and treatment of the two entities have been published^{41–44}. Janssen *et al.* employed various staining protocols to study acute and chronic hypoxia in head and neck tumors⁴². They showed that tumors contained on average 15% acute hypoxic (proliferating cells around non-perfused vessels) and around 30% chronic hypoxic areas (cells at a large distance from blood vessels). The two different areas did not overlap. Also *in vitro* studies showed that cells that had been under hypoxia for a short time, showed a different gene expression profile compared to cells that were hypoxic for a prolonged time³¹. As described by Lendahl *et al.* 4,047 genes were hypoxia-regulated in a colon carcinoma cell line. Only 52 genes were specific to the acute (1 or 2 h) hypoxia response. 144 genes were up- or downregulated by both acute and chronic (24 hour) hypoxia, whereas the majority of the genes (4,005) were chronic hypoxia specific³⁹. Nonetheless, in the past decades research has focused on generating a gene expression profile for ‘hypoxia’ in general, without distinction of acute and chronic. The fact that all four clinical

HNSCC hypoxia profiles correlated with chronic hypoxia could be due to the large excess of genes regulated by chronic hypoxia³⁹. Also, the methods used to generate the gene expression profile could be an explanation that the profiles correlated with chronic hypoxia. For example, the Toustrup *et al.* profile is based on Eppendorf probe measurements to find relevant genes. If indeed, the area of chronic hypoxia is on average twice the area of acute hypoxia, as reported by Janssen *et al.*, this could lead to a bias towards genes that are correlated with chronic hypoxia. Winter, Buffa and Eustace *et al.* started with 10 hypoxia ‘seed genes’ to develop their signatures. In our data, these 10 genes were not correlated with *in vitro* Seigneure acute hypoxia but showed correlations to the *in vitro* Seigneure chronic hypoxia profile (Supplementary Materials, Table S8).

Acute Hypoxia and Prognosis

The importance of acute hypoxia has been recognized for decades⁴⁵. For example, Chan *et al.* showed that a human lung squamous cell carcinoma cell line (H1299) becomes more radioresistant under acute hypoxia than under chronic hypoxia, with respective oxygen enhancement ratios of 1.96 and 1.37⁴⁶. Unfortunately, conclusive data on the separate and combined prognostic effects of acute and chronic hypoxia in HNSCC are lacking. This might be due to the fact that it is difficult to distinguish both types of hypoxia with immunohistochemistry. We found that patients with high acute or chronic hypoxia expression, had a 3.1 and 1.9 times higher risk of a local recurrence in the Pramana cohort, respectively. This was confirmed in the validation set of 174 patients with hazard ratios of 3.3 and 1.4 for local failure, for acute and chronic hypoxia, respectively. Although in both sets chronic hypoxia was not significant, the effect size appears comparable to previously reported hazard ratios for chronic hypoxia^{13,25–27}. For chronic hypoxia gene expression profiles, the general deduction is that more hypoxic tumors are approximately twice as likely to recur than less hypoxic tumors. This effect could be underestimated due to a division into two hypoxia groups according to the median. We therefore used the scores as continuous variables in the multivariable analysis, which shows that acute hypoxia is a stronger prognostic factor than chronic hypoxia. Literature has suggested that cells lacking functional p53 are more susceptible to genomic instability and potential tumorigenesis when experiencing reoxygenation after acute hypoxia compared to chronic hypoxia⁴⁷. This might also explain why the effect is more pronounced in the Van der Heijden cohort, which contains only HPV-negative HNSCC in which p53 mutations are highly prevalent. Since acute and chronic hypoxia have a different etiology, knowledge about which type of hypoxia causes radioresistance in a specific patient, could lead to the use and development of strategies targeting acute or chronic hypoxia.

Materials and Methods

Gene Expression Profiles

We selected four gene expression profiles for hypoxia that are clinically validated for head and neck cancer, namely, Winter *et al.*, Buffa *et al.*, Toustrup *et al.* and Eustace *et al.* Gene expression profiles generated by Seigneuric *et al.* were used for *in vitro* acute and chronic hypoxia. Since 2% oxygen is relatively close to oxygen concentration in normal tissue we used the <0.02% oxygen derived profiles as the *in vitro* acute and chronic hypoxia profiles. There is no clear cutoff for acute and chronic hypoxia. To simplify the complex spectrum of hypoxia, we used the Seigneuric 1 hour time point profile for acute hypoxia, while the Seigneuric 24 hour time point profile was used for chronic hypoxia⁶.

Datasets

We initially collected gene expression data of three different HNSCC patient cohorts (Pramana, De Jong 1 and De Jong 2), comprising 224 patients in total (Table 1). Extensive patient characteristics for each cohort are available in the Supplementary Materials, Tables S3–S6 and the original publications^{32–34}. Clinical and expression data of Pramana *et al.* has been previously published. Briefly, the dataset of Pramana *et al.* consisted of 91 HPV-positive and negative HNSCC patients all treated with definitive cisplatin based chemoradiotherapy³². Gene expression data were obtained from fresh-frozen pre-treatment material, analyzed using dual-channel Operon microarray slides. The full dataset is available at (GEO database) and extensive methods are available in the original publication.

Data of 52 matched patients of the De Jong 1 cohort was previously published³³. This dataset was expanded with an additional 47 samples that could not be used in the original study. Briefly, gene expression of 99 fresh-frozen laryngeal and oropharyngeal (HPV status unknown) carcinomas was determined using the Illumina beads microarray platform. All patients were treated with radiotherapy. The full dataset is available at (GEO database, We are in the process of submitting to GEO and will provide the numbers a.s.a.p.) and has been generated as described previously.

Clinical data from the De Jong 2 cohort are as described previously³⁴. The cohort consists of 34 laryngeal carcinomas treated with radiotherapy. Collection of tumor material and clinical data were approved by the Institutional Review Board of the Netherlands Cancer Institute and all patients signed informed consent. RNA was extracted from pre-treatment FFPE biopsies using the Roche High Pure miRNA Isolation Kit (REF: 05080576001). Quality and quantity of the total RNA was assessed with the 2100 Bioanalyzer using a Nano chip. TruSeq cDNA libraries were generated using the TruSeq RNA Library Preparation Kit v2

sample preparation kit (Illumina Cat.No RS-122-2001/2). The reads (51bp) were sequenced on a HiSeq2000 using V3 chemistry (Illumina Inc., San Diego CA, USA), aligned against the human genome, build 37, using Tophat (version 2.0.6). HTseq count was used to count the number of reads per gene. Only uniquely mapped reads were counted. Detailed methods are provided in the Supplementary Methods. Data are available at <http://ega-archive.org>.

To validate our findings, we used gene expression data collected and generated within a separate study as described by van der Heijden *et al.* and deposited at <http://ega-archive.org>^{35,36}. The “van der Heijden cohort” combines two independent cohorts, the NKI-CRAD and the DESIGN cohort, and comprises 174 HPV negative advanced stage HNSCC patients. All patients were treated with definitive cisplatin-based chemoradiotherapy. RNA was isolated from fresh-frozen material and gene-expression was determined using RNA-sequencing as described previously.

Data Analysis

All analyses were performed in R 3.4.3 using RStudio 1.1. The gene expression profiles consisted of genes exclusively upregulated under hypoxia. All expression data were normalized as described in the original publications. In the Pramana and both De Jong cohorts, per sample hypoxia scores were calculated for each gene expression profile using the mean normalized expression of all genes in the profile. In order to combine the three different patient cohorts, scores were rank-normalized per profile between 0 and 1 for each patient cohort. Gene-expression scores were calculated using Gene Set Variation Analysis(GSVA)⁴⁸ in the van der Heijden dataset. The median was used to split patients into “High” and “Low” hypoxia groups. Cox proportional hazard analyses were used to test associations with patient outcome (Overall Survival (OS), Progression Free Survival (PFS), Local Control (LC), Locoregional Control (LRC) and Distant Metastasis (DM)). Multivariable Cox proportional hazard analyses were performed in the Van der Heijden cohort and no cut-offs were used to test the continuous variables.

Conclusions

All four clinical HNSCC hypoxia profiles correlate with chronic and not with acute hypoxia expression profiles. However, the acute hypoxia profile has a stronger association with patient outcome after chemoradiotherapy in both the test and validation cohort. Multivariable analysis shows the value of acute hypoxia in addition to well-known prognostic factors for HNSCC. Acute hypoxia gene expression should therefore be incorporated into existing hypoxia-based prediction models.

References

1. Ferlay, J.; Shin, H.R.; Bray, F.; Forman, D.; Mathers, C.; Parkin, D.M. Estimates of worldwide burden of cancer in 2008: GLOBOCAN 2008. *Int. J. Cancer* **2010**, doi:10.1002/ijc.25516.
2. Groome, P.A.; Schulze, K.; Boysen, M.; Hall, S.F.; Mackillop, W.J.; O'Sullivan, B.; Irish, J.C.; Warde, P.R.; Schneider, K.M.; Mackenzie, R.G.; *et al.* A comparison of published head and neck stage groupings in laryngeal cancer using data from two countries. *J. Clin. Epidemiol.* **2002**, doi:10.1016/S0895-4356(02)00389-X.
3. Groome, P.A.; Schulze, K.M.; Mackillop, W.J.; Grice, B.; Goh, C.; Cummings, B.J.; Hall, S.F.; Liu, F.F.; Payne, D.; Rothwell, D.M.; *et al.* A comparison of published head and neck stage groupings in carcinomas of the tonsillar region. *Cancer* **2001**, *92*, 1484–1494, doi:10.1002/1097-0142(20010915)92:6<1484::AID-CNCR1473>3.0.CO;2-W.
4. Hall, S.F.; Groome, P.A.; Irish, J.; O'Sullivan, B. TNM-based stage groupings in head and neck cancer: Application in cancer of the hypopharynx. *Head Neck* **2009**, *31*, 1–8, doi:10.1002/hed.20917.
5. de Jong, M.C.; Pramana, J.; Kneijens, J.L.; Balm, A.J.M.; van den Brekel, M.W.M.; Hauptmann, M.; Begg, A.C.; Rasch, C.R. HPV and high-risk gene expression profiles predict response to chemoradiotherapy in head and neck cancer, independent of clinical factors. *Radiother. Oncol.* **2010**, *95*, 365–370, doi:10.1016/j.radonc.2010.02.001.
6. Janssen, H.L.; Haustermans, K.M.; Balm, A.J.; Begg, A.C. Hypoxia in head and neck cancer: How much, how important? *Head Neck* **2005**, *27*, 622–638, doi:10.1002/hed.20223.
7. Bristow, R.G.; Hill, R.P. Hypoxia and metabolism: Hypoxia, DNA repair and genetic instability. *Nat. Rev. Cancer* **2008**, *8*, 180–892, doi:10.1038/nrc2344.
8. Ma, N.Y.; Tinganelli, W.; Maier, A.; Durante, M.; Kraft-Weyrather, W. Influence of chronic hypoxia and radiation quality on cell survival. *J. Radiat. Res.* **2013**, *54*(Suppl. 1), i13–i22, doi:10.1093/jrr/rrs135.
9. Gray, L.H.; Conger, A.D.; Ebert, M.; Hornsey, S.; Scott, O.C. The concentration of oxygen dissolved in tissues at the time of irradiation as a factor in radiotherapy. *Br. J. Radiol.* **1953**, *26*, 638–648, doi:10.1259/0007-1285-26-312-638.
10. De Gonzalez, A.B.; Curtis, R.E.; Kry, S.F.; Gilbert, E.; Lamart, S.; Berg, C.D.; Stovall, M.; Ron, E. Proportion of second cancers attributable to radiotherapy treatment in adults: A cohort study in the US SEER cancer registries. *Lancet Oncol.* **2011**, *12*, 353–360, doi:10.1016/S1470-2045(11)70061-4.
11. Overgaard, J. Hypoxic modification of radiotherapy in squamous cell carcinoma of the head and neck—A systematic review and meta-analysis. *Radiother. Oncol.* **2011**, *100*, 22–32, doi:10.1016/j.radonc.2011.03.004.
12. Winter, S.C.; Shah, K.A.; Han, C.; Campo, L.; Turley, H.; Leek, R.; Corbridge, R.J.; Cox, G.J.; Harris, A.L. The Relation Between Hypoxia-Inducible Factor (HIF) -1 a and HIF-2 a Expression With Anemia and Outcome in Surgically Treated Head and Neck Cancer. *Cancer*. 2006; pp. 757–766, doi:10.1002/cncr.21983.
13. Toustrop, K.; Sørensen, B.S.; Nordsmark, M.; Busk, M.; Wiuf, C.; Alsner, J.; Overgaard, J. Development of a hypoxia gene expression classifier with predictive impact for hypoxic modification of radiotherapy in head and neck cancer. *Cancer Res.* **2011**, *71*, 5923–5931, doi:10.1158/0008-5472.CAN-11-1182.
14. Rademakers, S.E.; Hoogsteen, I.J.; Rijken, P.F.; Oosterwijk, E.; Terhaard, C.H.; Doornaert, P.A.; Langendijk, J.A.; van den Ende, P.; Takes, R.; De Bree, R.; *et al.* Pattern of CAIX expression is prognostic for outcome and predicts response to ARCON in patients with laryngeal cancer treated in a phase III randomized trial. *Radiother. Oncol.* **2013**, *108*, 517–522, doi:10.1016/j.radonc.2013.04.022.

15. Thomson, D.; Yang, H.; Baines, H.; Miles, E.; Bolton, S.; West, C.; Slevin, N. NIMRAD—A phase III trial to investigate the use of nimorazole hypoxia modification with intensity-modulated radiotherapy in head and neck cancer. *Clin. Oncol.* **2014**, *26*, 344–347, doi:10.1016/j.clon.2014.03.003.
16. Nordsmark, M.; Bentzen, S.M.; Rudat, V.; Brizel, D.; Lartigau, E.; Stadler, P.; Becker, A.; Adam, M.; Molls, M.; Dunst, J.; *et al.* Prognostic value of tumor oxygenation in 397 head and neck tumors after primary radiation therapy. An international multi-center study. *Radiother. Oncol.* **2005**, *77*, 18–24, doi:10.1016/j.radonc.2005.06.038.
17. Brizel, D.M.; Sibley, G.S.; Prosnitz, L.R.; Scher, R.L.; Dewhirst, M.W. Tumor hypoxia adversely affects the prognosis of carcinoma of the head and neck. *Int. J. Radiat. Oncol. Biol. Phys.* **1997**, *38*, 285–289, doi:10.1016/S0360-3016(97)00101-6.
18. Nordsmark, M.; Overgaard, J. A confirmatory prognostic study on oxygenation status and loco-regional control in advanced head and neck squamous cell carcinoma treated by radiation therapy. *Radiother. Oncol.* **2000**, *57*, 39–43, doi:10.1016/S0167-8140(00)00223-1.
19. Nordsmark, M.; Overgaard, M.; Overgaard, J. Pretreatment oxygenation predicts radiation response in advanced squamous cell carcinoma of the head and neck. *Radiother. Oncol.* **1996**, *41*, 31–39, doi:10.1016/S0167-8140(96)91811-3.
20. Kaanders, J.H.; Wijffels, K.I.; Marres, H.A.; Ljungkvist, A.S.; Pop, L.A.; van den Hoogen, F.J.; de Wilde, P.C.; Bussink, J.; Raleigh, J.A.; van der Kogel, A.J. Pimonidazole binding and tumor vascularity predict for treatment outcome in head and neck cancer. *Cancer Res.* **2002**, *62*, 7066–7074.
21. Schrijvers, M.L.; van der Laan, B.F.A.M.; de Bock, G.H.; Pattje, W.J.; Mastik, M.F.; Menkema, L.; Langendijk, J.A.; Kluin, P.M.; Schuur, E.; van der Wal, J.E. Overexpression of Intrinsic Hypoxia Markers HIF1 α and CA-IX Predict for Local Recurrence in Stage T1–T2 Glottic Laryngeal Carcinoma Treated With Radiotherapy. *Int. J. Radiat. Oncol. Biol. Phys.* **2008**, *72*, 161–169, doi:10.1016/j.ijrobp.2008.05.025.
22. Koukourakis, M.I.; Giatromanolaki, A.; Sivridis, E.; Simopoulos, C.; Turley, H.; Talks, K.; Gatter, K.C.; Harris, A.L. Hypoxia-inducible factor (HIF1A and HIF2A), angiogenesis, and chemoradiotherapy outcome of squamous cell head-and-neck cancer. *Int. J. Radiat. Oncol. Biol. Phys.* **2002**, *53*, 1192–1202, doi:10.1016/S0360-3016(02)02848-1.
23. Panek, R.; Welsh, L.; Dunlop, A.; Wong, K.H.; Riddell, A.M.; Koh, D.-M.; Schmidt, M.A.; Doran, S.; McQuaid, D.; Hopkinson, G.; *et al.* Repeatability and sensitivity of T2* measurements in patients with head and neck squamous cell carcinoma at 3T. *J. Magn. Reson. Imaging* **2016**, *44*, 72–80, doi:10.1002/jmri.25134.
24. Fleming, I.N.; Manavaki, R.; Blower, P.J.; West, C.; Williams, K.J.; Harris, A.L.; Domarkas, J.; Lord, S.; Baldry, C.; Gilbert, F.J. Imaging tumour hypoxia with positron emission tomography. *Br. J. Cancer* **2015**, *112*, 238–250, doi:10.1038/bjc.2014.610.
25. Winter, S.C.; Buffa, F.M.; Silva, P.; Miller, C.; Valentine, H.R.; Turley, H.; Shah, K.A.; Cox, G.J.; Corbridge, R.J.; Homer, J.J.; *et al.* Relation of a hypoxia metagene derived from head and neck cancer to prognosis of multiple cancers. *Cancer Res.* **2007**, *67*, 3441–3449, doi:10.1158/0008-5472.CAN-06-3322.
26. Buffa, F.M.; Harris, A.L.; West, C.M.; Miller, C.J. Large meta-analysis of multiple cancers reveals a common, compact and highly prognostic hypoxia metagene. *Br. J. Cancer* **2010**, *102*, 428–435, doi:10.1038/sj.bjc.6605450.
27. Eustace, A.; Mani, N.; Span, P.N.; Irlam, J.J.; Taylor, J.; Betts, G.N.; Denley, H.; Miller, C.J.; Homer, J.J.; Rojas, A.M.; *et al.* A 26-gene hypoxia signature predicts benefit from hypoxia-modifying therapy in laryngeal cancer but not bladder cancer. *Clin. Cancer Res.* **2013**, *19*, 4879–4788, doi:10.1158/1078-0432.CCR-13-0542.

TTCCATCAAGCCCTAGGGCTCTCTGTGGTGCTGGGAGTTGTAGTCTGAACGCTTCTATCTTGGGAGAAGGCTCTACGCTCCCCCTACGAGTCCCGGGTAATCTTAAAGACCTGACCGCCCCCGCGCTGTGAGAGGGGAG

28. Koukourakis, M.I.; Bentzen, S.M.; Giatromanolaki, A.; Wilson, G.D.; Daley, F.M.; Saunders, M.I.; Dische, S.; Sivridis, E.; Harris, A.L. Endogenous markers of two separate hypoxia response pathways (hypoxia inducible factor 2 alpha and carbonic anhydrase 9) are associated with radiotherapy failure in head and neck cancer patients recruited in the CHART randomized trial. *J. Clin. Oncol.* **2006**, *24*, 727–735, doi:10.1200/JCO.2005.02.7474.
29. Rademakers, S.E.; Lok, J.; van der Kogel, A.J.; Bussink, J.; Kaanders, J.H.A.M. Metabolic markers in relation to hypoxia; staining patterns and colocalization of pimonidazole, HIF-1 α , CAIX, LDH-5, GLUT-1, MCT1 and MCT4. *BMC Cancer* **2011**, *11*, 167, doi:10.1186/1471-2407-11-167.
30. Tawk, B.; Schwager, C.; Deffaa, O.; Dyckhoff, G.; Warta, R.; Linge, A.; Krause, M.; Weichert, W.; Baumann, M.; Herold-Mende, C.; *et al.* Comparative analysis of transcriptomics based hypoxia signatures in head- and neck squamous cell carcinoma. *Radiother. Oncol.* **2016**, *118*, 350–358, doi:10.1016/j.radonc.2015.11.027.
31. Seigneuric, R.; Starmans, M.H.W.; Fung, G.; Krishnapuram, B.; Nuyten, D.S.A.; van Erk, A.; Magagnin, M.G.; Rouschop, K.M.; Krishnan, S.; Rao, R.B.; *et al.* Impact of supervised gene signatures of early hypoxia on patient survival. *Radiother. Oncol.* **2007**, *83*, 374–382, doi:10.1016/j.radonc.2007.05.002.
32. Pramana, J.; Van den Brekel, M.W.M.; van Velthuysen, M.-L.F.; Wessels, L.F.A.; Nuyten, D.S.; Hofland, I.; Atsma, D.; Pimentel, N.; Hoebers, F.J.; Rasch, C.R.; *et al.* Gene expression profiling to predict outcome after chemoradiation in head and neck cancer. *Int. J. Radiat. Oncol. Biol. Phys.* **2007**, *69*, 1544–1552, doi:10.1016/j.ijrobp.2007.08.032.
33. De Jong, M.C.; Pramana, J.; Van Der Wal, J.E.; Lacko, M.; Peutz-Kootstra, C.J.; De Jong, J.M.; Takes, R.P.; Kaanders, J.H.; van der Laan, B.F.; Wachters, J.; *et al.* CD44 expression predicts local recurrence after radiotherapy in larynx cancer. *Clin. Cancer Res.* **2010**, *16*, 5329–5338, doi:10.1158/1078-0432.CCR-10-0799.
34. De Jong, M.C.; Ten Hoeve, J.J.; Grénman, R.; Wessels, L.F.; Kerkhoven, R.; Te Riele, H.; van den Brekel, M.W.; Verheij, M.; Begg, A.C. Pretreatment microRNA expression impacting on epithelial-to-mesenchymal transition predicts intrinsic radiosensitivity in head and neck cancer cell lines and patients. *Clin. Cancer Res.* **2015**, *21*, 5630–5638, doi:10.1158/1078-0432.CCR-15-0454.
35. Van der Heijden, M.; Essers, P.; Verheij, M.; Van den Brekel, M.; Vens, C. OC-0487: EMT signatures as a prognostic marker for metastasis in HPV-negative HNSCC. *Radiother. Oncol.* **2018**, *127*, S250–S251, doi:10.1016/S0167-8140(18)30797-7.
36. Van der Heijden, M.; Essers, P.; Verhagen, C.; Willems, S.; Sanders, J.; de Roest, R.; Vossen, D.M.; Leemans, C.R.; Verheij, M.; Brakenhoff, R.H.; *et al.* Epithelial-to-mesenchymal transition is a prognostic marker for patient outcome in advanced stage HNSCC patients treated with chemoradiotherapy. *Manuscript under review*.
37. Linge, A.; Lock, S.; Gudziol, V.; Nowak, A.; Lohaus, F.; Von Neubeck, C.; Jütz, M.; Abdollahi, A.; Debus, J.; Tinhofer, I.; *et al.* Low cancer stem cell marker expression and low hypoxia identify good prognosis subgroups in HPV(-) HNSCC after postoperative radiochemotherapy: A multicenter study of the DTK-ROG. *Clin. Cancer Res.* **2016**, *22*, 2639–2649, doi:10.1158/1078-0432.CCR-15-1990.
38. Roepman, P.; Kemmeren, P.; Wessels, L.F.A.; Slootweg, P.J.; Holstege, F.C.P. Multiple robust signatures for detecting lymph node metastasis in head and neck cancer. *Cancer Res.* **2006**, *66*, 2361–2366, doi:10.1158/0008-5472.CAN-05-3960.
39. Lendahl, U.; Lee, K.L.; Yang, H.; Poellinger, L. Generating specificity and diversity in the transcriptional response to hypoxia. *Nat. Rev. Genet.* **2009**, *10*, 821–832, doi:10.1038/nrg2665.
40. Bayer, C.; Shi, K.; Astner, S.T.; Maftai, C.A.; Vaupel, P. Acute versus chronic hypoxia: Why a simplified classification is simply not enough. *Int. J. Radiat. Oncol. Biol. Phys.* **2011**, *80*, 965–968, doi:10.1016/j.ijrobp.2011.02.049.

41. Bayer, C.; Vaupel, P. Acute versus chronic hypoxia in tumors: Controversial data concerning time frames and biological consequences. *Strahlenther. Onkol.* **2012**, *188*, 616–627, doi:10.1007/s00066-012-0085-4.
42. Janssen, H.L.K.; Haustermans, K.M.G.; Sprong, D.; Blommesteijn, G.; Hofland, I.; Hoebbers, F.J.; Blijweert, E.; Raleigh, J.A.; Semenza, G.L.; Varia, M.A.; *et al.* HIF-1 α , pimonidazole, and iododeoxyuridine to estimate hypoxia and perfusion in human head-and-neck tumors. *Int. J. Radiat. Oncol. Biol. Phys.* **2002**, *54*, 1537–1549, doi:10.1016/S0360-3016(02)03935-4.
43. Maftai, C.A.; Bayer, C.; Shi, K.; Vaupel, P. Intra- and intertumor heterogeneities in total, chronic, and acute hypoxia in xenografted squamous cell carcinomas: Detection and quantification using (immuno-)fluorescence techniques. *Strahlenther. Onkol.* **2012**, *188*, 606–615, doi:10.1007/s00066-012-0105-4.
44. Wijffels, K.I.; Kaanders, J.H.; Rijken, P.F.; Bussink, J.; van den Hoogen, F.J.; Marres, H.A.; de Wilde, P.C.; Raleigh, J.A.; van der Kogel, A.J. Vascular architecture and hypoxic profiles in human head and neck squamous cell carcinomas. *Br. J. Cancer* **2000**, *83*, 674–683, doi:10.1054/bjoc.2000.1325.
45. Brown, J.M. Evidence for acutely hypoxic cells in mouse tumours, and a possible mechanism of reoxygenation. *Br. J. Radiol.* **1979**, doi:10.1259/0007-1285-52-620-650.
46. Chan, N.; Koritzinsky, M.; Zhao, H.; Bindra, R.; Glazer, P.M.; Powell, S.; Belmaaza, A.; Wouters, B.; Bristow, R.G. Chronic hypoxia decreases synthesis of homologous recombination proteins to offset chemoresistance and radioresistance. *Cancer Res.* **2008**, *68*, 605–614, doi:10.1158/0008-5472.CAN-07-5472.
47. Pires, I.M.; Bencokova, Z.; Milani, M.; Folkes, L.K.; Li, J.A.; Stratford, M.R.; Harris, A.L.; Hammond, E.M. Effects of acute versus chronic hypoxia on DNA damage responses and genomic instability. *Cancer Res.* **2010**, *70*, 925–935, doi:10.1158/0008-5472.CAN-09-2715.
48. Hänzelmann, S.; Castelo, R.; Guinney, J. GSEA: Gene set variation analysis for microarray and RNA-Seq data. *BMC Bioinform.* **2013**, *14*, doi:10.1186/1471-2105-14-7.

Supplementary Information

Supplementary methods – mRNA extraction and sequencing of 34 larynx carcinomas.

mRNA extraction.

RNA was extracted from pre-treatment biopsies using the Roche High Pure miRNA Isolation Kit (REF: 05080576001). In summary, 5 sections of 5 µm thick each were deparaffinized and macrodissected, to guarantee that the sample contained of at least 50% tumor cells. The RNA was further purified according to the manufacturer's instructions.

Cleanup/Total RNA quality control

Quality and quantity of the total RNA was assessed with the 2100 Bioanalyzer using a Nano chip (Agilent, Santa Clara, CA). Following quantification, 1000 ng of each sample was subjected to an ethanol precipitation cleanup, ribosomal RNA depletion treatment and NGS library generation. Briefly, the samples were transferred to a 1.5-ml Eppendorf tube and the volume was adjusted to 180 µl, mixed with 18 µl of 3M Sodium Acetate with the addition of 1 µl of Glycogen (20 µg/µl) (Invitrogen, p/n 10814-010), followed by the addition of three volumes of ice-cold 100% ethanol , mixed and incubated overnight at -20 °C. The tubes were centrifuged at 13000 RCF for 1 hour and the supernatant discarded. The pellets were washed twice with 70% ethanol, air dried at room temperature for 10 minutes and dissolved in 30 µl of nuclease free water (Ambion p/n AM9937).

Ribo-Zero treatment

Ribosomal RNA sequences were depleted by treatment with the Ribo-Zero Magnetic Gold kit according to the manufacturer instructions (Epicentre, Cat no. MRZG12324), followed by an ethanol precipitation cleanup according to the manufacture instruction (Epicentre, Cat no. MRZG12324). The rRNA depleted fraction was subsequently used to generate NGS libraries suitable for the Illumina HiSeq2000 sequencing platform.

TruSeq library preparation

TruSeq cDNA libraries were generated using the TruSeq RNA Library Preparation Kit v2 sample preparation kit (Illumina Cat.No RS-122-2001/2) according to the manufacturer's instruction (Part # 15026495 Rev. D) with the following modifications. 1 µl of random primers (3 µg/µl), (Invitrogen p/n 48190-011) were added to the rRNA depleted fraction, followed

by an incubation at 65 °C for 5 minutes in a thermocycler and placed on ice, after which the manufacturer's protocol was further followed. Briefly, the RNA was reverse transcribed using SuperScript II Reverse Transcriptase (Invitrogen, part # 18064-014). Second strand synthesis was accomplished by using Polymerase I and RNaseH. The generated cDNA fragments were 3' end adenylated and ligated to Illumina Paired-end sequencing adapters and subsequently amplified by 15 cycles of PCR. The libraries were analyzed on a 2100 Bioanalyzer using a 7500 chip (Agilent, Santa Clara, CA), diluted and pooled equimolar into a 4-plex, 10 nM sequencing pool and stored at -20 °C.

Sequencing

The reads (51bp) were sequenced on a HiSeq2000 using V3 chemistry (Illumina Inc., San Diego), aligned against the human genome, build 37, using Tophat (version 2.0.6). Tophat allows to span exon-exon junctions. The samples were generated using a non-stranded library preparation protocol. Tophat was run with bowtie version 0.12.9 and supplied with a set of known gene models (GTF file, ensembl version 66) using the transcriptome index option. Other options supplied to Tophat were library typefr-unstranded, --prefilter-multihits and --no coverage. HTseq count was used to count the number of reads per gene. Only uniquely mapped reads were counted. Read counts of all samples were normalized to 10 million reads per sample. After normalization the expression values are log2 transformed. Upon log2 transformation, 1 was added to each expression value in order to avoid negative gene expression values.

TTCCATCAAGCCCTAGGGCTCTCTGTGGCTCTGGGAGTTGTAGTCTGAACGTTCTATCTTGGGAGAGAGCGCTACGCTCCCGCTACCGAGTCCCGCGGTAACTTTAAAGACCTGTACCGCCCCCGCCGCTGTGAGAGGGGAG

Supplementary Figures

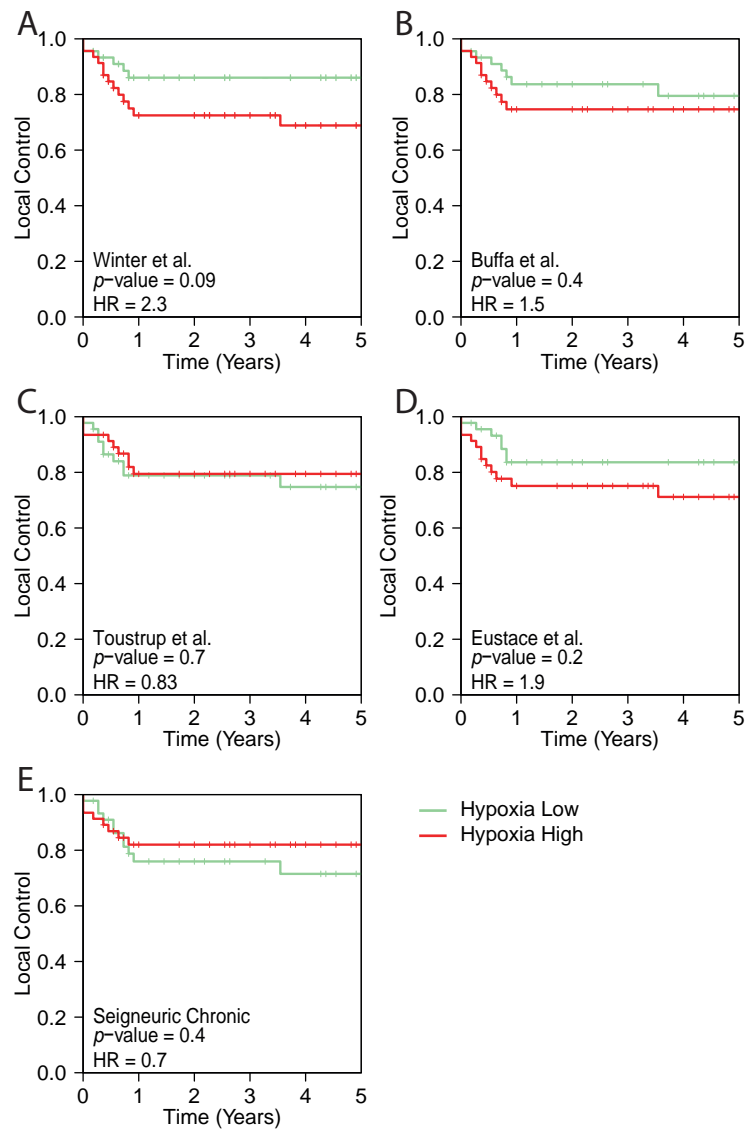


Figure S1. Univariable analyses of individual hypoxia profiles. A) Kaplan Meier curve of “High” versus “Low” hypoxia by Winter *et al.* in the Pramana cohort (N=91). The log-rank test was used to calculate the *p*-value. B, C, D, E depict the same as A except for Buffa, Toustrup, Eustace and Seigneure Chronic, respectively.

Acute Hypoxia Profile is a Stronger Prognostic Factor than Chronic Hypoxia

TTCCATCAAGCCCTAGGGTCTCTGTGGTGTGGGAGTTGTAGTGTGAAGCTTCTATCTTGGGAGAGGCTCTACGTCCTCCCTACCGAGTCCCGGGTAACTCTAAAGACCTGACCGCCCCCGCGCGCTGAGAGGGGAG

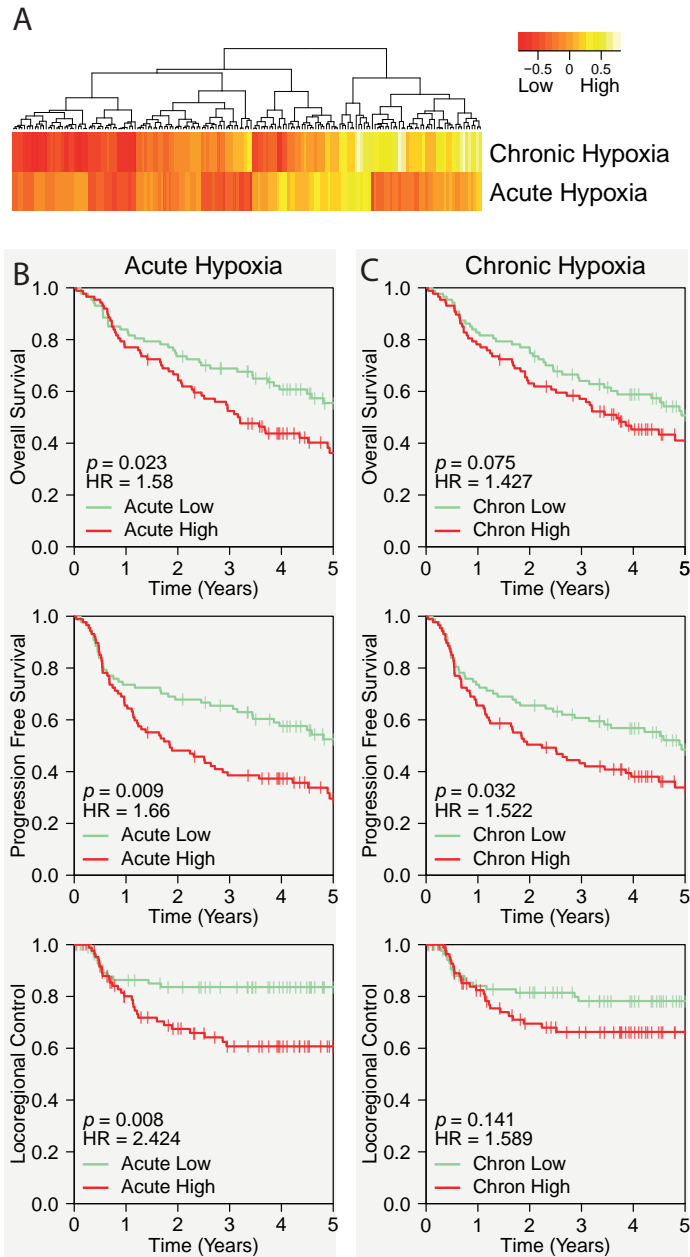


Figure S2. Distribution and outcome analyses of hypoxia in the Van der Heijden cohort. A) Heatmap of acute and chronic hypoxia scores. Seigneuric acute hypoxia represent acute hypoxia and chronic hypoxia is respresented by the joint chronic hypoxia score. B) Kaplan meier curves of Overall Survival, Progression Free Survival and Locoregional contol of acute hypoxa in the van der Heijden cohort (N=174). Statistics are calculated using a Cox proportional hazard model. A median split was used to divide patient into a “High” and a “Low” hypoxia group. C) Same as for B except for chronic hypoxia.

TTCCATCAAGCCCTAGGGTCTCTGTGGTCTGGGAGTTGTAGTCTGAAGCTTCTATCTTGGGAGAGAGGCTACGCTCCCCCTACCGAGTCCCGGGTAATCTTAAAGACCTGACCGCCCCCGCCGCTGAGAGGGGAG

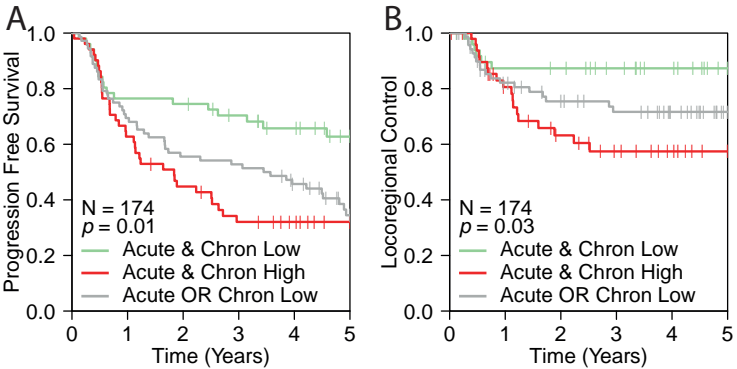


Figure S3. Combined acute and chronic hypoxia scores in the Van der Heijden cohort (N=174). A) Kaplan-Meier curve showing Progression Free Survival for 3 groups: (1) acute and chronic hypoxia both low, (2) acute or chronic hypoxia high, or (3) acute and chronic hypoxia both high. *P*-value represent the log-rank *p*-value. B) Same as A, except for locoregional control.

Supplementary Tables

Supplementary Table S1. Overview of the four clinical hypoxia gene expression profiles associated with patient outcome in HNSCC.

Year	Authors	Genes	Method used to identify the relevant genes
2007	Winter <i>et al.</i> ²⁵	99	Clustering of genes around 10 well-known hypoxia associated genes
2010	Buffa <i>et al.</i> ²⁶	51	Genes with similar expression pattern as well-known hypoxia associated genes in 3 HNSCC and 5 breast cancer series
2011	Toustrup <i>et al.</i> ¹³	15	<i>In vitro</i> experiments and correlation with Eppendorf hypoxia measurements
2013	Eustace <i>et al.</i> ²⁷	26	Reduced variant of Buffa <i>et al.</i> , validated in ARCON series

Supplementary Table S2. Genes in the used gene expression profiles

Winter	Buffa	Toustrup	Eustace	Seigneuric Acute	Seigneuric Chronic
ADORA2B	ACOT7	ADM	ALDOA	ACACA	ADAMTS13
AFARP1	ADM	ALDOA	ANGPTL4	ACOX1	ANGPTL4
AK3	AK3L1	ANKRD37	ANLN	ACSS1	ARID5B
ALDOA	ALDOA	BNIP3	BNC1	AIF1	ARL6IP1
ANGPTL4	ANKRD37	BNIP3L	C20orf20	AMH	C8orf4
ANKRD9	ANLN	C3orf28	CA9	ATF3	CASC2
ANLN	BNIP3	EGLN3	CDKN3	ATP5S	CDKN2D
AWR	C20orf20	KCTD11	COL4A6	ATXN7L1	CEP250
B4GALT2	CA9	LOX	DCBLD1	BACH1	CNOT2
BCAR1	CDKN3	NDRG1	ENO1	BET1L	CRABP2
BMS1L	CHCHD2	P4HA1	FAM83B	BTBD7	CSRP2
BNIP3	CORO1C	P4HA2	FOSL1	C12orf30	DDIT4
C14orf156	CTSL2	PDK1	GNAI1	C12orf53	DGCR8
C15orf25	DDIT4	PFKFB3	HIG2	CCL26	DOCK1
C20orf20	ENO1	SLC2A1	KCTD11	CCNH	DPYSL2
CA12	ESRP1		KRT17	CCT2	EHHADH
CA9	GAPDH		LDHA	CENTB2	EIF4E3
CDCA4P	GPI		MPRS17	COL6A3	ENO1
COL4A5	HIG2		P4HA1	CPEB4	ENO2
CORO1C	HK2		PGAM1	DKFZP761H1	FAM83A
CTEN	KIF20A		PGK1	DUSP3	FANCC
DKFZP564D166	KIF4A		SDC1	EIF4EBP2	FNDC3B
DPM2	LDHA		SLC16A1	FAM92A1	GPR75
EIF2S1	LRRC42		SLC2A1	GAS6	HIST1H4C
GAPD	MAD2L2		TPI1	GATAD2B	HNRPAO
GMFB	MAP7D1		VEGFA	GFOD2	INSIG1

Supplementary Table S2. Continued

Winter	Buffa	Toustrup	Eustace	Seigneuric Acute	Seigneuric Chronic
GSS	MCTS1			GRK6	JUN
HES2	MIF			GRM3	KALRN
HIG2	MRPL13			HIST1H2AL	KEAP1
HOMER1	MRPL15			HMMR	KIF1B
HSPC163	MRPS17			IFI6	KIRREL
IL8	NDRG1			IGF1R	KRT8
IMP-2	NP			IGFBP1	LOC645619
KCTD11	P4HA1			IGSF11	LOX
KIAA1393	PFKP			JMJD2A	MKNK2
KRT17	PGAM1			KIAA1219	MLL3
LDHA	PGK1			LOC64645 0	MXI1
LDLR	PSMA7			LOC728488	MYADM
LOC149464	PSRC1			LY86	NAV2
LOC56901	SEC61G			MAML2	NDRG1
LRP2BP	SHCBP1			NCAPH2	PCTK3
MGC14560	SLC16A1			NIT1	PDK1
MGC17624	SLC25A32			NSD1	PFKP
MGC2408	SLC2A1			OSTM1	PGK1
MGC2654	TPI1			PAPPA	PIM3
MIF	TUBA1B			PARG	PLOD1
MNAT1	TUBA1C			PCF11	PPP1R3B
MRPL14	TUBB6			PCSK1	PRKCB1
MRPS17	UTP11L			PEX14	PRKCBP1
MTX1	VEGFA			PHF10	PTGES
NDRG1	YKT6			PIK3R4	RAVER2
NME1				PRKAG2	SERINC5
NTMT1				PTPLAD1	SLC6A10P
NUDT15				RAB4B	SLC6A8
P4HA1				RBM4	SLCO4A1
PDZK11				RBPMS	SMEK1
PFKFB4				RHOBTB3	STRBP
PGAM1				RNASE4	SYTL2
PGF				SLC5A12	TCBA1
PGK1				SOX12	NEAT1
PLAU				SSH2	TSC22D2
PLEKHG3				ST3GAL1	UHRF1
PPARD				TERT	WSB1
PPP2CZ				TIMP2	ZFHX1B

Supplementary Table S2. Continued

Winter	Buffa	Toustrup	Eustace	Seigneuric Acute	Seigneuric Chronic
PPP4R1				TLE3	ZNF511
PSMA7				TP53111	
PSMB7				TRNT1	
PSMD2				TTLL5	
PTGFRN				ZNF117	
PVR				ZNF664	
PYGL					
RAN					
RNF24					
RNPS1					
RUVBL2					
S100A10					
S100A3					
SIP1					
SLC16A1					
SLC2A1					
SLC6A10					
SLC6A8					
SLCO1B3					
SMILE					
SNX24					
SPTB					
TEAD4					
TFAP2C					
TIMM23					
TMEM189					
TMEM30B					
TPBG					
TPD52L2					
TPI1					
TUBB2					
VAPB					
VEGF					
VEZT					
XPO5					

Supplementary Table S3. Patient characteristics Pramana cohort: 91 HNSCC stage III-IV chemoradiotherapy patients.

Variable	No of patients	%
Age (years)		
Mean	58	
Range	29-79	
Sex		
Male	61	67
Female	30	33
Tumor site		
Hypopharynx	20	22
Larynx	7	8
Oral cavity	13	14
Oropharynx	51	56
T-stage		
1	1	1
2	5	5
3	32	35
4	53	58
N-stage		
0	24	26
1	10	11
2	50	55
3	7	8
Local recurrences		
No	72	79
Yes	19	21
Median time (years)		
To local recurrence	0.5	
Follow-up without local recurrence	4.3	

Supplementary Table S4. Patient characteristics de Jong 1 cohort: 99 larynx/oropharynx radiotherapy patients.

Variable		No of patients	%
Age (years)			
	Mean	62.1	
	Range	42.3-85.5	
Sex			
	Male	84	85
	Female	15	15
Tumor site			
	Larynx	86	87
	Oropharynx	13	13
T-stage			
	1	22	22
	2	40	40
	3	35	35
	4	2	2
N-stage			
	0	83	84
	1	8	8
	2	6	6
	3	2	2

Supplementary Table S5. Patient characteristics de Jong 2 cohort: 34 larynx radiotherapy patients.

Variable		No of patients	%
Age (years)			
	Mean	67.2	
	Range	40.0-85.0	
Sex			
	Male	20	59
	Female	14	41
Tumor site			
	Larynx	34	100
T-stage			
	2	22	65
	3	12	35
N-stage			
	0	26	76
	1	1	3
	2	7	21

TTCCATCAAGCCCTAGGGCTCCTCTGTGGCTGCTGGGAGTTGTAGTCTGAACGCTTCTATCTTGGGAGAAAGCGCTACGCTCCCCCTACCGAGTCGCGGTAATCTTAAAGACCTGTACCGCCCCCGCGCTGTAGAGGGGAG

Supplementary Table S6. Patient characteristics of the van der Heijden cohort (NKI-CRAD and DESIGN combined): 174 advanced stage HPV-negative HNSCC patients treated with chemoradiotherapy. Age at diagnosis and follow-up time in years.

Variable	No of patients	%
Age (years)		
Median	61	
Range	40-80.2	
Sex		
Male	127	73
Female	47	27
Tumor site		
Larynx	31	18
Hypopharynx	66	38
Oropharynx	77	44
Stage		
I-III	35	20
IVA-IVB	139	80
T-stage		
T1-3	110	63
T4	64	37
N-stage		
N0-N2a	62	3
N2b-N3	112	64
Tumor Volume in cc. (N=149)		
Median	23,6	
Range	2.3-750	
Total cisplatin dose < 200 mg/m²		
Yes	63	37
No	107	63
Local Recurrences		
No	146	84
Yes	28	16
Follow up time (median years)		
To local recurrence	0.71	
Without local recurrence	4.00	

TTCCATCAAGCCCTAGGGCTCTCTGTGGTGTCTGGGAGTTGTAGTCTGAACGCTTCTATCTTGGGAGAAAGGCTACGCTCCCCCTACCGAGTCCCGGGTAATCTTAAAGACCTGACCGCCCCCGCGCGCTGTAGAGGGGAG

Supplementary Table S7. Univariate Cox proportional hazard analysis of parameters with patient outcome in the Van der Heijden cohort. OS: Overall Survival; PFS: Progression Free Survival; LC: Local Control; LRC: Locoregional Control; DM: Distant Metastasis Free Survival; HR: Hazard Ratio.

Variable	OS		PFS		LC		LRC		DM	
	HR	p-value	HR	p-value	HR	p-value	HR	p-value	HR	p-value
Gender										
Female	0,7	0,12	0,72	0,14	0,41	0,1	0,4	0,037	0,3	0,025
Male	REF		REF		REF		REF		REF	
Age at diagnosis	1,01	0,575	1	0,75	1,01	0,836	0,99	0,6	0,98	0,46
Tumor site										
Oropharynx	REF		REF		REF		REF		REF	
Hypopharynx	0,54	0,005	0,56	0,007	0,77	0,56	0,83	0,61	0,44	0,04
Larynx	0,57	0,07	0,8	0,43	1,6	0,31	1,7	0,18	0,94	0,87
Disease stage										
II-III	0,56	0,033	0,51	0,014	0,7	0,48	0,54	0,17	0,29	0,04
IVA-IVB	REF		REF		REF		REF		REF	
T classification										
T1-T3	REF		REF		REF		REF		REF	
T4	1,18	0,4	1,25	0,26	1,22	0,609	0,82	0,55	1,03	0,94
N classification										
N0-N2a	0,81	0,3	0,86	0,44	0,91	0,81	1,03	0,92	0,46	0,049
N2b-N3	REF		REF		REF		REF		REF	
Cumulative cisplatin < 200 mg/m²										
No	REF		REF		REF		REF		REF	
Yes	1,84	0,003	1,77	0,004	5,2	0,0001	2,74	0,002	1,2	0,58
Tumor Volume	1,006	5,81E-07	1,006	8,03E-06	1,006	0,26	1,005	0,26	1,009	0,022
Acute Hypoxia	2,36	0,029	2,66	0,009	6,07	0,008	5,51	0,003	3,7	0,035
Chronic Hypoxia	1,37	0,17	1,48	0,07	1,96	0,121	2,21	0,027	1,31	0,47

Supplementary Table S8. 10 hypoxia 'seed genes' and correlation with *in vitro* Seigneure acute and Seigneure chronic hypoxia profiles

Correlation with <i>in vitro</i> :	10 seed genes										Average correlation
	ADM	SLC2A1	PDK1	ENO1	HK2	PFKFB3	AK3	CCNG2	CA9	VEGF	
Seigneure Acute	-0.22	-0.17	-0.05	0.03	-0.18	-0.15	0.05	0.13	-0.05	-0.04	-0.06
Seigneure Chronic	0.48	0.45	0.23	0.48	0.36	0.17	-0.27	0.06	0.24	0.32	0.25



CHAPTER 3

TTCCCATCAACCCCTAGG99CTCCTCCT99CTGCTGGGAGTTGTAGTCTGAACTTCTATCTT99CGAGAA99C99CTAC99CTCC99CTACCGAGTCC99C99TAATCTTTAAAGCA99CTG99CA99CCCCCCCC99C99CCTG99

Epithelial-to-mesenchymal transition is a prognostic marker for patient outcome in advanced stage HNSCC patients treated with chemoradiotherapy

Martijn van der Heijden
Paul B.M. Essers
Caroline V.M. Verhagen
Stefan M. Willems
Joyce Sanders
Reinout H. de Roest
David M. Vossen
C. René Leemans
Marcel Verheij
Ruud H. Brakenhoff
Michiel W.M. van den Brekel
Conchita Vens

Abstract

Background: The prognosis of patients with HPV-negative advanced stage head and neck squamous cell carcinoma (HNSCC) remains poor. No prognostic markers other than TNM staging are routinely used in clinic. Epithelial-to-mesenchymal transition (EMT) has been shown to be a strong prognostic factor in other cancer types. The purpose of this study was to determine the role of EMT in HPV-negative HNSCC outcomes.

Methods: Pretreatment tumor material from patients of two cohorts, totaling 174 cisplatin-based chemoradiotherapy treated HPV-negative HNSCC patients, was RNA-sequenced. Seven different EMT gene expression signatures were used for EMT status classification and generation of HNSCC-specific EMT models using Random Forest machine learning.

Results: Mesenchymal classification by all EMT signatures consistently enriched for poor prognosis patients in both cohorts of 98 and 76 patients. Uni- and multivariate analyses show important HR of 1.6 to 5.8, thereby revealing EMT's role in HNSCC outcome. Discordant classification by these signatures prompted the generation of an HNSCC-specific EMT profile based on the concordantly classified samples in the first cohort (cross-validation AUC>0.98). The independent validation cohort confirmed the association of mesenchymal classification by the HNSCC-EMT model with poor overall survival (HR=3.39, $p<0.005$) and progression free survival (HR=3.01, $p<0.005$) in multivariate analysis with TNM. Analysis of primary HNSCC from PET-positive patients with metastatic disease prior to treatment further supports this relationship and reveals a strong link of EMT to the propensity to metastasize.

Conclusions: EMT in HPV-negative HNSCC co-defines patient outcome after chemoradiotherapy. The generated HNSCC-EMT prediction models can function as strong prognostic biomarkers.

Introduction

About half of head and neck squamous cell carcinoma (HNSCC) patients present with advanced stage disease. In the case of hypopharyngeal, laryngeal and HPV-negative oropharyngeal carcinomas, chemoradiotherapy is the current treatment of choice for these patients¹. Disease progression occurs in around 40%, stressing the need for additional treatment options for poor prognosis patients.

Tumor biology studies revealed an important role for HPV and hypoxia in HNSCC²⁻⁴. HPV-positive oropharyngeal tumors represent a distinct subgroup of HNSCC that is associated with a good prognosis whereas tumor hypoxia markers reveal patients with a poor prognosis⁵⁻⁷. Hypoxia biomarkers are currently being evaluated for standard clinical practice⁵. Yet, there is a need for additional prognostic markers in advanced stage HPV-negative HNSCC.

Epithelial to Mesenchymal Transition (EMT) is the process in which epithelial cells gain mesenchymal characteristics and has been linked to aggressive disease or disease progression in multiple cancer types^{8,9}. Many gene expression signatures have been developed that capture the gene expression profile differences between epithelial and mesenchymal tumor cells. Some have been generated by using cell material in pre-clinical studies, others by selecting genes highly correlated with well-known EMT markers¹⁰⁻¹⁶. A recent study that established a pan-cancer EMT signature helped to further characterize the role of EMT for clinical outcome¹². Despite this potential relevance, the role of EMT remains ill-defined in HNSCC. TCGA data analyses using a pan-cancer EMT signature show that EMT expression features are present in HNSCC¹². However, patient outcome studies have not been conducted^{11,12,17}. Assessed individually, decreased E-cadherin expression or increased expression of N-cadherin, β -catenin, Vimentin, TWIST and ZEB, all factors related to EMT, have been linked to a reduced overall survival in HNSCC¹⁸⁻²². However, others found no correlation²³, which questioned the role of EMT in patient outcome. Likewise, Chung *et al.* identified a poor prognosis expression profile which contains EMT-associated genes thereby indirectly implicating EMT in HNSCC outcomes²⁴.

Taken together, despite its recognized association with unfavorable tumor characteristics in other cancer types it is not clear to what extent EMT plays a role in HNSCC. Here we set out to determine the role of EMT and its prognostic value in advanced stage HPV-negative HNSCC patients treated with chemoradiotherapy. To this end we used RNA sequencing data of fresh frozen pretreatment HNSCC biopsy material and performed integrative analyses of multiple EMT gene expression profiles to assess their prognostic value in two HNSCC cohorts.

Material and Methods

Patients cohorts and material

Three different retrospective cohorts were used in this study. Specifics and patient characteristics are listed in the Sup-Information and Sup-Table S1. The NKI-CRAD cohort comprises 98 patients with advanced hypopharyngeal, laryngeal or HPV-negative oropharyngeal HNSCC who were treated with cisplatin-based chemoradiotherapy at the Netherlands Cancer Institute (NKI), Amsterdam, between 2001 to 2014. The validation cohort DESIGN comprises 76 HNSCC patients treated at the VUmc Cancer Center Amsterdam or NKI, between 2009-2014 and following identical inclusion and exclusion criteria to the NKI-CRAD cohort²⁵. NKI patients, diagnosed at the NKI between 2012 and 2017, with advanced stage HPV-negative head and neck cancer as in the other two cohorts but whose diagnostic PET-CT scan analyses showed distant metastasis prior to treatment. These cases were assigned to the NKI-MET cohort (N = 7). There is no overlap with the other cohorts. Treatment varied and distant metastasis at diagnosis was an exclusion criterion for the NKI-CRAD and DESIGN cohorts.

Biopsies and collection of fresh-frozen tumor material were approved by the Institutional Review Board of the Netherlands Cancer Institute. All patients signed informed consent for biopsy, HNSCC material collection and analysis. Sample tumor percentage was determined by a dedicated head and neck pathologist on H&E-stained slides. Samples with a tumor content of less than 40% or very low RNA quality were excluded from RNA sequencing and further analysis. HPV-status of all oropharyngeal carcinomas was determined by immunohistochemical assessment of p16 and RNA sequencing and HPV positive were excluded for this study²⁶.

RNA-sequencing and expression profile analyses

RNA was isolated and libraries were generated and sequenced generating 65 base single end reads on a HiSeq2500 (Illumina Inc., San Diego) which were mapped and filtered as detailed in the Sup-Information. All expression analyses were performed in R 3.4.3 using Rstudio 1.1. Seven different gene expression signatures were applied for epithelial or mesenchymal classifications of the tumor samples according to the author's protocols (specified and listed in Sup-Table S2)^{10–16}. RNA Sequencing data are available at <http://ega-archive.org> under study number EGAS00001004090 and datasets numbers EGAD00001005721, EGAD00001005715, EGAD00001005716, EGAD00001005717.

HNSCC-EMT prediction model generation

We applied different gene selection methods. Boruta²⁷ is a random forest based feature selection method. It validates the importance of each gene by comparing it to random

shuffled copies of the dataset. To increase robustness, this process is repeated multiple times. The “AUC” method calculates a ROC curve using the R-package pROC²⁸ for each gene, a measurement of how well the gene separates groups. With a pre-defined cutoff of 0.95, we selected all genes who classified 95% of the samples correctly. For each gene selection method a model was trained using random forest. Random forest is a machine learning algorithm which uses multiple random decision trees that result in an accurate prediction²⁹. The values as determined by the individual models were averaged and used when combining models. The HNSCC-EMT model is available as an R package at <https://github.com/M-tijn/HNSCCEMT>.

Tissue Microarray and Immunohistochemistry

Formalin Fixed Paraffin Embedded (FFPE) material from 92 patients was used to prepare tissue microarrays (TMA) with three cores per sample. Stroma content and the fraction of tumor cells positive for vimentin were assessed by a dedicated HNSCC pathologist on H&E and vimentin stained TMAs and averaged per patient sample.

Statistics

Survival analyses were performed using a cox-proportional hazard model with the “time” from the first day of treatment to the day the event was first detected. Tests were considered significant when $p < 0.05$. Primary tumor site and cumulative cisplatin dose were associated with patient outcome in univariate analyses (Sup-Table S3) and were therefore included in multivariate analyses. Wilcoxon Rank Sum tests and ANOVA were used to compare continuous variables. Correlations between continuous and categorical variables were assessed by Spearman correlation. Mean comparison between groups was assessed by a Student’s *t*-test.

Results

EMT has been shown to be associated with patient outcome in multiple cancer types, often by using EMT associated gene expression signatures. To characterize EMT in HNSCC and its association with outcome, we performed a transcriptomic analysis applying seven published EMT gene expression signatures (Sup-Table S2) on a cohort of 98 HPV-negative HNSCC patients treated with definitive cisplatin-based chemoradiotherapy from our institute (NKI-CRAD cohort, Sup-Table S1)^{10–16}. The distribution of each signature’s classification over the samples is depicted in Figure 1A and shows an overall poor agreement with few tumors being consistently classified as mesenchymal or epithelial respectively.

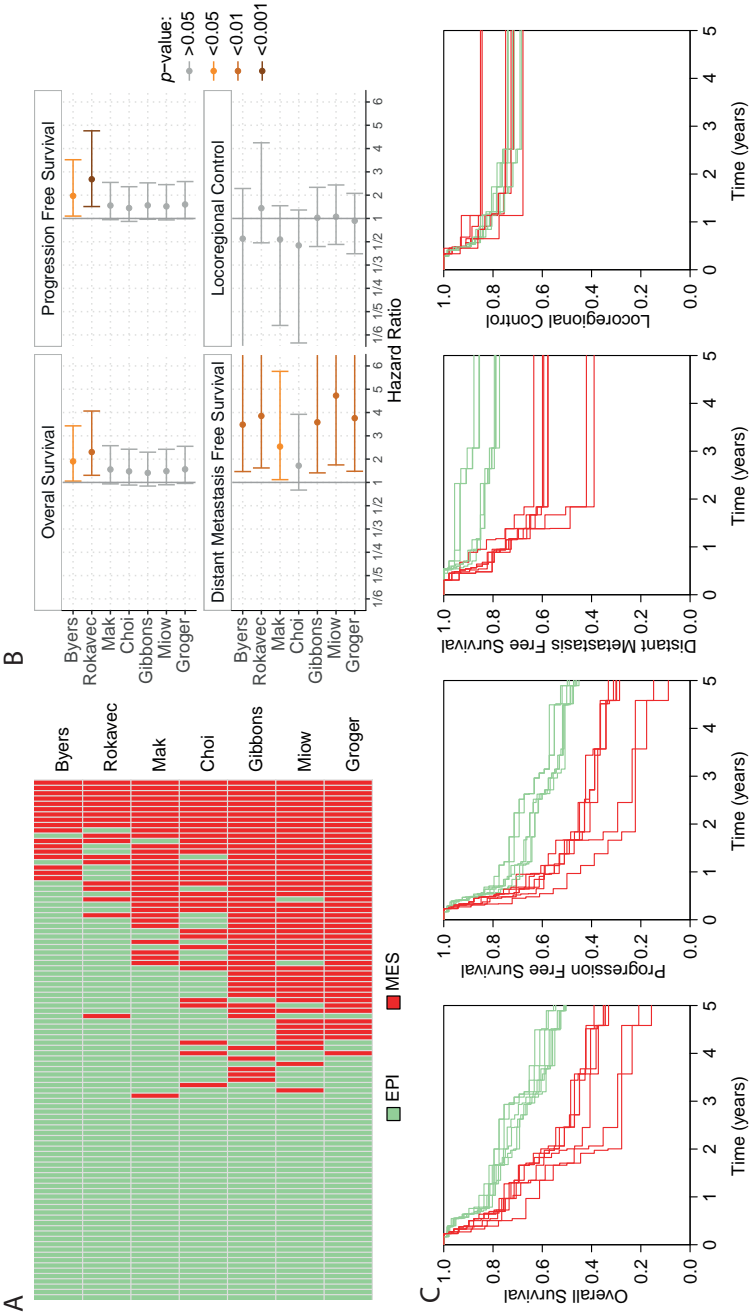


Figure 1. Non-HNSCC EMT signatures are associated with outcome in patients with HNSCC. A) Overview of patient sample assignment by each signature as indicated and specified in Materials and Methods (Table S2). Columns represent individual samples (N = 98) of patients with advanced stage HPV-negative HNSCC. Only 9 and 38 tumors were consistently classified as mesenchymal or epithelial respectively. Many of the remaining 51 tumors do however show a tendency towards either an epithelial or a mesenchymal classification. B) Forest plots with univariate cox proportional hazard analyses values for the epithelial versus the mesenchymal groups as determined by the indicated signature. Hazard ratios (HR) and 95% confidence intervals are shown for all recorded outcomes: overall survival (OS), progression free survival (PFS), and distant metastasis-free survival (DM), locoregional control (LRC) and are specified together with results from multivariable analyses in Table S4. Colors indicate statistical significance levels as shown to the right. C) Kaplan Meier OS, PFS, DM and LRC graphs comparing patients with mesenchymal (red) or epithelial (green) tumors as determined by the signatures and plotted together.

To test the association of EMT with patient outcome, we performed univariate cox proportional hazard analyses for each individual signature (Fig. 1B, Sup-Table S4) that show a statistically significant association with overall survival (OS) and progression free survival (PFS) for the Byers and Rokavec signatures. All signatures link mesenchymal classified tumors with poor outcome (OS, PFS and distant metastasis (DM)) and to a greater extent than TNM (Sup-Table S3 and Sup-Table S4). This indicates that EMT plays a role in HNSCC outcome (Fig. 1C). Multivariate analyses considering T-stage and N-stage (Sup-Table S4A), primary tumor site and cumulative cisplatin dose (Sup-Table S4B), confirmed that EMT is a significant prognostic factor for survival outcome (HR:1.3-5.8; p :0.001-0.219; Fig. S1, Sup-Table S4).

A recent study suggests that mesenchymal expression signals in bulk tumor material could in part be attributed to fibroblasts in the stroma^{30,31}. Correlations between EMT scores and tumor content of the samples, as determined on H&E stained slides, were however weak (Fig. S1C) and tumor content was also not associated with patient outcome ($p > 0.6$). Taken together, although created in different tumor types with varying methodologies, all EMT signatures consistently reveal a poor prognosis for the mesenchymal classified HNSCC groups.

Comparing different EMT signatures, many samples did not have a consistent epithelial or mesenchymal classification (Fig. 1A). It should however be acknowledged that the classification of individual tumor samples by these methodologies is cohort-dependent, i.e. a particular tumor sample, tested in a different cohort, could be classified differently. In order to approach a cohort-independent EMT assignment, we repeatedly selected 80% of all samples at random and re-classified these using the seven gene expression signatures. We find that all signatures provided relatively consistent results in each repetition round, except for Choi and Byers. (Fig. 2A, Sup-Table S2), indicating that the discrepancies were not a result of cohort size or composition. Two groups with a high consensus in the mesenchymal (C-MES, $N = 15$) or epithelial (C-EPI, $N = 39$) classification were identified based on these repeated analyses. Confirming the above revealed association, patients in the C-MES group have a significantly worse OS, PFS and DM free survival (Fig. 2B).

Prompted by the clinical relevance and the lack of consistent calling in nearly half of the samples, we next aimed to develop a HNSCC-specific EMT prediction model. Expression data of the C-MES ($N = 15$) and C-EPI ($N = 39$) samples were used to train a HNSCC-EMT prediction model as shown in Sup-Figure S2. Three methods were used to identify relevant genes (Sup-Table S5): Boruta²⁷, AUC or a selection based on literature ("keygenes") and a random forest²⁹ model was trained on each of these gene sets thereby obtaining three different models. A 10 times repeated 5-fold cross validation analyses reveal good performance of all three models (AUC > 0.98 , Sup-Table S6).

GTGGCAGGAGCCTCGAGGGGTTGATGGGATGGGGTTTTCCCTCCCTCATGTGCTCAAGACTGGCGCTAAAGTTTTGAGCTTCTCAAAAGTCTAGAGCCACCGTCCAGGGGAGGTAGTCTGGGTCCGGGGAACCTTGTCTGGG

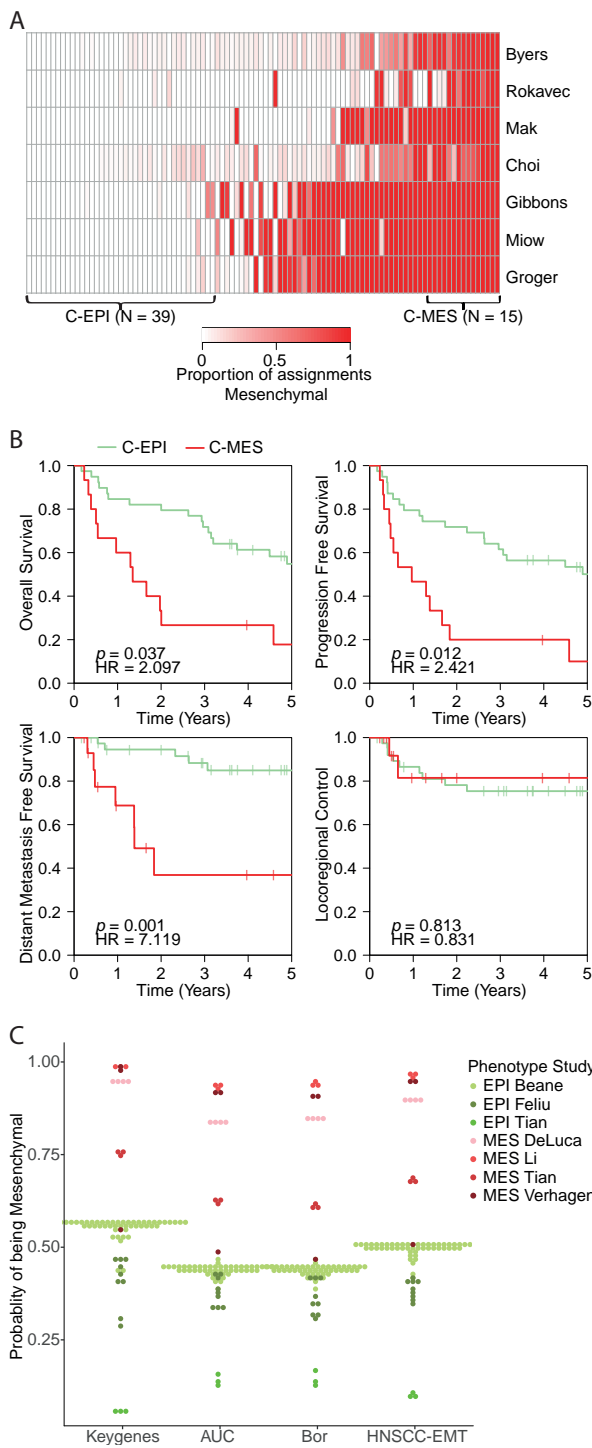


Figure 2. Cohort dependence of signature EMT assignments and generation of HNSCC-specific EMT models. A) Heatmap of mesenchymal (MES) classifications per sample after repeated application of each signature (on 1000 randomly selected sample sets constituting 80% of the original cohort). Two tumor groups were consistently classified in all repetitions and by all signatures. Consensus epithelial (C-EPI) or mesenchymal (C-MES) groups were established based on the proportion of “MES” assignments. Tumors with a minimum of 85% mesenchymal assignments were allocated to the C-MES tumor class, while C-EPI class samples were assigned EPI in more than 90% of the cases. B) Survival analysis of the C-EPI (N= 39) versus the C-MES group (N = 15). In-figure legends state univariate cox proportional hazard statistics values. C) Three EMT models (based on keygenes, AUC or Boruta selected gene selection as indicated) and the final HNSCC-EMT model (that combines all three) were tested. Model scores discriminate epithelial (green dots) or mesenchymal (red dots) cell lines. Publicly available expression data were used of cell lines of origin and study as indicated.

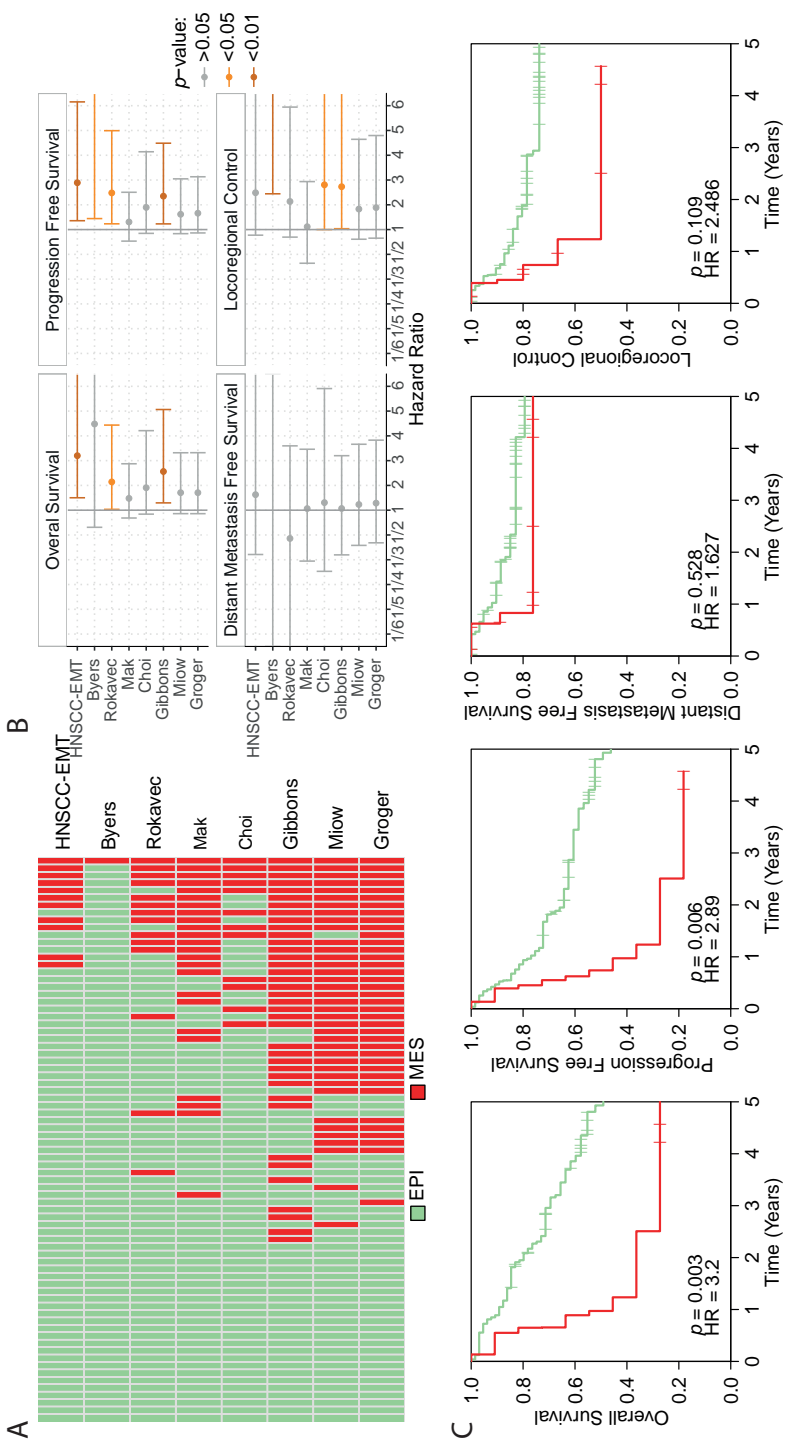
To evaluate whether these models still depict EMT, we tested their performance in independent cellular data sets with a verified EPI or MES status. All models were able to discriminate mesenchymal (MES) from epithelial (EPI) cells well when applied on publicly available RNA-sequencing data for several epithelial and mesenchymal cell lines^{32–37} (Fig. 2C). Unable to prioritize among the models, we constructed a final model that combines all three models by using the average value of all three models (HNSCC-EMT, Fig. 2C). As expected, when applied on the full NKI-CRAD cohort (N=98), the HNSCC-EMT model based classification is significantly associated with outcome in univariate analyses (OS: HR=1.9, $p=0.011$; PFS: HR=1.9, $p=0.011$; DM: HR=3.4, $p=0.003$; Sup-Fig. S3 and Table S4 for multivariate).

To summarize, an integrative EMT signature analysis provided a set of HNSCC samples with a congruent classification to train Random Forest models that generated a HNSCC-specific EMT prediction model.

