Development and validation of a combined hypoxia and immune prognostic classifier for head and neck cancer

Brooks, Jill; Antao Mobre De Menezes, Albert Rahul Eugene; Ibrahim, Maha; Archer, Lucinda; Lal, Neeraj; Bagnall, Chris; Zeidler, Sandra Ventorin von; Valentine, Helen; Spruce, Rachel; Batis, Nikolaos; Bryant, Jennifer; Hartley, Margaret; Kaul, Baksho; Ryan, Gordon; Bao, Riyue; Khattri, Arun; Lee, Steve; Ogbureke, Kalu U.E.; Middleton, Gary; Tennant, Daniel

DOI:
10.1158/1078-0432.CCR-18-3314
10.1158/1078-0432.CCR-18-3314

Document Version
Peer reviewed version

Citation for published version (Harvard):
Brooks, J, Antao Mobre De Menezes, ARE, Ibrahim, M, Archer, L, Lal, N, Bagnall, C, Zeidler, SVV, Valentine, H, Spruce, R, Batis, N, Bryant, J, Hartley, M, Kaul, B, Ryan, G, Bao, R, Khattri, A, Lee, S, Ogbureke, KUE, Middleton, G, Tennant, D, Beggs, A, Deeks, J, West, CML, Cazier, J-B, Willcox, B, Seiwert, TY & Mehanna, H 2019, 'Development and validation of a combined hypoxia and immune prognostic classifier for head and neck cancer', *Clinical Cancer Research*, vol. 25, no. 17, pp. 5315-5328. https://doi.org/10.1158/1078-0432.CCR-18-3314, https://doi.org/10.1158/1078-0432.CCR-18-3314

Development and Validation of a Combined Hypoxia and Immune Prognostic Classifier for Head and Neck Cancer

Brooks, J.M.[1,2], Menezes, A.N.[2], Ibrahim M.[2,3], Archer, L.[4], Lal, N.[5], Bagnall, C.J.[6], von Zeidler, S.V.[7], Valentine, H.[8], Spruce R.J.[1,2], Batis N.[1,2], Bryant, J.L[1,2]., Hartley, M.[1,2], Kaul, B.[9], Ryan, G.B.[1,2], Bao, R.[10], Khattri, A.[10], Lee, S.P.[5], Ogbureke, K.U.E.[11], Middleton G.W.[5], Tennant, D.A.[9], Beggs, A.D.[2], Deeks, J.[4,12], West C.M.[8], Cazier, J.B.[2], Willcox B.E.[5], Seiwert T.Y.[10], Mehanna, H.M[1,2].

[1] Institute of Head and Neck Studies and Education, University of Birmingham, UK

[2] Institute of Cancer and Genomic Sciences, University of Birmingham, UK

[3]South Egypt Cancer Institute, Assiut University, Egypt

[4] Institute of Applied Health Research, University of Birmingham, UK

[5] Institute of Immunology and Immunotherapy, University of Birmingham, UK

[6] Human Biomaterials Resource Centre, University of Birmingham, UK

[7]Department of Pathology, Federal University of Espírito Santo, Brazil

[8] Division of Cancer Sciences, University of Manchester, Christie Hospital, Manchester Academic Health Science Centre, UK

[9] Institute of Metabolism and Systems Research, University of Birmingham, UK

[10] The University of Chicago Medicine, Chicago, USA

[11] Department of Diagnostic and Biomedical Sciences, The University of Texas Health Science Center at Houston, Houston Texas, USA

[12] NIHR Birmingham Biomedical Research Centre, University of Birmingham and University Hospitals Birmingham NHS Foundation Trust, UK

Running title: hypoxia-immune prognostic classifier for HNC

Key words: head and neck cancer, prognostic classifier, hypoxia, immune profile, immunotherapy

Financial support: This work was funded by Queen Elizabeth Hospital Birmingham Charity, The Get A-head Charitable Trust and Cancer Research UK, with support from South Egypt Cancer Institute, Assiut; NIHR Research Methods Fellowship NIHR-RMFI-2016-07-28; the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Finance Code 001; Major Centre; US-UK Fulbright Commission, Distinguished Chair Award; The Center for Research Informatics of The University of Chicago Biological Science Division, The Institute for Translational Medicine/Clinical and Translational Award (NIH 5UL1TR002389-02), and The University of Chicago Comprehensive Cancer Center Support Grant (NIH P30CA014599); the NIHR Birmingham Biomedical Research Centre at the University Hospitals Birmingham NHS Foundation Trust; NIHR and the University of Birmingham.

Corresponding author: Prof Hisham Mehanna, Institute of Head and Neck Studies and Education, Institute of Cancer and Genomic Sciences, University of Birmingham, Birmingham, B15 2TT. Phone: +44 (0)121 414 8753. Email: h.mehanna@bham.ac.uk

Conflict of Interests disclosure: GEM has received funding from Plexxicon and served as a consultant for MSD, BMS and Roche; TYS has received research funding from BMS, Jounce, and

Merck and served as a consultant for Arduro/Novartis, AstraZeneca, Bayer, BMS, Celgene, Innate, Merck and Nanobiotix. HMM has received research funding from GSK, MSD and AstraZeneca. All other authors declare no potential conflicts of interest.

Word count: 5340
Abstract: 250
Number of Figure and Tables: 6

Translational relevance

Comprehensive genomic characterisation of head and neck cancer (HNC) has identified subgroups of tumours with distinct molecular and biological properties. Current standard-of-care treatment for HNC does not consider these fundamental differences. This largely reflects the absence of validated biomarkers for selection of standard and/or new targeted therapies based on individual tumour biology. Focusing on two biological properties: hypoxia status and immune profile, we developed and validated a prognostic transcriptional classifier which stratifies HNC patients into three distinct hypoxia-immune phenotypes and survival profiles: hypoxia^{high}/immune^{low}; hypoxia^{low}/immune^{high} and a mixed group. Multiplex immunofluorescence staining demonstrated an inverse spatial distribution of hypoxia- and immune-response markers, corroborating our subtype classification based on gene expression. Each of the three subtypes has distinct molecular, histologic and biologic characteristics, likely to be driven by distinct targetable pathways, for which potentially effective therapies (e.g. hypoxia modification, EGFR inhibition and immune checkpoint blockade) have been demonstrated in HNC.

Abstract

Purpose: Intratumoural hypoxia and immunity have been correlated with patient outcome in various tumour settings. However, these factors are not currently considered for treatment selection in head and neck cancer (HNC) due to lack of validated biomarkers. Here we sought to develop a hypoxia-immune classifier with potential application in patient prognostication and prediction of response to targeted therapy.

Experimental design: A 54-gene hypoxia-immune signature was constructed based on literature review. Gene expression was analysed *in silico* using the TCGA HNC dataset (n=275) and validated using two independent cohorts (n=130 and 123). Immunohistochemistry was used to investigate the utility of a simplified protein signature. The spatial distribution of hypoxia and immune markers was examined using multiplex immunofluorescence staining.

Results: Unsupervised hierarchical clustering of TCGA dataset (development cohort) identified three patient subgroups with distinct hypoxia-immune phenotypes and survival profiles: hypoxia^{low}/immune^{high}, hypoxia^{high}/immune^{low} and mixed, with 5-year overall survival (OS) rates of 71%, 51% and 49% respectively (p=0.0015). The prognostic relevance of the hypoxia-immune gene signature was replicated in two independent validation cohorts. Only PD-L1 and intratumoural CD3 protein expression were associated with improved OS on multivariate analysis. Hypoxia^{low}/immune^{high} and hypoxia^{high}/immune^{low} tumours were over-represented in 'inflamed' and 'immune-desert' microenvironmental profiles respectively. Multiplex staining demonstrated an inverse correlation between CA-IX expression and prevalence of intratumoural CD3+ T cells (r=-0.5464, p=0.0377), further corroborating the transcription-based classification.

