Received: 25 June 2021 | Revised: 18 August 2021 | Accepted: 31 August 2021

DOI: 10.1002/ijc.33842

CANCER GENETICS AND EPIGENETICS

IJC International Journal of Cancer

# Tumor DNA-methylome derived epigenetic fingerprint identifies HPV-negative head and neck patients at risk for locoregional recurrence after postoperative radiochemotherapy

Bouchra Tawk[1,2,3,4,5] | Ute Wirkner[1,2,3,4,5] | Christian Schwager[1,2,3,4,5] | Katrin Rein[1,2,3,4,5] | Karim Zaoui[6] | Philippe A. Federspil[6] | Sebastian Adeberg[1,2,7,8] | Annett Linge[3,9,10,11,12,13,14] | Ute Ganswindt[15,16] | Julia Hess[17,18,19] | Kristian Unger[17,18,19] | Ingeborg Tinhofer[20,21] | Volker Budach[20,21] | Fabian Lohaus[3,9,10,11,12,13,14] | Mechthild Krause[3,9,10,11,12,13,14] | Maja Guberina[22,23] | Martin Stuschke[10,23] | Panagiotis Balermpas[24,25] | Claus Rödel[24,26] | Anca L. Grosu[27,28,29] | Henning Schäfer[27,28] | Daniel Zips[30,31] | Stephanie E. Combs[15,32] | Steffi Pigorsch[15,32] | Horst Zitzelsberger[17,18,19] | Philipp Baumeister[18,33] | Thomas Kirchner[15,34] | Melanie Bewerunge-Hudler[35] | Wilko Weichert[15,36] | Jochen Hess[6,37] | Esther Herpel[1,2,38,39] | Claus Belka[15,17,18] | Michael Baumann[1,2,9,10,11] | Jürgen Debus[1,2,3,4,5] | Amir Abdollahi[1,2,3,4,5] | For the DKTK-ROG

[1]German Cancer Research Center (DKFZ), Heidelberg, Germany

[2]German Cancer Consortium (DKTK), Core Center Heidelberg, Heidelberg, Germany

[3]Clinical Cooperation Unit Translational Radiation Oncology, National Center for Tumor Diseases (NCT), Heidelberg University Hospital (UKHD), Heidelberg, Germany

[4]Division of Molecular and Translational Radiation Oncology, Department of Radiation Oncology, Heidelberg Faculty of Medicine (MFHD) and Heidelberg University Hospital (UKHD), Heidelberg Ion-Beam Therapy Center (HIT), Heidelberg, Germany

[5]Heidelberg Institute of Radiation Oncology (HIRO), National Center for Radiation Oncology (NCRO), Heidelberg University, Heidelberg, Germany

[6]Department of Otorhinolaryngology, Head and Neck Surgery, Heidelberg University Hospital, Heidelberg, Germany

[7]Department of Radiation Oncology, University Hospital Heidelberg, Heidelberg, Germany

[8]Heidelberg Institute of Radiation Oncology (HIRO), Heidelberg, Germany

[9]German Cancer Consortium (DKTK), Partner Site Dresden, Dresden, Germany

[10]Department of Radiotherapy and Radiation Oncology, Faculty of Medicine and University Hospital Carl Gustav Carus, Technische Universität Dresden, Dresden, Germany

**Abbreviations:** AUC, area under the curve; CNS, central nervous system; CSC, cancer stem like cells; DKFZ, German Cancer Research Center; DKTK-ROG, German Cancer Consortium-Radiation Oncology Group; DM, distant metastasis; DMFS, distant metastasis free survival; DMP, differentially methylated probes; DNA, deoxyribonucleic acid; ECE, extracapsular extension; EORTC, European Organization for Research and Treatment of Cancer; FFPE, formalin-fixed paraffin-embedded; FWER, family-wise error rate; Fx, fraction; Gy, gray; HICR, HPV-independent classifier of disease recurrence; HNSCC, head and neck squamous cell carcinoma; HPV, human papillomavirus; HR, hazard ratio; IHC, immunohistochemistry; KM, Kaplan-Meier; LASSO, least absolute shrinkage and selection operator; LC, local control; LET, linear energy transfer; LMU, Ludwig-Maximilians-University; LMU-KKG, LMU of Munich, Clinical Cooperation Group "Personalized Radiotherapy in Head and Neck Cancer"; LN, lymph nodes; LR, local recurrence; miRNA, microRNA; M-score, methylation score; OOB, out of bad error; OS, overall survival; PCR, polymerase chain reaction; PD, progressive disease; pLRT, *P*-log-rank-test; PORT-C, postoperative radiochemotherapy; R, resection margins; RCHT, radiochemotherapy; RF, random forest; RNA, ribonucleic acid; ROC, receiver operated curve; RT, radiotherapy; RTOG, Radiation Therapy Oncology Group; TILs, tumor-infiltrating lymphocytes; TSS, transcription start site; UICC, Union for International Cancer Control; WHO, World Health Organization

[11]OncoRay – National Center for Radiation Research in Oncology, Faculty of Medicine and University Hospital Carl Gustav Carus, Technische Universität Dresden, Dresden, Germany

[12]National Center for Tumor Diseases (NCT), Partner Site Dresden, Dresden, Germany

[13]Faculty of Medicine and University Hospital Carl Gustav Carus, Technische Universität Dresden, Dresden, Germany

[14]Helmholtz Association/Helmholtz-Zentrum Dresden—Rossendorf (HZDR), Dresden, Germany

[15]German Cancer Consortium (DKTK), Partner Site Munich, Munich, Germany

[16]Department of Radiation Oncology, Medical University of Innsbruck, Innsbruck, Austria

[17]Department of Radiation Oncology, University Hospital Ludwig-Maximilians-University of Munich, Munich, Germany

[18]Research Unit Radiation Cytogenetics, Helmholtz Zentrum München, German Research Center for Environmental Health GmbH, Neuherberg, Germany

[19]Clinical Cooperation Group "Personalized Radiotherapy in Head and Neck Cancer", Helmholtz Zentrum München, German Research Center for Environmental Health GmbH, Neuherberg, Germany

[20]German Cancer Consortium (DKTK), Partner Site Berlin, Berlin, Germany

[21]Department of Radiooncology and Radiotherapy, Charité University Hospital, Berlin, Germany

[22]German Cancer Consortium (DKTK), Partner Site Essen, Essen, Germany

[23]Department of Radiotherapy, Medical Faculty, University of Duisburg-Essen, Essen, Germany

[24]German Cancer Consortium (DKTK), Partner Site Frankfurt, Frankfurt, Germany

[25]Department of Radiation Oncology, University Hospital Zurich, Zurich, Switzerland

[26]Department of Radiotherapy and Oncology, Goethe-University Frankfurt, Frankfurt, Germany

[27]German Cancer Consortium (DKTK), Partner Site Freiburg, Freiburg, Germany

[28]Department of Radiation Oncology, Clinical Study Section, University of Freiburg, Freiburg, Germany

[29]Department of Radiation Oncology, University of Freiburg, Freiburg, Germany

[30]German Cancer Consortium (DKTK), Partner Site Tuebingen, Tuebingen, Germany

[31]Department of Radiation Oncology, Faculty of Medicine and University Hospital Tübingen, Eberhard Karls Universität Tübingen, Tübingen, Germany

[32]Department of Radiation Oncology, Technische Universität München, Munich, Germany

[33]Department of Otorhinolaryngology, Head and Neck Surgery, University Hospital, Ludwig-Maximilians-University of Munich, Munich, Germany

[34]Institute of Pathology, Faculty of Medicine, Ludwig-Maximilians-University of Munich, Munich, Germany

[35]Genomics and Proteomics Core Facility, German Cancer Research Center, Heidelberg, Germany

[36]Institute of Pathology, Technische Universität München, Munich, Germany

[37]Department of Molecular Mechanisms of Head and Neck Tumors, German Cancer Research Center, Heidelberg, Germany

[38]Tissue Bank, National Center for Tumor Diseases (NCT), Heidelberg, Germany

[39]Department of Applied Tumor Biology, Institute of Pathology, Heidelberg University Hospital, Heidelberg, Germany

**Correspondence**
Amir Abdollahi, CCU Translational Radiation Oncology (E210), German Cancer Consortium (DKTK), Core Center Heidelberg National Center for Tumor Diseases (NCT), German Cancer Research Center (DKFZ) and Heidelberg University Hospital (UKHD), Im Neuenheimer Feld 460, 69120 Heidelberg, Germany.
Email: a.amir@dkfz.de

**Funding information**
Intramural Funds of the National Center for Tumor Diseases (NCT) Heidelberg Radiation Oncology Program; Joint Funding Grant within the German Cancer Consortium (DKTK), which is funded as one of the National German Health Centers by the Federal German Ministry of Education (training cohort), the Helmholtz Cross-Program Initiative Personalized Medicine (iMed) project on "Multi-Scale Integrative Biology of HNSCC" (validation cohorts)