A multicenter study cohort with samples from the NKI (N = 34) and the VU Medical Center (N = 42) was used (N = 76) to validate the poor prognosis association²⁵. Patient characteristics and univariate outcome association analyses are provided in Sup-Table S1 and S7. Fig. 3A shows MES or EPI assignments over the 76 patient samples as classified by the HNSCC-EMT model. We find a significant association with OS (HR=3.2, $p=0.003$) and PFS (HR=2.89, $p=0.006$) (Fig. 3B and 3C) that is confirmed in multivariate analyses (Sup-Table S8). Consistent with the first cohort, this stresses the relevance of EMT, i.e. mesenchymal tumor characteristics, for poor outcome in chemoradiotherapy treated patients with HPV-negative HNSCC.

Remarkably, the previously discovered link to DM was not evident in this validation cohort. To assess whether this was due to classification issues by the generated HNSCC-EMT prediction model, we used the original seven EMT signatures and tested them individually.

Further confirming a role for EMT, we find a consistent trend towards lower OS, PFS and also LRC (Fig. 3B, Sup-Table S8). None of the signatures is however associated with DM free survival indicating a cohort-specific issue (Fig. 3 and Sup-Fig. S4A). Comparing the DESIGN cohort with the NKI-CRAD cohort, we however noted that consistently fewer patients were classified as mesenchymal (Fig. 3A, Sup-Table S8) indicating a lower proportion of such tumors in the validation cohort. Guidelines changed in The Netherlands and resulted in the frequent use of PET-CT for advanced stage tumors during the diagnostic work-up from 2011 on. The improved detection of DM prior to treatment can therefore result in the exclusion of a fraction of patients that present with early metastatic disease at time of diagnosis. Prior routine PET-CT, these patients may have been instead included in definitive chemoradiotherapy regimens with curative intent. In line with the decreased number of MES-classified patients,



the exclusion of such PET-CT-positive HNSCC patients may have however biased cohort composition against early DM cases and therefore reduced the power to reveal an EMT association with DM in the DESIGN cohort. Together, from this we predicted that samples from a PET-positive HNSCC cohort might show a higher incidence of mesenchymal tumors. To test this, we searched for patients who would have initially met the inclusion criteria but were excluded due to distant metastasis on the PET-CT scan during the diagnostic work-up (NKI-MET cohort). RNA-sequencing was performed on available fresh-frozen primary tumor material of these seven patients. When applying the HNSCC-EMT prediction model, these showed high HNSCC-EMT scores (Fig. 4), explaining the discrepancy between the two cohorts and underlining the role of EMT in DM.

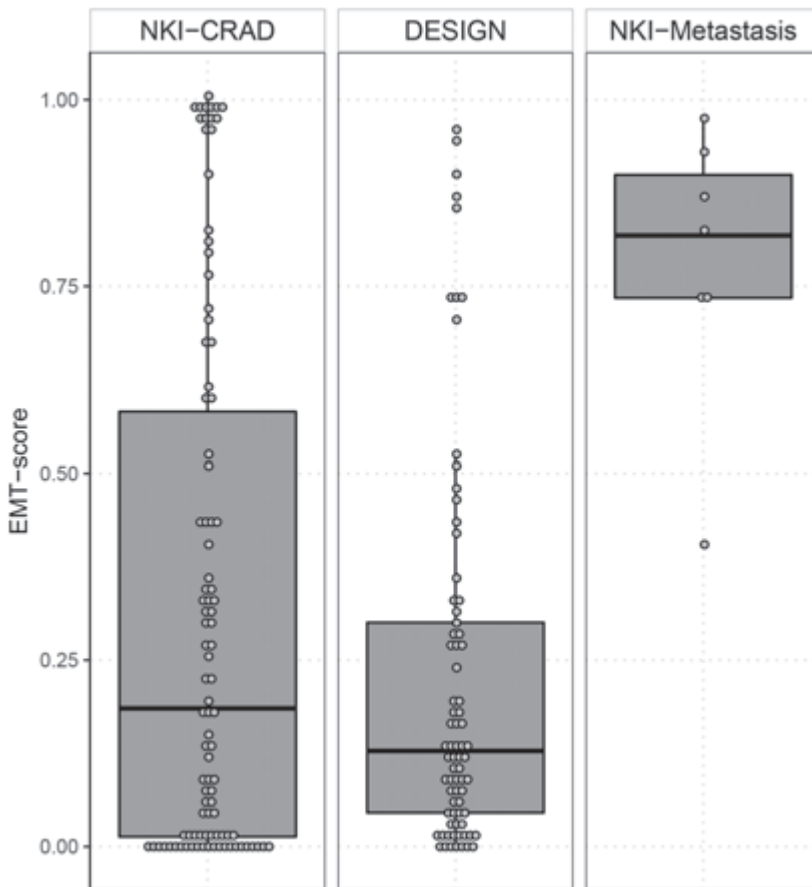


Figure 4. HNSCC-EMT model in a patient cohort with metastasized disease. A) A boxplot showing the distribution of the HNSCC-EMT prediction model scores in the NKI-CRAD, DESIGN and in primary HNSCC of the NKI-MET cohort with PET-scan positive patients prior to treatment.

When further evaluating important clinical parameters and potential discrepancies between the cohorts, we observed that DM free survival differs most in the oropharyngeal tumor site group (Sup-Fig. S5A). Overall survival is however similar in both cohorts and could indicate that the higher proportion of T4b tumors in the DESIGN cohort ($p=0.014$, Sup-Fig S5B) caused patients to decrease due to local tumor progression independent or prior to metastasis development or detection. Considering the observed link to EMT, this might have in part also reduced EMT-related DM events in the validation cohort.

Taken together, analyses in this DESIGN cohort validated the prognostic value of the HNSCC-EMT model. The high HNSCC-EMT prediction model scores of primary HNSCC in metastatic patients further highlights the strong poor prognosis association. This direct link implicates EMT in this process, beyond previously reported treatment resistance mechanisms.

The role of TNM, in contrast to EMT, in outcome is small and insignificant in both cohorts (Sup-Table S3 and S7), likely due to the advanced stage⁴. Multivariate analyses that include T-stage and N-stage provide increased HR and confirm the EMT association with OS and PFS in both cohorts ($p < 0.05$; Fig. 5; Sup-Table S4 and S8). We further tested for associations with clinical and pathological parameters (Sup-Fig. 7A) and find that T4 or those with non-cohesive growth pattern have higher HNSCC-EMT scores ($p=0.038$ and $p=0.019$ respectively). The presence of circulating tumor cells (CTC) has been linked to the occurrence of DM in many tumor types. We therefore applied the gene expression signature of Molloy *et al.*³⁸ that, based on transcriptomics data from primary breast cancers, identifies patients with high CTC numbers. CTC and HNSCC-EMT-scores correlate ($r=0.29$ and $p < 0.0001$) and the mesenchymal group, as classified by the HNSCC-EMT model, has higher CTC scores than the epithelial group ($p=0.0006$) (Sup-Fig. 7B-C). Together and consistent with our patient outcome findings and a higher propensity to metastasize, these data show that mesenchymal HNSCC have unfavorable characteristics such as high CTC markers, non-cohesive growth pattern and higher T-stage.

Tumor associated fibroblasts (TAF) are the largest stromal cellular component in HNSCC specimen. Importantly, a high TAF content has been previously associated with poor prognosis in HNSCC³⁹. We therefore questioned whether part of the poor outcome association was related to the mesenchymal expression signals from the stroma or whether it derived from mesenchymal tumor cells. To assess tumor EMT beyond stroma content associations (Fig. S1C), we determined the proportion of vimentin-positive tumor cells on immunohistochemistry (IHC) stained TMA (Fig. 6). RNA-Seq vimentin expression values correlate with the IHC-derived scores ($R: 0.359$, $p < 0.001$; Fig. 6C). The comparisons

confirm that HNSCC-EMT prediction model scores are not associated with stroma content ($p=0.567$; Fig. 6D) but instead with the proportion of vimentin-positive tumor cells in the sample ($p=0.053$; Fig. 6E), although not statistically significant. These analyses indicate that vimentin-positive tumor cells contribute to the HNSCC-EMT prediction model scores.

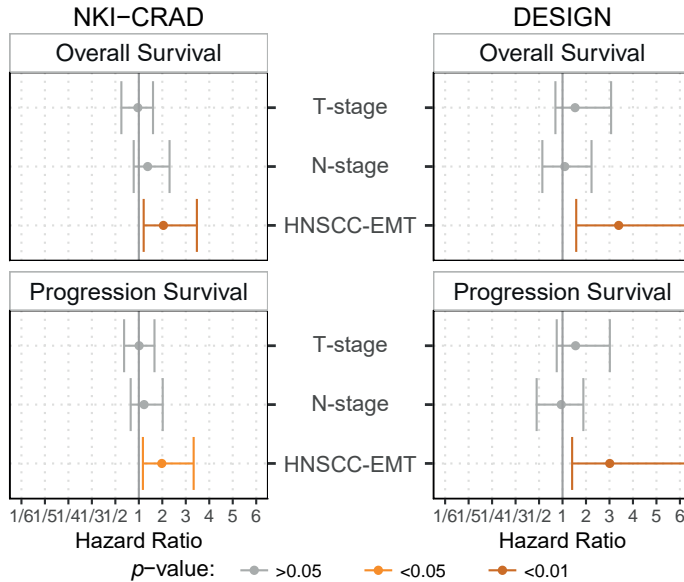


Figure 5. Multivariable outcome association analysis. Forest plots show the results of multivariate cox proportional hazard analyses with T-stage, N-stage and the HNSCC-EMT model classification in the NKI-CRAD and DESIGN cohort as indicated. Hazard ratios (HR) with 95% confidence intervals are shown for overall survival (OS) and progression free survival (PFS).

GTGCTAGGAGCCTCGAGGGGTTGATGGGATTGGGGTTTCCCTCCCATGTGTCTAAGACTGGCGCTAAAGTTTTGAGTTCTCAAAAGTCTAGAGCCACCGTCCAGGGAGAGGTAGTGTCTGGGTCCGGGGAACATTGTCTCGGG

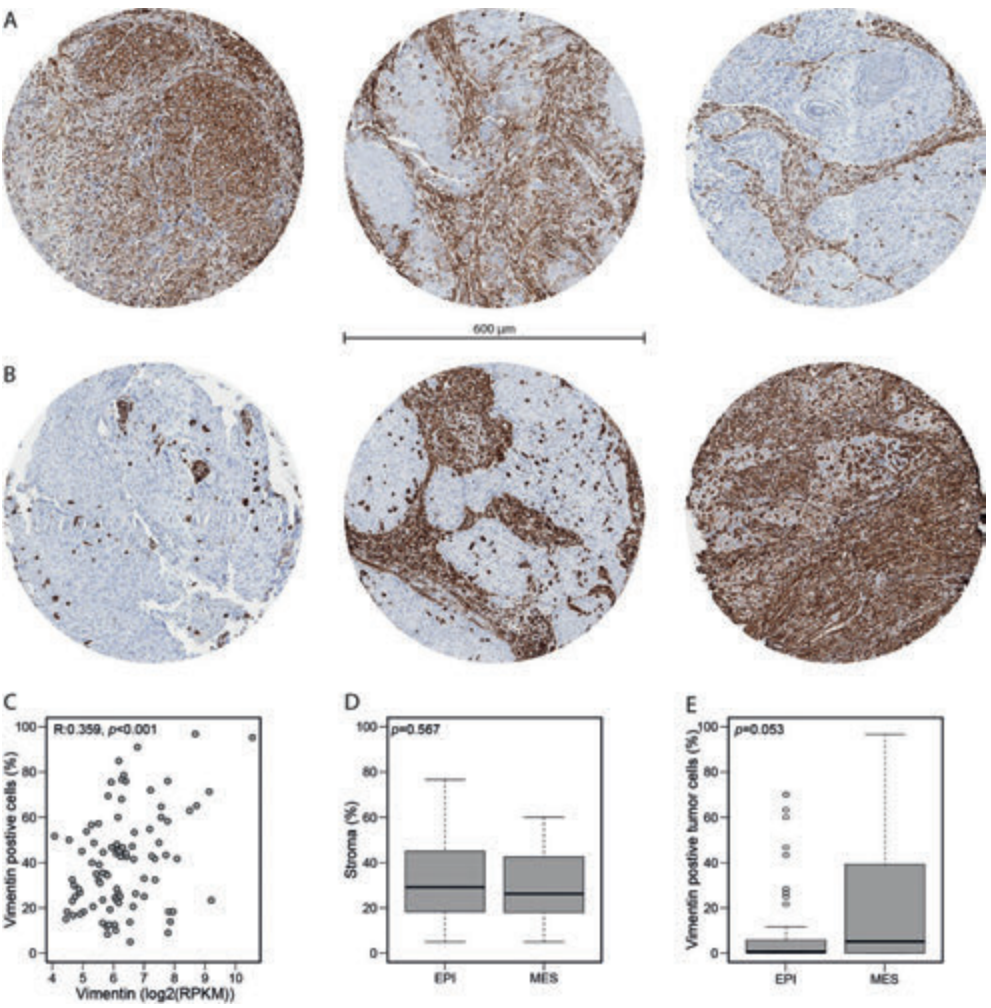


Figure 6. EMT and stromal contribution assessment by pathology A) Representative images of TMA samples stained for vimentin by immunohistochemistry (IHC). Examples of samples with more than 95% (left), around 50% (middle) and no (right) vimentin positive tumor cells are shown. B) Representative images of samples with different stroma contents. An example with no stroma content (left), 50% (middle) and >80% (right) stroma is shown. C) Correlation of the vimentin-positive cell scoring values by a pathologists with the vimentin expression data (in RPKM) from RNA-Seq showing that vimentin expression from the tumor cells largely contributes to the total sample expression values. D) A boxplot showing the distribution of the pathologist determined stroma content values in EPI and MES classified tumors as determined by the HNSCC-EMT prediction model. The lack of an association excludes that stroma content defines MES classification by the model. E) A boxplot showing the distribution of the percentage of vimentin-positive tumor cells in EPI and MES classified tumors as determined by the HNSCC-EMT prediction model. Student's *t*-test was used to calculate *p*-values.

Discussion

Here we set out to investigate the role of EMT in HPV-negative HNSCC. Transcriptomic analyses using seven published EMT signatures show that many tumors have mesenchymal characteristics. We show that EMT is associated with patient outcome in two cohorts of advanced HNSCC patients treated with chemoradiotherapy. A HNSCC-specific EMT prediction model was developed training a random forest model on samples with a consistent classification by all expression signatures. An independent validation cohort confirms that patient classification by this novel HNSCC-EMT model is associated with patient outcome. These data reveal the clinical importance of EMT in head and neck cancer. Increased scores in metastasized patients and associations with CTC markers and unfavorable growth pattern demonstrate an increased potential of mesenchymal HNSCC to metastasize.

EMT has been linked to poor prognosis in many cancer types^{9–12,18}. Evidence in HNSCC has been however scarce. Here we show that EMT, determined by multiple signatures and the novel HNSCC-EMT prediction model, is associated with patient outcome. The restriction to HPV-negative HNSCC, the use of transcriptomics and the application of machine learning techniques helped to reveal this link. Remarkable is the consistent poor prognosis association, independent of the used signatures and their origin and despite the discordance in sample classification. To improve EMT classification in HNSCC, we generated a HNSCC-specific EMT prediction model using random forest. An advantage of random forest is the computation of class probability values for individual patients that are therefore cohort-independent. Bound to impartial data presentation, we combined all three models but observe in retrospect that the performance of the three models is comparable and equally satisfactory in the validation cohort (Sup-Fig. S6). Importantly, we consistently observe that the HNSCC-EMT model performs well with respect to prognosis associations across the different cohorts. This also holds true when performing multivariate analyses using other clinical parameters such as TNM (Sup-Tables S4 and S8) and shows the benefit of the generation of tissue-specific (i.e. HNSCC-EMT) profiles. An improvement also observed for hypoxia biomarkers⁴⁰.

While the role for tumor associated fibroblasts (TAF) in HNSCC prognosis has been demonstrated in previous studies³⁹, here we were not able to reveal strong links to the stroma content. In comparison however, our study was strictly limited to HPV-negative and did not include oral cancers with a different prognostic genetic factor profile and treatment^{41–44}. It should be also noted that the restriction to minimal 40% tumor content for RNA-Seq analyses could have caused some bias against particularly TAF-rich tumors. In our study these constraints led to the exclusion of about 20% of the samples but remarkably none

of the identified PET-Scan positive patients had to be excluded, further supporting a, TAF-independent, role for mesenchymal tumor characteristics. The pathology data also indicate that the HNSCC-EMT prediction model captures features other than related to Vimentin, contributors to the robust and improved association with prognosis.

Allowing for organ preservation, chemoradiotherapy has become standard of care in advanced HNSCC in recent years. Unfortunately, none of the patients in the TCGA data have been treated with definitive chemoradiotherapy. In addition, only 25 and 5, HPV-negative oropharyngeal and hypopharyngeal respectively, tumors are currently present in the TCGA data set. Together this impedes any expression and association outcome studies for this patient population. Although limited to a total of 174 patients, our study is one of the largest full transcriptome RNA-sequencing studies in these HNSCC class. The cohorts are fairly unique in that the HNSCC patients were relatively uniformly treated with chemoradiotherapy which facilitates prognostic biomarker development within the context of contemporary treatment modalities. The two cohorts differ in terms of diagnosis and treatment years, which caused small variations in patient inclusion. The analysis of primary HNSCC material from patients with PET-CT confirmed early metastatic disease however confirmed the strong link to metastasis in a treatment independent manner. Lacking comparable data in other studies it is difficult to discern treatment resistance from poor prognosis. Our data however support the notion that mesenchymal HNSCC have a propensity to metastasize early, also prior treatment. The existence of highly scored patients and the association with locoregional control in the DESIGN cohort further suggests a prognostic value superior or at least complementary to PET-CT.

In conclusion, this study highlights the role of EMT for cancer progression and patient prognosis in HNSCC. Within advanced HPV-negative HNSCC, mesenchymal are pronounced and clinically relevant. Our data reveal a consistent and important poor prognosis association with several expression based EMT markers in three cohorts that motivated the generation of an HNSCC-specific EMT prediction model which performed well as prognostic biomarker in the validation cohort.

References

- Petersen JF, Timmermans AJ, van Dijk BAC, *et al.* Trends in treatment, incidence and survival of hypopharynx cancer: a 20-year population-based study in the Netherlands. *Eur Arch Oto-Rhino-Laryngology*. 2018;275(1):181-189. doi:10.1007/s00405-017-4766-6
- Leemans CR, Snijders PJF, Brakenhoff RH. The molecular landscape of head and neck cancer. *Nat Rev Cancer*. 2018;18(5):269-282. doi:10.1038/nrc.2018.11
- Baumann M, Krause M, Overgaard J, *et al.* Radiation oncology in the era of precision medicine. *Nat Rev Cancer*. 2016. doi:10.1038/nrc.2016.18
- Van Der Heijden M, Essers PBM, de Jong MC, *et al.* Biological determinants of chemo-radiotherapy response in HPV-negative head and neck cancer: a multicentric external validation. *Front Oncol*. 2019;9. doi:10.3389/fonc.2019.01470
- Toustrup K, Sørensen BS, Metwally MAH, *et al.* Validation of a 15-gene hypoxia classifier in head and neck cancer for prospective use in clinical trials. *Acta Oncol (Madr)*. 2016;55(9-10):1091-1098. doi:10.3109/0284186X.2016.1167959
- Li W, Thompson CH, O'Brien CJ, *et al.* Human papillomavirus positivity predicts favourable outcome for squamous carcinoma of the tonsil. *Int J Cancer*. 2003;106(4):553-558. doi:10.1002/ijc.11261
- van der Heijden M, de Jong MC, Verhagen CVM, *et al.* Acute Hypoxia Profile is a Stronger Prognostic Factor than Chronic Hypoxia in Advanced Stage Head and Neck Cancer Patients. *Cancers (Basel)*. 2019. doi:10.3390/cancers11040583
- Guinney J, Dienstmann R, Wang X, *et al.* The consensus molecular subtypes of colorectal cancer. *Nat Med*. 2015;21(11):1350-1356. doi:10.1038/nm.3967
- Tan TZ, Miow QH, Miki Y, *et al.* Epithelial-mesenchymal transition spectrum quantification and its efficacy in deciphering survival and drug responses of cancer patients. *EMBO Mol Med*. 2014;6(10):1279-1293. doi:10.15252/emmm.201404208
- Byers LA, Diao L, Wang J, *et al.* An epithelial-mesenchymal transition gene signature predicts resistance to EGFR and PI3K inhibitors and identifies Axl as a therapeutic target for overcoming EGFR inhibitor resistance. *Clin Cancer Res*. 2013;19(1):279-290. doi:10.1158/1078-0432.CCR-12-1558
- Mak MP, Tong P, Diao L, *et al.* A Patient-Derived, Pan-Cancer EMT Signature Identifies Global Molecular Alterations and Immune Target Enrichment Following Epithelial-to-Mesenchymal Transition. *Clin Cancer Res*. 2016;22(3):609-620. doi:10.1158/1078-0432.CCR-15-0876
- Gibbons DL, Creighton CJ. Pan-cancer survey of epithelial-mesenchymal transition markers across the Cancer Genome Atlas. *Dev Dyn*. 2018;247(3):555-564. doi:10.1002/dvdy.24485
- Gröger CJ, Grubinger M, Waldhör T, Vierlinger K, Mikulits W. Meta-Analysis of Gene Expression Signatures Defining the Epithelial to Mesenchymal Transition during Cancer Progression. *PLoS One*. 2012;7(12):1-10. doi:10.1371/journal.pone.0051136
- Rokavec M, Kaller M, Horst D, Hermeking H. Pan-cancer EMT-signature identifies RBM47 down-regulation during colorectal cancer progression. *Sci Rep*. 2017;7(1):1-15. doi:10.1038/s41598-017-04234-2
- Miow QH, Tan TZ, Ye J, *et al.* Epithelial-mesenchymal status renders differential responses to cisplatin in ovarian cancer. *Oncogene*. 2015;34(15):1899-1907. doi:10.1038/onc.2014.136
- Choi Y La, Bocanegra M, Kwon MJ, *et al.* LYN is a mediator of epithelial-mesenchymal transition and a target of dasatinib in breast cancer. *Cancer Res*. 2010;70(6):2296-2306. doi:10.1158/0008-5472.CAN-09-3141

17. Lawrence MS, Sougnez C, Lichtenstein L, *et al.* Comprehensive genomic characterization of head and neck squamous cell carcinomas. *Nature*. 2015;517(7536):576-582. doi:10.1038/nature14129
18. Nijkamp MM, Span PN, Hoogsteen IJ, Van Der Kogel AJ, Kaanders JHAM, Bussink J. Expression of E-cadherin and vimentin correlates with metastasis formation in head and neck squamous cell carcinoma patients. *Radiother Oncol*. 2011;99(3):344-348. doi:10.1016/j.radonc.2011.05.066
19. Zhu G, Song P, Zhou H, *et al.* Role of epithelial-mesenchymal transition markers E-cadherin, N-cadherin, β -catenin and ZEB2 in laryngeal squamous cell carcinoma. *Oncol Lett*. January 2018. doi:10.3892/ol.2018.7751
20. Luo Y, Yu T, Zhang Q, *et al.* Upregulated N-cadherin expression is associated with poor prognosis in epithelial-derived solid tumours: A meta-analysis. *Eur J Clin Invest*. 2018;(January):e12903. doi:10.1111/eci.12903
21. Zhang P, Hu P, Shen H, Yu J, Liu Q, Du J. Prognostic role of Twist or Snail in various carcinomas: A systematic review and meta-analysis. *Eur J Clin Invest*. 2014;44(11):1072-1094. doi:10.1111/eci.12343
22. Lefevre M, Rousseau A, Rayon T, *et al.* Epithelial to mesenchymal transition and HPV infection in squamous cell oropharyngeal carcinomas: The papillophar study. *Br J Cancer*. 2017;116(3):362-369. doi:10.1038/bjc.2016.434
23. Göppel J, Möckelmann N, Münscher A, Sauter G, Schumacher U. Expression of Epithelial-Mesenchymal Transition Regulating Transcription Factors in Head and Neck Squamous Cell Carcinomas. *Anticancer Res*. 2017;37(10):5435-5440. doi:10.21873/anticancer.11971
24. Chung CH, Parker JS, Ely K, *et al.* Gene expression profiles identify epithelial-to-mesenchymal transition and activation of nuclear factor- κ B Signaling as characteristics of a high-risk head and neck squamous cell carcinoma. *Cancer Res*. 2006;66(16):8210-8218. doi:10.1158/0008-5472.CAN-06-1213
25. Essers PBM, van der Heijden M, Verhagen CVM, *et al.* Drug sensitivity prediction models reveal link between DNA repair defects and poor prognosis in HNSCC. *Cancer Res*. 2019.
26. Lewis JS, Chernock RD, Ma XJ, *et al.* Partial p16 staining in oropharyngeal squamous cell carcinoma: Extent and pattern correlate with human papillomavirus RNA status. *Mod Pathol*. 2012. doi:10.1038/modpathol.2012.79
27. Kursu MB, Rudnicki WR. Feature Selection with the Boruta Package. *J Stat Softw*. 2010;36(11):1-13. doi:Vol. 36, Issue 11, Sep 2010
28. 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. doi:10.1186/1471-2105-12-77
29. Breiman L. Random Forrest. *Mach Learn*. 2001;5-32. doi:10.1023/A:1010933404324
30. Puram S V, Tirosh I, Parkh AS, *et al.* Single-Cell Transcriptomic Analysis of Primary and Metastatic Tumor Ecosystems in Head and Neck Cancer. *Cell*. 2017;172:1-14. doi:10.1016/j.cell.2017.10.044
31. Huelken J, Hanahan D. A Subset of Cancer-Associated Fibroblasts Determines Therapy Resistance. *Cell*. 2018. doi:10.1016/j.cell.2018.01.028
32. Li Y, Zhong C, Liu D, *et al.* Evidence for Kaposi sarcoma originating from mesenchymal stem cell through KSHV-induced mesenchymal-to-endothelial transition. *Cancer Res*. 2018;78(1):230-245. doi:10.1158/0008-5472.CAN-17-1961
33. Luca A De, Roma C, Gallo M, *et al.* RNA-seq analysis reveals significant effects of EGFR signalling on the secretome of mesenchymal stem cells. *Oncotarget*. 2014;5(21):10518-10528. doi:10.18632/oncotarget.2420

34. Tian B, Li X, Kalita M, *et al.* Analysis of the TGF β -induced program in primary airway epithelial cells shows essential role of NF-KB/RelA signaling network in type II epithelial mesenchymal transition. *BMC Genomics*. 2015;16(1). doi:10.1186/s12864-015-1707-x
35. Beane J, Mazzilli SA, Tassinari AM, *et al.* Detecting the presence and progression of premalignant lung lesions via airway gene expression. *Clin Cancer Res*. 2017;23(17):5091-5100. doi:10.1158/1078-0432.CCR-16-2540
36. Feliu N, Kohonen P, Ji J, *et al.* Next-generation sequencing reveals low-dose effects of cationic dendrimers in primary human bronchial epithelial cells. *ACS Nano*. 2015;9(1):146-163. doi:10.1021/nn5061783
37. Verhagen CVM, Vossen DM, Borgmann K, *et al.* Fanconi anemia and homologous recombination gene variants are associated with functional DNA repair defects in vitro and poor outcome in patients with advanced head and neck squamous cell carcinoma. *Oncotarget*. 2018;9(26):18198-18213. doi:10.18632/oncotarget.24797
38. Molloy TJ, Roepman P, Naume B, van't Veer LJ. A prognostic gene expression profile that predicts circulating tumor cell presence in breast cancer patients. *PLoS One*. 2012;7(2). doi:10.1371/journal.pone.0032426
39. Hanley CJ, Mellone M, Ford K, *et al.* Targeting the Myofibroblastic Cancer-Associated Fibroblast Phenotype Through Inhibition of NOX4. *J Natl Cancer Inst*. 2018. doi:10.1093/jnci/djx121
40. Halle C, Andersen E, Lando M, *et al.* Hypoxia-induced gene expression in chemoradioresistant cervical cancer revealed by dynamic contrast-enhanced MRI. *Cancer Res*. 2012. doi:10.1158/0008-5472.CAN-12-1085
41. Vossen DM, Verhagen CVM, Verheij M, Wessels LFA, Vens C, van den Brekel MWM. Comparative genomic analysis of oral versus laryngeal and pharyngeal cancer. *Oral Oncol*. 2018. doi:10.1016/j.oraloncology.2018.04.006
42. Vossen DM, Verhagen CVM, Van Der Heijden M, *et al.* Genetic factors associated with a poor outcome in head and neck cancer patients receiving definitive chemoradiotherapy. *Cancers (Basel)*. 2019. doi:10.3390/cancers11040445
43. Dogan S, Xu B, Middha S, *et al.* Identification of prognostic molecular biomarkers in 157 HPV-positive and HPV-negative squamous cell carcinomas of the oropharynx. *Int J cancer*. May 2019. doi:10.1002/ijc.32412
44. Gopinath D, Kunnath Menon R. Unravelling the molecular signatures in HNSCC: Is the homogenous paradigm becoming obsolete? *Oral Oncol*. 2018;82:195. doi:10.1016/j.oraloncology.2018.05.011

Supplementary Information

Supplementary figures

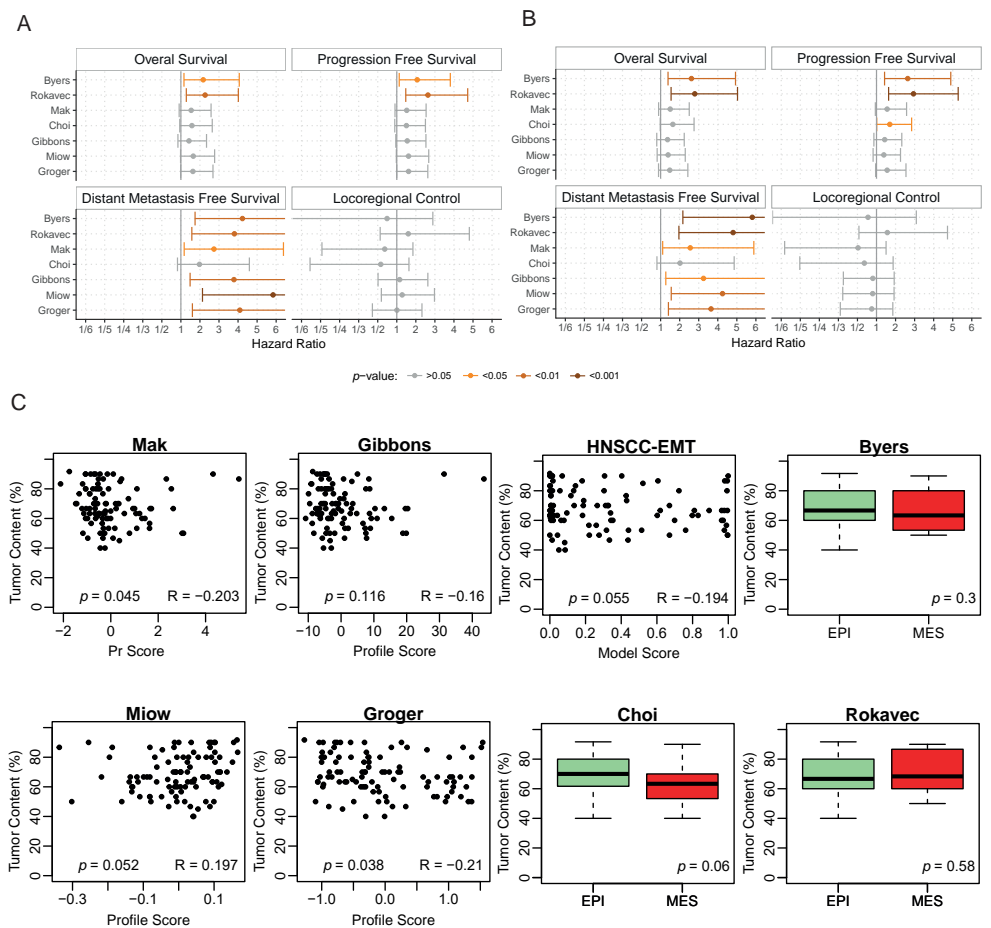


Figure S1. Multivariate analyses and tumor content associations for each EMT signature. A) Multivariate cox proportional hazard analyses with T-stage and N-stage are shown for each profile as represented in Fig 1A comparing the epithelial group versus the mesenchymal group for all recorded outcomes. Bars represent 95% confidence intervals and colors represent statistical significance level. B) Same as A, except with primary tumor site and cumulative cisplatin dose > 200 mg/m² as covariates in multivariate analysis. Statistics of A and B are listed in Supplementary Table S4. C) Tumor cell content of the samples was determined on H&E stained slides. The relation between tumor percentage and EMT classification by all EMT signatures and the HNSCC-EMT model was tested using Spearman correlations and Wilcoxon rank tests.

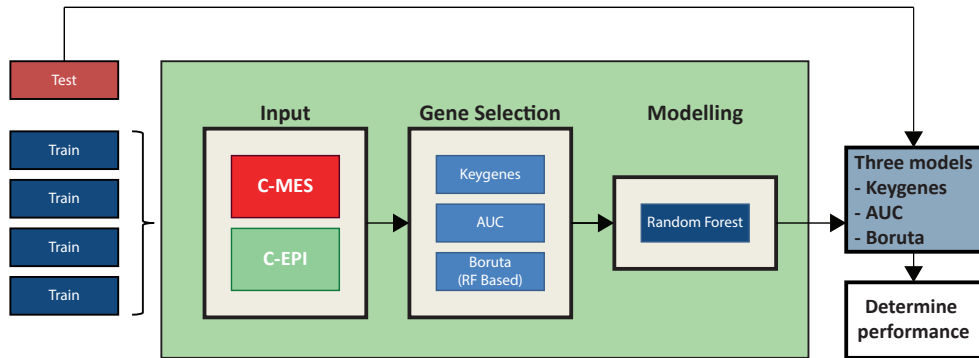


Figure S2. HNSCC-specific EMT model generation and cross validation. Expression profiles from the C-MES (N=15) and C-EPI (N=39) group samples served as gene selection and training set for HNSCC-specific EMT model generation. Using caret¹, the dataset was split into 5 folds, this was repeated 10 times, resulting in 50 different splits. Gene selection and model fitting was solely performed on the training folds.

To reduce the number of genes, three different selection techniques were employed. Boruta² generated permuted versions of variables and compares their importance to the original version, while selecting those variables where permutation causes a drop in importance. The “AUC” method calculates a ROC curve using the R-package pROC³ for each gene, a measurement of how well this gene separate the C-MES and C-EPI groups. With a pre-defined cutoff of 0.95, we selected all genes with an AUC of > 0.95. The selected genes are listed in Supplementary Table S5. Random forest has been reported to work well with high dimensional datasets and a random forest⁴ model was trained using caret. Cross validation results that estimate the performance of the models are provided in Supplementary Table S6.

1. Kuhn M. Building Predictive Models in R Using the caret Package. J Stat Softw. 2008;28(5):1-26. doi:10.1053/j.sodo.2009.03.002
2. Kursa MB, Rudnicki WR. Feature Selection with the Boruta Package. J Stat Softw. 2010;36(11):1-13. doi:Vol. 36, Issue 11, Sep 2010
3. 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. doi:10.1186/1471-2105-12-77
4. Breiman L. Random Forrest. Mach Learn. 2001:5-32. doi:10.1023/A:1010933404324

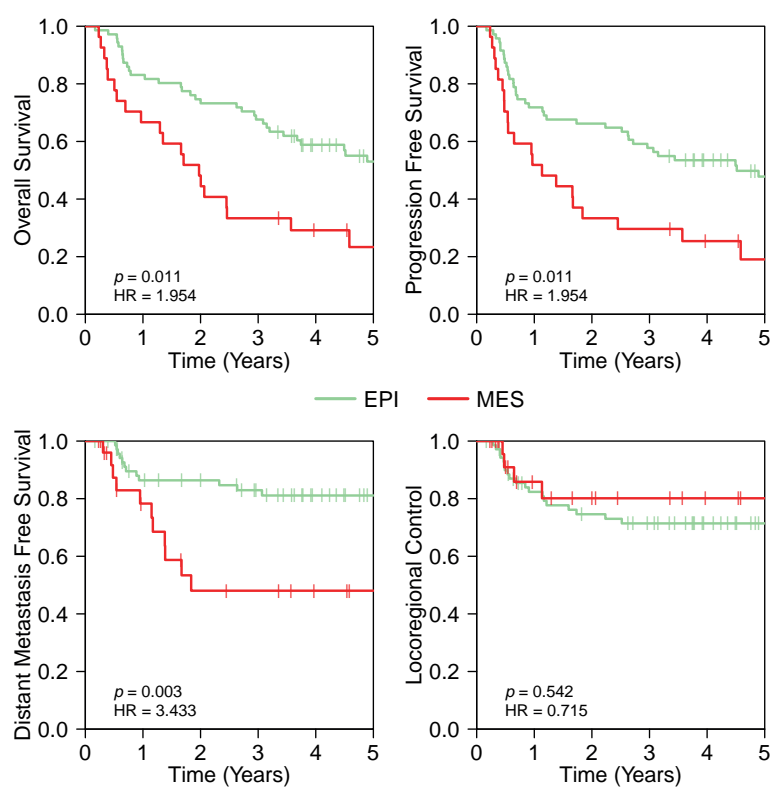


Figure S3. Kaplan Meier plots and survival analysis in the NKI-CRAD cohort as classified by the HNSCC-EMT model. Kaplan Meier curves are plotted for MES (red) or EPI (green) classified patients and for all recorded outcomes. Univariate cox proportional hazard statistics are shown. Univariate and multivariate statistic values are listed in Supplementary Table S4.

GACCTGGGTGTAGATGATGGGATGTTAGGACCATCCGAACTCAAGTTGAACGCTTAGGAGAGAGAGTGGAGCTTTGGGAACTTTGAGCCGGCTTAAAGCTATCTTTTGACATCCACCCGGTCTGGGCTAGGGAACTCCCTGAAT

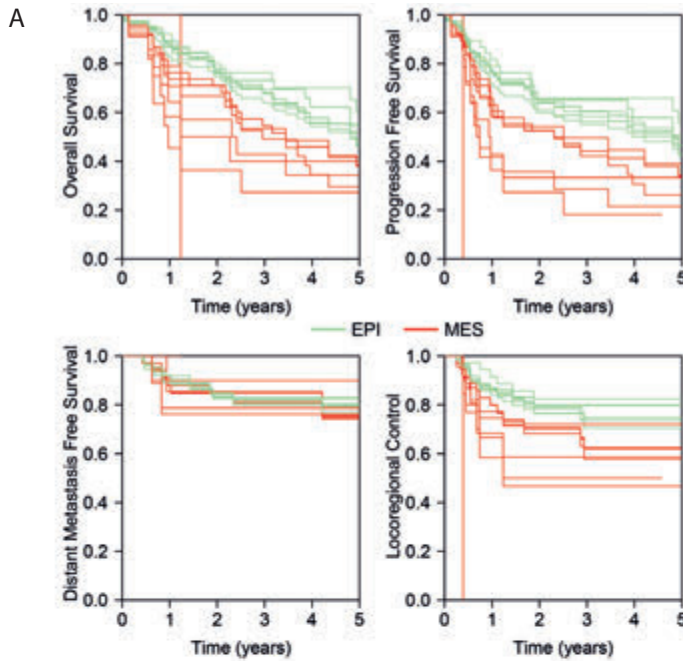
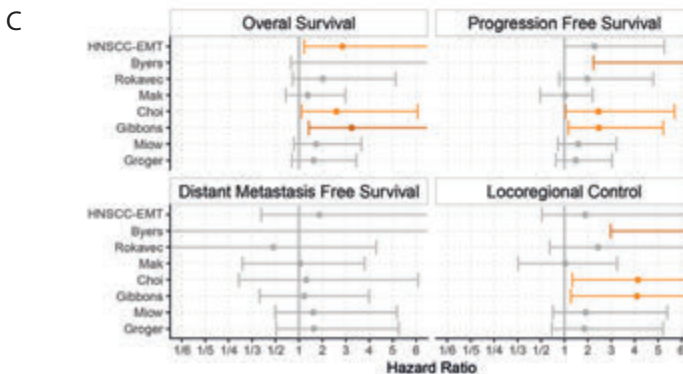
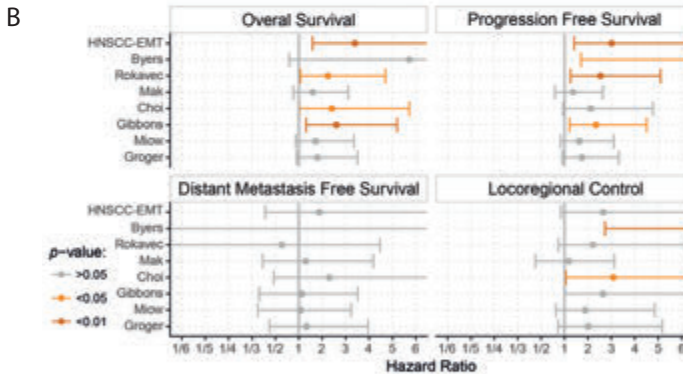


Figure S4. Kaplan Meier plots and survival analysis for all EMT signatures in the validation cohort A) Results on all recorded outcomes as indicated. Kaplan Meier curves as based on the individual signature classification are plotted. B) Multivariate cox proportional hazard analyses for the epithelial versus the mesenchymal groups were performed and results are shown. Multivariate analyses contain T-stage and N-stage as additional variables. Bars represent 95% confidence intervals and colors represent statistical significance level. C) Same as B except with primary tumor site and cumulative cisplatin dose as additional variables (statistics for B and C are also listed in Supplementary Table S7).



GTGCTAGAGAGCTCGAGGGGTTGATGGGATTGGGGTTTCCCTCCCTCATGTGCTCAAGACTGGCGCTAAAGTTTGTAGTTCTCAAAAGTCTAGAGCCACCGTCCAGGGAGCAGGTAGTGTCTGGGTCCGGGGAACATTTGTCTGGG

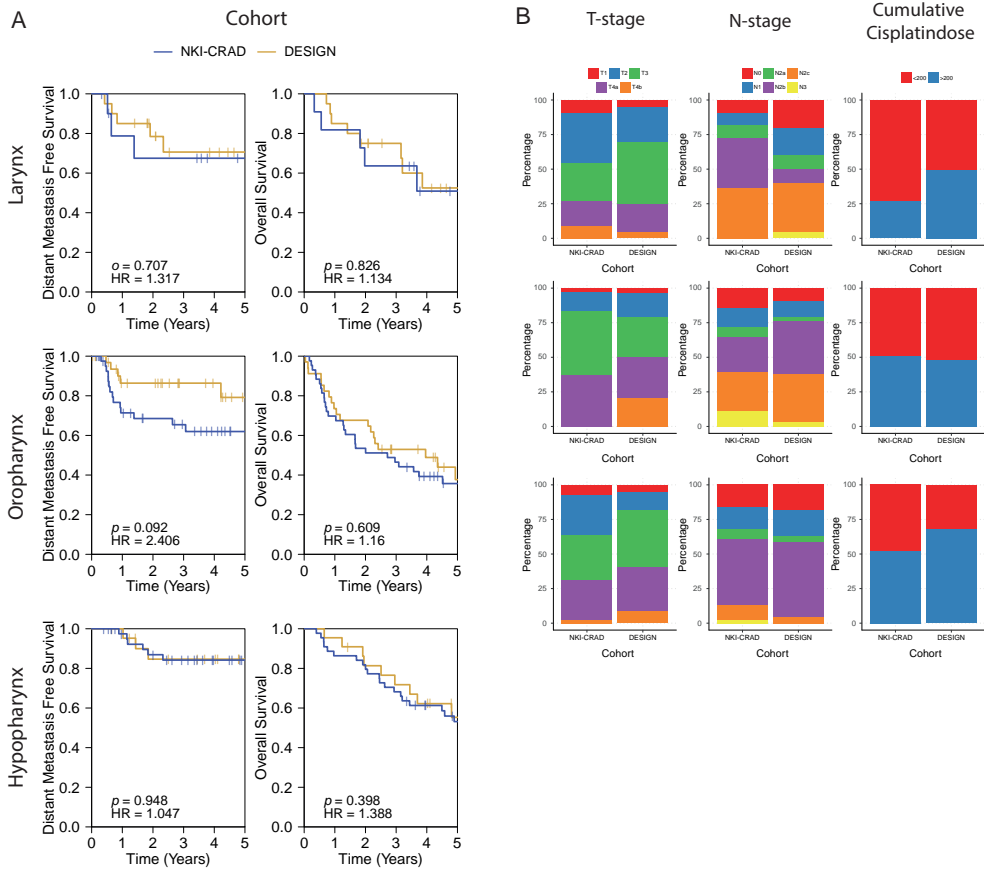


Figure S5. Evaluation of clinical parameters in the NKI-CRAD and DESIGN cohorts. A) Kaplan Meier curves for distant metastasis free survival and overall survival comparing both cohorts as indicated (NKI-CRAD blue, DESIGN ochreus) by primary tumor sites. In-figure legends state the univariable cox proportional hazard values. **B)** Overview showing T-stage and N-stage distribution over the samples and cohorts. Proportion of patients with high or low cumulative cisplatin doses in each subsite are shown for the two cohorts. Fisher's Exact tests were used to test the differences between the cohorts. From left to right: $p=0.859$, $p=0.859$, $p=0.275$ for larynx; $p=0.014$, $p=0.558$, $p=1$ for oropharynx and $p=0.470$, $p=0.965$, $p=0.277$ for hypopharynx, respectively.

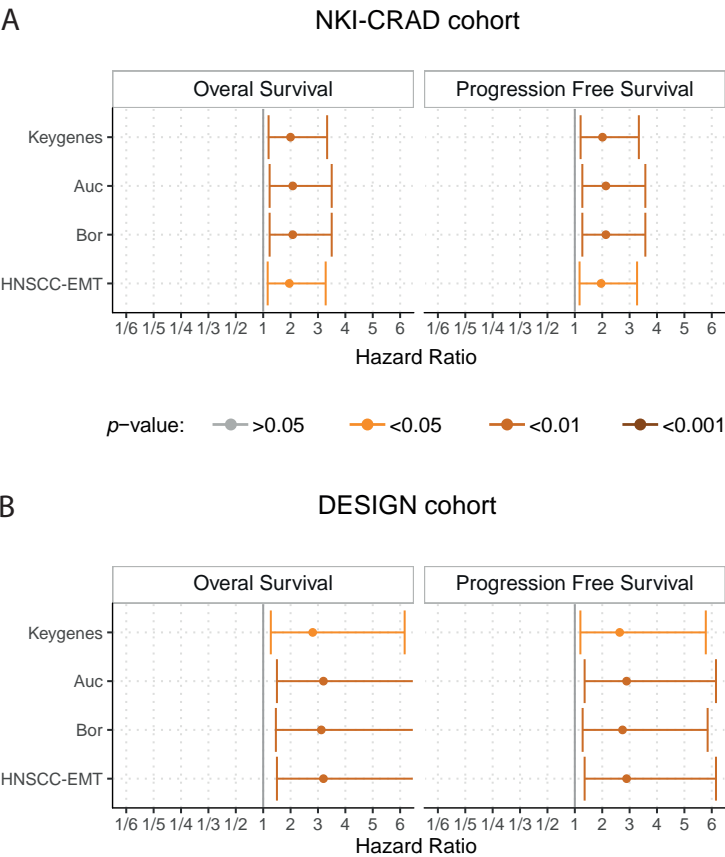


Figure S6. Performance of the individual models in the NKI-CRAD and DESIGN cohorts. A) Hazard ratio and confidence intervals resulting from classifications as based on the individual models in the NKI-CRAD cohort. The HNSCC-EMT score is the average of the MES probabilities of the three individual models combined. B) Same as in A, except showing results in the DESIGN cohort.

GTGCTAGAGGCTCTGAGGGGTTGATGGGATTGGGGTTTCCCTCCCAATGTCTCAAGACTGGCGCTAAAGTTTGTAGCTTCTCAAAAGTCTAGAGCCACCGTCCAGGGAGAGGTAGTCTCTGGGCTCCGGGGAACTTTGTCTCGGG

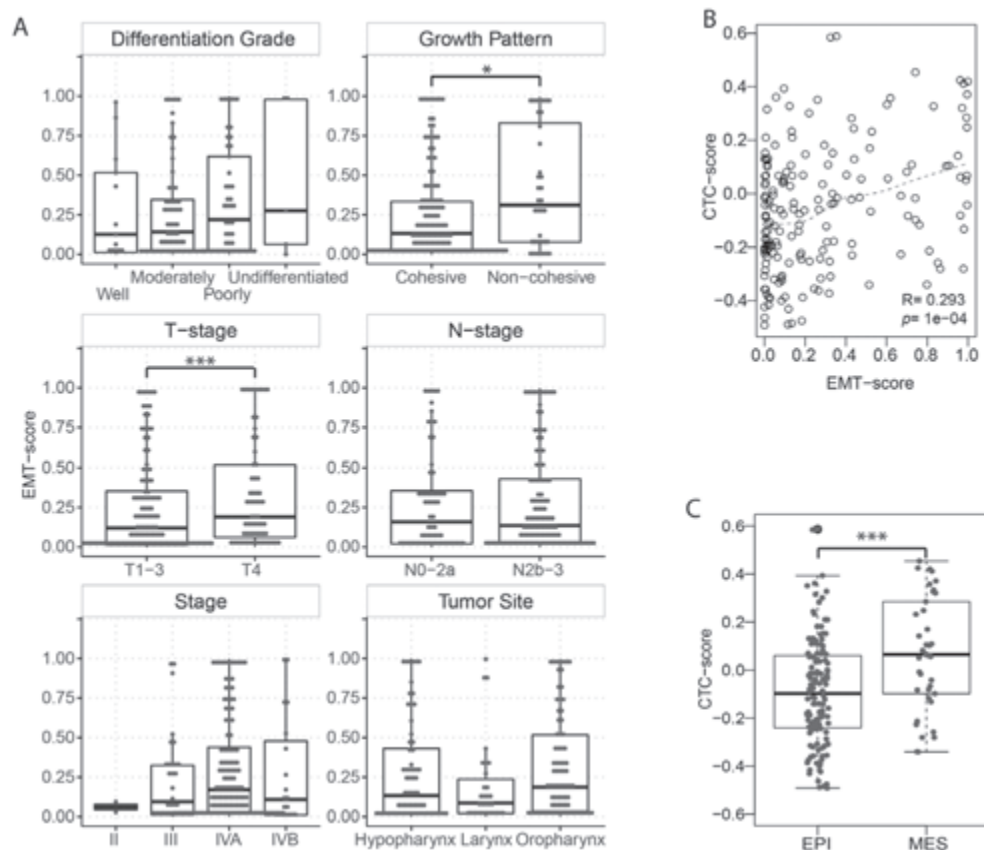


Figure S7. HNSCC-EMT model score association with clinical factors and circulating tumor cell scores. A) The relation of EMT scores with the clinical and pathological parameters differentiation grade ($p=0.49$), growth pattern, T-stage, N-stage ($p=0.82$), stage ($p=0.51$) and tumor site ($p=0.16$) was evaluated. Stars indicate significant changes in the EMT scores ($p=0.019$ for non-cohesive growth pattern and $p=0.038$ for T-stage). Wilcoxon Rank Sum tests were used to test growth pattern, T-stage and N-stage and ANOVA for differentiation grade, stage and tumor site. B) Circulating Tumor Cell-scores (CTC-score) as described in Molloy *et al.* against HNSCC-EMT scores. In-figure legend states spearman correlation ($R=0.29$, $p<0.0001$). C) Boxplots show CTC-scores in the epithelial or mesenchymal samples as classified by the HNSCC-EMT model. Stars indicate the t-test significance level ($p=0.0006$).

Supplementary tables

Supplementary Table 1. Patient and tumor characteristics of study cohorts. Age at diagnosis, follow-up and progression free survival (PFS) times in years.

Patient cohort	NKI-CRAD	DESIGN	
	N = 98	N = 76	
Variable	N (%)	N (%)	
<i>Gender</i>			
Female	26 (26.5)	21 (27.6)	<i>p</i> =1
Male	72 (73.5)	55 (72.4)	
<i>Median age at diagnosis (range)</i>	61 (40-78)	60 (44-80)	<i>p</i> =0.323
<i>Tumor site</i>			
Hypopharynx	44 (44.9)	22 (28.9)	<i>p</i> =0.154
Oropharynx	43 (43.9)	34 (44.7)	
Larynx	11 (11.2)	20 (26.3)	
<i>Disease stage</i>			
II	1 (1)	1 (1.3)	<i>p</i> =0.16
III	16 (16.3)	17 (22.4)	
IVA	74 (75.5)	46 (60.5)	
IVB	7 (7.1)	12 (15.8)	
<i>cT classification</i>			
T1-T3	65 (66.3)	45 (59.2)	<i>p</i> =0.420
T4	33 (33.7)	31 (40.8)	
<i>cN classification</i>			
N0-N2a	35 (35.7)	27 (35.5)	<i>p</i> =1
N2b-N3	63 (64.3)	49 (64.5)	
<i>Smoker</i>			
Former smoker	21 (21.4)	7 (9.2)	<i>p</i> =0.041
Never	1 (1)	3 (3.9)	
Unknown	5 (5.1)	1 (1.3)	
Yes	71 (72.4)	65 (85.5)	
<i>Alcohol consumption</i>			
Former alcoholic	16 (16.3)	4 (5.3)	<i>p</i> =0.023
Never	7 (7.1)	10 (13.2)	
Unknown	6 (6.1)	1 (1.3)	
Yes	69 (70.4)	61 (80.3)	
<i>Cumulative Cisplatin in mg/m² (range)</i>	200 (6-600)	270 (0-300)	<i>p</i> =0.838
<i>Number of PFS events (%)</i>	69 (70.4)	40 (52.6)	
<i>Number of DM events (%)</i>	23 (23)	13 (17)	
<i>Number of LRC events (%)</i>	23 (23)	19 (25)	
<i>Median PFS (95% CI)</i>	7.8 (5.6-10.5)	4.6 (4.3-5.6)	
<i>Median follow-up (95% CI)</i>	6.4 (5.6-10.1)	4.8 (4.3-5.6)	

Supplementary Table 2. Methods used for EMT classifications based on published signatures as indicated.

Origin of gene expression profile.			Application Parameters				
Profile	Cancer Types	Development - Signature Gene Selection Procedure	Platform	Input	Normalization	Method	Cut-off
Byers ¹⁰	Lung	Correlations between mRNA expression and the expression of four established EMT markers (<i>CDH1</i> , <i>CDH2</i> , <i>VIM</i> , <i>FN1</i>). Only genes with a bimodel index > 1.5 were selected.	Micro-array	RPKM	center and scale	Hierarchical Clustering with cuttree = 2.	Group with high <i>TWIST</i> expression = MES
	Rokavec ¹⁴ Pre-clinical and Clinical: Pan-cancer	Pre-clinical: differential expression between epithelial and mesenchymal cell-lines. Selected only $p<0.05$. Clinical: Spearman correlation of all mRNAs with <i>CDH1</i> or <i>VIM</i> . Only genes with $r>0.1$ or $r<-0.1$ were selected. Overlap of both, pre-clinical and clinical, data sets included in signature.	RNA-seq, CCLE / TCGA	RPKM	-	Kmeans clustering with K=2.	Group with high <i>TWIST</i> expression = MES
Mak ¹¹	Pan-Cancer	Computed Pearson correlations between all mRNAs and four established EMT markers (<i>CDH1</i> , <i>CDH2</i> , <i>VIM</i> , <i>FN1</i>). Obtaining a 77 gene signature.	RNA-seq, TCGA	RPKM	center and scale	Mean expression of mesenchymal genes minus mean expression of epithelial genes	Score > 0 = MES
Choi ¹⁶	Pre-clinical: Breast	Expression comparison of 5 mesenchymal breast cancer cell lines with 10 epithelial breast cancer cell lines. Top 100 overexpressed genes in the mesenchymal and epithelial breast cancer cell lines selected, obtaining a 200 gene signature.	Micro-array, 15 BC cell lines data	RPKM	center and scale	Hierarchical Clustering with cuttree = 2.	Group with high <i>TWIST</i> expression = MES
Gibbons ¹²	Pan-Cancer	Selection of 13 mesenchymal genes and 3 epithelial markers based on literature.	Literature	RPKM	center and scale	Sum of mesenchymal genes expressions minus sum of epithelial genes expressions.	Median

Supplementary Table 2. Continued

Origin of gene expression profile.			Application Parameters				
Profile	Cancer Types	Development - Signature Gene Selection Procedure	Platform	Input	Normalization	Method	Cut-off
Miow ¹⁵	Pre-clinical: Ovarian	Binary comparison 'epithelial' and 'mesenchymal' ovarian cancer cell lines. Selection based on with significance analysis (false-discovery rate q = 0) and receiver operating curves (ROC > 0.85).	Micro-array	Raw counts	GSVA, with option maseq=TRUE.	Principal Components Analysis. The final EMT score is the first principal component.	Median
Groger ¹³	Pre-clinical: Pan-cancer	Meta-analysis of 18 pre-clinical studies that performed differential expression analysis between epithelial and mesenchymal cell-lines. No HNSCC cell-lines.	Literature	Raw counts	GSVA, with option maseq=TRUE.	Score of upregulated genes minus score of downregulated genes in EMT.	Median

Supplementary Table 3. Univariate cox proportional hazard analysis of clinical parameters with patient outcome in the NKI-CRAD cohort.

OS: Overall Survival; PFS: Progression Free Survival; DM: Distant Metastasis Free Survival; LRC: Locoregional Control; HR: Hazard Ratio.

		OS		PFS		DM		LRC	
Variable		HR	p-value	HR	p-value	HR	p-value	HR	p-value
Gender									
	Female	REF		REF		REF		REF	
	Male	1.37	0.27	1.33	0.30	1.94	0.22	2.84	0.09
Age at diagnosis									
		0.99	0.71	0.99	0.61	0.94	0.04	0.99	0.95
Tumor site									
	Hypopharynx	REF		REF		REF		REF	
	Oropharynx	1.86	0.02	1.62	0.06	3.09	0.02	1.05	0.90
	Larynx	1.08	0.87	2.10	0.07	2.52	0.19	3.06	0.03
Disease stage									
	II	-		-		-		-	
	III	0.53	0.09	0.47	0.05	0.71	0.59	0.57	0.36
	IVA	REF		REF		REF		REF	
	IVB	1.51	0.31	1.26	0.56	2.79	0.06	0.52	0.53
cT classification									
	T1-T3	REF		REF		REF		REF	
	T4	1.07	0.80	1.13	0.62	0.70	0.46	0.41	0.11
cN classification									
	N0-N2a	REF		REF		REF		REF	
	N2b-N3	1.27	0.34	1.18	0.49	1.37	0.48	0.93	0.86
Smoker									
	Former smoker	0.64	0.19	0.60	0.13	0.60	0.41	0.54	0.32
	Never	-		-		-		-	
	Unknown	1.83	0.25	1.61	0.36	2.78	0.17	3.4x10 ⁻⁸	1
	Yes	REF		REF		REF		REF	
Alcohol consumption									
	Former alcoholic	1.64	0.11	1.48	0.20	2.08	0.13	2.60	0.04
	Never	0.42	0.17	0.39	0.12	1.37x10 ⁻⁸	1	1.35x10 ⁻⁸	1
	Unknown	2.46	0.06	2.14	0.11	2.96	0.15	1.35	0.76
	Yes	REF		REF		REF		REF	
Cumulative cisplatin > 200 mg/m ²									
	No	REF		REF		REF		REF	
	Yes	0.577	0.03	0.59	0.03	0.65	0.31	0.29	0.01

GACCTGGGTGTAGATGATGGGATGTAGGACCCATCCGAACTTAAAGTTTGAAAGCCCTAGGAGAGAGGAGTGGAGCTTTGGGGAACTTTGAGCCGGCTAAAGCGTACTTTTTCACATCCACACCGGTGTGGGGTGTAGGGAATCCCTGTGAAT

Supplementary Table 4. Cox proportional hazard model results for the profiles in NKI-CRAD cohort (Univariable and Multivariable).

Multivariable analysis A includes T-stage and N-stage

Multivariable analysis B includes cumulative cisplatin dose and tumorsite

* = number of samples classified as epithelial (EPI) or mesenchymal (MES).

= mean standard deviation values depict the distribution pattern of the classifications by the individual profiles over the 1000 repetitions.

Univariable	Overall Survival				Progression Free Survival				Distant Metastasis Free Survival				Locoregional control				# Mean SD			
	EPI*	MES*	p value	HR	Lower	Upper	p value	HR	Lower	Upper	p value	HR	Lower	Upper						
Byers	81	17	0.032	1.903	1.056	3.427	0.023	1.966	1.1	3.514	0.005	3.481	1.466	8.263	0.399	0.535	0.125	2.286	0.2168	
	80	18	0.004	2.303	1.305	4.064	0.001	2.645	1.49	4.694	0.002	3.852	1.621	9.156	0.508	1.441	0.489	4.249	0.0638	
	66	32	0.085	1.557	0.941	2.576	0.079	1.556	0.951	2.548	0.026	2.54	1.118	5.767	0.243	0.526	0.179	1.546	0.0256	
	65	33	0.125	1.476	0.898	2.425	0.142	1.444	0.885	2.356	0.199	1.718	0.753	3.923	0.162	0.464	0.158	1.363	0.2333	
	Gibbons	49	49	0.169	1.41	0.865	2.298	0.066	1.57	0.971	2.539	0.007	3.582	1.41	9.101	0.945	1.029	0.454	2.336	0.0502
	Mlow	49	49	0.115	1.483	0.908	2.423	0.083	1.528	0.947	2.466	0.002	4.727	1.751	12.758	0.865	1.074	0.473	2.438	0.0381
	Groger	49	49	0.073	1.564	0.959	2.551	0.05	1.612	0.999	2.6	0.005	3.759	1.479	9.553	0.821	0.909	0.398	2.077	0.0488
HNSCC-EMT Model	71	27	0.011	1.954	1.163	3.281	0.011	1.954	1.169	3.264	0.003	3.433	1.508	7.816	0.542	0.715	0.243	2.104		
Multivariable A	Overall Survival				Progression Free Survival				Distant Metastasis Free Survival				Locoregional control				# Mean SD			
	EPI*	MES*	p value	HR	Lower	Upper	p value	HR	Lower	Upper	p value	HR	Lower	Upper	p value	HR		Lower	Upper	
Byers	81	17	0.015	2.173	1.162	4.064	0.019	2.073	1.128	3.809	0.001	4.23	1.744	10.259	0.583	0.662	0.151	2.893		
	80	18	0.005	2.272	1.285	4.016	0.001	2.636	1.471	4.722	0.003	3.808	1.574	9.213	0.4	1.603	0.534	4.808		
	66	32	0.103	1.539	0.917	2.584	0.106	1.524	0.914	2.539	0.02	2.736	1.173	6.383	0.384	0.612	0.203	1.848		
	65	33	0.083	1.58	0.943	2.646	0.116	1.505	0.904	2.503	0.114	1.974	0.848	4.595	0.279	0.543	0.18	1.638		
	Gibbons	49	49	0.164	1.425	0.866	2.346	0.079	1.55	0.95	2.53	0.005	3.786	1.484	9.659	0.738	1.152	0.504	2.631	
	Mlow	49	49	0.058	1.65	0.983	2.771	0.062	1.617	0.976	2.677	0.001	5.838	2.133	15.976	0.562	1.283	0.553	2.979	
	Groger	49	49	0.054	1.631	0.991	2.684	0.051	1.619	0.998	2.628	0.003	4.095	1.604	10.454	0.983	1.009	0.439	2.321	
HNSCC-EMT	71	27	0.008	2.047	1.206	3.473	0.01	1.978	1.173	3.335	0.002	3.635	1.589	8.317	0.744	0.833	0.28	2.483		

Supplementary Table 4. Continued

Multivariable B	Overall Survival				Progression Free Survival				Distant Metastasis Free Survival				Locoregional control			
	EPI*	MES*	p value	HR	Lower	Upper	p value	HR	Lower	Upper	p value	HR	Lower	Upper	p value	HR
Byers	81	17	0.003	2.617	1.389	4.93	0.002	2.643	1.424	4.906	0	5.81	2.167	15.575	0.629	0.692
Rokavec	80	18	0.001	2.793	1.549	5.036	0	2.888	1.607	5.19	0.001	4.804	1.959	11.776	0.424	1.568
Mak	66	32	0.116	1.504	0.904	2.502	0.084	1.558	0.942	2.576	0.028	2.555	1.108	5.892	0.221	0.508
Choi	65	33	0.06	1.642	0.978	2.755	0.044	1.698	1.014	2.845	0.117	2.018	0.838	4.86	0.388	0.61
Gibbons	49	49	0.219	1.364	0.831	2.239	0.141	1.441	0.886	2.343	0.014	3.251	1.272	8.307	0.679	0.839
Mlow	49	49	0.193	1.393	0.846	2.293	0.178	1.401	0.858	2.286	0.005	4.247	1.557	11.59	0.668	0.833
Groger	49	49	0.121	1.48	0.902	2.428	0.065	1.589	0.972	2.599	0.008	3.647	1.409	9.438	0.603	0.8
HNSCC-EMT Model	71	27	0.006	2.118	1.246	3.6	0.005	2.125	1.254	3.601	0.002	3.777	1.638	8.713	0.631	0.763
																2.301

Supplementary Table 5. Genes selected via the three different methods.