Conclusion: We developed and validated a hypoxia-immune prognostic transcriptional classifier, which may have clinical application to guide the use of hypoxia modification and targeted immunotherapies for the treatment of HNC.

Introduction

Head and neck cancer (HNC) is the sixth most common cancer worldwide, with an annual incidence of around 600,000 cases. It has a poor outcome with a 5 year mortality rate of around fifty percent (1). The main risk factors include smoking, alcohol and infection with high-risk human papilloma viruses (HPV), the latter conferring considerably better survival outcomes than the former two (2).

Whilst many prognostic biomarkers have been described (2, 3), all include a combination of clinical factors and HPV status, and currently there is no widely accepted molecular classification. Importantly, there are no treatment response classifiers. As a result, patients receive treatment based on disease stage and patient/clinician preference, rather than on the biology of the tumour.

Recently, there have been advances in our understanding of the biologic and molecular characteristics of HNC. A comprehensive multi-platform genomic characterization by The Cancer Genome Atlas (TCGA) Network confirmed four previously defined subtypes: classical, mesenchymal, basal and atypical (4-6). A more recent study identified five subgroups: HPV positive mesenchymal, non-HPV mesenchymal, HPV classical, non-HPV classical and basal (non HPV) subtypes (7). When

combined, the findings of these studies suggest three main subgroups: an inflamed/mesenchymal subtype demonstrates high expression of immune response genes and mesenchymal markers. A classical subtype (regardless of HPV status) has increased genomic aberrations associated with tobacco use. Finally, a basal subtype is highly enriched for hypoxia signalling with a lack of immune related markers. Hypoxia is known to induce immunosuppression, both directly (e.g. via inhibition of T cell proliferation and effector cytokine production) and indirectly (e.g. through metabolic competition, upregulation of co-inhibitory receptors or recruitment/conversion of immunosuppressive cell populations) (8-14). Furthermore HIF-1α signalling can be switched on by EGFR signalling (15, 16), which was also increased in some basal cancers.

Using this as a basis for a molecular classification, we then researched the availability of potential therapeutic paradigms for each of the subgroups. Hypoxic modifiers have been shown to have significant effects on survival in HNC (17, 18). More recently, evidence has emerged for the efficacy of immune checkpoint inhibitors in HNC (19, 20).

Combining these together, we sought to develop and validate a prognostic molecular classifier, based on immune response and hypoxia status, first *in silico* from fresh frozen tissues, and then using targeted RNA sequencing of formalin fixed, paraffin embedded (FFPE) tissue samples to facilitate ease of clinical application. We then evaluated whether an immunohistochemistry-based signature could substitute for, and simplify, the developed molecular signature, to further facilitate clinical implementation and cost reduction.

Materials and Methods

Patient cohorts and samples

The *in silico* discovery cohort (characteristics and treatment detailed in Tables 1 & S1 respectively: TCGA) comprised samples from the HNC dataset in the TCGA, which contains whole transcriptome gene expression data from fresh-frozen samples. Since HPV-positivity is an independent prognostic factor for HNC (2), only cases with known HPV status were considered; this yielded a cohort of 275 patients, the majority being HPV-negative.

The *in-silico* validation cohort comprised microarray gene expression data for fresh-frozen samples from 134 patients with locoregionally advanced HNC ((7) detailed in Tables 1 & S1 (Chicago)), including both HPV-positive and -negative cases.

We then evaluated FFPE diagnostic biopsy or surgical resection samples from a retrospective cohort of 163 patients with oropharyngeal cancer, recruited to the PET-NECK or Predictr clinical studies Tables 1 & S1, Correlate). A cohort of 12 patients who underwent tonsillectomy for management of a non-malignant process (snoring), recruited via the Oromouth study, served as controls. Ethical approval for use of tissue samples in translational research was granted by North West - Preston Research Ethics Committee (Reference: 16/NW/0265). p16 expression was utilised as a surrogate marker for HPV status. Immunohistochemistry (IHC) staining for p16 was performed using the CINtec Histology Kit (Roche laboratories); samples with ≥70% strong diffuse nuclear and cytoplasmic staining of tumour cells were considered positive (21).

***In silico* development and validation of an RNA signature**

For *in silico* development, expression data for genes comprising the hypoxia (Eustace (22)) and immune signatures (CIRC (23)), plus additional genes of interest (including other immune-related genes and genes frequently mutated in HNC; Table S2) were filtered from HT-seq gene count files downloaded from the TCGA HNC dataset (6), https://portal.gdc.cancer.gov/projects/TCGA-HNSC. Each dataset was normalized by dividing the expression values by the sum of expression values of the analysed genes for each sample, then logs2 followed by Z-scores were calculated for these values using R v3.3.2 https://www.r-project.org/.

Cluster analyses for genes/samples were obtained from the log2 values matrix for each dataset using Spearman distance and Ward's criterion in R. These clusters were then used to plot heatmaps using the Z-score values matrix for colour intensities. Survival was assessed by Cox regression analyses using Kaplan-Meier curves to compare the three highest hierarchical sample groups on each heatmap. Additionally, a Cox regression multivariate was used to calculate the combined effect of HPV status and heatmap groups, using the Survival package in R https://cran.r-project.org/web/packages/survival/index.html. Outcome measure was overall survival (OS) from treatment end date to death or last follow-up and censure.

For *in silico* validation, data from a previously published cohort (7) were interrogated for expression of the above gene signatures. Data for four genes *COL4A6* (Eustace); *CD80*, *IFNG* and *PDCD1LG2* (CIRC) were unavailable.

Sample size calculation

Using the data from the development TCGA cohort, to identify a difference of 25% in 3-year survival (from 75% immune^{high} to 50% hypoxia^{high} and mixed groups) with a power of 80% and a two sided alpha of 0.05, we would require 138 cases; at 85% or 90% power we would require 158 or 185 cases respectively (using the Kelsey calculation method, and OpenEpi v3.0 software).

External Validation of RNA signature using FFPE samples*Targeted RNA Sequencing*

Gene expression was quantified using Illumina TruSeq® Targeted RNA technology. RNA was extracted from 3-8 FFPE sections of 10μm thickness using the PureLink™ FFPE RNA Isolation Kit (ThermoFisher Scientific), DNAse treated and then quantified using the Qubit® RNA BR Assay Kit (ThermoFisher Scientific). Libraries were prepared using a TruSeq Targeted RNA custom panel kit (Illumina), according to the manufacturer's instructions. The panel comprised 83 genes (see Table S2 for full details). The resultant libraries were pooled and sequenced using the Illumina MiSeq platform, with a MiSeq reagent kit v3 150 cycle and paired end reads.

Data analysis

Quality assessment and filtering of the reads were carried out using FastQC v0.11.2 http://www.bioinformatics.babraham.ac.uk/projects/fastqc and PrinSeq Lite v0.20.4 http://prinseq.sourceforge.net/ respectively. Reads were aligned to appropriate reference gene sequences obtained from GenBank (see Table S2 for reference numbers), using BWA samse v0.7.15 http://bio-bwa.sourceforge.net/. The obtained SAM files were then sorted and indexed using Samtools v1.2 http://samtools.sourceforge.net/ and read count by gene files was obtained using subread featureCounts v1.5.0-p1 http://subread.sourceforge.net/. The expression matrix comprising

genes within the analysed signatures was then submitted to a similar normalization and analysis as described for the discovery dataset.

Histology and Immunohistochemistry

All staining was performed on FFPE sections of 4-5μm thickness. Haematoxylin and Eosin (H&E) staining was carried out using a Leica Autostainer xl on program 1. Automated immunohistochemistry (IHC) staining was carried out using a Leica Bond Max with standard F protocol and the following primary antibodies: Novocastra Liquid Mouse Monoclonal CA-IX Antibody clone TH22 (Leica Biosystems), 1:100 dilution; PD-L1 (E1L3N®) XP® Rabbit monoclonal antibody (Cell Signalling technology), 1:200 dilution.

