Oral Oncology 128 (2022) 105862

Contents lists available at ScienceDirect

Oral Oncology

journal homepage: www.elsevier.com/locate/oraloncology

Correlation and colocalization of HIF-1α and pimonidazole staining for hypoxia in laryngeal squamous cell carcinomas: A digital, single-cell-based analysis

Justin E. Swartz [a,b,*], Hilde J.G. Smits [c], Marielle E.P. Philippens [c], Remco de Bree [b], Johannes H.A.M. Kaanders [d], Stefan M. Willems [e]

[a] Department of Otorhinolaryngology – Head and Neck Surgery, University Medical Center Utrecht, Utrecht, the Netherlands
[b] Department of Head and Neck Surgical Oncology, University Medical Center Utrecht, Utrecht, the Netherlands
[c] Department of Radiotherapy, University Medical Center Utrecht, Utrecht, the Netherlands
[d] Department of Radiation Oncology, University Medical Center Nijmegen, Nijmegen, the Netherlands
[e] Department of Pathology and Medical Biology, University Medical Center Groningen, Groningen, the Netherlands

ARTICLE INFO

Keywords:
Head and Neck Neoplasms
Image Processing, Computer-Assisted
Hypoxia
Nitroimidazoles
Hypoxia-Inducible Factor 1
Pathology
Microscopy
Reproducibility of Results

ABSTRACT

Objective: Tumor hypoxia results in worse local control and patient survival. We performed a digital, single-cell-based analysis to compare two biomarkers for hypoxia (hypoxia-inducible factor 1-alpha [HIF-1α] and pimonidazole [PIMO]) and their effect on outcome in laryngeal cancer patients treated with accelerated radiotherapy with or without carbogen breathing and nicotinamide (AR versus ARCON).

Materials and Methods: Immunohistochemical staining was performed for HIF-1α and PIMO in consecutive sections of 44 laryngeal cancer patients randomized between AR and ARCON. HIF-1α expression and PIMO-binding were correlated using digital image analysis in QuPath. High-density areas for each biomarker were automatically annotated and staining overlap was analyzed. Kaplan-Meier survival analyses for local control, regional control and disease-free survival were performed to predict a response benefit of ARCON over AR alone for each biomarker.

Results: 106 Tissue fragments of 44 patients were analyzed. A weak, significant positive correlation was observed between HIF-1α and PIMO positivity on fragment level, but not on patient level. A moderate strength correlation ($r = 0.705$, $p < 0.001$) was observed between the number of high-density staining areas for both biomarkers. Staining overlap was poor. HIF-1α expression, PIMO-binding or a combination could not predict a response benefit of ARCON over AR.

Conclusion: Digital image analysis to compare positive cell fractions and staining overlap between two hypoxia biomarkers using open-source software is feasible. Our results highlight that there are distinct differences between HIF-1α and PIMO as hypoxia biomarkers and therefore suggest co-existence of different forms of hypoxia within a single tumor.

Introduction

Tumor hypoxia results in worse local control and patient survival in patients with cancer, including those with head and neck squamous cell carcinoma (HNSCC) [1–3]. The best method to classify a tumor as hypoxic or normoxic is still a matter of debate. It may be done in both tissue and imaging-based techniques, including but not limited to Eppendorf pO_2 histography [4–6], ^{18}F-MISO or ^{18}F-AZA PET [7], or BOLD- or DW-MRI [8,9]. Tissue-based techniques include the use of biomarkers for hypoxia. These biomarkers may be detected in biopsies taken as a part of routine clinical work-up.