Selection Method	Ensembl_gene_id	External gene id	Description
Keygenes	ENSG00000019549	SNAI2	snail family zinc finger 2 [Source:HGNC Symbol;Acc:HGNC:11094]
Keygenes	ENSG00000026025	VIM	vimentin [Source:HGNC Symbol;Acc:HGNC:12692]
Keygenes	ENSG00000039068	CDH1	cadherin 1, type 1, E-cadherin (epithelial) [Source:HGNC Symbol;Acc:HGNC:1748]
Keygenes	ENSG00000115414	FN1	fibronectin 1 [Source:HGNC Symbol;Acc:HGNC:3778]
Keygenes	ENSG00000122691	Twist1	twist family bHLH transcription factor 1 [Source:HGNC Symbol;Acc:HGNC:12428]
Keygenes	ENSG00000124216	SNAI1	snail family zinc finger 1 [Source:HGNC Symbol;Acc:HGNC:11128]
Keygenes	ENSG00000148516	ZEB1	zinc finger E-box binding homeobox 1 [Source:HGNC Symbol;Acc:HGNC:11642]
Keygenes	ENSG00000169554	ZEB2	zinc finger E-box binding homeobox 2 [Source:HGNC Symbol;Acc:HGNC:14881]
Keygenes	ENSG00000185669	SNAI3	snail family zinc finger 3 [Source:HGNC Symbol;Acc:HGNC:18411]
Keygenes	ENSG00000233608	Twist2	twist family bHLH transcription factor 2 [Source:HGNC Symbol;Acc:HGNC:20670]
AUC	ENSG00000148848	ADAM12	ADAM metalloproteinase domain 12 [Source:HGNC Symbol;Acc:HGNC:190]
AUC	ENSG00000140937	CDH11	cadherin 11, type 2, OB-cadherin (osteoblast) [Source:HGNC Symbol;Acc:HGNC:1750]
AUC	ENSG00000170558	CDH2	cadherin 2, type 1, N-cadherin (neuronal) [Source:HGNC Symbol;Acc:HGNC:1759]
AUC	ENSG00000108821	COL1A1	collagen, type I, alpha 1 [Source:HGNC Symbol;Acc:HGNC:2197]
AUC	ENSG00000168542	COL3A1	collagen, type III, alpha 1 [Source:HGNC Symbol;Acc:HGNC:2201]
AUC	ENSG00000130635	COL5A1	collagen, type V, alpha 1 [Source:HGNC Symbol;Acc:HGNC:2209]
AUC	ENSG00000142156	COL6A1	collagen, type VI, alpha 1 [Source:HGNC Symbol;Acc:HGNC:2211]
AUC	ENSG00000163359	COL6A3	collagen, type VI, alpha 3 [Source:HGNC Symbol;Acc:HGNC:2213]
AUC	ENSG00000118523	CTGF	connective tissue growth factor [Source:HGNC Symbol;Acc:HGNC:2500]
AUC	ENSG00000164741	DLC1	DLC1 Rho GTPase activating protein [Source:HGNC Symbol;Acc:HGNC:2897]
AUC	ENSG00000140092	FBLN5	fibulin 5 [Source:HGNC Symbol;Acc:HGNC:3602]
AUC	ENSG00000077782	FGFR1	fibroblast growth factor receptor 1 [Source:HGNC Symbol;Acc:HGNC:3688]
AUC	ENSG00000115414	FN1	fibronectin 1 [Source:HGNC Symbol;Acc:HGNC:3778]
AUC	ENSG00000101335	MYL9	myosin, light chain 9, regulatory [Source:HGNC Symbol;Acc:HGNC:15754]
AUC	ENSG00000087303	NID2	nidogen 2 (osteonidogen) [Source:HGNC Symbol;Acc:HGNC:13389]
AUC	ENSG00000122707	RECK	reversion-inducing-cysteine-rich protein with kazal motifs [Source:HGNC Symbol;Acc:HGNC:11345]
AUC	ENSG00000123610	TNFAIP6	tumor necrosis factor, alpha-induced protein 6 [Source:HGNC Symbol;Acc:HGNC:11898]
AUC	ENSG00000140416	TPM1	tropomyosin 1 (alpha) [Source:HGNC Symbol;Acc:HGNC:12010]

Supplementary Table 5. Continued

Selection Method	Ensembl_gene_id	External gene id	Description
AUC	ENSG00000038427	VCAN	versican [Source:HGNC Symbol;Acc:HGNC:2464]
AUC	ENSG000000142227	EMP3	epithelial membrane protein 3 [Source:HGNC Symbol;Acc:HGNC:3335]
AUC	ENSG000000166147	FBN1	fibrillin 1 [Source:HGNC Symbol;Acc:HGNC:3603]
AUC	ENSG000000109099	PMP22	peripheral myelin protein 22 [Source:HGNC Symbol;Acc:HGNC:9118]
AUC	ENSG000000125384	PTGER2	prostaglandin E receptor 2 (subtype EP2), 53kDa [Source:HGNC Symbol;Acc:HGNC:9594]
AUC	ENSG000000163661	PTX3	pentraxin 3, long [Source:HGNC Symbol;Acc:HGNC:9692]
AUC	ENSG000000149591	TAGLN	transgelin [Source:HGNC Symbol;Acc:HGNC:11553]
AUC	ENSG000000167552	TUBA1A	tubulin, alpha 1a [Source:HGNC Symbol;Acc:HGNC:20766]
AUC	ENSG000000026025	VIM	vimentin [Source:HGNC Symbol;Acc:HGNC:12692]
AUC	ENSG000000058091	CDK14	cyclin-dependent kinase 14 [Source:HGNC Symbol;Acc:HGNC:8883]
AUC	ENSG000000066629	EML1	echinoderm microtubule associated protein like 1 [Source:HGNC Symbol;Acc:HGNC:3330]
AUC	ENSG000000163430	FSTL1	folliculin-like 1 [Source:HGNC Symbol;Acc:HGNC:3972]
AUC	ENSG000000134986	NREP	neuronal regeneration related protein [Source:HGNC Symbol;Acc:HGNC:16834]
AUC	ENSG000000039068	CDH1	cadherin 1, type 1, E-cadherin (epithelial) [Source:HGNC Symbol;Acc:HGNC:1748]
AUC	ENSG000000163435	ELF3	E74-like factor 3 (ets domain transcription factor, epithelial-specific) [Source:HGNC Symbol;Acc:HGNC:3318]
AUC	ENSG000000146904	EPHA1	EPH receptor A1 [Source:HGNC Symbol;Acc:HGNC:3385]
AUC	ENSG000000173801	JUP	junction plakoglobin [Source:HGNC Symbol;Acc:HGNC:6207]
AUC	ENSG000000149573	MPZL2	myelin protein zero-like 2 [Source:HGNC Symbol;Acc:HGNC:3496]
AUC	ENSG000000099204	ABLIM1	actin binding LIM protein 1 [Source:HGNC Symbol;Acc:HGNC:78]
AUC	ENSG000000118898	PPL	periaplain [Source:HGNC Symbol;Acc:HGNC:9273]
AUC	ENSG000000135378	PRRG4	proline rich Gla (G-carboxyglutamic acid) 4 (transmembrane) [Source:HGNC Symbol;Acc:HGNC:30799]
AUC	ENSG000000167601	AXL	AXL receptor tyrosine kinase [Source:HGNC Symbol;Acc:HGNC:905]
AUC	ENSG000000151715	TMEM45B	transmembrane protein 45B [Source:HGNC Symbol;Acc:HGNC:25194]
AUC	ENSG000000049283	EPN3	epsin 3 [Source:HGNC Symbol;Acc:HGNC:18235]
AUC	ENSG000000167880	EVPL	envoplakin [Source:HGNC Symbol;Acc:HGNC:3503]
AUC	ENSG000000130545	CRB3	crumbs family member 3 [Source:HGNC Symbol;Acc:HGNC:20237]
AUC	ENSG000000147676	MAL2	mal, T-cell differentiation protein 2 (gene/pseudogene) [Source:HGNC Symbol;Acc:HGNC:13634]

GACCTGGGTGTAGATGATGGGATGTAGGACCATCATCGAACTTAAAGTTGAAAGCTAGGAGAGAGAGAGTGGAGCTTTGGGGAACTTTGAGCCGGCTAAAGGTATCTTTTTCATCATCCACCCGGTGCTGGGGTGAAGGAACTCCCTGAAAT

Supplementary Table 5. Continued

Selection Method	Ensembl_gene_id	External gene id	Description
AUC	ENSG00000134317	GRHL1	grainyhead-like 1 (Drosophila) [Source:HGNC Symbol;Acc:HGNC:17923]
AUC	ENSG00000182795	C1orf116	chromosome 1 open reading frame 116 [Source:HGNC Symbol;Acc:HGNC:28667]
AUC	ENSG00000184292	TACSTD2	tumor-associated calcium signal transducer 2 [Source:HGNC Symbol;Acc:HGNC:11530]
AUC	ENSG00000065361	ERBB3	v-erb-b2 avian erythroblastic leukemia viral oncogene homolog 3 [Source:HGNC Symbol;Acc:HGNC:3431]
AUC	ENSG00000132698	RAB25	RAB25, member RAS oncogene family [Source:HGNC Symbol;Acc:HGNC:18238]
AUC	ENSG00000104413	ESRP1	epithelial splicing regulatory protein 1 [Source:HGNC Symbol;Acc:HGNC:25966]
AUC	ENSG00000128656	CHN1	chimerin 1 [Source:HGNC Symbol;Acc:HGNC:1943]
AUC	ENSG00000160588	MPZL3	myelin protein zero-like 3 [Source:HGNC Symbol;Acc:HGNC:27279]
AUC	ENSG00000109062	SLC9A3R1	solute carrier family 9, subfamily A (NHE3, cation proton antiporter 3), member 3 regulator 1 [Source:HGNC Symbol;Acc:HGNC:11075]
AUC	ENSG00000137269	LRRC1	leucine rich repeat containing 1 [Source:HGNC Symbol;Acc:HGNC:14307]
AUC	ENSG00000120149	MSX2	msh homeobox 2 [Source:HGNC Symbol;Acc:HGNC:7392]
AUC	ENSG00000223572	CKMT1A	creatine kinase, mitochondrial 1A [Source:HGNC Symbol;Acc:HGNC:31736]
AUC	ENSG00000166130	IKBIP	IKKB interacting protein [Source:HGNC Symbol;Acc:HGNC:26430]
AUC	ENSG00000103067	ESRP2	epithelial splicing regulatory protein 2 [Source:HGNC Symbol;Acc:HGNC:26152]
AUC	ENSG00000151388	ADAMTS12	ADAM metalloproteinase with thrombospondin type 1 motif, 12 [Source:HGNC Symbol;Acc:HGNC:14605]
AUC	ENSG00000087116	ADAMTS2	ADAM metalloproteinase with thrombospondin type 1 motif, 2 [Source:HGNC Symbol;Acc:HGNC:218]
AUC	ENSG00000106624	AEBP1	AE binding protein 1 [Source:HGNC Symbol;Acc:HGNC:303]
AUC	ENSG00000173068	BNC2	basonuclin 2 [Source:HGNC Symbol;Acc:HGNC:30988]
AUC	ENSG00000122786	CALD1	caldesmon 1 [Source:HGNC Symbol;Acc:HGNC:1441]
AUC	ENSG00000140931	CVTMM3	CKLF-like MARVEL transmembrane domain containing 3 [Source:HGNC Symbol;Acc:HGNC:19174]
AUC	ENSG00000119865	CNRIP1	cannabinoid receptor interacting protein 1 [Source:HGNC Symbol;Acc:HGNC:24546]
AUC	ENSG00000123500	COL10A1	collagen, type X, alpha 1 [Source:HGNC Symbol;Acc:HGNC:2185]
AUC	ENSG00000164692	COL1A2	collagen, type I, alpha 2 [Source:HGNC Symbol;Acc:HGNC:2198]
AUC	ENSG00000204262	COL5A2	collagen, type V, alpha 2 [Source:HGNC Symbol;Acc:HGNC:2210]
AUC	ENSG00000142173	COL6A2	collagen, type VI, alpha 2 [Source:HGNC Symbol;Acc:HGNC:2212]
AUC	ENSG00000144810	COL8A1	collagen, type VIII, alpha 1 [Source:HGNC Symbol;Acc:HGNC:2215]

Supplementary Table 5. Continued

Selection Method	Ensembl_gene_id	External gene id	Description
AUC	ENSG00000165617	DACT1	dishevelled-binding antagonist of beta-catenin 1 [Source:HGNC Symbol;Acc:HGNC:17748]
AUC	ENSG00000078098	FAP	fibroblast activation protein, alpha [Source:HGNC Symbol;Acc:HGNC:3590]
AUC	ENSG00000183098	GPC6	glypican 6 [Source:HGNC Symbol;Acc:HGNC:4454]
AUC	ENSG00000122641	INHBA	inhibin, beta A [Source:HGNC Symbol;Acc:HGNC:6066]
AUC	ENSG00000137809	ITGA11	integrin, alpha 11 [Source:HGNC Symbol;Acc:HGNC:6136]
AUC	ENSG00000134013	LOXL2	lysyl oxidase-like 2 [Source:HGNC Symbol;Acc:HGNC:6666]
AUC	ENSG00000172061	LRRCL5	leucine rich repeat containing 15 [Source:HGNC Symbol;Acc:HGNC:20818]
AUC	ENSG00000174099	MSRB3	methionine sulfoxide reductase B3 [Source:HGNC Symbol;Acc:HGNC:27375]
AUC	ENSG00000186310	NAP1L3	nucleosome assembly protein 1-like 3 [Source:HGNC Symbol;Acc:HGNC:7639]
AUC	ENSG00000162745	OLFML2B	olfactomedin-like 2B [Source:HGNC Symbol;Acc:HGNC:24558]
AUC	ENSG00000106333	PCOLCE	procollagen C-endopeptidase enhancer [Source:HGNC Symbol;Acc:HGNC:8738]
AUC	ENSG00000113721	PDGFRB	platelet-derived growth factor receptor, beta polypeptide [Source:HGNC Symbol;Acc:HGNC:8804]
AUC	ENSG00000133110	POSTN	periostin, osteoblast specific factor [Source:HGNC Symbol;Acc:HGNC:16953]
AUC	ENSG00000113140	SPARC	secreted protein, acidic, cysteine-rich (osteonectin) [Source:HGNC Symbol;Acc:HGNC:11219]
AUC	ENSG00000137573	SULF1	sulfatase 1 [Source:HGNC Symbol;Acc:HGNC:20391]
AUC	ENSG00000169554	ZEB2	zinc finger E-box binding homeobox 2 [Source:HGNC Symbol;Acc:HGNC:14881]
AUC	ENSG00000122691	TWIST1	twist family bHLH transcription factor 1 [Source:HGNC Symbol;Acc:HGNC:12428]
AUC	ENSG00000143515	ATP8B2	ATPase, aminophospholipid transporter, class I, type 8B, member 2 [Source:HGNC Symbol;Acc:HGNC:13534]
AUC	ENSG00000166780	C16orf45	chromosome 16 open reading frame 45 [Source:HGNC Symbol;Acc:HGNC:19213]
AUC	ENSG00000162694	EXTL2	exostosin-like glycosyltransferase 2 [Source:HGNC Symbol;Acc:HGNC:3516]
AUC	ENSG00000073712	FERMT2	fermitin family member 2 [Source:HGNC Symbol;Acc:HGNC:15767]
AUC	ENSG00000112320	SOBP	sine oculis binding protein homolog (Drosophila) [Source:HGNC Symbol;Acc:HGNC:29256]
AUC	ENSG00000111450	STX2	syntaxin 2 [Source:HGNC Symbol;Acc:HGNC:3403]
AUC	ENSG00000105137	SYDE1	synapse defective 1, Rho GTPase, homolog 1 (C. elegans) [Source:HGNC Symbol;Acc:HGNC:25824]
AUC	ENSG00000117385	LEPRE1	leucine proline-enriched proteoglycan (leprecan) 1 [Source:HGNC Symbol;Acc:HGNC:19316]
AUC	ENSG00000110811	LEPREL2	leprecan-like 2 [Source:HGNC Symbol;Acc:HGNC:19318]
AUC	ENSG00000050327	ARHGEF5	Rho guanine nucleotide exchange factor (GEF) 5 [Source:HGNC Symbol;Acc:HGNC:13209]

GACCTGGGTGTAGATGATGGGATGTAGGACCATCGTAAAGTTGAACGGCTAGGAGAGAGGTGGAGTTTGGGGAACTTTGAGCCGGCTAAAGGTATCTTTTTCACATCCACCGCTGGGGGTAGAGGATCCCTGAAAT

Supplementary Table 5. Continued

Selection Method	Ensembl_gene_id	External gene id	Description
AUC	ENSG00000175866	BAIAP2	BAI1-associated protein 2 [Source:HGNC Symbol;Acc:HGNC:947]
AUC	ENSG00000110723	EXPH5	exophilin 5 [Source:HGNC Symbol;Acc:HGNC:30578]
AUC	ENSG00000141738	GRB7	growth factor receptor-bound protein 7 [Source:HGNC Symbol;Acc:HGNC:4567]
AUC	ENSG00000102554	KLF5	Kruppel-like factor 5 (intestinal) [Source:HGNC Symbol;Acc:HGNC:6349]
AUC	ENSG00000112378	PERP	PERP, TP53 apoptosis effector [Source:HGNC Symbol;Acc:HGNC:17637]
AUC	ENSG00000137648	TMPRSS4	transmembrane protease, serine 4 [Source:HGNC Symbol;Acc:HGNC:11878]
AUC	ENSG00000100097	LGALS1	lectin, galactoside-binding, soluble, 1 [Source:HGNC Symbol;Acc:HGNC:6561]
AUC	ENSG00000149380	P4HA3	prolyl 4-hydroxylase, alpha polypeptide III [Source:HGNC Symbol;Acc:HGNC:30135]
Boruta	ENSG00000148848	ADAM12	ADAM metalloproteinase domain 12 [Source:HGNC Symbol;Acc:HGNC:190]
Boruta	ENSG00000170558	CDH2	cadherin 2, type 1, N-cadherin (neuronal) [Source:HGNC Symbol;Acc:HGNC:1759]
Boruta	ENSG00000108821	COL1A1	collagen, type I, alpha 1 [Source:HGNC Symbol;Acc:HGNC:2197]
Boruta	ENSG00000168542	COL3A1	collagen, type III, alpha 1 [Source:HGNC Symbol;Acc:HGNC:2201]
Boruta	ENSG00000130635	COL5A1	collagen, type V, alpha 1 [Source:HGNC Symbol;Acc:HGNC:2209]
Boruta	ENSG00000142156	COL6A1	collagen, type VI, alpha 1 [Source:HGNC Symbol;Acc:HGNC:2211]
Boruta	ENSG00000163359	COL6A3	collagen, type VI, alpha 3 [Source:HGNC Symbol;Acc:HGNC:2213]
Boruta	ENSG00000118523	CTGF	connective tissue growth factor [Source:HGNC Symbol;Acc:HGNC:2500]
Boruta	ENSG00000164741	DLC1	DLC1 Rho GTPase activating protein [Source:HGNC Symbol;Acc:HGNC:2897]
Boruta	ENSG00000077942	FBLN1	fibulin 1 [Source:HGNC Symbol;Acc:HGNC:3600]
Boruta	ENSG00000140092	FBLN5	fibulin 5 [Source:HGNC Symbol;Acc:HGNC:3602]
Boruta	ENSG00000077782	FGFR1	fibroblast growth factor receptor 1 [Source:HGNC Symbol;Acc:HGNC:3688]
Boruta	ENSG00000115414	FN1	fibronectin 1 [Source:HGNC Symbol;Acc:HGNC:3778]
Boruta	ENSG00000087245	MMP2	matrix metalloproteinase 2 (gelatinase A, 72kDa gelatinase, 72kDa type IV collagenase) [Source:HGNC Symbol;Acc:HGNC:7166]
Boruta	ENSG00000101335	MYL9	myosin, light chain 9, regulatory [Source:HGNC Symbol;Acc:HGNC:15754]
Boruta	ENSG00000087303	NID2	nidogen 2 (osteonidogen) [Source:HGNC Symbol;Acc:HGNC:13389]
Boruta	ENSG00000123610	TNFAIP6	tumor necrosis factor, alpha-induced protein 6 [Source:HGNC Symbol;Acc:HGNC:11898]
Boruta	ENSG00000140416	TPM1	tropomyosin 1 (alpha) [Source:HGNC Symbol;Acc:HGNC:12010]
Boruta	ENSG00000038427	VCAN	versican [Source:HGNC Symbol;Acc:HGNC:2464]

Supplementary Table 5. Continued

Selection Method	Ensembl_gene_id	External gene id	Description
Boruta	ENSG00000166147	FBN1	fibrillin 1 [Source:HGNC Symbol;Acc:HGNC:3603]
Boruta	ENSG00000109099	PMP22	peripheral myelin protein 22 [Source:HGNC Symbol;Acc:HGNC:9118]
Boruta	ENSG00000125384	PTGER2	prostaglandin E receptor 2 (subtype EP2), 53kDa [Source:HGNC Symbol;Acc:HGNC:9594]
Boruta	ENSG00000149591	TAGLN	transgelin [Source:HGNC Symbol;Acc:HGNC:11553]
Boruta	ENSG00000167552	TUBA1A	tubulin, alpha 1a [Source:HGNC Symbol;Acc:HGNC:20766]
Boruta	ENSG00000026025	VIM	vimentin [Source:HGNC Symbol;Acc:HGNC:12692]
Boruta	ENSG00000058091	CDK14	cyclin-dependent kinase 14 [Source:HGNC Symbol;Acc:HGNC:8883]
Boruta	ENSG00000163430	FSTL1	folliculin-like 1 [Source:HGNC Symbol;Acc:HGNC:3972]
Boruta	ENSG00000134986	NREP	neuronal regeneration related protein [Source:HGNC Symbol;Acc:HGNC:16834]
Boruta	ENSG00000039068	CDH1	cadherin 1, type 1, E-cadherin (epithelial) [Source:HGNC Symbol;Acc:HGNC:1748]
Boruta	ENSG00000146904	EPHA1	EPH receptor A1 [Source:HGNC Symbol;Acc:HGNC:3385]
Boruta	ENSG00000149573	MPZL2	myelin protein zero-like 2 [Source:HGNC Symbol;Acc:HGNC:3496]
Boruta	ENSG00000099204	ABLIM1	actin binding LIM protein 1 [Source:HGNC Symbol;Acc:HGNC:78]
Boruta	ENSG00000118898	PPL	periplakin [Source:HGNC Symbol;Acc:HGNC:9273]
Boruta	ENSG00000167601	AXL	AXL receptor tyrosine kinase [Source:HGNC Symbol;Acc:HGNC:905]
Boruta	ENSG00000151715	TMEM45B	transmembrane protein 45B [Source:HGNC Symbol;Acc:HGNC:25194]
Boruta	ENSG00000049283	EPN3	epsin 3 [Source:HGNC Symbol;Acc:HGNC:18235]
Boruta	ENSG00000167880	EVPL	envoplakin [Source:HGNC Symbol;Acc:HGNC:3503]
Boruta	ENSG00000130545	CRB3	crumbs family member 3 [Source:HGNC Symbol;Acc:HGNC:20237]
Boruta	ENSG00000147676	MAL2	mal, T-cell differentiation protein 2 (gene/pseudogene) [Source:HGNC Symbol;Acc:HGNC:13634]
Boruta	ENSG00000134317	GRHL1	grainyhead-like 1 (Drosophila) [Source:HGNC Symbol;Acc:HGNC:17923]
Boruta	ENSG00000152766	ANKRD22	ankyrin repeat domain 22 [Source:HGNC Symbol;Acc:HGNC:28321]
Boruta	ENSG00000184292	TACSTD2	tumor-associated calcium signal transducer 2 [Source:HGNC Symbol;Acc:HGNC:11530]
Boruta	ENSG00000065361	ERBB3	v-erb-b2 avian erythroblastic leukemia viral oncogene homolog 3 [Source:HGNC Symbol;Acc:HGNC:3431]
Boruta	ENSG00000132698	RAB25	RAB25, member RAS oncogene family [Source:HGNC Symbol;Acc:HGNC:18238]
Boruta	ENSG00000104413	ESRP1	epithelial splicing regulatory protein 1 [Source:HGNC Symbol;Acc:HGNC:25966]
Boruta	ENSG00000173705	SUSD5	sushi domain containing 5 [Source:HGNC Symbol;Acc:HGNC:29061]
Boruta	ENSG00000128656	CHN1	chimerin 1 [Source:HGNC Symbol;Acc:HGNC:1943]

Supplementary Table 5. Continued

Selection Method	Ensembl_gene_id	External gene id	Description
Boruta	ENSG00000160588	MPZL3	myelin protein zero-like 3 [Source:HGNC Symbol;Acc:HGNC:27279]
Boruta	ENSG00000109062	SLC9A3R1	solute carrier family 9, subfamily A (NHE3, cation proton antiporter 3), member 3 regulator 1 [Source:HGNC Symbol;Acc:HGNC:11075]
Boruta	ENSG00000073350	LLGL2	lethal giant larvae homolog 2 (Drosophila) [Source:HGNC Symbol;Acc:HGNC:6629]
Boruta	ENSG00000166130	IKBIP	IKKB interacting protein [Source:HGNC Symbol;Acc:HGNC:26430]
Boruta	ENSG00000103067	ESRP2	epithelial splicing regulatory protein 2 [Source:HGNC Symbol;Acc:HGNC:26152]
Boruta	ENSG00000151388	ADAMTS12	ADAM metalloproteinase with thrombospondin type 1 motif, 12 [Source:HGNC Symbol;Acc:HGNC:14605]
Boruta	ENSG00000087116	ADAMTS2	ADAM metalloproteinase with thrombospondin type 1 motif, 2 [Source:HGNC Symbol;Acc:HGNC:218]
Boruta	ENSG00000106624	AEBP1	AE binding protein 1 [Source:HGNC Symbol;Acc:HGNC:303]
Boruta	ENSG00000173068	BNC2	basonuclin 2 [Source:HGNC Symbol;Acc:HGNC:30988]
Boruta	ENSG00000122786	CALD1	caldesmon 1 [Source:HGNC Symbol;Acc:HGNC:1441]
Boruta	ENSG00000140931	CVTM3	CKLF-like MARVEL transmembrane domain containing 3 [Source:HGNC Symbol;Acc:HGNC:19174]
Boruta	ENSG00000119865	CNRIP1	cannabinoid receptor interacting protein 1 [Source:HGNC Symbol;Acc:HGNC:24546]
Boruta	ENSG00000164692	COL1A2	collagen, type I, alpha 2 [Source:HGNC Symbol;Acc:HGNC:2198]
Boruta	ENSG00000204262	COL5A2	collagen, type V, alpha 2 [Source:HGNC Symbol;Acc:HGNC:2210]
Boruta	ENSG00000142173	COL6A2	collagen, type VI, alpha 2 [Source:HGNC Symbol;Acc:HGNC:2212]
Boruta	ENSG00000144810	COL8A1	collagen, type VIII, alpha 1 [Source:HGNC Symbol;Acc:HGNC:2215]
Boruta	ENSG00000165617	DACT1	dishevelled-binding antagonist of beta-catenin 1 [Source:HGNC Symbol;Acc:HGNC:17748]
Boruta	ENSG00000078098	FAP	fibroblast activation protein, alpha [Source:HGNC Symbol;Acc:HGNC:3590]
Boruta	ENSG00000183098	GPC6	glypican 6 [Source:HGNC Symbol;Acc:HGNC:4454]
Boruta	ENSG00000122641	INHBA	inhibin, beta A [Source:HGNC Symbol;Acc:HGNC:6066]
Boruta	ENSG00000137809	ITGA11	integrin, alpha 11 [Source:HGNC Symbol;Acc:HGNC:6136]
Boruta	ENSG00000134013	LOXL2	lysyl oxidase-like 2 [Source:HGNC Symbol;Acc:HGNC:6666]
Boruta	ENSG00000172061	LRRC15	leucine rich repeat containing 15 [Source:HGNC Symbol;Acc:HGNC:20818]
Boruta	ENSG00000174099	MSRB3	methionine sulfoxide reductase B3 [Source:HGNC Symbol;Acc:HGNC:27375]
Boruta	ENSG00000162745	OLFML2B	olfactomedin-like 2B [Source:HGNC Symbol;Acc:HGNC:24558]
Boruta	ENSG00000106333	PCOLCE	procollagen C-endopeptidase enhancer [Source:HGNC Symbol;Acc:HGNC:8738]
Boruta	ENSG00000113721	PDGFRB	platelet-derived growth factor receptor, beta polypeptide [Source:HGNC Symbol;Acc:HGNC:8804]

Supplementary Table 5. Continued

Selection Method	Ensembl_gene_id	External gene id	Description
Boruta	ENSG00000133110	POSTN	periostin, osteoblast specific factor [Source:HGNC Symbol;Acc:HGNC:16953]
Boruta	ENSG00000113140	SPARC	secreted protein, acidic, cysteine-rich (osteonectin) [Source:HGNC Symbol;Acc:HGNC:11219]
Boruta	ENSG00000137573	SULF1	sulfatase 1 [Source:HGNC Symbol;Acc:HGNC:20391]
Boruta	ENSG00000122691	TWIST1	twist family bHLH transcription factor 1 [Source:HGNC Symbol;Acc:HGNC:12428]
Boruta	ENSG00000143515	ATP8B2	ATPase, aminophospholipid transporter, class I, type 8B, member 2 [Source:HGNC Symbol;Acc:HGNC:13534]
Boruta	ENSG00000073712	FERMT2	fermitin family member 2 [Source:HGNC Symbol;Acc:HGNC:15767]
Boruta	ENSG00000112320	SOBP	sine oculis binding protein homolog (Drosophila) [Source:HGNC Symbol;Acc:HGNC:29256]
Boruta	ENSG00000111450	STX2	syntaxin 2 [Source:HGNC Symbol;Acc:HGNC:3403]
Boruta	ENSG00000105137	SYDE1	synapse defective 1, Rho GTPase, homolog 1 (C. elegans) [Source:HGNC Symbol;Acc:HGNC:25824]
Boruta	ENSG00000117385	LEPRE1	leucine proline-enriched proteoglycan (leprecan) 1 [Source:HGNC Symbol;Acc:HGNC:19316]
Boruta	ENSG00000110811	LEPRE2	leprecan-like 2 [Source:HGNC Symbol;Acc:HGNC:19318]
Boruta	ENSG00000050327	ARHGEF5	Rho guanine nucleotide exchange factor (GEF) 5 [Source:HGNC Symbol;Acc:HGNC:13209]
Boruta	ENSG00000175866	BAIAP2	BAI1-associated protein 2 [Source:HGNC Symbol;Acc:HGNC:947]
Boruta	ENSG00000086544	ITPKC	inositol-trisphosphate 3-kinase C [Source:HGNC Symbol;Acc:HGNC:14897]
Boruta	ENSG00000102554	KLF5	Kruppel-like factor 5 (intestinal) [Source:HGNC Symbol;Acc:HGNC:6349]
Boruta	ENSG00000137648	TMPRSS4	transmembrane protease, serine 4 [Source:HGNC Symbol;Acc:HGNC:11878]
Boruta	ENSG00000100097	LGALS1	lectin, galactoside-binding, soluble, 1 [Source:HGNC Symbol;Acc:HGNC:6561]
Boruta	ENSG00000149380	P4HA3	prolyl 4-hydroxylase, alpha polypeptide III [Source:HGNC Symbol;Acc:HGNC:30135]

Supplementary Table 6. Performance of each individual model in a 50x cross validation

	AUC	Accuracy	Specificity	Sensitivity
Keygenes	0.99	0.99	0.98	1
Boruta	1	1	1	1
AUC	1	1	1	1

Supplementary Table 7. Univariate cox proportional hazard analysis of clinical parameters with patient outcome in the DESIGN cohort.

OS: Overall Survival; PFS: Progression Free Survival; DM: Distant Metastasis Free Survival; LRC: Locoregional Control; HR: Hazard Ratio; REF: Reference.

		OS		PFS		DM		LRC	
Variable		HR	<i>p</i> -value	HR	<i>p</i> -value	HR	<i>p</i> -value	HR	<i>p</i> -value
<i>Gender</i>									
	Female	REF		REF		REF		REF	
	Male	1.54	0.28	1.49	0.29	>100	0.99	2.17	0.22
<i>Age at diagnosis</i>		1.03	0.20	0.99	0.97	1.05	0.21	0.98	0.46
<i>Tumor site</i>									
	Hypopharynx	REF		REF		REF		REF	
	Oropharynx	1.97	0.09	2.27	0.04	1.32	0.70	1.43	0.53
	Larynx	1.26	0.64	1.42	0.46	2.09	0.31	1.54	0.48
<i>Disease stage</i>									
	II	-	-	-	-	-	-	-	-
	III	0.73	0.48	0.67	0.36	<0.01	0.998	0.59	0.42
	IVA	REF		REF		REF		REF	
	IVB	2.33	0.04	1.77	0.15	1.95	0.27	1.43	0.54
<i>cT classification</i>									
	T1-T3	REF		REF		REF		REF	
	T4	1.45	0.27	1.46	0.24	1.99	0.21	1.51	0.37
<i>cN classification</i>									
	N0-N2a	REF		REF		REF		REF	
	N2b-N3	1.15	0.68	1.12	0.73	7.7	0.05	0.99	0.99
<i>Smoker</i>									
	Former smoker	0.80	0.71	0.70	0.55	0.71	0.75	0.48	0.48
	Never	0.46	0.42	0.38	0.35	<0.01	1.0	<0.01	0.99
	Unknown	<0.01	0.98	<0.01	0.98	<0.01	1.0	<0.01	0.99
	Yes	REF		REF		REF		REF	
<i>Alcohol consumption</i>									
	Former alcoholic	1.69	0.39	1.56	0.46	2.88	0.72	2.20	0.30
	Never	0.66	0.49	0.85	0.76	<0.01	0.99	1.44	0.57
	Unknown	1.52	0.68	1.38	0.75	4.40	0.16	1.16	0.99
	Yes	REF		REF		REF		REF	
Cumulative cisplatin > 200 mg/m ²									
	No	REF		REF		REF		REF	
	Yes	0.47	0.04	0.59	0.11	1.19	0.76	0.49	0.15

Supplementary Table 8. Cox proportional hazard models results for the HNSCC-EMT model and published non-HNSCC specific profiles in the DESIGN cohort (Univariable and Multivariable)

Multivariable analysis A includes T-stage and N-stage

Multivariable analysis B includes cumulative cisplatin dose and tumor site

* = number of samples classified as epithelial (EPI) or mesenchymal (MES).

Univariable	Overall Survival			Progression Free Survival			Distant Metastasis Free Survival			Locoregional control								
	EPI*	MES*	p value	HR	Lower	Upper	p value	HR	Lower	Upper	p value	HR	Lower	Upper				
HNSCC-EMT Model	65	11	0.003	3.2	1.502	6.819	0.006	2.89	1.358	6.154	0.528	1.627	0.359	7.383	0.109	2.486	0.816	7.577
	75	1	0.147	4.48	0.592	33.927	0.021	11.99	1.444	99.62	0.999	0	0	Inf	0.006	23.496	2.444	225.887
	62	14	0.04	2.142	1.035	4.435	0.011	2.48	1.232	4.992	0.465	0.467	0.061	3.597	0.147	2.135	0.767	5.944
	52	24	0.249	1.478	0.76	2.874	0.42	1.307	0.682	2.505	0.917	1.064	0.327	3.461	0.824	1.116	0.424	2.937
	64	12	0.111	1.906	0.863	4.211	0.109	1.895	0.867	4.14	0.729	1.305	0.288	5.909	0.05	2.806	1	7.873
	38	38	0.007	2.561	1.295	5.064	0.01	2.349	1.229	4.486	0.906	1.068	0.357	3.2	0.042	2.731	1.035	7.201
	38	38	0.115	1.707	0.878	3.318	0.137	1.617	0.858	3.047	0.71	1.23	0.413	3.661	0.205	1.828	0.719	4.643
	38	38	0.116	1.707	0.877	3.321	0.117	1.66	0.88	3.131	0.652	1.285	0.432	3.827	0.182	1.887	0.743	4.795
Multivariable A	Overall Survival			Progression Free Survival			Distant Metastasis Free Survival			Locoregional control								
	EPI*	MES*	p value	HR	Lower	Upper	p value	HR	Lower	Upper	p value	HR	Lower	Upper				
HNSCC-EMT	65	11	0.002	3.392	1.583	7.268	0.004	3.011	1.409	6.434	0.417	1.875	0.41	8.571	0.088	2.654	0.865	8.145
	75	1	0.102	5.713	0.707	46.141	0.015	14.98	1.703	131.7	0.999	0	0	Inf	0.005	29.334	2.731	315.118
	62	14	0.032	2.241	1.07	4.696	0.009	2.531	1.256	5.099	0.599	0.577	0.075	4.468	0.13	2.21	0.791	6.177
	52	24	0.18	1.585	0.808	3.112	0.349	1.368	0.711	2.633	0.681	1.281	0.393	4.177	0.741	1.179	0.444	3.129
	64	12	0.046	2.409	1.015	5.717	0.069	2.122	0.943	4.777	0.296	2.299	0.483	10.95	0.038	3.084	1.065	8.928
	38	38	0.007	2.605	1.305	5.2	0.011	2.343	1.217	4.51	0.818	1.142	0.37	3.521	0.051	2.646	0.997	7.019
	38	38	0.114	1.717	0.879	3.352	0.135	1.633	0.858	3.11	0.885	1.084	0.363	3.234	0.192	1.879	0.728	4.851
	38	38	0.088	1.794	0.917	3.51	0.089	1.746	0.919	3.318	0.616	1.323	0.443	3.954	0.146	2.016	0.784	5.179

Supplementary Table 8. Continued

Multivariable B		Overall Survival			Progression Free Survival			Distant Metastasis Free Survival			Locoregional control							
EPI*	MES*	p value	HR	Lower	Upper	p value	HR	Lower	Upper	p value	HR	Lower	Upper					
HNSCC-EMT Model	65	11	0.015	2.849	1.226	6.623	0.051	2.288	0.995	5.264	0.441	1.865	0.383	9.095	0.342	1.883	0.51	6.951
Byers	75	1	0.089	6.887	0.745	63.704	0.008	22.75	2.241	231	0.999	0	0	Inf	0.006	44.01	2.977	650.638
Rokavec	62	14	0.142	2.013	0.79	5.127	0.13	1.981	0.818	4.793	0.509	0.477	0.053	4.294	0.205	2.439	0.615	9.664
Mak	52	24	0.411	1.381	0.64	2.98	0.918	1.04	0.491	2.203	0.936	1.054	0.293	3.799	0.939	1.046	0.337	3.248
Choi	64	12	0.028	2.595	1.11	6.066	0.037	2.45	1.053	5.699	0.733	1.306	0.281	6.085	0.014	4.137	1.338	12.789
Gibbons	38	38	0.005	3.248	1.417	7.447	0.018	2.468	1.167	5.218	0.743	1.22	0.373	3.99	0.018	4.102	1.279	13.154
Mlow	38	38	0.15	1.734	0.82	3.667	0.201	1.587	0.782	3.22	0.43	1.604	0.496	5.182	0.226	1.902	0.671	5.388
Groger	38	38	0.206	1.625	0.766	3.449	0.272	1.49	0.732	3.035	0.411	1.633	0.507	5.259	0.249	1.843	0.652	5.211

Supplementary material and methods

Patients NKI-CRAD cohort

The NKI-CRAD cohort is a retrospective cohort that comprises 98 HNSCC patients with hypopharyngeal, laryngeal or HPV-negative oropharyngeal HNSCC. All patients were treated with cisplatin-based chemoradiotherapy at the Netherlands Cancer Institute, Amsterdam, between 2001 to 2014. Previous treatment with radiotherapy in the head and neck area, chemotherapy or the lack of frozen pretreatment biopsy material were exclusion criteria. Three different cisplatin regimens were included. Patients received cisplatin daily (25x6 mg/m²) or 3-weekly (3x100 mg/m²). Eight patients received intra-arterial cisplatin (4x150 mg/m²) according to the RADPLAT trial¹. All patients received and completed radiotherapy, either conventional 35x2Gy over 7 weeks or following the DAHANCA scheme (35x2Gy fractions over 6 weeks). Not all patients completed chemotherapy and the individual total cumulative cisplatin dose was recorded. Patients were categorized into high or low cumulative cisplatin dose (cutoff 200 mg/m²) according to Al-Mamgani *et al.*² Patient characteristics are provided in Supplementary Table S1.

Patients DESIGN cohort

The validation cohort comprises patients included in the Dutch multicenter DESIGN study (KWF-A6C7072). The DESIGN study was initiated to explore and validate biomarkers for patients with advanced stage HPV-negative HNSCC and treated with chemoradiotherapy. Patient selection criteria for the DESIGN study followed inclusion and exclusion criteria identical to the first NKI-CRAD cohort. All patients were treated at the VUmc Cancer Center Amsterdam or in The Netherlands Cancer Institute, both situated in Amsterdam, between 2009-2014. Minimal 40% tumor content values and good quality RNA sequencing data were achieved for 76 patient samples. Patient characteristics are provided in Supplementary Table S1.

Patients NKI-MET cohort

This cohort of patients with metastatic disease were diagnosed and treated at the Netherlands Cancer Institute (NKI) between 2012 and 2017. These patients were initially analyzed for an advanced stage HPV-negative head and neck cancer, however, diagnostic PET-CT scan analyses showed distant metastasis prior to treatment. As a result these patients were not treated with regimens standard to those with curative intent. Pretreatment fresh-fresh frozen biopsy material of the primary tumor was available for 7 patients and processed and sequenced as described above. There were three oropharyngeal tumors patients, laryngeal and two hypopharyngeal tumors.

Patient material

Biopsies and collection of fresh-frozen tumor material were approved by the Institutional Review Board of the Netherlands Cancer Institute. All patients signed informed consent for biopsy, HNSCC material collection and analysis. Sample tumor percentage was determined by a dedicated head and neck pathologist on H&E-stained slides. Samples with a tumor content of less than 40% were excluded from RNA sequencing and further analysis.

RNA-sequencing

RNA was isolated using the DNA AllPrep DNA/RNA mini kit (Qiagen). Samples with RIN>7, as determined in the 2100 Bioanalyzer with a Nano chip (Agilent, Santa Clara, CA), were subjected to library generation. Strand-specific libraries were prepared using the TruSeq Stranded mRNA preparation kit (Illumina Inc., San Diego, RS-122-2101/2) according to the manufacturer's instructions (Illumina, Part # 15031047 Rev. E) and analyzed on a 2100 Bioanalyzer using a 7500 chip (Agilent, Santa Clara, CA). Libraries were sequenced generating 65 base single end reads on a HiSeq2500 using V4 chemistry (Illumina Inc., San Diego). Reads were mapped against the GRCh38 human genome with TopHat2.1³, using the options 'fr-firststrand', 'transcriptome-index' and 'prefilter multi-hits'. Read counts were determined using HTSeq-count⁴ with the options 'stranded' and the mode 'union'. Raw read counts were used to calculate the reads per kilobase million (RPKM).

HPV status determination

HPV-status of all oropharyngeal carcinomas was determined by immunohistochemical assessment of p16 by a dedicated head and neck pathologist⁵. RNA sequence alignment against the HPV subtypes 16, 18, 33 and 35 was performed and confirmed the HPV status of p16 positive samples for exclusion.

Expression analysis

All expression analyses were performed in R 3.4.3 using Rstudio 1.1. Seven different gene expression profiles were applied for epithelial or mesenchymal classifications of the tumor samples. We applied these profiles according to the author's protocols (specified and listed in Supplementary Table S2)⁶⁻¹². In brief, for profiles from Byers *et al.* and Choi *et al.*, a hierarchical clustering was performed based on the genes in their profile. The clustering tree was cut into two groups. TWIST1 expression levels served to identify mesenchymal and epithelial groups. For profiles by Mak *et al.* and Gibbons *et al.*, the data was centered and scaled. In case of the Mak profile, the mean expression of the epithelial genes was subtracted from the mean

expression of the mesenchymal genes. Samples with values above 0, the cut-off provided by Mak *et al.*, were classified as mesenchymal, below epithelial. In case of the Gibbons profile, the sum of the expression of epithelial genes was subtracted from the sum of the expression of the mesenchymal genes. A cutoff was not provided by Gibbons *et al.* and therefore the median was used. For Rokavec *et al.*, k-means clustering was performed thereby obtaining two groups. Groger *et al.* and Miow *et al.* did not provide a method on how to apply their profile. For each sample the score of the EMT-profiles was calculated using GSVA¹³, a Bioconductor package for R, on raw read counts. For Miow *et al.*, a principal component analysis was performed and the first principal component was used applying a median cutoff. The median of the GSVA score was used in the Groger profile to define epithelial and mesenchymal groups.

Gene selection and model generation

To simplify modelling we applied different gene selection methods. Boruta¹⁴ is a random forest based feature selection method. It validates the importance of each gene by comparing it to random shuffled copies of the dataset. To increase robustness, this process is repeated multiple times. The “AUC” method calculates a ROC curve using the R-package pROC¹⁵ for each gene, a measurement of how well the gene separates groups. With a pre-defined cutoff of 0.95, we selected all genes who classified 95% of the samples correctly. For each gene selection method a model was trained using random forest. Random forest is a machine learning algorithm which uses multiple random decision trees that result in an accurate prediction¹⁶. The values as determined by the individual models were averaged and used when combining models.

Tissue Microarray and Immunohistochemistry

TMA's were made automated using the TMA Grandmaster (Sysmex/3D Histech). The area where to punch out a core from the tumourblock was marked by a pathologist on a representative H&E stained slide. Per tumor block 3 cores with a diameter of 0.6mm each were taken out and put in the TMA. Immunohistochemistry of the TMA was performed on a BenchMark Ultra autostainer (Ventana Medical Systems). Briefly, paraffin sections were cut at 3 μ m, heated at 75°C for 28 minutes and deparaffinized in the instrument with EZ prep solution (Ventana Medical Systems). Heat-induced antigen retrieval was carried out using Cell Conditioning 1 (CC1, Ventana Medical Systems) for 32 minutes at 950C. Vimentin was detected using clone Vim3B4 (1/5000, 32 minutes at 370C, Agilent / DAKO). Signal amplification was applied using the Optiview Amplification Kit (4 minutes, Ventana Medical Systems). Bound antibody was detected using the OptiView DAB Detection Kit (Ventana Medical Systems). Slides were counterstained with Hematoxylin and Bluing Reagent (Ventana Medical Systems).

References

1. Balm AJM, Rasch CRN, Schornagel JH, *et al.* High-dose superselective intra-arterial cisplatin and concomitant radiation (radplat) for advanced head and neck cancer. *Head Neck*. 2004;26(6):485-493. doi:10.1002/hed.20006
2. 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 Oto-Rhino-Laryngology*. 2017. doi:10.1007/s00405-017-4687-4
3. Trapnell C, Pachter L, Salzberg SL. TopHat: Discovering splice junctions with RNA-Seq. *Bioinformatics*. 2009;25(9):1105-1111. doi:10.1093/bioinformatics/btp120
4. Anders S, Pyl PT, Huber W. HTSeq-A Python framework to work with high-throughput sequencing data. *Bioinformatics*. 2015;31(2):166-169. doi:10.1093/bioinformatics/btu638
5. Lewis JS, Chernock RD, Ma XJ, *et al.* Partial p16 staining in oropharyngeal squamous cell carcinoma: Extent and pattern correlate with human papillomavirus RNA status. *Mod Pathol*. 2012. doi:10.1038/modpathol.2012.79
6. Byers LA, Diao L, Wang J, *et al.* An epithelial-mesenchymal transition gene signature predicts resistance to EGFR and PI3K inhibitors and identifies Axl as a therapeutic target for overcoming EGFR inhibitor resistance. *Clin Cancer Res*. 2013;19(1):279-290. doi:10.1158/1078-0432.CCR-12-1558
7. Mak MP, Tong P, Diao L, *et al.* A Patient-Derived, Pan-Cancer EMT Signature Identifies Global Molecular Alterations and Immune Target Enrichment Following Epithelial-to-Mesenchymal Transition. *Clin Cancer Res*. 2016;22(3):609-620. doi:10.1158/1078-0432.CCR-15-0876
8. Gibbons DL, Creighton CJ. Pan-cancer survey of epithelial-mesenchymal transition markers across the Cancer Genome Atlas. *Dev Dyn*. 2018;247(3):555-564. doi:10.1002/dvdy.24485
9. Gröger CJ, Grubinger M, Waldhör T, Vierlinger K, Mikulits W. Meta-Analysis of Gene Expression Signatures Defining the Epithelial to Mesenchymal Transition during Cancer Progression. *PLoS One*. 2012;7(12):1-10. doi:10.1371/journal.pone.0051136
10. Rokavec M, Kaller M, Horst D, Hermeking H. Pan-cancer EMT-signature identifies RBM47 down-regulation during colorectal cancer progression. *Sci Rep*. 2017;7(1):1-15. doi:10.1038/s41598-017-04234-2
11. Miow QH, Tan TZ, Ye J, *et al.* Epithelial-mesenchymal status renders differential responses to cisplatin in ovarian cancer. *Oncogene*. 2015;34(15):1899-1907. doi:10.1038/onc.2014.136
12. Choi Y La, Bocanegra M, Kwon MJ, *et al.* LYN is a mediator of epithelial-mesenchymal transition and a target of dasatinib in breast cancer. *Cancer Res*. 2010;70(6):2296-2306. doi:10.1158/0008-5472.CAN-09-3141
13. Hänzelmann S, Castelo R, Guinney J. GSEA: Gene set variation analysis for microarray and RNA-Seq data. *BMC Bioinformatics*. 2013;14. doi:10.1186/1471-2105-14-7
14. Kursu MB, Rudnicki WR. Feature Selection with the Boruta Package. *J Stat Softw*. 2010;36(11):1-13. doi:Vol. 36, Issue 11, Sep 2010
15. 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. doi:10.1186/1471-2105-12-77
16. Breiman L. Random Forrest. *Mach Learn*. 2001;5-32. doi:10.1023/A:1010933404324



CHAPTER 4

TTCCCATCAACCCCTAGG99CTCCTCCTT99CTGCTGGGAGTTGTAGTCTGAACTTCTATCTT99CAGAGAGCCCTAC9CTCCCCCTACCGAGTCCCGCGGTAAATTCCTTAAAGCAGCTGCTACCGCCCCCCCCCCCCCGCTGCTG

Drug sensitivity prediction models reveal a link between DNA repair defects and poor prognosis in HNSCC

Paul B.M. Essers

Martijn van der Heijden[‡]

Caroline V.M. Verhagen[‡]

Emily M. Ploeg, Reinout H. de Roest

C. René Leemans, Ruud H. Brakenhoff

Michiel W.M. van den Brekel

Harry Bartelink

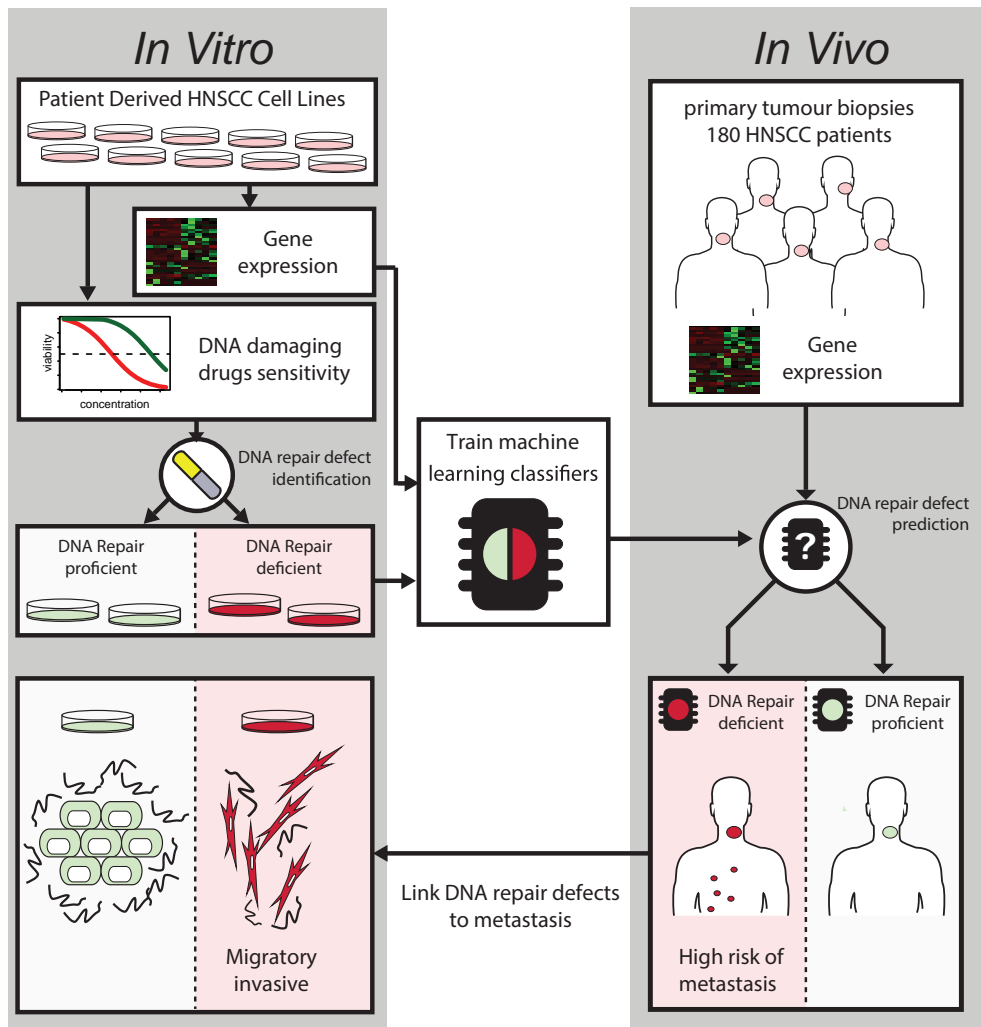
Marcel Verheij and Conchita Vens

[‡] These authors contributed equally to this work

Cancer Research, November 2019, 79 (21): 5597–5611.

<https://doi.org/10.1158/0008-5472.CAN-18-3388>

Graphical abstract



Machine learning classifiers developed to identify DNA repair defects in cells reveal associations with metastasis in patients and link repair defects to migration and invasiveness.

Abstract

Head and neck squamous cell carcinoma (HNSCC) is characterized by the frequent manifestation of DNA crosslink repair defects. We established novel expression-based DNA repair defect markers to determine the clinical impact of such repair defects.

Using hypersensitivity to the DNA crosslinking agents mitomycin C and olaparib as proxies for functional DNA repair defects in a panel of 25 HNSCC cell lines, we applied machine learning to define gene expression models that predict repair defects. The expression profiles established predicted hypersensitivity to DNA damaging agents and were associated with mutations in crosslink repair genes, as well as downregulation of DNA damage response and repair genes, in two independent datasets.

The prognostic value of the repair defect prediction profiles was assessed in two retrospective cohorts with a total of 180 patients advanced HPV-negative HNSCC, who were treated with cisplatin-based chemo-radiotherapy. DNA repair defects, as predicted by the profiles, were associated with poor outcome in both patient cohorts. The poor prognosis association was particularly strong in normoxic tumor samples and was linked to an increased risk of distant metastasis. *In vitro*, only crosslink repair defective HNSCC cell lines are highly migratory and invasive. This phenotype could also be induced in cells by inhibiting rad51 in repair competent and reduced by DNA-PK inhibition.

In conclusion, DNA crosslink repair prediction expression profiles reveal a poor prognosis association in HNSCC.

Significance: This study uses innovative machine learning-based approaches to derive models that predict the effect of DNA repair defects on treatment outcome in HNSCC.

Introduction

DNA repair pathways, such as homologous recombination (HR), Fanconi Anemia (FA) and associated cellular repair pathways, govern genomic integrity and maintain chromosomal stability. Recent genetic studies highlight the identification of an increasing number of DNA damage response (DDR) and DNA repair gene defects in cancer¹, but the role of these defects in cancer outcome or treatment response is not yet fully understood. Early *in vitro* studies showed how defects in DDR and DNA repair pathways cause increased sensitivities to chemotherapeutic agents or radiation. Thus, given the DNA damaging nature of many cancer treatment agents, tumor repair defects are expected to be beneficial for overall patient outcome. However, evidence, mostly assessed in breast cancer, is inconclusive². Studies in other cancer types are less numerous and are greatly hampered by the lack of biomarkers that identify functional repair defects. Here we set out to develop DNA repair defect markers to determine the clinical impact of DNA repair defects in head and neck squamous cell carcinoma (HNSCC).

DNA repair gene variants may be used to depict DNA repair defects, such as in the case of pathogenic BRCA1/2 variants in breast or ovarian cancer^{1,3}. Although mutations in such genes have been shown to predict platinum⁴ and olaparib⁵ response, determining which mutations affect the function of a given protein or pathway remains difficult⁶. Additionally, despite the high number of variants and mutations, the frequency of specific variants is very low. Other attempts successfully applied indirect repair defect measures such as “genomic scars”⁷. The genomic scar biomarkers were developed in BRCA1/2-mutant breast cancer and show promise for patient stratification for olaparib treatment in clinical trials in breast cancer⁸. Such scar-based biomarkers depict past DNA repair defect events. Multiple resistance mechanisms have recently been described and some show evidence of DNA repair capacity restoration^{9,10}.

The quest to assess the clinical impact of DDR and repair defects in other cancer types faces many challenges. Numerous DDR genes have been shown to be mutated in human cancer¹¹. However, the functional significance of most of these mutations is largely unknown. A multitude of genes are affected and the mutated genes or pathways vary largely among different cancer types. The oncogenetic context in these cancers also impacts DNA repair pathway engagement or performance. Detection of DNA repair defects therefore requires broadening the scope beyond breast- and ovarian cancer-specific BRCA mutations. A gene expression profile for mutations that phenocopy p53 inactivation was developed as a proxy for a deficiency in DNA damage response, and showed an association with poor prognosis in some cancers¹². However, an exclusive focus on TP53 mutations limits the prediction to

altered p53 response rather than DNA repair defect prediction. Functional analyses, such as the inability of cells to recover from DNA crosslinker exposure, reliably reveal DNA crosslink repair defects, independent of the underlying genetic cause. Similar to genomic scars, functional repair endpoint studies aim to improve the association with true yet undetected DNA repair defects^{13–15}. Previous studies by us and others exposed functional DNA repair defects in the HR/FA pathway in HNSCC^{16–18}. These defects were identified by determining sensitivity to the two DNA-damaging agents: mitomycin C (MMC) or olaparib in patient-derived HNSCC cell lines¹⁶. MMC is a potent DNA crosslinking agent. Cellular survival after exposure to DNA crosslinking agents requires the concerted action of multiple repair pathways, including nucleotide excision repair, FA and HR¹⁹. MMC hypersensitivity is therefore a hallmark of FA pathway defects and has been used in the identification of many members of the FA and HR pathway, such as the BRCA1/2 genes²⁰. The PARP inhibitor olaparib interferes in PARP-mediated DNA repair, primarily singlestrand break repair and inhibitor-induced trapped PARP poses replication blocks²¹. Together with the increased load of unresolved DNA damage, such blocks also contribute to the induction of secondary DNA double-strand breaks²². Cells with HR/FA defects and with high levels of replication stress struggle to survive PARP inhibition. As with MMC, hypersensitivity to PARP inhibition can therefore reveal functional HR and FA pathway defects, also of unknown genetic causes. Notably, a shift in the engagement from HR toward other DSB repair pathways can circumvent defects that cause PARP inhibitor sensitivity, while maintaining crosslinker sensitivity^{9,20,23}. Functional pathway defects were accompanied by genetic variants in multiple FA and HR genes in some of the HNSCC cell lines, but not all, in our *in vitro* study, further arguing for functional endpoint-based assays¹⁶. HNSCC are in particular interesting in that they have been shown to exhibit functional DNA repair defects that are caused by a multitude of genetic aberrations in different DNA repair pathways. FA and HR gene variants were identified in larger HNSCC sequencing efforts¹⁸ and found to be common in young patients with HNSCC²⁴, but the relation to treatment outcome is less well explored. A small study compared HR pathway engagement in biopsies of 13 patients with HNSCC before and after the start of chemoradiotherapy by analyzing the ability to form Rad51 foci; this points to a poorer patient outcome of carriers of DNA repair-affected tumors¹⁴. In line with such an influence, a recent study highlights the association between chromosomal instability and lymph node status in HNSCC²⁵. Larger studies investigating the effect of functional DNA repair defects on patient outcome in HNSCC are lacking. In this study, we aimed to determine the impact of such DNA repair defects on outcome and treatment response in HNSCC. HR and FA repair defects provoke replication stress in proliferating cells and a network of various compensatory DNA repair options are engaged to assure cellular survival. We

therefore posit that DNA repair-defective cells will show a distinct transcriptional pattern that indirectly reflects the DNA damage response to unrepaired damage from endogenous sources. To embrace the variety of repair defect types and genetic contexts, we employed a large panel of patient-derived HNSCC cell lines, of which, a subset showed functional repair defects as determined by hypersensitivity to MMC and/or the PARP inhibitor olaparib and used machine learning models to predict a biological characteristic, DNA crosslink repair status. Independent *in vitro* datasets were used to test the performance in identifying DNA repair defects. The impact of DNA repair defects, as identified by these expression models, in clinical outcome was determined in two independent HNSCC patient cohorts uniformly receiving standard-of-care cisplatin-based chemoradiotherapy.

Materials and Methods

Cell culture

UT-SCC are primary tumor cell lines from patients with HNSCC and were established by Prof. Grenman (University of Turku, Turku, Finland) from whom we obtained the cell lines. Cell lines were genotyped (Eurofins) and further characterized by photographs, cell doubling times, and assayed at low passage numbers. Cells were cultured in DMEM þ GlutaMAX (Gibco, Life Technologies), supplemented with 10% FBS (Gibco, Life Technologies), 1% penicillin-streptomycin (Gibco, Life Technologies) and 1 MEM non-essential amino acids solution (NEAA; Gibco, Life Technologies) at 37C, 5% CO₂, and 3% O₂. Cell lines were routinely tested for Mycoplasma contamination.

Expression analysis

RNA was isolated using the DNA AllPrep DNA/RNA Mini Kit (Qiagen). Strand-specific libraries were generated using the TruSeq Stranded mRNA Sample Preparation Kit (Illumina Inc., RS122-2101/2) according to the manufacturer's instructions. Reads were mapped against the GRCh38 human genome with TopHat2.1, using the options fr-firststrand, transcriptome-index, and prefilter multi-hits. Gene expression was determined using HTSeq-count with options stranded and mode union. Differential expression was performed using DESeq2. Genes with log₂-fold change higher than one were classified as differentially expressed, with a false discovery rate of 10%. Data are available at the European Genome-phenome Archive (EGA).

Predictive modeling

For modelling, RPKM values were calculated from the read counts and were scaled to have mean 0 and SD 1. Only genes whose expression was measured in the CGP²⁶ and Peng²⁷

datasets (15,383 genes) were considered to allow model validation in these datasets. In addition, expressions of 13,886 signatures from the Enrichr website (Supplementary Table S1) were calculated using the GSVA BioConductor package. Variables were selected as explained in Supplementary Fig. S1A. Models were fit using the train function from the Caret R package. Options used were trainControl (boot = 10), tuneLength = 10, and reprocessing (centre, scale). pROC package was used to calculate performance. Each of the best performing combinations of feature selection and classifier was trained on the whole dataset 20 times. For predictions, class probabilities were calculated as the average of 20 model predictions. For the model ensemble predictions, class probabilities of the selected models of each type (mmcOnly, olaOnly, or threeClass) were averaged to obtain the final prediction value. The mmcOnly model is available as an R package at bitbucket.org/paulesers/dna-repair-model-ensemble/.

Patient data and material

Biopsies and collection of fresh-frozen tumor material were approved by the Institutional Review Board and all patients granted written informed consent. The retrospective “NKI-CRAD” cohort provided pretreatment tumor biopsy material from 98 patients with hypopharyngeal, laryngeal, or HPV-negative oropharyngeal HNSCC who were treated at the Netherlands Cancer Institute between 2001 and 2014. Samples with tumor percentage <40%, as determined by hematoxylin and eosin (H&E) staining, were excluded. HPV status was determined using p16 and p53 IHC and validated using PCR and RNA-sequencing. HPV-positive patients were excluded. All patients were treated with cisplatin-based chemoradiotherapy. Radiotherapy was administered according to a conventional (35 fractions of 2 Gy over 7 weeks) or DAHANCA scheme (35 fractions of 2 Gy over 6 weeks). Chemotherapy consisted of daily (25 times 6 mg/m²), weekly (6 x 40 mg/m²) or 3-weekly (3 x 100 mg/m²) intravenous administration of cisplatin. Eight patients received intra-arterial cisplatin (4 x 150 mg/m²) according to the RADPLAT trial²⁸. Not all patients completed chemotherapy and the individual total cumulative cisplatin dose was recorded. Overall survival was calculated from the start of the treatment until the event or censored at the time of the last follow-up. Metastasis and locoregional recurrence-free survival were additionally censored at the time of death. The DESIGN cohort consisted of 82 advanced HPV-negative patients with HNSCC who were included in the Dutch multicenter DESIGN study (KWF-A6C7072) and were treated at the VUmc Cancer Center Amsterdam or in the Netherlands Cancer Institute (Amsterdam, the Netherlands). Patient selection criteria were equal to those of the NKI-CRAD cohort; patients who were previously treated with radiotherapy in the head and neck area or chemotherapy were excluded.

Scratch assays

For migration assays, cells were seeded in DMEM + GlutaMAX in an IncuCyte ImageLock 96-wells Plate (Essen Bioscience) at 100% confluence. After 24 hours, a homogeneous scratch was made with the use of the IncuCyte WoundMaker (Essen Bioscience) and washed twice with culture medium. Relative wound density was measured every 4 hours using the IncuCyte (Essen Bioscience). For invasion assays, the same protocol was used with modifications: 100 mg/mL growth factor reduced (GFR) Matrigel (Corning) was added to an IncuCyte ImageLock 96-well plate. After 24 hours, nonsolidified Matrigel was aspirated and cells seeded. After 4 hours, the cells were scratched and washed twice and 8 mg/mL GFR Matrigel was added. Thirty minutes later, culture medium was added and plates were placed in the IncuCyte. Rad51 experiments were performed by adding 500 nmol/L Rad51 inhibitor (B02, SML0364, Sigma) after cell seeding and again after the scratch was made. Inhibitors of DNA-PK NU7026, ATM Ku-55933, and WNT XAV-939 (all from SelleckChem) were added after the scratch was made.

Transwell assays

For migration assays, multiple UT-SCC cell lines were seeded (1.11×10^4 cells/insert) in culture medium with low serum concentration (0.5% FBS) in the top of an 8-mm pore size Transwell insert (Falcon, Thermo Fisher Scientific). Culture medium with 10% FBS was placed in the bottom of a 24-well plate. After 24 hours, membranes were fixed for 2 minutes at room temperature with 3.7% formaldehyde and washed twice with PBS. Cells were permeabilized with 100% methanol for 20 minutes at room temperature, washed twice with PBS, and stained with Hoechst (16 mg/mL; Thermo Fisher Scientific) for 10 minutes at room temperature. The membrane was washed with PBS and cells were removed from the bottom or top of the membrane with a cotton swab. Images were made (10 objective) with the Axiovert 200M (Zeiss) microscope and analyzed using ImageJ software. For invasion experiments, 125 mg/mL GFR Matrigel was added to the top of the insert and left open overnight, allowing the Matrigel to dry onto the membrane. The Matrigel was reconstituted with culture medium for 1 hour.

Statistical analyses

Analyses were performed in the R statistical computing environment. Scripts used for visualizing patient survival are available at <https://github.com/PaulEssers/SurvivalPlots/>. Spindle index, scratch assays, and transwell assays were analyzed by linear-mixed effects models using the nlme R package. Models were fit with “Group” as the fixed effect and the nested term “1|cellLine/Experiment” as the random effect. *P* values reported are those for the fixed effects.

Results

Generation of DNA repair defect predictive expression models using different functional endpoints

A dual drug sensitivity analysis in patient-derived HNSCC cell lines revealed three classes of HNSCC: i) MMC and olaparib sensitive, consistent with pronounced crosslink repair, i.e. HR and FA, pathway defects; ii) MMC sensitive that lack concomitant olaparib sensitivity, likely due to compensatory repair mechanisms and a shift in replication associated repair pathway preference and iii) HNSCC that do not show strong repair defects as identified by drug sensitivity (Fig 1A)¹⁶. We will refer to these groups as *SensMO (i)*, *SensM (ii)* and *Normal (iii)*, respectively. Upon further analysis, HR and FA gene mutations were found in the MMC sensitive HNSCC cell lines, providing the potential genetic causes of the functional repair defects (Supplementary Fig S1)^{16,29}.

We reasoned that changes in the transcriptome will reflect cellular DNA repair status via activation of the DNA damage stress response and downstream mediators. We therefore used RNA-sequencing data from HNSCCs assigned to these three different functional endpoint classes. By pairwise comparisons, 82 genes were found to be differentially expressed between the groups. However, expression levels of these genes varied considerably within groups, and there were no genes that were expressed exclusively and ubiquitously in each group (Fig 1B). Consequently, these traditional gene signatures, derived from these differentially expressed genes performed poorly as predictors for drug sensitivity, as determined by the AUC (the area under the receiver operating characteristic (ROC) curves); Fig 1C). Therefore, we used machine learning algorithms to capture more complex relationships between gene expression and the functional DNA repair endpoint.

We generated a predictive model, exploiting the combined drug sensitivity data, so as to be able to distinguish between the three groups (threeClass). Alternatively, we generated models that were based on the sensitivity to each drug separately with the goal of predicting olaparib or MMC sensitivity only (mmcOnly and olaOnly). Next to gene expression values, we also calculated expression values for established gene signatures³⁰ (Supplementary Table S1). Gene expression values, signature expression values or both were then used as input to train the models. As our ultimate aim was to identify and investigate the effect of DNA crosslink repair defects on patient outcome, we considered a variety of methods to minimize potential method-specific bias and repair-unrelated associations. Relevant genes and gene signatures were preselected using several techniques and were used to train multiple predictive models (see Supplementary Fig S2 for schematic representation and description).

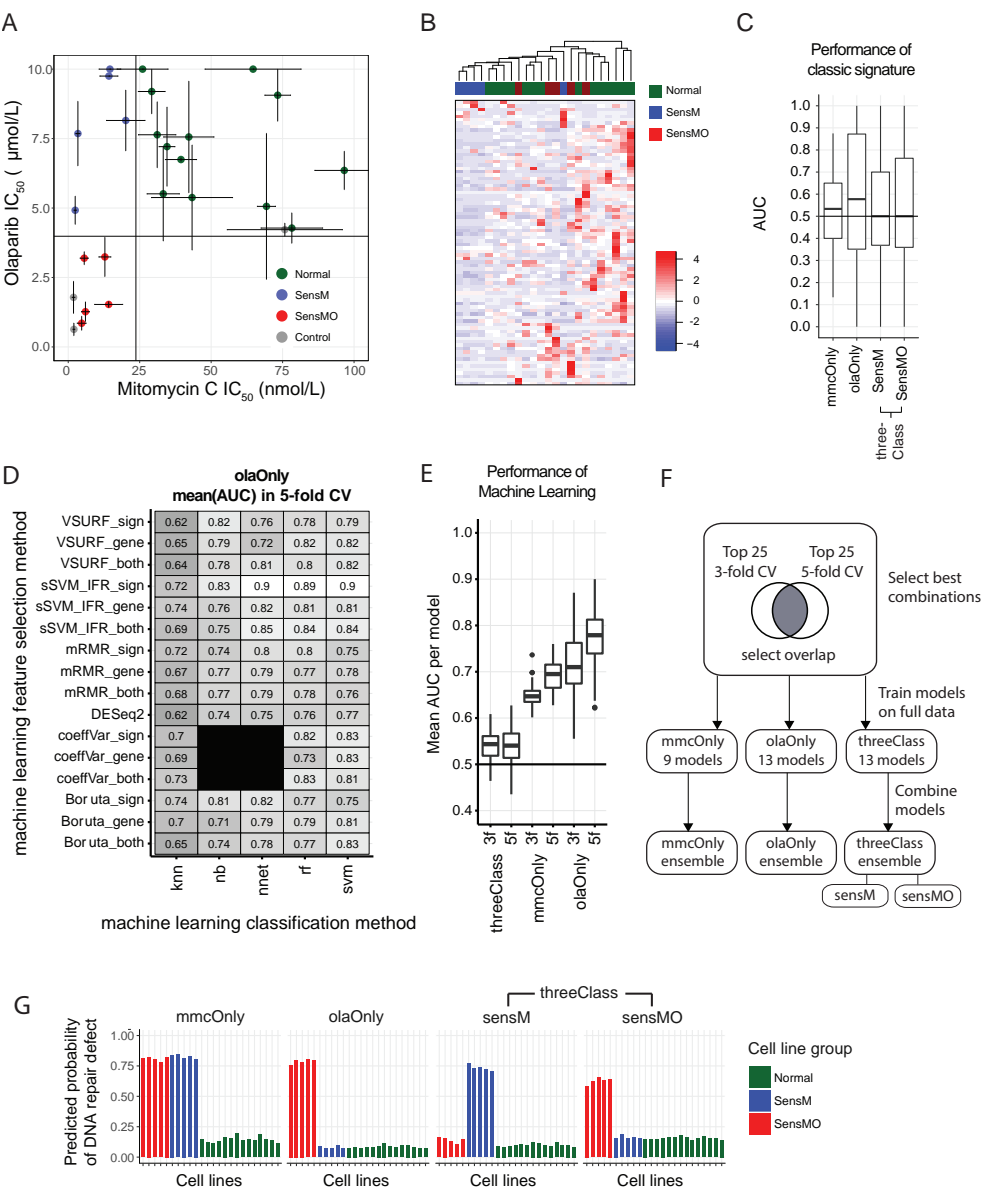


Figure 1. Continued

Error bars are SEM on 3-5 independent experiments. **B)** Heat map of differentially expressed genes, combined from the three possible one-on-one comparisons. **C)** Cross-validation for the traditional signature approach. Cell lines were randomly divided into three groups, two of which were combined to be the training set and one was used as the test set. Differentially expressed genes were identified in the training set. For cell lines in the test set, a score was calculated by adding the normalized (center-scaled) expressions of overexpressed genes and subtracting the expression of under-expressed genes. Area under the ROC (AUC) values were then calculated using these scores. This process was repeated with each of the three groups taking the role of test set, a process known as k-fold cross-validation, with k=3 in this case. Cross-validation was repeated 16 times, resulting in a total of 48 iterations. **D)** Multiple feature selection methods were combined with multiple machine learning methods to determine which combination performs optimally in predicting DNA repair defect classes in our cell lines. Either gene expression (`_gene`), signature expression (`_sign`) or both (`_both`) were used as input. Displayed AUC values are the average of 10 repeats of 5-fold cross-validation for `olaOnly` models. **E)** Mean performance values of all 80 possible combinations for the indicated cross-validations (3fold or 5fold). Each box contains all the values for one cross-validation approach (i.e. `olaOnly_5f` contains all the values shown in D, where each combination is a single data point. For threeClass models, the mean AUC of each class versus the rest was reported. **F)** Overview of the strategy for building the model ensembles. **G)** Performance in the training set to visualize the distinctive group identification of each model ensemble. Note that all AUCs are 1.

To assess the performance of all the different models derived from the individual combinations of feature selection method and classifier, we computed the mean AUC (area under the receiver operating characteristic curves) values from 10 rounds of a 5-fold cross-validation, or 16 rounds of a 3-fold cross-validation of each (Fig 1D and E, Supplementary Table S2). Nearly all combinations predicting olaparib or MMC sensitivity (i.e. `olaOnly` and `mmcOnly`) performed well. We assessed whether the IC_{50} cutoffs that were used to classify the cell lines were crucial for model performance and found that minor variations had little effect on the AUC (Supplementary Fig S3). The threeClass approach performed less well in this analysis (Fig 1E), however this could be a result of low sample size when dividing the data into three groups. This therefore does not necessarily preclude an ability to identify repair defects. Consequently, the threeClass combinations were included in the subsequent validation studies.

For further external validation, we selected the best performing combinations of feature selection and classifier, defined as those within the top 25 of both the 3- and 5-fold cross-validations, resulting in 9 `mmcOnly`, 13 `olaOnly` and 13 `threeClass` combinations (Fig 1F, Supplementary Table S2). We used these to train models on the full HNSCC cell line expression dataset and combined the resulting models into ensembles (Fig 1F). Scores of these ensembles that represent the probability of each cell line to fall into any of the sensitivity groups (normal, *SensM*, *SensMO*) are shown in Fig 1G.

Validation in external datasets

To verify that our models can be used to predict DNA repair defects, we next determined whether our models are predictive for sensitivity of DNA damaging agents in other datasets. HNSCC cell line responses to a large number of compounds were determined as part of the Cancer Genome Project (CGP)²⁶. The predictions from both MMC sensitivity relevant ensembles, mmcOnly and threeClass (*SensM* and *SensMO* combined), correlate with MMC IC₅₀ in this dataset (Fig 2A). This also translated into an ability to identify DNA repair defective cell lines, as classified by their MMC sensitivity (Fig 2B). We did not observe a correlation between the olaOnly ensemble scores and MMC sensitivity. The olaOnly model ensemble was consequently unable to distinguish MMC sensitive cells. For olaparib sensitivity, we did not observe a correlation with model prediction scores for any ensemble, likely due to uncertainty in the computationally inferred IC₅₀ values in the CGP HNSCC cell lines (Supplementary Fig S4A-C). Importantly, all three model ensembles also predicted the presence of mutations in DNA crosslink repair genes (Supplementary Fig S4B; Supplementary Table S3), with the mmcOnly and threeClass model ensembles showing the best performance (Fig 2C).

As a complementary validation strategy to show that these models predict DNA repair defects beyond drug sensitivity, we determined the capability of our models to detect alterations in DDR and DNA repair. In a study by Peng *et al.*²⁷, components of the DDR and DNA repair pathways were knocked down using shRNA in breast cancer cells and gene expression was measured using micro-arrays. We tested whether the model ensembles can distinguish cells treated with no or control shRNA from those treated with the tested shRNAs against DDR genes. We find that the ensembles accurately predicted DDR defects caused by the loss of ATM, ATR, CHK1 CHK2 and 53BP1 (Fig 2D). The performance for BRCA1 and Rad51 knockdown was more heterogeneous (Fig 2D), a characteristic that could be explained by the differences in downstream signaling and transcriptomics that are caused by the complete loss of the protein compared to deleterious gene mutations. Interestingly, loss of the DNA damage response mediator MCPH1 (BRIT1) results in a consistently and greatly reduced signal for DNA repair defects, suggesting an essential role of the chromatin remodeling cyclin response in establishing the distinctive gene expression patterns that are detected by these models.

On the whole, we conclude that both the mmcOnly and the threeClass model ensembles capture features of DDR and DNA repair defects, i.e. drug sensitivity and DDR gene mutations, and repair defects resulting from DDR gene knockdowns.

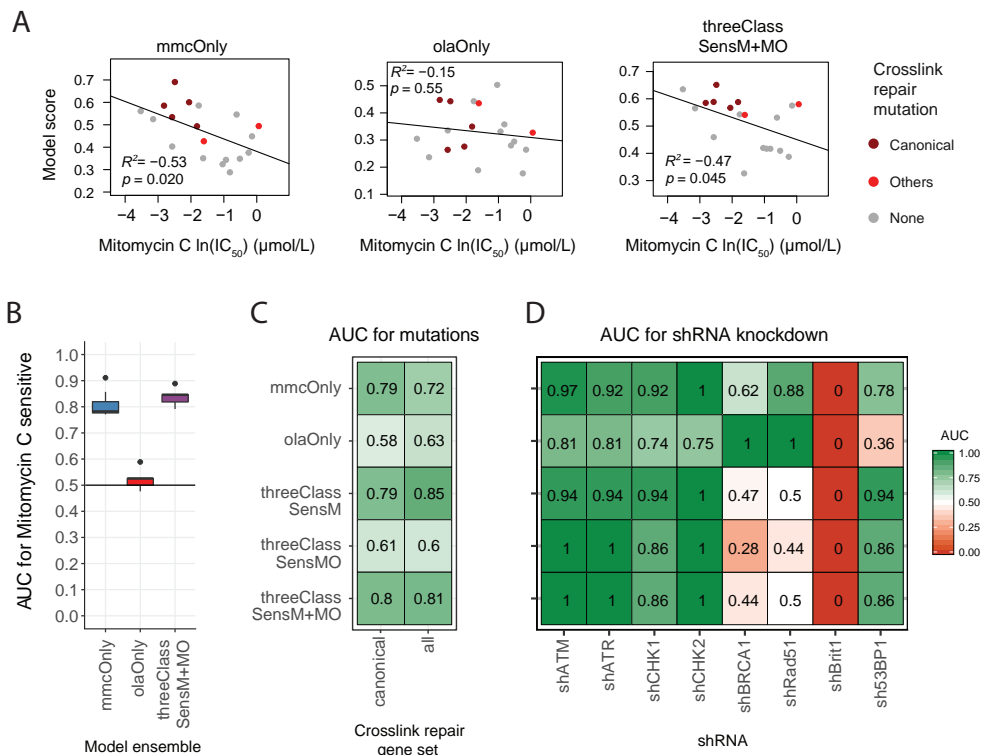


Figure 2. External validation of model ensemble performance. A) Correlation between HNSCC cell line MMC IC_{50} values and their model scores in the CGP dataset. For the threeClass model ensemble, the sum of predicted probabilities for the SensM and SensMO group to predict MMC sensitivity was used, as both these groups are MMC sensitive. Cell lines with reported mutations in crosslink repair pathway genes (Supplementary Table S3, 27 HR/FA genes for the canonical gene set) are highlighted in red. B) AUC for DNA repair defect prediction of the indicated model ensembles as classified by MMC sensitivity. AUC are shown for a range of IC_{50} cutoffs (from 20 to 50% of the maximum tested dose; see Supplementary Fig S4) to avoid reliance on a distinct arbitrary cutoff. C) AUC for detecting mutations in crosslink repair genes in the HNSCC CGP data set (gene list in Supplementary Table S3). D) Performance of the model ensembles in the Peng *et al.* dataset (27) to predict DNA damage response and repair defects. Samples treated with shRNA against the respective DNA repair genes were considered defective (n=4 for BRCA1, RAD51 and BRIT1, n=6 for ATM, ATR, CHK1, CHK2 and 53BP1), while scrambled shRNA (n=3-4) or non-treated samples (n=3, not used in BRCA1, RAD51 and BRIT1 experiments) were considered proficient.

Predicted DNA repair defect association with poor clinical outcome

Using these two validated model ensembles, we determined whether survival is different in patients with tumors that show repair defect-associated gene expression patterns. For this purpose, we used pre-treatment tumor biopsy gene expression data from our in-house HNSCC cohort (NKI-CRAD) of 98 patients with advanced stage HPV-negative tumors

(characteristics in Supplementary Table S4), who underwent radiotherapy with cisplatin based treatment regimens. Consistent with previous reports, tumor site³¹ and cumulative cisplatin dose³² are prognostic factors in our NKI-CRAD cohort. DNA repair defect scores were established using the mmcOnly model ensemble for all tumor samples. These scores did not correlate with any of the potentially prognostic clinical factors (Supplementary Fig S5A and Supplementary Table S5A). We then divided the patient population into two equal-sized groups, based on the ranking of the mmcOnly ensemble scores, and found a trend towards poorer prognosis for the patients with the higher scores (Fig 3A).