For manual IHC staining, sections were deparaffinised in xylene then rehydrated in graded concentrations of alcohol. Following heat induced epitope retrieval, slides were stained using the Novolink™ Polymer Detection System (Leica Biosystems) and the following primary antibodies: FLEX ready-to-use polyclonal rabbit anti-human CD3 (Dako); mouse monoclonal anti-human LAG3 Antibody clone 17B4 (LSBio), 1:250 dilution.

Immunohistochemistry scoring

IHC slides were scored independently by two trained calibrated pathologists; when scoring was discrepant slides were re-scored and a consensus reached. H&E slides were scored for tumour infiltrating lymphocytes (TILs) as described by Ward et al (3). CD3 was semi-quantitatively estimated (range 1-4) using the mean score from three representative high power fields, where 1 = no, or sporadic CD3+ cells, 2 = moderate numbers of CD3+ cells, 3 = abundant occurrence of CD3+ cells and 4 = highly abundant occurrence of CD3+ cells; and was evaluated separately for three regions: intratumoural, peritumoural stroma and tumour advancing margin (adapted from (24)). Median values were used for survival analysis dichotomisation. Cases were also assigned an immune profile (desert, excluded or inflamed (25)) based on CD3 expression using the following cut-offs: immune desert - score of <2 in all regions; immune excluded - difference of >1 between stromal versus intratumoural score; immune inflamed - score of ≥2 in two or more regions.

PD-L1 was evaluated on both tumour cells and immune infiltrating cells (morphologically identified as lymphocytes and macrophages/dendritic cells; adapted from (26, 27)). For tumour cells, the percentage of positive cells was estimated for whole sections; four different cut-offs defining positivity were considered: ≥1%, ≥5%, ≥25% and ≥50% (28). The pattern of staining was also recorded, as either constitutive (most/all of tumour cells positive) or inductive (peripheral staining on the interface between tumour and stroma or discrete regions of staining within tumour nests). Immune cells within tumour nests or within the peritumoural stroma were semi-quantitatively scored as follows: 0 = negative, 1 = low abundance, 2 = high abundance. Tumour and inflammatory cell expression were also pooled to give a combined positive score (CPS). The presence/absence of immune cells surrounding the tumour and forming a PD-L1 positive cordon was also noted.

LAG3 expression on tumour infiltrating immune cells was scored as the mean percentage of positive cells from 10 high power fields for two regions: intra-tumoural and peri-tumoural. CA-IX was evaluated for tumour cells only; the percentage of positive cells was estimated on whole sections.

For both LAG3 and CA-IX, four different cut-offs defining positivity were considered: ≥1%, ≥5%, ≥25% and ≥50%.

Immunohistochemistry data analysis

Univariable Cox regression was used to identify immunohistochemical factors associated with OS. Follow-up was censored at 60 months in all analyses. Variables were analysed as linear continuous measures and dichotomised according to pre-stated thresholds. Survival by category for each marker is displayed in Kaplan-Meier plots. Multivariable Cox regression models assessed whether prognostic value was altered with adjustments for important clinical factors (age, gender, T stage, N stage, p16 status, smoking status and alcohol consumption). Analysis was conducted as complete case (123 of 161 patients), and using multiple imputation with chained equations (24 imputations) to impute missing data. All analysis was performed using Stata 15.

Multiplex staining

Multi-colour immunofluorescent staining was carried out on the Leica Bond Rx using an Opal™ 7-colour fIHC kit and the following primary antibody fluorochrome combinations: CD3/Opal 520, CA-IX/Opal 570, PD-L1/Opal 650 and CK (monoclonal mouse anti-human cytokeratin (concentrate) antibody clone AE1/AE3)/Opal 650. Staining was imaged with the Vectra Automated Quantitative Pathology Imaging System, and 7-11 representative fields per case quantitatively analysed using Inform Advanced Image Analysis software v 2.3 (Perkin Elmer).

Results

Development of prognostic classifier

TCGA data from 275 HNC cases (detailed in Tables 1 & S1) were interrogated for expression of hypoxia- and immune response-related genes using a combination of two previously developed signatures: a hypoxia signature comprised of 26 genes (Eustace (22)) and an immune response signature containing 28 genes (CIRC (23)); individual genes are detailed in Table S2. Bioinformatics analysis, using unsupervised two dimensional hierarchical clustering, identified distinct HNC patient subgroups with co-ordinate expression of hypoxia- and immune response-related genes (Fig 1A). Three subgroups were classified as follows: (i) hypoxia^{high}/immune^{low}, comprised almost exclusively of HPV-negative cases, (ii) hypoxia^{low}/immune^{high}, enriched for HPV-positive cases, and (iii) a mixed subgroup, containing both HPV-positive and negative cases (Fig. 1A; gene ID is provided in Fig. S1). Kaplan-Meier analysis revealed that the three subgroups have different survival profiles, with the hypoxia^{low}/immune^{high} subgroup having the best prognosis, corresponding to a 5-year OS of 71.4% (Fig. 1B). The hypoxia^{high}/immune^{low} and mixed subgroups have similar profiles, with 5-year OS figures of 50.8% and 48.6% respectively (p=0.0015, HPV-adjusted p=0.0005; adjustment for other standard prognostic variables is considered in Table S3). Of note, the hypoxia and immune gene signatures were not independently prognostic (Fig. S2). Furthermore, consideration of HPV-negative cases alone, within the predefined hypoxia-immune subgroups, maintained comparable survival profiles (p=0.001).

Validation of prognostic classifier

We then sought to validate the hypoxia-immune signature, firstly *in silico* using data for 134 HNC cases comprising the previously published Chicago cohort ((7); detailed in Tables 1 & S1). Bioinformatics analysis identified comparable patient subgroups to those described above

(hypoxia^{high}/immune^{low}, hypoxia^{low}/immune^{high} and mixed; Fig. 1C). As previously, the hypoxia^{high}/immune^{low} subgroup was predominantly HPV-negative. Although HPV-positive cases were again over-represented in the hypoxia^{low}/immune^{high} subgroup, approximately one third of cases in this group were HPV-negative. Likewise, the mixed class comprised both HPV-positive and negative cases, suggesting that the hypoxia-immune signature classifies beyond HPV status. Correlation of gene expression profiles with survival confirmed that subgroups have distinct outcomes. Here again, the hypoxia^{low}/immune^{high} subgroup have the best prognosis, 5-year OS 76.2%, whereas the hypoxia^{high}/immune^{low} and mixed classes have inferior outcomes with 5-year OS of 65.1% and 60.0% respectively, (p=0.012, HPV adjusted p=0.0025).

The second validation cohort comprised 163 cases of oropharyngeal squamous cell carcinoma (OPSCC); consistent with their anatomic location a higher percentage of cases were HPV-positive (Tables 1 & S1: Correlate; flow of patient samples through study is summarised in Fig. S3). FFPE tissue samples were available from this cohort enabling comparative gene and protein expression analysis. Gene expression was quantified using Illumina TruSeq® Targeted RNA technology. Quality control measures, based on a minimum number of mapped reads of 70,000, reduced the number of informative cases to 123; the heat map for these and the 12 normal control samples is shown in Fig. 1E. As previously, hierarchical clustering identified three patient subgroups, two with co-ordinate expression of hypoxia and immune signature genes. Normal controls clustered with the hypoxia^{low}/immune^{high} subgroup, where p16-positive cases were also over-represented. Conversely, the small hypoxia^{high}/immune^{low} subgroup was comprised of a majority of p16-negative cases. The mixed subgroup again included both HPV-positive and -negative cases, supporting the proposal that the hypoxia-immune signature classifies beyond HPV status (Fig. 1E). Despite the smaller sub-group size, Kaplan-Meier analysis revealed that the three subgroups have distinct outcome profiles, mirroring the TCGA results (Fig. 1F), with 5-year OS rates of 79.5%, 64.4% and 55% for the hypoxia^{low}/immune^{high}, mixed and hypoxia^{high}/immune^{low} subgroups respectively (p=0.097, HPV-adjusted p=0.0006).