## Abstract

Biomarkers with relevance for loco-regional therapy are needed in human papillomavirus negative aka HPV(−) head and neck squamous cell carcinoma (HNSCC). Based on the premise that DNA methylation pattern is highly conserved, we sought to develop a reliable and robust methylome-based classifier identifying HPV(−) HNSCC patients at risk for loco-regional recurrence (LR) and all-event progression after post-operative radiochemotherapy (PORT-C). The training cohort consisted of HPV-DNA negative HNSCC patients (n = 128) homogeneously treated with PORT-C in frame of the German Cancer Consortium—Radiation Oncology Group (DKTK-ROG) multicenter biomarker trial. DNA Methylation analysis was performed using Illumina 450 K and 850 K-EPIC microarray technology. The performance of the classifier was integrated with a series of biomarkers studied in the training set namely hypoxia-, 5-microRNA (5-miR), stem-cell gene-expression signatures and immunohistochemistry (IHC)-based immunological characterization of tumors (CD3/CD8/PD-L1/PD1). Validation occurred in an independent cohort of HPV(−) HNSCC patients, pooled from two German centers (n = 125). We identified a 38-methylation probe-based HPV(−)

Independent Classifier of disease Recurrence (HICR) with high prognostic value for LR, distant metastasis and overall survival ($P < 10^{-9}$). HICR remained significant after multivariate analysis adjusting for anatomical site, lymph node extracapsular extension (ECE) and size (T-stage). HICR high-risk tumors were enriched for younger patients with hypoxic tumors (15-gene signature) and elevated 5-miR score. After adjustment for hypoxia and 5-miR covariates, HICR maintained predicting all endpoints. HICR provides a novel mean for assessing the risk of LR in HPV(−) HNSCC patients treated with PORT-C and opens a new opportunity for biomarker-assisted stratification and therapy adaptation in these patients.

**KEYWORDS**

disease recurrence, DNA methylation, head and neck cancers, radiotherapy, stratification

**What's new?**

New biomarkers are urgently needed for the stratification of patients with HPV-negative head and neck squamous cell carcinoma (HNSCC). In this study, the authors used DNA-methylation patterns to develop and validate a biomarker system called the "DNA methylation-based, HPV-Independent, Classifier of disease Recurrence" (HICR). The signature was able to identify those tumors that increased the risk of loco-regional recurrence and disease progression, and decreased survival. This held true regardless of treatment, microRNA, or hypoxia status. HICR may therefore provide a valuable prognostic biomarker panel to guide treatment in HNSCC.

## 1 | INTRODUCTION

Beyond p16-IHC as a surrogate for human papillomavirus (HPV)-status in oropharyngeal tumors, biomarkers are direly needed for the stratification of patients with HNSCC. Most biomarker approaches fail to consider dose-distribution constraints of loco-regional radiotherapy (RT) and are thus unable to discriminate between local events vs systemic progression and potential interaction with local therapy. The RadPlanBio platform was established by the German cancer consortium (DKTK) to integrate tumor biology with clinical information, imaging data and RT structure sets thus allowing a faithful assessment of in field-recurrences.[1] Successful implementation in multicenter trials allowed for discovery of biomarkers associated with LR in patients with HNSCC treated with postoperative radiotherapy and cisplatin based chemotherapy (PORT-C).[2-9]

In the retrospective multicenter postoperative cohort of the DKTK-ROG (PORT-C), direct assessment of viral load (HPV-DNA) through PCR-based array was a prominent biomarker of local control (LC). Patients with HPV-DNA positive tumors had a 89.6% LC rate 2 years post-RT.[9] Hypoxia is a known mechanism for radioresistance in HNSCC and gene expression signatures have attempted to assess tumoral hypoxia in bulk tissue.[10] In the same cohort, tumors with high expression of hypoxia-associated gene signatures showed worsened outcomes.[11] Additionally, high expression of genes associated with cancer stem like cell (CSC) phenotype (eg, CD44 protein and *SLC3A2* gene) was negatively prognostic of LC and overall survival (OS).[4,11,12] By contrast, the presence of CD8-positive tumor-infiltrating lymphocytes (TILs) and PD1-positive cells were associated with improved LC, DM free survival (DMFS) and OS, independently of HPV status.[5,8] Finally, expression of a 5-miRNA signature was prognostic in HPV-DNA-negative tumors in this cohort.[2]

DNA methylation has revolutionized neuropathological diagnostics.[13] Methylation-based biomarkers are attractive, given that DNA biomaterial is robustly assayed from FFPE tissue, for example, in contrast to RNA that is often subject to degradation. Additionally, epigenetic modulation of DNA methylation pattern is well-conserved and subject to long-term reprograming whereas the transcriptional response can be rapidly perturbed by transient pathophysiological parameters, for example, clamping time of vessel during the surgery, and so on.

In our study, we aimed to show that tumor DNA methylation pattern can also predict radiotherapy-specific outcomes in HPV negative HNSCC. We developed and validated an HPV-Independent Classifier of disease Recurrence (HICR) identifying patients at risk for LR after PORT-C. Prospective validation of HICR is currently ongoing in PORT-C arm of the multicenter HNprädBio study of the DKTK-ROG (www.clinicaltrials.gov, NCT02059668).

## 2 | MATERIALS AND METHODS

### 2.1 | Patient cohorts

Our study includes three patient cohorts with retrospectively collected clinical data. The training cohort, the PORT-C cohort of the

DKTK-ROG, is a retrospective multicenter cohort of patients treated with postoperative RT and concurrent cisplatin-based chemotherapy. Inclusion criteria, data collection and handling were previously described in detail.[9] Briefly, it consists of 221 patients, with histologically proven squamous cell carcinoma of the oropharynx, oral cavity and hypopharynx, treated between 2004 and 2012, originating from eight DKTK partner sites in Germany.[9] The minimum of one of following criteria mandated inclusion in the study: pT4 stage, more than three positive lymph nodes (LN), positive microscopic resection margins or extracapsular extension (ECE). All patients received RT to the surgically resected tumor bed and regional LNs with concurrent weekly cisplatin per standard protocols.[9] Tumor beds and ECE-positive regional LN-levels received 50 Gray (Gy) in 2 Gy fractions (fx) with a 16 Gy boost in 2 Gy fx. Electively irradiated LN-levels received 50 Gy/2 Gy fx. Genomic DNA from 128 HPV-DNA negative tumors was profiled for our study.

The validation cohort consists of two retrospective monocentric cohorts from the University Hospital Dresden (n = 152) and the Ludwig-Maximilians-University (LMU) of Munich (n = 243). The Dresden retrospective cohort (DD) was treated between 1999 and 2006.[14] Inclusion criteria were proven HNSCC of the oral cavity, oropharynx or hypopharynx and treatment with standard postoperative RT (see above) with or without concurrent chemotherapy. Seventy-two HPV DNA negative tumors were profiled for methylation. No patient from this cohort was included in the retrospective PORT-C cohort of the DKTK-ROG. The Munich retrospective cohort (LMU-KKG, LMU of Munich, Clinical Cooperation Group "Personalized Radiotherapy in Head and Neck Cancer") consists of 243 patients treated with postoperative RT with or without chemotherapy between 2008 and 2013 at the LMU Department of Radiation Oncology.[15] Inclusion criteria were UICC stage III HNSCC with ± surgical or close margins (<5 mm). R1-resected tumor beds and regional ECE + LNs received total doses of 66 Gy/56 Gy in 2 Gy fx. Uninvolved regional LNs received electively 50 Gy/2 Gy as described previously.[2] Methylation profiling occurred for 53 primary HPV-DNA negative tumors.

## 2.2 | Clinical endpoints

Clinical outcomes (time to death, LR and DM) were defined as time (in months) from first day of RT treatment to time of first event. The primary clinical endpoint was progression/progressive disease (PD). Time to PD was defined as time to LR or DM, whichever occurred first. Censoring occurred at time of death or last follow-up. Review of RT treatment plans and clinical imaging was performed centrally for the training cohort and locally in Dresden and Munich for the validation cohorts to confirm LRs within previously irradiated volumes.[9,15]

## 2.3 | HPV status determination

For the retrospective PORT-C cohort of the DKTK-ROG and the monocentric Dresden cohort, HPV DNA status of samples was determined at the DKTK partner site Dresden as previously described.[9,14] Genomic DNA extraction was performed on 5-µm FFPE sections using QIAmp DNA FFPE tissue kit (Qiagen, Venlo, NL). The LCD-Array HPV 3.5 Kit (CHIPRON GmbH, Berlin, DE) was used for analysis of HPV-DNA status and genotyping as per manufacturer instructions.

For the LMU-KKG cohort, genomic DNA extraction was performed using AllPrep DNA/RNA FFPE kit (Qiagen). HPV DNA status was determined by quantitative real-time qPCR detecting the HPV16 L1 gene as previously described.[2]

## 2.4 | p16 immunohistochemistry staining

For all patient samples, the CINtec Histology kit (Roche mtm laboratories AG, Basel, CH) was used to assess p16 status at the DKTK partner site Dresden for the retrospective PORT-C cohort of the DKTK-ROG and the Dresden cohort and at DKTK partner site Munich for the LMU-KKG cohort, as described previously for each cohort.[9,14,15] Tumors with strong and diffuse nuclear and cytoplasmic staining in ≥70% of tumor cells were considered as p16 positive.[9,14,15]

## 2.5 | Methylation profiling

DNA methylation profiling was performed using Illumina microarrays based on Infinium HumanMethylation 450 (450K) array for the retrospective postoperative cohort of the DKTK-ROG and the retrospective cohort of the LMU (LMU-KKG) or the EPIC (850K) array for the retrospective Dresden cohort as per the manufacturer instructions (Illumina, San Diego, California) at the Genomics and Proteomics Core Facility of the German Cancer Research Center (DKFZ). Functional normalization of DNA methylation was conducted by performing background correction, dye bias correction and probe type bias correction using "minfi" package in R.[16,17] Cohorts were processed separately. No batch control measures were performed. Probes mapping to sex chromosomes, SNPs, cross-reactive probes and probes that were not detected on the array were filtered out leaving 431 537 probes for analysis.