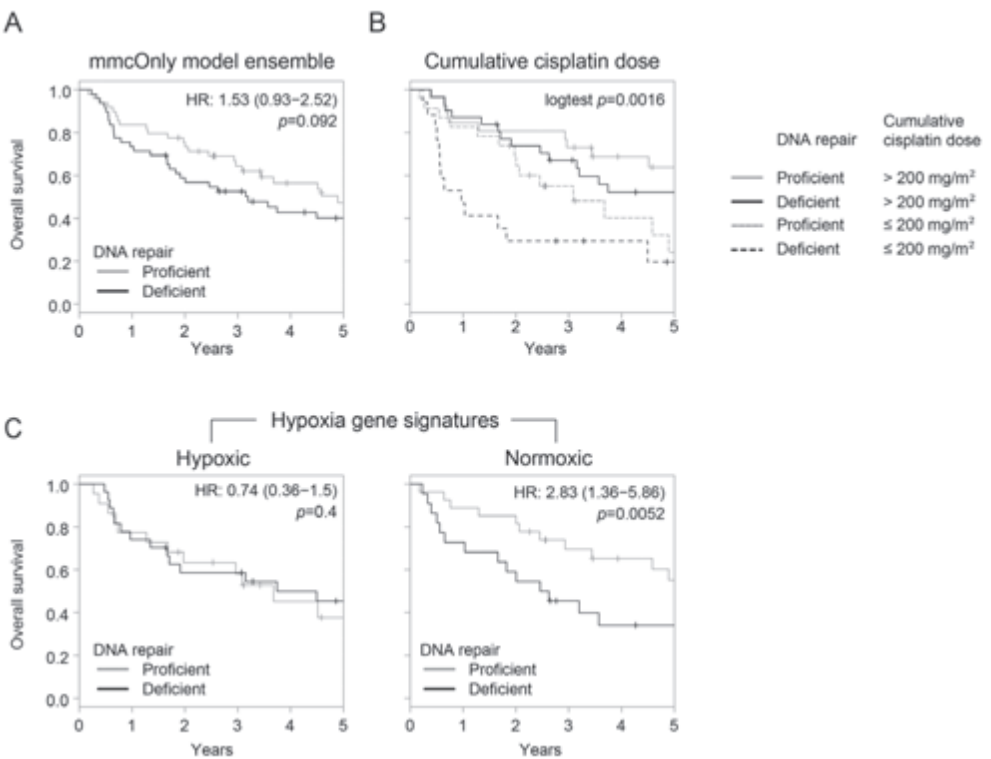


Figure 3. DNA repair defects are prognostic for chemo-radiotherapy treated HNSCC patients. A) Overall survival of advanced HPV-negative HNSCC patients (NKI-CRAD cohort) based on DNA repair defect prediction. The patient cohort was divided into two equal-sized groups based on the median mmcOnly model ensemble score. The high scoring group was classified as DNA repair deficient. **B)** The cumulative cisplatin dose of >200 mg/m² was used to discriminate patients who received high cisplatin doses from those with less to evaluate the survival association with DNA repair defect prediction. **C)** mmcOnly model ensemble prediction association with overall survival in patients with hypoxic or normoxic tumours. Insets in A-C show univariate HR and associated p -values.

Hypoxia and cumulative cisplatin dose are independent prognostic factors in HNSCC (Supplementary Fig S6A-B)³³. To consider the hypoxia status, the average expression of five published hypoxia signatures³⁴ was calculated and used to divide the patient population into equally-sized hypoxic and normoxic subgroups. Subsequent DNA repair defect classification by the mmcOnly models in these subgroups shows that the poor prognosis association is particularly strong in normoxic tumor samples (Fig 3C). This interaction is statistically significant ($p=0.011$) and was therefore included in subsequent multivariate analyses. Upon closer inspection, we find that overall gene expression changes from repair defects (high in DNA repair defect score) under normoxia correlate highly with those from hypoxia exposure (as determined by the hypoxia score) in repair proficient tumors in both cohorts, even though the actual patient classification does not (Supplementary Fig S7A-D). Pathway enrichment analyses revealed that the downstream effects of hypoxia and DNA repair defects are partially overlapping (Supplementary Fig S7E-G). This may explain the lack of an endpoint association within the hypoxic tumor group and points to common repair defect sources or downstream signaling effects.

Tumors harboring DNA crosslink repair defects are expected to be sensitive to cisplatin treatment. Variations in the cumulative cisplatin dose in our patients, as a result of discontinuation due to systemic toxicities, gave us the opportunity to investigate this hypothesis. The patients were assigned to a high (>200 mg/m²) or low (≤ 200 mg/m²) cumulative cisplatin dose group according to previous reports, resulting in a univariate HR of 2.3 for the low dose patients ($p=0.0014$). Overall, patients with predicted DNA repair defective tumors benefitted most from the high cumulative cisplatin dose (Fig 3B and Supplementary Fig S8A-B).

To consider both the prognostic clinical factors and the herein uncovered interaction with hypoxia, we combined predicted DNA repair status, hypoxia status, cisplatin dose and tumor site in a multivariate model (Fig 4A). Our analysis shows that expression profiles associated with DNA repair defects were strongly indicative of poor prognosis (HR=4.01, $p=0.00091$). The interaction between DNA repair and hypoxia status was also significant in this model (HR=0.19, $p=0.0066$), confirming that hypoxia and DNA repair defects are not additive (Fig 4A, left). To independently test this poor patient prognosis association, we next computed DNA repair defect model scores for tumor samples from a second independent cohort with 82 advanced and HPV negative HNSCC patients who were chemo-radiotherapy treated and from which pre-treatment material was collected within the multicenter DESIGN project (characteristics in Supplementary Table S4). This independent validation, shown in Figure 4A (right), confirms that predicted DNA repair defects, as determined by our model ensemble, are significantly associated with poor prognosis (HR=3.2, $p=0.023$).

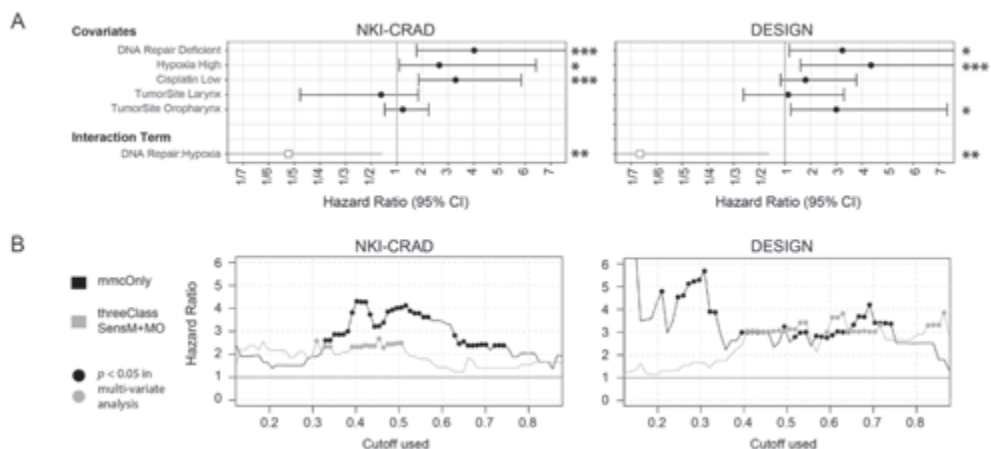


Figure 4. Independent validation of the prognostic value of the DNA repair defect models. A) Forest plots for multivariate Cox analysis in the NKI-CRAD and the DESIGN HNSCC patient cohorts, showing the DNA repair defect model (mmcOnly ensemble) association with survival in multi-variate analyses that also include tumour site, cumulative cisplatin dose, hypoxia status and the interaction term between DNA repair defects and hypoxia as indicated. The hazard ratios for patients with tumors that are both DNA repair deficient and hypoxic, are 2.03 (HRDNA_repair_defect * HRhypoxia * HRinteraction = $4.01 \times 2.64 \times 0.19$) in the NKI-CRAD cohort and 2.12 ($= 3.24 \times 4.36 \times 0.15$) in the DESIGN cohort. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. **B)** Repeated multivariate Cox HR analyses in unequally-sized patient groups as indicated on the X-axis. Used cutoffs for the “DNA repair deficient” group comprised the 85% (left axis) to the 15% (right axis) top scoring patients. Multivariate Cox HR, with the same variables as in Fig 4A, are shown on the y-axis for each of these splits and model ensemble as indicated. Note that the values shown in Fig 3C correspond to the 0.5 cutoff (dotted grey line). Dots represent HR values with p -value under 0.05.

By splitting the cohorts at their median scores into two equal-sized groups, as is common practice in biomarker studies, the above analyses stipulate that DNA repair defects are present in half of the HNSCC. Yet there is little support for this suggestion (50% DNA repair-defective tumors). Furthermore, prevalence can also differ between cohorts. To verify that our findings are not biased by certain cutoff points, we evaluated the multivariate hazard ratio at a range of cutoffs (Fig 4B). The varied cutoffs for repair defect status assignment are based on the proportion of patients in one or the other category as defined by their model score ranking, resulting in different score cutoffs for both cohorts in this analysis. Doing so, we find that DNA repair defects called by the mmcOnly model ensemble were significantly associated with poor prognosis at a broad range of cutoffs in both the NKI-CRAD cohort and the second independent DESIGN cohort.

We next assessed the DNA repair defect / poor prognosis association by using the threeClass model ensemble. We find that this ensemble (threeClass model) showed highly similar results, confirming a DNA repair defect / poor prognosis association and that it also predicts

poor survival in both advanced HNSCC patient cohorts largely independent of the chosen cutoffs (Fig 4B and Supplementary Fig S9). As a comparison, clinical stage was not found to be significantly associated with poor survival in either cohort (Supplementary Fig S10).

Finally, to further reveal the origin of the poor prognosis association and cisplatin responsiveness, we analyzed the association with different outcome endpoints. We find little evidence for a role in locoregional control (Fig 5A). However, we observed significantly increased hazard ratios for the occurrence of distant metastasis, in particular in the NCI-CRAD cohort (Fig 5A), which suggests that DNA repair defective tumors may be more prone to metastasize. As for overall survival (Fig 3C), this effect was most pronounced in normoxic tumors that also metastasize less, unless repair defective (Fig 5B). Given the lack of endpoint association in hypoxic tumors, we further investigated the effect of cumulative cisplatin dose in normoxic tumors only. Hazard ratios for DNA repair defects were higher in the patient group treated with low dose cisplatin at all cutoffs (Fig 5CD). Although only statistically significant at one cutoff, likely due to low patient numbers, it points to the benefit of high dose cisplatin to reduce metastasis occurrence in patients with DNA repair defective tumors.

DNA repair defect association with migratory and invasive behavior.

These findings prompted us to explore the link between DNA repair defects and metastasis in *in vitro* studies with the confirmed crosslink repair impaired HNSCCs. We selected several representative cell lines for each sensitivity class (*Normal*, *SensM* and *SensMO*; described in Fig 1A). To investigate the link between the observed metastatic behavior and DNA repair defects, we used two assays to evaluate migratory behavior in these cell lines. In scratch assays, a confluent layer of cells is scratched and the rate at which the remaining cells close the resulting gap is determined³⁵. In trans-well assays, the ability of cells to migrate through a porous membrane is measured³⁵. We found that cell lines that show MMC sensitivity (*SensM* and *SensMO*) migrated significantly more often and faster than *Normal* classified cell lines in both the scratch assays (Fig 6AB) and trans-well migration assays (Fig 6C). To evaluate invasive capacity *in vitro*, both assays were modified by depositing extracellular matrix proteins (matrigel) in the scratch or on the membrane. We observed that in the scratch and trans-well invasion assays, the *SensMO* cell lines were able to invade through matrigel, while this was not evident in the *SensM* or *Normal* cell lines (Fig 6D-F). To further test whether this is DNA repair defect related, we inhibited HR/FA DNA repair in the repair proficient *Normal* cell lines using the RAD51 inhibitor B02 for 24h. This resulted in strong induction of migratory ($p=1.49 \times 10^{-5}$) and invasive behavior ($p=1.8 \times 10^{-6}$ in ANOVA; Fig 6GH), further supporting a possible direct link between DNA repair defects and invasiveness in HNSCC.

CTTTGGGAGGCGGAGCGGGGGATTACTTGGAGTAGAGTCTTCAGACAGCGTGGCCAACTGTGTGAATCCCGCTCTCTATTAATAAATAACAAAAATTAAGCTGGGCGTGGTGGGTGCTGTATTCGACGCTATTGGGAGGTGAGGAG

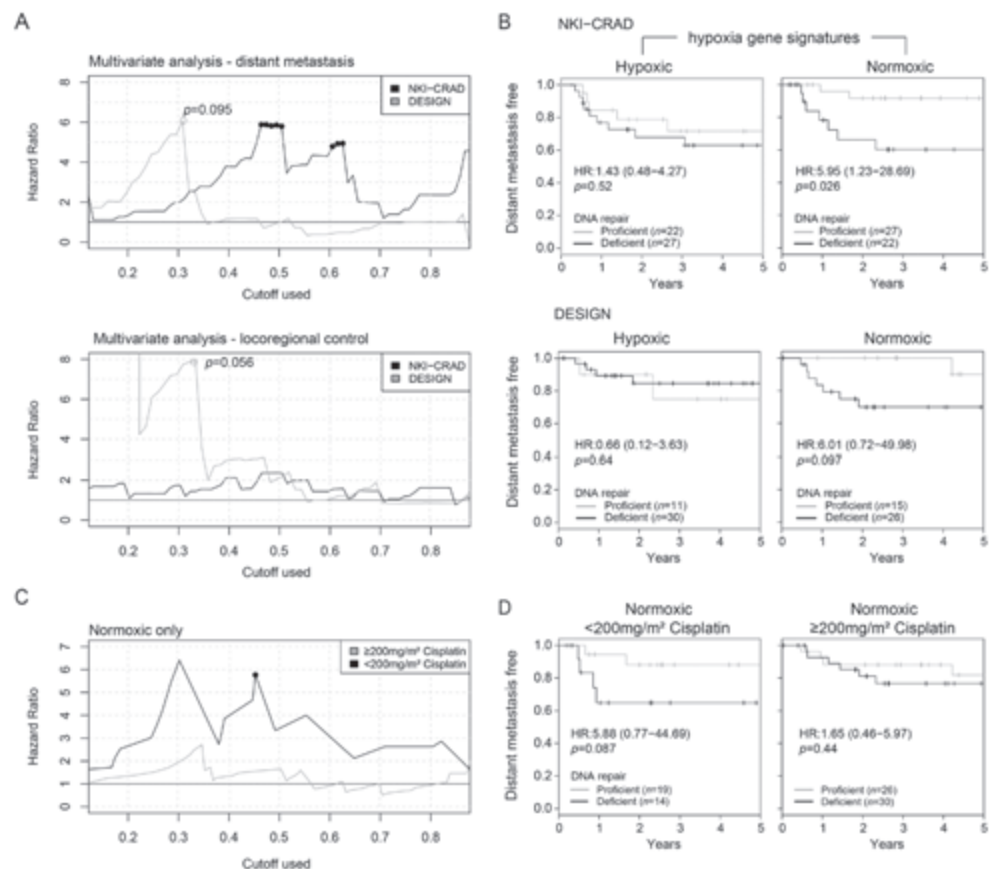


Figure 5. Association of DNA repair defects with metastasis and cisplatin response. **A)** Patients were repeatedly divided into two groups with the “DNA repair deficient” group comprising the top scoring 85% (left) to 15% (right) of patients as based on their predicted probability of DNA repair defects by the mmcOnly model ensemble. Multivariate Cox HR, with the same variables as in Fig 4B, are shown on the y-axis for each possible split, with distant metastasis or locoregional recurrence as clinical outcome. Dots represent splits where the associated p -value was under 0.05. **B)** Kaplan-Meier plots at best splits (median in NKI-CRAD and 0.31 in DESIGN) are shown for DNA repair defects (mmcOnly ensemble) within the hypoxic or normoxic classified tumour patient groups as indicated. Insets show univariate HR and associated p values. **C)** HR plots for metastasis occurrence as in A in all patients (combining the NKI-CRAD and DESIGN cohort) according to their received cumulative cisplatin dose. **D)** Kaplan-Meier plots showing distant metastasis free survival in the high or low cumulative cisplatin patient group and data according to DNA repair status based on the median in all patients with normoxic tumors. Insets show univariate HR and associated p values.

Interestingly, the quick induction also indicates a lack of a requirement for genetic adaptation. Treatment with equal concentrations of Rad51 inhibitor was lethal for both *SensM* and *SensMO* cell lines (Supplementary Fig S11), precluding further epistasis experiments. These *in vitro* data support the notion derived from the clinical data that functional FA/HR repair pathway defects are associated with a more aggressive, metastasis-prone nature.

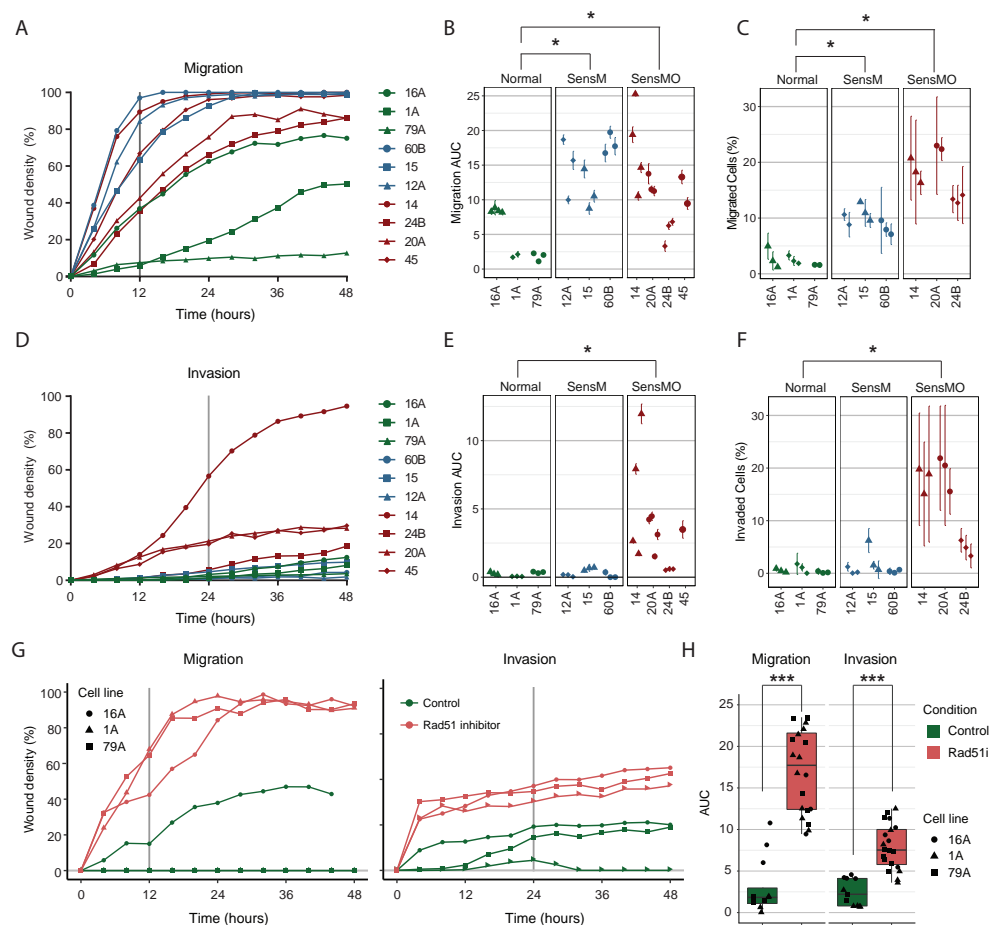


Figure 6. DNA repair defective cell lines have a strong migratory and invasive phenotype. **A)** Migration as determined in scratch assays of SensMO (red), SensM (blue) and normal (green) cell lines. After the scratch was made, cell density in the scratch area was monitored for 48 hours. Values shown are the average of multiple experiments. **B)** Areas under the curves (AUC) over the first 12h were determined for the individual independent experiments (from A) that were performed on separate days and are shown as individual points, with error bars representing SD from the 5 technical replicates. Asterisks point to statistical significances between the groups, as determined by linear mixed effects models. **C)** Proportion of migrated cells after 24 hours, as determined by trans-well migration assays. Individual points represent independent experiments. Error bars are the SD ($\text{sdratio} = \sqrt{(\sigma^2 + \sigma^2)} \cdot \sqrt{n}$) on the ratios of cell numbers on the bottom to the top of the membranes ($n=2-3$). **D and E)** As in A and B with the modification that wells were covered with extracellular matrix after the scratch was made. AUC were calculated over the first 24h. **F)** As in C, with the modification that membranes were covered in matrigel prior to the assay. **G)** Migration and invasion scratch assays for Normal cell lines treated with B02. Values shown are the average of multiple experiments. **H)** Area under the curve for each experiment in G. ANOVA was used to determine statistical significance using the formula: “ \sim Inhibitor*CellLine + Experiment”; the interaction term is added to account for cell-line specific response. The p -value for the inhibitor is reported.

Mechanistic links between DNA repair defect and migration

In search for potential mechanistic links and prompted by the high specificity in identifying DNA damage response deregulation by ATM knockdown (Fig 2D, Supplementary Fig S7), we hypothesized that the activation of DNA damage response pathways induces migration. This is in line with our initial position that unrepaired damage from oxidative and endogenous sources would provoke a DNA damage response and accompanying transcriptional changes that allows us to identify repair defects. Indeed, ATM depletion was previously reported to reduce migration and invasion. This was shown to occur through IL-8 signaling regulated by ATM activation and induced by oxidative stress³⁶. We find that *CXCL8* (IL-8) is consistently upregulated in the migrating but not in the invading cell lines (Fig 7A). Inhibition of the DNA damage response protein ATM reduced migration mildly in some cell lines, suggesting that this is not specific to DNA repair defective status (Fig 7B, Supplementary Fig S12). A second key player in the DNA damage response network is the DNA-dependent protein kinase, DNA-PK. DNA-PK has been shown to promote migration and invasion in prostate cancer, possibly through DNA-PK-mediated gene expression changes³⁷. Others postulated DNA-PK influences invasion and migration through the control of the secretion of multiple metastasis associated proteins³⁸. When tested in our cell lines, DNA-PK inhibition reduced migration in the *BRCA1*-mutated UT-SCC-60B (Fig 7C, Supplementary Fig S12). Together our data suggest that DNA-PK activation and IL8-expression mediated processes may be responsible for the migration of cells within the migratory SensM group.

As the invasive SensMO group did not show such mechanistic links, we next evaluated the presence of known metastasis promoting features. Metastasis is partly associated with epithelial to mesenchymal transition (EMT)³⁹. DNA damage, such as induced by radiation, has been reported to promote EMT and migration in cell lines. We therefore tested whether DNA repair defective cell lines show mesenchymal characteristics, elongated cell body and expression of key transcription factors⁴⁰. Inspection by light microscopy revealed that both *SensM* and *SensMO* cell lines are more spindle-shaped than *Normal* cell lines, with *SensMO* cells showing the most elongated phenotype ($p < 10^{-16}$ for both; Fig 7D, Supplementary Fig S13A). However, while some migratory and invasive cell lines exhibited mesenchymal characteristics,

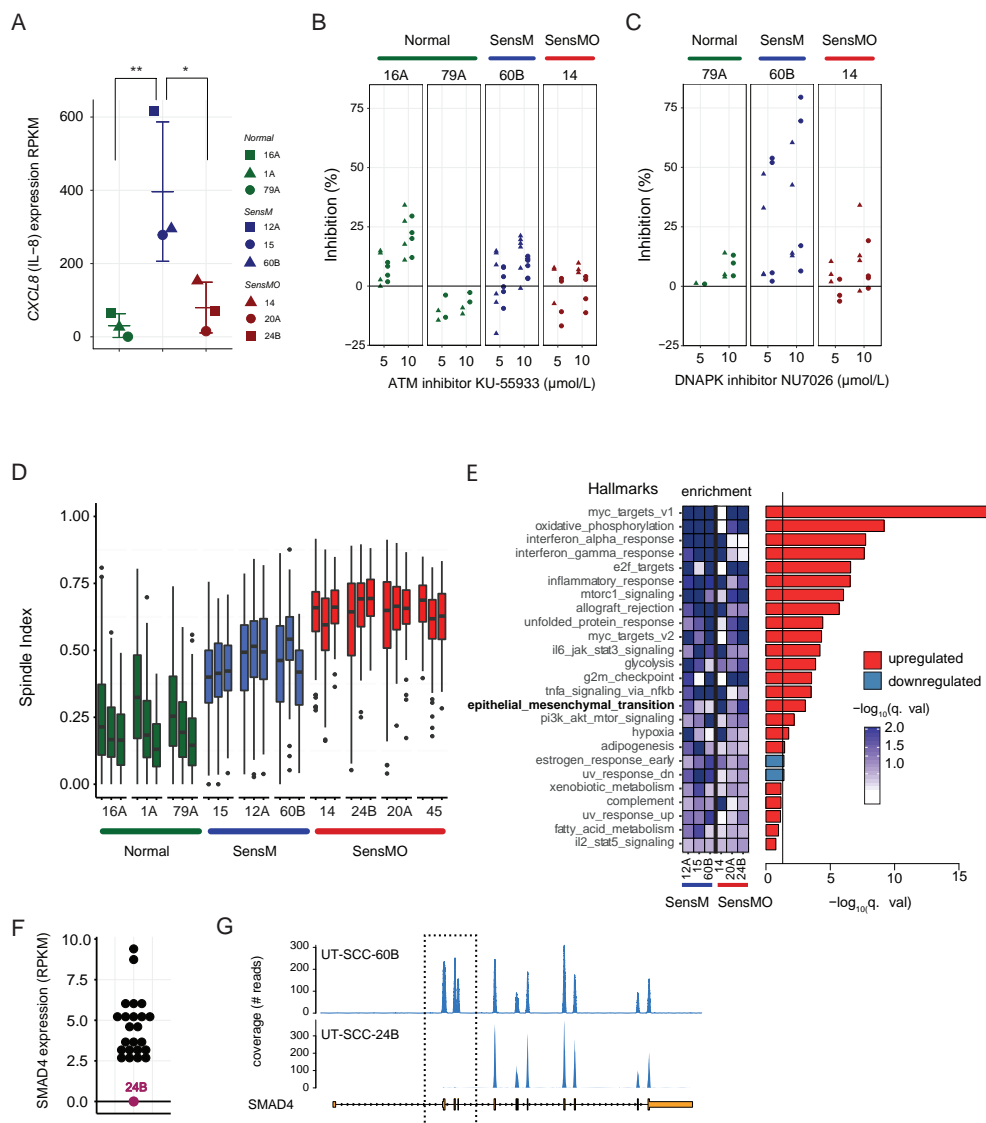


Figure 7. Mechanisms linking DNA repair defects to increased migration. **A**) CXCL8 (IL8) expression levels in the three DNA repair classes (Normal, SensM, SensMO) with pairwise t test p -values of * <0.05 and ** <0.01 . **B**) Migration inhibition values (%) from ATM inhibition by KU-55933 as determined in the scratch assay at 12 (triangles) and 24 (filled circles) hours in the indicated HNSCC cell lines. Symbols show mean values of individual experiments. **C**) Migration inhibition values from DNA-PK inhibition by NU-7026 as determined in the scratch assay at 12 (triangles) and 24 (filled circles) hours in the indicated HNSCC cell lines. **D**) Distributions of spindle index values in the HNSCC cell lines from the three different DNA repair classes. The spindle index is defined as $1 - (\text{width}/\text{length})$, such that a perfectly round cell ($\text{width}=\text{length}$) has index 0 and increases as cells are more elongated. Data from $n=3$ independent experiments with 100 cells each are shown. **E**) Geneset enrichment analysis using MSigDB hallmarks and comparing the migrating to the Normal cell lines. **F**) SMAD4 expression levels in all analysed cell lines. **G**) DNA sequencing read coverage plots in the UT-SCC-24B and UT-SCC-60B indicating a loss of exons 1-3 in UT-SCC-24B.

we did not find consistent differences between the cell line classes for expression of the classical EMT markers or EMT related signatures⁴¹ when using all cell lines. Assuming repair proficient cell lines may also exhibit EMT (unrelated or not caused by DNA repair defects), we performed gene set enrichment analysis (GSEA) using only the cell lines for which migration and invasion data was available. EMT related genes were significantly enriched (Fig 7E), pointing to a possible mechanistic role for EMT. When testing EMT related processes^{41,42}, we found cell adhesion and migration genes to be enriched (Supplementary Fig S13B) and most migrating cells to express mesenchymal markers (Supplementary Fig S13C). EMT and cellular invasive behavior as mediated by TGF β requires SMAD4^{43,44}. SMAD4 loss, that can cause spontaneous head and neck cancer in mice, however, promotes squamous cell carcinoma metastasis and was shown to influence metastasis formation in co-operation with E-cadherin in other models^{45,46}. Using DNA capture sequencing analysis¹⁶, we find that UT-SCC-24B lacks SMAD4 expression due to a genomic deletion of the first three exons (Fig 7FG). This finding links HR/FA repair defects with SMAD4 and invasive behavior in one of the cell lines.

In conclusion, while we find evidence that repair defects can be linked to migratory behavior through the engagement of key DNA damage response players and their links to migration as reported by others, mechanisms in the invasive repair defective class may depend on genetic context and may be more multifaceted and complex.

Discussion

In this study, we set out to investigate the effect of DNA repair defects on clinical outcome in HNSCC. We defined DNA repair defects by functional outcomes, hypersensitivity to DNA damaging agents, which integrates the proficiency of all DNA damage response and DNA repair pathways, including adaptations that have arisen to any defects in these pathways. We reasoned that the transcriptome reflects the sum of all failed or successful repair activities on endogenously generated damage and therefore encompasses defects and the activation of backup repair pathways and other coping mechanisms. Re-establishment of DNA repair, through compensatory repair activities, is a mechanism for acquiring treatment resistance^{9,47}, which does not revert genomic scars or mutations but is presumably reflected in the transcriptome. Together, we consider this an important addition to the current arsenal of DNA repair biomarkers.

Our strategy allowed us to investigate the effect of DNA repair defects in patients with HNSCC. This revealed that patients with tumors predicted to harbor DNA repair defects

showed worse overall survival in two independent patient cohorts. We conclude that they constitute a HNSCC group with a more aggressive behavior. Hypoxia is known to alter cellular repair capacity and pathway choice and is a poor prognosis biomarker⁴⁸. We find a noticeable interaction with hypoxia that suggests that, in patients with hypoxic tumors, prognosis is not further decreased by repair defects. Hypoxia has been shown to promote metastasis and to downregulate HR/FA repair processes^{49–51}. A considerable overlap in gene expression changes to the tumors with DNA repair defects indicated common repair defect sources or downstream signaling effects. Alternatively, the models may be unable to predict repair defects in hypoxic tumors, since the model ensembles were generated under nonhypoxic conditions. The RNA expression patterns in the DNA repair defective cells were likely caused (in part) by oxidative stress. In contrast to our findings, in breast, ovarian and lung cancer, DNA repair defects in tumors were found to be associated with a favorable prognosis, possibly due to the treatment with DNA damaging agents^{52–55}. Current treatments for HNSCC may hence not fully exploit the cellular sensitivities of tumors with DNA crosslink repair defects. The apparent increased benefit of high cumulative cisplatin dose in our clinical data supports this interpretation. Given the risk, such patients could be candidates for additional systemic treatment with crosslinkers, PARP inhibitors or other novel therapies targeting such defects.

The *in vitro* migratory and invasive capacity that was exclusive to the *SensMO* cell lines confirmed a proposed propensity of tumors with DNA repair defects to metastasize. In addition, we found that the induction of DNA repair defects, in cell lines that were otherwise repair proficient, results in rapid acquisition of this invasive phenotype. This suggests that a direct link and signaling mechanism underlies this effect, rather than genetic contextual adaptations to such repair defects. In line with these findings, a recent study showed that chromosomal instability induces invasive behavior through activation of NF-κB signaling²⁵. Others have reported that knockdown of Fanconi anemia genes in HNSCC cell lines results in increased invasion through activation of the NHEJ repair pathway and via DNA-PK and Rac1 signaling⁵⁶. Our DNA-PK inhibition data and differential pathway expression analyses support the engagement of these migration promoting mechanisms in DNA repair deficient HNSCC cell lines and provides a mechanistic link to a higher propensity to metastasize. DNA-PK activation mediated migration is also in line with an increased engagement of NHEJ as indicated by the PARP inhibitor resistance and connects repair defects to the reports from others that show a role for these proteins in migration, invasion and metastasis. Consistent with the notion that loss of FA/HR DNA repair may induce metastatic behavior, knockdown of BRCA1 and FANCD2 was reported to cause dedifferentiation and induce EMT

in human mammary epithelial cells *in vitro*⁵⁶. Cells that have undergone EMT downregulate apoptosis⁵⁷ and show increased resistance to DNA damage⁵⁸. At first glance, our data seems inconsistent with such observations. Upon closer inspection, we find hints of an increased mesenchymal phenotype in the gene expression data and cellular morphology of the invasive cell lines. In recent years, however, it has become clear that cancer cells tend not to be in a solely epithelial or mesenchymal state, but that many intermediate states exists⁵⁹. Also, DNA repair defective cell lines may show a partial EMT phenotype, this as part of a defense mechanism against further DNA damage and the effects thereof.

In the cell line panel we observed that invasive behavior and DNA repair defects overlap to a large degree. Considering the recent data from Bakhoum *et al.*²⁵, our data further support a causal relationship between defective DNA repair that promotes chromosomal instability and migratory behavior. However, given this overlap in features, in patient material and based on our transcriptome analysis alone, a clear distinction between DNA repair deficiency and the sole and independent presentation of invasive features is therefore formally not possible. The association with cisplatin response and the inability to predict survival changes in hypoxic samples or to identify all metastasis prone cases points however to repair defect features in contrast to pure metastasis features. For more accurate identification of tumor DNA repair defects in individual patients, integration of transcriptomics, genomic rearrangements, pathogenic mutations and phospho-proteomics is desirable. More cohorts of homogenously treated HSNCC patients, for whom full transcriptomics, genomic data and carefully recorded clinical data are available, are required to further optimize the identification of patients with tumors harboring DNA repair defects with the purpose to provide alternative treatment options to this poor prognosis group.

In summary, we find that DNA repair defective cell lines are more motile and invasive *in vitro* and we have developed gene expression-based models to detect DNA repair defects in HNSCC. Importantly, using these models, we find that these defects are associated with poor overall survival in two independent patient cohorts thereby confirming their prognostic value and the detrimental role of tumor repair defects.

References

1. Chae YK, Anker JF, Carneiro BA, *et al.* Genomic landscape of DNA repair genes in cancer. *Oncotarget*. 2016;7(17):23312-23321. doi:10.18632/oncotarget.8196
2. Van Den Broek AJ, Schmidt MK, Van 't Veer LJ, Tollenaar RAEM, Van Leeuwen FE. Worse breast cancer prognosis of BRCA1/BRCA2 mutation carriers: What's the evidence? A systematic review with meta-analysis. *PLoS One*. 2015;10(3):1-29. doi:10.1371/journal.pone.0120189
3. Mersch J, Jackson MA, Park M, *et al.* Cancers associated with BRCA1 and BRCA2 mutations other than breast and ovarian. *Cancer*. 2015;121(2):269-275. doi:10.1002/cncr.29041
4. Pennington KP, Walsh T, Harrell MI, *et al.* Germline and somatic mutations in homologous recombination genes predict platinum response and survival in ovarian, fallopian tube, and peritoneal carcinomas. *Clin Cancer Res*. 2014;20(3):764-775. doi:10.1158/1078-0432.CCR-13-2287
5. Robson M, Im S-A, Senkus E, *et al.* Olaparib for Metastatic Breast Cancer in Patients with a Germline BRCA Mutation. *N Engl J Med*. 2017;377(6):523-533. doi:10.1056/NEJMoa1706450
6. Lattimore V, Currie M, Lintott C, Sullivan J, Robinson BA, Walker LC. Meeting the challenges of interpreting variants of unknown clinical significance in BRCA testing. *N Z Med J*. 2015;128(1419):56-61.
7. Watkins J a, Irshad S, Grigoriadis A, Tutt ANJ. Genomic scars as biomarkers of homologous recombination deficiency and drug response in breast and ovarian cancers. *Breast Cancer Res*. 2014;16(3):211. doi:10.1186/bcr3670
8. Telli ML, Jensen KC, Vinayak S, *et al.* Phase II study of gemcitabine, carboplatin, and iniparib as neoadjuvant therapy for triple-negative and BRCA1/2 mutation-associated breast cancer with assessment of a tumor-based measure of genomic instability: PrECOG 0105. *J Clin Oncol*. 2015;33(17):1895-1901. doi:10.1200/JCO.2014.57.0085
9. Jaspers JE, Kersbergen A, Boon U, *et al.* Loss of 53BP1 causes PARP inhibitor resistance in BRCA1-mutated mouse mammary tumors. *Cancer Discov*. 2013;3(1):68-81. doi:10.1158/2159-8290.CD-12-0049
10. Henneman L, van Miltenburg MH, Michalak EM, *et al.* Selective resistance to the PARP inhibitor olaparib in a mouse model for BRCA1-deficient metaplastic breast cancer. *Proc Natl Acad Sci*. 2015;112(27):8409-8414. doi:10.1073/pnas.1500223112
11. Jeggo PA, Pearl LH, Carr AM. DNA repair, genome stability and cancer: A historical perspective. *Nat Rev Cancer*. 2016;16(1):35-42. doi:10.1038/nrc.2015.4
12. Knijnenburg TA, Wang L, Zimmermann MT, *et al.* Genomic and Molecular Landscape of DNA Damage Repair Deficiency across The Cancer Genome Atlas. *Cell Rep*. 2018;23(1):239-254.e6. doi:10.1016/j.celrep.2018.03.076
13. Willers H, Gheorghiu L, Liu Q, *et al.* DNA Damage Response Assessments in Human Tumor Samples Provide Functional Biomarkers of Radiosensitivity. *Semin Radiat Oncol*. 2015;25(4):237-250. doi:10.1016/j.semradonc.2015.05.007
14. Bhide SA, Thway K, Lee J, *et al.* Delayed DNA double-strand break repair following platin-based chemotherapy predicts treatment response in head and neck squamous cell carcinoma. *Br J Cancer*. 2016;115(7):825-830. doi:10.1038/bjc.2016.266
15. Naipal KAT, Verkaik NS, Ameziane N, *et al.* Functional ex vivo assay to select homologous recombination-deficient breast tumors for PARP inhibitor treatment. *Clin Cancer Res*. 2014;20(18):4816-4826. doi:10.1158/1078-0432.CCR-14-0571

16. Verhagen CVM, Vossen DM, Borgmann K, *et al.* Fanconi anemia and homologous recombination gene variants are associated with functional DNA repair defects in vitro and poor outcome in patients with advanced head and neck squamous cell carcinoma. *Oncotarget*. 2018;9(26):18198-18213. doi:10.18632/oncotarget.24797
17. Stoepker C, Ameziane N, Van Der Lelij P, *et al.* Defects in the Fanconi anemia pathway and chromatid cohesion in head and neck cancer. *Cancer Res*. 2015;75(17):3543-3553. doi:10.1158/0008-5472.CAN-15-0528/651997/AM/DEFECTS-IN-THE-FANCONI-ANEMIA-PATHWAY-AND
18. Romick-Rosendale LE, Lui VWY, Grandis JR, Wells SI. The Fanconi anemia pathway: Repairing the link between DNA damage and squamous cell carcinoma. *Mutat Res - Fundam Mol Mech Mutagen*. 2013;743-744:78-88. doi:10.1016/j.mrfmmm.2013.01.001
19. Stinge J, Bellelli R, Boulton SJ. Mechanisms of DNA-protein crosslink repair. *Nat Rev Mol Cell Biol*. 2017;18(9):563-573. doi:10.1038/nrm.2017.56
20. Lopez-Martinez D, Liang C-C, Cohn MA. Cellular response to DNA interstrand crosslinks: the Fanconi anemia pathway. *Cell Mol Life Sci*. 2016;73(16):3097-3114. doi:10.1007/s00018-016-2218-x
21. Pommier Y, O'Connor MJ, De Bono J. Laying a trap to kill cancer cells: PARP inhibitors and their mechanisms of action. *Sci Transl Med*. 2016;8(362):1-8. doi:10.1126/scitranslmed.aaf9246
22. O'Connor MJ. Targeting the DNA Damage Response in Cancer. *Mol Cell*. 2015;60(4):547-560. doi:10.1016/j.molcel.2015.10.040
23. Chaudhuri AR, Callen E, Ding X, *et al.* Replication fork stability confers chemoresistance in BRCA-deficient cells. *Nature*. 2016;535(7612):382-387. doi:10.1038/nature18325
24. Chandrasekharappa SC, Chinn SB, Donovan FX, *et al.* Assessing the spectrum of germline variation in Fanconi anemia genes among patients with head and neck carcinoma before age 50. *Cancer*. 2017;123(20):3943-3954. doi:10.1002/cncr.30802
25. Bakhoum SF, Ngo B, Laughney AM, *et al.* Chromosomal instability drives metastasis through a cytosolic DNA response. *Nature*. 2018;553(7689):467-472. doi:10.1038/nature25432
26. Garnett MJ, Edelman EJ, Heidorn SJ, *et al.* Systematic identification of genomic markers of drug sensitivity in cancer cells. *Nature*. 2012;483(7391):570-575. doi:10.1038/nature11005
27. Peng G, Chun-Jen Lin C, Mo W, *et al.* Genome-wide transcriptome profiling of homologous recombination DNA repair. *Nat Commun*. 2014;5:1-21. doi:10.1038/ncomms4361
28. Balm AJM, Rasch CRN, Schornagel JH, *et al.* High-dose superselective intra-arterial cisplatin and concomitant radiation (radplat) for advanced head and neck cancer. *Head Neck*. 2004;26(6):485-493. doi:10.1002/hed.20006
29. Vossen DM, Verhagen CVM, Grénman R, *et al.* Role of variant allele fraction and rare SNP filtering to improve cellular DNA repair endpoint association. *PLoS One*. 2018;13(11):e0206632. doi:10.1371/journal.pone.0206632
30. Chen EY, Tan CM, Kou Y, *et al.* Enrichr: interactive and collaborative HTML5 gene list enrichment analysis tool. *BMC Bioinformatics*. 2013;14(1):128. doi:10.1186/1471-2105-14-128
31. Leemans CR, Braakhuis BJM, Brakenhoff RH. The molecular biology of head and neck cancer. *Nat Rev Cancer*. 2011;11(1):9-22. doi:10.1038/nrc2982
32. 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 Oto-Rhino-Laryngology*. 2017. doi:10.1007/s00405-017-4687-4
33. Leemans CR, Snijders PJF, Brakenhoff RH. The molecular landscape of head and neck cancer. *Nat Rev Cancer*. 2018;18(5):269-282. doi:10.1038/nrc.2018.11

34. van der Heijden M, de Jong MC, Verhagen CVM, *et al.* Acute Hypoxia Profile is a Stronger Prognostic Factor than Chronic Hypoxia in Advanced Stage Head and Neck Cancer Patients. *Cancers (Basel)*. 2019. doi:10.3390/cancers11040583
35. Justus CR, Leffler N, Ruiz-Echevarria M, Yang L V. In vitro cell migration and invasion assays. *J Vis Exp*. 2014;(88):1-8. doi:10.3791/51046
36. Chen W, Ebelt ND, Stracker TH, Xhemalce B, Van Den Berg CL, Miller KM. ATM regulation of IL-8 links oxidative stress to cancer cell migration and invasion. *Elife*. 2015;4:1-21. doi:10.7554/eLife.07270
37. Goodwin JF, Knudsen KE. Beyond DNA repair: DNA-PK function in cancer. *Cancer Discov*. 2014;4(10):1126-1139. doi:10.1158/2159-8290.CD-14-0358
38. Kotula E, Berthault N, Agrario C, *et al.* DNA-PKcs plays role in cancer metastasis through regulation of secreted proteins involved in migration and invasion. *Cell Cycle*. 2015;14(12):1961-1972. doi:10.1080/15384101.2015.1026522
39. Lamouille S, Xu J, Derynck R. Molecular mechanisms of epithelial–mesenchymal transition. *Nat Rev Mol Cell Biol*. 2014;15(3):178-196. doi:10.1038/nrm3758
40. Koo V, El Mekabaty A, Hamilton P, *et al.* Novel in vitro assays for the characterization of EMT in tumorigenesis. *Cell Oncol*. 2010;32(1-2):67-76. doi:10.3233/CLO-2009-0501
41. Gröger CJ, Grubinger M, Waldhör T, Vierlinger K, Mikulits W. Meta-Analysis of Gene Expression Signatures Defining the Epithelial to Mesenchymal Transition during Cancer Progression. *PLoS One*. 2012;7(12):1-10. doi:10.1371/journal.pone.0051136
42. Byers LA, Diao L, Wang J, *et al.* An epithelial-mesenchymal transition gene signature predicts resistance to EGFR and PI3K inhibitors and identifies Axl as a therapeutic target for overcoming EGFR inhibitor resistance. *Clin Cancer Res*. 2013;19(1):279-290. doi:10.1158/1078-0432.CCR-12-1558
43. Zhao S, Venkatasubbarao K, Lazor JW, *et al.* Inhibition of STAT3Tyr705 phosphorylation by Smad4 suppresses transforming growth factor β -mediated invasion and metastasis in pancreatic cancer cells. *Cancer Res*. 2008;68(11):4221-4228. doi:10.1158/0008-5472.CAN-07-5123
44. Vincent T, Neve EPA, Johnson JR, *et al.* A SNAIL1-SMAD3/4 transcriptional repressor complex promotes TGF- β mediated epithelial-mesenchymal transition. *Nat Cell Biol*. 2009;11(8):943-950. doi:10.1038/ncb1905
45. White RA, Neiman JM, Reddi A, *et al.* Epithelial stem cell mutations that promote squamous cell carcinoma metastasis. *J Clin Invest*. 2013;123(10):4390-4404. doi:10.1172/JCI65856
46. Park JW, Jang SH, Park DM, *et al.* Cooperativity of E-cadherin and Smad4 Loss to Promote Diffuse-Type Gastric Adenocarcinoma and Metastasis. *Mol Cancer Res*. 2014;12(8):1088-1099. doi:10.1158/1541-7786.mcr-14-0192-t
47. Sakai W, Swisher EM, Karlan BY, *et al.* Secondary mutations as a mechanism of cisplatin resistance in BRCA2-mutated cancers. *Nature*. 2008;451(7182):1116-1120. doi:10.1038/nature06633
48. Bristow RG, Hill RP. Hypoxia and metabolism: Hypoxia, DNA repair and genetic instability. *Nat Rev Cancer*. 2008. doi:10.1038/nrc2344
49. Chan N, Pires IM, Bencokova Z, *et al.* Contextual synthetic lethality of cancer cell kill based on the tumor microenvironment. *Cancer Res*. 2010;70(20):8045-8054. doi:10.1158/0008-5472.CAN-10-2352
50. Rankin EB, Giaccia AJ. Hypoxic control of metastasis. *Science*. 2016;352(6282):175-180. doi:10.1126/science.aaf4405

CTTTGGGAGGCGGAGGCGGGGGGATTACTTTGAGGATAGGAGTTCAGACAGGTGGCCAACTGGTGAAATCCCGTCTCTACTAAAAATACAAAAATTAGCTGGGCTGGTGGGTGCTGTATCCAGCTATTGGGAGGTGAGGAG

51. Bindra RS, Schaffer PJ, Meng A, *et al.* Down-Regulation of Rad51 and Decreased Homologous Recombination in Hypoxic Cancer Cells Down-Regulation of Rad51 and Decreased Homologous Recombination in Hypoxic Cancer Cells. *Mol Cell Biol.* 2004;24(19):8504-8518. doi:10.1128/MCB.24.19.8504
52. Vollebergh MA, Lips EH, Nederlof PM, *et al.* Genomic patterns resembling BRCA1- and BRCA2-mutated breast cancers predict benefit of intensified carboplatin-based chemotherapy. *Breast Cancer Res.* 2014;16(3):1-13. doi:10.1186/bcr3655
53. Wang ZC, Birkbak NJ, Culhane AC, *et al.* Profiles of genomic instability in high-grade serous ovarian cancer predict treatment outcome. *Clin Cancer Res.* 2012;18(20):5806-5815. doi:10.1158/1078-0432.CCR-12-0857
54. Kang J, D'Andrea AD, Kozono D. A DNA repair pathway-focused score for prediction of outcomes in ovarian cancer treated with platinum-based chemotherapy. *J Natl Cancer Inst.* 2012;104(9):670-681. doi:10.1093/jnci/djs177
55. Pitroda SP, Pashtan IM, Logan HL, *et al.* DNA Repair Pathway Gene Expression Score Correlates with Repair Proficiency and Tumor Sensitivity to Chemotherapy. *Sci Transl Med.* 2014;6(229):229ra42-229ra42. doi:10.1126/scitranslmed.3008291
56. Romick-Rosendale LE, Hoskins EE, Privette Vinnedge LM, *et al.* Defects in the fanconi anemia pathway in head and neck cancer cells stimulate tumor cell invasion through DNA-PK and Rac1 signaling. *Clin Cancer Res.* 2016;22(8):2062-2073. doi:10.1158/1078-0432.CCR-15-2209
57. Kurrey NK, Jalgaonkar SP, Joglekar A V., *et al.* Snail and slug mediate radioresistance and chemoresistance by antagonizing p53-mediated apoptosis and acquiring a stem-like phenotype in ovarian cancer cells. *Stem Cells.* 2009;27(9):2059-2068. doi:10.1002/stem.154
58. De Jong MC, Ten Hoeve JJ, Grénman R, *et al.* Pretreatment microRNA expression impacting on epithelial-to-mesenchymal transition predicts intrinsic radiosensitivity in head and neck cancer cell lines and patients. *Clin Cancer Res.* 2015;21(24):5630-5638. doi:10.1158/1078-0432.CCR-15-0454
59. Nieto MA, Huang RYYJ, Jackson RAA, Thiery JPP. Emt: 2016. *Cell.* 2016;166(1):21-45. doi:10.1016/j.cell.2016.06.028

Supplementary Information

Supplementary figures

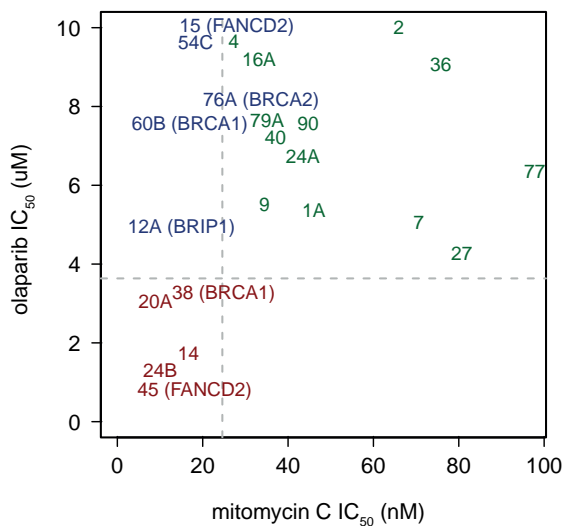


Figure S1. Genetic alterations in the cell line panel.

A) Drug sensitivity plot with UT-SCC cell line numbers at their approximate olaparib and MMC sensitivities. These cell lines were previously tested for genetic alterations in DNA repair genes and found to harbor homozygous, rare and potentially disruptive variants in the indicated genes¹.

¹ Verhagen CVM, Vossen DM, Borgmann K, Hageman F, Grénman R, Verwijs-Janssen M *et al.* Fanconi anemia and homologous recombination gene variants are associated with functional DNA repair defects in vitro and poor outcome in patients with advanced head and neck squamous cell carcinoma. *Oncotarget* 2018;

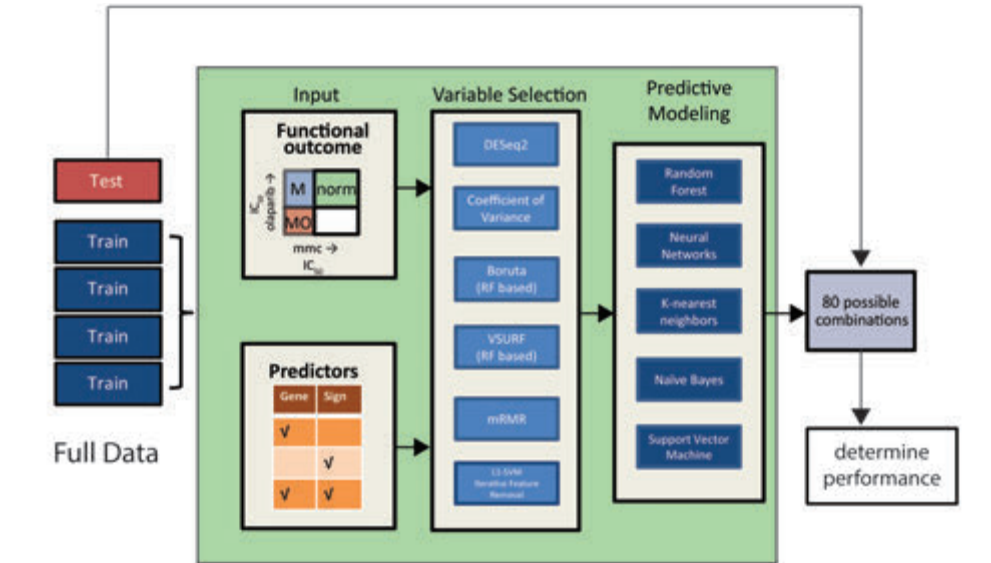


Figure S2. Strategy for internal cross validation of prediction performance.

We selected a broad range of algorithms, to ensure that our conclusion is independent of the method used. K-nearest neighbour (knn) and support vector machines (svm) are considered to work well with low sample numbers, although knn cannot handle high dimensional data. Likewise, the naïve Bayes (nb) algorithm works with low numbers of samples, and because each variable is assumed to be independent, is considered not to suffer from high dimensionality. Random forests have been reported to work well with high dimensionality datasets. Neural networks were also found to work well with small sample sizes, only if the groups are perfectly separable, although they do suffer from high dimensionality in the data. To reduce the dimensionality of the dataset, we employed several selection techniques. Both Boruta² and VSURF³ select variables based on random forest predictions. VSURF selects variables that have consistently high variable importance across multiple iterations of training. Boruta generates permuted versions of variables and compares their importance to the original version, selecting those variables where permutation causes a drop in importance. Minimum redundancy, maximum relevance (mRMR) aims to find a minimal set of genes that are correlated to the outcome parameter but not collinear. Multiple runs of each of these algorithms tend not to identify the same genes repeatedly (data not shown). In contrast, differentially expression analysis, using DESeq2 package, results in the same set of genes each time. We re-implemented the sparse SVM iterative feature removal (sSVM-IFR) technique described by ref.⁴ in R. In this method, a penalized SVM is trained repeatedly on the full dataset and after each iteration the selected features are recorded and removed from the dataset. This process is repeated until the performance drops below the threshold, at which point the full set of features that is relevant to the outcome variable has been identified. The full dataset was either split into 5 folds, 10 times, resulting in 50 different splits of the data, or into 3 folds, 16 times, resulting in 48 splits. Five and three-fold splits of the data were chosen so that the olaparib sensitive group (n=5) would be either divided 4-1 or 3-2 over the training and test splits. Variable selection and model fitting was performed on the training folds only. With this strategy, all considered variable selection and predictive modelling techniques were trained and evaluated on the same data splits.

² Kursa MB, Rudnicki WR. Feature Selection with the Boruta Package. J Stat Softw 2010; 36

³ Genuer R, Poggi J-M, Tuleau-Malot CT-M. VSURF: An R Package for Variable Selection Using Random Forests. R J 2015; 7.:2: 19–33

⁴ O’Hara S, Wang K, Slayden RA, Schenkel AR, Huber G, O’Hern CS *et al.* Iterative feature removal yields highly discriminative pathways. BMC Genomics 2013; 14.

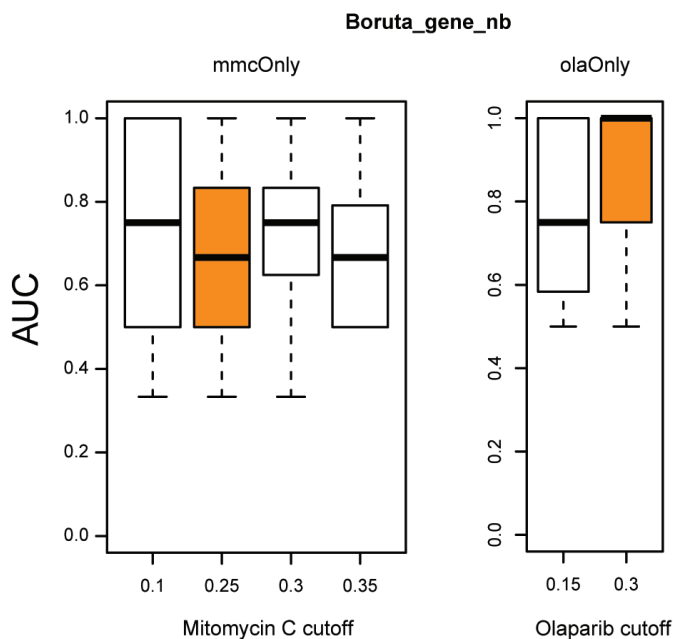


Figure S3. Cross validation at different drug sensitivity cutoffs.

To assess performance at different cutoffs for mitomycin C and olaparib sensitivity, we chose the overall best-performing combination in the mmcOnly and olaOnly models, Boruta variable selection coupled to a naïve Bayes classifier. The orange boxes represent the cut offs used throughout the study. The performance is not highly cutoff dependent. From this we conclude that widening or narrowing repair defect cell line classification to generate the models did not have a large impact, thereby also further showing that our models capture general and group defined features rather than highly individual cell line based features.

CTTTGGGAGGCGGAGGCGGGGGATTACTTTGAGGATAGGAGTCTCCAGACAGCGTGGCCCAAGTGGTGAATCCCGCTCTACTAAAAAATACAAAAATTAGCTGGCGTGGTGGGTGCTGTGAATCCCGAGCTATTGGGAGGGTGAGGAGAG

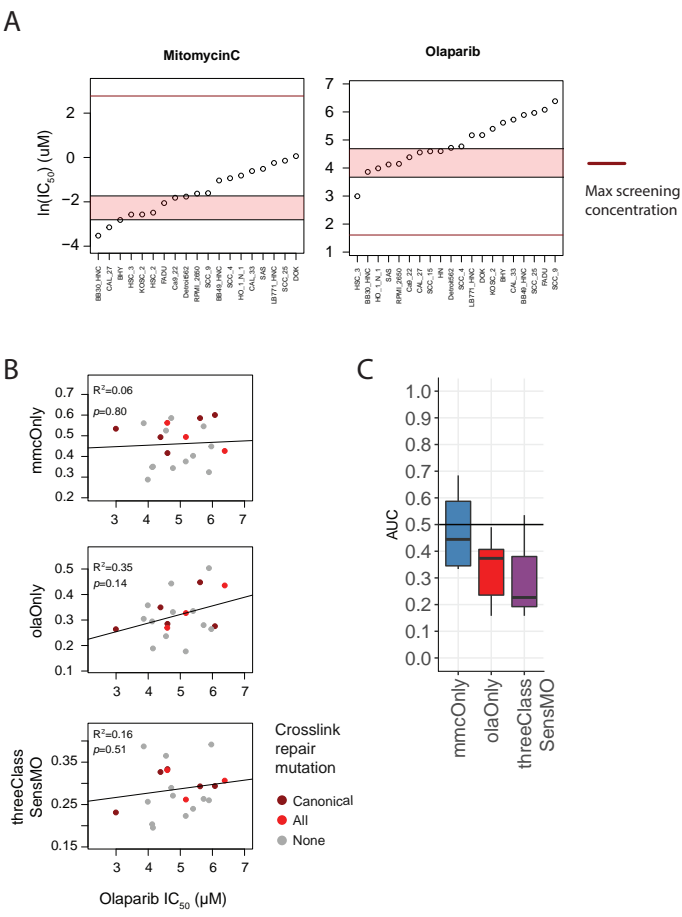


Figure S4. Mitomycin C and olaparib sensitivity in the CGP dataset.
A) IC_{50} values for cell lines tested in the CGP dataset. Pink boxes denote the range of concentrations that was used to split the cell lines into the sensitive groups in Fig 2B. Note that all reported olaparib IC_{50} values are higher than the maximum screening concentration (dark red line) and thus computationally inferred. CGP IC_{50} values were determined at early time points (72h) and independent of cell line doubling times (DT). This is in contrast to those in our study (5xDT with DTs ranging from 10 to 60 hours). Furthermore, the CGP chosen metabolic readout has suboptimal association with survival, particularly for inhibition of PARP, which regulates metabolic activity⁵. As a result, the CGP IC_{50} values cannot match ours. **B)** Correlation of model ensemble scores with olaparib sensitivity in HNSCC cell lines. **C)** AUC for distinguishing DNA repair defects based on olaparib sensitivity. Rather than applying an arbitrary MMC or olaparib hypersensitivity assignment, we tested the performance at multiple cutoffs for repair defect definition, between 20 and 50% of the range (pink box in (a)) and report the median performance over this range. Overall, the performance of the olaparib sensitivity predicting models cannot be confirmed or invalidated due to the computationally inferred olaparib IC_{50} values that were not experimentally tested.

⁵ Fouquerel E, Goellner EM, Yu Z, Gagné J-P, Barbi de Moura M, Feinstein T *et al.* ARTD1/PARP1 negatively regulates glycolysis by inhibiting hexokinase 1 independent of NAD⁺ depletion. Cell Rep 2014; 8: 1819–1831.

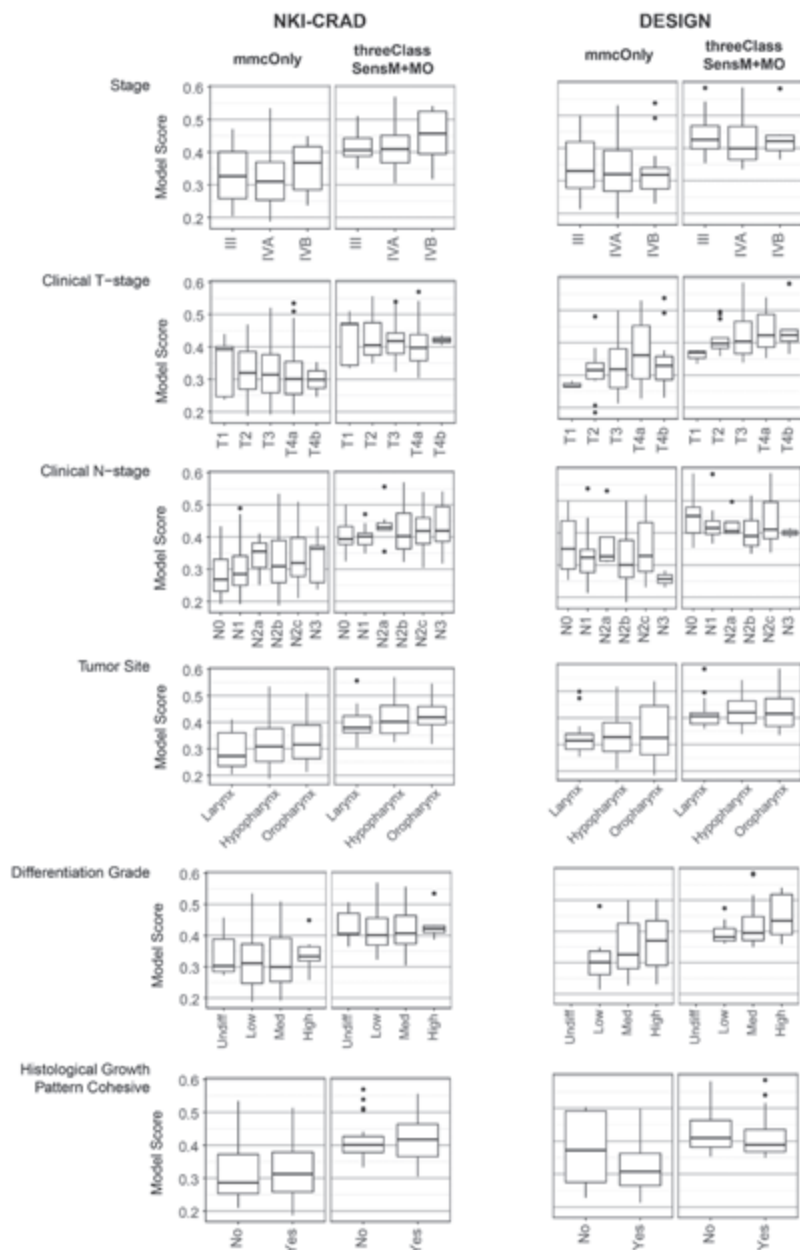


Figure S5. Relationship between model predictions and clinical parameters. The clinical impact of predicted DNA repair defects could have been a result of an association with known clinical parameters, such as lymph node infiltration (clinical N-stage). We therefore tested the associations between DNA repair defect probabilities and clinical variables. Model scores distribution was plotted in relation to multiple clinical variables. To assess statistical significance, pairwise student's *t*-tests were used and reported in Table S5.

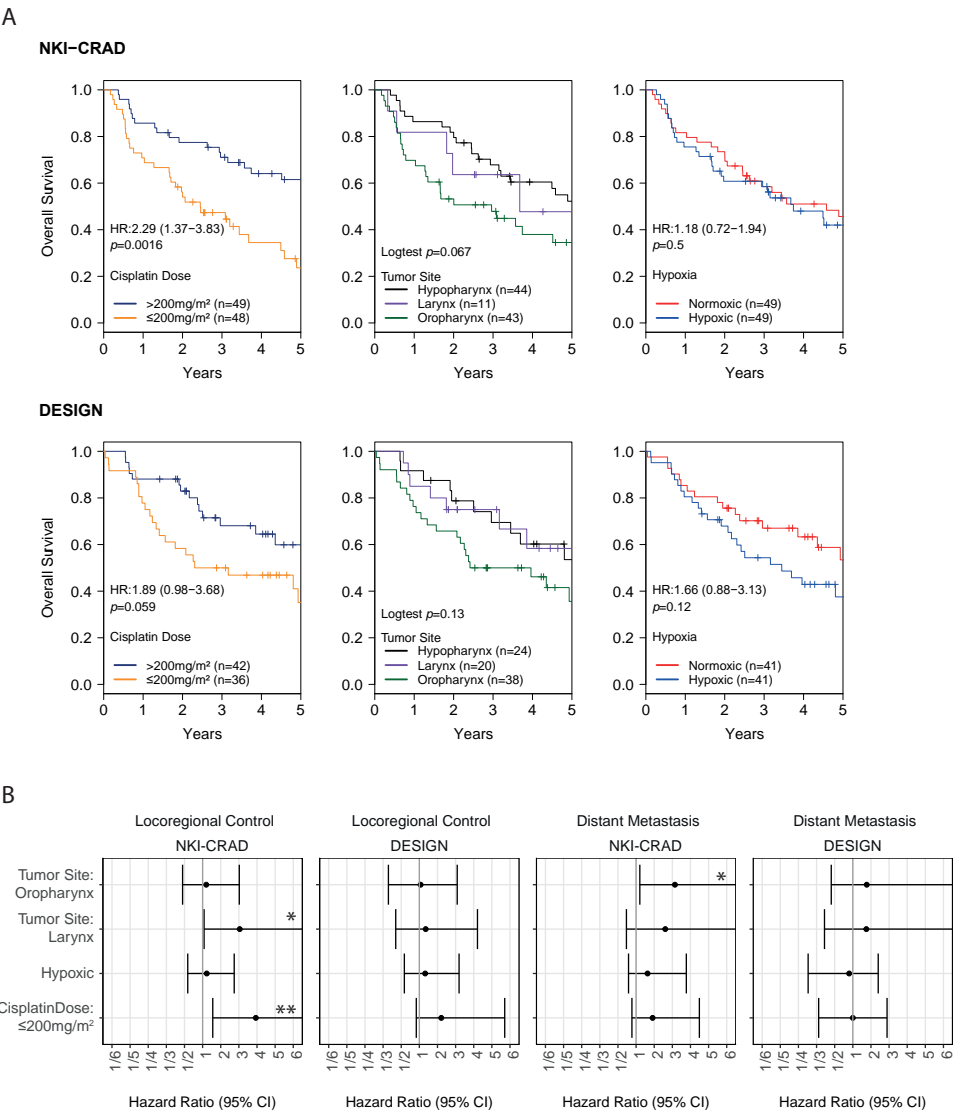


Figure S6. Univariate analysis of variables used in multi-variate analyses.
A) Kaplan-Meier plots of univariate analysis showing the effect of cumulative cisplatin dose, tumor site and hypoxia status on overall survival in both cohorts. **B)** Forest plots with univariate analysis values of the same variables as in A, for locoregional control and distant metastasis.

Drug sensitivity prediction models reveal a link between DNA repair defects and poor prognosis in HNSCC

CTTTGGGAGGCGGAGGCGGGGGGATCTTCTGAGGATAGGATCTCTAGACAGGCTGGGCAACGTGGTGAATCCCGCTCTACTAAATAATCAAAATATAGCTGGGCTGGTGGGTGCTGTATCCGAGCTATTGGGAGGGTGAGGAGCT

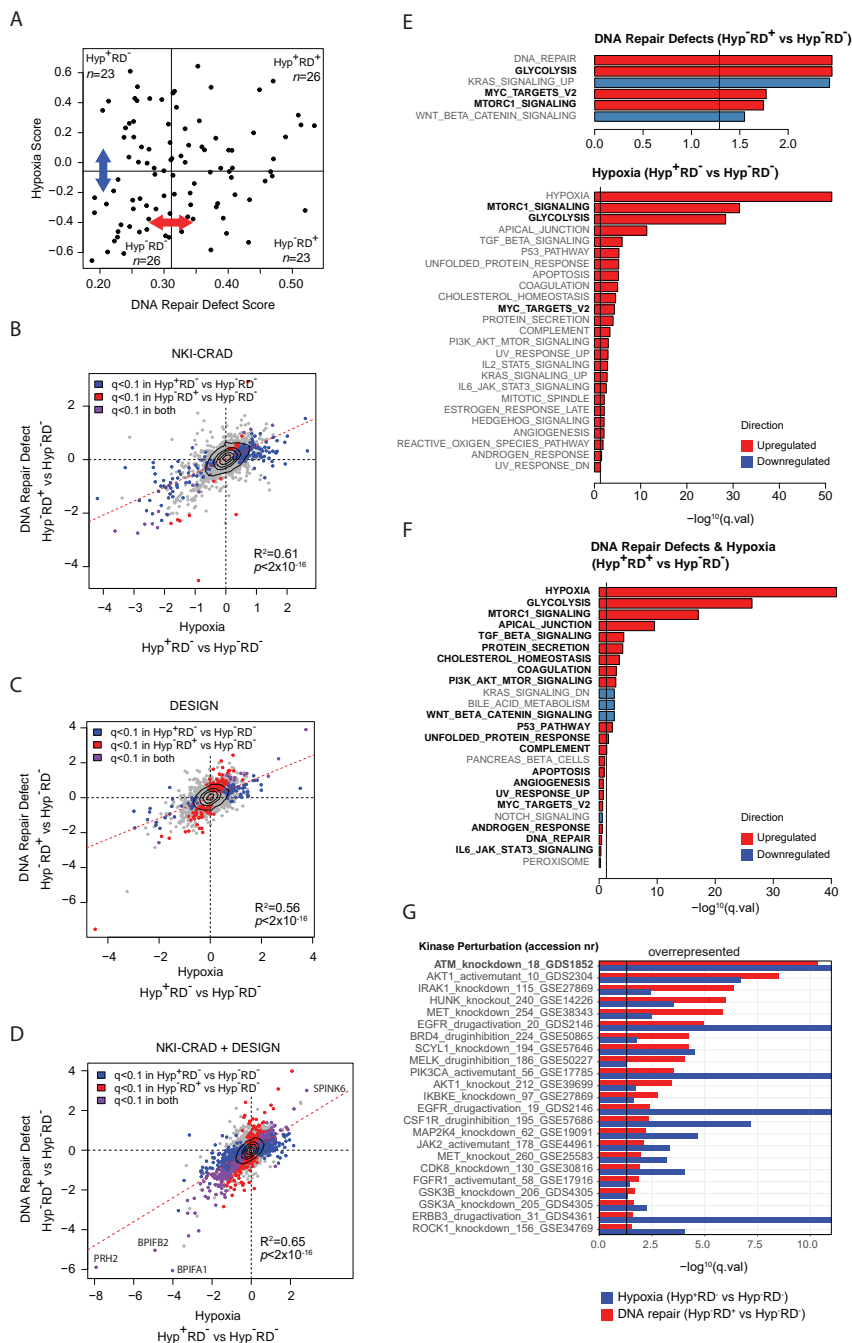


Figure S7 DNA repair defects and hypoxia share downstream effector pathways.

A) Patients in the NKI-CRAD cohort were divided into equal-sized groups, based on both their score for DNA repair defects (mmcOnly model) and for hypoxia: Normoxic (non-hypoxic) and DNA Repair

Figure S7 Continued

Proficient (low repair defect score) (Hyp^{RD+}), Hypoxic and Repair Proficient (Hyp^{RD+}), Normoxic and Repair Defected (Hyp^{RD+}) and Hypoxic and Repair Defected (Hyp^{RD+}). **B-C**) Differential gene expression associated with DNA repair defects (vs repair proficient within the normoxic tumor group, red arrow in A) were compared to those associated with hypoxia (hypoxic vs normoxic in the repair proficient tumor group, blue arrow in A). The log2 fold changes per gene are shown on the axis. Transcriptional changes associated with hypoxia and DNA repair defects are highly correlated ($p < 2 \times 10^{-16}$) in the NKI (B) and DESIGN cohort (C). **D**) Combining both cohorts strongly increased statistical power, exemplified by the increased number of differentially expressed genes that reach significance. Among the genes that were significantly upregulated by hypoxia and a repair defect, the most up-regulated gene is SPINK6. SPINK6 is a serine peptidase inhibitor, a secreted potent inhibitor of epidermal proteases and has been implicated in poor prognosis and metastasis in HNSCC⁶. **E**) Gene set enrichment analyses (GSEA) were performed for both comparisons in the combined cohorts, using the MSigDB Hallmarks genesets (DOI: 10.1093/bioinformatics/btr260). The analyses also pointed to commonalities in the biological processes regulated by hypoxia or DNA repair defects. Shared hallmarks are indicated in bold. The horizontal line indicates $q = 0.05$. **F**) GSEA was also performed for the hypoxic and DNA repair defective tumor group (Hyp^{RD+} vs Hyp^{RD-} in A), to identify any additional regulated pathways unique to this group. Most pathways are also found in the comparisons shown in E (indicated in bold lettering) confirming the similar transcriptional response pattern. **G**) In order to get more insight into the shared pathways, GSEA was performed for both comparisons, using the “Kinase_Perturbations_from_GEO_up” geneset library from Enrichr. Genesets that were significantly up regulated in both comparisons are shown and contain genes that were found to be upregulated after the stated perturbation was performed. The most prominent result is the geneset “ATM knockdown” further pointing to a similar transcriptional response from hypoxia or repair defect in HNSCC tumors.

⁶ Zheng LS, Yang JP, Cao Y, Peng LX, Sun R, Xie P *et al*. SPINK6 Promotes Metastasis of Nasopharyngeal Carcinoma via Binding and Activation of Epithelial Growth Factor Receptor. *Cancer Res*. 2017; 77.

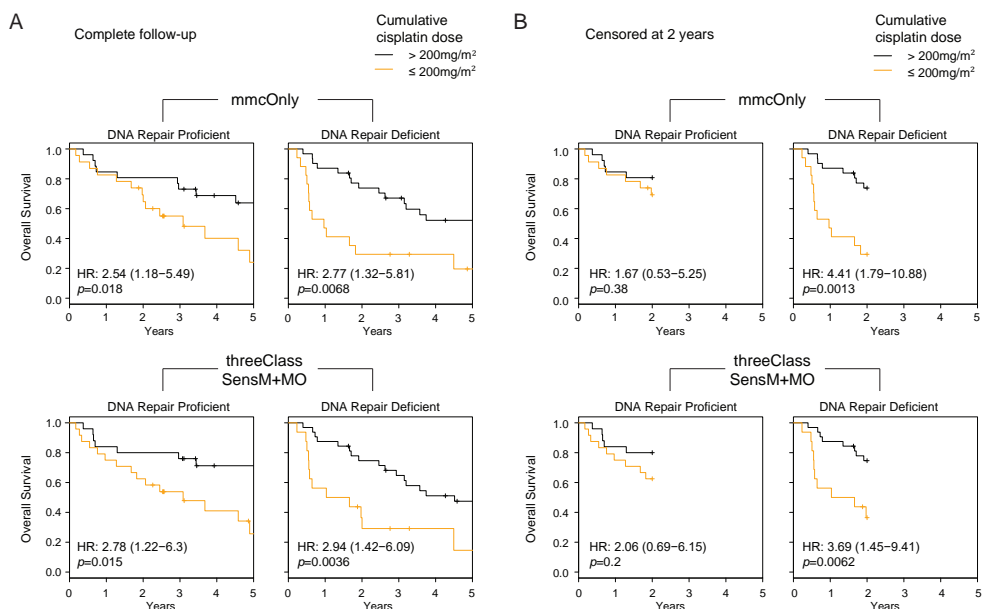


Figure S8. Relation of DNA repair defects and cisplatin treatment.