Hypoxia may result from rapid tumour growth with inadequate neovascularisation. To address the possibility that hypoxia-immune subgroup classification was significantly influenced by tumour volume the distribution of T stage within the three subgroups was examined. A correlation with T stage was not consistently observed (Fig. S4).

Distribution of individual hypoxia- and immune-related gene expression within the three subgroups

The analyses described above identified three distinct HNC patient subgroups with co-ordinate expression of hypoxia- and immune-related gene signatures. However, within each group individual genes may differ in their expression and contribution to the endotype. Subsequent exploratory analyses therefore mapped expression of individual genes within the three subgroups in an attempt to understand their distinct contributions to each endotype. This analysis focused on the TCGA dataset where whole transcriptome data is available, enabling interrogation of genes of interest not included in the Chicago microarray or Correlate targeted sequencing datasets. As illustrated in Figs. 2 and S5A, the pattern of expression for individual hypoxia-related genes was relatively broad with considerable overlap between subgroups, particularly the hypoxia^{high}/immune^{low} and mixed classes. However, a subset of genes showed greater differential expression (hypoxia^{high}/immune^{low} > mixed >

hypoxia^{low}/immune^{high}) including *CA9*, *SLC2A1* (GLUT1) and *SLC16A1* (MCT1), the products of which are involved in pH regulation and glucose metabolism (Fig. 2 top panels and S5A). Twenty three out of the 26 genes represented in the hypoxia signature showed significant differential expression between the hypoxia^{low}/immune^{high} and hypoxia^{high}/immune^{low} or mixed subgroups.

Immune response-related genes described by the CIRC signature were all differentially expressed, albeit to varying extents, according to the expected order: hypoxia^{low}/immune^{high} > mixed > hypoxia^{high}/immune^{low} (Fig. 2 and S5B). The majority of immune genes of interest outside of the CIRC signature also demonstrated a similar pattern of differential expression (Figs. 2 and S5C). This was true of genes whose products would be expected to mediate anti- or pro-tumoural roles; for example, markers of cytotoxic and regulatory T cell function. Genes associated with the myeloid lineage were also expressed at significantly higher levels within the hypoxia^{low}/immune^{high} subgroup; although markers of effector/suppressor function (e.g. *ARG1*, *NOS2* (iNOS)) were not uniformly co-expressed. Other genes of particular interest included those encoding activating receptors (e.g. CD40, CD137, OX40) or inhibitory receptors (e.g. CTLA-4, LAG3, PD-1) the ectonucleotidases CD39/CD73 and indoleamine 2,3-dioxygenase 1 (IDO); as therapies targeting these markers are currently in clinical practice, trials or in development. IDO is one of six genes comprising an 'Interferon-γ (IFNγ) signature' that has shown predictive value for response to anti-PD-1 monoclonal antibody therapy in HNC and other cancers (19). The other five genes - all part of the CIRC signature - display a similar expression pattern: hypoxia^{low}/immune^{high} >> mixed > hypoxia^{high}/immune^{low}; these include IFNγ itself and the chemokines CXCL9/CXCL10 which may be important for T cell homing to tumours. Finally, EGFR was significantly over-expressed in the hypoxia^{high}/immune^{low} and mixed subgroups relative to the hypoxia^{low}/immune^{high} group.

Protein expression classifier

Whilst gene signatures potentially better reflect the complex interactions between immune cells and the hypoxic tumour microenvironment, simplified IHC signatures may be more clinically applicable. Furthermore, gene expression may not correlate with end protein expression due to post-transcriptional or translational control and protein degradation. Finally, gene signatures neither discriminate between expression within tumour or stromal compartments, nor identify cell type-specific expression. Therefore, we next examined whether an IHC signature comprising TILs, CD3, PD-L1 and LAG3 as immune-response markers and CA-IX as a marker of hypoxia could stratify patients into the corresponding subgroups with different prognoses.

Results for individual markers are summarised in Table S4; representative images showing the different patterns of expression and their association with OS are shown in Figs. 3 and S6 with the corresponding statistical analyses in Table S5. In agreement with previous studies, high TIL status (assessed on H&E staining) correlated with p16 status (3), and patients with the highest TIL numbers had a significantly better OS (hazard ratio (HR) 0.27, 95% Confidence interval (CI) 0.12-0.66, p = 0.004 on univariate analysis). Similarly, high numbers of T cells (measured by CD3 staining) correlated with improved OS (combined CD3: HR 0.39 (0.2-0.77), p = 0.007 on univariate analysis). Since the precise location of T cells within the tumour microenvironment may impact on their prognostic significance (29), T cell frequencies were independently evaluated for three locations: within tumour nests, within the peri-tumoural stroma and at the tumour margin. An elevated frequency of intratumoural T cells showed the greatest prognostic value (HR 0.33 (0.16-0.67), p =

0.002) and this was independent of clinical factors and p16 status (adjusted HR (aHR) 0.43 (0.19-0.97), p = 0.043) (Fig. 3 and Table S5).

The prognostic impact of PD-L1 expression was dependent on both the cell types considered and the cut-off selected (Figs. 3C, 3H and S6). Higher expression on tumour cells correlated with improved survival using one of the four tested cut-offs (25%; aHR 0.33 (0.13-0.89), p = 0.025). Similarly, a trend for improved survival was observed for higher expression on intra-/peritumoural immune cells (HR 0.45 (0.19 – 1.05), p = 0.066). In addition, the presence of a PD-L1 positive immune cordon around tumour nests was associated with better OS (aHR 0.19 (0.05-0.68), p = 0.011). Using a CPS, which considers PD-L1 expression on both tumour and inflammatory cells, strengthened prediction of OS (Table S5). Thus CPS based on ≥1%, ≥5% and ≥25% cut-offs all predicted OS on multivariate analysis; only the ≥50% cut-off did not.

Results for the checkpoint receptor LAG3 expressed on inflammatory cells showed a trend for improved survival with higher expression (Fig. 3D/I), particularly when these cells were located within tumour nests (25% cut-off: HR 0.34 (0.12-0.96), p = 0.042). However, these were not significant when adjusted for other tumour factors.

Two thirds of tumours contained hypoxic regions as measured by CA-IX expression (≥10% cut-off). There are conflicting reports in the literature regarding the prognostic value of CA-IX expression in HNC (reviewed in (30). Here, CA-IX expression was not prognostic in the complete case analysis (Fig. 3J), although there was a trend for improved survival in the subset of patients with very low or absent expression when multiple imputation analysis was used (HR 3.89 (0.95 to 16.00), p=0.060; adjusted HR 9.02 (1.96-41.59), p = 0.005; Fig. S6).

Since both RNASeq and IHC data were available for the majority of samples from the Correlate cohort, we compared normalised gene versus protein expression levels and found a significant positive correlation for all examined hypoxia or immune markers (Fig. S7). Subsequent analyses evaluated levels of IHC marker expression within the three heat map gene cluster-defined subgroups (hypoxia^{high}/immune^{low}, hypoxia^{low}/immune^{high} and mixed). All immune-related markers showed significantly higher levels of protein expression in the hypoxia^{low}/immune^{high} versus hypoxia^{high}/immune^{low} subgroups; expression levels for the mixed subgroup were intermediate. As expected, CA-IX protein expression was significantly decreased in the hypoxia^{low}/immune^{high} versus hypoxia^{high}/immune^{low} subgroups and intermediate in the mixed group (Fig. 4A).