## 2.6 | Model building and statistical analysis

### 2.6.1 | Identification of differentially methylated probes

Statistical analyses were performed using R (version 3.5.1)[18] in Rstudio environment.[19] We built our classifier in a step-wise series of models. First, bumphunter was used to identify differentially methylated probes (DMPs) (at Family Wise Error Rate [FWER] <0.05) between tumors with progression (n = 43) vs no progression (n = 85) using the "minfi" package.[17] Logistic regression using progression as a binary classification (progression vs none) was then performed for every DMP using "glmnet" package in R. Probes with *P*-log-rank-test (pLRT) ≤.05 were included in the bootstrap analysis LASSO.[20]

**TABLE 1** Demographic characteristics of the study cohorts

| Characteristics | Training (N = 128) | Validation (N = 125) | *P*-value |
|---|---|---|---|
| Gender | | | .8 |
| Male | 99 (77%) | 102 (82%) | |
| Female | 29 (23%) | 23 (18%) | |
| Age, median (years) | 56 | 57 | .8 |
| Anatomical site | | | ***.002*** |
| Oral cavity | 46 (36%) | 72 (58%) | |
| Oropharynx | 60 (47%) | 36 (29%) | |
| Hypopharynx | 22 (17%) | 17 (14%) | |
| Tobacco use | | | .08 |
| Current smoker | 82 (64%) | 83 (66%) | |
| Never smoker | 6 (5%) | 15 (12%) | |
| Missing | 40 (31%) | 30 (22%) | |
| Alcohol consumption | | | .7 |
| Ever drinker | 63 (49%) | 80 (64%) | |
| Never drinker | 12 (9%) | 12 (9%) | |
| Missing | 53 (42%) | 33 (27%) | |
| p16 IHC | | | ***<.0001*** |
| Positive | 19 (15%) | 1 (1%) | |
| Negative | 108 (84%) | 124 (99%) | |
| Missing | 1 (1%) | 0 (0%) | |
| pT-stage | | | .6 |
| T1-T2 | 75 (59%) | 78 (62%) | |
| T3-T4 | 53 (41%) | 47 (38%) | |
| pN-stage | | | ***.04*** |
| N0-N1 | 34 (27%) | 49 (39%) | |
| N2-N3 | 94 (73%) | 76 (61%) | |
| UICC stage | | | .3 |
| II | 4 (3%) | 6 (5%) | |
| III | 23 (18%) | 31 (25%) | |
| IV | 101 (79%) | 88 (70%) | |
| Resection margins | | | ***.001*** |
| R0 | 74 (58%) | 91 (73%) | |
| R1 | 54 (42%) | 27 (22%) | |
| Missing | 0 (0%) | 7 (5%) | |
| ECE | | | ***.01*** |
| Positive | 63 (49%) | 42 (34%) | |
| Negative | 65 (51%) | 83 (66%) | |
| Chemotherapy | | | ***<.00001*** |
| Cisplatin-based | 128 (100%) | 30 (24%) | |
| Other | 0 (0%) | 24 (19%) | |
| None | 0 (0%) | 71 (57%) | |
| RT | | | |
| Median dose (Gy) | 64 | 64 | .8 |
| Range (Gy) | 56-68.4 | 56-70 | .35 |
| OS (events) | 50 (39%) | 74 (59%) | ***.002*** |
| Months to death: Median [range] | 18.2 [2.5-74.5] | 15.2 [2-129] | .8 |
| LR (events) | 27 (21%) | 31 (25%) | .6 |
| Months to LR: Median [range] | 9.6 [2–68] | 8.4 [2.73-119] | |
| DM (Events) | 28 (22%) | 24 (19%) | .6 |
| Months to DM: Median [range] | 8.3 [2-42] | 11.7 [1-40] | .3 |
| PD (events) | 43 (34%) | 44 (35%) | .9 |
| Months to PD: Median [range] | 9.6 [2–68] | 10.2 [1-119] | .3 |

*Note:* Statistically significant *P*-values are in bold and italics.

### 2.6.2 | Boostrap LASSO and feature selection

DMPs that were significant on logistic regression were then included in a leave-5%-out boostrap LASSO for 250 iterations using "glmnet" package.[20] Sensitivity and specificity of prediction were calculated for models derived from the 250 iterations. Only models with the highest sensitivity and specificity were retained for feature selection and used to calculate the average coefficient for every DMP. Per patient, the risk score was calculated by summing the product of the methylation value of a given probe multiplied by its LASSO-derived coefficient. The cutoff with highest accuracy for classification into high- vs low-risk groups was identified by building a Receiver-Operated Curve (ROC) using "ROCR" package.[21] Kaplan-Meier (KM) curves were compared to estimate the difference in progression, OS, LR and DM between high- vs low-risk groups using pLRT < .05 for significance in "survival"[22] and "ggsurv"[23] packages. Median time to clinical event and 95% confidence intervals were calculated in "survival" package.

### 2.6.3 | Random forest classification, prediction

Methylation data for the training set was derived using Illumina Methylation 450K microarrays. However, in 72/125 (58%) of our validation set, the newer Illumina Methylation EPIC 850K microarray technology was used, where 2 out of 40-signature probes lacked representation. Consequently, we recalculated patient risk scores based on the 38 remaining probes, with adjustment of model cutoff. To circumvent thresholding issues for our numerical cutoff across microarray platforms and tissue batches, we opted for random forest (RF) classification in the training set using "randomForest."[24] The forest was trained using LASSO-derived risk assignment for 500 trees (seed set to 1) and out of bag error (OOB) was calculated. The forest then predicted risk groups for patients in the validation cohort.

Multivariate Cox regression was performed to adjust for clinical parameters. Permuted Fisher's test and Student's *t*-test tested for

differences in clinical parameters across cohorts (training vs validation) and within cohorts across risk groups (high- vs low-risk).

### 2.6.4 | Integration of previously reported biomarkers

The prognostic impact of two hypoxia gene signatures, expression of Cancer Stem Cell (CSC) markers (CD44 protein, *SLC3A2* gene/CD98H), expression of a 5-microRNA signature and staining of immune markers by immunohistochemistry (CD3, CD8, PD1 and PD-L1) have been reported for the PORT-C cohort of the DKTK-ROG.[2,5,8,11] We tested for associations between assignment using the methylation classifier HICR and assignment using the parameters (15-gene, 26-gene hypoxia-associated signatures, CD44 protein, *SLC3A2*/CD98H, 5-miRNA signature, CD3, CD8, PD1 and PD-L1) in the training cohort. Fisher's test tested for enrichments between high- vs low-risk groups. Multivariate Cox regressions were thereafter performed to adjust for those differentially enriched parameters. Finally, univariate regressions tested impact of biomarkers on PD and OS in HICR high- vs low-risk groups separately.

## 3 | RESULTS

### 3.1 | Patient cohorts

Clinicopathologic characteristics of the training (n = 128) and validation (n = 125) cohorts are displayed in Table 1. Cohorts were balanced for age, gender, smoking history, pT-stage and UICC stage. Compared to the training cohort, the validation cohort was enriched in oral cavity tumors (58% vs 36%, $P < .002$) with fewer oropharyngeal tumors (29% vs 47%, $P < .002$) and p16 positive tumors (1% vs 15%, $P < .0001$), higher incidence of N0-N1 staging (39% vs 27%, $P = .04$), clean margins following surgical resection (R0, 73% vs 58%, $P = .001$) and ECE-negativity (66% vs 51%, $P = .01$). Fewer patients in the validation cohort received concurrent chemotherapy (43% vs 100%, $P < .00001$). There was no difference in RT dose prescribed or in the rates of LR (25% vs 21%, $P = .6$); DM (19% vs 22%, $P = .3$) or all event PD (35% vs 34%, $P = .9$). A significantly higher number of deaths occurred in the validation cohort (59% vs 39%, $P = .002$) although median time to death was comparable in both cohorts (15 vs 18 months, $P = .8$).