Patients were divided into DNA repair proficient and deficient groups, based on either the mmcOnly or the threeClass (*SensM+SensMO*) model ensemble. Within these groups, univariate HR for cisplatin dose were calculated using the complete follow-up available (in A) or using a maximum of 2 years follow-up (in B). Despite losing the benefit of a stronger cisplatin treatment, patients that ultimately received lower cisplatin doses tend to be also less fit. This may therefore affect long term overall survival data. Although this effect is not maintained over the full follow-up time, the data show that the low cisplatin dose patients in the DNA repair deficient tumour group have a larger and more significant HR in the first 2 years.

CTTTGGGAGGCGGAGGCGGGGGATTACTTGGAGATAGAGTCTCCAGACAGCGTGCGCCAACTGGTGAAATCCCGCTCTACTAAAAAATACAAAAATTAGCTGGCGTGCGTGGTGGCTGTGTATCCAGCTATTGGGAGGTTGAGGAG

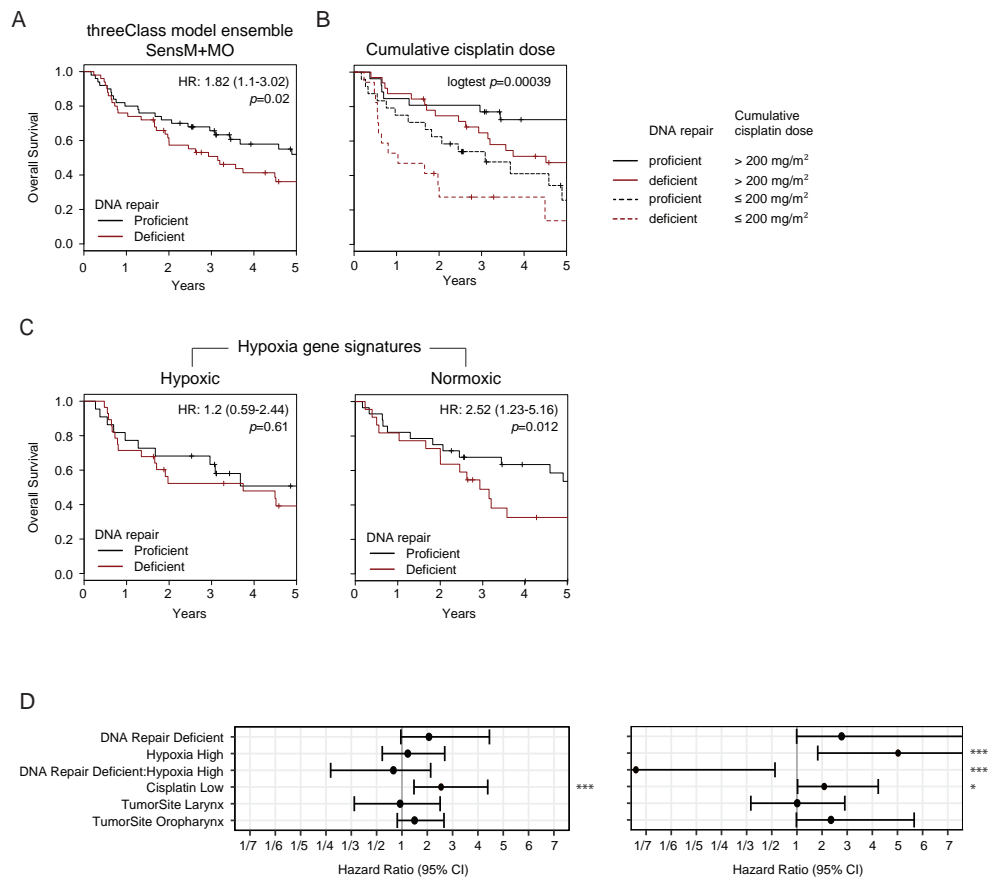


Figure S9. Prognostic value of the threeClass model ensemble.

A) Patients were divided into two equal-sized groups based on their predicted probability by the threeClass (*SensM+SensMO*) model ensemble. Univariate Cox regression hazard ratios (HR) and associated p -value is reported. **B)** Patients were divided into a group with hypoxic and normoxic tumours, based on the average expression of five published gene signatures for hypoxia. Univariate HR and associated p -values are shown. **C)** Patients were divided by the cumulative cisplatin dose received during treatment. **D)** Forest plot for multivariate Cox analysis of the threeClass (*SensM+SensMO*) model ensemble prediction that also include hypoxia, cisplatin dose and tumor site for the NKI-CRAD and the DESIGN cohorts. * $p<0.05$, ** $p<0.01$, *** $p<0.001$.

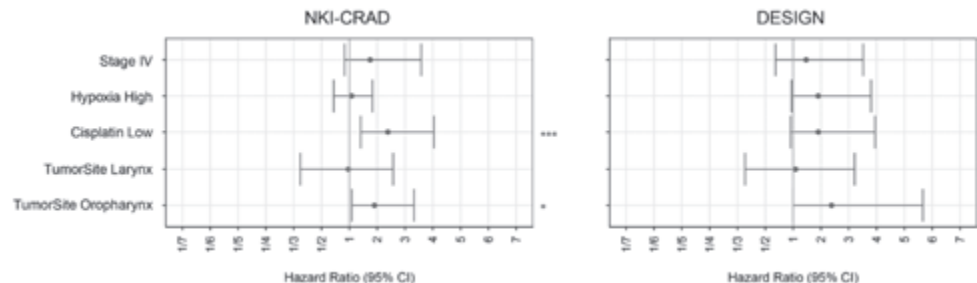
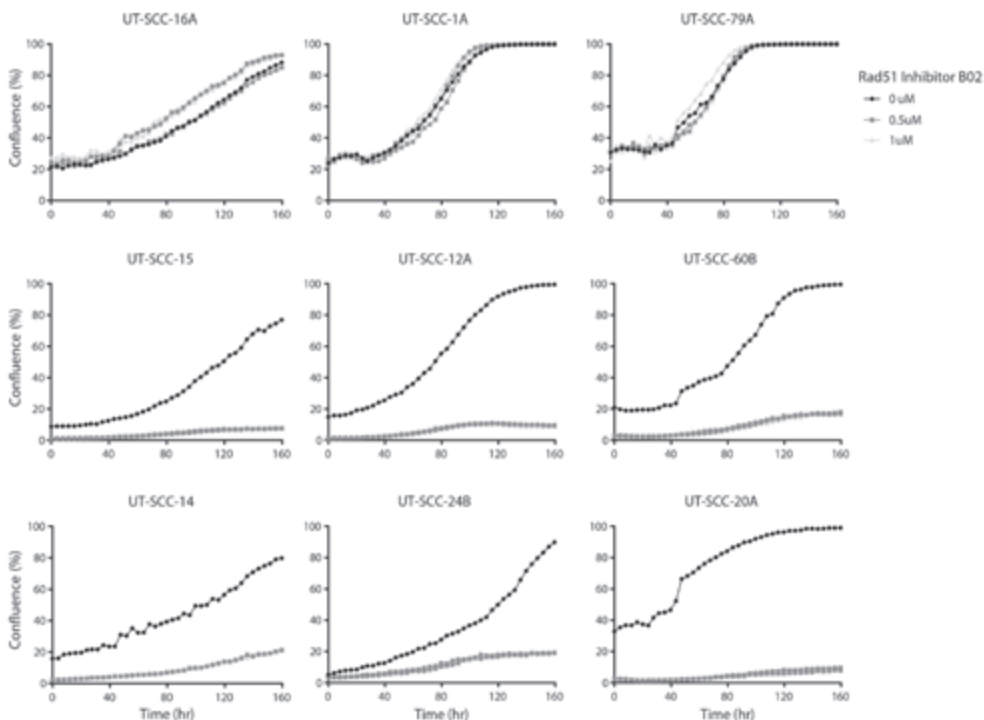


Figure S10. Multivariate analysis using disease stage as predictor. Clinical stage does not have the same prognostic value as DNA repair defects (Fig 4A and Fig S6) in both cohorts.

4



FigureS11. Rad51 inhibitor induced growth inhibition in DNA repair defective HNSCC cell lines. Cells were seeded in a 48-well plate (Corning, 353230) at 20% confluence. Culture medium was added with or without RAD51 inhibitor as indicated (B02, SML0364, SIGMA) and confluence was determined using the IncuCyte (Essen BioScience, Ltd., Hertfordshire, United Kingdom).

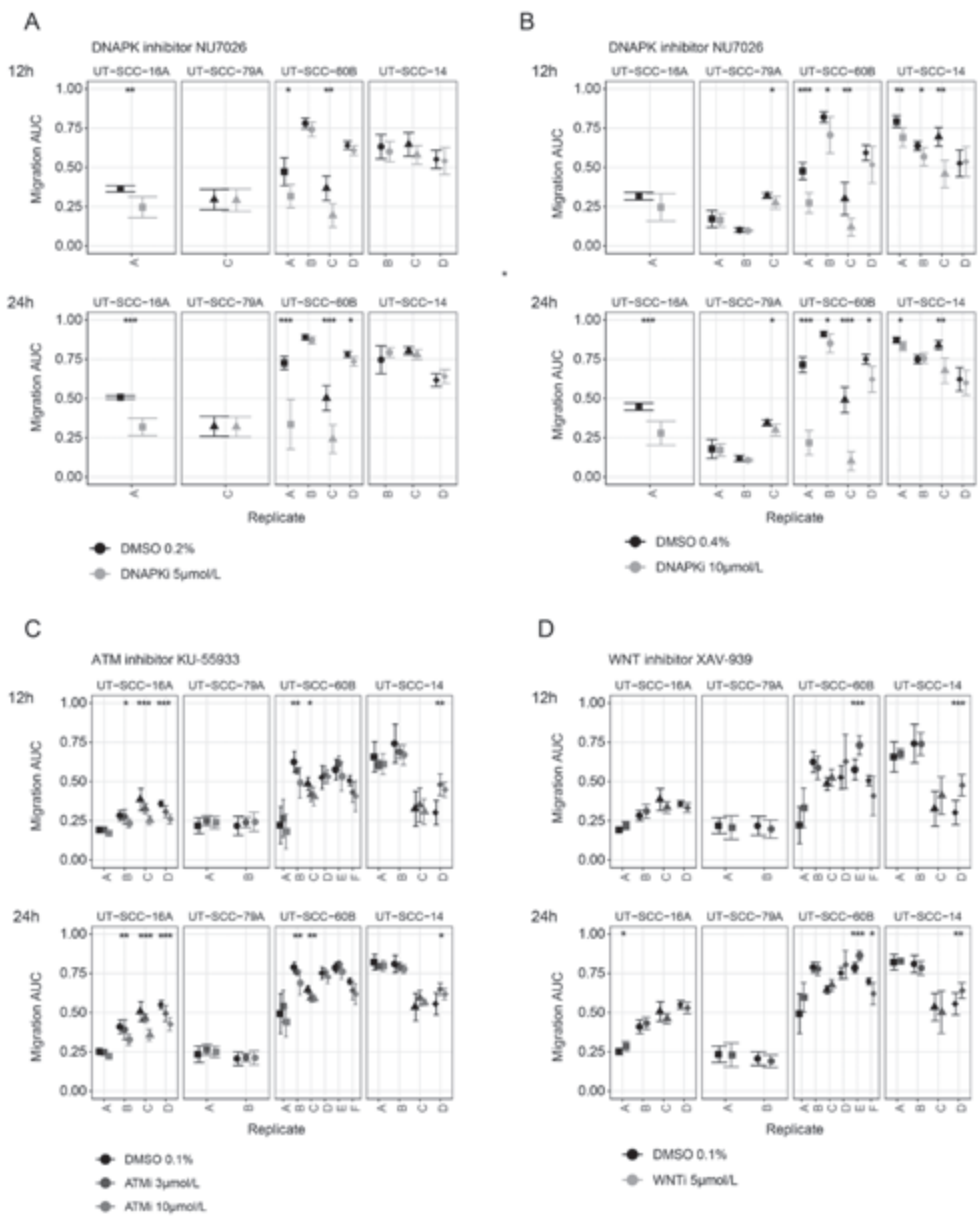


Figure S12. Inhibition of migration in scratch assays.
A-D) Areas under the curve (AUC) at 12 hours and 24 hours in scratch assays. This data was used to compute inhibition values shown Fig 7B and C. Averages of individual experiments are shown with standard deviation of 4-6 technical replicates each. Experiments were performed with DNA-PK (AB), ATM (C) or WNT (D) inhibitor at the concentrations indicated.



A) Representative light microscopy images of each cell line used in Fig 4. **B)** GSEA analysis with selected EMT process gene sets comparing the migrating to non-migrating cells as in A. **C)** Expression of mesenchymal markers and E-Cadherin (CDH1) in the analysed cell line set.

Supplementary tables

Table S1. List of Enrichr Signatures collections. GSVA scores for signatures in these collections were used as input to those machine learning models with “_sign” or “_both” in their name.

BioCarta_2016
ChEA_2015
Chromosome_Location
CORUM
ENCODE_and_ChEA_Consensus_TFs_from_ChIP-X
ENCODE_TF_ChIP-seq_2015
GO_Biological_Process_2015
GO_Cellular_Component_2015
GO_Molecular_Function_2015
Humancyc_2016
KEA_2015
KEGG_2016
Kinase_Perturbations_from_GEO_up
Ligand_Perturbations_from_GEO_up
Ligand_Perturbations_from_GEO_down
Reactome_2016
Single_Gene_Perturbations_from_GEO_up
Single_Gene_Perturbations_from_GEO_down
TargetScan_microRNA
TF-LOF_Expression_from_GEO
WikiPathways_2016

Table S2. Cross validation performance. Mean performance for all models tested: threeClass, olaOnly and mmcOnly, in 5-fold and 3-fold cross validation. Combinations used in the ensembles are highlighted in green

Type	Input	FeatureSelection	Classifier	AUC_3Fold	AUC_5Fold	UsedInEnsemble
mmcOnly	gene	coeffVar	rf	0.74	0.76	Yes
mmcOnly	gene	Boruta	nb	0.65	0.74	Yes
mmcOnly	gene	Boruta	nnet	0.68	0.73	Yes
mmcOnly	gene+sign	coeffVar	svmRadial	0.65	0.71	Yes
mmcOnly	gene+sign	coeffVar	rf	0.68	0.69	Yes
mmcOnly	gene	coeffVar	knn	0.64	0.67	Yes
mmcOnly	gene+sign	Boruta	nb	0.68	0.66	Yes
mmcOnly	sign	Boruta	nb	0.65	0.65	Yes
mmcOnly	sign	Boruta	svmRadial	0.66	0.63	Yes
mmcOnly	gene+sign	Boruta	svmRadial	0.64	0.76	No
mmcOnly	gene	Boruta	knn	0.63	0.75	No
mmcOnly	gene+sign	sSVM-IFR	nb	0.68	0.75	No
mmcOnly	gene+sign	mRMR	svmRadial	0.65	0.74	No
mmcOnly	gene+sign	mRMR	nnet	0.65	0.73	No
mmcOnly	gene	DESeq2	svmRadial	0.68	0.73	No
mmcOnly	gene	coeffVar	svmRadial	0.62	0.73	No
mmcOnly	sign	mRMR	nb	0.69	0.73	No
mmcOnly	gene	DESeq2	rf	0.68	0.73	No
mmcOnly	gene	Boruta	svmRadial	0.65	0.72	No
mmcOnly	gene+sign	Boruta	knn	0.63	0.72	No
mmcOnly	gene+sign	mRMR	knn	0.61	0.72	No
mmcOnly	gene	DESeq2	nnet	0.65	0.72	No
mmcOnly	gene	DESeq2	knn	0.66	0.72	No
mmcOnly	gene	VSURF	rf	0.69	0.72	No
mmcOnly	gene	VSURF	nnet	0.66	0.72	No
mmcOnly	gene+sign	mRMR	rf	0.64	0.72	No
mmcOnly	gene	VSURF	svmRadial	0.70	0.71	No
mmcOnly	gene	Boruta	rf	0.65	0.71	No
mmcOnly	gene	VSURF	knn	0.66	0.71	No
mmcOnly	gene	VSURF	nb	0.63	0.71	No
mmcOnly	sign	mRMR	svmRadial	0.64	0.71	No
mmcOnly	gene+sign	Boruta	nnet	0.63	0.71	No
mmcOnly	gene	mRMR	svmRadial	0.66	0.71	No
mmcOnly	gene+sign	VSURF	rf	0.68	0.70	No
mmcOnly	gene	mRMR	nb	0.65	0.70	No
mmcOnly	gene+sign	sSVM-IFR	nnet	0.64	0.70	No
mmcOnly	gene+sign	Boruta	rf	0.64	0.70	No

Table S2. Continued

Type	Input	FeatureSelection	Classifier	AUC_3Fold	AUC_5Fold	UsedInEnsemble
mmcOnly	gene	sSVM-IFR	knn	0.62	0.70	No
mmcOnly	gene	sSVM-IFR	nnet	0.60	0.70	No
mmcOnly	gene	mRMR	rf	0.65	0.70	No
mmcOnly	gene+sign	VSURF	nb	0.62	0.70	No
mmcOnly	gene+sign	mRMR	nb	0.64	0.70	No
mmcOnly	sign	coeffVar	knn	0.64	0.70	No
mmcOnly	sign	mRMR	nnet	0.65	0.69	No
mmcOnly	gene+sign	VSURF	svmRadial	0.65	0.69	No
mmcOnly	sign	VSURF	knn	0.62	0.69	No
mmcOnly	gene	sSVM-IFR	svmRadial	0.61	0.69	No
mmcOnly	gene	mRMR	knn	0.63	0.69	No
mmcOnly	gene+sign	sSVM-IFR	rf	0.70	0.69	No
mmcOnly	gene+sign	VSURF	nnet	0.65	0.69	No
mmcOnly	sign	Boruta	nnet	0.66	0.68	No
mmcOnly	gene	sSVM-IFR	rf	0.65	0.68	No
mmcOnly	gene	DESeq2	nb	0.67	0.68	No
mmcOnly	sign	VSURF	svmRadial	0.62	0.67	No
mmcOnly	gene+sign	sSVM-IFR	svmRadial	0.63	0.67	No
mmcOnly	sign	VSURF	rf	0.65	0.67	No
mmcOnly	gene+sign	sSVM-IFR	knn	0.64	0.67	No
mmcOnly	gene+sign	VSURF	knn	0.66	0.67	No
mmcOnly	gene+sign	coeffVar	knn	0.62	0.67	No
mmcOnly	sign	mRMR	rf	0.67	0.66	No
mmcOnly	gene	sSVM-IFR	nb	0.65	0.66	No
mmcOnly	sign	Boruta	rf	0.65	0.66	No
mmcOnly	sign	VSURF	nb	0.65	0.66	No
mmcOnly	gene	mRMR	nnet	0.68	0.65	No
mmcOnly	sign	Boruta	knn	0.64	0.65	No
mmcOnly	sign	mRMR	knn	0.63	0.65	No
mmcOnly	sign	sSVM-IFR	svmRadial	0.62	0.64	No
mmcOnly	sign	coeffVar	svmRadial	0.64	0.64	No
mmcOnly	sign	coeffVar	rf	0.65	0.64	No
mmcOnly	sign	sSVM-IFR	knn	0.64	0.64	No
mmcOnly	sign	VSURF	nnet	0.64	0.64	No
mmcOnly	sign	sSVM-IFR	nnet	0.63	0.64	No
mmcOnly	sign	sSVM-IFR	rf	0.65	0.63	No
mmcOnly	sign	sSVM-IFR	nb	0.62	0.63	No
mmcOnly	gene+sign	coeffVar	nb			No
mmcOnly	gene+sign	coeffVar	nnet			No

Table S2. Continued

Type	Input	FeatureSelection	Classifier	AUC_3Fold	AUC_5Fold	UsedInEnsemble
mmcOnly	gene	coeffVar	nb			No
mmcOnly	gene	coeffVar	nnet			No
mmcOnly	sign	coeffVar	nb			No
mmcOnly	sign	coeffVar	nnet			No
olaOnly	sign	sSVM-IFR	svmRadial	0.82	0.90	Yes
olaOnly	sign	sSVM-IFR	nnet	0.81	0.90	Yes
olaOnly	sign	sSVM-IFR	rf	0.84	0.89	Yes
olaOnly	sign	coeffVar	svmRadial	0.82	0.83	Yes
olaOnly	sign	sSVM-IFR	nb	0.74	0.83	Yes
olaOnly	sign	coeffVar	rf	0.80	0.82	Yes
olaOnly	gene	Boruta	svmRadial	0.78	0.81	Yes
olaOnly	sign	mRMR	nnet	0.75	0.80	Yes
olaOnly	gene	Boruta	nnet	0.75	0.79	Yes
olaOnly	gene+sign	Boruta	nnet	0.74	0.78	Yes
olaOnly	gene+sign	Boruta	rf	0.76	0.77	Yes
olaOnly	gene+sign	Boruta	nb	0.76	0.74	Yes
olaOnly	gene	Boruta	nb	0.77	0.71	Yes
olaOnly	gene+sign	sSVM-IFR	nnet	0.69	0.85	No
olaOnly	gene+sign	sSVM-IFR	rf	0.70	0.84	No
olaOnly	gene+sign	sSVM-IFR	svmRadial	0.67	0.84	No
olaOnly	gene	coeffVar	svmRadial	0.65	0.83	No
olaOnly	gene+sign	Boruta	svmRadial	0.75	0.83	No
olaOnly	gene+sign	coeffVar	rf	0.76	0.83	No
olaOnly	sign	VSURF	nb	0.72	0.82	No
olaOnly	gene	sSVM-IFR	nnet	0.70	0.82	No
olaOnly	gene	VSURF	rf	0.71	0.82	No
olaOnly	sign	Boruta	nnet	0.74	0.82	No
olaOnly	gene	VSURF	svmRadial	0.77	0.82	No
olaOnly	gene+sign	VSURF	svmRadial	0.68	0.82	No
olaOnly	gene	sSVM-IFR	svmRadial	0.68	0.81	No
olaOnly	gene+sign	VSURF	nnet	0.73	0.81	No
olaOnly	gene+sign	coeffVar	svmRadial	0.69	0.81	No
olaOnly	sign	Boruta	nb	0.76	0.81	No
olaOnly	gene	sSVM-IFR	rf	0.68	0.81	No
olaOnly	sign	mRMR	rf	0.77	0.80	No
olaOnly	gene+sign	VSURF	rf	0.72	0.80	No
olaOnly	gene	VSURF	nb	0.75	0.79	No
olaOnly	gene	Boruta	rf	0.72	0.79	No
olaOnly	gene	mRMR	nnet	0.66	0.79	No

Table S2. Continued

Type	Input	FeatureSelection	Classifier	AUC_3Fold	AUC_5Fold	UsedInEnsemble
olaOnly	sign	VSURF	svmRadial	0.70	0.79	No
olaOnly	gene+sign	mRMR	nnet	0.67	0.79	No
olaOnly	gene+sign	mRMR	rf	0.70	0.78	No
olaOnly	gene+sign	VSURF	nb	0.87	0.78	No
olaOnly	sign	VSURF	rf	0.78	0.78	No
olaOnly	gene	mRMR	svmRadial	0.69	0.78	No
olaOnly	gene	mRMR	nb	0.67	0.77	No
olaOnly	sign	Boruta	rf	0.71	0.77	No
olaOnly	gene	DESeq2	svmRadial	0.77	0.77	No
olaOnly	gene+sign	mRMR	nb	0.65	0.77	No
olaOnly	gene	mRMR	rf	0.72	0.77	No
olaOnly	gene+sign	mRMR	svmRadial	0.68	0.76	No
olaOnly	sign	VSURF	nnet	0.77	0.76	No
olaOnly	gene	sSVM-IFR	nb	0.70	0.76	No
olaOnly	gene	DESeq2	rf	0.78	0.76	No
olaOnly	sign	Boruta	svmRadial	0.74	0.75	No
olaOnly	gene+sign	sSVM-IFR	nb	0.66	0.75	No
olaOnly	sign	mRMR	svmRadial	0.76	0.75	No
olaOnly	gene	DESeq2	nnet	0.78	0.75	No
olaOnly	gene	DESeq2	nb	0.76	0.74	No
olaOnly	sign	mRMR	nb	0.77	0.74	No
olaOnly	sign	Boruta	knn	0.56	0.74	No
olaOnly	gene	sSVM-IFR	knn	0.66	0.74	No
olaOnly	gene	coeffVar	rf	0.68	0.73	No
olaOnly	gene+sign	coeffVar	knn	0.71	0.73	No
olaOnly	gene	VSURF	nnet	0.74	0.72	No
olaOnly	sign	mRMR	knn	0.65	0.72	No
olaOnly	sign	sSVM-IFR	knn	0.61	0.72	No
olaOnly	sign	coeffVar	knn	0.69	0.70	No
olaOnly	gene	Boruta	knn	0.61	0.70	No
olaOnly	gene+sign	sSVM-IFR	knn	0.67	0.69	No
olaOnly	gene	coeffVar	knn	0.64	0.69	No
olaOnly	gene+sign	mRMR	knn	0.61	0.68	No
olaOnly	gene	mRMR	knn	0.60	0.67	No
olaOnly	gene	VSURF	knn	0.60	0.65	No
olaOnly	gene+sign	Boruta	knn	0.64	0.65	No
olaOnly	gene+sign	VSURF	knn	0.56	0.64	No
olaOnly	gene	DESeq2	knn	0.70	0.62	No
olaOnly	sign	VSURF	knn	0.65	0.62	No

Table S2. Continued

Type	Input	FeatureSelection	Classifier	AUC_3Fold	AUC_5Fold	UsedInEnsemble
olaOnly	gene+sign	coeffVar	nb			No
olaOnly	gene+sign	coeffVar	nnet			No
olaOnly	gene	coeffVar	nb			No
olaOnly	gene	coeffVar	nnet			No
olaOnly	sign	coeffVar	nb			No
olaOnly	sign	coeffVar	nnet			No
threeClass	gene	coeffVar	rf	0.55	0.61	Yes
threeClass	gene+sign	Boruta	knn	0.55	0.59	Yes
threeClass	gene	sSVM-IFR	rf	0.56	0.58	Yes
threeClass	gene+sign	coeffVar	rf	0.54	0.58	Yes
threeClass	sign	VSURF	knn	0.50	0.56	Yes
threeClass	sign	Boruta	svmRadial	0.53	0.55	Yes
threeClass	gene+sign	VSURF	knn	0.61	0.54	Yes
threeClass	gene+sign	VSURF	rf	0.52	0.53	Yes
threeClass	gene+sign	sSVM-IFR	svmRadial	0.55	0.53	Yes
threeClass	gene+sign	coeffVar	svmRadial	0.55	0.53	Yes
threeClass	sign	Boruta	nb	0.52	0.52	Yes
threeClass	gene	VSURF	rf	0.54	0.49	Yes
threeClass	gene	VSURF	knn	0.56	0.45	Yes
threeClass	gene	sSVM-IFR	nb	0.58	0.62	No
threeClass	sign	coeffVar	svmRadial	0.55	0.61	No
threeClass	gene	DESeq2	rf	0.54	0.61	No
threeClass	gene	DESeq2	nnet	0.57	0.60	No
threeClass	gene+sign	sSVM-IFR	nb	0.57	0.58	No
threeClass	sign	Boruta	nnet	0.52	0.58	No
threeClass	gene	Boruta	rf	0.50	0.58	No
threeClass	gene	sSVM-IFR	nnet	0.50	0.58	No
threeClass	sign	coeffVar	rf	0.55	0.58	No
threeClass	sign	sSVM-IFR	knn	0.56	0.58	No
threeClass	gene	DESeq2	svmRadial	0.53	0.56	No
threeClass	sign	mRMR	svmRadial	0.54	0.56	No
threeClass	gene+sign	sSVM-IFR	rf	0.59	0.56	No
threeClass	gene+sign	VSURF	svmRadial	0.55	0.56	No
threeClass	sign	coeffVar	knn	0.51	0.56	No
threeClass	gene	DESeq2	knn	0.54	0.56	No
threeClass	gene	Boruta	nnet	0.52	0.55	No
threeClass	gene+sign	Boruta	svmRadial	0.52	0.55	No
threeClass	gene+sign	mRMR	svmRadial	0.49	0.55	No
threeClass	gene+sign	VSURF	nnet	0.58	0.55	No

Table S2. Continued

Type	Input	FeatureSelection	Classifier	AUC_3Fold	AUC_5Fold	UsedInEnsemble
threeClass	gene+sign	Boruta	nnet	0.53	0.55	No
threeClass	sign	VSURF	rf	0.54	0.55	No
threeClass	sign	VSURF	svmRadial	0.55	0.55	No
threeClass	gene	coeffVar	knn	0.55	0.54	No
threeClass	sign	VSURF	nnet	0.53	0.54	No
threeClass	gene+sign	Boruta	rf	0.51	0.54	No
threeClass	sign	Boruta	knn	0.55	0.54	No
threeClass	gene+sign	coeffVar	knn	0.51	0.54	No
threeClass	gene	sSVM-IFR	svmRadial	0.51	0.54	No
threeClass	sign	sSVM-IFR	nnet	0.53	0.54	No
threeClass	gene	sSVM-IFR	knn	0.54	0.53	No
threeClass	gene	mRMR	knn	0.52	0.53	No
threeClass	sign	Boruta	rf	0.54	0.53	No
threeClass	gene	Boruta	svmRadial	0.49	0.53	No
threeClass	gene+sign	Boruta	nb	0.49	0.52	No
threeClass	sign	sSVM-IFR	rf	0.51	0.52	No
threeClass	sign	mRMR	rf	0.55	0.52	No
threeClass	sign	sSVM-IFR	svmRadial	0.52	0.52	No
threeClass	sign	VSURF	nb	0.51	0.52	No
threeClass	gene	Boruta	knn	0.49	0.52	No
threeClass	sign	mRMR	nb	0.53	0.52	No
threeClass	gene	Boruta	nb	0.48	0.51	No
threeClass	gene+sign	VSURF	nb	0.49	0.51	No
threeClass	sign	mRMR	nnet	0.52	0.51	No
threeClass	gene+sign	sSVM-IFR	knn	0.50	0.51	No
threeClass	gene+sign	mRMR	rf	0.54	0.50	No
threeClass	sign	mRMR	knn	0.52	0.50	No
threeClass	gene	mRMR	nb	0.52	0.50	No
threeClass	gene+sign	sSVM-IFR	nnet	0.59	0.49	No
threeClass	gene	mRMR	svmRadial	0.47	0.49	No
threeClass	gene	VSURF	nnet	0.59	0.49	No
threeClass	gene	mRMR	nnet	0.49	0.49	No
threeClass	gene	DESeq2	nb	0.48	0.47	No
threeClass	gene	VSURF	nb	0.54	0.47	No
threeClass	gene+sign	mRMR	nb	0.48	0.46	No
threeClass	gene	VSURF	svmRadial	0.54	0.46	No
threeClass	gene	mRMR	rf	0.48	0.46	No
threeClass	sign	sSVM-IFR	nb	0.50	0.46	No
threeClass	gene+sign	mRMR	nnet	0.45	0.44	No

Table S2. Continued

Type	Input	FeatureSelection	Classifier	AUC_3Fold	AUC_5Fold	UsedInEnsemble
threeClass	gene+sign	mRMR	knn	0.51	0.43	No
threeClass	gene	coeffVar	svmRadial	0.48	0.41	No
threeClass	gene+sign	coeffVar	nb			No
threeClass	gene+sign	coeffVar	nnet			No
threeClass	gene	coeffVar	nb			No
threeClass	gene	coeffVar	nnet			No
threeClass	sign	coeffVar	nb			No
threeClass	sign	coeffVar	nnet			No

Table S3: List of DNA crosslink repair genes.

canonical	Additional crosslink repair associated genes
FANCA	ATM
FANCB	ATR
FANCC	BLM
BRCA2	ERCC1
FANCD2	ERCC2
FANCE	ERCC3
FANCF	ERCC4
FANCG	MRE11A
FANCI	NBN
BRIP1	RAD18
FANCL	RAD21
FANCM	RPA1
PALB2	UBE2A
RAD51C	UBE2B
SLX4	UBE2T
RAD51	USP1
TP53BP1	WDR48
BRCA1	
FAAP100	
RAD50	
RAD51B	
RAD51D	
RAD52	
RAD54B	
RAD54L	
XRCC3	
XRCC2	

Table S4. Cohort Characteristics and univariate analysis. Hazard ratios and *p*-values were calculated for overall survival.

		NKI-CRAD			DESIGN		
		#	HR	<i>p</i> value	#	HR	<i>p</i> value
Sex	female	26	0.78	0.40	21	0.80	0.55
	male	72			61		
Age at Diagnosis	Under 65	67			63		
	Over 65	31	1.10	0.74	19	1.09	0.82
Stage	II	1	0.00	1.00	1	5.10	0.12
	III	17	0.65	0.24	18	0.65	0.32
	IVA	74	reference		50	reference	
	IVB	6	1.45	0.40	13	1.86	0.12
Tstage	T1/T2	28	0.69	0.26	17	1.22	0.65
	T3	38	reference		32	reference	
	T4	32	0.88	0.65	33	1.46	0.29
Nstage	N0/N1	29	0.64	0.12	24	0.74	0.39
	N2+	69	reference		58	reference	
Tumor Site	Hypopharynx	44	reference		24	reference	
	Larynx	11	1.20	0.71	20	1.01	0.98
	Oropharynx	43	1.85	0.021	38	1.91	0.0902
CisplatinDose	≤200mg/m ²	48	2.29	0.00163	36	1.89	0.0594
	>200mg/m ²	49	reference		42	reference	
	unknown	1			4		
Tobacco	Current	70	reference		59	reference	
	Former	21	0.57	0.11	19	1.08	0.84
	Never	2	6.85	0.0105	0		
	unknown	5			4		
Alcohol	Current	68	reference		58	reference	
	Former	16	1.63	0.12	13	1.40	0.40
	Never	8	0.59	0.31	5	0.94	0.93
	unknown	6			6		

Table S5. *p*-values for *t*-tests of model scores between different clinical variables. Companion to Figure S4.

NKI-CRAD					DESIGN						
Stage					Stage						
mmcOnly					mmcOnly						
	III	IVA				III	IVA				
IVA	0.83				IVA	0.86					
IVB	0.54	0.41			IVB	0.69	0.76				
threeClass SensM+MO					threeClass SensM+MO						
	III	IVA				III	IVA				
IVA	0.88				IVA	0.21					
IVB	0.45	0.41			IVB	0.89	0.34				
Tstage					Tstage						
mmcOnly					mmcOnly						
	T1	T2	T3	T4a		T1	T2	T3	T4a		
T2	0.72				T2	0.03672					
T3	0.61	0.80			T3	0.00081	0.61				
T4a	0.61	0.80	0.99		T4a	0.00011	0.0602	0.10			
T4b	0.54	0.65	0.71	0.72	T4b	0.02707	0.43	0.64	0.44		
threeClass SensM+MO					threeClass SensM+MO						
	T1	T2	T3	T4a		T1	T2	T3	T4a		
T2	0.97				T2	0.21					
T3	0.84	0.67			T3	0.10	0.54				
T4a	0.51	0.23	0.36		T4a	0.056	0.27	0.51			
T4b	0.92	0.89	0.98	0.75	T4b	0.05	0.21	0.37	0.73		
Nstage					Nstage						
mmcOnly					mmcOnly						
	N0	N1	N2a	N2b	N2c		N0	N1	N2a	N2b	N2c
N1	0.47					N1	0.40				
N2a	0.17	0.43				N2a	0.81	0.40			
N2b	0.09	0.39	0.82			N2b	0.21	0.81	0.29		
N2c	0.11	0.40	0.86	0.95		N2c	0.74	0.53	0.63	0.28	
N3	0.31	0.62	0.86	0.98	0.96	N3	0.12	0.27	0.13	0.30	0.15
threeClass SensM+MO					threeClass SensM+MO						
	N0	N1	N2a	N2b	N2c		N0	N1	N2a	N2b	N2c
N1	0.90					N1	0.39				
N2a	0.20	0.15				N2a	0.52	0.97			
N2b	0.31	0.18	0.49			N2b	0.0472	0.33	0.55		
N2c	0.21	0.11	0.66	0.73		N2c	0.54	0.72	0.78	0.10	
N3	0.51	0.47	0.89	0.79	0.91	N3	0.30	0.56	0.63	0.88	0.44
Differentiation Grade					Differentiation Grade						
mmcOnly					mmcOnly						
	Undiff	Low	Med			Low	Med				
Low	0.63				Med	0.23					
Med	0.56	0.82			High	0.21	0.70				
High	0.96	0.56	0.48								

Table S5. Continued

NKI-CRAD				DESIGN		
threeClass SensM+MO				threeClass SensM+MO		
	Undiff	Low	Med		Low	Med
Low	0.58			Med	0.20	
Med	0.67	0.77		High	0.17	0.49
High	0.92	0.46	0.54			
Histological Growth Pattern						
mmcOnly						
	Cohesive					
Non-cohesive	0.90					
threeClass SensM+MO						
	Cohesive					
Non-cohesive	0.99					
Tumor Site				Tumor Site		
mmcOnly				mmcOnly		
	Hypopharynx	Larynx			Hypopharynx	Larynx
Larynx	0.24			Larynx	0.52	
Oropharynx	0.63	0.14		Oropharynx	0.79	0.35
threeClass SensM+MO				threeClass SensM+MO		
	Hypopharynx	Larynx			Hypopharynx	Larynx
Larynx	0.41			Larynx	0.54	
Oropharynx	0.39	0.18		Oropharynx	0.73	0.32



Martijn van der Heijden
Paul B.M. Essers
Monique de Jong
Reinout H. de Roest
Sebastian Sanduleanu
Caroline V.M. Verhagen
Olga Hamming-Vrieze
Frank Hoebers
Philippe Lamin
Harry Bartelink
C. René Leemans
Marcel Verheij
Ruud H. Brakenhoff
Michiel W.M. van den Brekel
Conchita Vens

Abstract

Purpose: Tumor markers that are related to hypoxia, proliferation, DNA damage repair and stem cell-ness, have a prognostic value in advanced stage HNSCC patients when assessed individually. Here we aimed to evaluate and validate this in a multifactorial context and assess interrelation and the combined role of these biological factors in determining chemoradiotherapy response in HPV-negative advanced HNSCC.

Methods: RNA sequencing data of pre-treatment biopsy material from 197 HPV-negative advanced stage HNSCC patients treated with definitive chemoradiotherapy was analyzed. Biological parameter scores were assigned to patient samples using previously generated and described gene expression signatures. Locoregional control rates were used to assess the role of these biological parameters in radiation response and compared to distant metastasis data. Biological factors were ranked according to their clinical impact using bootstrapping methods and multivariate Cox regression analyses that included clinical variables. Multivariate Cox regression analyses comprising all biological variables were used to define their relative role among all factors when combined.

Results: Only few biomarker scores correlate with each other, underscoring their independence. The different biological factors do not correlate or cluster, except for the two stem cell markers CD44 and SLC3A2 ($r=0.4$, $p<0.001$) and acute hypoxia prediction scores which correlated with T-cell infiltration score, CD8⁺ T cell abundance and proliferation scores ($r=0.52$, 0.56 and 0.6 respectively with $p<0.001$). Locoregional control association analyses revealed that chronic (Hazard Ratio (HR)=3.9) and acute hypoxia (HR=1.9), followed by stem cell-ness (CD44/SLC3A2; HR=2.2/2.3), were the strongest and most robust determinants of radiation response. Furthermore, multivariable analysis, considering other biological and clinical factors, reveal a significant role for EGFR expression (HR=2.9, $p<0.05$) and T-cell infiltration (CD8⁺T-cells: HR=2.2, $p<0.05$; CD8⁺T-cells/Treg: HR=2.6, $p<0.01$) signatures in locoregional control of chemoradiotherapy-treated HNSCC.

Conclusion: Tumor acute and chronic hypoxia, stem cell-ness, and CD8⁺ T-cell parameters are relevant and largely independent biological factors that together contribute to locoregional control. The combined analyses illustrate the additive value of multifactorial analyses and support a role for EGFR expression analysis and immune cell markers in addition to previously validated biomarkers. This external validation underscores the relevance of biological factors in determining chemoradiotherapy outcome in HNSCC.

In this study we set out to perform multifactorial analyses to gain understanding of the role and dependence of biological factors that have shown to influence tumor radiation response in preclinical studies and to be associated with radiotherapy response in clinical studies^{1,2}. Chemo-radiotherapy is the primary treatment option for advanced head and neck squamous cell carcinoma (HNSCC). Cure and locoregional failure rates of around 50% and 25%, respectively, facilitate the evaluation of biological determinants of radiation response. Using biological characteristics of the tumors, outcome association studies revealed many potential determinants of prognosis and treatment response in HNSCC³⁻⁷. This study evaluates complementarity and hierarchy of radiation response determining “HNSCC biology”.

Early radiobiology studies revealed determinants of tumor radiation response. Hypoxia, repopulation, driven by tumor cell proliferation, tumor stem cell density (i.e. clonogenic cell density) and cellular radiosensitivity (as for example determined by cellular DNA damage repair capacity) were shown to be among the most relevant biological factors that affect radiation or fractionated radiotherapy response in preclinical models of different cancers^{1,2}. In recent years, increased interest emerged in immune response related markers and immune cells due to novel immunotherapeutic options^{17–21}. A series of preclinical and clinical studies highlight the potential relevance of immune-related markers in HNSCC (reviewed in^{5,6,19,22–24}).

HNSCC outcome association studies using many different biomarkers, demonstrated the clinical importance of some of these pre-clinically assessed tumor biology parameters^{1,5-7}. HPV and hypoxia are indeed the best studied biology related prognostic markers in HNSCC. Within the HPV-negative patients, tumor hypoxia marks patients with a poor prognosis²⁵⁻²⁹. Confirming its role above marking poor prognosis patients, hypoxia biomarkers also predict response to hypoxia modification therapy^{25,30-33}. Elaborating on a gene expression profile that captures the cellular changes caused by acute hypoxia, we recently showed the relevance of acute hypoxia in addition to chronic hypoxia²⁹. As predicted by the process they capture, these two classifications did not necessarily overlap in the samples and also reveal different outcome associations in HNSCC that result from a prominent role of acute hypoxia.

While the success of accelerated radiotherapy schedules³⁴ highlight the important role of tumor repopulation in HNSCC, there is a lack of biomarker data showing a link to cellular proliferation³⁵. Based on genetic mutation data, we find a small role for co-occurring CCND1 and CDKN2A mutations in HPV-negative chemo-radiotherapy treated HNSCC that was however not visible in the locoregional control endpoints³⁶. Yet, the combination of radiotherapy with the epidermal growth factor receptor (EGFR) binding antibody cetuximab has shown efficacy and EGFR expression has been associated with poor survival, preferentially in non-accelerated schedules arguing for a role in tumor repopulation³⁷⁻⁴⁰. The role of EGFR and cellular proliferation in radiotherapy response needs to be further elaborated⁴¹⁻⁴³. However proliferation, as determined by the proliferation marker by Starmans *et al.*, has been linked to aggressive disease or disease progression in multiple cancer types; unfortunately this was not assessed in HNSCC⁴⁴.

Originated from CD44 expression data from de Jong *et al.*⁴⁵ in laryngeal cancer and confirmed in resected and chemo-radiotherapy treated HNSCC for CD44 and SLC3A2^{27,46} in subsequent studies, it also became clear that tumor “stem cell-ness” is important in radiotherapy outcomes since these stem cell related biomarkers were associated with poor prognosis^{35,47-49}.

The consistent effect of CD8⁺ T cell depletion on radiation induced tumor growth delays in preclinical studies expose the relevance of certain immune cell populations in radiation response and resistance^{50,51}. Evidence in clinic of a possible interaction is less strong and current studies focus on strategies to optimize combinations with immune response modulators to improve radiotherapy outcomes^{6,18,20,21,27,52-55}. Interestingly, Mandal *et al.* recently showed that markers for regulatory T cells (Treg), NK cells and CD8⁺ T cells are prognostic in head and neck cancer⁵⁶ in the TCGA dataset. Despite these interesting initial

reports, the prognostic value of these gene expression based immune markers is still unknown for chemo-radiotherapy treated patients since all patients in the TCGA dataset have been treated with primary surgery. Immunohistochemically (IHC) determined high CD8⁺ T-cell counts are associated with good prognosis in postoperative chemo-radiotherapy treated patients, further indicating its relevance for HNSCC⁵⁷. A good prognosis association with IHC CD8⁺ TIL density was found in patients with oropharyngeal squamous cell carcinoma treated with surgery or (chemo) radiotherapy and in a similarly mixed treatment cohort of hypopharyngeal SCC patients^{58–60}.

Our previous studies emphasized the important role of functional and genetic DNA crosslink repair defects in HNSCC^{61,62} and provided the basis for machine learning generated models that predicted such DNA repair defects in clinical samples⁶³. The expression based DNA repair defect prediction models revealed an association with metastasis in HNSCC and linked DNA repair defects to migratory and invasive behavior in HNSCC cell lines⁶³. Given the relevance of Epithelial to Mesenchymal Transition (EMT) in many cancer types, we also developed a HNSCC-specific EMT model that classifies HNSCC according to epithelial or mesenchymal characteristics⁶⁴. The strong prognostic value of this HNSCC-EMT model also suggests an important role in radiation response.

Taken together, the individual roles of some of these biological factors important in radiation response have not been validated and the interrelation of these biological factors has not been investigated in the clinical setting. We therefore studied the role of the aforementioned biological factors in the context of head and neck cancer and chemo-radiotherapy. Previously published gene expression based signatures were used to detect these factors. In a set of nearly 200 patients with advanced stage HPV-negative HNSCC treated with chemo-radiotherapy, we used univariate and multivariate outcome analyses to examine these factors while also considering correlation and dependence to delineate their relative roles.

Materials and Methods

Patient Data and Material

This retrospective study included material and data from patients that were diagnosed between 2001-2014 and treated with definitive cisplatin-based chemo-radiotherapy within three centers: the Netherlands Cancer Institute (Amsterdam, NL), the Amsterdam University Medical Center (Amsterdam, NL) or the MAASTRO clinic/MUMC+ (Maastricht, NL). Selection criteria for this gene expression study cohort were (i) concomitant radiotherapy and

cisplatin treatment of unresected HNSCC, (ii) hypopharyngeal, laryngeal or HPV-negative oropharyngeal (iii) no prior treatment with chemotherapy or radiotherapy in the head and neck area. To minimize the number of variables, AJCC disease staging, summarizing TNM stage, was used to classify HNSCC patients after determining whether this classification also represented N-staging and its known association with survival well (Supplementary Figure S1). Received radiotherapy regimens were 70 Gy over 35 fractions (up to 77 Gy in ARTFORCE patients) in 7 or 6 weeks (DAHANCA scheme). All patients were treated with either of four different cisplatin regimens: daily (25x6mg/m² Body Surface Area (BSA)), weekly (7x40mg/m² BSA) or 3-weekly (3x100mg/m² BSA) intravenous administration or weekly intra-arterial administration (4x150mg/m² BSA, for 8 patients according to the RADPLAT trial protocol⁶⁵). Not all patients completed the full chemotherapy scheme. Therefore, cumulative cisplatin doses were calculated and patients were classified into < or ≥ or 200mg/m² BSA cisplatin, according to literature^{66,67}. Survival data was calculated from the start of treatment until the first event was detected. The primary outcome measure is loco-regional control (LRC) and implies absence of recurrences in the radiotherapy targeted regions of the head and neck area. Patient characteristics are provided in Table S1. Institutional Review Boards at the Netherlands Cancer Institute, the Amsterdam University Medical Center and the MAASTRO clinic/MUMC+ approved biopsies and collection of fresh-frozen HNSCC tumor material and the use of genetic and clinical data from patients at their respective centers after anonymization. All patients granted written informed consent for biopsy, material use and data use. Pre-treatment tumor biopsy material available for the DESIGN study or collected from the NKI ARTFORCE⁶⁸ or RADPLAT trial patients were used for RNA preparation and sequencing. HPV-status of all oropharyngeal carcinomas was determined by immunohistochemical assessment of p16 by a dedicated head and neck pathologist⁶⁹ followed by a HPV DNA test on the p16-immunopositive cases and/or confirmed using RNA-sequencing data.

Material preparation and RNA-sequencing

Fresh-frozen tumor samples were sectioned, collected for RNA preparation and in part subjected to tumor percentage evaluation by revision of HE stained coupes by senior head and neck pathologist Dr. S.M. Willems. Only samples with a tumor percentage of >40% proceeded to RNA-sequencing. RNA was isolated using the AllPrep DNA/RNA mini kit (Qiagen). Quality and quantity of total RNA was assessed by the 2100 Bioanalyzer using a Nano chip (Agilent, Santa Clara, CA). Only total RNA samples having RIN>7 were used for library preparation. Strand-specific libraries were generated using the TruSeq Stranded mRNA sample preparation kit (Illumina Inc., San Diego, RS-122-2101/2) according to the

Expression and Patient Outcome Analyses

In order to obtain a robust cut off when transforming a continuous variable into a dichotomous variable we used the bootstrap procedure as described in Linge *et al.*²⁸. In brief, 197 sample values were randomly assigned into one bootstrap cohort (from the cohort of 197 patients) while data from the same patient could be chosen multiple times. This procedure was repeated to obtain 10.000 randomized cohorts. At each possible cutoff value of the marker of interest, the individual cohorts were split into a “low” and “high” group and Cox proportional hazards models were fit based on these splits. These models included, next to the newly grouped marker of interest, all clinical variables that were found to be significantly associated with the outcome of interest (Locoregional Control (LRC), Distant Metastasis (DM), Overall Survival (OS) or Progression Free Survival (PFS)). The fraction of cohorts for which the marker of interest was significantly associated with survival

($p < 0.05$) was recorded for each cutoff. The values of nine adjacent cutoffs were averaged to smoothen the data. The cutoff with the highest fraction of significant associations was chosen for further analysis. Cutoffs that would result in patient subgroups with less than 10% of the patients were not considered to maintain statistical power. Note that, this analysis was repeated for each endpoint resulting in different cutoffs.

To reduce the number of possible variables included in multivariable analysis we used a backward selection procedure. The most frequent level of each variable was used as the reference level for this analysis. A Cox proportional hazard model was fit containing all biological markers and clinical variables. Then, each individual variable was removed from the model and improvements in model performance by this process were assessed using the Akaike Information Criterion (AIC) from the 'stats' package in R. The best model (lowest AIC) was selected for further analysis in the multivariate Cox regression analysis. This process was repeated until removing variables from the model did no longer result in an improved model.

Results

Role of clinical factors and patient characteristics in chemo-radiotherapy outcome

In this retrospective multicenter study, 197 patients met all inclusion criteria and had sufficient tumor material available. All patients were treated with definitive cisplatin-based chemo-radiotherapy for advanced stage HPV-negative oropharyngeal, hypopharyngeal or laryngeal carcinoma. Patient characteristics are shown in Table 1. The median age in this patient cohort is 62 years and there is a male: female ratio of 3:1. Most patients reported ongoing or a history of alcohol and/or tobacco use. The largest subsite representation is oropharyngeal tumors with 85 patients, then hypopharyngeal with 78 and laryngeal carcinoma with 34 patients. Except for two patients, all patients had stage III/IV classified tumors. As expected, outcomes and survival curves differ according to stage (Supplementary Figure S2). Tumor volume data as determined by delineation on RT planning CT images were available for 166 patients with a median volume of 23.2 cm³. Not all patients finished chemotherapy, but 126 patients (63%) received a cumulative dose of and above 200mg/m² body surface area. Locoregional recurrences occurred in 23.8% ($N = 49$) of cases and distant metastasis in 19,8% ($N = 39$).

Clinical factors were tested for their association with locoregional control and other survival outcomes (Supplementary Figure S1, Supplementary Table S1). Consistent with previous reports we find that locoregional control (LRC) is influenced by cumulative cisplatin dose levels^{66,67}. The cumulative cisplatin dose of < 200 mg/m² BSA was significantly associated with LRC failure (HR=2.57, $p=0.0012$). Female sex shows a trend towards better locoregional

control (HR=0.52, $p=0.072$). This could however been confounded by the less prominent alcohol consumption characteristics or other differences in lifestyle in this particular patient group. More female patients reported to abstain from alcohol compared to male patients (21.8% vs 7.4%, $p=0.019$), which was however not the case for tobacco use ($p=0.66$). Heavy past or ongoing alcohol consumption was associated with an increased risk for LRC failure (HR=2.16, $p=0.041$). Interestingly, age, tobacco, tumor subsite and AJCC stage is not significantly associated with LRC in our patient cohort. The other clinical outcomes (DM, PFS or OS) showed significant associations with sex, tumor volume, stage and cisplatin (Supplemental Figure S1, Supplementary Table S1).

Table 1. HNSCC patients and tumor characteristics

Variable		# Patients (total 197)
Median Age at diagnosis (Range)		61 years (40-80.23)
Sex	female	55
	male	142
Alcohol	no	22
	former alcoholic	22
	yes	146
	missing	7
Tobacco	never	5
	former smoker	30
	yes	156
	missing	6
Tumour subsite	Larynx	34
	Hypopharynx	78
	Oropharynx	85
Median tumour volume (Range)		23.2 cm ³ (1.03-752.2)
Stage	missing	31
	IVB	20
	III	40
	II	2
	IVA	135
Cumulative cisplatin dose	<200mg/m ²	67
	≥200mg/m ²	126
	missing	4
Median Follow-up		5.24 years (4.59-5.86)
Locoregional Recurrences		49
Distant Metastasis		39

Tumor biology assessment and (in)dependence in HPV-negative HNSCC

Preclinical radiobiology studies and clinical biomarker studies exposed many different determinants of radiation response. The number of variables that can be included in statistical analyses are however limited by the cohort size and number of events. Thus, in order to evaluate the relative role of different tumor biology parameters in clinic, we prioritized those with a reported clinical outcome association. The following twelve gene expression signatures were therefore selected to characterize the clinical samples using pretreatment HNSCC transcriptomic data: The Toustrup (1) and Seigneuriac (2) expression signatures were used to assess the level of chronic (1) and acute (2) hypoxia respectively^{25,73}. Linked to tumor stem cell richness SLC3A2 (3) and CD44 (4) gene expression values (in TPM) were included since both have been reported to be associated with outcome in chemo-radiotherapy treated patients⁷⁴. (5) EGFR expression (in TPM) and (6) the Starmans *et al.* ‘proliferation’ expression signature⁴⁴ were selected as cellular proliferation markers which could influence tumor repopulation between radiotherapy fractions. To cover immune-related factors we further included expression signatures from Senbabaoglu *et al.*⁷⁵ that originated from Bindea *et al.*⁷⁶ and assess (7) ‘T-cell infiltration score’ (TIS), (8) CD8⁺ T-cells, (9) CD56^{dim} natural killer (NK) cells abundance while considering the (10) CD8⁺ versus T regulatory (Treg) cell ratios. This immune status gene expression signature selection is based on the reported outcome association in resected HNSCC⁵⁶. Our own studies conducted in HPV-negative advanced HNSCC revealed an important role for EMT and DNA crosslink (CL) repair defects in treatment outcome and these prediction models for mesenchymal characteristics and tumor cell DNA crosslink repair defects, (11) ‘HNSCC-EMT’ and (12) ‘MMConly’, were therefore included in this analysis and are referred to as ‘EMT’ and ‘DNA CL repair’ in this manuscript^{63,64}. Most of these biological factors have been tested individually, predominantly in univariable analyses and in different settings in previous studies; however their mutual correlations and possible dependence between them are unknown.

The goal of this study is to pinpoint biological factors that are important for (chemo)-radiotherapy treatment failure and thus might validate their independent role in radioresistance. We therefore calculated scores for all aforementioned markers. Supplemental Figure S3 shows the frequency of the scores and their distribution over the patient cohort. Next, we performed hierarchical clustering to investigate the presence of HNSCC subsets as defined by these characteristics. Tumor volume was included in this analysis as it promotes chronic hypoxia or may be associated with high proliferation scores. Despite the coexistence and correlation of some factors this does however not reveal any prominent clusters (Figure 1A and B). Surprisingly, we find that the acute hypoxia profile score correlates with the Starmans proliferation score ($r=0.58$, $p<0.001$) (Figure 1C) but also with the T-cell Infiltration Score (TIS) and the CD8⁺ T cell scores ($r=0.51$ and $r=0.54$, $p<0.001$).

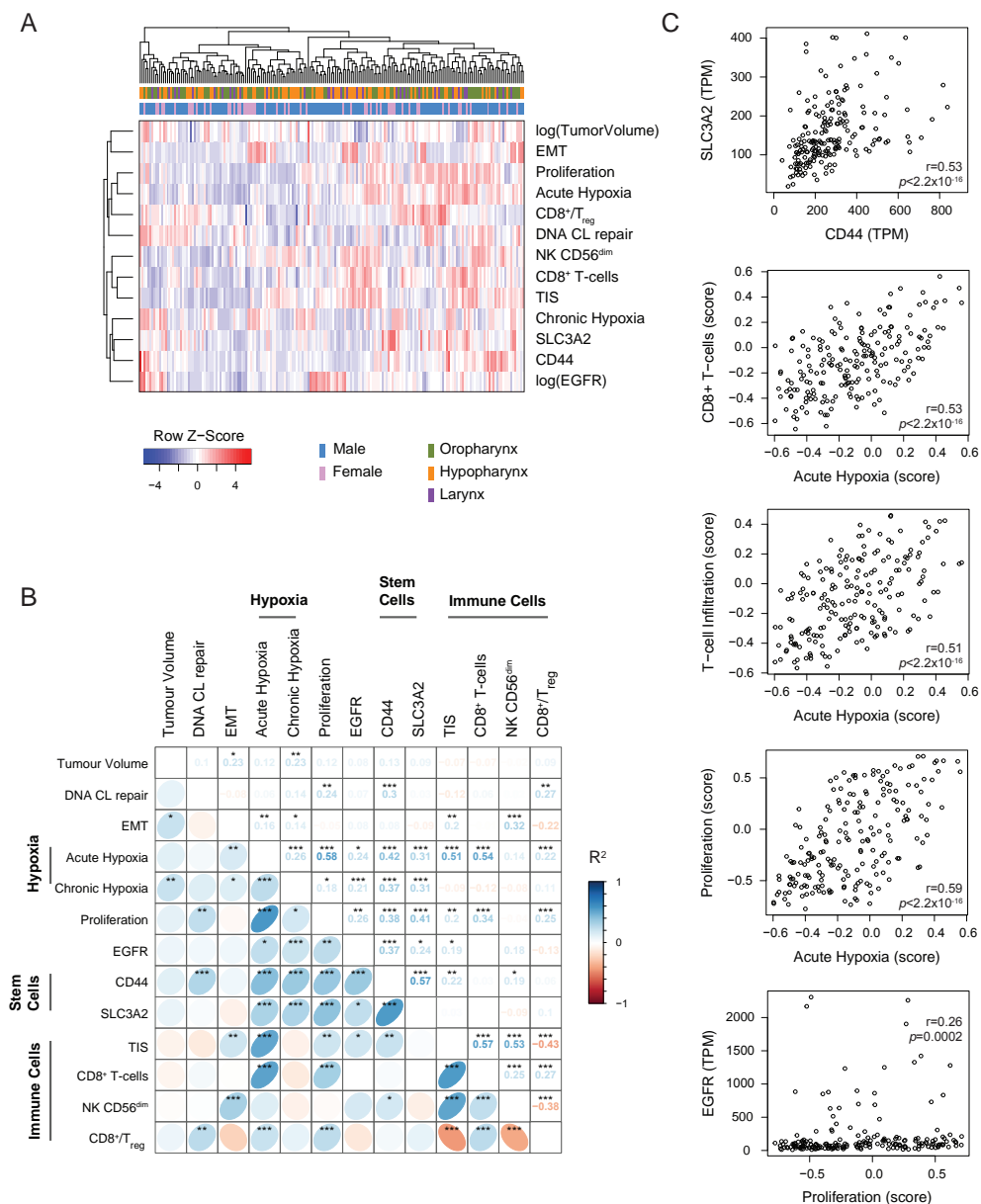


Figure 1. Interrelation and dependence of biological markers. A) Heatmap showing hierarchical clustering of patients based on the selected biological markers. Tumor volume and EGFR expression were log-transformed to prevent clustering to be dominated by few samples with high values. **B)** Correlation plot with spearman's ranked correlation values between all biological marker scores. Markers were grouped by tumor biology class where appropriate. * $p<0.05$, ** $p<0.01$, *** $p<0.001$, not corrected for multiple testing. **C)** Scatterplots with Spearman's coefficients and p -values for correlations of interest.

Within their own category, stem cell related markers, CD44 and SLC3A2 ($r=0.57$, $p<0.001$), and the immune cell related markers correlate with each other. While the correlation of acute and chronic hypoxia is significant, it is fairly weak ($r=0.26$, $p<0.001$) and was in line with previous reports²⁹. EGFR expression and the proliferation score are correlated to some extent ($r=0.26$, $p<0.01$). The CD8⁺ T cell to regulatory T cell ratio (CD8⁺/Treg) as determined by the expression signature scores is negatively associated with the abundance of CD56^{dim} natural killer (NK) cells and the TIS signature. While the link between tumor volume and chronic hypoxia ($r=0.23$, $p<0.01$) is expected, tumor volume is also associated with EMT ($r=0.23$, $p<0.05$). With a maximum variance inflation factor value of 3.3, correlations were not strong enough to exclude parameters from subsequent analyses. None of these markers show strong associations with any of the clinical factors. Among all, we find that the most independent tumor characteristics are the presence of DNA CL repair defects and tumor EMT status (and tumor volumes).

Role of individual biological factors in locoregional control by chemo-radiotherapy in HNSCC

Since we aimed to evaluate tumor characteristics with respect to radiation resistance and response, we initially focused on locoregional control outcome values that are mainly determined by the success of the 'local' radiotherapy treatment. Given the lack of strong correlations, all markers were individually tested for their association with locoregional failure. A 10.000 times bootstrapping method was employed to (a) determine a potential role for the biomarker across different cutoffs and (b) to identify a clinically robust cutoff for each so to compare the biomarkers among each other. In brief, each marker was tested for their association with the selected survival outcome for all possible cutoffs. This analysis was performed using a multivariable Cox proportional hazard model with all relevant clinical factors included, as determined above. Consequently, clinical variables were included according to outcome type: sex and cumulative cisplatin dose for LRC; sex, subsite and cumulative cisplatin dose for OS; stage, subsite and cisplatin dose for PFS; and sex and alcohol use for DM. Based on the results of these 10.000 bootstrap repeats (Figure 2A), we find that the hypoxia and stem cell related markers are most robustly associated with LRC across different score cutoffs. Proliferation, EGFR and immune cell signatures merely provide significant associations with LRC in a fraction of the randomly created cohorts and tested cutoffs.

Cutoffs with the most stable clinical association were selected for each biomarker for further analysis as depicted in Figure 2A and listed in Supplementary Table S2. These analyses confirm that both, chronic and acute hypoxia, are strongly associated with locoregional control. Using these calculated cutoffs in multivariable analyses with clinical factors, we find that among all chronic hypoxia is most strongly associated with a failure of locoregional control (HR=3.95, $p=0.0038$) followed by acute hypoxia (HR=1.9, $p=0.03$) and stem cell related, SLC3A2 (HR=2.31,

$p=0.026$) and CD44 (HR=2.03, $p=0.043$; Figure 2B and 3 and Supplementary Table S3). Although not significantly, larger tumor volumes showed a trend towards worse locoregional control with a hazard ratio (HR=1.63, $p=0.11$) that is comparable to those previously reported by others²⁷. It should be however noted that most tumors in this cohort are relatively large and stage III/IV. This and the fact that the LRC measure also includes regional recurrences, may together affect the specific HR values. Trends towards a worse LRC prognosis were observed in patient groups with tumors with high proliferation and CD8+ T cell scores (HR=1.89, $p=0.067$ and HR=2.35, $p=0.071$, respectively) (Supplementary Figure S4 and Supplementary Table S3).

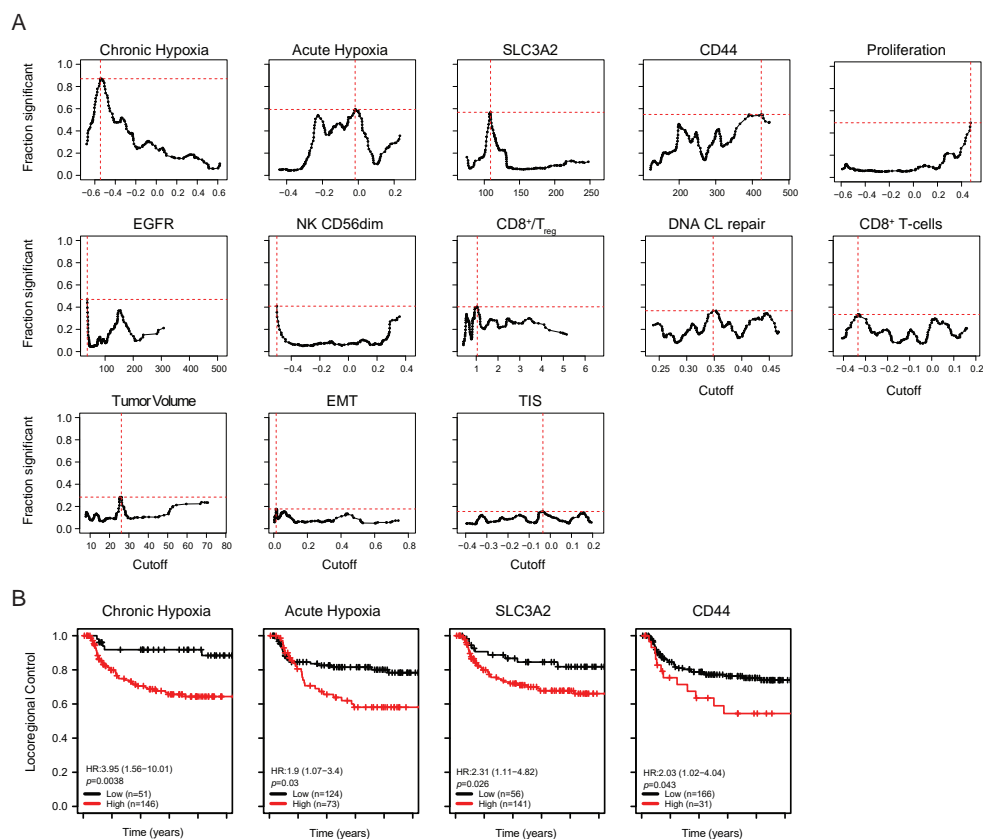


Figure 2. Role of individual biological factors for locoregional control. A) Results of individual bootstrap analysis (see methods) including important clinical factors. The fraction of randomized cohorts with a significant association with locoregional control is shown for each cut off and biological marker. Markers are ordered by the magnitude of the maximum fraction of significant Cox proportional hazard tests at the best cutoff. The best cut off is indicated with dotted red lines. **B)** For each marker, the cohort was split into a high and low group at the best cut off determined in A. Hazard ratios for recurrences and corresponding p -values were obtained with a multivariate Cox proportional hazard analysis using the same variables as those used to determine the cutoff.

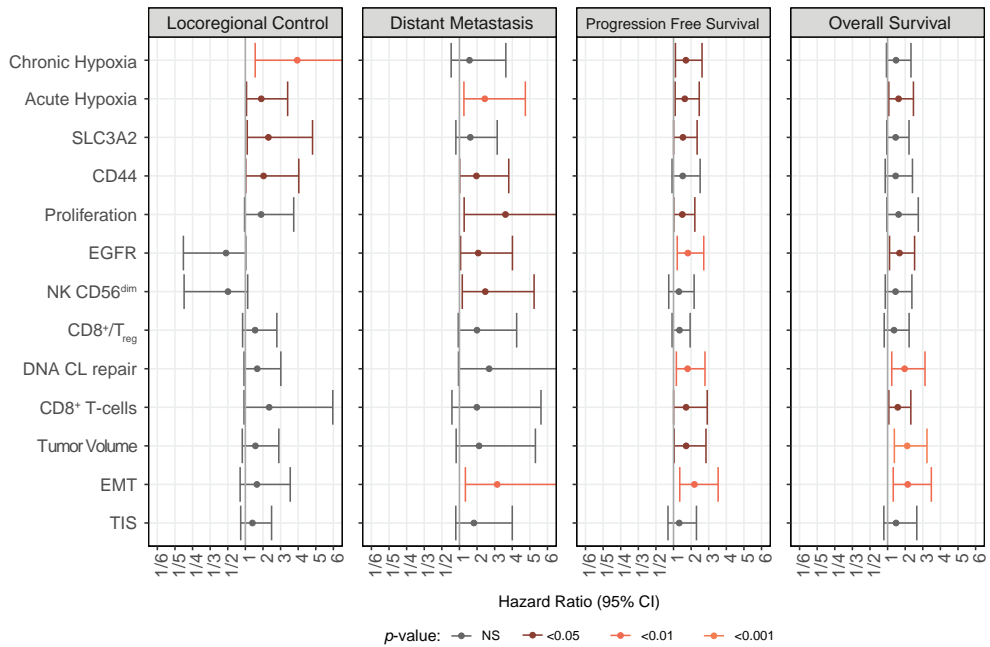


Figure 3. Relevance of biological markers for HNSCC chemo-radiotherapy outcomes. Forest plots comparing individual biological markers across different outcome endpoints are shown. Hazard Ratios for locoregional recurrences, distant metastasis, disease progression and death are from multivariate Cox proportional hazard analyses of the dichotomized cohorts using bootstrap analysis defined cut offs for all outcomes analyzed in this study. For each marker the “low” score or low expression group was used as reference. Each individual model contained the same clinical variables as those used to determine the cut off. Only the biological markers are shown.

To delineate this from general poor prognosis pattern and to investigate the radiation response link further, we repeated these analyses and compared the role of these biomarkers for overall survival, progression free survival and distant metastasis (Figure 3). Cutoff values were defined by the bootstrapping method described above for each biomarker; and multivariable Cox proportional hazard analyses with clinical factors were performed. Notably, gene expression signature scores or expression value cut offs, as determined by their potential relevance in the 10.000x bootstrapping method, resulted to be different in some of the biomarkers such as ‘acute hypoxia’, ‘chronic hypoxia’, ‘EGFR’, ‘TIS’, ‘NK CD56^{dim}’ and ‘CD8+/Treg’ (Supplementary Figures S4 to S6 and Supplementary Table S2).

Most markers show an association with several of the outcome parameters. We find that distant metastasis is associated with EMT (HR=3.14, $p=0.0086$), acute hypoxia score (HR=2.44, $p=0.0086$), NK CD56^{dim} score (HR=2.47, $p=0.019$) and EGFR expression (HR=2.07, $p=0.032$) (Supplementary Figure S5, Supplementary Table S3). Poor overall survival is

associated with increased tumor volume (HR=2.12, $p=0.00054$), EMT score (HR=2.15, $p=0.002$), DNA CL repair defect (HR=1.97, $p=0.0043$), acute hypoxia (HR=1.62, $p=0.023$) and EGFR expression (HR=1.68, $p=0.014$) (Supplementary Figure S6, Supplementary Table S3). High EMT (HR=2.19, $p=0.0014$), EGFR (HR=1.82, $p=0.0038$), acute hypoxia (HR=1.64, $p=0.017$), DNA CL repair defect (HR=1.8, $p=0.0085$) and chronic hypoxia scores (HR=1.7, $p=0.015$) and tumor volumes (HR=1.72, $p=0.036$) are associated with a worse progression free survival (PFS) (Supplementary Figure S7; Supplementary Table S3). Interestingly, when comparing the distant metastasis and locoregional control failure data, locoregional control is increased in tumors with higher EGFR expression or containing few CD56^{dim} NK cells while high values in both result in an increased risk of DM. It should be noted, however, that the bootstrapping defined cut offs were different in both. Yet, as evident from the bootstrapping data chronic hypoxia was not linked to DM but LRC at many cut offs. On the contrary, the similar shape of the results from the acute hypoxia bootstrapping supports its relevance in both, LRC and DM (Figure 2 and Supplementary Fig S5).

Taken together, for biomarkers which have been previously reported to be prognostic in HNSCC these analyses validate their role in an independent data set. Most biological factors as determined by the selected biomarkers are significantly associated with PFS thereby confirming their relevance. Overall, we find a prominent role for acute and chronic hypoxia and CD44 and SCL3A2 in our cohorts. We show that, from all, chronic hypoxia appears to be the most specific to LRC. In contrast, HR values from EMT and proliferation based splits are greater when assessing DM. Furthermore, these data reveal a role for the immune cell and proliferation related biomarkers in HNSCC outcome after definitive chemo-radiotherapy.

The relative role of biological factors in chemo-radiotherapy outcomes in HNSCC

The biological markers have been tested independently of each other and most are significantly associated with patient outcome thereby supporting their role in HNSCC and treatment response. However, tumor biology and determinants of radioresistance are multifactorial and may depend on the context and relation to each other. We therefore aimed to identify the most relevant markers in a multivariable analysis. To this end we used a backward selection method. This method creates a Cox proportional hazard model using all available factors. It then iteratively eliminates the least relevant factor until no further decrease in AIC, a measure of model performance, is possible. From these analyses (Supplementary Table S4), we conclude that chronic hypoxia, EGFR expression, CD8+/Treg, T-cell infiltration and CD44 are the most relevant biological factors that are associated with locoregional control. Multivariable analyses (Figure 4) also demonstrate that they are

independent from relevant clinical factors such as cumulative cisplatin dose or sex. Cisplatin dose, age and sex are the clinical factors most associated with locoregional control in this cohort (Figure 4 and Supplementary Table S4). Broadly consistent with the results from the multivariate analyses that were performed on the biomarkers on an individual basis, EGFR and immune cell related factors remain important in instituting an increased risk for distant metastasis, while chronic hypoxia and CD44 are less relevant. Instead, tumor EMT and proliferation affects progression free survival most profoundly and independent of other important factors such as tumor volume or cisplatin dose (Supplementary Figure S8 and Supplementary Table S4). The consistent worse prognosis and distant metastasis association of patients with tumors that score high in the CD8⁺ T cells and related gene expression signatures is remarkable. High CD8⁺ T cell scores, as determined by these signatures or by immunohistochemistry, have been reported to be linked to good prognosis in other heterogeneous HNSCC cohorts^{6,56,77} and prompted us to analyze this further (Supplementary Figure S9, S10). These analyses suggest that the lack of a good prognosis association could be based on the absence of HPV-positive HNSCC which show overall higher CD8 expression and CD8⁺ T-cell signature scores in the TCGA cohort (Supplementary Figure S9A). Within the HPV-positive group, high CD8 expression is strongly associated with good prognosis (Supplementary Figure S9B). Notably, CD8A/B expression and CD8⁺ T-cell signature values do not correlate well (Supplementary Figure S10). Interestingly, the observed outcome associations in HPV-negative HNSCC appear to be dependent on cumulative cisplatin dose (Supplementary Figure S10).