Immune profile

It has been suggested that cancers, including HNC, can be assigned one of three immune profiles, based on the frequency and distribution of T cells within the tumour and microenvironment (25). Thus 'immune deserts' are characterised by minimal T cell presence; 'immune-excluded' tumours are surrounded by abundant peritumoural T cells, but permit limited intra-tumoural penetration and 'inflamed tumours' contain relatively high numbers of T cells in both compartments – here T cell function is apparently compromised. To explore a possible relationship between tumour immune profile and the three gene cluster-defined subgroups as above, all cases within the Correlate cohort were assigned an immune phenotype (desert/excluded/inflamed) based on the magnitude and pattern of CD3 staining. As illustrated in Fig. 4B, hypoxia^{high}/immune^{low} tumours were over-

represented in the 'immune desert' category and significantly reduced in the inflamed subgroup. The reverse pattern was observed for hypoxia^{low}/immune^{high} tumours; whilst tumours classified as mixed based on hypoxia-immune gene expression were more evenly distributed across all three immune profiles.

Spatial interactions of immune cells and hypoxia in the tumour microenvironment

In the final series of experiments, we examined the spatial distribution of different cell populations within the tumour microenvironment, in particular the relationship between regions of tumour hypoxia (as measured by CA-IX expression), immune cell localisation and PD-L1 expression. Fifteen cases, representative of the three hypoxia-immune subgroups identified in the gene expression analyses, were stained using Opal Multiplex IHC assays. Representative images are shown in Figs. 5 and S8. Staining for cytokeratin (as a marker for tumour cells) confirmed that CA-IX expression is restricted to tumour cells (Fig. S8A). A subset of tumours with high levels of CA-IX expression displayed an 'immune desert' phenotype, with few T cells present in either the stromal or tumoural compartments (Fig. 5A hypoxia^{high}/immune^{low}). In other cases abundant T cells were observed within the peritumoural stroma, but minimal numbers appeared to have infiltrated tumour nests, consistent with an 'immune excluded' phenotype (Fig. 5B mixed and Figs. S8B & C). The final category comprised 'inflamed' tumours with abundant T cells present in the peritumoural stroma and significant intratumoural infiltration, particularly in the absence of hypoxia as indicated by no/minimal CA-IX expression (Fig. 5C hypoxia^{low}/immune^{high}). Overall, we observed a significant inverse correlation between CA-IX positivity and the prevalence of CD3^{+} T cells within tumour nests (Fig. 5D left panel; r=-0.5464, p=0.0377,) but not within the stromal compartment alone (Fig. 5D middle panel, r=-0.2393, p=0.3982) or the combined tumour environment (Fig.5D right panel, r=-0.4571, p=0.0889), consistent with hypoxia-mediated inhibition of T cell migration and/or survival within tumour nests. No significant correlation was observed between tumoural expression of CA-IX and PD-L1 (Fig. S8D), which is unsurprising given that hypoxia (via HIF1α) is only one of multiple potential pathways leading to PD-L1 expression. Excepting 'constitutive' cases, strong tumoural PD-L1 expression was restricted to cells at the stromal interface, found in close proximity to CD3-positive T cells (Fig. 5B and C; yellow triangles). The positive correlation between stromal PD-L1 and CD3 T cells (Fig.5E) is consistent with a dominant role for IFNγ-mediated expression.

Discussion

Recent studies have advanced our understanding of the biological and molecular characteristics of HNC (4-7). However, there remains an unmet need for clinically relevant classifications to guide therapy selection. Key features associated with the biologically distinct subtypes include immune landscape, expression of EGFR/HER and hypoxia (7, 31). Given the strong evidential basis for efficacy of hypoxia modification (reviewed in (18)) and immune checkpoint inhibitors (19, 20) in the treatment of HNC, we explored a combination of hypoxia and immune status classifiers. Use of the hypoxia-immune signature in an unsupervised cross-cohort manner identified three distinct HNC subtypes. Subgroup classification correlated with survival, with the hypoxia^{low}/immune^{high} subgroup having the best prognosis. The hypoxia^{high}/immune^{low} and mixed subgroups have similar survival profiles; however, they have distinct transcriptional patterns and immune profiles, consistent with activated pathways that could be targeted by different therapeutic interventions, hence the importance of separating them.

Given the widespread use of IHC for diagnostic purposes in routine pathology practice, we evaluated whether a simplified IHC-based signature could substitute for the developed hypoxia-immune gene expression signature. Expression of individual immune response markers notably intratumoural CD3-positive T cells and PD-L1 expression (assessed using a CPS) correlated with superior outcome, but did not improve the prognostic value of clinical factors alone. CA-IX, as a marker of hypoxia, did not have prognostic power.

Finally, consistent with hypoxia-driven immune suppression, multiplex IHC staining identified an inverse relationship between CA-IX and infiltration of CD3 T cells into tumour regions. The observed spatial distribution of hypoxia and immune response markers within the tumour microenvironment correlated strongly with our classification based on gene expression, confirming its validity.

An inverse correlation between hypoxia-related gene expression and anti-tumoural immune responses is consistent with previous studies in HNC (7, 31). As mentioned above, hypoxia drives immune suppression via multiple mechanisms. It may be achieved through HIF1α-mediated activation of inhibitory pathways, including CD47 (32), the adenosine-generating enzymes CD73 and CD39 (33) and immune checkpoint receptors such as PD-L1 (11, 12). Perhaps paradoxically, many inhibitory markers are more highly expressed in 'inflamed' tumours (hypoxia^{low}/immune^{high}), where they may be indicative of an ongoing 'exhausted' anti-tumour immune response (34), with the potential for reactivation via immune checkpoint blockade. Tumour cell metabolic adaptations to hypoxia, such as increased glucose uptake and lactate production, also act to promote and perpetuate an immunosuppressive microenvironment (9, 10). Due to differences in their metabolic programmes these factors would be expected to have greater impact on effector versus regulatory T cells (35). Alternatively, since most hypoxic tumour areas contained minimal T cells (immune desert or excluded profile), immune suppression may be mediated mainly by mechanisms that inhibit T cell migration into tumours (e.g. VEGF, CXCL12) or promote T cell death (e.g. FasL) (reviewed in (36)).

As described here and reported previously (31), 'inflamed' (hypoxia^{low}/immune^{high}) tumours demonstrate robust myeloid gene expression profiles. The balance of anti-tumoural versus immunosuppressive function of myeloid cells - including myeloid derived suppressor cells (MDSCs) and tumour associated macrophages (TAMs) - may be modulated by tumour hypoxia, which favours suppression (8, 11, 13). Mechanistically, hypoxia (predominantly via HIF signalling) can tip the balance of cytokine/chemokine/effector molecule production towards those with pro-tumoural functions (e.g. IL-10, VEGF, arginase, iNOS) and also upregulate expression of T cell inhibitory ligands (e.g. PD-L1).

Hypoxia can upregulate EGFR expression (37, 38) and promote ligand-independent EGFR signalling (39), both mechanisms potentially augmenting tumour glycolytic metabolism and consequent metabolic competition. EGFR overexpression is observed in a high percentage of HNC, particularly HPV-negatives cases, and is associated with poor prognosis; other mechanisms of upregulation include mutations, gene duplications and protein stabilisation (reviewed in (40)). As previously mentioned, EGFR signalling may lead to hypoxia-independent stabilisation of HIF1α and consequent upregulation of glycolytic metabolism (15, 16); providing an alternative (not mutually exclusive) explanation for low intra-tumoural T cell infiltration (31). However, it is noteworthy that hypoxic

modification appears most beneficial for patients with HPV-negative tumours (41, 42), consistent with an important role for hypoxia-mediated HIF1α upregulation. Furthermore, it has been reported that oropharyngeal tumours, especially HPV-positive tumours are less hypoxic (43).