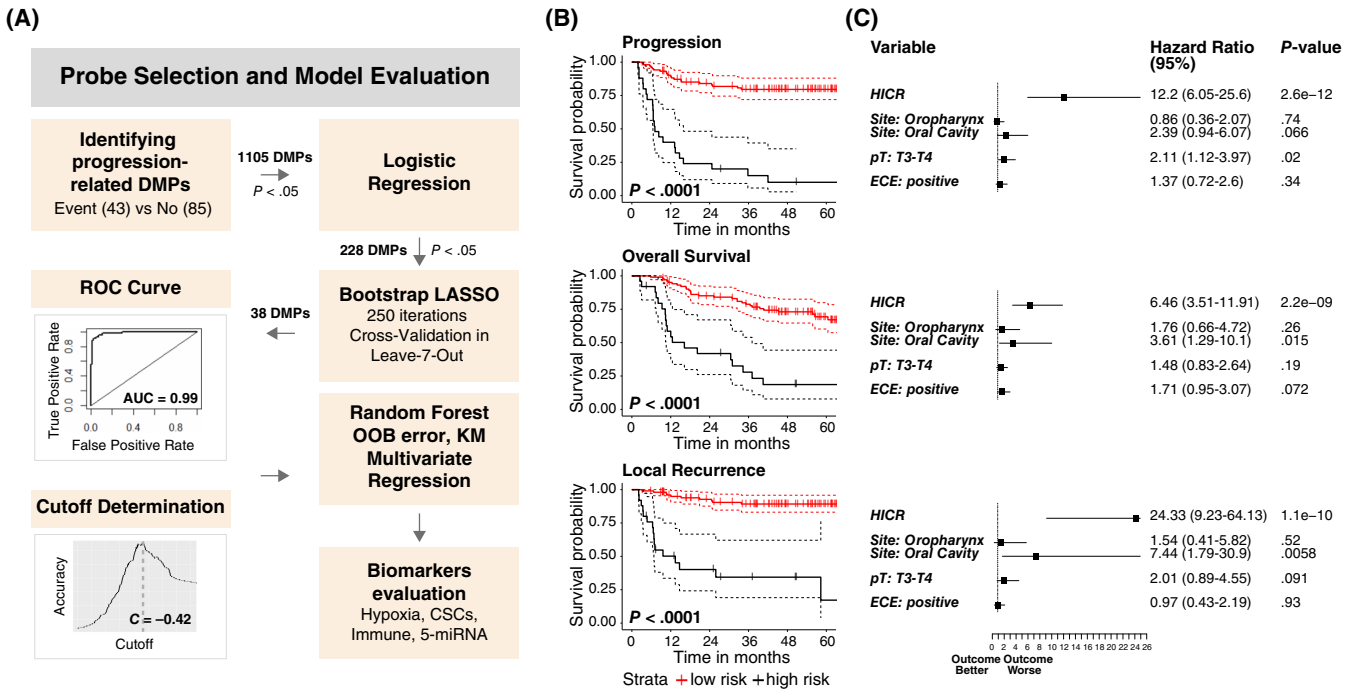


**FIGURE 1** Development of the HPV-independent Classifier of Recurrence (HICR). (A) Schematic workflow. HICR was developed in a step-wise series after identifying differentially methylated probes (DMPs) between tumors with PD vs no PD after PORT-C in the DKTK-ROG training cohort. A bootstrap LASSO selected 40 features for a linear model, where the sum of methylation values multiplied by LASSO-coefficients resulted in a model with the highest sensitivity and specificity for predicting PD. Two probes were removed from the model due to lack of representation in 850K Illumina methylation platform. The resulting 38-methylation probe model attributed a risk score for every patient. ROC analysis identified the cutoff as −0.42 for risk stratification (low-risk < −0.42 or high-risk ≥ −0.42). A Random Forest was then trained to yield a binary classification (HICR high vs low). HICR classes were predicted in an independent retrospective validation cohort. KM survival curves and multivariate cox regressions were used to adjust for the role of clinical parameters. The impact of 5-miR-score, hypoxia- and CSCs gene-signatures and immune profiling parameter on HICR classified high- vs low-risk groups was further evaluated. (B) KM curves of OS, PD, LR among patients with HPV-DNA negative HNSCC in the retrospective arm of the DKTK-ROG PORT-C cohort. Patients are stratified by HICR into high-risk (n = 25) vs low-risk (n = 103) groups. *P* values comparing risk groups were calculated using Mantel-Cox log-rank test. (C) Forest plot adjusting for clinicopathological variables, HICR remains significantly associated with worsened all-event PD, OS and LR ($P < .05$). Advanced T-stage (pT3-T4 tumors) is significantly associated with worsened all-event PD and shows a trend toward increased LR rates. ECE shows a trend towards worsened OS [Color figure can be viewed at wileyonlinelibrary.com]

### 3.2 | Feature selection

The schematic workflow for the development of the HICR classifier is shown in Figure 1. Between tumors with PD (n = 43) vs no PD (n = 85), 1105 differentially methylated probes (DMPs) were statistically significant at a FWER < 0.05 in the PORT-C cohort of the DKTK-ROG (n = 128) (Appendix S1: statistical output). Two hundred and twenty-eight remained statistically significant for PD upon logistic regression ($P < .05$). Bootstrap LASSO regression identified 40 probes for classifier building. Two out of 40 probes were missing from the more recent 850 K-EPIC array and were consequently removed from the model. Thereby, a 38 probe-containing signature was retained. Among these, 26 had functional annotations according to the Illumina manifest. Twelve probes lacked functional annotations. Two probes mapped to *PCDHB4* transcription start site (TSS). Two probes mapped to *INPP5A* gene body. From nonannotated probes, three probes mapped to a region near *ROPN1B* and two probes near *ESM1*.

### 3.3 | Building of the LASSO model

Methylation scores (M-scores) were computed per patient sample (range −4.66 to 3.13). AUC was 0.99 for PD (yes vs none) and an optimal cutoff was determined at −0.42. Patients with risk scores less than −0.42 had better clinical outcomes. Using M-scores, 88 patients were classified as low-risk (5/88 recurrence events [5.6%]) vs 40 patients with a high-risk score (38/40 events [95%]). KM curves are shown in Figure S1.

### 3.4 | Random forest classification

To circumvent cutoff thresholding issues across microarray platforms, a random forest (HICR) was trained on the 38 DMPs using the −0.42 cutoff. OOB error was 14.8%. The forest classified 25/128 patients as "high-risk" (20%) vs 103/128 patients as "low-risk" (80%). Patients in low-risk group had significantly lower rates of disease recurrence (20/103 [19.4%] vs 23/25 [92%], $P < 10^{-9}$),

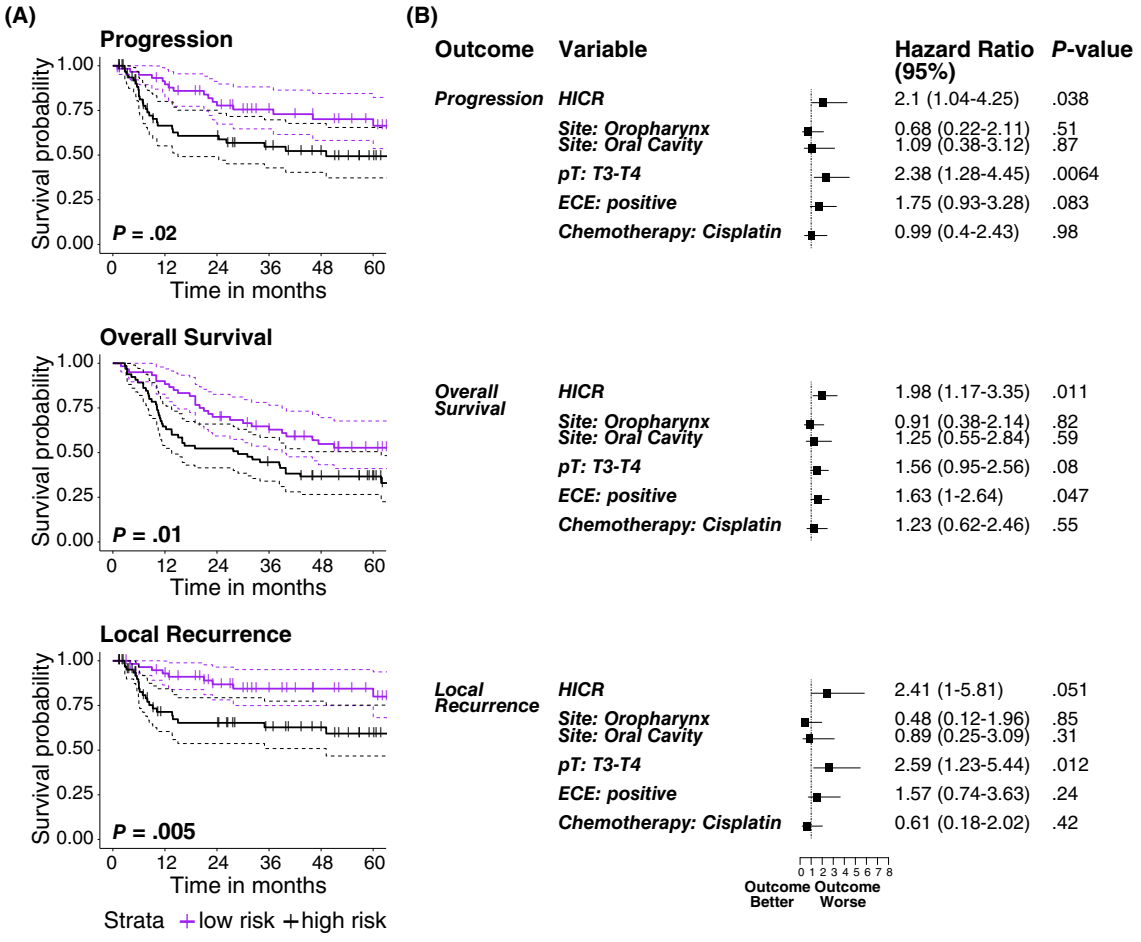


**FIGURE 2** HICR performance in the validation cohort. (A) Kaplan-Meier curves of OS, PD and LR for HPV-HNSCC patients treated with PORT-C in the validation cohort. Patients are stratified by HICR high (n = 65) vs low (n = 60). (B) Forest plot adjusting for clinicopathological variables, HICR remains significantly associated with worsened all-event PD, OS and LR ($P < .05$). Advanced T-stage is significantly associated with worsened all-event PD and LR and shows a trend towards worsened OS. ECE is significantly associated with worsened OS and shows a trend towards worsened all-event PD [Color figure can be viewed at wileyonlinelibrary.com]

LR (10/103 [9.7%] vs 17/25 [68%], $P < 10^{-9}$), DM (13/103 [12.6%] vs 15/25 [60%], $P < 10^{-9}$) and death (30/103 [29%] vs 20/25 [80%], $P < 10^{-9}$). KM curves are shown in Figure 1 and for DM in Figure S2.