The obvious divergence in the biomarker associations with local treatment outcomes compared to DM development risks prompted us to investigate this further. Unable to classify patients according to ‘true’ biological parameter classes, we relied on bootstrapping methods to provide cut offs for each outcome endpoints. As described above those resulted to be largely different in some cases such as for EGFR expression and pointed to a different influence in the respective biological mechanisms. We therefore compared the biological markers with respect to their influence in locoregional control or DM risk in a less cutoff-dependent manner by computing the AUC of the hazard ratio plots from multivariable regression analysis with clinical variables (Supplementary Figure S11). Figure 5 shows an overview of the impact of the individual biological parameters on locoregional control or distant metastasis risk. This analysis highlights the difference in the collection of the most relevant survival determinants for each outcome endpoint. Notably, locoregional control is mainly determined by chronic hypoxia, but also acute hypoxia. CD44 expression and CD8⁺ T-cell/Treg ratio are more relevant to LRC than DM, whereas distant metastasis is predominantly influenced by EMT, acute hypoxia, proliferation and EGFR status.

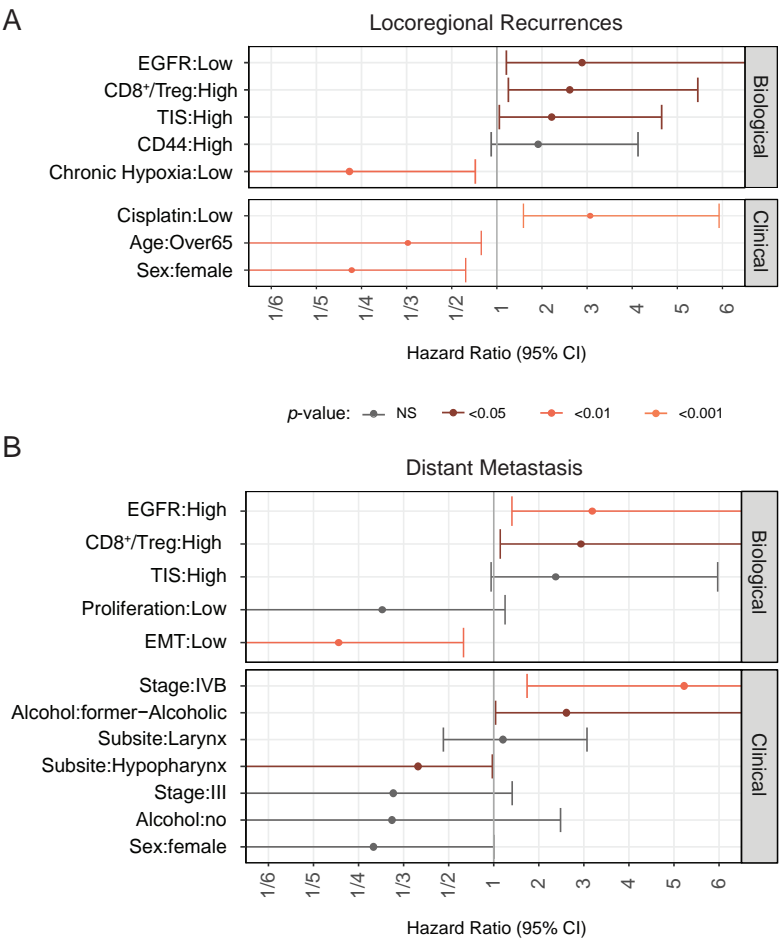


Figure 4. Relative role of biological markers in combined analyses. Forest plots with results from full multivariate Cox Proportional Hazards models are shown. The model was generated using a backward selection procedure. The most frequent level was used as the reference level for this analysis. A Cox proportional hazard model was fit that included all biological markers and clinical variables. Then, each variable was individually eliminated from the model and improvements in model performance were assessed. This process was repeated with the best performing models until the removal of variables did no longer improve the models. Hazard ratios for locoregional recurrences **(A)** and distant metastasis **(B)** as determined by the final model are shown.

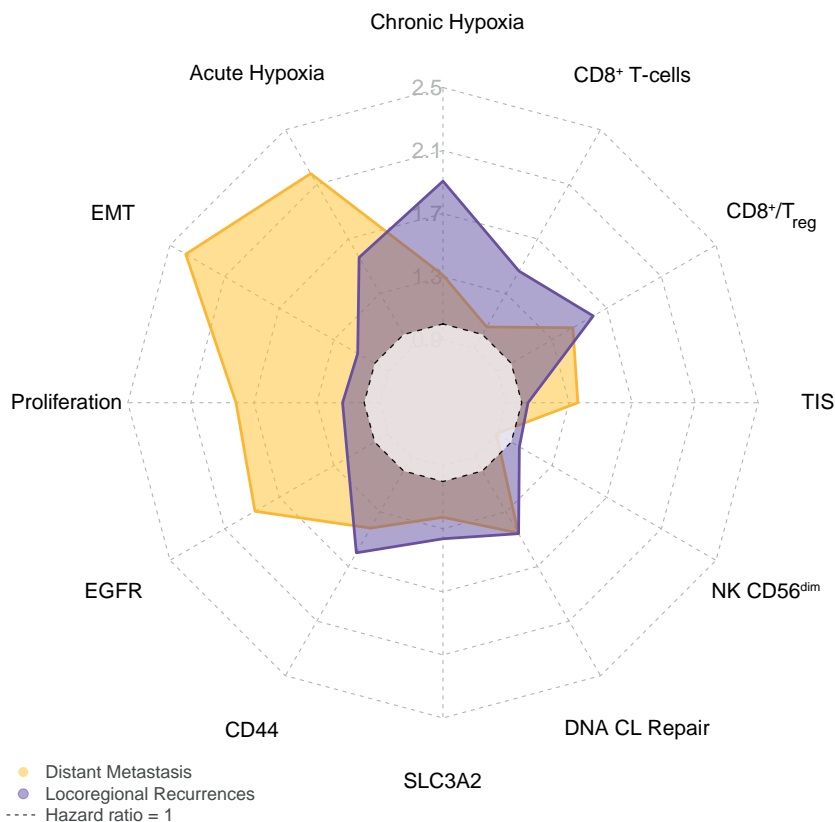


Figure 5. Divergence in biological parameter role for locoregional or distant control. The spider plot depicts the average hazard ratio as obtained after testing all possible cutoffs in a multivariable Cox proportional hazard analysis that included relevant clinical factors as shown in Supplementary Figure S8 and described in Materials and Methods. Hazard ratios for locoregional recurrences are shown in purple and distant metastasis hazard ratios in yellow.

Discussion

Here, we aimed to evaluate the relevance and interrelation of biological factors known to influence radiation response as determined in preclinical studies. Limited by the size of the study cohort, we restricted the study to markers for which discriminative power has been reported in clinical data in HNSCC and added clinical or biological factors that have shown an important association with radiotherapy outcomes. Using RNA-sequencing data from a large and relative uniform cohort of 197 HPV-negative advanced stage HNSCC patients that were all treated with cisplatin-based chemo-radiotherapy, we find an important role of immune cell (T cell) markers in locoregional control which suggests a role in radiation response. We also show that chronic and acute hypoxia are robustly associated with locoregional control.

Similarly, we validated the equally important role of CD44 and SCL3A2, in part related to stem cells, in our study cohort. When assessed in combination, hypoxia, immune cells, EGFR are the most discriminating independent factors in LRC. For DM, those also include EMT. Overall, considered in the context of clinical factors and each other, our study underscores the relevance of many of these biological factors in HNSCC chemo-radiotherapy outcomes.

To advance previous findings on determinants of chemo-radiotherapy outcomes and to prioritize prognostic markers for multi-parametric prediction models, we focused on (i) the validation of expression-based prognostic markers in the chemo-radiotherapy setting and (ii) the evaluation of their complementarity and iii) the assessment of any dependence to important clinical or other biologic factors. Many HNSCC studies highlight the role of biology for outcome^{3-6,78}. A major drawback however for many of such studies is the heterogeneity of the HNSCC patient cohorts or a lack of contextual analyses⁷⁸. If focused on tumor site they often encompass many different treatments or if focused on treatment they combine different tumor sites, HPV-negative and positive. Large multicentric studies are therefore valuable contributions to the field^{12,13,15,16,57} that provide insights to the biology of HNSCC and its link to patient outcome^{5,6,27,28,79,80}. Clinical factors are important^{81,82} but are often not considered in multivariable analyses^{78,83}. A lack of tumor volume data for example, even though clearly linked to LRC^{81,84-86}, impedes the assessment of a role for or a bias from tumor volumes in such analyses. To minimize such treatment or tumor site related bias due to possible interactions; we deliberately excluded the biologically distinct oral cavity and HPV-positive oropharyngeal HNSCC^{14,16,87-90} in our study. Our cohort also solely comprises definitive chemo-radiotherapy-treated advanced HNSCC. In contrast to predetermined gene sets in nanostring technologies, the availability of full transcriptomic data by RNA-Seq allowed us to test selected gene expressions and signatures related to the biological processes that we queried. Together we were able to show that most of the selected markers or marker categories are not related, are independently linked to outcome and that outcome associations are not based on links with known important clinical factors. Overall, we observed little or no influence or interactions with clinical factors, with the notable exception of tumor volume and cumulative cisplatin dose, factors often not accounted for in other biomarker studies. While correlations within the immune markers were expected, here we reveal an association with acute hypoxia scores which in turn appears to be linked to proliferation. Such relations or complementarities can alter the prognostic value or impede a discrimination of the true source of the observed outcome relevance. It however highlights the importance to study such markers in the context of each other and within the same cohort. Overall, our study pinpoints expression markers that should be considered as valuable contributors of future multi-parametric prediction models that combine clinical,

radiologic, pathological and genetic variables for improved prognosis in advanced HPV-negative HNSCC^{91,92}. It is difficult to discern factors that determine tumor radioresistance⁸³. A comparison of similar patient cohorts treated without or with different doses of radiotherapy would be required to strengthen such a link. Since cisplatin-based chemo-radiotherapy has become a standard treatment for HNSCC to improve quality of life by achieving organ preservation, surgically resected HNSCC patients with similar clinical tumor characteristics are rare, impeding such comparisons. However, in the absence of a comparable but non-radiotherapy treated study cohort, differences in LRC (mostly achieved by radiotherapy) as defined by the biomarker classification, can suggest a role in radiation response. In our study, we assured that important clinical factors that impact patient outcomes have been considered to limit bias or dependence. DM events may have occurred prior to LRC events and could have masked a greater impact in radiation response, such as in the case of acute hypoxia that also shows a strong association with DM. The comparison of LRC with DM further helped to discern a more radiation response specific role from a role in metastasis. Our data here and those reported by us and others do indeed confirm the role of hypoxia in determining radiation response as reflected by the LRC rates^{25,27,29}. In addition, hypoxia has been also implicated in tumor cell invasiveness, facilitating dissemination, and has been therefore associated with metastasis formation, a role that is also evident from our DM analyses. Similarly, we have recently shown that HNSCC cell lines with DNA crosslink repair defects are more migratory and invasive⁶³, a feature that may explain the association with DM prognosis but could also result in a greater regional spread and failure of locoregional control.

After the initial EGFR studies in clinic and the success of cetuximab combinations^{40,93,94}, cetuximab in HNSCC and the role of EGFR amplification and expression have been disputed since then⁹⁵. Most of these studies focused on the very high expressing or used a median cutoff to detect an association with the clinical endpoints analyzed. Here we see a clear role for EGFR in the outcome data when also considering hypoxia and other factors in multivariable analyses. Average to high EGFR expression, is linked to improved LRC when analyzed individually. The association of a low EGFR expressing group with poor LRC however becomes much clearer in the combined multivariable analyses that integrated all relevant biomarkers. It epitomizes the importance of combined analysis, as the prevalence of other, also clinical, factors in the different EGFR expression classified groups may have shifted or masked a possible influence in other studies if not accounted for, as revealed here. Given its role in promoting cell cycle progression, it is conceivable that increased EGFR levels mediate an increase in tumor repopulation between fractions; a radiotherapy response determining process that is counteracted by radiotherapy treatment acceleration or concurrent

chemotherapy. This process is therefore limited in our patient population in contrast to some earlier studies that analyzed the influence of EGFR⁹⁶. Notably, the association with improved LRC is still discernible (HR=0.57, $p=0.067$) when reanalyzing the data using the higher EGFR expression cutoff that was used for the DM data. DM HR values however drop to 0.57 ($p=0.2$) showing a DM link only in the top 25% EGFR expression group (HR=3.19, $p=0.0056$). This more aggressive nature of highly EGFR expressing tumors is consistent with other reports in HNSCC and other cancer types⁹⁷.

Our study is limited by statistical constraints due to the cohort size. This enforced us to limit the biological variables and apply selection processes such as the bootstrapping analyses. Yet, it becomes evident that the prognostic value of many of the factors could be validated in our cohort and withstood multivariable analyses with the important clinical variables. Among the clinical variables, we observe a trend towards poor outcomes in current smokers, however this does not reach significance in our cohort. Low numbers in the former smoker category but also the lack of more accurate smoking status values may have decreased the power to reveal the reported association with smoking^{79,98,99}. Since we focused on known determinants of radiation response, other biomarkers were not included despite their relevance or prognostic value in HNSCC^{5,6,16,100–106}. Some, such as tumor mutational burden (TMB) are prevalent in laryngeal and HPV-negative pharyngeal HNSCC¹⁴ but require DNA sequencing data. TMB was found to be associated with poor prognosis in HPV-negative chemo-radiotherapy treated patients in our previous study³⁶ and more strongly so in a cohort of patients that also included oral cavity cancers and HPV-positive oropharyngeal¹⁰⁷. Interestingly, low immune cell infiltration or CD8+ T cell values, as assessed by gene expression, have been assigned to HNSCC high in TMB or mutational signatures related to smoking^{56,107}.

Other limitations result from technical challenges. Here we detect different biological processes and factors in clinical samples by using published and validated expression signatures - that are linked to these processes. These gene expression signatures may not be perfect identification tools for the specific biology in question⁸³; however they often reflect the abundance of certain biological elements well^{108,109}. The DNA CL repair defect prediction model has for example been generated using functional endpoints and then validated in independent cell line panels or by genetic modification. On the other hand, markers such as CD44 are less clear defined. CD44 expression is associated with stem cell-ness in tumor cells¹¹⁰, but it is also expressed under hypoxic conditions or in epithelial cells and is a marker for effector memory T cells⁴⁵. Therefore it is particularly interesting to observe the correlation with SLC3A2 another stem cell related marker in our samples which confirms its link to

tumor stem cell abundance. Notably, we find a correlation between acute hypoxia and TIS or CD8⁺ T-cell scores, suggesting a higher T-cell content in acute hypoxic areas or tumors which could be proposed to be driven by hypoxia induced inflammatory cytokine release¹¹¹. This T-cell/acute hypoxia correlation may in part also be responsible for the consistent poor outcome association of the CD8⁺ T-cell gene expression signatures. Reiterating the role of technical limitations, it should be noted that these gene expression signatures were based on transcriptional profiles of purified immune cell subsets. Through multiple adaptation steps, they evolved to markers that allowed further discrimination in the context of colorectal carcinoma and HNSCC^{56,75,76}. In terms of identification accuracy there are potential challenges with such technical approaches that can also explain discrepancies with immunohistochemistry determined factors. It is evident that the tumor context affects gene expression of the immune cells and, on the other hand, tumor gene expression features, if present in these signatures, can compound the identification. For instance, the CD8⁺ T cell signature includes ZEB1 expression, a protein involved in EMT and a poor prognostic factor in HNSCC^{56,76,112–115}. We therefore assessed CD8A and B gene expression in our samples as a simple surrogate for CD8⁺ T cells and show its limited complementarity with the CD8⁺ T cell signature score and associations with outcome. The better LRC outcome of patients with CD8 positive tumors in the low cumulative cisplatin patient category is in line with previous report based on IHC¹¹⁶. The lack of an association with outcome in patients that received high cisplatin doses however demonstrates treatment dependence and explains the discrepancy to other studies^{6,77} when considering this clinical variable in our cisplatin treated cohort. Despite a significant but weak correlation with CD8A expression, high CD8⁺ T cell signature values are associated with poor outcomes, demonstrating the influence of the other features in this discriminating signature. Immune cell identification by gene expression may not be flawless. Yet, together, our data indicate a prognosis association that is linked to this particular patient treatment. One could speculate that hematologic toxicities associated with cisplatin administration could contribute to this pattern by abolishing the benefit from an immune cell rich tumor status in these individuals. On the other hand, recent studies suggest an enhancement of antitumor immunity by cisplatin that could also diminish the impact of the pre-treatment immune status^{117,118}. While the primary emphasis for prognostic biomarkers lays in the discriminatory power to predict patient outcome, the focus of biomarkers for targeting opportunities is the achievement of an accurate representation of the marked biological process or elements. The signatures used here were selected based on their reported association with both immune cell infiltration and prognosis in HNSCC^{56,76}. Yet the question remains whether they reflect CD8⁺ T cell infiltration well.

Interestingly, we did find a seemingly independent and consistent role for CD8⁺, non-regulatory, T cells in our study cohort. Observed for resected HNSCC in overall survival outcome data before, here we show an association with both, LRC and DM, in chemo-radiotherapy treated HNSCC patients indicating those with a high abundance of such T cells to have a worse prognosis. To our knowledge this poor prognosis association with radiation response has not been reported previously^{6,77,119}. As detailed above, this discrepancy with other studies is only in part explained by the used technology^{116,120} (IHC CD8 expression versus gene expression signatures) since the signatures showed a good prognosis association in the Mandal *et al.* study⁵⁶. Careful inspection of the TCGA data revealed increased CD8⁺ T-cell gene expression signature scores in the HPV-positive oropharyngeal that drive the good prognosis association. A pattern observed in other studies as well^{56,77,121–124}. Mandal *et al.* adjusted for HPV-associated outcome differences, which does not account for a possible interaction between the two variables⁵⁶. The CD8A and B expression HR plots in our analyses however suggest a stronger effect in the HPV-positive subgroup. Despite obvious evaluation challenges when using the different techniques and associated cutoffs, a similar argument applies to other studies based on immunohistochemistry determined CD8⁺ T cell infiltration values. A significant HPV status association got lost in multivariable analyses that indicated a good prognosis association of CD8⁺ T cells in oropharyngeal squamous cell carcinoma patients⁵⁹. Yet, some studies also show a good prognosis association with TIS or CD8⁺ T cells in HPV-negative patients using other scorings, cut offs and expression signatures¹²⁵. Since the effect size can be small, patient treatment associations with survival are often not significant in small studies. Treatment could however alter prognosis in subsets of patients. For instance, patients with tumors with DNA crosslink repair defects benefit most from a high cumulative cisplatin dose⁶³. Similarly, possible immune cell infiltration links could depend on treatment. Despite worse PFS in cases that lack or show minimal CD8A or CD8B expression, we cannot observe the previously reported poor prognosis link in CD8⁺ T cell signature low patients in our cohort. No associations between TIS or CD8⁺ scores and clinical variables were found; and outcome association links derived from the correlation with acute hypoxia should have been accounted for by the multivariable analyses. Together, our data suggest a role for HNSCC treatment, in particular cisplatin, in immune cell infiltration determined outcomes. Early cancer immunotherapy trials in HNSCC with immune checkpoint inhibitors demonstrate a benefit and underscore the potential value of immune response and chemo-radiotherapy relevant biomarkers to identify patients that will benefit from such treatments^{24,126–134}. Larger comparative studies are therefore needed to disentangle the role of CD8⁺ T cells in the individual genetic HNSCC context and the important clinical variables connected to its role in patient outcome^{78,83,135}.

Our patient cohort is fairly unique in that it consists of definitive chemoradiotherapy treated advanced HNSCC patients. Based exclusively on resected HNSCC, these cases are unfortunately not present in the TCGA data. Supported by the detailed clinical data and follow up, this allowed us to elaborate on the role of biological determinants of chemoradiotherapy response. A quarter of the patients suffered from loco regional recurrences after treatment; a treatment success rate that further stresses the relevance of the biological factors found to determine treatment failure. This study does not provide or test clinically applicable prognostic markers. It was designed to compare the individual factors in relation to each other to assess and understand their influence in HNSCC outcome. Optimal cutoffs identified by the bootstrapping method and illustrated in the hazard ratio plots require validation for further development into true prognostic markers. Based on our results, future studies should focus on the elaboration of prognostic models that incorporate these biological markers together with important clinical factors. The multivariable outcome association results and the lack of correlations suggest that these future models should include all biological factors. Discrepancy in the optimal cutoff values further points to the value of non-dichotomized variables in such efforts and also reveals a possible cause of incongruent outcome associations in previous studies. The value of the clinical factors is exemplified by the fact that some biological markers (i.e. DNA CL repair or CD8⁺ T cells) lose their strength in patients groups with a high cumulative cisplatin dose.

While tumor stem cell targeting agents are still under development, some of the other biological factors are targetable. Next to high-dose alkylating agents, PARP inhibitors may help to exploit DNA CL repair defects⁶² and different immunotherapy options are currently being tested in the HNSCC setting¹³⁶. The value of such biological markers in personalized treatments remains to be determined; however our study demonstrates that those patients are in need of improved therapy options.

In conclusion, this multicentric external validation study confirms the important and independent role of biological factors that embody hypoxia, stem cell-ness, tumor growth, EMT and DNA repair for locoregional control in chemoradiotherapy treated patients. The multifactorial analyses results highlight the need to consider these biomarkers in the context of each other and also revealed an important role for immune cell abundance in HNSCC treatment outcome.

References

1. Baumann M, Krause M, Overgaard J, *et al.* Radiation oncology in the era of precision medicine. *Nat Rev Cancer*. 2016. doi:10.1038/nrc.2016.18
2. van der Kogel A. *Basic Clinical Radiobiology Fourth Edition.*; 2009. doi:10.1201/b13224
3. Alsahafi E, Begg K, Amelio I, *et al.* Clinical update on head and neck cancer: molecular biology and ongoing challenges. *Cell Death Dis*. 2019;10(8):540. doi:10.1038/s41419-019-1769-9
4. Leemans CR, Braakhuis BJM, Brakenhoff RH. The molecular biology of head and neck cancer. *Nat Rev Cancer*. 2011;11(1):9-22. doi:10.1038/nrc2982
5. Economopoulou P, de Bree R, Kotsantis I, Psyrri A. Diagnostic Tumor Markers in Head and Neck Squamous Cell Carcinoma (HNSCC) in the Clinical Setting. *Front Oncol*. 2019;9:827. doi:10.3389/fonc.2019.00827
6. Budach V, Tinhofer I. Novel prognostic clinical factors and biomarkers for outcome prediction in head and neck cancer: a systematic review. *Lancet Oncol*. 2019;20(6):e313-e326. doi:10.1016/S1470-2045(19)30177-9
7. Kang H, Kiess A, Chung CH. Emerging biomarkers in head and neck cancer in the era of genomics. *Nat Rev Clin Oncol*. 2015;12(1):11-26. doi:10.1038/nrclinonc.2014.192
8. Petersen JF, Timmermans AJ, van Dijk BAC, *et al.* Trends in treatment, incidence and survival of hypopharynx cancer: a 20-year population-based study in the Netherlands. *Eur Arch Oto-Rhino-Laryngology*. 2018;275(1):181-189. doi:10.1007/s00405-017-4766-6
9. Timmermans AJ, van Dijk BAC, Overbeek LIH, *et al.* Trends in treatment and survival for advanced laryngeal cancer: A 20-year population-based study in The Netherlands. *Head Neck*. 2015. doi:10.1002/hed.24200
10. Price KAR, Cohen EE. Current treatment options for metastatic head and neck cancer. *Curr Treat Options Oncol*. 2012. doi:10.1007/s11864-011-0176-y
11. Huang SH, O'Sullivan B. Overview of the 8th Edition TNM Classification for Head and Neck Cancer. *Curr Treat Options Oncol*. 2017. doi:10.1007/s11864-017-0484-y
12. Lawrence MS, Sougnez C, Lichtenstein L, *et al.* Comprehensive genomic characterization of head and neck squamous cell carcinomas. *Nature*. 2015;517(7536):576-582. doi:10.1038/nature14129
13. Stransky N, Egloff AM, Tward AD, *et al.* The mutational landscape of head and neck squamous cell carcinoma. *Science (80-)*. 2011;333(6046):1157-1160. doi:10.1126/science.1208130
14. Vossen DM, Verhagen CVM, Verheij M, Wessels LFA, Vens C, van den Brekel MWM. Comparative genomic analysis of oral versus laryngeal and pharyngeal cancer. *Oral Oncol*. 2018. doi:10.1016/j.oraloncology.2018.04.006
15. Agrawal N, Frederick MJ, Pickering CR, *et al.* Exome sequencing of head and neck squamous cell carcinoma reveals inactivating mutations in NOTCH1. *Science (80-)*. 2011;333(6046):1154-1157. doi:10.1126/science.1206923
16. Dogan S, Xu B, Middha S, *et al.* Identification of prognostic molecular biomarkers in 157 HPV-positive and HPV-negative squamous cell carcinomas of the oropharynx. *Int J cancer*. May 2019. doi:10.1002/ijc.32412
17. Theelen WSME, Peulen HMU, Lalezari F, *et al.* Effect of Pembrolizumab After Stereotactic Body Radiotherapy vs Pembrolizumab Alone on Tumor Response in Patients With Advanced Non-Small Cell Lung Cancer. *JAMA Oncol*. 2019;1-7. doi:10.1001/jamaoncol.2019.1478
18. Ngwa W, Irabor OC, Schoenfeld JD, Hesser J, Demaria S, Formenti SC. Using immunotherapy to boost the abscopal effect. *Nat Rev Cancer*. 2018. doi:10.1038/nrc.2018.6

19. Tinhofer I, Budach V, Johrens K, Keilholz U. The rationale for including immune checkpoint inhibition into multimodal primary treatment concepts of head and neck cancer. *Cancers head neck*. 2016;1:8. doi:10.1186/s41199-016-0009-6
20. Bristow RG, Alexander B, Baumann M, *et al*. Combining precision radiotherapy with molecular targeting and immunomodulatory agents: a guideline by the American Society for Radiation Oncology. *Lancet Oncol*. 2018;19(5):e240-e251. doi:10.1016/S1470-2045(18)30096-2
21. Swart M, Verbrugge I, Beltman JB. Combination Approaches with Immune-Checkpoint Blockade in Cancer Therapy. *Front Oncol*. 2016;6:233. doi:10.3389/fonc.2016.00233
22. Canning M, Guo G, Yu M, *et al*. Heterogeneity of the Head and Neck Squamous Cell Carcinoma Immune Landscape and Its Impact on Immunotherapy. *Front cell Dev Biol*. 2019;7:52. doi:10.3389/fcell.2019.00052
23. Economopoulou P, Agelaki S, Perisanidis C, Giotakis EI, Psyrris A. The promise of immunotherapy in head and neck squamous cell carcinoma. *Ann Oncol Off J Eur Soc Med Oncol*. 2016;27(9):1675-1685. doi:10.1093/annonc/mdw226
24. Solomon B, Young RJ, Rischin D. Head and neck squamous cell carcinoma: Genomics and emerging biomarkers for immunomodulatory cancer treatments. *Semin Cancer Biol*. 2018;52(Pt 2):228-240. doi:10.1016/j.semcancer.2018.01.008
25. Toustrup K, Sørensen BS, Nordsmark M, *et al*. Development of a hypoxia gene expression classifier with predictive impact for hypoxic modification of radiotherapy in head and neck cancer. *Cancer Res*. 2011. doi:10.1158/0008-5472.CAN-11-1182
26. Tawk B, Schwager C, Deffaa O, *et al*. Comparative analysis of transcriptomics based hypoxia signatures in head- and neck squamous cell carcinoma. *Radiother Oncol*. 2016;118(2):350-358. doi:10.1016/j.radonc.2015.11.027
27. 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. *Radiother Oncol*. 2016;121(3):364-373. doi:10.1016/j.radonc.2016.11.008
28. Linge A, Lock S, Gudziol V, *et al*. Low Cancer Stem Cell Marker Expression and Low Hypoxia Identify Good Prognosis Subgroups in HPV(-) HNSCC after Postoperative Radiochemotherapy: A Multicenter Study of the DTK-ROG. *Clin Cancer Res*. 2016;22(11):2639-2649. doi:10.1158/1078-0432.CCR-15-1990
29. van der Heijden M, de Jong MC, Verhagen CVM, *et al*. Acute Hypoxia Profile is a Stronger Prognostic Factor than Chronic Hypoxia in Advanced Stage Head and Neck Cancer Patients. *Cancers (Basel)*. 2019. doi:10.3390/cancers11040583
30. Rademakers SE, Hoogsteen IJ, Rijken PF, *et al*. Pattern of CAIX expression is prognostic for outcome and predicts response to ARCON in patients with laryngeal cancer treated in a phase III randomized trial. *Radiother Oncol*. 2013. doi:10.1016/j.radonc.2013.04.022
31. Thomson D, Yang H, Baines H, *et al*. NIMRAD - A phase III trial to investigate the use of nimorazole hypoxia modification with intensity-modulated radiotherapy in head and neck cancer. *Clin Oncol*. 2014. doi:10.1016/j.clon.2014.03.003
32. Kaanders JHAM, Bussink J, Van der Kogel AJ. ARCON: A novel biology-based approach in radiotherapy. *Lancet Oncol*. 2002. doi:10.1016/S1470-2045(02)00929-4
33. Overgaard J. Hypoxic modification of radiotherapy in squamous cell carcinoma of the head and neck - A systematic review and meta-analysis. *Radiother Oncol*. 2011. doi:10.1016/j.radonc.2011.03.004

34. Overgaard J, Hansen HS, Specht L, *et al.* Five compared with six fractions per week of conventional radiotherapy of squamous-cell carcinoma of head and neck: DAHANCA 6&7 randomised controlled trial. *Lancet*. 2003. doi:10.1016/S0140-6736(03)14361-9
35. Bentzen SM, Thames HD. Clinical evidence for tumor clonogen regeneration: interpretations of the data. *Radiother Oncol*. 1991. doi:10.1016/0167-8140(91)90019-D
36. Vossen DM, Verhagen CVM, Van Der Heijden M, *et al.* Genetic factors associated with a poor outcome in head and neck cancer patients receiving definitive chemoradiotherapy. *Cancers (Basel)*. 2019. doi:10.3390/cancers11040445
37. Ang KK. Multidisciplinary Management of Locally Advanced SCCHN: Optimizing Treatment Outcomes. *Oncologist*. 2008. doi:10.1634/theoncologist.2007-0157
38. Bonner JA, Harari PM, Giralt J, *et al.* Radiotherapy plus cetuximab for locoregionally advanced head and neck cancer: *Lancet Oncol*. 2010. doi:10.1016/S1470-2045(09)70311-0
39. Bentzen SM, Atasoy BM, Daley FM, *et al.* Epidermal growth factor receptor expression in pretreatment biopsies from head and neck squamous cell carcinoma as a predictive factor for a benefit from accelerated radiation therapy in a randomized controlled trial. *J Clin Oncol*. 2005. doi:10.1200/JCO.2005.06.411
40. Bossi P, Resteghini C, Paielli N, Licitra L, Pilotti S, Perrone F. Prognostic and predictive value of EGFR in head and neck squamous cell carcinoma. *Oncotarget*. 2016. doi:10.18632/oncotarget.11413
41. Xu MJ, Johnson DE, Grandis JR. EGFR-targeted therapies in the post-genomic era. *Cancer Metastasis Rev*. 2017;36(3):463-473. doi:10.1007/s10555-017-9687-8
42. Juergens RA, Bratman S V, Tsao M-S, *et al.* Biology and patterns of response to EGFR-inhibition in squamous cell cancers of the lung and head & neck. *Cancer Treat Rev*. 2017;54:43-57. doi:10.1016/j.ctrv.2017.01.003
43. Taberna M, Oliva M, Mesia R. Cetuximab-Containing Combinations in Locally Advanced and Recurrent or Metastatic Head and Neck Squamous Cell Carcinoma. *Front Oncol*. 2019;9:383. doi:10.3389/fonc.2019.00383
44. Starmans MHW, Krishnapuram B, Steck H, *et al.* Robust prognostic value of a knowledge-based proliferation signature across large patient microarray studies spanning different cancer types. *Br J Cancer*. 2008. doi:10.1038/sj.bjc.6604746
45. De Jong MC, Pramana J, Van Der Wal JE, *et al.* CD44 expression predicts local recurrence after radiotherapy in larynx cancer. *Clin Cancer Res*. 2010;16(21):5329-5338. doi:10.1158/1078-0432.CCR-10-0799
46. Linge A, Schmidt S, Lohaus F, *et al.* Independent validation of tumour volume, cancer stem cell markers and hypoxia-associated gene expressions for HNSCC after primary radiochemotherapy. *Clin Transl Radiat Oncol*. 2019;16:40-47. doi:10.1016/j.ctr.2019.03.002
47. Baumann M, Krause M, Hill R. Exploring the role of cancer stem cells in radioresistance. *Nat Rev Cancer*. 2008. doi:10.1038/nrc2419
48. Krause M, Dubrovskaya A, Linge A, Baumann M. Cancer stem cells: Radioresistance, prediction of radiotherapy outcome and specific targets for combined treatments. *Adv Drug Deliv Rev*. 2017;109:63-73. doi:10.1016/j.addr.2016.02.002
49. Krause M, Yaromina A, Eicheler W, Koch U, Baumann M. Cancer stem cells: targets and potential biomarkers for radiotherapy. *Clin Cancer Res*. 2011;17(23):7224-7229. doi:10.1158/1078-0432.CCR-10-2639
50. Chen H yan, Xu L, Li L feng, Liu X xing, Gao J xin, Bai Y rui. Inhibiting the CD8 + T cell infiltration in the tumor microenvironment after radiotherapy is an important mechanism of radioresistance. *Sci Rep*. 2018. doi:10.1038/s41598-018-30417-6

51. Gupta A, Probst HC, Vuong V, *et al.* Radiotherapy Promotes Tumor-Specific Effector CD8 + T Cells via Dendritic Cell Activation . *J Immunol.* 2012. doi:10.4049/jimmunol.1200563
52. Ran X, Yang K. Inhibitors of the PD-1/PD-L1 axis for the treatment of head and neck cancer: current status and future perspectives. *Drug Des Devel Ther.* 2017;11:2007-2014. doi:10.2147/DDDT.S140687
53. Moskovitz J, Moy J, Ferris RL. Immunotherapy for Head and Neck Squamous Cell Carcinoma. *Curr Oncol Rep.* 2018;20(2):22. doi:10.1007/s11912-018-0654-5
54. Economopoulou P, Kotsantis I, Psyrri A. Checkpoint Inhibitors in Head and Neck Cancer: Rationale, Clinical Activity, and Potential Biomarkers. *Curr Treat Options Oncol.* 2016;17(8):40. doi:10.1007/s11864-016-0419-z
55. Van Limbergen EJ, De Ruyscher DK, Pimentel VO, *et al.* Combining radiotherapy with immunotherapy: The past, the present and the future. *Br J Radiol.* 2017. doi:10.1259/bjr.20170157
56. Mandal R, Şenbabaoğlu Y, Desrichard A, *et al.* The head and neck cancer immune landscape and its immunotherapeutic implications. *JCI Insight.* 2016;1(17):1-18. doi:10.1172/jci.insight.89829
57. Balermipas P, Rodel F, Rodel C, *et al.* CD8+ tumour-infiltrating lymphocytes in relation to HPV status and clinical outcome in patients with head and neck cancer after postoperative chemoradiotherapy: A multicentre study of the German cancer consortium radiation oncology group (DKTK-ROG). *Int J cancer.* 2016;138(1):171-181. doi:10.1002/ijc.29683
58. Ono T, Azuma K, Kawahara A, *et al.* Pre-treatment CD8(+) tumour-infiltrating lymphocyte density predicts distant metastasis after definitive treatment in patients with stage III/IV hypopharyngeal squamous cell carcinoma. *Clin Otolaryngol.* 2018;43(5):1312-1320. doi:10.1111/coa.13171
59. De Meulenaere A, Vermassen T, Aspeslagh S, *et al.* Tumor PD-L1 status and CD8(+) tumor-infiltrating T cells: markers of improved prognosis in oropharyngeal cancer. *Oncotarget.* 2017;8(46):80443-80452. doi:10.18632/oncotarget.19045
60. De Meulenaere A, Vermassen T, Aspeslagh S, *et al.* Prognostic markers in oropharyngeal squamous cell carcinoma: focus on CD70 and tumour infiltrating lymphocytes. *Pathology.* 2017;49(4):397-404. doi:10.1016/j.pathol.2017.02.002
61. Verhagen CVM, Vossen DM, Borgmann K, *et al.* Fanconi anemia and homologous recombination gene variants are associated with functional DNA repair defects in vitro and poor outcome in patients with advanced head and neck squamous cell carcinoma. *Oncotarget.* 2018;9(26):18198-18213. doi:10.18632/oncotarget.24797
62. Verhagen CVM, De Haan R, Hageman F, *et al.* Extent of radiosensitization by the PARP inhibitor olaparib depends on its dose, the radiation dose and the integrity of the homologous recombination pathway of tumor cells. In: *Radiotherapy and Oncology.* ; 2015. doi:10.1016/j.radonc.2015.03.028
63. Essers PBM, van der Heijden M, Verhagen CVM, *et al.* Drug sensitivity prediction models reveal link between DNA repair defects and poor prognosis in HNSCC. *Cancer Res.* 2019.
64. Van der Heijden M, Essers P, Verheij M, Van den Brekel M, Vens C. OC-0487: EMT signatures as a prognostic marker for metastasis in HPV-negative HNSCC. *Radiother Oncol.* 2018;127:S250-S251. doi:10.1016/S0167-8140(18)30797-7
65. Balm AJM, Rasch CRN, Schornagel JH, *et al.* High-dose superselective intra-arterial cisplatin and concomitant radiation (radplat) for advanced head and neck cancer. *Head Neck.* 2004;26(6):485-493. doi:10.1002/hed.20006
66. 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 Oto-Rhino-Laryngology.* 2017. doi:10.1007/s00405-017-4687-4

67. Spreafico A, Huang SH, Xu W, *et al.* Impact of cisplatin dose intensity on human papillomavirus-related and -unrelated locally advanced head and neck squamous cell carcinoma. *Eur J Cancer*. 2016. doi:10.1016/j.ejca.2016.08.013
68. Heukelom J, Hamming O, Bartelink H, *et al.* Adaptive and innovative Radiation Treatment FOR improving Cancer treatment outcome (ARTFORCE); a randomized controlled phase II trial for individualized treatment of head and neck cancer. *BMC Cancer*. 2013. doi:10.1186/1471-2407-13-84
69. Lewis JS, Chernock RD, Ma XJ, *et al.* Partial p16 staining in oropharyngeal squamous cell carcinoma: Extent and pattern correlate with human papillomavirus RNA status. *Mod Pathol*. 2012. doi:10.1038/modpathol.2012.79
70. Kim D, Pertea G, Trapnell C, Pimentel H, Kelley R, Salzberg SL. TopHat2: Accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. *Genome Biol*. 2013. doi:10.1186/gb-2013-14-4-r36
71. Anders S, Pyl PT, Huber W. HTSeq-A Python framework to work with high-throughput sequencing data. *Bioinformatics*. 2015;31(2):166-169. doi:10.1093/bioinformatics/btu638
72. Hänzelmann S, Castelo R, Guinney J. GSEA: Gene set variation analysis for microarray and RNA-Seq data. *BMC Bioinformatics*. 2013;14. doi:10.1186/1471-2105-14-7
73. Seigneuric R, Starmans MHW, Fung G, *et al.* Impact of supervised gene signatures of early hypoxia on patient survival. *Radiother Oncol*. 2007. doi:10.1016/j.radonc.2007.05.002
74. Lohaus F, Linge A, Tinhofer I, *et al.* HPV16 DNA status is a strong prognosticator of loco-regional control after postoperative radiochemotherapy of locally advanced oropharyngeal carcinoma: Results from a multicentre explorative study of the German Cancer Consortium Radiation Oncology Group (. *Radiother Oncol*. 2014;113(3):317-323. doi:10.1016/j.radonc.2014.11.011
75. Senbabaoglu Y, Winer AG, Gejman RS, *et al.* The Landscape of {T} Cell Infiltration in Human Cancer and Its Association with Antigen Presenting Gene Expression.; 2015. doi:10.1101/025908
76. Bindea G, Mlecnik B, Tosolini M, *et al.* Spatiotemporal dynamics of intratumoral immune cells reveal the immune landscape in human cancer. *Immunity*. 2013;39(4):782-795. doi:10.1016/j.immuni.2013.10.003
77. De Meulenaere A, Vermassen T, Aspeslagh S, Vandecasteele K, Rottey S, Ferdinande L. TILs in Head and Neck Cancer: Ready for Clinical Implementation and Why (Not)? *Head Neck Pathol*. 2017;11(3):354-363. doi:10.1007/s12105-016-0776-8
78. Kim KY, McShane LM, Conley BA. Designing biomarker studies for head and neck cancer. *Head Neck*. 2014;36(7):1069-1075. doi:10.1002/hed.23444
79. Lassen P, Lacas B, Pignon J-P, *et al.* Prognostic impact of HPV-associated p16-expression and smoking status on outcomes following radiotherapy for oropharyngeal cancer: The MARCH-HPV project. *Radiother Oncol*. 2018;126(1):107-115. doi:10.1016/j.radonc.2017.10.018
80. 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-S): a multicentre cohort study. *Lancet Oncol*. 2016;17(4):440-451. doi:10.1016/S1470-2045(15)00560-4
81. Hoebbers F, Rios E, Troost E, *et al.* Definitive radiation therapy for treatment of laryngeal carcinoma: impact of local relapse on outcome and implications for treatment strategies. *Strahlentherapie und Onkol Organ der Dtsch Rontgengesellschaft . [et al]*. 2013;189(10):834-841. doi:10.1007/s00066-013-0414-2
82. Silva P, Homer JJ, Slevin NJ, *et al.* Clinical and biological factors affecting response to radiotherapy in patients with head and neck cancer: a review. *Clin Otolaryngol*. 2007;32(5):337-345. doi:10.1111/j.1749-4486.2007.01544.x

83. Lucs A, Saltman B, Chung CH, Steinberg BM, Schwartz DL. Opportunities and challenges facing biomarker development for personalized head and neck cancer treatment. *Head Neck*. 2013;35(2):294-306. doi:10.1002/hed.21975
84. Welch ML, McIntosh C, Haibe-Kains B, *et al.* Vulnerabilities of radiomic signature development: The need for safeguards. *Radiother Oncol*. 2019;130:2-9. doi:10.1016/j.radonc.2018.10.027
85. Strongin A, Yovino S, Taylor R, *et al.* Primary tumor volume is an important predictor of clinical outcomes among patients with locally advanced squamous cell cancer of the head and neck treated with definitive chemoradiotherapy. *Int J Radiat Oncol Biol Phys*. 2012;82(5):1823-1830. doi:10.1016/j.ijrobp.2010.10.053
86. Knegjens JL, Hauptmann M, Pameijer FA, *et al.* Tumor volume as prognostic factor in chemoradiation for advanced head and neck cancer. *Head Neck*. 2011;33(3):375-382. doi:10.1002/hed.21459
87. Gopinath D, Kunnath Menon R. Unravelling the molecular signatures in HNSCC: Is the homogenous paradigm becoming obsolete? *Oral Oncol*. 2018;82:195. doi:10.1016/j.j.oraloncology.2018.05.011
88. Pai SI, Westra WH. Molecular pathology of head and neck cancer: implications for diagnosis, prognosis, and treatment. *Annu Rev Pathol*. 2009;4:49-70. doi:10.1146/annurev.pathol.4.110807.092158
89. van Monsjou HS, Balm AJM, van den Brekel MM, Wreesmann VB. Oropharyngeal squamous cell carcinoma: a unique disease on the rise? *Oral Oncol*. 2010;46(11):780-785. doi:10.1016/j.oraloncology.2010.08.011
90. Dogan V, Rieckmann T, Münscher A, Busch CJ. Current studies of immunotherapy in head and neck cancer. *Clin Otolaryngol*. 2018. doi:10.1111/coa.12895
91. Caudell JJ, Torres-Roca JF, Gillies RJ, *et al.* The future of personalised radiotherapy for head and neck cancer. *Lancet Oncol*. 2017;18(5):e266-e273. doi:10.1016/S1470-2045(17)30252-8
92. Lambin P, Zindler J, Vanneste BGL, *et al.* Decision support systems for personalized and participative radiation oncology. *Adv Drug Deliv Rev*. 2017. doi:10.1016/j.addr.2016.01.006
93. Cripps C, Winquist E, Devries MC, Stys-Norman D, Gilbert R. Epidermal growth factor receptor targeted therapy in stages III and IV head and neck cancer. *Curr Oncol*. 2010.
94. Grandis JR, Melhem MF, Gooding WE, *et al.* Levels of TGF- α and EGFR protein in head and neck squamous cell carcinoma and patient survival. *J Natl Cancer Inst*. 1998. doi:10.1093/jnci/90.11.824
95. Hutchinson L. Drug therapy: Cetuximab or cisplatin in HNSCC? *Nat Rev Clin Oncol*. 2016. doi:10.1038/nrclinonc.2015.235
96. Ang KK, Berkey BA, Tu X, *et al.* Impact of epidermal growth factor receptor expression on survival and pattern of relapse in patients with advanced head and neck carcinoma. *Cancer Res*. 2002.
97. Nicholson R., Gee JM., Harper M. EGFR and cancer prognosis. *Eur J Cancer*. 2001. doi:10.1016/s0959-8049(01)00231-3
98. Gillison ML, Zhang Q, Jordan R, *et al.* Tobacco smoking and increased risk of death and progression for patients with p16-positive and p16-negative oropharyngeal cancer. *J Clin Oncol*. 2012;30(17):2102-2111. doi:10.1200/JCO.2011.38.4099
99. Gronhoj C, Jensen JS, Wagner S, *et al.* Impact on survival of tobacco smoking for cases with oropharyngeal squamous cell carcinoma and known human papillomavirus and p16-status: a multicenter retrospective study. *Oncotarget*. 2019;10(45):4655-4663. doi:10.18632/oncotarget.27079

100. Hess J, Unger K, Maihoefer C, *et al.* A Five-MicroRNA Signature Predicts Survival and Disease Control of Patients with Head and Neck Cancer Negative for HPV Infection. *Clin Cancer Res.* 2019;25(5):1505-1516. doi:10.1158/1078-0432.CCR-18-0776
101. Glorieux M, Dok R, Nuyts S. Novel DNA targeted therapies for head and neck cancers: clinical potential and biomarkers. *Oncotarget.* 2017;8(46):81662-81678. doi:10.18632/oncotarget.20953
102. Bossi P, Alfieri S, Strojan P, *et al.* Prognostic and predictive factors in recurrent and/or metastatic head and neck squamous cell carcinoma: A review of the literature. *Crit Rev Oncol Hematol.* 2019;137:84-91. doi:10.1016/j.critrevonc.2019.01.018
103. Foy J-P, Bazire L, Ortiz-Cuaran S, *et al.* A 13-gene expression-based radioresistance score highlights the heterogeneity in the response to radiation therapy across HPV-negative HNSCC molecular subtypes. *BMC Med.* 2017;15(1):165. doi:10.1186/s12916-017-0929-y
104. Scott JG, Berglund A, Schell MJ, *et al.* A genome-based model for adjusting radiotherapy dose (GARD): a retrospective, cohort-based study. *Lancet Oncol.* 2017;18(2):202-211. doi:10.1016/S1470-2045(16)30648-9
105. Hsieh JC-H, Wang H-M, Wu M-H, *et al.* Review of emerging biomarkers in head and neck squamous cell carcinoma in the era of immunotherapy and targeted therapy. *Head Neck.* 2019;41 Suppl 1:19-45. doi:10.1002/hed.25932
106. Lim SM, Cho SH, Hwang IG, *et al.* Investigating the Feasibility of Targeted Next-Generation Sequencing to Guide the Treatment of Head and Neck Squamous Cell Carcinoma. *Cancer Res Treat.* 2019;51(1):300-312. doi:10.4143/crt.2018.012
107. Eder T, Hess AK, Kanschak R, *et al.* Interference of tumour mutational burden with outcome of patients with head and neck cancer treated with definitive chemoradiation: a multicentre retrospective study of the German Cancer Consortium Radiation Oncology Group. *Eur J Cancer.* 2019;116:67-76. doi:10.1016/j.ejca.2019.04.015
108. Hoogsteen IJ, Marres HAM, Bussink J, van der Kogel AJ, Kaanders JHAM. Tumor microenvironment in head and neck squamous cell carcinomas: predictive value and clinical relevance of hypoxic markers. A review. *Head Neck.* 2007;29(6):591-604. doi:10.1002/hed.20543
109. Spiegelberg L, Houben R, Niemans R, *et al.* Hypoxia-activated prodrugs and (lack of) clinical progress: The need for hypoxia-based biomarker patient selection in phase III clinical trials. *Clin Transl Radiat Oncol.* 2019;15:62-69. doi:10.1016/j.ctro.2019.01.005
110. Baumann M, Krause M. CD44: A cancer stem cell-related biomarker with predictive potential for radiotherapy. *Clin Cancer Res.* 2010. doi:10.1158/1078-0432.CCR-10-2244
111. Eltzschig HK, Carmeliet P. Hypoxia and Inflammation. *N Engl J Med.* 2011. doi:10.1056/nejmra0910283
112. Zhang P, Sun Y, Ma L. ZEB1: at the crossroads of epithelial-mesenchymal transition, metastasis and therapy resistance. *Cell Cycle.* 2015;14(4):481-487. doi:10.1080/15384101.2015.1006048
113. Schmalhofer O, Brabletz S, Brabletz T. E-cadherin, beta-catenin, and ZEB1 in malignant progression of cancer. *Cancer Metastasis Rev.* 2009;28(1-2):151-166. doi:10.1007/s10555-008-9179-y
114. Chen C, Zimmermann M, Tinhofer I, Kaufmann AM, Albers AE. Epithelial-to-mesenchymal transition and cancer stem(-like) cells in head and neck squamous cell carcinoma. *Cancer Lett.* 2013;338(1):47-56. doi:10.1016/j.canlet.2012.06.013
115. Smith A, Teknos TN, Pan Q. Epithelial to mesenchymal transition in head and neck squamous cell carcinoma. *Oral Oncol.* 2013;49(4):287-292. doi:10.1016/j.oraloncology.2012.10.009
116. De Meulenaere A, Vermassen T, Creyten D, *et al.* Importance of choice of materials and methods in PD-L1 and TIL assessment in oropharyngeal squamous cell carcinoma. *Histopathology.* 2018;73(3):500-509. doi:10.1111/his.13650

117. Tran L, Allen CT, Xiao R, *et al.* Cisplatin Alters Antitumor Immunity and Synergizes with PD-1/PD-L1 Inhibition in Head and Neck Squamous Cell Carcinoma. *Cancer Immunol Res.* 2017;5(12):1141-1151. doi:10.1158/2326-6066.CIR-17-0235
118. Kroon P, Frijlink E, Iglesias-Guimaraes V, *et al.* Radiotherapy and Cisplatin Increase Immunotherapy Efficacy by Enabling Local and Systemic Intratumoral T-cell Activity. *Cancer Immunol Res.* 2019;7(4):670-682. doi:10.1158/2326-6066.CIR-18-0654
119. 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. doi:10.1080/2162402X.2017.1356148
120. Hendry S, Salgado R, Gevaert T, *et al.* Assessing Tumor-Infiltrating Lymphocytes in Solid Tumors: A Practical Review for Pathologists and Proposal for a Standardized Method from the International Immuno-Oncology Biomarkers Working Group: Part 2: TILs in Melanoma, Gastrointestinal Tract Carcinom. *Adv Anat Pathol.* 2017. doi:10.1097/PAP.0000000000000161
121. Spector ME, Bellile E, Amlani L, *et al.* Prognostic Value of Tumor-Infiltrating Lymphocytes in Head and Neck Squamous Cell Carcinoma. *JAMA Otolaryngol Head Neck Surg.* September 2019. doi:10.1001/jamaoto.2019.2427
122. Balermipas P, Rodel F, Krause M, *et al.* The PD-1/PD-L1 axis and human papilloma virus in patients with head and neck cancer after adjuvant chemoradiotherapy: A multicentre study of the German Cancer Consortium Radiation Oncology Group (DKTK-ROG). *Int J cancer.* 2017;141(3):594-603. doi:10.1002/ijc.30770
123. 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. doi:10.1038/bjc.2013.639
124. 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. doi:10.1038/bjc.2015.277
125. Hess A-K, Johrens K, Zakarneh A, *et al.* Characterization of the tumor immune micromilieu and its interference with outcome after concurrent chemoradiation in patients with oropharyngeal carcinomas. *Oncoimmunology.* 2019;8(8):1614858. doi:10.1080/2162402X.2019.1614858
126. Ghanizada M, Jakobsen KK, Grønhøj C, von Buchwald C. The effects of checkpoint inhibition on head and neck squamous cell carcinoma: A systematic review. *Oral Oncol.* 2019. doi:10.1016/j.oraloncology.2019.01.018
127. 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. doi:10.1016/S1470-2045(16)30066-3
128. Alfieri S, Cavalieri S, Licitra L. Immunotherapy for recurrent/metastatic head and neck cancer. *Curr Opin Otolaryngol Head Neck Surg.* 2018;26(2):152-156. doi:10.1097/MOO.0000000000000448
129. 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 (London, England).* 2019;393(10167):156-167. doi:10.1016/S0140-6736(18)31999-8
130. Lee YS, Johnson DE, Grandis JR. An update: emerging drugs to treat squamous cell carcinomas of the head and neck. *Expert Opin Emerg Drugs.* 2018;23(4):283-299. doi:10.1080/14728214.2018.1543400

131. Sato K, Ono T, Sato F, *et al.* Different responses to nivolumab therapy between primary and metastatic tumors in a patient with recurrent hypopharyngeal squamous cell carcinoma. *Oral Oncol.* July 2019. doi:10.1016/j.oraloncology.2019.07.009
132. Ferris RL, Blumenschein GJ, Fayette J, *et al.* Nivolumab for Recurrent Squamous-Cell Carcinoma of the Head and Neck. *N Engl J Med.* 2016;375(19):1856-1867. doi:10.1056/NEJMoa1602252
133. Zandberg DP, Algazi AP, Jimeno A, *et al.* Durvalumab for recurrent or metastatic head and neck squamous cell carcinoma: Results from a single-arm, phase II study in patients with $\geq 25\%$ tumour cell PD-L1 expression who have progressed on platinum-based chemotherapy. *Eur J Cancer.* 2019;107:142-152. doi:10.1016/j.ejca.2018.11.015
134. Chow LQM, Haddad R, Gupta S, *et al.* Antitumor Activity of Pembrolizumab in Biomarker-Unselected Patients With Recurrent and/or Metastatic Head and Neck Squamous Cell Carcinoma: Results From the Phase Ib KEYNOTE-012 Expansion Cohort. *J Clin Oncol.* 2016;34(32):3838-3845. doi:10.1200/JCO.2016.68.1478
135. Bedognetti D, Hendrickx W, Ceccarelli M, Miller LD, Seliger B. Disentangling the relationship between tumor genetic programs and immune responsiveness. *Curr Opin Immunol.* 2016;39:150-158. doi:10.1016/j.coi.2016.02.001
136. Ferris RL. Immunology and immunotherapy of head and neck cancer. *J Clin Oncol.* 2015. doi:10.1200/JCO.2015.61.1509

	stage			
	II	III	IVA	IVB
N-stage				
N0	2	15	11	1
N1	0	26	3	2
N2a	0	0	11	0
N2b	0	0	70	3
N2c	0	0	38	6
N3	0	0	0	9

Figure S1. Continued

Figure S1. AJCC stage relation to TNM staging and outcomes

AJCC stages incorporate most prognostic information contained in TNM stages. A) Prognostic value of T-stage in our cohort, for all measured outcomes. The log-test results indicate that T-stage is not relevant for prognosis. B) Prognostic value of N-stage in our cohort, for all measured outcomes. The log-test results indicate that N-stage is relevant for prognosis of overall survival, progression free survival and distant metastasis. C) Distribution of N-stages over the AJCC stages and prognostic value of N-stage within these stages. Within each AJCC stage, N-staging provides little additional prognostic value, while a trend towards poor survival of N0 or N1 but IVA staged patients can be captured in the AJCC stages. Note that the few numbers in Stage IVB preclude an appropriate statistical evaluation. D) N-stage distribution over the AJCC stages.

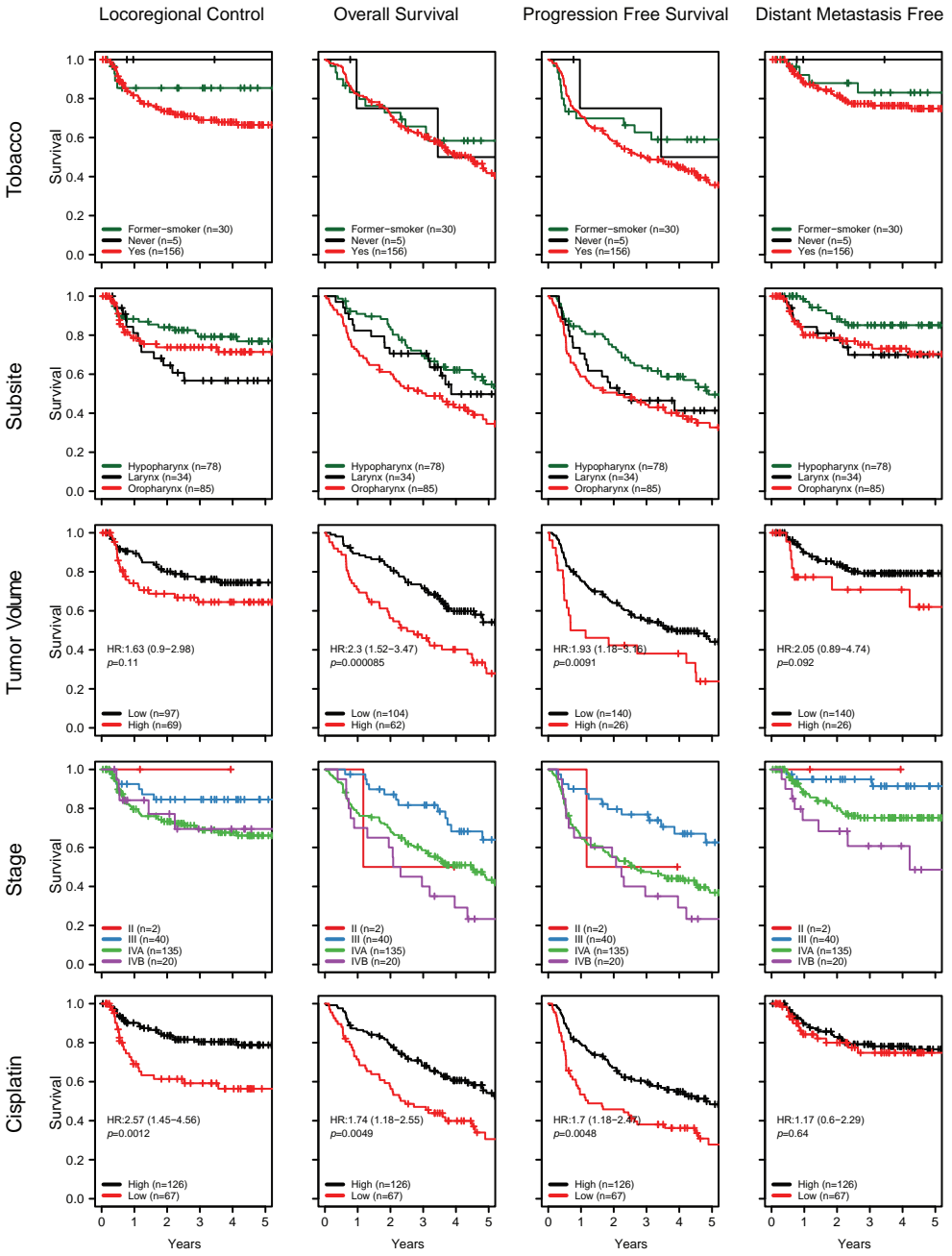


Figure S2. Continued

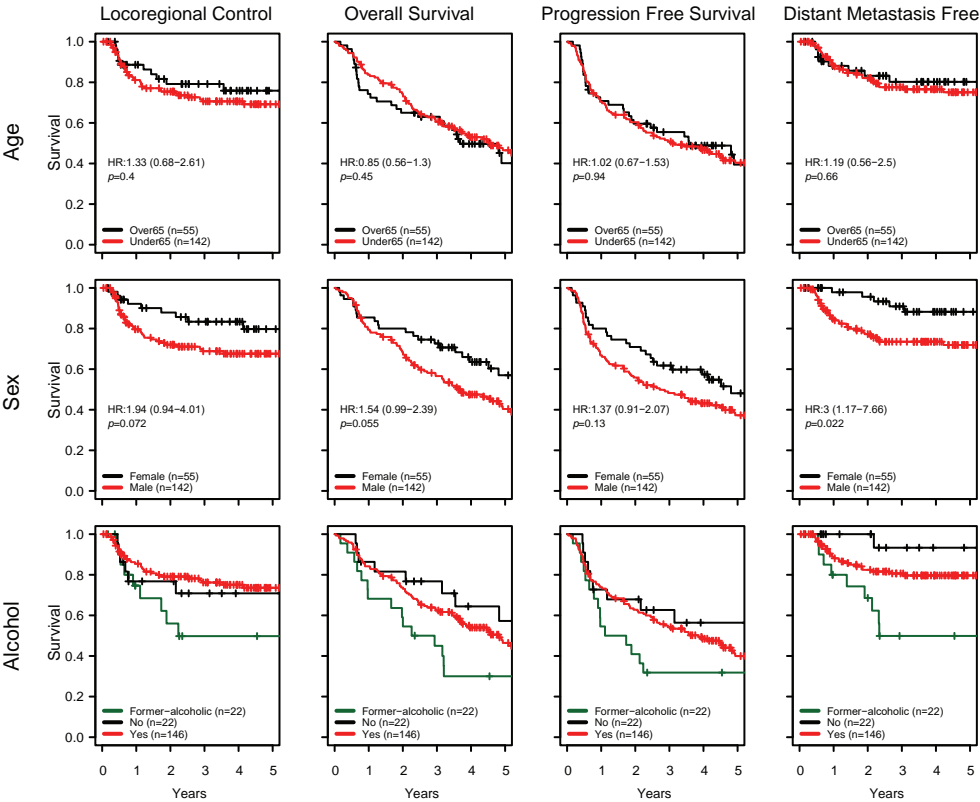


Figure S2. Patient outcome analysis by clinical factors

Each clinical factor (as indicated to the left) was tested for their association with locoregional control, overall survival, progression free survival and distant metastasis free survival. Hazard ratios and p-values for these outcome endpoints were obtained with Cox proportional hazard analyses and are as in Table S1.

ATTGTTTATAGGTGTGAGGAGTCTAGTCTCTGATCTTTCTCTGTATGGAAATCCAGTTATCTGTCTCCACTTGTGAAATAGGCTTCCTTTCTCTACTGAATGCTTTTAAATTTAATTAATTTACAGTTGGAGTATAGGGTTACCA

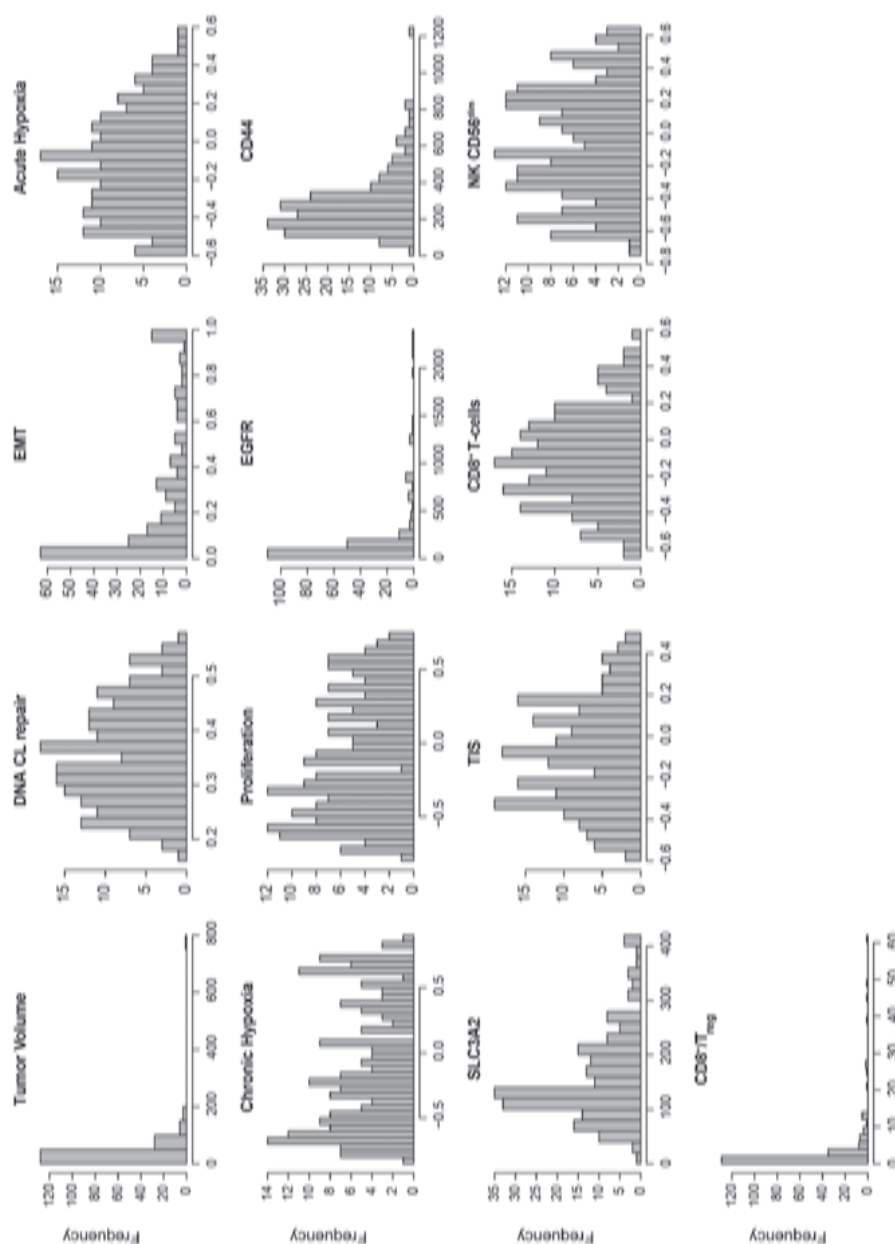


Figure S3. Distribution of biological markers
Distribution of the marker scores for each biological factor as indicated. The frequency of the individual marker scores (as indicated on the x-axis) within the patient sample cohort is shown.

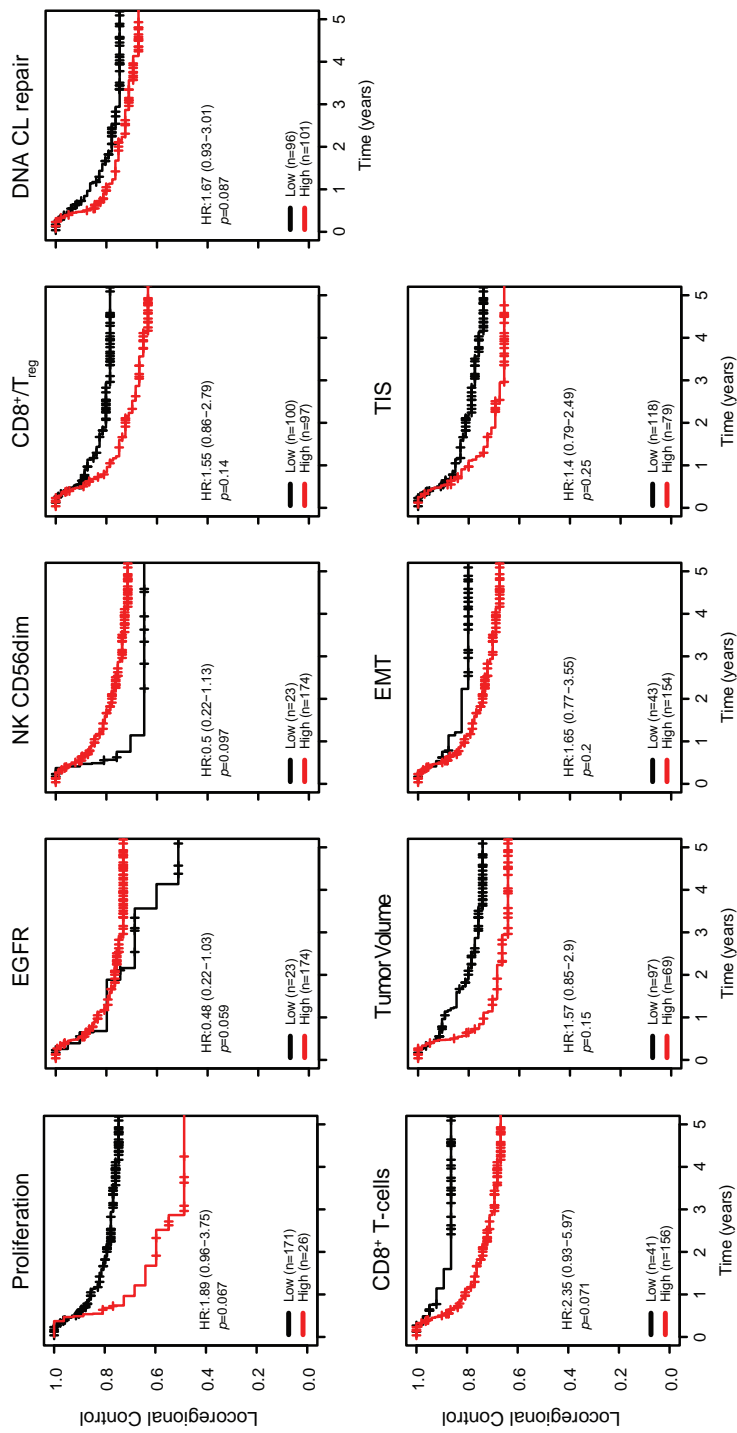
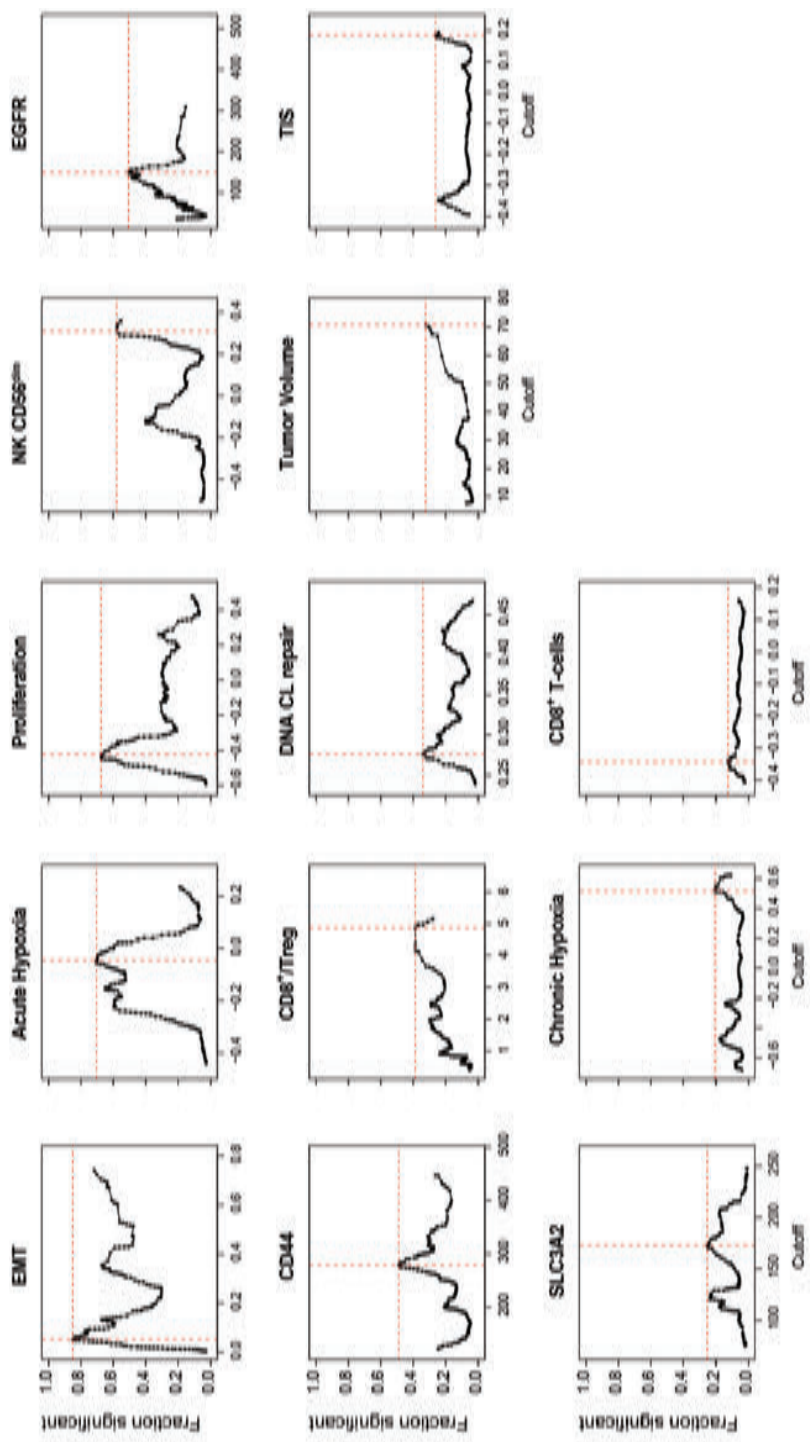


Figure S4. Association biological markers with locoregional control
Kaplan Meier graphs showing the results from the individual biological marker analyses. The study cohort was split into high and low marker score classified groups based on the cutoff determined in Figure 1A and analyzed as in Figure 1B. This supplementary figure shows the Kaplan Meier curves of the remaining and non-significant markers. Hazard ratios for locoregional recurrences and associated p-values were obtained from multivariate Cox proportional hazard analyses using the same clinical variables as those used to determine the cutoff (listed in Materials and Methods).