T cell infiltration of tumours (especially by CD8+ cytotoxic T lymphocytes) has been associated with a favourable prognosis in several tumour types, including HNC (3, 24, 29, 44). Likewise, PD-L1 expression, particularly within immune cells in the tumour microenvironment, can have both prognostic and predictive importance (26, 27). In our study, higher numbers of TILs and $CD3^+$ T cells - particularly within tumour nests - correlated with improved survival. PD-L1 expression was also linked with better prognosis (dependent on cut-off), especially when both tumour and immune cell expression were considered as a CPS. The differential prognostic impact of intermediate (CPS>1% and <50%) versus high (≥50%) PD-L1 expression might reflect alternative mechanisms controlling expression. Intermediate expression (found on immune cells and tumour cells at the stromal interface) may be indicative of adaptive immune resistance in the face of an ongoing anti-tumoural immune response (45, 46); whereas high levels are associated with tumour intrinsic expression (e.g. loss of PTEN or EGFR activation (47, 48)).

Previous studies in diverse tumour settings have correlated response to immune checkpoint blockade (particularly involving the PD-1/PD-L1 axis) with 'cancer-immune phenotype' (25).Thus 'inflamed' tumours exhibit higher response rates compared with those having 'immune desert' or 'immune excluded' profiles, although even in the former cases a response is not assured. Our combined analysis of gene and protein expression in HNC suggests that hypoxia may be an important factor distinguishing hot (inflamed) and cold (immune desert or excluded) tumours. In agreement with this, it has recently been reported that targeted hypoxia reduction restores intratumoural T cell infiltration in a mouse model of prostate cancer (49).

We recognise certain limitations in our study, including the potential for bias due to different tumour:stroma ratios within samples; for example, TCGA samples comprise >80% tumour cells, therefore reducing analysis of the stromal compartment. Minor inter-cohort variations in gene clustering and prognostic value of the classifier (as measured by 5-year OS) may reflect (i) different RNA quantification platforms (microarray versus whole transcriptome sequencing versus targeted sequencing), (ii) fresh/frozen versus FFPE tissue, (iii) differences in case mix (anatomic site, TNM/clinical stage, smoking, alcohol consumption, HPV status) or treatment. Of note, the consistent performance of the classifier despite the heterogeneous nature of the study cohorts and assays is an indication of the strength of the signature Currently, subgroup classification is based on gene clustering (three highest order clusters) rather than defined cut-offs. For clinical application, future work will need to define appropriate cut-offs, for example using median values (22) and/or develop a clinically-applicable continuous scoring system.

The identification of our hypoxia-immune prognostic classifier for HNC suggests that differential treatment approaches might be required for patient subgroups. Although hypoxic modification (nimorazole) and immune checkpoint therapies (PD-1/PD-L1 inhibitors) have shown single agent activity in HNC, the response rate is relatively low; for example, only 13-18% for PD-1 in the recurrent/metastatic setting (19, 20). Our data suggest that combinations or sequential use of hypoxia-modifying and/or immunomodulatory drugs may be beneficial. For example, in patients

with a hypoxia^{high}/immune^{low} or mixed signature correlating with poor prognosis, treatment with hypoxia modifiers may sensitize to chemo/radiotherapy, and EGFR inhibitors or agents targeting MDSCs/TAMs may reverse the immunosuppressive environment. Sequential treatment with immune checkpoint inhibitors may then be required to prevent inhibition or dampening of the emerging immune response. Cases that are hypoxia^{high}/immune^{high} might benefit from concurrent treatment with hypoxia modifiers and immune oncology agents. Importantly, our data on expression of individual genes also indicate potential for immunotherapy treatment strategies beyond PD-L1, for example combinations of PD-L1 and anti-LAG3 treatments for the hypoxia^{low}/immune^{high} subgroup.

In conclusion, we developed and validated a prognostic molecular classifier based on hypoxia and immune status. This classifier may have clinical application to guide the use of hypoxia modification and targeted immunotherapies such as checkpoint inhibitors. We would recommend validation of its prognostic value and assessment of any potential predictive power (e.g. using treatment with nimorazole, EGFR inhibitors or anti-PD-1/PD-L1) in a prospective setting.

References

1. Ferlay J, Soerjomataram I, Dikshit R, Eser S, Mathers C, Rebelo M, et al. Cancer incidence and mortality worldwide: sources, methods and major patterns in GLOBOCAN 2012. Int J Cancer. 2015;136(5):E359-86.
2. Ang KK, Harris J, Wheeler R, Weber R, Rosenthal DI, Nguyen-Tan PF, et al. Human papillomavirus and survival of patients with oropharyngeal cancer. N Engl J Med. 2010;363(1):24-35.
3. Ward MJ, Thirdborough SM, Mellows T, Riley C, Harris S, Suchak K, et al. Tumour-infiltrating lymphocytes predict for outcome in HPV-positive oropharyngeal cancer. Br J Cancer. 2014;110(2):489-500.
4. Chung CH, Parker JS, Karaca G, Wu J, Funkhouser WK, Moore D, et al. Molecular classification of head and neck squamous cell carcinomas using patterns of gene expression. Cancer Cell. 2004;5(5):489-500.
5. Walter V, Yin X, Wilkerson MD, Cabanski CR, Zhao N, Du Y, et al. Molecular subtypes in head and neck cancer exhibit distinct patterns of chromosomal gain and loss of canonical cancer genes. PLoS One. 2013;8(2):e56823.
6. Cancer Genome Atlas N. Comprehensive genomic characterization of head and neck squamous cell carcinomas. Nature. 2015;517(7536):576-82.
7. Keck MK, Zuo Z, Khattri A, Stricker TP, Brown CD, Imanguli M, et al. Integrative analysis of head and neck cancer identifies two biologically distinct HPV and three non-HPV subtypes. Clin Cancer Res. 2015;21(4):870-81.
8. Chouaib S, Noman MZ, Kosmatopoulos K, Curran MA. Hypoxic stress: obstacles and opportunities for innovative immunotherapy of cancer. Oncogene. 2017;36(4):439-45.
9. Ho PC, Bihuniak JD, Macintyre AN, Staron M, Liu X, Amezquita R, et al. Phosphoenolpyruvate Is a Metabolic Checkpoint of Anti-tumor T Cell Responses. Cell. 2015;162(6):1217-28.
10. Chang CH, Qiu J, O'Sullivan D, Buck MD, Noguchi T, Curtis JD, et al. Metabolic Competition in the Tumor Microenvironment Is a Driver of Cancer Progression. Cell. 2015;162(6):1229-41.
11. Noman MZ, Desantis G, Janji B, Hasmim M, Karray S, Dessen P, et al. PD-L1 is a novel direct target of HIF-1alpha, and its blockade under hypoxia enhanced MDSC-mediated T cell activation. J Exp Med. 2014;211(5):781-90.
12. Barsoum IB, Smallwood CA, Siemens DR, Graham CH. A mechanism of hypoxia-mediated escape from adaptive immunity in cancer cells. Cancer Res. 2014;74(3):665-74.