## 3.5 | Signature validation

The HICR classifier was used to predict risk groups for the validation cohort (n = 125) based on the methylation status of 38 probes.

**TABLE 2** Clinical and biological parameters in high- vs low-risk HICR groups in the training and validation cohorts

| Cohort | Training (n = 128) | | | Validation (n = 125) | | |
|---|---|---|---|---|---|---|
| Risk group | High (n = 25) | Low (n = 103) | *P*-value | High (n = 65) | Low (n = 60) | *P*-value |
| *Clinical characteristics* | | | | | | |
| Gender | | | .99 | | | .10 |
| Male | 20 (80%) | 79 (77%) | | 57 (88%) | 45 (75%) | |
| Female | 5 (20%) | 24 (23%) | | 8 (12%) | 15 (25%) | |
| Age, median (years) | 55 | 57 | ***.03*** | 52.7 | 60.5 | ***.0005*** |
| Anatomical site | | | .9 | | | ***.001*** |
| Oral cavity | 10 (40%) | 36 (35%) | | 47 (72%) | 25 (42%) | |
| Oropharynx | 11 (44%) | 49 (48%) | | 14 (21%) | 22 (37%) | |
| Hypopharynx | 4 (16%) | 18 (17%) | | 4 (6%) | 13 (22%) | |
| Smoking history | | | .7 | | | .09 |
| Current | 20 (80%) | 62 (60%) | | 45 (69%) | 38 (63%) | |
| Never smoker | 1 (4%) | 5 (5%) | | 4 (6%) | 11 (18%) | |
| Missing | 4 (16%) | 36 (35%) | | 16 (25%) | 33 (19%) | |
| pT-stage | | | .1 | | | .9 |
| T1-T2 | 11 (44%) | 64 (62%) | | 41 (63%) | 37 (62%) | |
| T3-T4 | 14 (56%) | 39 (38%) | | 24 (37%) | 23 (38%) | |
| pN-stage | | | .2 | | | .9 |
| N0-N1 | 4 (16%) | 30 (29%) | | 26 (40%) | 23 (38%) | |
| N2-N3 | 21 (84%) | 73 (71%) | | 39 (60%) | 37 (62%) | |
| UICC stage | | | .9 | | | .1 |
| II | 0 (0%) | 4 (4%) | | 1 (2%) | 5 (8%) | |
| III | 4 (16%) | 19 (18%) | | 19 (29%) | 12 (20%) | |
| IV | 21 (84%) | 80 (78%) | | 45 (69%) | 43 (72%) | |
| p16 IHC | | | ***.02*** | | | .9 |
| Positive | 0 (0%) | 19 (18%) | | 1 (1%) | 0 (0%) | |
| Negative | 25 (100%) | 83 (81%) | | 64 (99%) | 60 (100%) | |
| Missing | 0 (0%) | 1 (1%) | | 0 (0%) | 0 (0%) | |
| p53 IHC | | | .3 | | | NA |
| Overexpressed | 16 (64%) | 53 (51%) | | NA | NA | |
| Negative | 9 (36%) | 50 (49%) | | NA | NA | |
| Resection margins | | | .5 | | | .7 |
| R0 | 16 (64%) | 58 (56%) | | 49 (75%) | 42 (70%) | |
| R1 | 9 (36%) | 45 (44%) | | 13 (20%) | 14 (23%) | |
| Missing | 0 (0%) | 0 (0%) | | 3 (5%) | 4 (0%) | |
| ECE | | | .2 | | | .6 |
| Positive | 16 (64%) | 54 (52%) | | 20 (31%) | 22 (37%) | |
| Negative | 9 (36%) | 49 (48%) | | 45 (69%) | 38 (63%) | |
| Chemotherapy | | | .9 | | | ***.0005*** |
| Cisplatin-based | 40 (100%) | 88 (100%) | | 4 (6%) | 26 (43%) | |
| Other | 0 (0%) | 0 (0%) | | 11 (17%) | 13 (22%) | |
| None | 0 (0%) | 0 (0%) | | 50 (77%) | 21 (35%) | |
| RT | | | .4 | | | 0.2 |
| Median dose | 64.8 | 64 | | 64 | 64 | |
| Range (Gy) | 56-66 | 59.4-68.4 | | 60-66 | 56–70 | |

**TABLE 2** (Continued)

| Cohort | Training (n = 128) | | |
|---|---|---|---|
| Risk group | High (n = 25) | Low (n = 103) | *P*-value |
| *Biological parameters* | | | |
| Hypoxia (15-gene) | | | ***.05*** |
| High | 19 (76%) | 55 (53%) | |
| Low | 6 (24%) | 48 (47%) | |
| Hypoxia (26-gene) | | | .6 |
| High | 19 (76%) | 72 (70%) | |
| Low | 5 (20%) | 29 (28%) | |
| Missing | 1 (4%) | 2 (2%) | |
| CD44 protein | | | .7 |
| High | 23 (92%) | 87 (84%) | |
| Low | 1 (4%) | 9 (9%) | |
| Missing | 1 (4%) | 7 (7%) | |
| SLC3A2 mRNA/CD98H | | | .3 |
| High | 21 (84%) | 76 (74%) | |
| Low | 3 (12%) | 25 (24%) | |
| Missing | 1 (4%) | 2 (2%) | |
| CD8 IHC (≥6) | | | .3 |
| Positive | 4 (22%) | 28 (37%) | |
| Negative | 14 (78%) | 47 (63%) | |
| Missing | 7 (28%) | 28 (27%) | |
| CD3 IHC (≥6) | | | .4 |
| Positive | 4 (16%) | 27 (26%) | |
| Negative | 14 (56%) | 48 (47%) | |
| Missing | 7 (28%) | 28 (27%) | |
| PD1 staining | | | .8 |
| Positive | 6 (24%) | 30 (29%) | |
| Negative | 12 (48%) | 45 (44%) | |
| Missing | 7 (28%) | 28 (27%) | |
| PD-L1 staining | | | .4 |
| Positive | 4 (16%) | 25 (24%) | |
| Negative | 14 (56%) | 50 (49%) | |
| Missing | 7 (28%) | 28 (27%) | |
| 5-miRNA signature | | | ***.01*** |
| High-risk | 14 (56%) | 27 (26%) | |
| Low-risk | 4 (16%) | 36 (35%) | |
| Missing | 7 (28%) | 40 (39%) | |

*Note:* Statistically significant *P*-values are in bold and italics.

Sixty-five out of one hundred and twenty-five patients (52%) were classified as high-risk vs 60/125 (48%) patients as low-risk. Impact of predicted risk status on clinical outcomes is shown in Figure 2. Predicted low-risk patients had a significantly lower probability of disease recurrence compared to predicted high-risk patients (16/60 [26.7%] vs 28/65 [43%], $P = .02$), death (29/60 [48.3%] vs 45/65 [69.2%], $P = .01$) and LR (9/60 [15%] vs 22/65 [35%], $P = .005$). DM rates were similar for low- vs high-risk groups (11/60 [18.3%] vs 13/65 [20%], $P = .3$; Figure S2).

## 3.6 | Evaluation of clinical parameters

Univariate cox regression evaluated the impact of clinicopathologic parameters on LR, PD and OS in both cohorts, as seen in Tables S1 to S3. Evaluated parameters were HICR-class (high- vs low-risk), pT-stage (T3-T4 vs T1-T2), pN-stage (N2-N3 vs N0-N1), anatomical site (oropharynx, oral cavity, hypopharynx), ECE and resection status. Additionally, impact of p16 immunohistochemistry (positive vs negative) was investigated in the training cohort (Tables S1–S3; Figure S3).