ATTGTTTATGGTGTGAGGTGTAGGAGTCTAGTCTCTGGTATCTTTCTCTGTATGGAATCCAGTTATTCTGTCTCCACTTGTGAAATAGGCTTCCTTTCTCTACTGAAATGCTTTTAAATTTTAAATATTTTACAGTTGGAGTATAGGGTTACCA

A



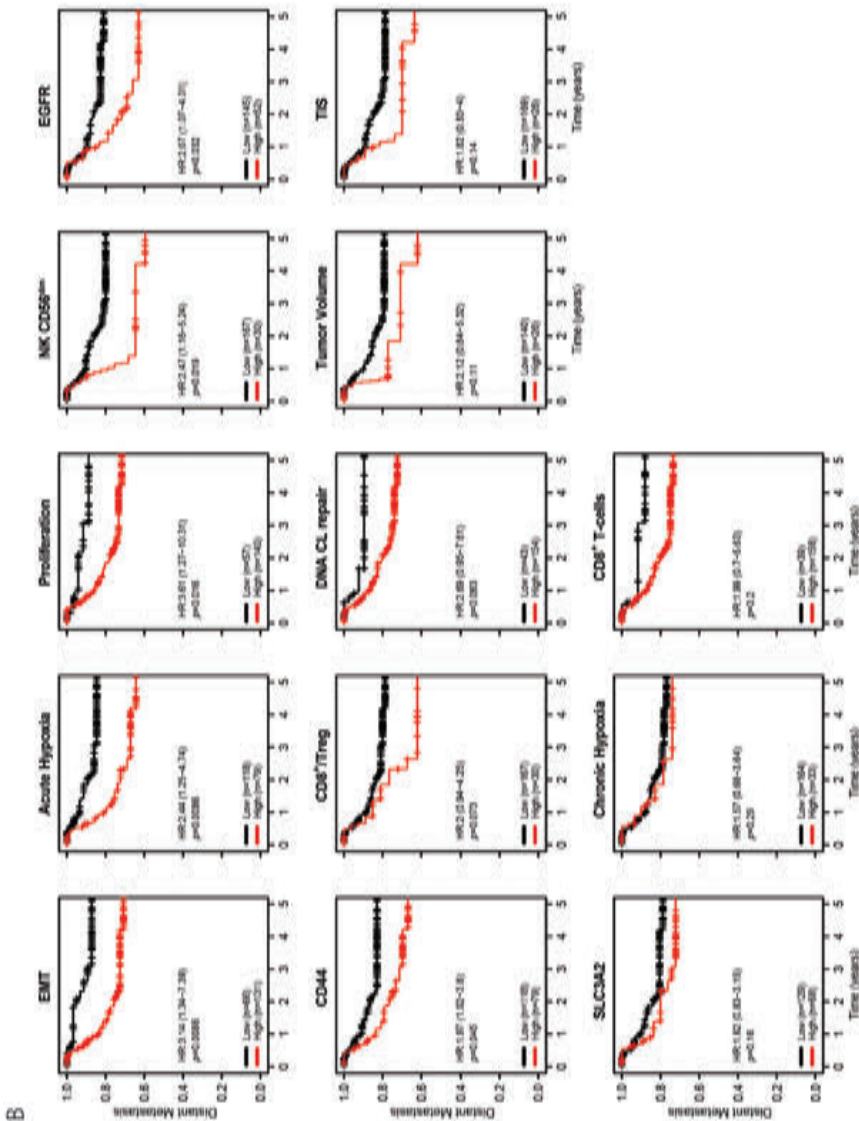


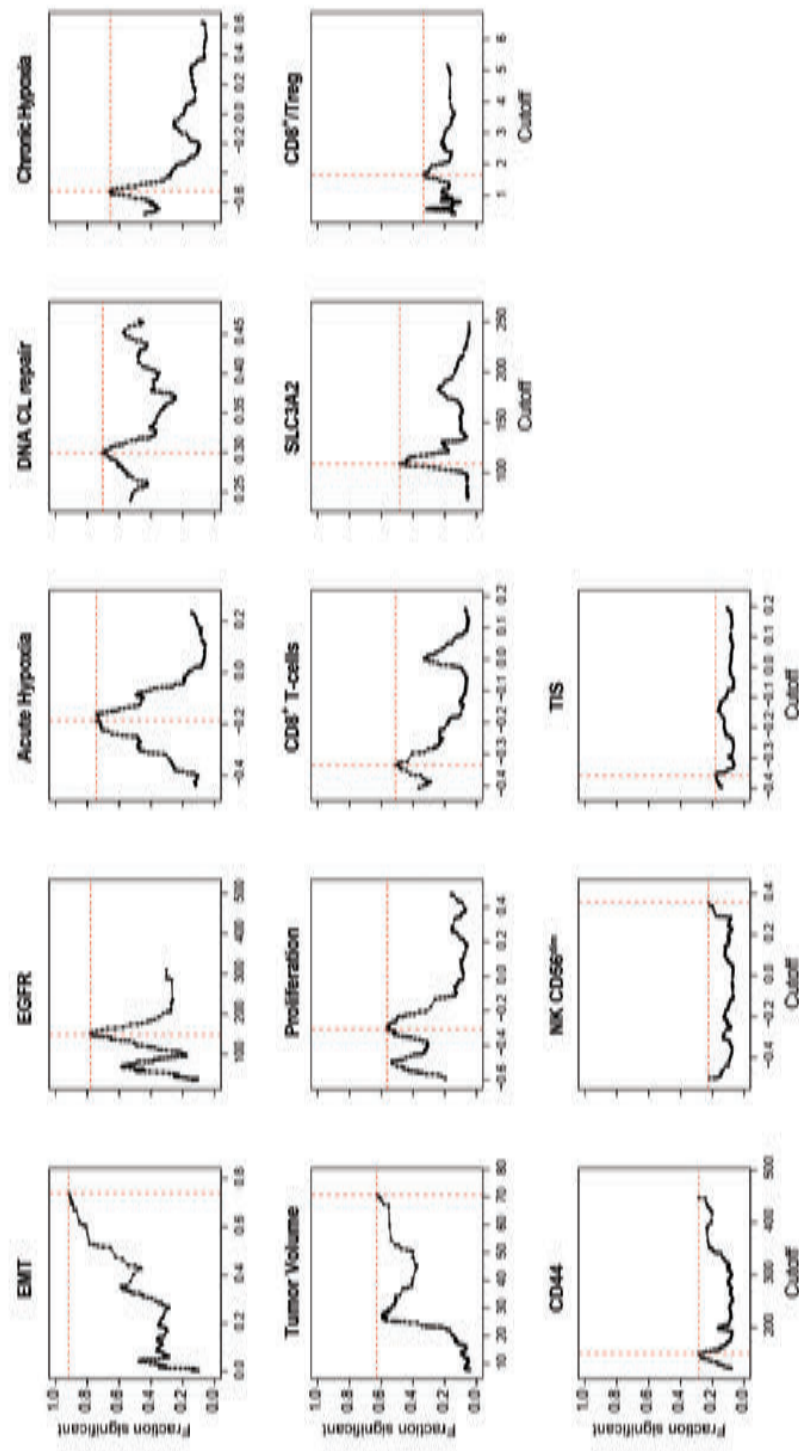
Figure S5. Role of individual biological factors in distant metastasis risk

(A) Results of bootstrap analysis (see methods). The fraction of randomized cohorts with a significant association with overall survival is plotted for each score cutoff and biological marker. Markers are ordered by the maximum fraction of significant Cox proportional hazard tests at the best cutoff, indicated with dotted red lines.

(B) For each marker, the patient cohort was split into high and low scoring groups using the cutoffs determined in A. Hazard ratios for distant metastasis and p-values were obtained from multivariate Cox proportional analyses using the same clinical variables as those used to determine the cutoff (listed in Materials and Methods).

ATTGTTTATGGTGTGAGGTAGAGTCTAGTCTCTGGTATCTTCTCTGTATGGAAATCCAGTTATTCTGTCTCCACTTGTGAAATAGGCTTCTTCTCTACTGAATGCTTTTAAATTTTAAATATTATACAGTTGGAGTATAGGGTACCA

A
220



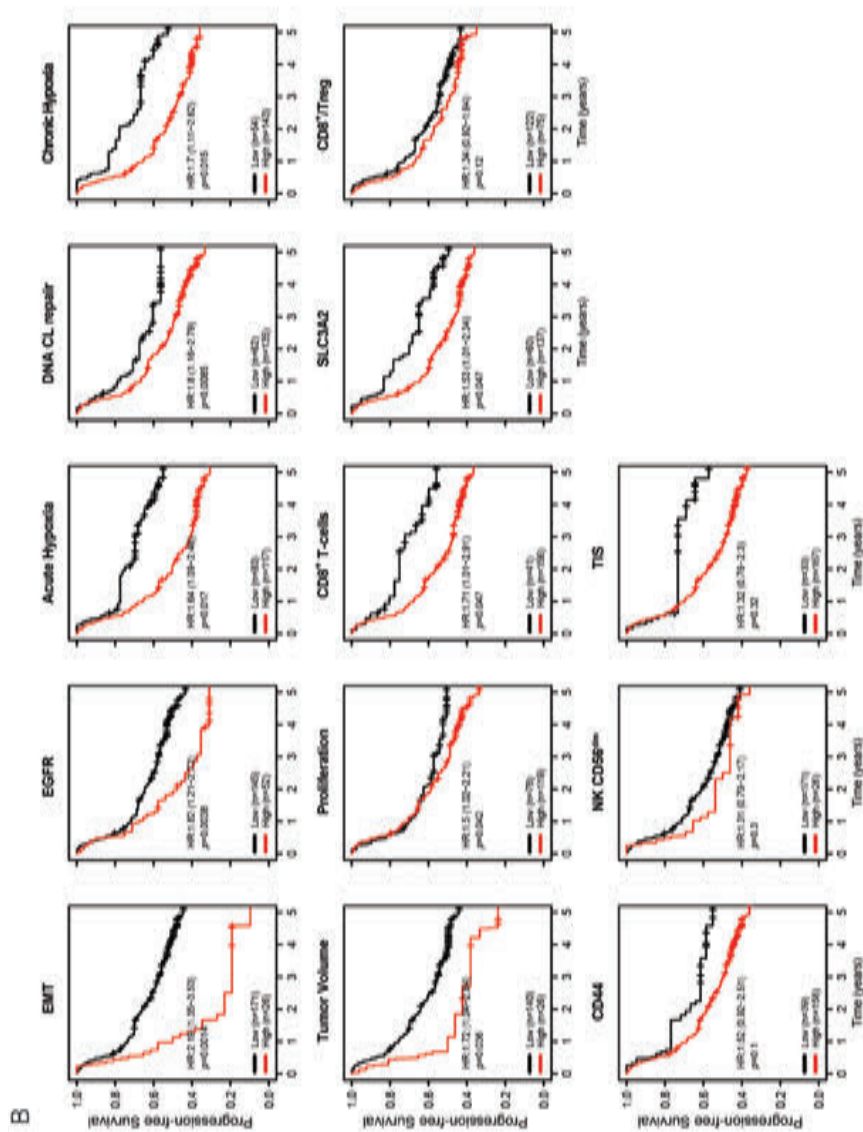
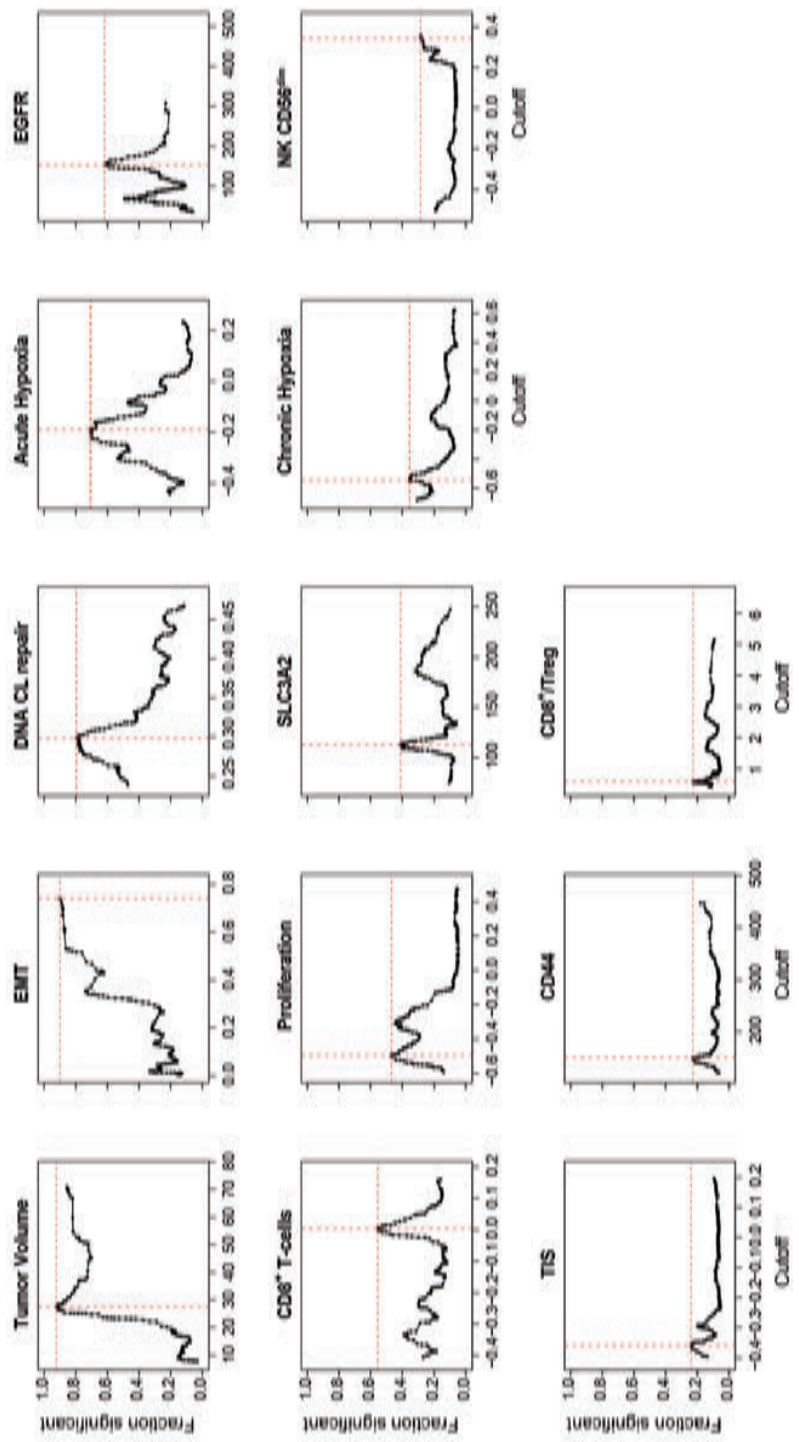


Figure S6. Role of individual biological factors in progression free survival
(A) Results of bootstrap analysis (see methods). The fraction of randomized cohorts with a significant association with progression free survival is plotted for each score cutoff and biological marker. Markers are ordered by the maximum fraction of significant Cox proportional hazard tests at the best cutoff, indicated with dotted red lines.
(B) For each marker, the patient cohort was split into high and low scoring groups using the cutoffs determined in A. Hazard ratios for progression and p -values were obtained from multivariate Cox proportional analyses using the same clinical variables as those used to determine the cutoff (listed in Materials and Methods).

ATTGTTTATGGTGTGAGGTGATAGGATCAGTCTCTGGATCTTCTCTGTATGGAAATCCAGTTATTCTGTCTCCACTTGTGAAATAGGCTTCCTTCTCTACTGAAATGCTTTAAATTTTAAATATTTTACAGTTGGAGTATAGGGTACCA

A



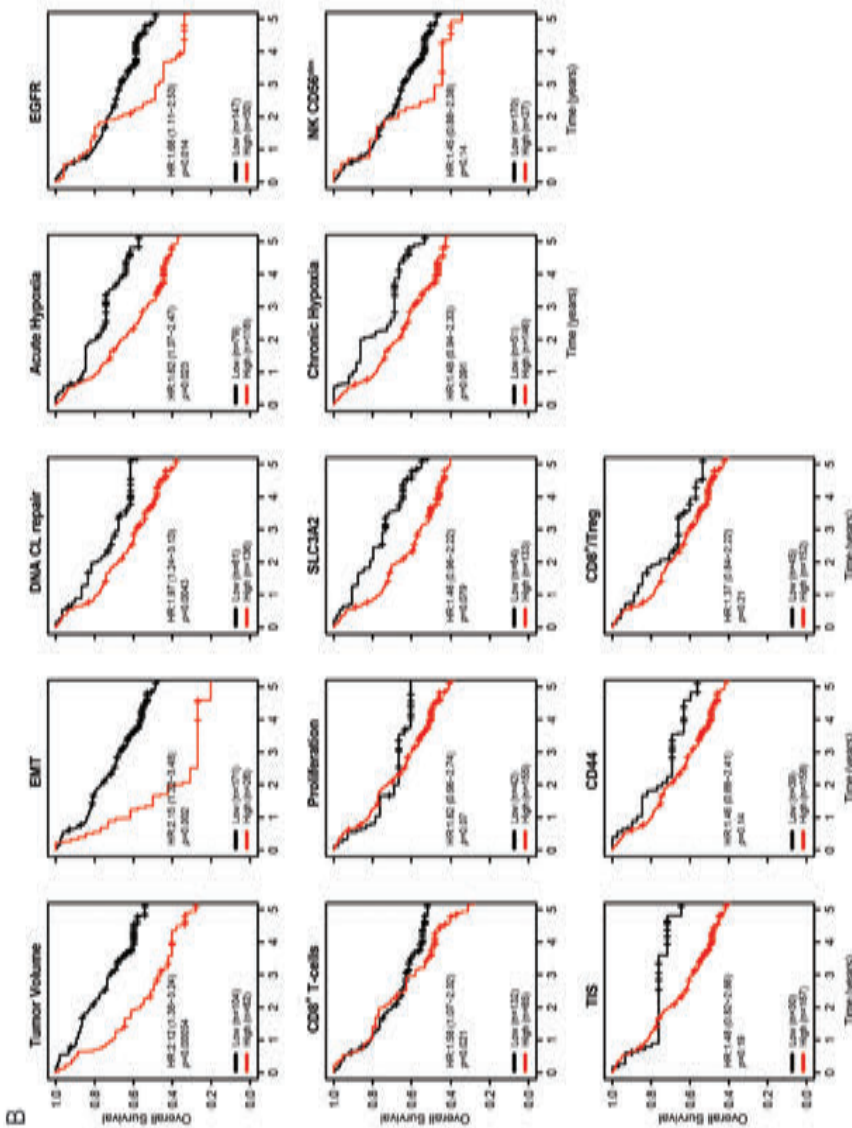


Figure S7. Role of individual biological factors in overall survival

A) Results of bootstrap analysis (see methods). The fraction of randomized cohorts with a significant association with overall survival is plotted for each score cutoff and biological marker. Markers are ordered by the maximum fraction of significant Cox proportional hazard tests at the best cutoff, indicated with dotted red lines.

(B) For each marker, the patient cohort was split into high and low groups using the cutoffs determined in A. Hazard ratios for death and p-values were obtained with a multivariate Cox proportional analysis using the same clinical variables as those used to determine the cutoff (listed in Materials and Methods).

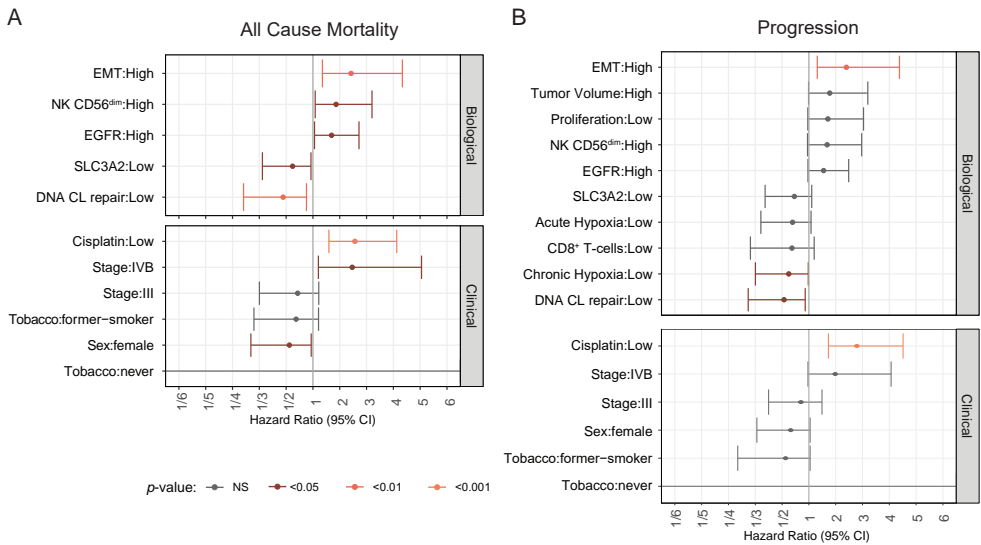


Figure S8. Multivariate Cox Proportional Hazards models for overall survival and progression free survival The models were generated using a backward selection procedure as described in Materials and Methods. In brief, a Cox proportional hazard model was fitted that included all biological markers and clinical variables. Possible improvements in model performance were assessed after eliminating individual variables, a process repeated until elimination of variables did not further improve the model. Hazard ratios for death (overall survival endpoint) or for progression from the final model are shown.

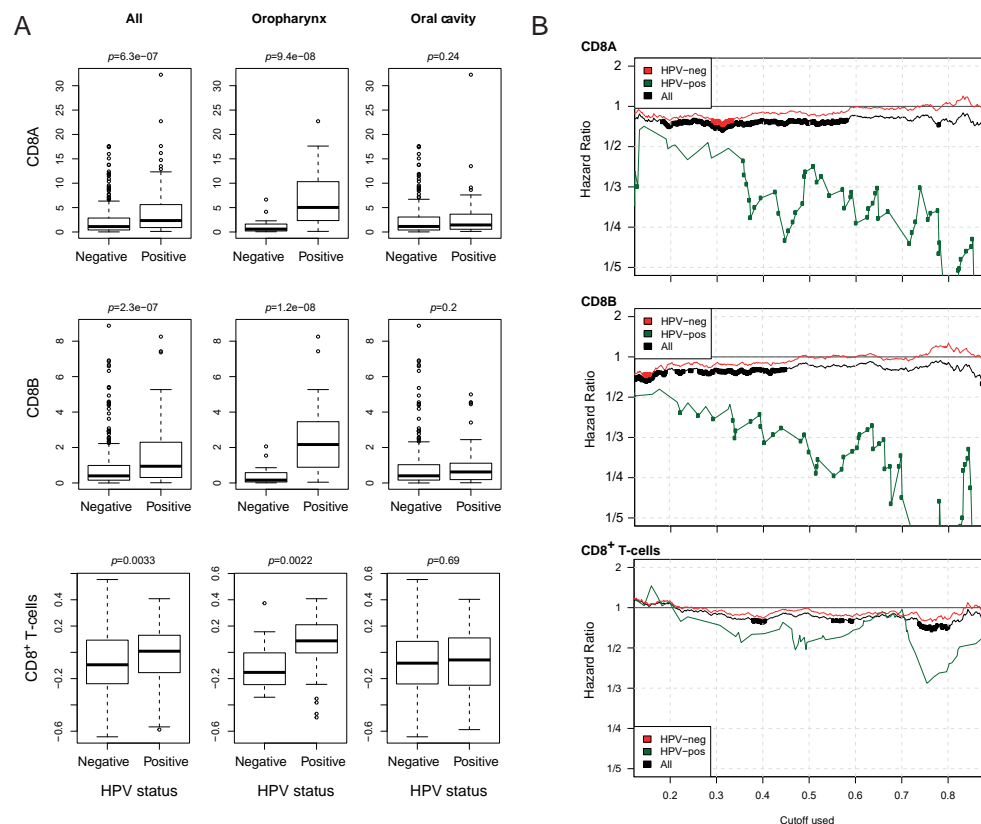


Figure S9. Relation between CD8⁺ T-cell markers and HPV status in TCGA data A) CD8A, CD8B and the CD8⁺ T-cell signature are higher in HPV-positive samples in the TCGA-HNSC cohort. This effect is predominantly observed in oropharynx patients. B) HR-plots by CD8A, CD8B expression and the CD8⁺ T-cell signature scoring. Dots indicate HR for which $p < 0.05$. The improved prognosis observed for patients with high CD8A, CD8B expression or CD8⁺ T-cell signature scores in the whole cohort (black lines) is mostly driven by a good prognosis association in HPV-positive patients (green line); hazard ratios for HPV-negative patients (red line) are close to 1.

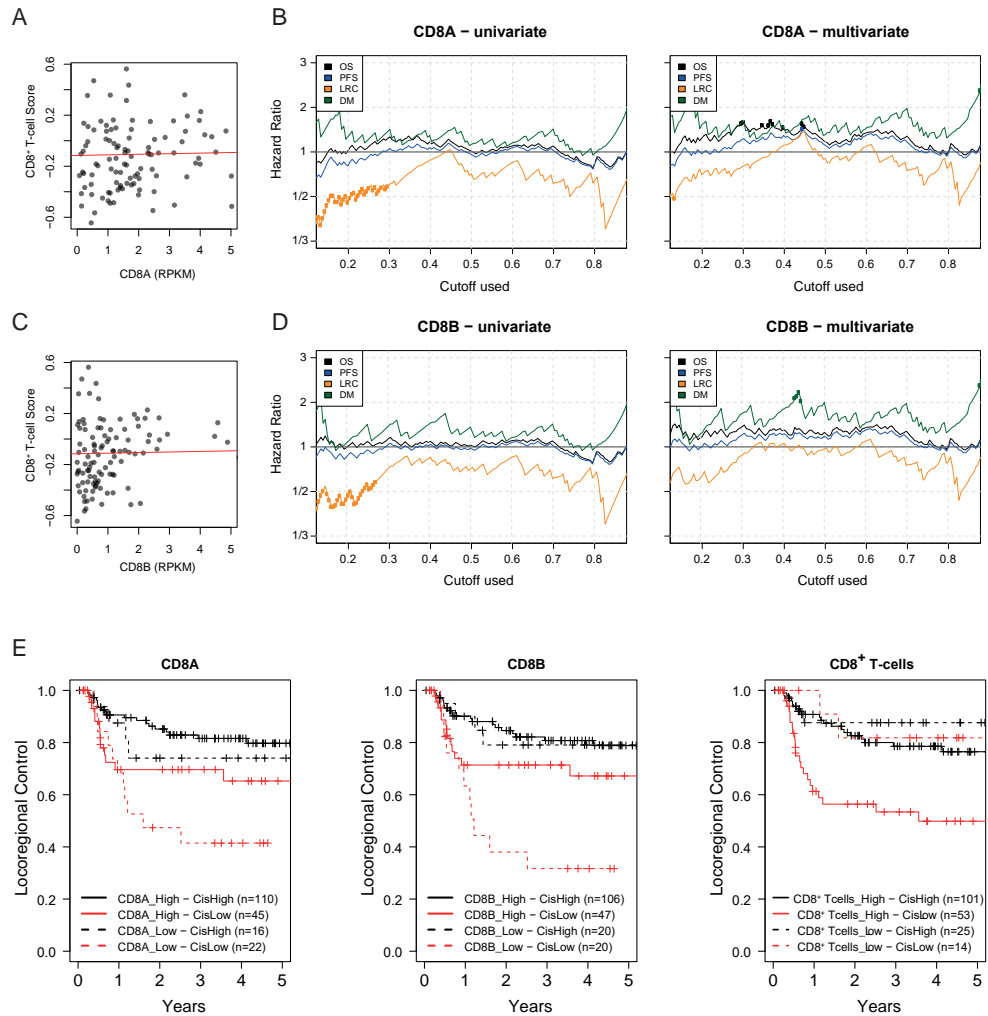


Figure S10. Relation between CD8⁺ T-cell signature and CD8 expression A) CD8⁺ T-cell signature scoring correlates poorly with CD8A expression. B) Univariate and multivariate (including cisplatin and tumor site) HR-plots for CD8A. The top 75-85% highest CD8A expressing patients have better prognosis than the 15-25% lowest CD8A expressing patients. However, this effect is lost if differences in cumulative cisplatin dose and tumor site are accounted for in multivariate analyses. Dots indicate HR for which $p < 0.05$. C) Equivalent to A, for CD8B. D) Equivalent to B, for CD8B. E) Kaplan-Meijer plots showing the interaction between CD8 expression and cisplatin dose. A 20-80 split for CD8A and CD8B. Cutoff as identified in the bootstrapping method was applied for the CD8⁺ T-cell signature. Patients grouped according to CD8A / CD8B or the signature scores seem to have opposite prognosis within those who received a low cumulative cisplatin dose. Note, the interactions are not statistically significant.

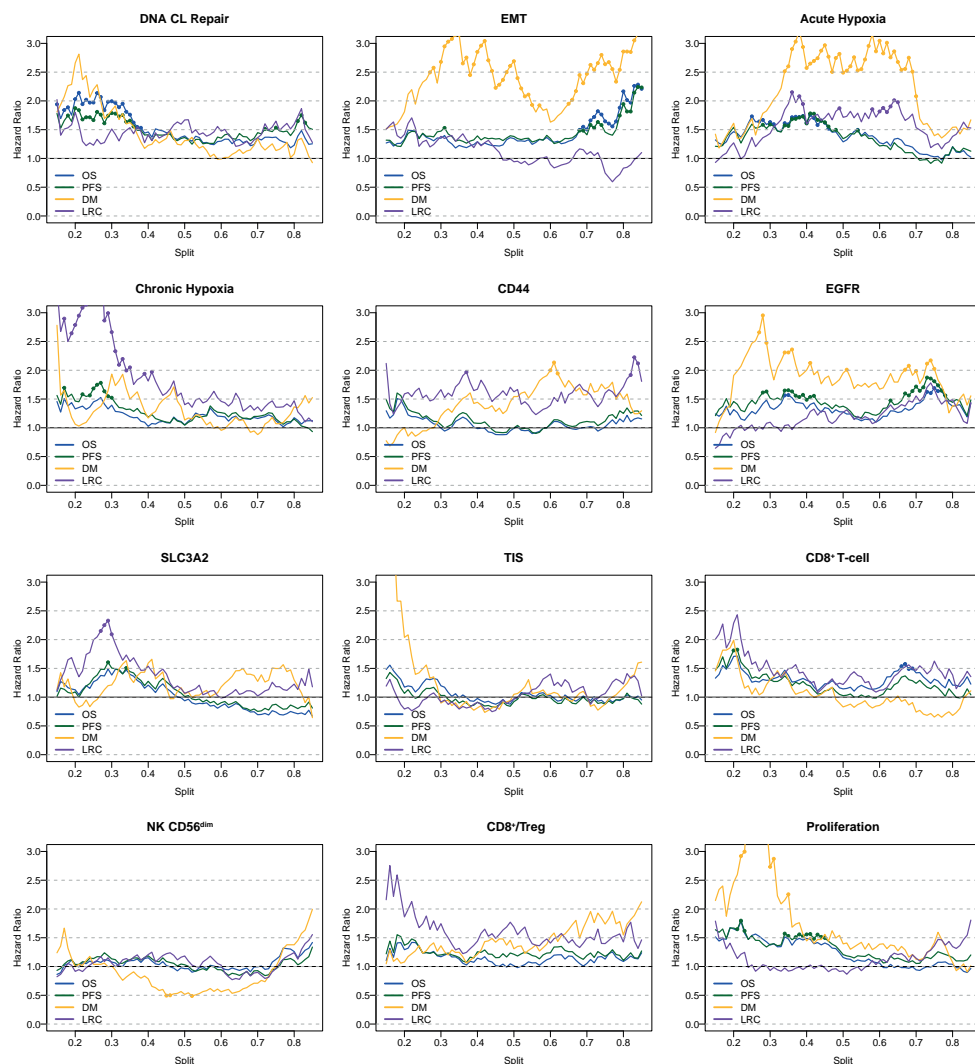


Figure S11. Hazard ratio plots for all outcomes and all biological factors Repeated multivariate Cox HR analyses in unequally-sized patients groups with splits as indicated on the X-axis. Multivariable analyses were performed as before with: sex and cisplatin for locoregional control (LRC); sex, subsite and cisplatin for overall survival (OS); stage, subsite and cisplatin for progression free survival (PFS) and sex and alcohol for distant metastasis free survival (DM). Multivariate Cox proportional hazard ratios (HR) for locoregional recurrences (LRC), distant metastasis (DM), disease progression (PFS) and death (OS) are shown on the y-axis for each split. Dots are HR values with $p < 0.05$.

Supplementary tables

Supplementary Table 1. Univariate Cox proportional hazard analyses for all clinical variables considered in this study.
REF: reference group

Variable	Locoregional Regression				Mortality			Progression			Distant Metastasis	
	N	HR (95% CI)	p-value		HR (95% CI)	p-value		HR (95% CI)	p-value		HR (95% CI)	p-value
Age at diagnosis												
	≥65	55	0.749 (0.38-1.47)	0,4	1.176 (0.77-1.8)	0,45		0.984 (0.65-1.49)	0,94		0.844 (0.4-1.78)	0,66
	<65	142	REF		REF			REF			REF	
Sex												
	female	55	0.515 (0.25-1.06)	0,072	0.65 (0.42-1.01)	0,055		0.729 (0.48-1.1)	0,13		0.334 (0.13-0.85)	0,022
	male	142	REF		REF			REF			REF	
Alcohol												
	no	22	1.235 (0.52-2.95)	0,64	0.769 (0.4-1.49)	0,44		0.813 (0.43-1.52)	0,52		0.243 (0.03-1.79)	0,17
	former	22	2.158 (1.03-4.52)	0,041	1.76 (1.05-2.95)	0,032		1.59 (0.95-2.66)	0,076		2.68 (1.25-5.73)	0,011
	alcoholic											
	yes	146	REF		REF			REF			REF	
	missing	7										
Tobacco												
	never	5	0 (0-Inf)	1	0.724 (0.18-2.94)	0,65		0.575 (0.14-2.33)	0,44		0 (0-Inf)	1
	former	30	0.463 (0.17-1.29)	0,14	0.657 (0.37-1.16)	0,15		0.594 (0.34-1.05)	0,071		0.65 (0.23-1.83)	0,42
	smoker											
	yes	156	REF		REF			REF			REF	
	missing	6										
Tumour subsite												
	Larynx	34	1.47 (0.73-2.96)	0,28	0.615 (0.35-1.09)	0,094		0.822 (0.49-1.38)	0,46		1.026 (0.47-2.25)	0,95
	Hypopharynx	78	0.708 (0.37-1.37)	0,3	0.517 (0.34-0.78)	0,0017		0.555 (0.37-0.83)	0,0039		0.436 (0.2-0.93)	0,032
	Oropharynx	85										
	≥25.4 cc	70	1.754 (0.96-3.19)	0,066	2.147 (1.42-3.24)	0,00027		1.832 (1.23-2.72)	0,0027		1.513 (0.76-3)	0,23
Tumour volume												
	<25.4 cc	96	REF		REF			REF			REF	
	missing	31										

Supplementary Table 1. Continued

Variable		Locoregional Regression			Mortality			Progression			Distant Metastasis	
		N	HR (95% CI)	p-value	HR (95% CI)	p-value	HR (95% CI)	p-value	HR (95% CI)	p-value	HR (95% CI)	p-value
Stage	IVB	20	0.896 (0.35-2.28)	0,82	1.803 (1.06-3.06)	0,029	1.438 (0.85-2.42)	0,17	2.093 (0.95-4.6)	0,066		
	III	40	0.443 (0.19-1.05)	0,064	0.607 (0.36-1.03)	0,063	0.521 (0.31-0.88)	0,014	0.306 (0.09-1.01)	0,051		
	II	2	0 (0-Inf)	1	1.133 (0.16-8.17)	0,9	0.819 (0.11-5.9)	0,84	0 (0-Inf)	1		
	IVA	135	REF		REF		REF		REF			
Cumulative	<200mg/m ²	67	2.569 (1.45-4.56)	0,0012	1.737 (1.18-2.55)	0,0049	1.704 (1.18-2.47)	0,0048	1.175 (0.6-2.29)	0,64		
	≥200mg/m ²	126	REF		REF		REF		REF			
cisplatin dose	missing	4										

Supplementary Table S2. Cutoffs as determined by the bootstrap method.

Marker	Outcome	Cutoff	Fraction($p < 0.05$)
DNA CL repair	Overall Survival	0,298	0,796
	Progression Free Survival	0,299	0,703
	Locoregional Control	0,348	0,368
	Distant Metastasis	0,276	0,341
EMT	Overall Survival	0,741	0,901
	Progression Free Survival	0,741	0,916
	Locoregional Control	0,015	0,177
	Distant Metastasis	0,053	0,849
Acute Hypoxia	Overall Survival	-0,189	0,707
	Progression Free Survival	-0,186	0,745
	Locoregional Control	-0,018	0,593
	Distant Metastasis	-0,046	0,705
Chronic Hypoxia	Overall Survival	-0,545	0,355
	Progression Free Survival	-0,524	0,655
	Locoregional Control	-0,545	0,870
	Distant Metastasis	0,516	0,205
CD44	Overall Survival	150,247	0,228
	Progression Free Survival	150,247	0,284
	Locoregional Control	424,349	0,548
	Distant Metastasis	278,756	0,489
EGFR	Overall Survival	154,421	0,620
	Progression Free Survival	148,863	0,780
	Locoregional Control	36,609	0,469
	Distant Metastasis	148,863	0,506
SLC3A2	Overall Survival	112,566	0,411
	Progression Free Survival	109,439	0,482
	Locoregional Control	108,867	0,568
	Distant Metastasis	172,793	0,255
TIS	Overall Survival	-0,358	0,240
	Progression Free Survival	-0,358	0,179
	Locoregional Control	-0,035	0,155
	Distant Metastasis	0,186	0,265
CD8 ⁺ T cells	Overall Survival	0,002	0,556
	Progression Free Survival	-0,334	0,509
	Locoregional Control	-0,334	0,334
	Distant Metastasis	-0,342	0,125
NK CD56 ^{dim}	Overall Survival	0,341	0,286
	Progression Free Survival	0,358	0,224
	Locoregional Control	-0,501	0,409
	Distant Metastasis	0,308	0,579

Supplementary Table 2. Continued

Marker	Outcome	Cutoff	Fraction($p<0.05$)
CD8 ⁺ /T _{reg}	Overall Survival	0,580	0,226
	Progression Free Survival	1,649	0,334
	Locoregional Control	1,048	0,403
	Distant Metastasis	4,869	0,387
Proliferation	Overall Survival	-0,495	0,468
	Progression Free Survival	-0,305	0,563
	Locoregional Control	0,476	0,492
	Distant Metastasis	-0,423	0,674
Tumor Volume	Overall Survival	27,500	0,923
	Progression Free Survival	70,700	0,627
	Locoregional Control	26,096	0,284
	Distant Metastasis	70,700	0,324

Multivariable analyses used cutoffs as listed in Supplementary Table S2 for each marker and incorporated all clinical variables that resulted to be significantly associated with the respective outcome endpoint.

REF: reference group

Variable	Locoregional Control				Overall Survival				Progression				Distant Metastasis			
	Value	N	HR (95% CI)	p-value	N	HR (95% CI)	p-value	N	HR (95% CI)	p-value	N	HR (95% CI)	p-value	N	HR (95% CI)	p-value
DNA CL Repair	High	101	1.67 (0.93-3.01)	0,087	136	1.97 (1.24-3.13)	0,00426	135	1.8 (1.16-2.79)	0,00847	154	2.69 (0.95-7.61)	0,0628	154	2.69 (0.95-7.61)	0,0628
	Low	96	REF		61	REF		62	REF		43	REF		43	REF	
EMT	High	154	1.65 (0.77-3.55)	0,196	26	2.15 (1.32-3.48)	0,00202	26	2.19 (1.35-3.53)	0,00138	131	3.14 (1.34-7.39)	0,00861	131	3.14 (1.34-7.39)	0,00861
	Low	43	REF		171	REF		171	REF		66	REF		66	REF	
Acute Hypoxia	High	73	1.9 (1.07-3.4)	0,0296	118	1.62 (1.07-2.47)	0,0231	117	1.64 (1.09-2.46)	0,017	79	2.44 (1.25-4.74)	0,00857	79	2.44 (1.25-4.74)	0,00857
	Low	124	REF		79	REF		80	REF		118	REF		118	REF	
Chronic Hypoxia	High	146	3.95 (1.56-10.01)	0,00381	146	1.48 (0.94-2.33)	0,0913	143	1.7 (1.11-2.62)	0,0152	33	1.57 (0.68-3.64)	0,29	33	1.57 (0.68-3.64)	0,29
	Low	51	REF		51	REF		54	REF		164	REF		164	REF	
CD44	High	31	2.03 (1.02-4.04)	0,0426	158	1.46 (0.88-2.41)	0,143	158	1.52 (0.92-2.51)	0,0999	79	1.97 (1.02-3.8)	0,0447	79	1.97 (1.02-3.8)	0,0447
	Low	166	REF		39	REF		39	REF		118	REF		118	REF	
EGFR	High	174	0.48 (0.22-1.03)	0,0585	50	1.68 (1.11-2.53)	0,014	52	1.82 (1.21-2.72)	0,00378	52	2.07 (1.07-4.01)	0,0317	52	2.07 (1.07-4.01)	0,0317
	Low	23	REF		147	REF		145	REF		145	REF		145	REF	
SLC3A2	High	141	2.31 (1.11-4.82)	0,0257	133	1.46 (0.96-2.22)	0,0791	137	1.53 (1.01-2.34)	0,0468	68	1.62 (0.83-3.15)	0,157	68	1.62 (0.83-3.15)	0,157
	Low	56	REF		64	REF		60	REF		129	REF		129	REF	
TTIS	High	79	1.4 (0.79-2.49)	0,248	167	1.48 (0.82-2.66)	0,188	167	1.32 (0.76-2.3)	0,322	28	1.82 (0.83-4)	0,136	28	1.82 (0.83-4)	0,136
	Low	118	REF		30	REF		30	REF		169	REF		169	REF	
CD8 ⁺ T-cells	High	156	2.35 (0.93-5.97)	0,0715	65	1.58 (1.07-2.32)	0,0208	156	1.71 (1.01-2.91)	0,0473	158	1.99 (0.7-5.63)	0,196	158	1.99 (0.7-5.63)	0,196
	Low	41	REF		132	REF		41	REF		39	REF		39	REF	
NK CD56 ^{dim}	High	174	0.5 (0.22-1.13)	0,0967	27	1.45 (0.88-2.38)	0,142	26	1.31 (0.79-2.17)	0,302	30	2.47 (1.16-5.24)	0,0188	30	2.47 (1.16-5.24)	0,0188
	Low	23	REF		170	REF		171	REF		167	REF		167	REF	
CD8 ⁺ /T _{reg}	High	97	1.55 (0.86-2.79)	0,141	152	1.37 (0.84-2.22)	0,21	75	1.34 (0.92-1.94)	0,123	30	2 (0.94-4.25)	0,0732	30	2 (0.94-4.25)	0,0732
	Low	100	REF		45	REF		122	REF		167	REF		167	REF	
Proliferation	High	26	1.89 (0.96-3.75)	0,0668	155	1.62 (0.96-2.74)	0,07	119	1.5 (1.02-2.21)	0,0418	140	3.61 (1.27-10.31)	0,0162	140	3.61 (1.27-10.31)	0,0162
	Low	171	REF		42	REF		78	REF		57	REF		57	REF	

Supplementary Table S4. Multivariate Cox proportional hazard analysis results.

Multivariable analyses incorporated all clinical and biological variables remaining after the backward selection procedure.

Outcome	variable	Hazard Ratio	p-value
Mortality	Sex: female	0.53 (0.3-0.94)	0.03
	Tobacco: former-smoker	0.61 (0.31-1.21)	0.16
	Tobacco: never	0 (0-Inf)	1
	Stage: III	0.64 (0.33-1.22)	0.17
	Stage: IVB	2.47 (1.2-5.05)	0.014
	Cisplatin: Low	2.57 (1.6-4.13)	1,00E-04
	DNA CL Repair: Low	0.47 (0.28-0.81)	0.0058
	EMT: High	2.42 (1.35-4.34)	0.0029
	EGFR: High	1.7 (1.06-2.72)	0.028
	SLC3A2: Low	0.57 (0.35-0.93)	0.025
	NK CD56 ^{dim} : High	1.86 (1.08-3.2)	0.024
Progression	Sex: female	0.6 (0.34-1.05)	0.071
	Tobacco: former-smoker	0.53 (0.27-1.05)	0.067
	Tobacco: never	0 (0-Inf)	1
	Stage: III	0.77 (0.4-1.49)	0.44
	Stage: IVB	1.98 (0.97-4.07)	0.062
	Cisplatin: Low	2.79 (1.73-4.51)	2.8e-05
	DNA CL Repair: Low	0.52 (0.31-0.88)	0.015
	EMT: High	2.4 (1.31-4.38)	0.0044
	Acute Hypoxia: Low	0.62 (0.36-1.08)	0.092
	Chronic Hypoxia: Low	0.57 (0.33-0.98)	0.041
	EGFR: High	1.55 (0.97-2.48)	0.069
	SLC3A2: Low	0.65 (0.38-1.11)	0.11
	CD8 ⁺ T-cells: Low	0.61 (0.31-1.2)	0.15
	NK CD56 ^{dim} : High	1.68 (0.95-2.97)	0.074
	Proliferation: Low	1.71 (0.96-3.04)	0.068
	Tumor Volume: High	1.78 (0.99-3.2)	0.055
	Age Over 65	0.34 (0.15-0.74)	0.007
Locoregional Recurrences	Sex: female	0.24 (0.1-0.59)	0.002
	Cisplatin: Low	3.07 (1.59-5.93)	0.00084
	Chronic Hypoxia: Low	0.23 (0.08-0.68)	0.0073
	CD44: High	1.92 (0.89-4.13)	0.097
	EGFR: Low	2.89 (1.21-6.89)	0.017
	TIS: High	2.21 (1.05-4.65)	0.036
	CD8 ⁺ /T _{reg} : High	2.62 (1.26-5.45)	0.01

Supplementary Table 4. Continued

Outcome	variable	Hazard Ratio	p-value
Distant Metastasis	Sex: female	0.27 (0.07-1)	0.05
	Alcohol: former-Alcohol: ic	2.61 (1.04-6.56)	0.041
	Alcohol: no	0.31 (0.04-2.48)	0.27
	Stage: III	0.31 (0.07-1.41)	0.13
	Stage: IVB	5.22 (1.74-15.69)	0.0032
	Subsite: Hypopharynx	0.37 (0.14-0.97)	0.043
	Subsite: Larynx	1.2 (0.47-3.07)	0.7
	EMT: Low	0.23 (0.08-0.6)	0.0028
	EGFR: High	3.19 (1.4-7.23)	0.0056
	TIS: High	2.37 (0.94-5.97)	0.066
	CD8 ⁺ /T _{reg} : High	2.94 (1.14-7.53)	0.025
	Proliferation: Low	0.29 (0.07-1.25)	0.097



CHAPTER 6

TTCOCATCAACCCCTAGG99CTCCTCCTT99CTGCTGGGAGTTGTAGTCTGAACTTCTATCTT99CGAGAAACCCCTAC9CTCCCCCTACCGAGTCCC99GTAAATCTTTAAAGCACTGCAAC9CCCCCCCC999999CTG

General discussion

General discussion

This thesis focuses on the development of prognostic biomarkers to support the prediction of outcome of HPV-negative HNSCC patients treated with chemoradiotherapy. The ultimate goal is to reliably identify patients with a tumor that will respond poorly to chemoradiotherapy, leading to treatment failure. Upfront identification allows for the adaptation of treatment by e.g. surgery and ultimately better patient outcomes. This thesis shows that tumor biology plays an important role in the outcome of patients and also identifies targets for the development of alternative treatments.

Prognostic biomarkers using RNA-sequencing

HPV

As stated in the introduction, the most well-known prognostic biomarker in HNSCC is HPV status¹⁻³. HPV-positive HNSCC is biologically and clinically very different^{4,5}. Survival rates of HPV-positive and HPV-negative oropharyngeal HNSCC patients differ significantly. The addition of a separate TNM classification for HPV-positive oropharyngeal carcinoma underlines the clinical impact of HPV status. HPV-negative tumors have a worse prognosis compared to HPV-positive HNSCC, and reliable biomarkers besides TNM will be required to stratify HPV-negative tumors in relevant prognostic groups. The distinction between HPV-positive or -negative was however not done in most earlier biomarker research, which might result in severely biased and less robust outcome associations^{6,7}. To limit this confounding factor, we excluded HPV-positive HNSCC in all our studies. This homogenization of our data increased the reliability of our analyses (see the section Sample size and confounders in the discussion below), and the possibilities for future clinical implementation.

Hypoxia

Hypoxia is the most extensively investigated prognostic factor in HNSCC, specifically in the context of (chemo)radiotherapy. Hypoxic HNSCCs have a poor prognosis and hypoxia reduction can be a major improvement in HNSCC patient outcome, although initial results did not have enough impact yet (see below). Hypoxia measurements are invasive and therefore many researchers apply RNA gene expression profiles as a surrogate. Many hypoxia RNA gene profiles have been described. In this thesis, in Chapter 2, we showed that, although the four clinical hypoxia gene signatures have little overlapping genes, they classify patients concordantly. Based on the profiles from Seigneuric *et al.* we aimed to examine the impact of acute and chronic hypoxia. As it appears, current clinical hypoxia signatures for

GCTGACCAACATGGTGAACCCCGTTCTAAATAAAATACAAAATCAGGTGGGACGGTGGCTACGGCTGTATCCAGCACTTTGGGAGGCGAGGTGGGAGATCAAGAGGTAGGAGTTCAAGACAGCGCTGACCAATGTGTGA

HNSCC are closely related to chronic hypoxia and not acute hypoxia profiles. Consistent with previous reports, patients in our cohort with chronic hypoxic tumors showed worse survival rates compared to the more normoxic ones. Acute hypoxic cells are more radiotherapy resistant than chronic hypoxic cells in the preclinical setting, but the effect of acute hypoxia in the clinical setting is unclear^{8–10}. All our patients received a combination of radio- and chemotherapy. In line with the preclinical studies, in Chapter 2 we show that patients with acute hypoxic tumors, as determined by the Seigneure expression profiles, showed worse locoregional control, progression-free survival and overall survival.

Hypoxia promotes invasion, immune escape and metastasis, and has been described as a poor prognostic marker in solely surgically treated patients. Hypoxia also causes less DNA damage by radiation, and consequently hypoxia biomarkers are not only prognostic but also predictive. Several trials have demonstrated the predictive potential of hypoxia biomarkers through their association with poor radiotherapy response and hypoxia modification therapy. The ARCON trial showed that regional control rates in HNSCC were better for hypoxic tumors when treated with carbogen and nicotinamide, but unfortunately, no difference in local tumor control was found¹¹. Hypoxia modification using nimorazole has also shown significant improvement in combination with radiotherapy as compared to radiotherapy alone in hypoxic tumors¹². A large meta-analysis of clinical trials showed a favorable impact of hypoxia modification during radiotherapy for hypoxic HNSCC¹³. Currently, the NIMRAD study, a randomized clinical trial for hypoxia modification, uses the Buffa *et al.* gene expression profile to classify hypoxia¹⁴. However, whether the analyzed hypoxia signatures in this thesis are predictive biomarkers, in the strict definition of the term, for radiotherapy response cannot be easily assessed. For this purpose, outcome associations will have to be compared to cohorts of patients treated differently, such as with different RT modalities or surgery.

EMT

In Chapter 3, we showed the impact of EMT in advanced-stage HPV-negative HNSCC patients treated with chemoradiation. Its robust involvement prompted us to develop a new HNSCC-specific EMT profile. In 2017, Puram *et al.* published data supporting the important role of EMT in HNSCC¹⁵. Using single-cell transcriptomic analysis on resected oral cavity squamous cell carcinomas, they identified tumor cells with mesenchymal characteristics. However, these tumor cells maintained mostly epithelial characteristics and they therefore call it partial EMT (pEMT). Where at first EMT was thought of as describing distinctive states it is now believed to be a gradient where tumor cells can have more or less properties of both

mesenchymal and epithelial cells^{15,16}. Importantly, they also show that the classical EMT signal in bulk tumors can be largely attributed to cancer-associated fibroblasts (CAFs).

Although our EMT profile, which uses RNA-sequencing data on bulk tumor material, did not correlate to the amount of stroma in the sample (Chapter 3, Figure 6), we cannot rule out that our EMT profile is also partially associated with the proportion of CAFs. In the meantime, several deconvolution methods for bulk sequencing have been described¹⁷. Tyler *et al.* specifically focused on EMT¹⁷. They revealed that the EMT signal in bulk tumor is attributed mainly to the amount of CAF, and developed a pEMT signature for bulk RNA-sequencing. Although conventional EMT gene expression profiles are associated with outcome in many cancer types, this pEMT signature is not associated with clinical factors in most cancer types, except for HNSCC. Tested in the heterogeneous TCGA data, the pEMT signature is associated with lymph node metastasis, therapy resistance/patient outcome in HNSCC¹⁷. Their CAF signature shows moderate correlations with clinical factors in many cancer types, including HNSCC, underlining the general importance of CAFs in cancer. In bulk EMT gene expression profiles, pEMT and CAF signaling are hard to distinguish. In our study however, when comparing vimentin-positive tumor cell content by immunohistochemistry with EMT classification by our models, we find a reasonable association and an enrichment of vimentin-rich tumors in the mesenchymal classified (Chapter 3, Figure 6). This indicates that these models may be able to distinguish signals from CAF and tumor cells. Future endeavours will show which is most important, or whether bulk EMT gene expression will suffice in guiding treatment direction.

The negative impact of tumor EMT status on patient outcome is in line with previous studies conducted, which have been published since 2017^{18,19}. All patients in our cohort are treated with chemoradiation. EMT is associated with radiotherapy resistance²⁰. EMT is also associated with chemoresistance in many cancer types including HNSCC^{21–25}. Interestingly and in line with our analysis of locoregional outcome associations (Chapter 3), the German Cancer Consortium Radiation Oncology Group published data showing that EMT is associated with worse locoregional control in patients treated with post-operative chemoradiotherapy²⁶.

EMT is a complex process that is induced via many different pathways for example TGF- β , WNT and the Notch pathway²⁰. Hypoxia is an already known important prognostic factor in HNSCC and is linked to EMT. Hypoxia induces EMT signalling pathways such as TGF- β , WNT and Notch. Hypoxia also directly modulates EMT-related transcription factors. HIF1 α and β have been shown to induce EMT via upregulation of proteins such as TWIST1, SNAIL1 and ZEB^{16,27,28}. Although not strong, Chapter 3 Figure 1B, shows a correlation between

GCTGACCAACATGGTGAACCCCGTTCTAAATAAAATACAAAATCAGGTTGGGACGGTGGCTACGGCTGTAAATCCAGACATTTGGGAGGCGAGGTGGGAGATCAAGAGTCAGGAGTTCAAGACAGCGCTGACCAATGTGGTGA

EMT and both acute and chronic hypoxia. Although this association does not necessarily mean causation, hypoxia modifiers may be a strategy to target EMT¹⁶. Besides hypoxia, all previously mentioned pathways provide potential targets to reduce EMT. Notch targeting, for instance, has been shown to increase radiosensitivity by blocking EMT^{29,30}. A potential problem obviously is the fact that *NOTCH1* is a genuine tumor suppressor gene in HNSCC. *SNAI1* targeting in oral squamous cell carcinoma cell lines has shown increased radiosensitivity³¹. TGF β is among the most strongly related pEMT inducers, and inhibitors will also target pEMT¹⁶.

DNA-repair

DNA-repair deficiency in tumors, and the homologous recombination/Fanconi anemia (HR/FA) DNA repair pathways in particular, is associated with poor survival^{32,33}. Fanconi anemia patients with germline mutations in these genes are often confronted with HNSCC at a young age. However, the role of HR/FA pathway defects is not fully understood in HNSCC. Several studies have shown the relevance of HR/FA in resolving DNA damage caused by crosslinkers and cisplatin. Pre-clinical studies in HNSCC cell lines have shown that functional defects in the HR/FA pathways determine crosslinker and cisplatin sensitivity^{32,34}. Recently, Burcher *et al.* showed in 2021, that pathogenic mutations in DNA repair genes are present in up to 17.6% of patients with HNSCC. In line with our data, they show that patients with mutations in DNA repair genes are associated with poorer outcomes³⁵.

Some DNA-sequencing-based strategies have been deployed to assess homologous recombination deficiency (HRD) in tumors by generating predictors based on mutational signatures or copy number aberration profiles. Some of these so-called “HRD scar” based signatures originated from breast cancer studies and identify genomic aberrations specific for HRD as caused by BRCA1 or 2 (HR/FA members) mutations³⁶. Association with cisplatin sensitivity is assessed in large pan-cancer analysis that also includes HNSCC³⁶. Originated and validated in the BRCA mutational context these biomarkers may however not capture genomic scars resulting from other HR/FA member mutations such as those identified in HNSCC^{32,33}. Another downside of these HRD scar signatures is that they measure the effects of possible HRD and not the current HRD status. In our study in Chapter 4, we aimed to develop a RNA-sequencing-based profile to identify tumors with FA/HR deficiency. This profile proved to be associated with patient survival. Since then, only a few studies have studied the association of DNA repair defects, as determined by transcriptomics to patient outcomes in HNSCC. Several studies have shown that low expression of individual DNA-repair genes, for example, ATM and BRCA1, are associated with poor patient outcome

in HNSCC^{37,38}. Ming *et al.* constructed a gene expression profile of 13 DNA repair-related genes³⁹. Their profile was associated with patient outcome in HNSCC patients from the TCGA database in univariate and multivariable analyses. However, HPV was not included in any analysis. In addition, hypoxia interaction or cisplatin response were not analyzed. The authors, however, did show an association with immune response and tumor mutational burden³⁹. Different from our study (Chapter 4), those studies assumed that repair activities are linked with DNA repair gene expression levels. Those levels are, however, largely cell cycle regulated and impacted by hypoxia^{9,40–42}. Moreover, expression of many of these proteins are regulated at the post-transcriptional level.

We show that the DNA-repair deficient tumors have a poor prognosis, but this appears to be only evident in normoxic tumors. Hypoxic tumors already have a comparable poor prognosis and we do not find an additional poor prognostic value of our DNA-repair gene expression profile in these tumors. This could raise the question of whether our “DNA-repair defect prediction models” are still able to identify repair defects in hypoxic tumors. Possible explanations for this observation are: 1) the models were generated from cells grown in 5% oxygen and not yet adapted to hypoxic conditions. The transcriptome changes largely under hypoxic conditions, thereby possibly altering discriminatory features that the models rely on. 2) Several studies showed that DNA repair mechanisms, especially HR/FA pathways, are impaired by prolonged and severe hypoxia. They suggest that most hypoxic tumors are more or less DNA repair deficient. This is underlined by the number of shared downstream differentially expressed genes as shown in Chapter 4, supplemental Figure S7. 3) Hypoxia already affects DNA repair capacity and because of the similar downstream expression effects, survival is possibly not further impaired by DNA deficiency. This would result in no additional value of our DNA repair deficiency profile in a fraction of the patients for prognosis but can still be important to guide therapy.

In Chapter 4, our data also suggest a link between DNA repair and increased invasiveness. This link is in line with the conclusions described by Bakhoum *et al.*⁴³. In a subsequent analysis (Chapter 4, Figure 7 and Figure S13) we show a relation between DNA repair deficiency and EMT. Recent studies reveal a crucial link between ZEB1, an EMT-related transcription factor, and the regulation of DDR signaling and DSB repair pathways^{44–46}. Others suggest a link between NBS1 levels, a protein involved in DNA repair, and increased levels of SNAIL, another EMT-related transcription factor, in HNSCC tumors^{47,48}. The exact mechanism or link between DNA repair and EMT remains unclear and future studies will reveal their exact relation.

Challenges in biomarker research in HNSCC

Heterogeneity

Cancer is a heterogeneous disease^{49–52}. Heterogeneity is found in several layers of the disease. Intertumor heterogeneity means that there are differences between individual tumors. The obvious example is between HPV-positive and HPV-negative oropharyngeal cancers. But even a laryngeal tumor in one patient is not the same as a laryngeal tumor in another patient. Even though etiological factors are often similar, smoking, for instance, the underlying mutational pattern can be very different. This results in different geno- and phenotypes and ultimately in different responses to therapy. In this thesis, we tried to unravel part of these biological differences to explain treatment response, which can ultimately lead to alternative treatments.

Intratumor heterogeneity describes differences within a tumor. Tumors have regions with different mutations or micro-environmental conditions. Head and neck squamous cell carcinoma is no exception and is a heterogeneous disease with hypoxic regions and usually multiple clones present⁵². Intratumor heterogeneity can lead to sampling error when a biopsy sample is used to characterize the whole tumor. Previous analyses of expression data conducted in our institute show that biopsies from a single tumor, cluster in more than 80% of the cases together and not with samples from other tumors, indicating that intertumor heterogeneity is far greater than intratumor heterogeneity⁵³. Toustrup *et al.* addressed the question by looking at hypoxia classification of 2–4 samples from a single tumor⁵⁴. In 70% of tumors, there was full concordance and samples were similarly classified. Tumor percentage in the biopsy seemed to play a role in classification. The higher the tumor percentage, the higher the concordance. Similar results are available for lung carcinoma showing similar proportions of concordance in EMT gene expression profile classifications⁵⁵, confirming that heterogeneity across different individual tumors is greater than within a tumor.

Heterogeneity itself can also be an important feature of the tumor and has been quantified via several methods, mostly DNA sequencing-based⁵⁶. Determining heterogeneity using RNA-sequencing is limited to single-cell RNA-sequencing, and quantitative analysis methods using single or bulk RNA-sequencing data are not available yet. Mroz *et al.* describe the tumor DNA sequence-based MATH method in HNSCC⁴⁹. The mutant-allele tumor heterogeneity (MATH) score is calculated as the distribution of mutant allele fractions of individual mutations. A high MATH score, i.e. very heterogeneous tumor, is associated with a worse prognosis and related to chemoradiotherapy response in head and neck squamous cell carcinoma^{57,58}. Several other studies revealed similar associations, where more heterogeneous HNSCCs

have a worse prognosis^{59,60}. Many links between DNA repair deficiency and intratumor heterogeneity have been described⁶¹.

In this thesis, we used single biopsy samples and selected samples with a tumor content of more than 40%. This cutoff was chosen to increase the likelihood of analyzing tumor signal and to decrease but not to exclude the signal of stromal/immune and normal cells in the bulk-sequenced material. Most studies use a minimal tumor percentage by excluding highly stroma-enriched tumor sections. As a consequence, there is a risk of a selection bias with largely stroma-rich tumors being excluded, excluding important stromal compartments such as cancer-associated fibroblasts or immune cells. Analysis of our data set, which is limited to the 0-60% stroma content specimens, however, showed no correlation between stroma content and the mesenchymal classification (Chapter 3, Figure 7). In-depth heterogeneity analysis opportunities are limited for bulk RNA-sequencing data and were therefore not included in our analysis. While single-cell sequencing and data analysis may describe tumor heterogeneity and the relevance of particular clonal evolutions or tumor cell relations well, its opportunities to establish prognostic or predictive biomarkers are currently still limited as this requires numerous sequencing sources within and across different tumors to establish robust links with outcome.

Sample size and confounders

In this thesis, we collected two datasets of 98 and 76 HPV-negative HNSCC cases, respectively. By collaborating with other centers, we were able to increase patient numbers to improve power. It was calculated that these four centers were responsible for roughly 50% of patients in the Netherlands, which would account for 420 patients in the four years in which the initial retrospective discovery set was collected. In the end, we included 379 patients for clinical analysis⁶². Greatly due to limited availability of frozen tumor biopsy material, but also due to exclusion by a too low tumor percentage in available biopsies or poor quality of the material after RNA isolation and/or sequencing, we were able to analyze 98 patients in the discovery cohort and 76 in the validation cohort. Despite these limitations and compared to other transcriptomic studies, the total number of patients used in our studies is relatively large^{63,64}. The total number of patients for the biomarker studies is sufficient for most univariate analyses or limited but important multivariate analyses.

Importantly and as also highlighted above, analyses need to be placed in the context of other important clinical and biological factors. This is well illustrated by our additional TCGA analysis in Chapter 5. High CD8⁺ T cells have been associated with a good prognosis in the TCGA data⁴. Upon closer inspection, this previously reported association of good prognosis can be attributed to the many HPV-positive oropharyngeal carcinomas in the TCGA database

GCTTGACCAACATGGTGAACCGCTTCTAAATAAAATAAAAAATAGGTTGGGACGGTGGCTACGGCTGTAAATCCAGCACTTTGGGAGGCGAGGTGGGAGATCAAGAGGTGAGGAGTTAAGACAGCTGACCAATGTGTGA

(Chapter 5, Supplemental Figure 9). We found a poor association of high CD8⁺ T cells with overall survival in HPV-negative tumors. This further underlines the different cancer biology between HPV-positive and HPV-negative HNSCC and the need for adjusting for other determinants of outcome.

This example illustrates the importance of the interpretation of data in a contextual background and the need for multivariable testing. However, this is only possible to a certain extent, since many clinical and/or biological factors that determine outcome in HNSCC are still unknown. Nevertheless, multivariate analyses are important to account for known confounding factors, whether clinical or biological. Yet, such analyses increase the need for larger datasets as the number of events should support the number of variables in a multivariate analysis to enable sufficient statistical power to depict significant differences. The possibility to increase the size of the dataset is often limited, especially in relatively rare tumors such as head and neck cancer in the Netherlands.

The robustness of biomarkers is of vital importance and validation in independent cohorts is key in clinical evaluation. Many biomarker studies using omics data fail in later validation stages⁶⁵. Recent omics-based biomarker studies use external validation, which is an increasingly important aspect of validation⁶⁵. The *MammaPrint*[™], the most well-known and FDA-approved prognostic signature, was developed using 117 patients and validated externally multiple times in thousands of patients. This shows the effort that has to be put in, to validate an omics test. The limited case numbers and data available in HNSCC is an extra hurdle to take when validating HNSCC-specific biomarkers. But it is possible as also the N-stage RNA expression profile was convincingly validated in a Dutch nation-wide effort⁶⁶.

In this thesis, we primarily used RNA-sequencing data to determine the clinical impact of certain biological processes. This was dictated by earlier studies in our group that determined the, partly limited prognostic value of individual and sufficiently frequent tumor-specific gene mutations, as determined by DNA-sequencing⁶⁷. Limited HNSCC case numbers that largely exclude the analysis of HNSCC typical mutations with a frequency of less than 20% obstruct a broader study on outcome associations. Copy number variation (CNV) profiles, as determined by shallow DNA-sequencing, however, do show some prognostic value, particularly for the more “silent” tumors with low numbers of DNA breaks⁶⁸. RNA-sequencing provides several advantages. RNA-sequencing data reflects the dynamic biological and micro-environmental and contextual states of the tumor cell, better than a genetically confined state. Transcriptomic analyses enable pathway analysis to identify important cellular processes. Another advantage is that prognostic modeling with transcriptomic data usually shows better performance compared to mutational or CNV data^{69,70}.

Future research emphasizes the combination of several different data sources, where for instance RNA-sequencing, DNA-sequencing and radiomic data will be combined to more accurately predict patient response to certain treatments or will be related to patient-reported outcome measures (PROMs)^{70–72}. These so-called multi-omics analyses require much bigger and more homogeneous cohorts and larger datasets that are difficult to establish due to limited access to patients, funds and resources. To address these limitations, tools have been developed to calculate sample sizes for such multi-omics projects⁷³. Large consortia and multi-national and multi-center collaborations are necessary while certain databases such as the TCGA and the Hartwig Foundation in the Netherlands can help to enlarge datasets^{74,75}. Good clinical data are essential and national registries such as the DHNA (Dutch Head and Neck Audit) show the growing attention to standardized methods of data collection and provide easier access to clinical data⁷⁶. Combining the DHNA with other non-clinical databases such as the pathology PALGA database, IKNL or genetics data in the Hartwig Foundation will be the next step. Computational advancements such as artificial intelligence, machine learning and bioinformatics will help to overcome previously mentioned cohort size problems and maximize robustness and reproducibility. Technological advancements will reduce costs, for example, sequencing costs, and thereby increase possibilities for larger datasets.

Limitations and strengths

A major advantage of this study is that treatment for advanced-stage head and neck cancer is standardized. Patient selection, treatment modality and intensity and follow-up are the same for all clinics throughout the Netherlands. This makes data from different centers comparable and provides the opportunity to combine data from different centers to acquire larger datasets. Therefore, our datasets are fairly homogenous concerning cancer type (all oro-, hypopharyngeal or laryngeal HPV-negative cancers), stage (all advanced stage) and treatment (all cisplatin-based chemoradiotherapy). By homogenizing the datasets as much as possible we tried to eliminate the impact of these confounders. To accurately evaluate the true impact of an investigated factor, it would be best to compare groups that are equal in all factors except for the factor one wants to test. Except for randomized controlled trials, in clinical research, this is an illusion, although many efforts are undertaken to homogenize the study groups. Moreover, it will make extrapolation of results to other patients much more difficult. Our conclusions and the interpretation of the results in this thesis should thus be placed in the context of several strengths and limitations.

The Netherlands has a well-organized health-care system where head and neck cancer care is centralized in 14 centers, and treatment and follow-up are protocolled. Due to protocolled follow-up, very little data was missing and few patients were lost to follow-up which

results in good quality clinical data and limited selection bias. Together with a relatively high number of patients, this allowed us to do multivariable analyses. We were able to test the robustness and reliability of novel and reported findings with increased confidence. The quality of the RNA of our fresh-frozen material was also good, which resulted in little exclusion and good RNA-sequencing data. Compared to other similar datasets, we have a relatively large dataset. Using machine learning and cross-validation the reproducibility and robustness (Chapter 3 and 4) could be increased. Ultimately, findings were validated in independent cohorts or pre-clinical models or both (all Chapters).

As explained before, heterogeneity is a complicating factor in clinical research. Although we have a large cohort that is fairly homogeneous compared to other datasets, our studies were still affected by heterogeneity to a certain degree. For example, all patients received cisplatin-based chemoradiotherapy, but the final cumulative cisplatin dose that patients received differed due to treatment disruptions caused by toxicities. This consequently required the inclusion of this variable in our analyses, as cumulative cisplatin dose impacts survival outcomes and was found to be important in determining the prognostic strength of some of the biomarkers we tested (e.g. DNA repair defect)⁷⁷.

We restricted the variables that were to be included in our multivariate analyses to those that were repeatedly and significantly associated with the investigated endpoint as the number of variables in our multivariable analysis was restricted by the number of events and patients in our cohorts. Yet, ideally and supposedly enabled by sufficiently large cohorts, additional variables and biomarkers should have been analyzed in multivariate analyses to establish their influence in the context and with respect to each other.

Another limitation of our study is the single biopsy use as tumors themselves are also heterogeneous. Biopsies are only a limited sample of the tumor. While some data, as outlined above, suggest it may suffice for some biomarkers for prognosis, it is unknown whether these single biopsy samples can be representative of the whole tumor for all.

Conclusion and personal future perspective

In this thesis, we show that cancer biology plays an important role in patient outcome. Using RNA sequencing we show an important role for hypoxia, EMT and DNA repair, all significantly and repeatedly impacting HPV-negative HNSCC outcomes in our Dutch chemoradiotherapy-treated cohorts. These biological attributes are and will be exploitable by different types of drugs soon. Hypoxia modifiers are currently already under investigation in clinical trials. HPV, DNA-repair inhibitors and immunotherapy are following closely.

I believe that many therapies are not being used because they are either not effective or not cost-effective in the absence of adequate biomarkers to select responding patients.

Future research will point toward more accurate patient selection for personalized treatment. Discovery studies to identify novel biomarkers will require large datasets, a relatively homogeneous patient population and multi-omics analysis. Collaborations between institutes and perhaps nations are inevitable to gain the required power for these analyses. Pre-clinical studies will provide essential insights to develop hypothesis-driven studies and will lay the foundations for clinical drug trials. In the Netherlands, head and neck cancer is a less frequent disease so patients and materials are scarce. It is of utmost importance that all patients will be offered opportunities to participate in clinical trials. In future clinical trials for anti-cancer drugs, it should be obligated to collect tissue samples and endeavour treatment-specific biomarker research such as the IMCISION study⁷⁸. Perhaps future clinical trials will not only be powered to identify a potential drug benefit but also to enable biomarker research. It is the clinician's responsibility to gain the most healthcare benefit for the patient but also for society. Since patients and patient material is scarce and valuable, it is the clinician's job to make sure these resources are maximally utilized. Future research will further integrate medicine, biology, bioinformatics, pharmacology and technological advancements. For future clinicians, it is impossible to attain all the required skills themselves and they more and more serve as a "connector between physicians, laboratory experts and researchers in the basic, computer, and clinical sciences"⁷⁹.