13. Corzo CA, Condamine T, Lu L, Cotter MJ, Youn JI, Cheng P, et al. HIF-1alpha regulates function and differentiation of myeloid-derived suppressor cells in the tumor microenvironment. J Exp Med. 2010;207(11):2439-53.
14. Facciabene A, Peng X, Hagemann IS, Balint K, Barchetti A, Wang LP, et al. Tumour hypoxia promotes tolerance and angiogenesis via CCL28 and T(reg) cells. Nature. 2011;475(7355):226-30.
15. Phillips RJ, Mestas J, Gharaee-Kermani M, Burdick MD, Sica A, Belperio JA, et al. Epidermal growth factor and hypoxia-induced expression of CXC chemokine receptor 4 on non-small cell lung cancer cells is regulated by the phosphatidylinositol 3-kinase/PTEN/AKT/mammalian target of rapamycin signaling pathway and activation of hypoxia inducible factor-1alpha. J Biol Chem. 2005;280(23):22473-81.
16. Peng XH, Karna P, Cao Z, Jiang BH, Zhou M, Yang L. Cross-talk between epidermal growth factor receptor and hypoxia-inducible factor-1alpha signal pathways increases resistance to apoptosis by up-regulating survivin gene expression. J Biol Chem. 2006;281(36):25903-14.
17. Overgaard J, Hansen HS, Overgaard M, Bastholt L, Berthelsen A, Specht L, 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):135-46.
18. 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(1):22-32.
19. Seiwert TY, Burtness B, Mehra R, Weiss J, Berger R, Eder JP, et al. Safety and clinical activity of pembrolizumab for treatment of recurrent or metastatic squamous cell carcinoma of the head and neck (KEYNOTE-012): an open-label, multicentre, phase 1b trial. Lancet Oncol. 2016;17(7):956-65.
20. Ferris RL, Blumenschein G, Fayette J, Guigay J, Colevas AD, Licitra L, et al. Nivolumab for Recurrent Squamous-Cell Carcinoma of the Head and Neck. New England Journal of Medicine. 2016;375(19):1856-67.
21. Jordan RC, Lingen MW, Perez-Ordonez B, He X, Pickard R, Koluder M, et al. Validation of methods for oropharyngeal cancer HPV status determination in US cooperative group trials. Am J Surg Pathol. 2012;36(7):945-54.
22. Eustace A, Mani N, Span PN, Irlam JJ, Taylor J, Betts GN, 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(17):4879-88.
23. Lal N, Beggs AD, Willcox BE, Middleton GW. An immunogenomic stratification of colorectal cancer: Implications for development of targeted immunotherapy. Oncoimmunology. 2015;4(3):e976052.
24. Dahlin AM, Henriksson ML, Van Guelpen B, Stenling R, Oberg A, Rutegard J, et al. Colorectal cancer prognosis depends on T-cell infiltration and molecular characteristics of the tumor. Mod Pathol. 2011;24(5):671-82.
25. Chen DS, Mellman I. Elements of cancer immunity and the cancer-immune set point. Nature. 2017;541(7637):321-30.
26. Herbst RS, Soria JC, Kowanetz M, Fine GD, Hamid O, Gordon MS, et al. Predictive correlates of response to the anti-PD-L1 antibody MPDL3280A in cancer patients. Nature. 2014;515(7528):563-7.
27. Kim HR, Ha SJ, Hong MH, Heo SJ, Koh YW, Choi EC, et al. PD-L1 expression on immune cells, but not on tumor cells, is a favorable prognostic factor for head and neck cancer patients. Sci Rep. 2016;6:36956.
28. De Meulenaere A, Vermassen T, Aspeslagh S, Huvenne W, Van Dorpe J, Ferdinande L, et al. Turning the tide: Clinical utility of PD-L1 expression in squamous cell carcinoma of the head and neck. Oral Oncol. 2017;70:34-42.

29. Galon J, Costes A, Sanchez-Cabo F, Kirilovsky A, Mlecnik B, Lagorce-Pages C, et al. Type, density, and location of immune cells within human colorectal tumors predict clinical outcome. Science. 2006;313(5795):1960-4.
30. Swartz JE, Pothen AJ, Stegeman I, Willems SM, Grolman W. Clinical implications of hypoxia biomarker expression in head and neck squamous cell carcinoma: a systematic review. Cancer Med. 2015;4(7):1101-16.
31. Ottensmeier CH, Perry KL, Harden EL, Stasakova J, Jenei V, Fleming J, et al. Upregulated Glucose Metabolism Correlates Inversely with CD8+ T-cell Infiltration and Survival in Squamous Cell Carcinoma. Cancer Res. 2016;76(14):4136-48.
32. Zhang H, Lu H, Xiang L, Bullen JW, Zhang C, Samanta D, et al. HIF-1 regulates CD47 expression in breast cancer cells to promote evasion of phagocytosis and maintenance of cancer stem cells. Proc Natl Acad Sci U S A. 2015;112(45):E6215-23.
33. Allard B, Longhi MS, Robson SC, Stagg J. The ectonucleotidases CD39 and CD73: Novel checkpoint inhibitor targets. Immunol Rev. 2017;276(1):121-44.
34. Pauken KE, Wherry EJ. Overcoming T cell exhaustion in infection and cancer. Trends Immunol. 2015;36(4):265-76.
35. Siska PJ, Rathmell JC. T cell metabolic fitness in antitumor immunity. Trends Immunol. 2015;36(4):257-64.
36. Joyce JA, Fearon DT. T cell exclusion, immune privilege, and the tumor microenvironment. Science. 2015;348(6230):74-80.
37. Franovic A, Gunaratnam L, Smith K, Robert I, Patten D, Lee S. Translational up-regulation of the EGFR by tumor hypoxia provides a nonmutational explanation for its overexpression in human cancer. Proc Natl Acad Sci U S A. 2007;104(32):13092-7.
38. Misra A, Pandey C, Sze SK, Thanabalu T. Hypoxia activated EGFR signaling induces epithelial to mesenchymal transition (EMT). PLoS One. 2012;7(11):e49766.
39. Wang Y, Roche O, Xu C, Moriyama EH, Heir P, Chung J, et al. Hypoxia promotes ligand-independent EGF receptor signaling via hypoxia-inducible factor-mediated upregulation of caveolin-1. Proc Natl Acad Sci U S A. 2012;109(13):4892-7.
40. Mirghani H, Amen F, Moreau F, Guigay J, Hartl DM, Lacau St Guily J. Oropharyngeal cancers: relationship between epidermal growth factor receptor alterations and human papillomavirus status. Eur J Cancer. 2014;50(6):1100-11.
41. Lassen P, Eriksen JG, Hamilton-Dutoit S, Tramm T, Alsner J, Overgaard J, et al. HPV-associated p16-expression and response to hypoxic modification of radiotherapy in head and neck cancer. Radiother Oncol. 2010;94(1):30-5.
42. Rischin D, Young RJ, Fisher R, Fox SB, Le QT, Peters LJ, et al. Prognostic significance of p16INK4A and human papillomavirus in patients with oropharyngeal cancer treated on TROG 02.02 phase III trial. J Clin Oncol. 2010;28(27):4142-8.
43. Tawk B, Schwager C, Deffaa O, Dyckhoff G, Warta R, Linge A, et al. Comparative analysis of transcriptomics based hypoxia signatures in head- and neck squamous cell carcinoma. Radiother Oncol. 2016;118(2):350-8.
44. Balermpas P, Michel Y, Wagenblast J, Seitz O, Weiss C, Rodel F, et al. Tumour-infiltrating lymphocytes predict response to definitive chemoradiotherapy in head and neck cancer. Br J Cancer. 2014;110(2):501-9.
45. Tumeh PC, Harview CL, Yearley JH, Shintaku IP, Taylor EJ, Robert L, et al. PD-1 blockade induces responses by inhibiting adaptive immune resistance. Nature. 2014;515(7528):568-71.
46. Lyford-Pike S, Peng SW, Young GD, Taube JM, Westra WH, Akpeng B, et al. Evidence for a Role of the PD-1:PD-L1 Pathway in Immune Resistance of HPV-Associated Head and Neck Squamous Cell Carcinoma. Cancer Research. 2013;73(6):1733-41.
47. Parsa AT, Waldron JS, Panner A, Crane CA, Parney IF, Barry JJ, et al. Loss of tumor suppressor PTEN function increases B7-H1 expression and immunoresistance in glioma. Nat Med. 2007;13(1):84-8.