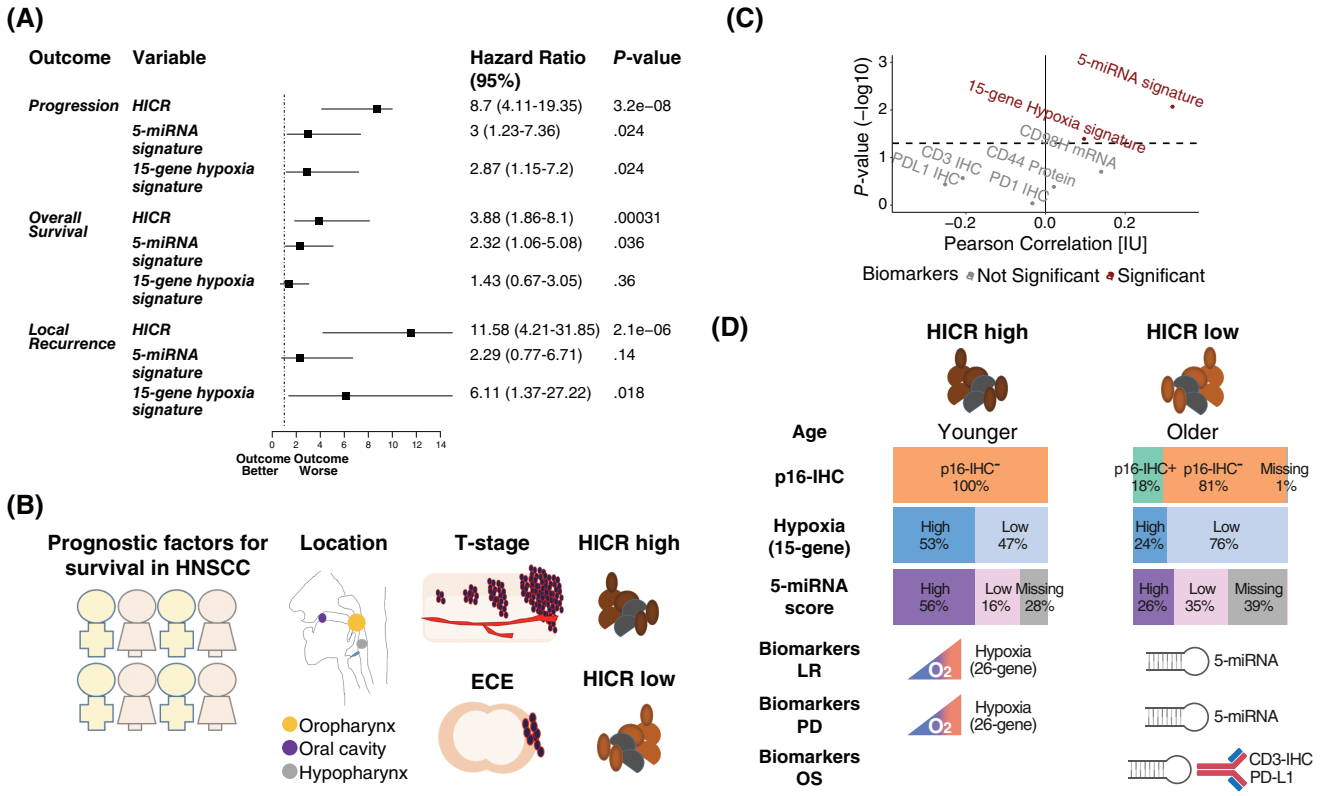


**FIGURE 3** Multiscale biomarker correlation and characterization of HICR. (A) Multivariate analysis revealed that HICR, the 15-gene hypoxia signature and the 5-miRNA signature are independently associated with PD ($P < .05$). Additionally, the 5-miRNA signature was associated with worsened OS whereas the 15-gene hypoxia signature was associated with worsened LR. (B) Graphical representation of clinicopathological factors that were univariate predictors of PD (HICR, T-stage; ECE), OS (HICR, Hypopharyngeal tumors, T-stage, ECE) or LR (HICR, oropharyngeal tumors, T-Stage, ECE) in the DKTK-ROG training cohort. (C) Volcano plot displaying Pearson's correlation (*x*-axis: correlation, *y*-axis: −log(*P*-value) between HICR and hypoxia signatures, 5-miRNA signature, CD44 protein, SLC3A2 gene expression and immune cell infiltration. Classification by HICR and the 15-gene hypoxia signature were weakly correlated ($r = .1$, $P = .04$). A moderate correlation was found between HICR and the 5-miRNA signature ($r = .32$, $P = .008$). The dashed line represents the threshold for statistical significance ($P < .05$ corresponding to (−log[*P*-value] > 1.3) and statistically significant correlations are in brown. (D) Characteristics of HICR tumors. Patients with HICR high tumors are more likely to be younger, with tumors that are hypoxic, with a high-risk 5-miRNA score. On univariate analysis, the 26-hypoxia gene signature is an independent predictor of LR and PD in HICR high tumors. For HICR low-risk tumors, 5-miRNA signature and immune cell infiltration (CD3-IHC, PD-L1) are predictors of OS, PD and LR on univariate analysis [Color figure can be viewed at wileyonlinelibrary.com]

Impact of chemotherapy agents (cisplatin vs others vs none) was investigated in the validation cohort.

In the training cohort, HICR class, localization, pT-stage and ECE were significant on univariate analysis for LR, PD and OS. p16 status and pN-stage were not significant. In the validation cohort, HICR class was significant for all outcomes. pT-stage were significantly associated with PD, LR but not OS.

On multivariate analysis, HICR class retained its prognostic significance after adjusting for anatomical location, T-stage, ECE and additionally in the case of the validation cohort, chemotherapy type for PD, LR and OS. Results are shown in Figures 1 and 2. HICR also retained its prognostic significance after adjusting for tobacco smoking on multivariate analysis, as seen in Figures S4 and S5.

Finally, we tested for associations between high- vs low-risk groups and clinicopathological parameters. Age at diagnosis was lower in high-risk groups in both cohorts (training: median age 55 vs 57, $P = .03$; validation: 52.7 vs 60.5, $P = .0005$). In the training cohort, patients in the high-risk group were less likely to be p16 IHC positive (0% vs 18%, $P = .02$). In the validation cohort, patients in the high-risk group were more likely to have oral cavity tumors (72% vs 42%, $P = .0006$) and less likely to have hypopharyngeal tumors (6% vs 22%, $P = .02$) and less patients received concurrent chemotherapy (77% vs 35%, $P = .0005$). High- vs low-risk groups were otherwise balanced for all other clinicopathological parameters, as seen in Table 2.

## 3.7 | Impact of hypoxia, CSCs, microRNA signature and immune marker assignment

Within the training cohort, we evaluated the relationship between HICR risk group and patient assignment according to previously reported DKTK-ROG biomarkers. Upon stratification by HICR, tumors in the high-risk group had higher gene expression for the 15-gene

hypoxia signature (76% vs 53%, $P = .05$) and 5-miRNA signature (56% vs 26%, $P = .01$; see Table 2). Risk groups were otherwise balanced for assignment using the 26-gene hypoxia signature, the CSCs marker CD44 protein and *SCL3A2* gene as well as immune-parameter CD3/CD8 influx and PD1 and PD-L1 status.

Similarly, HICR was significantly correlated with the 5-miRNA signature ($r = .32$, $P = .009$) and the 15-gene hypoxia signature ($r = .1$, $P = .04$) (Figure 3). Given these findings, we performed multivariate analyses adjusting for the 15-gene hypoxia and miRNA signatures. The HICR classifier remained prognostic for all outcomes (PD: HR = 8.7, $P < 3 \times 10^{-8}$; OS: HR = 3.88, $P < .003$; LR: HR = 11.58, $P < 2 \times 10^{-6}$; DM: HR = 7.8, $P < 4 \times 10^{-7}$). The miRNA score was also prognostic (PD: HR = 3.0, $P = .02$; OS: HR = 2.32, $P = .04$; DM: HR = 4.3, $P = .03$) but not for LR. The 15-gene hypoxia signature was significant for PD (HR = 2.87, $P = .02$) and LR (HR = 6.1, $P = .02$) as seen in Figure 3 but not for OS nor DM (Figure S2).

We performed univariate Cox regressions for each parameter in the HICR high-risk (n = 25) and low-risk (n = 103) groups separately (Tables S5 and S6). In the HICR high-risk group, tumor hypoxia determined by the 26-gene signature was prognostic for PD (HR = 5.57, $P = .009$), LR (HR = 8.32, $P = .04$) with a trend for OS (HR = 3.21, $P = .07$). The 15-gene signature showed a trend for PD (HR = 2.65, $P = .07$) but was not significant for OS ($P = .3$).

In the low-risk group, the 5-miRNA signature was prognostic for PD (HR = 6.07, $P = .006$) and LR (HR = 9.44, $P = .04$) while hypoxia, CSCs and immune markers were not. OS was significantly associated with higher expression of the 5-miRNA signature (HR = 3.07, $P = .02$), CD8 (HR = 0.33, $P = .04$), CD3 (HR = 0.28, $P = .02$) and PD-L1 (HR = 0.24, $P = .02$) markers.

## 4 | DISCUSSION

Our study reports the discovery of a novel and robust DNA methylation-based classifier of HPV-negative HNSCC treated with PORT-C. The classifier was an independent prognostic factor, after adjusting for clinical parameters or previously reported biomarkers. There were several advantages for using the retrospective PORT-C cohort of the DKTK-ROG as our training set. This multicentric cohort consisted of patients homogeneously treated with standard postoperative cisplatin-based radiochemotherapy (RCHT). In comparison, patients in the validation cohort were treated prior to the standard of cisplatin-based RCHT (EORTC 22931[25] and RTOG 950[25]). Nevertheless, HICR successfully stratified patients in both cohorts despite treatment heterogeneities, highlighting its robustness.

Direct testing for HPV-DNA by PCR-based methods was used instead of p16 immunohistochemistry because HPV-DNA was the superior prognosticator of locoregional control in the retrospective PORT-C cohort of the DKTK-ROG.[9] From a technical viewpoint, PCR-based HPV-DNA assay provides a direct detection method for HPV subtypes and has higher sensitivity compared to p16-IHC staining.[26] Therefore, exclusion based on HPV-DNA positivity is more stringent than exclusion by p16-IHC.