References

1. Leemans CR, Snijders PJF, Brakenhoff RH. The molecular landscape of head and neck cancer. *Nat Rev Cancer*. 2018;18(5):269-282. doi:10.1038/nrc.2018.11
2. Sabatini ME, Chiocia S. Human papillomavirus as a driver of head and neck cancers. *Br J Cancer*. 2020;122:306-314. doi:10.1038/s41416-019-0602-7
3. Kian Ang K, Harris J, Wheeler R, *et al*. Human Papillomavirus and Survival of Patients with Oropharyngeal Cancer A BS TR AC T. *N Engl J Med*. 2010;363:24-35. doi:10.1056/NEJMoa0912217
4. Mandal R, Senbabaoglu Y, Desrichard A, *et al*. The head and neck cancer immune landscape and its immunotherapeutic implications. *JCI insight*. 2016;1(17):e89829. doi:10.1172/jci.insight.89829
5. Li S, Wang Y, Sun R, *et al*. Single-Cell Transcriptome Analysis Reveals Different Immune Signatures in HPV-and HPV+Driven Human Head and Neck Squamous Cell Carcinoma. *Hindawi J Immunol Res*. 2022;2022. doi:10.1155/2022/2079389
6. Chung CH, Parker JS, Ely K, *et al*. Gene expression profiles identify epithelial-to-mesenchymal transition and activation of nuclear factor- κ B Signaling as characteristics of a high-risk head and neck squamous cell carcinoma. *Cancer Res*. 2006;66(16):8210-8218. doi:10.1158/0008-5472.CAN-06-1213
7. Pramana J, Van den Brekel MWM, van Velthuysen MLF, *et al*. Gene Expression Profiling to Predict Outcome After Chemoradiation in Head and Neck Cancer. *Int J Radiat Oncol Biol Phys*. 2007;69(5):1544-1552. doi:10.1016/j.ijrobp.2007.08.032
8. Brown JM. Evidence for acutely hypoxic cells in mouse tumours, and a possible mechanism of reoxygenation. *Br J Radiol*. 1979. doi:10.1259/0007-1285-52-620-650
9. Chan N, Koritzinsky M, Zhao H, *et al*. Chronic hypoxia decreases synthesis of homologous recombination proteins to offset chemoresistance and radioresistance. *Cancer Res*. 2008. doi:10.1158/0008-5472.CAN-07-5472
10. Bayer C, Vaupel P. Acute versus chronic hypoxia in tumors. *Strahlentherapie und Onkol*. 2012. doi:10.1007/s00066-012-0085-4
11. Janssens GO, Rademakers SE, Terhaard CH, *et al*. Accelerated radiotherapy with carbogen and nicotinamide for laryngeal cancer: Results of a phase III randomized trial. *J Clin Oncol*. 2012;30(15):1777-1783. doi:10.1200/JCO.2011.35.9315
12. Overgaard J, Hansen HS, Overgaard M, *et al*. A randomized double-blind phase III study of nimorazole as a hypoxic radiosensitizer of primary radiotherapy in supraglottic larynx and pharynx carcinoma. Results of the Danish Head and Neck Cancer Study (DAHANCA) Protocol 5-85. *Radiother Oncol*. 1998;46(2):135-146. doi:10.1016/S0167-8140(97)00220-X
13. Overgaard J. Hypoxic modification of radiotherapy in squamous cell carcinoma of the head and neck - A systematic review and meta-analysis. *Radiother Oncol*. 2011. doi:10.1016/j.radonc.2011.03.004
14. Thomson D, Yang H, Baines H, *et al*. NIMRAD - A phase III trial to investigate the use of nimorazole hypoxia modification with intensity-modulated radiotherapy in head and neck cancer. *Clin Oncol*. 2014. doi:10.1016/j.clon.2014.03.003
15. Puram S V, Tirosch I, Parikh AS, *et al*. Single-Cell Transcriptomic Analysis of Primary and Metastatic Tumor Ecosystems in Head and Neck Cancer. *Cell*. 2017;172:1-14. doi:10.1016/j.cell.2017.10.044
16. Pal A, Thomas •, Barrett F, *et al*. Partial EMT in head and neck cancer biology: a spectrum instead of a switch. *Oncogene*. 2021;40:5049-5065. doi:10.1038/s41388-021-01868-5

17. Tyler M, Tirosh I. Decoupling epithelial-mesenchymal transitions from stromal profiles by integrative expression analysis. *Nat Commun.* 2021;12(1):1-13. doi:10.1038/s41467-021-22800-1
18. Wan Y, Liu H, Zhang M, *et al.* Prognostic value of epithelial-mesenchymal transition-inducing transcription factors in head and neck squamous cell carcinoma: A meta-analysis. *Head Neck.* 2020;42(5):1067-1076. doi:10.1002/HED.26104
19. Mayhew GM, Uronis JM, Hayes DN, Zevallos JP. Mesenchymal gene expression subtyping analysis for early-stage human papillomavirus-negative head and neck squamous cell carcinoma reveals prognostic and predictive applications. *Front Oncol.* 2022;12:4401. doi:10.3389/FONC.2022.954037/BIBTEX
20. Zhou S, Zhang M, Zhou C, Wang W, Yang H, Ye W. The role of epithelial-mesenchymal transition in regulating radioresistance. *Crit Rev Oncol Hematol.* 2020;150(150):102961. doi:10.1016/j.critrevonc.2020.102961
21. Baumeister P, Zhou J, Canis M, Gires O. Epithelial-to-mesenchymal transition-derived heterogeneity in head and neck squamous cell carcinomas. *Cancers (Basel).* 2021;13(21). doi:10.3390/CANCERS13215355
22. Fischer KR, Durrans A, Lee S, *et al.* Epithelial-to-mesenchymal transition is not required for lung metastasis but contributes to chemoresistance. *Nature.* 2015;527(7579):472-476. doi:10.1038/nature15748
23. Zheng X, Carstens JL, Kim J, *et al.* Epithelial-to-mesenchymal transition is dispensable for metastasis but induces chemoresistance in pancreatic cancer. *Nature.* 2015;527(7579):525-530. doi:10.1038/nature16064
24. Papadaki MA, Stoupis G, Theodoropoulos PA, Mavroudis D, Georgoulas V, Agelaki S. Companion Diagnostic, Pharmacogenomic, and Cancer Biomarkers Circulating Tumor Cells with Stemness and Epithelial-to-Mesenchymal Transition Features Are Chemoresistant and Predictive of Poor Outcome in Metastatic Breast Cancer. doi:10.1158/1535-7163.MCT-18-0584
25. Fustaino V, Presutti D, Colombo T, *et al.* Characterization of epithelial-mesenchymal transition intermediate/hybrid phenotypes associated to resistance to EGFR inhibitors in non-small cell lung cancer cell lines. www.impactjournals.com/oncotarget. Accessed November 11, 2022.
26. Patil S, Tawk B, Grosser M, *et al.* Analyses of molecular subtypes and their association to mechanisms of radioresistance in patients with HPV-negative HNSCC treated by postoperative radiochemotherapy. *Radiother Oncol.* 2022;167:300-307. doi:10.1016/J.RADONC.2021.12.049
27. Hapke RY, Haake SM. Hypoxia-induced epithelial to mesenchymal transition in cancer. *Cancer Lett.* 2020;487:10-20. doi:10.1016/J.CANLET.2020.05.012
28. Saxena K, Jolly MK, Balamurugan K. Hypoxia, partial EMT and collective migration: Emerging culprits in metastasis. *Transl Oncol.* 2020;13(11):100845. doi:10.1016/j.tranon.2020.100845
29. Zhou S, Zhang M, Zhou C, Wang W, Yang H, Ye W. The role of epithelial-mesenchymal transition in regulating radioresistance. *Crit Rev Oncol Hematol.* 2020;150. doi:10.1016/J.CRITREVONC.2020.102961
30. Wang J, Kang M, Wen Q, *et al.* Berberine sensitizes nasopharyngeal carcinoma cells to radiation through inhibition of Sp1 and EMT. *Oncol Rep.* 2017;37(4):2425-2432. doi:10.3892/OR.2017.5499/HTML
31. Jiang F, Zhou L, Wei C, Zhao W, Yu D. Slug inhibition increases radiosensitivity of oral squamous cell carcinoma cells by upregulating PUMA. *Int J Oncol.* 2016;49(2):709-719. doi:10.3892/IJO.2016.3570/HTML

GCTGACCAACATGGTGAACCCGCTTCTAAATAAAATAAAAAATAGGTGGGACGGTGGTCAAGCTTAAATCCAGCACTTTGGGAGGCGAGGTGGGAGATCAAGAGTCAAGGAGTTAAGACAGCTGACCAATGTGGTGA

32. Verhagen CVM, Vossen DM, Borgmann K, *et al.* Fanconi anemia and homologous recombination gene variants are associated with functional DNA repair defects in vitro and poor outcome in patients with advanced head and neck squamous cell carcinoma. *Oncotarget*. 2018;9(26):18198-18213. doi:10.18632/oncotarget.24797
33. Stoepker C, Ameziane N, Van Der Lelij P, *et al.* Defects in the Fanconi anemia pathway and chromatid cohesion in head and neck cancer. *Cancer Res*. 2015;75(17):3543-3553. doi:10.1158/0008-5472.CAN-15-0528
34. Burkitt K, Ljungman M. Compromised Fanconi anemia response due to BRCA1 deficiency in cisplatin-sensitive head and neck cancer cell lines. *Cancer Lett*. 2007;253(1):131-137. doi:10.1016/J.CANLET.2007.01.017
35. Burcher KM, Fauchoux AT, Lantz JW, *et al.* Prevalence of DNA Repair Gene Mutations in Blood and Tumor Tissue and Impact on Prognosis and Treatment in HNSCC. *Cancers* 2021, Vol 13, Page 3118. 2021;13(13):3118. doi:10.3390/CANCERS13133118
36. van der Wiel AMA, Schuitmaker L, Cong Y, *et al.* Homologous Recombination Deficiency Scar: Mutations and Beyond—Implications for Precision Oncology. *Cancers (Basel)*. 2022;14(17). doi:10.3390/cancers14174157
37. Wang YC, Lee KW, Tsai YS, *et al.* Downregulation of ATM and BRCA1 Predicts Poor Outcome in Head and Neck Cancer: Implications for ATM-Targeted Therapy. *J Pers Med* 2021, Vol 11, Page 389. 2021;11(5):389. doi:10.3390/JPM11050389
38. Bold IT, Specht AK, Droste CF, *et al.* DNA Damage Response during Replication Correlates with CIN70 Score and Determines Survival in HNSCC Patients. *Cancers* 2021, Vol 13, Page 1194. 2021;13(6):1194. doi:10.3390/CANCERS13061194
39. Ming R, Wang E, Wei J, Shen J, Zong S, Xiao H. The Prognostic Value of the DNA Repair Gene Signature in Head and Neck Squamous Cell Carcinoma. *Front Oncol*. 2021;11(July). doi:10.3389/fonc.2021.710694
40. Bindra RS, Schaffer PJ, Meng A, *et al.* Down-Regulation of Rad51 and Decreased Homologous Recombination in Hypoxic Cancer Cells. *Mol Cell Biol*. 2004;24(19):8504-8518. doi:10.1128/MCB.24.19.8504-8518.2004/ASSET/A098FCF4-8BDC-44DD-8DD1-35E041763B9E/ASSETS/GRAPHIC/ZMB0190444190008.JPEG
41. Begg K, Tavassoli M. Inside the hypoxic tumour: reprogramming of the DDR and radioresistance. *Cell Death Discov* 2020 61. 2020;6(1):1-15. doi:10.1038/s41420-020-00311-0
42. Kumareswaran R, Ludkovski O, Meng A, Sykes J, Pintilie M, Bristow RG. Chronic hypoxia compromises repair of DNA double-strand breaks to drive genetic instability. *J Cell Sci*. 2012;125(1):189-199. doi:10.1242/JCS.092262
43. Bakhoum SF, Ngo B, Laughney AM, *et al.* Chromosomal instability drives metastasis through a cytosolic DNA response. *Nature*. 2018;553(7689):467-472. doi:10.1038/nature25432
44. Zhang P, Wei Y, Wang L, *et al.* ATM-mediated stabilization of ZEB1 promotes DNA damage response and radioresistance through CHK1. *Nat Cell Biol*. 2014;16(9). doi:10.1038/ncb3013
45. Zhang X, Zhang Z, Zhang Q, *et al.* ZEB1 confers chemotherapeutic resistance to breast cancer by activating ATM. *Cell Death Dis* 2018 92. 2018;9(2):1-15. doi:10.1038/s41419-017-0087-3
46. Moyret-Lalle C, Prodhomme MK, Burlet D, *et al.* Role of EMT in the DNA damage response, double-strand break repair pathway choice and its implications in cancer treatment. *Cancer Sci*. 2022;113(7):2214-2223. doi:10.1111/CAS.15389
47. Nathansen J, Meyer F, Müller L, Schmitz M, Borgmann K, Dubrovskaya A. Beyond the Double-Strand Breaks: The Role of DNA Repair Proteins in Cancer Stem-Cell Regulation. *Cancers* 2021, Vol 13, Page 4818. 2021;13(19):4818. doi:10.3390/CANCERS13194818

48. Yang MH, Chang SY, Chiou SH, *et al.* Overexpression of NBS1 induces epithelial–mesenchymal transition and co-expression of NBS1 and Snail predicts metastasis of head and neck cancer. *Oncogene* 2007 2610. 2006;26(10):1459-1467. doi:10.1038/sj.onc.1209929
49. 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):211-215. doi:10.1016/J.ORALONCOLOGY.2012.09.007
50. Liu Q, Li L, Wang X. MYTH: An algorithm to score intratumour heterogeneity based on alterations of DNA methylation profiles. *Clin Transl Med.* 2021;11(10). doi:10.1002/CTM2.611
51. Dentre SC, Leshchiner I, Haase K, *et al.* Characterizing genetic intra-tumor heterogeneity across 2,658 human cancer genomes. *Cell.* 2021;184(8):2239-2254.e39. doi:10.1016/J.CELL.2021.03.009
52. Andor N, Graham TA, Jansen M, *et al.* Pan-cancer analysis of the extent and consequences of intratumor heterogeneity. *Nat Med.* 2016;22(1):105-113. doi:10.1038/NM.3984
53. Pramana J, Pimentel N, Hofland I, *et al.* Heterogeneity of gene expression profiles in head and neck cancer. *Head Neck.* 2007;29(12):1083-1089. doi:10.1002/HED.20621
54. Toustrup K, Sørensen BS, Metwally MAH, *et al.* Validation of a 15-gene hypoxia classifier in head and neck cancer for prospective use in clinical trials. *Acta Oncol (Madr).* 2016. doi:10.3109/0284186X.2016.1167959
55. Lee W-C, Diao L, Wang J, *et al.* Multiregion gene expression profiling reveals heterogeneity in molecular subtypes and immunotherapy response signatures in lung cancer. *Mod Pathol.* 2018;31:947-955. doi:10.1038/s41379-018-0029-3
56. Abécassis J, Hamy AS, Laurent C, *et al.* Assessing reliability of intra-tumor heterogeneity estimates from single sample whole exome sequencing data. *PLoS One.* 2019;14(11). doi:10.1371/JOURNAL.PONE.0224143
57. Mroz EA, Tward ADM, Hammon RJ, *et al.* Intra-tumor Genetic Heterogeneity and Mortality in Head and Neck Cancer : Analysis of Data from The Cancer Genome Atlas. *PLoS Med.* 2015;12(2):1-27. doi:10.7908/C1VH5KV4
58. Mroz EA, Tward AD, Pickering CR, Myers JN, Ferris RL, Rocco JW. High intra-tumor genetic heterogeneity is related to worse outcome in head and neck squamous cell carcinoma. *Cancer.* 2013;119(16):3034-3042. doi:10.1002/cncr.28150
59. Morris LGT, Riaz N, Desrichard A, *et al.* Pan-cancer analysis of intratumor heterogeneity as a prognostic determinant of survival. *Oncotarget.* 2016;7(9):10051-10063. doi:10.18632/ONCOTARGET.7067
60. Schrank TP, Lenze N, Landess LP, *et al.* Genomic heterogeneity and copy number variant burden are associated with poor recurrence-free survival and 11q loss in human papillomavirus-positive squamous cell carcinoma of the oropharynx. *Cancer.* 2021;127(15):2788-2800. doi:10.1002/CNCR.33504
61. Burrell RA, McGranahan N, Bartek J, Swanton C. The causes and consequences of genetic heterogeneity in cancer evolution. *Nat* 2013 5017467. 2013;501(7467):338-345. doi:10.1038/nature12625
62. de Roest RH, van der Heijden M, Wesseling FWR, *et al.* Disease outcome and associated factors after definitive platinum based chemoradiotherapy for advanced stage HPV-negative head and neck cancer. *Radiother Oncol.* 2022;175:112-121. doi:10.1016/j.radonc.2022.08.013
63. Tonella L, Giannoccaro M, Alfieri S, Canevari S, De Cecco L. Gene Expression Signatures for Head and Neck Cancer Patient Stratification: Are Results Ready for Clinical Application? *Curr Treat Options Oncol.* 2017;18(5). doi:10.1007/s11864-017-0472-2

GCCTGACCAACATGGTGAAACCCGTTCTAAATAAAATAAAAAATAGGTTGGGACGGTGGCTACGGCTGTAAATCCAGACATTTGGGAGGCGAGGTGGGAGATCAAGAGGTAGGAGTTAAGACAGCTGACCAATGTGGTGA

64. Serafini MS, Lopez-Perez L, Fico G, Licitra L, Cecco L De, Resteghini C. Transcriptomics and Epigenomics in head and neck cancer: available repositories and molecular signatures. *Cancers Head Neck*. 2020;5(1). doi:10.1186/S41199-020-0047-Y
65. Glaab E, Rauschenberger A, Banzi R, Gerardi C, Garcia P, Demotes J. Biomarker discovery studies for patient stratification using machine learning analysis of omics data: a scoping review. *BMJ Open*. 2021;11(12):e053674. doi:10.1136/BMJOPEN-2021-053674
66. Van Hooff SR, Leusink FKJ, 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):4104-4110. doi:10.1200/JCO.2011.40.4509
67. Vossen DM, Verhagen CVM, Van Der Heijden M, et al. Genetic factors associated with a poor outcome in head and neck cancer patients receiving definitive chemoradiotherapy. *Cancers (Basel)*. 2019. doi:10.3390/cancers11040445
68. Essers PBM, van der Heijden M, Vossen D, et al. Ovarian cancer-derived copy number alterations signatures are prognostic in chemoradiotherapy-treated head and neck squamous cell carcinoma. *Int J Cancer*. 2020;147(6):1732-1739. doi:10.1002/IJC.32962
69. Gómez-Rueda H, Martínez-Ledesma E, Martínez-Torteya A, Palacios-Corona R, Trevino V. Integration and comparison of different genomic data for outcome prediction in cancer. *BioData Min*. 2015;8(1):1-12. doi:10.1186/S13040-015-0065-1/FIGURES/4
70. Kim D, Shin H, Song YS, Kim JH. Synergistic effect of different levels of genomic data for cancer clinical outcome prediction. *J Biomed Inform*. 2012;45(6):1191-1198. doi:10.1016/j.jbi.2012.07.008
71. He J, Abdel-Wahab O, Nahas MK, et al. Integrated genomic DNA/RNA profiling of hematologic malignancies in the clinical setting. *Blood*. 2016;127(24):3004-3014. doi:10.1182/BLOOD-2015-08-664649
72. Willems SM, Abeln S, Feenstra KA, et al. The potential use of big data in oncology. *Oral Oncol*. 2019;98(March):8-12. doi:10.1016/j.oraloncology.2019.09.003
73. Tarazona S, Balzano-Nogueira L, Gómez-Cabrero D, et al. Harmonization of quality metrics and power calculation in multi-omic studies. *Nat Commun* 2020 111. 2020;11(1):1-13. doi:10.1038/s41467-020-16937-8
74. Priestley P, Baber J, Lolkema MP, et al. Pan-cancer whole-genome analyses of metastatic solid tumours. *Nat* 2019 5757781. 2019;575(7781):210-216. doi:10.1038/s41586-019-1689-y
75. Cancer Genome Atlas Research Network, Weinstein JN, Collisson EA, et al. The Cancer Genome Atlas Pan-Cancer analysis project. *Nat Genet*. 2013;45(10):1113-1120. doi:10.1038/ng.2764
76. Lydia FJ Van O, Robert P. T, Ludi E S, et al. The Dutch Head and Neck Audit: The First Steps. *J Head Neck Surg*. 2018;1(1):1-8. doi:10.36959/605/528
77. 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 Oto-Rhino-Laryngology*. 2017. doi:10.1007/s00405-017-4687-4
78. Vos JL, Elbers JBW, Krijgsman O, et al. Neoadjuvant immunotherapy with nivolumab and ipilimumab induces major pathological responses in patients with head and neck squamous cell carcinoma. doi:10.1038/s41467-021-26472-9
79. van Karnebeek CDM, Wortmann SB, Tarailo-Graovac M, et al. The role of the clinician in the multi-omics era: are you ready? *J Inherit Metab Dis*. 2018;41(3):571-582. doi:10.1007/S10545-017-0128-1



CHAPTER 7

Summary

Summary

Head and neck cancer transcriptomics. Profiling tumor biology for chemoradiotherapy response.

Patients with head and neck cancer often present with advanced-stage disease. Patients with advanced stage head and neck squamous cell carcinomas (HNSCC) have poor prognosis with a 5-year overall survival rate of around 50%. Advanced-stage oropharyngeal, hypopharyngeal or laryngeal carcinomas are often treated with organ-preserving concomitant cisplatin-based chemoradiotherapy. Surgery is an invasive and not an organ-preserving treatment alternative. Upfront prediction of likely treatment failure is necessary for patients to be eligible for alternative treatment strategies. Biomarkers can help to select patients who are at high risk of treatment failure. An already well-known biomarker is human papilloma virus (HPV). Tumors resulting from earlier HPV infection (HPV-positive tumors) have a good prognosis compared to HPV-negative tumors. Especially in HPV-negative HNSCC, there is a need for prognostic biomarkers to identify patients who will fail current standard treatments. The DESIGN consortium was designed to identify biomarkers to predict chemoradiotherapy response in HPV-negative HNSCC. This multicenter consortium consists of the AmsterdamUMC, the Netherlands Cancer Institute, the UMC Utrecht, the MAASTRO clinic and the Maastricht University. Each partner brought in their specific expertise and The Netherlands Cancer Institute and therefore this thesis focuses on the discovery and prognostic impact of genetic and transcriptomic biomarkers on the outcome of advanced stage HPV negative oropharyngeal, hypopharyngeal and laryngeal HNSCC patients treated with cisplatin-based chemoradiotherapy. This thesis uses bulk RNA-sequencing data to perform pathway analysis. Machine learning techniques were used to generate models representing head and neck cancer biology and to examine their clinical impact.

In **Chapter 2** we have a closer look at a well-known prognostic biomarker in HNSCC, hypoxia. Hypoxic tumors have a poorer prognosis compared to more normoxic tumors. Tumor cells can become hypoxic by chronic and acute mechanisms. Currently, there are four clinical transcriptomic hypoxia profiles published. It is unclear what their mutual relationships are and how they are related to chronic and acute hypoxia. We also set out to determine the clinical impact of acute and chronic hypoxia in several datasets.

The four clinical hypoxia profiles showed little overlap in the genes used for each profile. Combining three initial datasets, comprising 224 HPV patients in total, analyses of the four clinical hypoxia profiles showed that the profiles are strongly correlated to each other and chronic hypoxia. The four clinical profiles were averaged to construct a joint chronic hypoxia

score. In a subset of 91 patients, HPV-positive and HPV-negative tumors, chronic hypoxia was, although not significant, associated with local control, and acute hypoxia profile was significantly associated with local control. Analysis of an additional set of 174 HPV-negative HNSCC treated with chemoradiotherapy confirmed these findings. Univariate analyses showed that acute hypoxia was significantly associated with overall survival, progression-free survival, local and locoregional control. Chronic hypoxia was significantly associated with progression-free survival and showed a trend with overall survival, but not with locoregional control. Patients with high scores of both acute and chronic hypoxia showed the worst overall survival and local control. Multivariable analysis, in a limited subgroup of 149 patients, shows that acute hypoxia is significantly associated with overall survival, progression-free survival, local control and locoregional control. Chronic hypoxia was not associated with outcome in this cohort.

In **Chapter 3** we focus on epithelial-to-mesenchymal transition (EMT), a less well-known biomarker in HNSCC. EMT is the process where epithelial cells gain mesenchymal characteristics. EMT has been associated with outcome in many other cancer types but there is limited evidence in HNSCC. To assess the impact of EMT in HNSCC we extrapolated seven EMT gene expression profiles which were originally constructed for other cancer types. We applied these seven profiles in a training set of 98 patients with advanced-stage HPV-negative HNSCC from laryngeal, oro- and hypopharyngeal origin, treated with cisplatin-based chemoradiotherapy. All seven EMT profiles showed associations with overall survival, progression-free survival and distant metastasis-free survival in the training set. By comparing the most epithelial and most mesenchymal HNSCCs and using machine learning, we created a new HNSCC-specific EMT profile. The impact on outcome of the seven non HNSCC-specific models and the HNSCC-EMT model was confirmed in the validation set, which consisted of 78 patients with identical inclusion and exclusion criteria. Strong associations of EMT are seen in overall survival and progression-free survival. However, in the validation set EMT was not associated with distant metastasis-free survival but with locoregional control. To investigate the previously suggested link between distant metastasis and EMT, we analyzed additional primary tumors of patients who first presented with metastatic HNSCC. The primary tumor of patients with metastatic disease showed high HNSCC-EMT values, indicating an association with EMT. Multivariable analyses with TNM-stage in both cohorts showed that the impact of the HNSCC-EMT model is significant in overall and progression-free survival. A high HNSCC-EMT was correlated with T4-stage tumors and with a non-cohesive growth pattern. Finally, literature suggested that the EMT signal in bulk RNA-sequencing can be attributed to cancer-associated fibroblasts. We could not confirm

this finding. Using immunohistochemistry staining for vimentin, we could not correlate the HNSCC-EMT model scores to stroma content. This study showed that EMT is associated with outcome in advanced-stage HNSCC patients treated with chemoradiotherapy.

DNA crosslink repair defects are frequent manifestations in HNSCC and have been associated with outcomes in many cancer types. In **Chapter 4** we generated models to identify tumors with DNA crosslink repair defects in HNSCC cell lines and patients. Models were generated by machine learning algorithms using transcriptomic data of cell lines and their sensitivity to mitomycin C and olaparib. Sensitivity to these cross-linking agents was used as a proxy for functional DNA repair defects. Multiple feature selection methods were combined with multiple machine learning methods to determine which combination performs optimally in predicting DNA repair defect classes in these cell lines. Cross-validation was used to measure the robustness of the methods used. We tested the best-performing models in several datasets. First, in the Cancer Genome Project, many different HNSCC cell lines were subjected to a large number of compounds. The generated DNA-repair models correlated well with mitomycin C IC_{50} values of the HNSCC cell lines and predicted mutations in DNA crosslink repair genes. In a separate dataset, the models predicted the knockdown of DNA crosslink repair genes. This confirmed that the models could identify crosslink repair deficiency. Next, we evaluated the impact of the models on clinical outcome in two datasets of HPV negative advanced stage HNSCC patients treated with cisplatin-based chemoradiotherapy. DNA repair defected tumors, as classified by the models, showed a trend towards worse overall survival in both cohorts. The poor prognosis association was particularly strong in normoxic tumor samples and was linked to an increased risk of distant metastasis. Subsequent pathway analyses showed that the downstream effects of hypoxia and DNA repair defects are partially overlapping. This may explain the difference in outcome association of the models between normoxic and hypoxic tumors. The association of our models with distant metastasis prompted us to go back to *in vitro* studies and examine their migratory capabilities. In scratch and transwell assays, crosslink repair-defective HNSCC cell lines showed to be more migratory and invasive. In conclusion, the models are able to identify crosslink repair deficient tumors. DNA crosslink repair deficiency, according to the models, is associated with a poor prognosis in chemoradiotherapy-treated HNSCC.

To assess the links between the aforementioned biomarkers and biomarkers published in literature we employed a validation study combining the available data and most important biomarkers (hypoxia (chronic and acute), proliferation, DNA repair, stem-cell markers (SLC3A2 and CD44), EMT, EGFR and immune (T-cell infiltration score, CD8+ T-cells, CD56dim natural killer and CD8+ versus T regulatory cell ratios)) and their impact on radiation

response. In **Chapter 5**, we tested the mutual relations between these biomarkers and their clinical impact on locoregional control, as a marker for radiation response. For this purpose, we used a dataset of 197 patients with advanced stage HPV negative HNSCC treated with cisplatin-based chemoradiotherapy. Strong correlations were observed between the stem-cell markers and between acute hypoxia and T-cell infiltration score, CD8+ T-cell abundance and proliferation scores. A 10.000 times bootstrapping method containing multiple clinical variables was used to determine robust cutoffs for each marker. Chronic and acute hypoxia, EGFR expression and stem cellness markers were most robustly associated with locoregional control in multivariable analyses with only clinical variables. Multivariable analysis, considering both other biological and clinical markers, showed a significant role for chronic hypoxia, EGFR expression, CD8+ versus T regulatory cell ratios and T-cell infiltration signatures in locoregional control. In contrast to locoregional control, distant metastasis is predominantly influenced by EMT, EGFR and CD8+ versus T regulatory cell ratios. This external validation underscores the relevance of biological factors in determining chemoradiotherapy outcome in HNSCC.

Chapter 6 concludes this thesis with a general discussion, describing our findings in relation to recent literature. Challenges in biomarkers research are also described and we conclude with future perspectives in biomarker research.



CHAPTER 8

TTCCCAATCAACCCCTAGG99CTCCTCCTT99CTGCTGGGAGTTGTAGTCTGAACTTCTATCTT99CGAGAACTCCCTAC9CTCCCCCTACCGAGTCCC99GTAAATCTTTAAAGCACTGCAAC9CCCCCCCC999999CTG

Nederlandse samenvatting

Nederlandse samenvatting

Transcriptomics van hoofd-halskanker: profileren van tumor biologie voor de respons op chemoradiotherapie.

Patiënten met hoofd-halskanker (HHC) presenteren zich vaak in een gevorderd stadium van de ziekte. Ze hebben een slechte prognose met een 5-jaarsoverleving van ongeveer 50%. Gevorderde orofarynx-, hypofarynx- of larynxcarcinomen worden vaak behandeld met orgaansparende chemoradiotherapie op basis van cisplatinum. Chirurgie is een invasieve en geen orgaan sparend alternatief. Om patiënten in aanmerking te laten komen voor alternatieve behandelingen is het identificeren van patiënten waarbij de behandeling waarschijnlijk niet aanslaat noodzakelijk. Biomarkers kunnen helpen bij het selecteren van patiënten die een hoog risico lopen op falen van de behandeling. Een reeds bekende biomarker is infectie met het humaan papillomavirus (HPV). Tumoren die het gevolg zijn van een eerdere HPV-infectie (HPV-positieve tumoren) hebben een goede prognose in vergelijking met HPV-negatieve tumoren. De nood voor prognostische biomarkers is dan ook vooral groot bij HPV-negatieve HHC. Het DESIGN-consortium is samengebracht om biomarkers te identificeren om de respons te voorspellen op chemoradiotherapie in HPV-negatieve orofaryngeale, hypofaryngeale en laryngeale HHC-patiënten in een gevorderd stadium. Dit multicenter consortium bestaat uit het AmsterdamUMC, het Nederlands Kanker Instituut/Antoni van Leeuwenhoek Ziekenhuis, het UMC Utrecht, de MAASTRO kliniek en de Universiteit van Maastricht. Ieder kliniek had zijn eigen expertise. In het Antoni van Leeuwenhoek, en het onderwerp van dit proefschrift, hebben we de gezocht naar de prognostische impact van genetische en transcriptoom biomarkers voor deze patiënten groep. Dit proefschrift maakt gebruik van bulk-RNA-sequencing data. Machine learning technieken werden gebruikt om modellen te genereren die de biologie van hoofd-halskanker omvatten. De prognostische impact van deze modellen, en daarmee de onderliggende tumorbiologie, werden onderzocht.

In **hoofdstuk 2** gaan we dieper in op een bekende prognostische biomarker in HHC, hypoxie. Hypoxische tumoren hebben een slechtere prognose in vergelijking met meer normoxische tumoren. Tumorcellen kunnen hypoxisch worden door chronische en acute mechanismen. Momenteel zijn er vier klinische hypoxie profielen op basis van RNA gen expressie. Het is onduidelijk wat hun onderlinge relaties zijn en hoe ze verband houden met chronische en acute hypoxie. We wilden ook de klinische impact van acute en chronische hypoxie bepalen in verschillende datasets.

De vier klinische hypoxieprofielen vertoonden weinig overlap in de genen die voor elk profiel werden gebruikt. Door drie initiële datasets te combineren, met in totaal 224 HPV-

patiënten, toonden analyses van de vier klinische hypoxieprofielen aan dat de profielen sterk gecorreleerd zijn met elkaar en met chronische hypoxie. De vier klinische profielen werden gemiddeld om een gezamenlijke chronische hypoxiescore te construeren. In een subgroep van 91 patiënten, HPV-positieve en -negatieve tumoren, was chronische hypoxie, hoewel niet significant, geassocieerd met lokaal recidief, en acute hypoxie was significant geassocieerd met lokale controle. Analyse in aanvullende set van 174 HPV-negatieve HHC behandeld met chemoradiotherapie bevestigde deze bevindingen. Univariabele analyses toonden aan dat acute hypoxie significant geassocieerd was met algehele overleving, progressievrije overleving, lokale en locoregionale controle. Chronische hypoxie was significant geassocieerd met progressievrije overleving en vertoonde een algemene overlevingstrend, maar niet met locoregionale controle. Patiënten met hoge scores van zowel acute en chronische hypoxie hadden de slechtste overall survival en lokale controle. Multivariabele analyse, in een beperkte subgroep van 149 patiënten, laat zien dat acute hypoxie significant geassocieerd is met algehele overleving, progressievrije overleving, lokale controle en locoregionale controle. Chronische hypoxie was niet geassocieerd met hypoxie in dit cohort.

In **hoofdstuk 3** richten we ons op epitheel-naar-mesenchymale transitie (EMT), een minder bekende biomarker in HHC. EMT is het proces waarbij epitheelcellen mesenchymale kenmerken krijgen. EMT is in verband gebracht overleving bij veel andere soorten kanker, maar er is weinig wetenschappelijk bewijs in HHC. Om de impact van EMT in HHC te bepalen, hebben we zeven EMT gen expressie profielen geëxtrapoleerd die oorspronkelijk waren gemaakt voor andere kanker soorten. We hebben deze zeven profielen toegepast in een training set van 98 patiënten met gevorderde HPV-negatieve HHC van laryngeale, oro- en hypofaryngeale oorsprong, behandeld met op cisplatinum gebaseerde chemoradiotherapie. Alle zeven EMT-profielen vertoonden associaties met algehele overleving, progressievrije overleving en metastasevrije overleving in de training set. Door de meest epitheliale en meest mesenchymale tumoren uit onze dataset te vergelijken en machine learning te gebruiken, hebben we een nieuw HHC-specifiek EMT-model gemaakt. De impact op de uitkomst van de zeven niet-HHC-specifieke modellen en het HHC-EMT-model werd bevestigd in de validatieset, die bestond uit 78 patiënten met identieke in- en exclusiecriteria. Sterke associaties met EMT werden gezien in de algehele overleving en progressievrije overleving. In de validatie set was EMT echter niet geassocieerd met metastasevrije overleving, maar met locoregionale controle. Om het eerder gesuggereerde verband tussen metastase op afstand en EMT te onderzoeken, analyseerden we aanvullende primaire tumoren van patiënten die zich voor het eerst presenteerden met gemetastaseerde HHC. De primaire

tumor van patiënten met gemetastaseerde ziekte vertoonde hoge waarden in het HHC-EMT model, wat wijst op een associatie met EMT. Multivariabele analyses met TNM-stadium in beide cohorten toonden aan dat het HHC-EMT-model significant geassocieerd is met algehele en progressievrije overleving. Een hoge HHC-EMT was gecorreleerd met tumoren in het T4-stadium en met een sprieterig groeipatroon. Ten slotte suggereerde de literatuur dat het EMT-signaal bij bulk-RNA-sequencing kan worden toegeschreven aan kanker-geassocieerde fibroblasten. We konden deze bevinding niet bevestigen. Met behulp van immunohistochemische kleuring voor vimentin konden we de HHC-EMT-modelscores niet correleren met stroma aanwezigheid. Deze studie toonde aan dat EMT geassocieerd is met de uitkomst bij HHC-patiënten in een gevorderd stadium die worden behandeld met chemoradiotherapie.

Defecten in DNA-crosslink reparatie zijn frequente manifestaties bij HHC en zijn in verband gebracht met overleving bij veel soorten kanker. In **hoofdstuk 4** hebben we modellen gegenereerd om tumoren met defecten in DNA-crosslink reparatie te identificeren en te bepalen wat de klinische impact is. Met behulp van RNA-sequencing data van cellijnen en hun gevoeligheid voor mitomycine C en olaparib, hebben we met gebruik van machine learning technieken modellen gemaakt. Gevoeligheid voor, olaparib en mitomycine C, werd gebruikt als een proxy voor functionele defecten in DNA-crosslink reparatie. Verschillende genselectiemethoden en machine learning technieken werden gecombineerd in modellen om te voorspellen welke cellijnen defecten in DNA-crosslink reparatie hadden. Kruisvalidatie werd gebruikt om de robuustheid van de gebruikte methoden te meten. Vervolgens hebben we de best presterende modellen getest in verschillende datasets. In het Cancer Genome Project (CGP) zijn verschillende HHC-cellijnen onderworpen aan veel verschillende medicatie. In de HCC-cellijnen van de CGP correleerden de modellen goed met mitomycine C IC50-waarden en de modellen voorspelde mutaties in DNA-crosslink-reparatie genen correct. In een aparte dataset voorspelden de modellen de knockdown van DNA-crosslink reparatie genen. Dit bevestigde dat de modellen cellen met defecten in DNA-crosslink reparatie konden identificeren. Vervolgens evalueerden we de impact van de modellen op de klinische uitkomst in twee datasets van HPV-negatieve HHC-patiënten in een gevorderd stadium die werden behandeld met chemoradiotherapie. Tumoren met defecten in DNA-crosslink reparatie, zoals geclassificeerd door de modellen, vertoonden een trend naar een slechtere algehele overleving in beide cohorten. De slechte prognose-associatie was bijzonder sterk in normoxische tumoren en was gekoppeld aan een verhoogd risico op metastasen op afstand. Daaropvolgende analyses toonden aan dat de effecten van hypoxie en DNA-reparatiedefecten elkaar gedeeltelijk overlappen. Dit kan het verschil in uitkomst

van de modellen tussen normoxische en hypoxische tumoren verklaren. De associatie van onze modellen met metastase op afstand bracht ons ertoe om terug te gaan naar *in vitro* studies om hun migratiemogelijkheden te onderzoeken. In scratch- en transwell-assays bleken HHC-cellijnen met defecten in DNA-crosslink-reparatie meer te migreren en waren ook meer invasief. Concluderend zijn de modellen in staat om deficiëntie van DNA-crosslink reparatie te identificeren in HHC. Defecten in DNA-crosslink reparatie, volgens de modellen, zijn geassocieerd met een slechte prognose voor HHC patiënten behandeld met chemoradiotherapie.

Om de verbanden tussen de bovengenoemde biomarkers en in de literatuur gepubliceerde biomarkers te beoordelen, hebben we een validatiestudie verricht die de beschikbare gegevens combineert met de belangrijkste biomarkers (hypoxie (chronisch en acuut), proliferatie, DNA-herstel, stamcelmarkers (SLC3A2 en CD44), EMT, EGFR en immuun (T-cel infiltratiescore, CD8+ T-cellen, CD56dim natural killer en CD8+ versus T regulerende celverhoudingen)) en hun impact op respons op radiotherapie. In **hoofdstuk 5** hebben we de onderlinge relaties tussen deze biomarkers en hun klinische impact op locoregionale controle getest als marker voor respons op radiotherapie. Hiervoor hebben we een RNA sequencing dataset gebruikt van 197 patiënten met gevorderde HPV-negatieve HHC die werden behandeld met op cisplatinum gebaseerde chemoradiotherapie. Er werden sterke correlaties waargenomen tussen de stamcelmarkers en tussen acute hypoxie en T-celinfiltratiescore, CD8+ T-celabundantie en proliferatiescores. Een 10.000-voudige bootstrapping-methode met meerdere klinische variabelen werd gebruikt om robuuste grenswaarden voor elke marker te bepalen. Chronische en acute hypoxie, EGFR-expressie en markers van stamcellen waren het sterkst geassocieerd met locoregionale controle in multivariabele analyses met alleen klinische variabelen. Multivariabele analyse, rekening houdend met zowel andere biologische als klinische markers, toonde een significante rol aan voor chronische hypoxie, EGFR-expressie, CD8+/Treg cel verhoudingen en T-cel infiltratie bij locoregionale controle. In tegenstelling tot locoregionale controle, wordt metastase op afstand voornamelijk beïnvloed door EMT-, EGFR- en CD8+- versus T-reg cel verhoudingen. Deze externe validatie onderstreept de relevantie van biologische factoren bij het bepalen van de uitkomst van chemoradiotherapie bij HHC.

Hoofdstuk 6 sluit dit proefschrift af met een algemene discussie, waarin onze bevindingen worden beschreven in relatie tot recente literatuur. Uitdagingen worden ook beschreven en we sluiten af met toekomstperspectieven voor het biomarkeronderzoek in HHC.



Curriculum Vitae

Authors and affiliations

CV Conchita Vens	Division of Cell Biology; Department of Radiation Oncology, The Netherlands Cancer Institute, Amsterdam, The Netherlands.
CVV Caroline V.M. Verhagen	Division of Cell Biology; Department of Head and Neck Oncology and Surgery, The Netherlands Cancer Institute, Amsterdam, The Netherlands
DV David M. Vossen	Division of Cell Biology; Department of Head and Neck Oncology and Surgery, The Netherlands Cancer Institute, Amsterdam, The Netherlands
EP Emily M. Ploeg	Division of Cell Biology, The Netherlands Cancer Institute, Amsterdam, the Netherlands
FH Frank Hoebbers	Department of Radiation Oncology (MAASTRO), GROW – School for Oncology and Developmental Biology, Maastricht University Medical Centre, The Netherlands
HB Harry Bartelink	Department of Radiation Oncology, The Netherlands Cancer Institute, Amsterdam, the Netherlands.
JS Joyce Sanders	Department of Pathology, The Netherlands Cancer Institute, Amsterdam
MB Michiel W.M. van den Brekel	Department of Head and Neck Oncology and Surgery, The Netherlands Cancer Institute, Amsterdam, The Netherlands; Department of Oral and Maxillofacial Surgery, Amsterdam UMC, Academic Medical Center, Amsterdam, The Netherlands
MH Martijn van der Heijden	Division of Cell Biology; Department of Head and Neck Oncology and Surgery, The Netherlands Cancer Institute, Amsterdam, The Netherlands
MJ Monique C. de Jong	Department of Radiation Oncology, The Netherlands Cancer Institute, Amsterdam, The Netherlands
MV Marcel Verheij	Division of Cell Biology; Department of Radiation Oncology, The Netherlands Cancer Institute, Amsterdam, The Netherlands; Department of Radiation Oncology, Radboud University Medical Center, Nijmegen, The Netherlands
OH Olga Hamming-Vrieze	Department of Radiation Oncology, The Netherlands Cancer Institute, Amsterdam, Netherlands
PE Paul B.M. Essers	Division of Cell Biology; Department of Radiation Oncology, The Netherlands Cancer Institute, Amsterdam, The Netherlands
PL Philippe Lambin	The D-Lab and The M-Lab, Department of Precision Medicine, GROW—School for Oncology and Developmental Biology, Maastricht University, Maastricht, Netherlands
RB Ruud H. Brakenhoff	Otolaryngology/Head & Neck Surgery, Cancer Center Amsterdam, Amsterdam UMC, Vrije Universiteit Amsterdam, Amsterdam, The Netherlands;
RL C. René Leemans	Otolaryngology/Head & Neck Surgery, Cancer Center Amsterdam, Amsterdam UMC, Vrije Universiteit Amsterdam, Amsterdam, The Netherlands;
RR Reinout H. de Roest	Otolaryngology/Head & Neck Surgery, Cancer Center Amsterdam, Amsterdam UMC, Vrije Universiteit Amsterdam, Amsterdam, The Netherlands;

TTCCATCAAGCCCTAGGGCTCCTCTGTGGCTGCTGGGAGTTGTAGTCTGAACGCTTCTATCTTGGGAGAAAGCGCTACGCTCCCCCTACCGAGTCCCGCGGTAACTTTAAAGCACTGACCGCCCCCGCGCGCTGAGAGGGGAGAG

SS| Sebastian Sanduleanu

Department of Radiation Oncology (MAASTRO), GROW – School for Oncology and Developmental Biology, Maastricht University Medical Centre, Maastricht, The Netherlands

SW| Stefan M. Willems

Department of Pathology, University Medical Center Utrecht, Utrecht, The Netherlands.

Author contributions

Chapter 2 Acute hypoxia profile is a stronger prognostic factor than chronic hypoxia in advanced stage head and neck cancer patients.

Study concepts and design	MB, MH, MJ, MV
Resources	CV, FH, MB, MV, RB, RL
Data acquisition	CVV, MH, MJ, RR, SS
Data analysis and interpretation	MH, MJ
Statistical analysis	MH, MJ
Manuscript preparation	MH, MJ
Manuscript editing and review	CV, CVV, FH, MB, MV, RB, RL, RR, SS,

Chapter 3 Epithelial-to-mesenchymal transition is a prognostic marker for patient outcome in advanced stage HNSCC patients treated with chemoradiotherapy.

Study concepts and design	CV, MB, MH, MV
Resources	CV, MB, MV, RB, RL
Data acquisition	CVV, JS, MH, MJ, RR, SW
Data analysis and interpretation	CV, DV, JS, MH, MB, PE, SW
Statistical analysis	CV, DV, MH, PE
Manuscript preparation	CV, MH
Manuscript editing and review	CVV, DV, JS, MB, MV, PE, RB, RL, RR, SW

Chapter 4 Drug Sensitivity Prediction Models Reveal a Link between DNA Repair Defects and Poor Prognosis in HNSCC Predicting DNA Repair Defects in HNSCC.

Study concepts and design	CV, HB, MB, PE
Resources	CV, HB, MB, MV, RB, RL
Data acquisition	CV, CVV, EP, HB, MH, MV, PE, RB, RL, RR
Data analysis and interpretation	CV, CVV, MB, MH, PE
Statistical analysis	CV, CVV, MB, MH, PE
Manuscript preparation	CVV, MH, PE
Manuscript editing and review	CV, HB, MB, MV, RB, RL, RR

Chapter 5 Biological determinants of chemo-radiotherapy response in HPV-negative head and neck cancer: a multicentric external validation.

Study concepts and design	CV, MH, MJ, PE
Resources	CV, FH, MB, MV, PL, RL, RB
Data acquisition	CV, CVV, MH, FH, HB, MB, MV, PL, OH, RL, RR, RB
Data analysis and interpretation	CV, MH, MJ, PE, SS
Statistical analysis	MH, PE
Manuscript preparation	CV, MH, PE
Manuscript editing and review	CVV, FH, HB, MB, MV, PL, OH, RL, RR, RB

TTCCATCAAGCCCTAGGGCTCCTCGTGGCTGCTGGGAGTTGTAGTCTGAACGCTTCTATCTTGGGAGAAAGCGCTACGCTCCCCCTACCGAGTCCCGCGGTAACTTTAAAGCACTGACCGCCCCCGCGCGCTGAGAGGGGCGAG

Funding

This thesis was funded by the Dutch Cancer Society-Alpe d'Huzes (Alp 7072, acronym DESIGN), the EU 7th Framework Program (ARTFORCE - n 257144) and financial support from Brunel and Verwelius.

PhD portfolio

Naam: Martijn van der Heijden

PhD period: November 2015 – October 2023

Promotion committee:

Promotoren: Prof. Dr. M.W.M. van den Brekel, Prof. Dr. R.H. Brakenhoff

Co-promotoren: Dr. C. Vens, Dr. P.B.M. Essers

PhD Training

General Courses

	Date	EC
Introduction to R and statistics, OOA, NKI-AVL	2015	1.3
Bioinformatics Sequence Analysis, AMC graduate school	2016	1.1
Experimental Oncology, OOA, NKI-AVL	2016	6
Course Clin.Epid.4: Systematic Review, AMC graduate school	2017	0.7
BROK (AMC)	2017	1.5
Scientific Writing in English, AMC graduate school	2017	1.5
Scientific Integrity, ACTA	2018	2
Histopathology of Human Tumors, OOA, AMC graduate school	2018	0.6
Boerhaave Nascholing: Chirurgische anatomie van het hoofd-halsgebied, LUMC	2022	1

Conferences/seminars/meetings

ATOS GPR (Provox) Course, NKI-AVL	2015	0.5
Weekly multidisciplinary head and neck surgery and oncology meeting, NKI-AVL	2016-2018	3.5
Weekly scientific meeting biological stress response/Vens/verheij group, NKI-AVL	2016-2018	3.5
Annual OOA Retreat, Renesse	2016-2018	6
Annual DESIGN research meeting	2016-2019	2
Semi-annual national ENT conference, Nieuwegein	2016-2023	8
ESSR/NVRB meeting, Amsterdam	2016	1
Next Generation Sequencing in Molecular Pathology, NKI-AVL	2016	0.25
13 th International Netherlands Cancer Institute Head and Neck Cancer Symposium	2017-04	0.5
7 th Jonge onderzoekersdag, NWHHT, Rotterdam	2017-04	0.5
23 ^e symposium voor verpleegkundigen en paramedici, Amsterdam	2017-06	0.25
37 th ESTRO Conference, Barcelona, Spain	2018-04	1
8 th Jonge onderzoekersdag, NWHHT, Den Haag	2018-05	0.5
6 th World Congress of the International Federation of Head and Neck Oncologic Societies (IFHNOS), Buenos Aires, Argentina	2019-09	1

Poster/Oral Presentations

OOA Annual Progress meeting	2017-2018
Quarterly workdiscussion Biological Stress Response/Vens/Verheij group.	2016-2018
Annual DESIGN Scientific meeting	2016-2019
OOA (Poster)	2016-10
"Identification of prognostic gene mutations in advanced stage HNSCC."	

TTCCATTAAGCCCTAGGGTCTCTGTGGTGTGGGAGTTGTAGTCTGAACGTTCTATCTTGGGAGAGAGGCTCTACGTTCCCTTACCGAGTCCCGGGTAACTTTAAAGACCTGACCGCCCCCGCGCTGAGAGGGGAG

7 th Jonge onderzoekersdag (Oral)	2017-04	
"Epitheel naar mesenchymaal transitie als een prognostische marker voor metastase bij HPV-negatieve hoofd-hals kanker."		
23 ^e symposium voor verpleegkundigen en paramedici. (Oral)	2017-06	
"(genetische) Biomarkers in personalized treatment voor hoofd-halskanker."		
OOA Retreat (Oral)		
"EMT as a prognostic marker for the existence of DM in HNSCC."		
Nederlandse vereniging KNO, najaarsvergadering. Nieuwegein (Oral)	2017-11	
"Epitheliaal Mesenchymaal Transitie als een prognostische marker voor hoofd hals kanker."		
ESTRO (Oral)	2018-04	
"Epithelial to Mesenchymal Transition, a marker for outcome in HPV- negative head and neck cancer."		
UMCU Pathology department scientific meeting	2018-06	
"Epithelial to Mesenchymal Transition, a marker for outcome in HPV-negative head and neck cancer."		
IFHNOS (Oral, Award for best Scientific Presentation)	2018-09	
"Epithelial to Mesenchymal Transition is related to immune checkpoint expression and is a prognostic marker in advanced stage HPV negative head and neck cancer."		
OOA Retreat (Oral, Chairman)	2018-10	
"Prognostic value of hypoxia gene expression profiles in HNSCC."		
Nederlandse vereniging KNO, voorjaarsvergadering. Nieuwegein (Oral)	2019-04	
"DNA-repair defect model is geassocieerd met prognose en de respons op cisplatin van hoofd-hals kanker patiënten."		

Miscellaneous

DESIGN study Protocols: "Patient material record form", "Tissue sectioning", "Tissue Transport", "Blood sample"	2016	
Biobank protocol admission and approval: B17GEN	2017	
Guest Examiner HBO In-Holland	2018	0.25
Invited Peer Review Int J Gen Med	2021	0.25

TTCCCATCAAGCCCTAGGGCTCTCTGTGGTGCTGGGAGTTGTAGTCTGAACGCTTCTATCTTGGGAGAAAGCGCTACGCTCCCGCTACCGAGTCCCGGGTAATCTTAAAGACCTGTACCGCCCCCCCCGCGCTGTAGAGGGGAG

List of publications

- 2022 **Disease outcome and associated factors after definitive platinum based chemoradiotherapy for advanced stage HPV-negative head and neck cancer.** De Roest RH, van der Heijden M, Wesseling FWR, de Ruiter EJ, Heymans MW, Terhaard C, Vergeer MR, Buter J, Devriese LA, de Boer JP, Navran A, Hoebe A, Vens C, van den Brekel M, Brakenhoff RH, Leemans CR, Hoebers F. *Radiother Oncol.* 2022 Oct;175:112-121.
- 2020 **Ovarian cancer-derived copy number alterations signatures are prognostic in chemoradiotherapy-treated head and neck squamous cell carcinoma.** Essers PBM, van der Heijden M, Vossen D, de Roest RH, Leemans CR, Brakenhoff RH, van den Brekel MWM, Bartelink H, Verheij M, Vens C. *Int J Cancer.* 2020 Sep 15;147(6):1732-1739
- 2020 **Epithelial-to-mesenchymal transition is a prognostic marker for patient outcome in advanced stage HNSCC patients treated with chemoradiotherapy.** van der Heijden M, Essers PBM, Verhagen CVM, Willems SM, Sanders J, de Roest RH, Vossen DM, Leemans CR, Verheij M, Brakenhoff RH, van den Brekel MWM, Vens C. *Radiother Oncol.* 2020 Jun;147:186-194.
- 2020 **Computed tomography-derived radiomic signature of head and neck squamous cell carcinoma (peri)tumoral tissue for the prediction of locoregional recurrence and distant metastasis after concurrent chemo-radiotherapy.** Keek S, Sanduleanu S, Wesseling F, de Roest R, van den Brekel M, van der Heijden M, Vens C, Giuseppina C, Licitra L, Scheckenbach K, Vergeer M, Leemans CR, Brakenhoff RH, Nauta I, Cavalieri S, Woodruff HC, Poli T, Leijenaar R, Hoebers F, Lambin P. *PLoS One.* 2020 May 22;15(5):e0232639.
- 2020 **Privacy-preserving distributed learning of radiomics to predict overall survival and HPV status in head and neck cancer.** Bogowicz M, Jochems A, Deist TM, Tanadini-Lang S, Huang SH, Chan B, Waldron JN, Bratman S, O'Sullivan B, Riesterer O, Studer G, Unkelbach J, Barakat S, Brakenhoff RH, Nauta I, Gazzani SE, Calareso G, Scheckenbach K, Hoebers F, Wesseling FWR, Keek S, Sanduleanu S, Leijenaar RTH, Vergeer MR, Leemans CR, Terhaard CHJ, van den Brekel MWM, Hamming-Vrieze O, van der Heijden MA, Elhalawani HM, Fuller CD, Guckenberger M, Lambin P. *Sci Rep.* 2020 Mar 11;10(1):4542.
- 2020 **Biological Determinants of Chemo-Radiotherapy Response in HPV-Negative Head and Neck Cancer: A Multicentric External Validation.** van der Heijden M, Essers PBM, de Jong MC, de Roest RH, Sanduleanu S, Verhagen CVM, Hamming-Vrieze O, Hoebers F, Lambin P, Bartelink H, Leemans CR, Verheij M, Brakenhoff RH, van den Brekel MWM, Vens C. *Front Oncol.* 2020 Jan 10;9:1470.
- 2019 **Drug Sensitivity Prediction Models Reveal a Link between DNA Repair Defects and Poor Prognosis in HNSCC.** Essers PBM, van der Heijden M, Verhagen CVM, Ploeg EM, de Roest RH, Leemans CR, Brakenhoff RH, van den Brekel MWM, Bartelink H, Verheij M, Vens C. *Cancer Res.* 2019 Nov 1;79(21):5597-5611.
- 2019 **Acute Hypoxia Profile is a Stronger Prognostic Factor than Chronic Hypoxia in Advanced Stage Head and Neck Cancer Patients.** van der Heijden M, de Jong MC, Verhagen CVM, de Roest RH, Sanduleanu S, Hoebers F, Leemans CR, Brakenhoff RH, Vens C, Verheij M, van den Brekel MWM. *Cancers (Basel).* 2019 Apr 25;11(4):583.
- 2019 **Genetic Factors Associated with a Poor Outcome in Head and Neck Cancer Patients Receiving Definitive Chemoradiotherapy.** Vossen DM, Verhagen CVM, van der Heijden M, Essers PBM, Bartelink H, Verheij M, Wessels LFA, van den Brekel MWM, Vens C. *Cancers (Basel).* 2019 Mar 29;11(4):445.

TTCCCATCAAGCCCTAGGGCTCCTCTGTGGCTCTGGGAGTTGTAGTCTGAACGTTCTATCTTGGGAGAAAGGCTCTACGCTCCCCCTACCGAGTCCCGGGTAATCTTAAAGCACTGACCGCCCCCGCGCGCTGAGAGGGGAG

- 2018 **Fanconi anemia and homologous recombination gene variants are associated with functional DNA repair defects in vitro and poor outcome in patients with advanced head and neck squamous cell carcinoma.** Verhagen CVM, Vossen DM, Borgmann K, Hageman F, Grénman R, Verwijs-Janssen M, Mout L, Kluin RJC, Nieuwland M, Severson TM, Velds A, Kerkhoven R, O'Connor MJ, **van der Heijden M**, van Velthuysen ML, Verheij M, Wreesmann VB, Wessels LFA, van den Brekel MWM, Vens C. *Oncotarget*. 2018 Apr 6;9(26):18198-18213.
- 2016 **Treatment outcome of supraglottoplasty vs. wait-and-see policy in patients with laryngomalacia.** **van der Heijden M**, Dikkers FG, Halmos GB. *Eur Arch Otorhinolaryngol*. 2016 Jun;273(6):1507-13.
- 2015 **The groningen laryngomalacia classification system--based on systematic review and dynamic airway changes.** **van der Heijden M**, Dikkers FG, Halmos GB. *Pediatr Pulmonol*. 2015 Dec;50(12):1368-73.
- 2015 **The influence of mechanical bowel preparation on long-term survival in patients surgically treated for colorectal cancer.** van't Sant HP, Kamman A, Hop WC, **van der Heijden M**, Lange JF, Contant CM. *Am J Surg*. 2015 Jul;210(1):106-10.
- 2012 **Serotype distribution and antimicrobial resistance of invasive pneumococcal disease strains in the Comunidad Valenciana, Spain, during the winter of 2009-2010: low PCV7 coverage and high levofloxacin resistance.** Gant CM, Rosingh AW, López-Hontangas JL, **van der Heijden M**, González-Morán F, Bijlsma JJ, Canton E; RedMiva (Network of Microbiological Vigilance of Comunidad Valenciana). *Antimicrob Agents Chemother*. 2012 Sep;56(9):4988-9.

Dankwoord

Dit proefschrift is tot stand gekomen met de hulp van velen. Enkelen hiervan wil ik graag in het bijzonder bedanken voor hun bijdrage.

Mijn copromotor, **Dr. C. Vens**, beste **Conchita**, niet voor niets sta jij als eerste in mijn dankwoord. Als geen ander heb je mij begeleid in de wetenschap. Jouw brede kennis en ervaring is nagenoeg onuitputtelijk en ik kijk met een goed gevoel terug op onze besprekingen en belevenissen. Vele uren hebben we samen doorgebracht om over wetenschap, carrière maar ook over privé zaken te praten. Ik vond het dan ook heel erg leuk dat je op ons huwelijk aanwezig was. Jouw deur stond altijd open en ik weet zeker dat ik daar in de toekomst nog gebruik van ga maken. Bedankt voor al je wijze lessen, betrokkenheid en begeleiding.

Mijn promotor van het eerst uur, **Prof. M.W.M. van den Brekel**, beste **Michiel**, als promotor overzag je het grote plaatje. De verbinder van vele projecten. Je bent voor mij een natuurlijke motivator geweest. Als ik er even door heen zat of niet wist wat de volgende stap moest zijn dan ging ik naar jou. Zonder dat jij het zelf door had, denk ik, kwam ik altijd met veel energie van onze besprekingen. Klaar om er weer mee aan de slag te gaan. Een gave! Daarnaast heb je me ook veel geholpen met mijn verdere carrière. Ik ben trots dat jij mijn promotor bent en hoewel de toekomst nog ongewis is, hoop ik dat we nog veel samen kunnen werken.

Copromotor, **Dr. P.B.M. Essers**, beste **Paul**, met jouw komst naar onze groep braken onze hoogtij dagen aan. Vreemde festivals, fietsen en vooral heel erg veel humor waren schering en inslag bij ons. Met jouw ervaring en relaxte houding heb je mij enorm veel geholpen. Helaas heb je de wetenschap gedag gezegd, maar je kan als copromotor niet ontbreken vandaag!

Promotor **Prof. R.H. Brakenhoff**, beste **Ruud**, hoewel je al vanaf het begin betrokken was bij al mijn projecten, ben je later pas officieel mijn promotor geworden. Als drijvende kracht achter het DESIGN consortium dacht je veel mee met organisatorische en logistieke vraagstukken. Tevens kwam je kritische blik altijd van pas. Ik vind het een eer dat je mijn promotor wil zijn.

Geachte leden van de promotie commissie, **prof. dr. L.E. Smeele**, **prof. dr. M.J. van de Vijver**, **dr. M.F. Bijlsma**, **prof. dr. C.L. Zuur**, **prof. dr. J. Bussink** en **prof. dr. R.P. Takes**, bedankt voor het beoordelen van mijn proefschrift.

Lieve **Hanneke**, **Dr. van Eden**, wat ben ik blij dat jij vandaag mijn paranimf bent. Onze vriendschap is onlosmakelijk verbonden aan onze PhD's. Wat begon bij van Mechelen,

groeide via de skireizen, Stelvio for life en de RVV uit tot een echte vriendschap. Hoewel de toenemende reisafstand en agenda management niet altijd meewerken, weet ik zeker dat we nog nieuwe avonturen aan bovenstaande rijtje gaan toevoegen.

Beste **Arne, Dr. de Niet**, 'Arnold', mijn paranimf, ceremoniemeester en vooral vriend. Ondanks de grote fysieke afstand heb ik juist het gevoel dat we de laatste jaren naar elkaar toe zijn gegroeid. We doen veel ons best om elkaar te zien, al lukt dat te weinig, en we levelen op veel vlakken. Ik heb veel bewondering voor je sociale skills, je begrip en rust. Je moest gewoon naast me staan vandaag.

Collega's van de **Vens/Verheij groep/B5/H3**, zonder jullie was deze dag er niet gekomen. Beste **David**, het was een genot om jaren naast je te hebben mogen zitten. Je hebt mij een andere kijk op de wereld gegeven die vaak erg verfrissend is en ook een goede voedingsbodem is gebleken voor onze humor. Ik ben dankbaar dat ik jouw paranimf mocht zijn en leuk dat je op mijn huwelijksfeest was. Ik ben blij dat je ook je carrière pad hebt gevonden. Lieve **Shuraila**, de verhuizing naar B5 heeft ons veel goed gedaan. We hebben elkaar daar pas echt leren kennen. Bedankt voor al je humor, diepe gesprekken en steun. Ik wens je alle succes toe. Ik hoop dat onze wegen elkaar nog eens kruisen. **Prof. Dr. M. Verheij**, bedankt voor de ruimte, begeleiding en inzichten tijdens de vele besprekingen. **Manon**, bedankt voor alle jaren gezelligheid. **Ben**, helaas kun je er vandaag niet bij zijn maar ik wil je toch bedanken voor de prettige samenwerking en je nuchterheid. **Hans**, hoewel we na de verhuizing na B5 elkaar minder zagen, denk ik op de racefiets nog vaak aan onze lunches in het AvL. Bedankt voor al je steun en wijsheden over het fietsen. Beste **Thea**, bedankt voor al je vrolijkheid en hulp door de jaren heen. **Caroline**, hoewel wij technisch gezien niet samen hebben gewerkt ben je wel een van de grondleggers van dit project. We hebben toch veel over de data gesproken en ik ben blij dat je ook je PhD hebt afgerond! We gaan elkaar in de KNO wereld zeker nog tegen komen.

Rebecca, wanneer gaan we nu eens eindelijk bowlen dan? **Marije, Ann-Jean, Jos, Joris, Viola, Minke, Ravi, Sarah, Charlotte, Joost, Willem, Marnix, Tessa, Fleur, Merijn, Max, Elise, Rosa, Esther, Sophie** en al mijn andere OIO collega's die ik ongetwijfeld nog vergeet. Bedankt voor alle borrels in van Mechelen, Ebeling, feestje, skireisjes, OOA retreats en weet ik wat nog meer. Zonder al deze afleiding was het half niet zo leuk geweest.

Mijn PhD is onderdeel geweest van het **DESIGN-consortium**. Naast Conchita, Michiel en Ruud zijn er nog vele anderen betrokken geweest bij de totstandkoming van dit project. **Prof. Dr. S.M. Willems**, beste **Stefan**, bedankt voor het wegwijs maken in de wondere wereld van de pathologie. Ik kijk met een goed gevoel terug op de vele uren die we samen

TTCCATCAAGCCCTAGGGTCTCTGTGGTGCTGGGAGTTGTAGTCTGAACGTTCTATCTTGGGAGAAGGCTCAAGTCCCCCTACCGAGTCCCGGGTAATCTTAAAGACCTGTACCGCCCCCGCGGCTGTAGAGGGGAG

achter de microscoop hebben doorgebracht. Verder wil ik **Prof. Dr. C.R. Leemans, Prof Dr. C.H.J. Terhaard, Dr. F. Hoebers, Prof. Dr. P. Lambin, Dr. A. Berlanga** bedanken voor de mogelijkheden die mij dankzij dit project geboden zijn, het lezen van mijn manuscripten en voor alle wetenschappelijk input tijdens de jaarlijkse bijeenkomsten. **Reinout**, ik beschouw jouw als mijn onderzoeksmaatje, maar dan van het VUMC. Nu zijn we inmiddels beiden in opleiding tot KNO-arts en ik hoop dat we jouw promotie binnenkort ook kunnen vieren.

Beste **Sten, Linda en Annegien** en andere collega's van de **pathologie**. Bedankt voor alle tijd en energie die jullie in dit project hebben gestoken. De vele uren die jullie besteed hebben aan het prepareren van de DNA/RNA zijn cruciaal geweest. Dit kan niet onderschat worden en mede dankzij jullie is dit tot een goed einde gebracht.

Ook de **Genomics Core facility, Ron, Roel en Arno**, wil ik bedanken voor het sequencen van het RNA/DNA en het meedenken over vragen hieromtrent. Jullie bijdrage aan onze manuscripten is cruciaal geweest.

Ook wil ik alle hoofd-hals chirurgen en fellows bedanken. **Michiel, Lotje, Peter, Richard, Pim, Ludi, Luc, Baris, Marin, Lilly-Ann, Xander en Frank**. Zonder het verzamelen van al het materiaal was dit onderzoek niet mogelijk geweest. Daarnaast hebben jullie mij begeleid bij de eerste stappen in de hoofd-hals chirurgie op de polikliniek en is mijn passie voor dit vak alleen maar toegenomen. In dat kader wil ik ook **Olga, Arash, Abraham, Charlotte en Bas** bedanken. Ik kijk dan er dan ook naar uit om een keer terug te keren.

Uiteraard ben ik ook alle **patiënten** dankbaar die toestemming hebben gegeven voor het gebruik van het materiaal.

Naast al deze bovenstaande personen zijn er veel mensen die geen idee hebben van wat mijn promotie inhoudt, maar niet onbelangrijk geweest zijn voor het voltooien van mijn promotie.

Collega AIOS uit het EMC, **Tessa, Stephanie, Eveline, Carlijn, Florence, Arta, Edela, Christina, Jacob, Simone, Rens, Yassine, Danique, Maarten, Ajay, Victor**, en **Anke** en arts-onderzoekers **Kira, Boyd, Tjeerd, Stefanie, Anouk, Diako en Hanneke**. Uiteraard kan ik de oud-assistenten **Stephanie, Anne, Emilie** en **Roeland** ook niet vergeten. Bedankt voor alle nodige afleidingen in de vorm van borrels, jaardiners, skiweekenden en assistentenweekenden. Dat ik er nog veel mee mag maken. **Rens**, squashen, boksen, voetbal, maar altijd onder het genot van een biertje. Zeker zo therapeutisch. Wanneer gaan we voor een squash rematch? **Tes** en **Flo**, ik kijk uit naar onze momenten in de Garden wanneer het eens niet over onze onderzoeken gaat en dat we ook jullie PhD's hebben kunnen vieren. Dat zal vast niet lang duren.

KNO-artsen uit het EMC, **Bas Pullens, Floris van Zijl, Frank Datema, Jose Hardillo, Hylke van de Toom, Jet de Gier, Marc van de Schroef, Dominiek Monserez, Robert-Jan Pauw, Simone Bernard, Stijn Keereweert, Paul Nagtegaal, Laura Veder, Hamid Dixon, Professor Rob Baatenburg de Jong, Professor Bernd Kremer en Mick Metselaar.** KNO-artsen uit het Maasstad Ziekenhuis, **Paul ten Koppel, Anton Stoop, Edward Grooters, Ivar Tabink en Nienke Grootenhuis.** KNO-artsen uit het Amphia Ziekenhuis, **Gijs van Wermeskerken, Antoon van de Rijt, Jasper Companjen, Eric Janssen, Geert Küppers, Ferdinand Timmer, Hayk Baazil en Bernard Vonck.** Bedankt voor de kans die jullie mij hebben gegeven om KNO-arts te worden. Hoewel ik pas in de laatste fase van mijn proefschrift bij jullie kwam, ben ik blij dat jullie mij de ruimte en tijd hebben gegeven om deze af te maken. **Paul,** ik hoop dat we nog vele tochten op de racefiets gaan maken de komende jaren.

Heren van **Cojones, Arthur, Fokko** of **Dr. Huizing, Gerben, Joost, Sander, Sjoerd** en **Thomas.** Studentenmaatjes van het eerste uur. Talloze uren hebben we samen doorgebracht tijdens feestjes, gala's, vakanties, sporten, huwelijken en weekendjes weg. Dat er nog velen mogen volgen.

MHGDOP'ers. Van Groningen, tot Thailand en Turkije, er zijn tal van momenten geweest die zorgde voor de nodige afleiding en ontspanning. Hoewel er nog vele huwelijken op de wachtlijst staan, is vandaag weer een goede gelegenheid om bij elkaar te komen en te zingen. **Thomas,** een van mijn beste maten en niet voor niets mijn getuige. Jij hebt relativeren tot kunst verheven en bent altijd in voor een borrel. Hoewel je nu zelf ook in Rotterdam woont, is het er nog niet vaak van gekomen. Nu mijn promotie afgerond is, komen er hopelijk weer veel momenten om samen te borrelen en lachen. Ik kijk er naar uit. **Scott,** de filosoof van het stel en onze BABS. We hebben wat filosofische avonden aan de bar gezeten en ik hoop dat er nog vele zullen volgen. Ik heb bewondering voor de keuzes die je maakt in het leven en ben blij dat ik je mijn vriend mag noemen. **Tam,** vreemd hoe je tegelijk de meest volwassen bent en soms ook juist niet. Je trekt je eigen plan en je houdt er ook aan. Ik bewonder je doorzettingsvermogen en eigenzinnigheid en je staat altijd voor iedereen klaar. **Joost,** de afgelopen jaren hebben we elkaar steeds vaker opgezocht. Gymdag was helaas niet blijvend, maar ik hoop dat het luieren in de avondzon van Scheveningen dat wel is. We hebben veel aan je professionele luisterende oor gehad. Blijf vooral die energieke clown die je bent! **Maikel,** de stille verbinder van dit stel. Mooi dat je er altijd bij bent en een beetje kleur toevoegt aan de groep. Ik hoop dat je wensen uitkomen. **Louren, Dr. Goedhart,** kwamen alle 5-jaren plannen maar zo uit als die van jou! Gewoon bewonderenswaardig. Je recente promotie was een voorbeeld en motiverend om de laatste loodjes af te maken. **Jeroen,** mooi om te zien hoe jij je eigen weg kiest en daarmee fantastische avonturen

TTCCATTAAGCCCTAGGGTCTCTGTGGTGCTGGGAGTTGTAGTCTGAACGTTCTATCTTGGGAGAAGGCGCTACGTCCTCCCTACCGAGTCCCGGGTAATCTTAAAGACCTGACCGCCCCCGCGCGCTGAGAGGGGAG

beleeft. Soms jaloersmakend. **Marnix**, mede B&Ber, het wordt dan eindelijk tijd om je zelf gebrouwen biertjes maar eens te komen proeven. **Bart**, stiekem ben ik altijd bang als je mee fietst, maar ik geniet er wel van. **Daan**, oud-huisgenootje! Mooi dat je nu je eigen praktijk hebt. De patiënten mogen maar blij zijn met jou als huisarts. **Paul**, mede B&Ber, ik kan altijd genieten van de vieze verhalen die je in je vak meemaakt.

Lieve schoonfamilie. **Jeroen en Tilly**, bedankt voor het tweede thuis dat jullie me hebben gegeven en alle gevraagde en ongevraagde hulp. Ik moet jullie ook bedanken dat jullie mij op de ski's hebben geduwd en voor alle volgende vakanties. Ik weet dat ik altijd bij jullie terecht kan. **Floor en Jur**, hoewel de laatste maanden hectisch waren en soms wat etentjes hebben moeten wijken, kijk ik wel uit naar onze volgende dinerdate. Ik waardeer heel erg hoe jullie mij/onze rol zien ten opzichte van jullie kroost. **Sam en Nore**, jullie zijn heerlijke kids. **Kars en Charlotte**, helemaal aan de andere kant van de wereld maakt contact onderhouden wat lastig, maar in gedachte er altijd bij. Ik hoop jullie en jullie fantastische kinderen, **Pepijn, Mees en Mette** snel weer te zien.

Lieve **Chris**, als Brit in Nederland zijn de eerste jaren niet gemakkelijk geweest, maar je mag trots zijn waar je nu staat. Ik vind het heerlijk hoe jij ons Nederlanders een spiegel voor kan houden, hoe pijnlijk soms ook, en dat je ons verrijkt met je eigen cultuur. Precies wat ik af en toe nodig had. Sláinte!

Lief zusje, **Marleen**, hoewel we wel eens slaande ruzie hadden als kinderen zijn we daarna juist steeds meer naar elkaar gegroeid. Je speech bij ons huwelijk geeft aan dat je me door en door kent, al vereffenen we deze rekening op een later tijdstip! Daarvoor is deze ruimte te beperkt. De liefde is geheel wederzijds.

Lieve **pap en mam**, ik kan jullie niet genoeg bedanken voor de onbezorgde en liefdevolle jeugd en het weerstaan van al mijn grillen door de jaren heen. Hoewel de vraag hoe het met mijn promotie ging af en toe pijnlijk was, heb ik altijd jullie betrokkenheid gevoeld. Jullie hebben mij geleerd hard te werken, door te zetten en mijn verantwoordelijkheid te nemen. Zonder deze eigenschappen wat dit boekje niet tot stand gekomen. Bedankt en ik kijk uit naar ons volgende familieweekend.

En als laatste natuurlijk mijn liefste **Cath**. Je vindt van jezelf dat je geen credits verdient voor alles wat hiervoor opgetekend is, maar dat is natuurlijk niet zo. Integendeel. Je verdient alles. Je bent te streng voor jezelf en ik kan me geen betere partner in het leven wensen. We zijn door diepe dalen en vele hoogtepunten gegaan maar je stond er altijd als ik je nodig had. Ik kijk uit naar alle avonturen die we samen nog gaan beleven. Ik hou van je!

Curriculum Vitae

Martijn van der Heijden was born on the 17th of September in 1988 in Nijmegen, the Netherlands where he grew-up in a loving family together with his younger sister. After graduating from the Stedelijk Gymnasium in Nijmegen, he started studying Medicine at the Rijksuniversiteit Groningen in 2006. During his studies he was attracted to Ear-, Nose- and Throat (ENT) surgery and Head and Neck Oncology in particular. His interest in science started during an internship in Microbiology in Valencia, Spain, where he also learned Spanish. After finishing his study in 2013, he started working as a resident not in training (ANIOS) at the surgery department of the IKAZIA hospital and later briefly at the intensive care unit of the Slotervaart Hospital.



In 2015, he started his PhD candidacy at the Head and Neck Surgery and Oncology department of the Netherlands Cancer Institute – Antoni van Leeuwenhoek (NKI-AvL), Amsterdam, under the supervision of Prof. Dr. M.W.M. van den Brekel, Prof. Dr. R.H. Brakenhoff, Dr. C. Vens and Dr. P.B.M. Essers. His focus was to identify genetic biomarkers for chemoradiotherapy response in advanced stage head and neck cancer patients as part of the DESIGN study, a multi-center study in collaboration with the Free University, UMCU and Maastricht.

After over 3 years of full-time research, in February 2019, he started as a resident not in training at the department of ENT surgery and head and neck oncology at the Erasmus MC, Rotterdam, the Netherlands and was accepted into training later in 2019, under the supervision of Prof. Dr. R.J. Baatenburg de Jong and Dr. R.M. Metselaar and later Prof. Dr. B. Kremer. He developed a special interest in endoscopic sinus/skull base surgery and benign and malignant head and neck surgery. During his training he worked at the Maastricht Hospital in Rotterdam and Amphia Hospital in Breda and he will complete his training in 2024.

He is married to Cathelijne Mud and they are living together in Rotterdam.