48. Concha-Benavente F, Srivastava RM, Trivedi S, Lei Y, Chandran U, Seethala RR, et al. Identification of the Cell-Intrinsic and -Extrinsic Pathways Downstream of EGFR and IFNgamma That Induce PD-L1 Expression in Head and Neck Cancer. Cancer Res. 2016;76(5):1031-43.
49. Jayaprakash P, Ai M, Liu A, Budhani P, Bartkowiak T, Sheng J, et al. Targeted hypoxia reduction restores T cell infiltration and sensitizes prostate cancer to immunotherapy. J Clin Invest. 2018;128(11):5137-49.
50. Moons KG, Altman DG, Reitsma JB, Ioannidis JP, Macaskill P, Steyerberg EW, et al. Transparent Reporting of a multivariable prediction model for Individual Prognosis or Diagnosis (TRIPOD): explanation and elaboration. Ann Intern Med. 2015;162(1):W1-73.

Table 1: Baseline characteristics of patients in the discovery and validation cohorts

	TCGA cohort n = 275 Number (%)	Chicago cohort n = 134 Number (%)[#]	Correlate cohort n = 163 Number (%)
Age at diagnosis (years)			
median (range)	62 (19 – 90)	57 (33 – 81)	57 (35-84)
Gender			
Male	199 (72.4)	107 (82.3)	116 (71.2)
Female	76 (27.6)	23 (17.7)	46 (28.2)
Not known	0 (0.0)	0 (0.0)	1 (0.6)
Tobacco use			
Never	51 (18.6)	19 (14.6)	36 (22.1)
Light	34 (12.4)	35 (26.9)	29 (17.8)
Heavy	183 (66.5)	73 (56.2)	85 (52.1)
Not known	7 (2.5)	3 (2.3)	13 (8.0)
Alcohol consumption			
Never	85 (30.9)	15 (11.5)	19 (11.7)
Light	184 (66.9)	54 (41.5)	85 (52.1)
Heavy		60 (46.2)	37 (22.7)
Not known	6 (2.2)	1 (0.8)	22 (13.5)
Anatomic site			
Oral cavity	169 (61.5)	25 (19.2)	1 (0.6)
Oropharynx	32 (11.6)	73 (56.1)	162 (99.4)
Hypopharynx	2 (0.7)	8 (6.2)	0 (0.0)
Larynx	71 (25.8)	24 (18.5)	0 (0.0)
Other	1 (0.4)	0 (0.0)	0 (0.0)
T-stage (TNM7)			
T1	13 (4.7)	8 (6.2)	33 (20.2)
T2	84 (30.6)	37 (28.4)	59 (36.2)
T3	82 (29.8)	29 (22.3)	35 (21.5)
T4	96 (34.9)	55 (42.3)	28 (17.2)
Not known	0 (0.0)	1 (0.8)	8 (4.9)
N-stage (TNM7)			
N0	136 (49.4)	12 (9.2)	17 (10.4)
N1	49 (17.8)	10 (7.7)	17 (10.4)
N2	81 (29.5)	92 (70.7)	107 (65.6)
N3	6 (2.2)	14 (10.8)	11 (6.8)
N4	0 (0.0)	1 (0.8)	0 (0.0)
Not known	3 (1.1)	1 (0.8)	11 (6.8)
Clinical stage			
I	9 (3.3)	2 (1.5)	7 (4.3)
II	57 (20.7)	0 (0.0)	5 (3.1)
III	65 (23.6)	3 (2.3)	17 (10.4)
IV	144 (52.4)	124 (95.4)	125 (76.7)
Not known	0 (0.0)	1 (0.8)	9 (5.5)
HPV status*			
Positive	36 (13.1)	57 (42.5)	93 (57.1)
Negative	239 (86.9)	77 (57.5)	70 (42.9)

#clinical data not available for 4 cases

* HPV status for Correlate based on p16 expression

Figure Legends

Figure 1. Development and validation of hypoxia-immune prognostic classifier**.** Two-dimensional unsupervised hierarchical clustering identifies distinct HNC patient subgroups with co-ordinate expression of hypoxia- and immune response-related genes (red, high gene expression; white, intermediate gene expression; blue, low gene expression). Hypoxia (Eustace) signature genes are coloured purple and immune (CIRC) signature genes yellow, for (A) TCGA, (C) Chicago and (E) Correlate datasets. Alcohol consumption, smoking status and HPV/p16 status are represented per patient. Correlate dataset (C) includes 12 patients who underwent tonsillectomy for management of a non-malignant process (snoring) coloured cyan (controls). (B, D, F) Kaplan-Meier survival plots for OS stratified according to the combined hypoxia-immune signature for (B) TCGA (n = 275), (D) Chicago (n = 130) and (F) Correlate (n = 123) patient cohorts. Data are censored at 5 years. Log-rank P-values are displayed for non-adjusted and HPV/p16 adjusted analyses.

Figure 2. Expression of individual hypoxia- or immune-response related genes within the heat-map defined subgroups. Box-and-whisker plots showing normalised z scores for selected genes of interest, acquired from TCGA dataset. The hypoxia^{low}/immune^{high}, hypoxia^{high}/immune^{low} and mixed subgroups (as defined by two dimensional unsupervised hierarchical clustering of gene expression) are represented as light grey, dark grey and white respectively. Data were analysed using Kruskal-Wallis test with Dunn's correction; p values are reported as: ns non-significant *<0.05, **<0.005, ***<0.0005, ****<0.0001.

Figure 3. Immunohistochemical staining of selected immune- or hypoxia-marker expression and correlation with survival. Representative images showing (A) high and low numbers of TILs (H&E); (B) high and low expression of CD3 within different regions of the tumour (within tumour nests, the stroma proximal to the tumour and at the tumour periphery); (C) constitutive versus inductive expression of PD-L1 on tumour cells (top panels) and distinct patterns of PD-L1 expression on immune cells within the tumour microenvironment (bottom panels). (D) high and low expression of LAG3 on infiltrating immune cells (E); high and low CA-IX expression within tumour cells. Main images: original magnification x10; inset: x40. (F-J) Kaplan-Meier curves for overall survival according to individual immune or hypoxia marker expression. Data were dichotomised into low (blue line) and high (red line) expression subgroups based on the indicated cut-offs (for CD3 median value used) and censored at 5 years.

Figure 4. Relationship between heat map groups, IHC marker expression and immune profile. (A) All cases where both protein and gene expression data were available (n = 123) were sub-divided into the three gene cluster defined subgroups and IHC marker expression scatterplots generated using raw scores (TILs, PD-L1 on immune cells (IC)) or z scores (CD3, PD-L1 tumour, LAG3, CA-IX). Bars represent mean with SD. Data were analysed using Kruskal-Wallis test with Dunn's correction; exact p values are reported. (B) The 123 cases were assigned an immune profile (desert, excluded or inflamed) based on their CD3 IHC score. Graph shows percentage of samples within each immune phenotype belonging to the three heat map gene cluster-defined subgroups. Data were analysed using Fisher's exact test.

Figure 5. Spatial interactions of immune cells and hypoxia in the tumour microenvironment. (A-C) Representative phenochart images of Opal multiplex IHC staining exemplifying the three hypoxia-

immune subgroups, hypoxia^{high}/immune^{low} (A), mixed (B) and hypoxia^{low}/immune^{high} (C). Images for individual markers: CD3 (green), CA-IX (red), PD-L1 (blue) are shown (left) and composite images with scale bars (right). S: stroma; T: tumour; white arrows: intratumoural CD3+ T cells; red stars stromal PD-L1+/CD3+ T cells; yellow triangles: PD-L1+ 'cordon' at tumour-stroma interface. (D and E) Quantitation of CD3, CA-IX and PD-L1 staining demonstrates an inverse correlation between tumoural CA-IX expression and prevalence of CD3-positive T cells (D) and a positive correlation between CD3 T cells and PD-L1 expression (E). Graphs display percentage of cells expressing the marker of interest out of total nucleated cells evaluated and located within the specified compartment (tumour, stroma or combined) for 15 cases. Data were analysed using Spearman's correlation.