DNA methylation-based classifiers have high translational potential for a spectrum of reasons. Methylation can be profiled using genomic DNA derived from routine FFPE samples with only nanogram amounts.[13] In contrast to gene expression profiling, DNA methylation patterns require long-term reprograming and are less influenced by surgical variables such as clamping time. Furthermore, it is robust and reproducible even in samples showing lower quality DNA, that is, higher degrees of cross-links and degradation. Moreover, in our study, the HICR classifier performed well on two different Illumina microarray platform-types (450K and 850K-EPIC), despite losing representation of two classifier probes on the newer EPIC arrays.

Among genes associated with the 38-classifier probes, hypermethylation vs hypomethylation was already reported in different tumor types (with methylation changes being in the same direction as in our study). For instance, hypomethylation of ZNF154 was associated with worsened OS in pancreatic cancer[27] while TTC28 hypomethylation, a target of TP53, was associated with genetic instability and L1-element transpositions in colorectal cancer.[28] Likewise, DCC was reportedly stably hypermethylated in HNSCC, independent of smoking status.[29] Disease recurrence in bladder cancer was reported for hypermethylation of EOMES, a gene involved in developmental processes, T-cell differentiation and whose silencing is attributed to CD8+ T-cell exhaustion and dysfunction.[30] SIM1 hypermethylation was reported in cervical tumors (cell free DNA; brushings)[31] and lung adenocarcinoma.[32] HOXC9 was reported as silenced via methylation in breast cancer cell lines.[33]

There were no reported aberrant patterns of methylation for INPP5A or ESM1 genes, each represented with two probes among the 38 HICR set. Endothelial cell specific molecule 1 (ESM-1 or endocan) is associated with tumor angiogenesis (endothelial tip cell formation), inflammation and VEGF-dependent vascular permeability.[34] ESM-1 is upregulated in a number of cancer entities and is associated with cell proliferation, survival, epithelial to mesenchymal transition (EMT) and invasion in colorectal cancer via interaction with, for example, NF-κB signaling.[35] Higher expression of ESM1 was associated with worsened survival in triple negative breast cancer and gastric cancer as well as increased radioresistance in breast cancer- and HNSCC cells.[36-38]

Inositol polyphosphate 5-phosphatase (INPP5A) seems to elicit tumorsupressive effects. In one study, its silencing was shown to increase intracellular inositol 1,4,5-trisphosphate and inositol 1,3,4,5-tetrakisphosphate, leading to cell transformation and tumor formation in nude mice models[39] Furthermore, loss of INPP5A expression was associated with decreased OS in recurrent/metastatic cutaneous squamous cell carcinoma (n = 52).[40] Description of genes associated with the classifier probes is in Table S4.

Additionally, genes mapping to the 1105 DMPs correlating with PD were examined in more detail. There was a trend toward hypomethylation, as hypomethylation was associated with a higher risk in 925/1105 probes (84%). The top hypomethylated probes mapped to metallothionein-1M (MT1M) and T-cell leukemia homeobox protein 1 (TLX1). Septin 9 (SEPT9), matrix metalloprotein-9 (MMP9) and oral

cancer overexpressed 1 (ORAOV1) were among the most strongly hypermethylated probes. Therefore, in addition to its role as a robust and reliable classifier, HICR may provide a causal link to functional biological processes that govern tumor radioresistance and local failure after PORT-C. Taken together, there were prominent genes among the DMPs that performed congruent functions with multiple components of the matrix metalloproteinase pathway prominent genes (ESM1) and genes associated with invasion and tissue remodeling such as CUX1, TBX2, SEPT9, MT1A, MT1E, MT1M, MMP2 and MMP9. Likewise, tumor stem cell markers like TLX1-3 and CD133 (PROM1) were identified among DMPs.

Within the context of tumor biology, the classifier remained significant after adjusting for previously reported biomarkers. Equally interestingly, the impact of these biomarkers varied as a function of HICR risk groups. In the HICR high-risk group, hypoxia was adversely associated with local recurrence and OS (Table S5). In the HICR low-risk group, higher levels of CD3 and PD-L1 markers by IHC were prognostic of better treatment outcomes (Table S6). Though HICR was independently prognostic, it may also be instructive to preselect patients for biologically individualized therapy, given the evidence coming from RCTs for the favorable role for hypoxia sensitizers/modifiers[41] and immune check point inhibitors[42] (such as PD-1/PD-L1 antibodies) in improving OS in patients with HNSCC. In practical terms, HICR high-risk groups might benefit from hypoxia imaging and radiotherapy dose intensification or high-linear energy transfer (LET) irradiation.[10,43,44] Additionally, HICR low-risk tumors may benefit from additional stratification using the 5-miR-signature for identification of patients at risk for progressive disease. The complementary and independent prognostic value of the 5-miRNA signature for DMFS, PD and OS is noteworthy. Together, these results highlight the relevance of epigenetic and posttranslational control of gene expression in tumor biology and development of novel biomarkers.

## 5 | CONCLUSION

In our study, we propose a 38-probe DNA methylation signature to classify HPV negative HNSCC with potential to stratify patients for therapy intensification/de-escalation strategies. Prospective validation of this methylation classifier is ongoing in the PORT-C arm of the currently recruiting HNprädBio study of the DKTK-ROG (www.clinicaltrials.gov, NCT02059668).

### ACKNOWLEDGMENTS

We would like to thank Carmen Klein, Barbara Schwager and Claudia Rittmüller for their excellent technical assistance, the DKFZ Genomics & Proteomics Core Facility for providing excellent support for microarray experiments as well as all members of the DKTK-ROG and related multidisciplinary departments at Heidelberg Core and all DKTK-partner sites who contributed to the success of this trial. Our study was supported by a Joint Funding Grant within the German Cancer Consortium (DKTK), which is funded as one of the National German Health Centers by the Federal German Ministry of Education (training cohort), the Helmholtz Cross-Program Initiative Personalized Medicine (iMed) project on "Multi-Scale Integrative Biology of HNSCC" (validation cohorts) and Intramural Funds of the National Center for Tumor Diseases (NCT) Heidelberg Radiation Oncology Program. Open Access funding enabled and organized by Projekt DEAL.

### CONFLICT OF INTEREST

The authors declare no conflict of interest in connection with this paper. In the past 5 years, Dr. Michael Baumann received funding for his research projects and for educational grants to the University of Dresden by Bayer AG (2016-2018), Merck KGaA (2014-open) and Medipan GmbH (2014-2018). He is on the supervisory board of HI-STEM gGmbH (Heidelberg) for the German Cancer Research Center (DKFZ, Heidelberg) and also member of the supervisory body of the Charité University Hospital, Berlin. As former chair of OncoRay (Dresden) and present CEO and Scientific Chair of the German Cancer Research Center (DKFZ, Heidelberg), he has been or is responsible for collaborations with a multitude of companies and institutions, worldwide. In this capacity, he has discussed potential projects and signed contracts for research funding and/or collaborations with industry and academia for his institute(s) and staff, including but not limited to pharmaceutical companies such as Bayer, Boehringer Ingelheim, Bosch, Roche and other companies such as Siemens, IBA, Varian, Elekta, Bruker, etc. In this role, he was/is also responsible for the commercial technology transfer activities of his institute(s), including the creation of start-ups and licensing. This includes the DKFZ-PSMA617 related patent portfolio [WO2015055318 (A1), ANTIGEN (PSMA)] and similar IP portfolios. Dr. Baumann confirms that, to the best of his knowledge, none of the above funding sources were involved in the preparation of this article.

### DATA AVAILABILITY STATEMENT

DNA Methylome data have been deposited in the ArrayExpress database at EMBL-EBI (http://www.ebi.ac.uk/arrayexpress) under the following accession numbers (E-MTAB-10577 for the retrospective multicenter postoperative cohort of the DKTK-ROG (PORT-C), E-MTAB-10576 for the validation cohort samples from the Munich retrospective cohort (LMU-KKG) and E-MTAB-10578 for the validation cohort samples from the Dresden retrospective cohort." Further details and other data that support the findings of our study are available from the corresponding author upon request.

### ETHICS STATEMENT

Local ethics committees at all eight DKTK partner sites including Heidelberg University Hospital and University Hospital Carl Gustav Carus Dresden and at Ludwig-Maximilians-Universität München granted approval for retrospective data collection and analysis of clinical and biological data (EA 312-12, EA 448-13, EA 17-116, Dresden EK397102014).

### ORCID

*Karim Zaoui* https://orcid.org/0000-0002-6676-2273
*Ingeborg Tinhofer* https://orcid.org/0000-0002-0512-549X
*Amir Abdollahi* https://orcid.org/0000-0002-8016-5897

## REFERENCES


1. Skripcak T, Belka C, Bosch W, et al. Creating a data exchange strategy for radiotherapy research: towards federated databases and anonymised public datasets. *Radiother Oncol*. 2014;113:303-309.
2. 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;25:1505-1516.
3. Schmidt S, Linge A, Zwanenburg A, et al. Development and validation of a gene signature for patients with head and neck carcinomas treated by postoperative radio(chemo)therapy. *Clin Cancer Res*. 2018; 24:1364-1374.
4. Linge A, Löck 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 DKTK-ROG. *Clin Cancer Res*. 2016;22(11):2639–2649. https://doi.org/10.1158/1078-0432.ccr-15-1990
5. Balermpas P, Rödel F, Rödel 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;138:171-181.
6. De-Colle C, Menegakis A, Mönnich D, et al. SDF-1/CXCR4 expression is an independent negative prognostic biomarker in patients with head and neck cancer after primary radiochemotherapy. *Radiother Oncol*. 2018;126:125-131.
7. Stangl S, Tontcheva N, Sievert W, et al. Heat shock protein 70 and tumor-infiltrating NK cells as prognostic indicators for patients with squamous cell carcinoma of the head and neck after radiochemotherapy: a multicentre retrospective study of the German cancer consortium radiation oncology Gro. *Int J Cancer*. 2018;142:1911-1925.
8. Balermpas P, Rödel 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;141:594-603.
9. 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 (DKTK-ROG). *Radiother Oncol*. 2014; 113:317-323.
10. Tawk B, Schwager C, Deffaa O, et al. Abdollahi a. comparative analysis of transcriptomics based hypoxia signatures in head- and neck squamous cell carcinoma. *Radiother Oncol*. 2016;118:350-358.
11. Linge A, Lohaus F, Löck S, et al. HPV status, cancer stem cell marker expression, hypoxia gene signatures and tumour volume identify good prognosis subgroups in patients with HNSCC after primary radiochemotherapy: a multicentre retrospective study of the German cancer consortium radiation oncology group (DKTK-ROG). *Radiother Oncol*. 2016;121:364-373.
12. Digomann D, Kurth I, Tyutyunnykova A, et al. The CD98 heavy chain is a marker and regulator of head and neck squamous cell carcinoma radiosensitivity. *Clin Cancer Res*. 2019;25:3152-3163.
13. Capper D, Jones DTW, Sill M, et al. DNA methylation-based classification of central nervous system tumours. *Nature*. 2018;555: 469-474.
14. Linge A, Löck S, Krenn C, et al. Independent validation of the prognostic value of cancer stem cell marker expression and hypoxia-induced gene expression for patients with locally advanced HNSCC after postoperative radiotherapy. *Clin Transl Radiat Oncol*. 2016;1:19-26.
15. Maihoefer C, Schüttrumpf L, Macht C, et al. Postoperative (chemo) radiation in patients with squamous cell cancers of the head and neck—clinical results from the cohort of the clinical cooperation group "personalized radiotherapy in head and neck cancer". *Radiat Oncol*. 2018;13:123.
16. Fortin J-P, Labbe A, Lemire M, et al. Functional normalization of 450k methylation array data improves replication in large cancer studies. *Genome Biol*. 2014;15:503.
17. Aryee MJ, Jaffe AE, Corrada-Bravo H, et al. Minfi: a flexible and comprehensive bioconductor package for the analysis of infinium DNA methylation microarrays. *Bioinformatics*. 2014;30:1363-1369.
18. R Core Team. *R: A Language and Environment for Statistical Computing [Internet]*; July 2018. https://www.r-project.org/
19. Huber W, Carey VJ, Gentleman R, et al. Orchestrating high-throughput genomic analysis with bioconductor. *Nat Methods*. 2015; 12:115-121.
20. Friedman J, Hastie T, Tibshirani R. Regularization paths for generalized linear models via coordinate descent. *J Stat Softw*. 2010;33:1-22.
21. Sing T, Sander O, Beerenwinkel N, Lengauer T. ROCR: visualizing classifier performance in R. *Bioinformatics*. 2005;21:3940-3941.
22. Therneau TM, Grambsch PM. *Modeling Survival Data: Extending the Cox Model*. vol. 1, New York, NY: Springer; 2000. ISBN 978-0-387-98784-2
23. Alboukadel K, Marcin K, Przemyslaw B, Scheipl F. *Drawing Survival Curves using "ggplot2" [R package survminer version 0.4.3]. R Packag version 043 [Internet]*; July 2018. https://cran.r-project.org/web/packages/survminer/index.html
24. Liaw A, Wiener M. Classification and regression by randomForest. *R News*. 2002;2:18-22.
25. Bernier J, Cooper JS, Pajak TF, et al. Defining risk levels in locally advanced head and neck cancers: a comparative analysis of concurrent postoperative radiation plus chemotherapy trials of the EORTC (#22931) and RTOG (# 9501). *Head Neck*. 2005;27:843-850.
26. Prigge E-S, Arbyn M, von Knebel DM, Reuschenbach M. Diagnostic accuracy of p16 $^{INK4a}$ immunohistochemistry in oropharyngeal squamous cell carcinomas: a systematic review and meta-analysis. *Int J Cancer*. 2017;140:1186-1198.
27. Wiesmueller F, Kopke J, Aust D, et al. Silenced ZNF154 is associated with longer survival in resectable pancreatic cancer. *Int J Mol Sci*. 2019;20(21):5437. https://doi.org/10.3390/ijms20215437
28. Pitkänen E, Cajuso T, Katainen R, et al. Frequent L1 retrotranspositions originating from TTC28 in colorectal cancer. *Oncotarget*. 2014;5:853-859.
29. Zhou C, Ye M, Ni S, et al. DNA methylation biomarkers for head and neck squamous cell carcinoma. *Epigenetics*. 2018;13:398-409.
30. Reinert T, Borre M, Christiansen A, Hermann GG, Ørntoft TF, Dyrskjøt L. Diagnosis of bladder cancer recurrence based on urinary levels of EOMES, HOXA9, POU4F2, TWIST1, VIM, and ZNF154 hypermethylation. *PLoS One*. 2012;7:e46297.
31. Kim H-J, Kim CY, Jin J, et al. Aberrant *single-minded homolog 1* methylation as a potential biomarker for cervical cancer. *Diagn Cytopathol*. 2018;46:15-21.
32. Daugaard I, Dominguez D, Kjeldsen TE, et al. Identification and validation of candidate epigenetic biomarkers in lung adenocarcinoma. *Sci Rep*. 2016;6:35807.
33. Gevaert O, Tibshirani R, Plevritis SK. Pancancer analysis of DNA methylation-driven genes using MethylMix. *Genome Biol*. 2015; 16:17.
34. Rocha SF, Schiller M, Jing D, et al. Esm1 modulates endothelial tip cell behavior and vascular permeability by enhancing VEGF bioavailability. *Circ Res*. 2014;115:581-590.
35. Kang YH, Ji NY, Han SR, et al. ESM-1 regulates cell growth and metastatic process through activation of NF-κB in colorectal cancer. *Cell Signal*. 2012;24(10):1940-1949. https://doi.org/10.1016/j.cellsig.2012.06.004
36. Xu H, Chen X, Huang Z. Identification of ESM1 overexpressed in head and neck squamous cell carcinoma. *Cancer Cell Int*. 2019;19:118.

37. Jin H, Rugira T, Ko YS, et al. ESM-1 overexpression is involved in increased tumorigenesis of radiotherapy-resistant breast cancer cells. *Cancers*. 2020;12(6):1363. https://doi.org/10.3390/cancers12061363
38. Lin L-Y, Yeh Y-C, Chu C-H, et al. Endocan expression is correlated with poor progression-free survival in patients with pancreatic neuroendocrine tumors. *Medicine (Baltimore)*. 2017;96:e8262.
39. Speed CJ, Neylon CB, Little PJ, Mitchell CA. Underexpression of the 43 kDa inositol polyphosphate 5-phosphatase is associated with spontaneous calcium oscillations and enhanced calcium responses following endothelin-1 stimulation. *J Cell Sci*. 1999;112:669-679.
40. Cumsky HJL, Costello CM, Zhang N, et al. The prognostic value of inositol polyphosphate 5-phosphatase in cutaneous squamous cell carcinoma. *J Am Acad Dermatol*. 2019;80:626-632.
41. Overgaard J. Hypoxic radiosensitization: adored and ignored. *J Clin Oncol*. 2007;25:4066-4074.
42. Cohen EE, Harrington KJ, Le Tourneau C, et al. Pembrolizumab (pembro) vs standard of care (SOC) for recurrent or metastatic head and neck squamous cell carcinoma (R/M HNSCC): Phase 3 KEYNOTE-040 trial. *Ann Oncol*. 2017;28(Supplement5):v628. https://doi.org/10.1093/annonc/mdx440.040
43. Chiblak S, Tang Z, Lemke D, et al. Carbon irradiation overcomes glioma radioresistance by eradicating stem cells and forming an antiangiogenic and immunopermissive niche. *JCI Insight*. 2019;4(2):e123837. https://doi.org/10.1172/jci.insight.123837
44. Zips D, Zöphel K, Abolmaali N, et al. Exploratory prospective trial of hypoxia-specific PET imaging during radiochemotherapy in patients with locally advanced head-and-neck cancer. *Radiother Oncol*. 2012;105:21-28.

## SUPPORTING INFORMATION

Additional supporting information may be found in the online version of the article at the publisher's website.

**How to cite this article:** Tawk B, Wirkner U, Schwager C, et al. Tumor DNA-methylome derived epigenetic fingerprint identifies HPV-negative head and neck patients at risk for locoregional recurrence after postoperative radiochemotherapy. *Int. J. Cancer*. 2022;150(4):603-616. doi:10.1002/ijc.33842