While multiple hypoxia biomarkers are available, each marker has distinct differences and an 'ideal hypoxia biomarker' is not available. Two types of biomarkers are distinguished: endogenous and exogenous biomarkers. Hypoxia will trigger a cellular response to improve cellular survival under hypoxic circumstances [10]. Proteins upregulated under hypoxia may be used as endogenous biomarkers for hypoxia. One of the best described endogenous hypoxia biomarkers is the transcription

* Corresponding author at: Heidelberglaan 100, Post-box 85500, 3508 AB Utrecht, the Netherlands.
E-mail address: j.e.swartz@umcutrecht.nl (J.E. Swartz).

https://doi.org/10.1016/j.oraloncology.2022.105862
Received 22 February 2022; Received in revised form 31 March 2022; Accepted 12 April 2022
Available online 18 April 2022

factor Hypoxia-inducible factor 1-alpha (HIF-1α) [11]. This protein is constitutively expressed, but is quickly degraded under normoxic circumstances. Under hypoxic circumstances HIF-1α accumulates in cells and together with the HIF-1β subunit this transcription factor binds to hypoxia-responsive elements (HREs) on the DNA to upregulate cellular survival mechisms [12]. Downstream targets of HIF-1α include carbonic anhydrase IX (CA-IX) and glucose-transporter-1 (GLUT-1), which are also considered biomarkers for hypoxia. For all these markers, it has been shown that high expression is related to worse outcome in HNSCC patients [2].

Exogenous hypoxia biomarkers are drugs that are administered to patients prior to biopsy. Hypoxia-activated pro-drugs are a subgroup of exogenous hypoxia biomarkers that are activated when tissue pO_2 reaches below a certain threshold [13]. These drugs are selectively metabolized or bound by hypoxic cells [14]. Pimonidazole (PIMO), a drug of the 2-nitroimidazole class of antibiotics, is irreversibly bound to proteins in the cytoplasm below pO_2 levels of 10 mmHg. After administering PIMO to patients intravenously before biopsy, PIMO binding may be detected using immunohistochemistry [15].

Radiotherapy is an important treatment modality in HNSCC, but its effectivity is reduced under hypoxia [16]. Therefore, treatment modifications have been developed to improve tumors' sensitivity. One of these modifications is accelerated radiotherapy (AR) with carbogen breathing and nicotinamide (ARCON). Carbogen (a gas mixture of 98% O_2 and 2% CO_2) and nicotinamide (a vasoactive agent) improve tumor oxygenation. Therefore, ARCON is thought to increase the sensitivity to radiotherapy [17]. A trial in laryngeal squamous cell carcinoma (LSCC) patients has shown improved regional control and a trend toward improved disease-free survival in patients treated with ARCON compared to patients treated with AR alone in tumors with high PIMO binding [18]. As there are distinct differences between HIF-1α and PIMO as hypoxia markers, it raises the question how HIF-1α and PIMO correlate and whether HIF-1α alone or a combination of PIMO and HIF-1α is better able to identify patients that benefit from the addition of ARCON over AR alone.

The comparison of biomarker expression may be done using visual scoring by a pathologist [19]. This is currently seen as the reference standard, even though this method may suffer considerable inter-observer variability depending on the staining [20,21]. Alternatively, digital image analysis has been performed for hypoxia markers in several studies [22,23]. These studies compared both markers based on the number of positive pixels or positive staining area. Although this method is more objective than visual scoring, it does not take into account cellular features such as size, nor whether the staining was present in the nucleus or cytoplasm. QuPath is a free, open-source software package that performs biomarker analysis by identifying biomarker expression in individual cells [24]. Moreover, it detects staining positivity in the nuclear and cytoplasmic cell compartments separately so that only the biologically relevant cellular compartment is considered.

The goal of the present study was to compare HIF-1α and PIMO expression in LSCC patients participating in a randomized controlled trial of ARCON versus AR. This was done using a digital image analysis using QuPath where we compared not only positive cell counts, but also the location of their expression within the tissue and staining overlap. Finally, the effect of HIF-1α and PIMO staining on survival and the benefit of ARCON over AR for tumors positive for these hypoxia biomarkers were investigated.

Patients and methods

Patient cohort

Laryngeal tumor biopsies from 58 patients who participated in a phase III randomized trial were used in this study (Figure 1) [18]. In this trial 345 patients with LSCC were randomized to AR or ARCON. Inclusion criteria were classification T2b and higher, any N-stage, no distant metastases, WHO performance status 0 or 1, age above 18 years and written informed consent. Institutional review board approval was obtained from the Radboud University Centre Nijmegen (Radboudumc) Medical Research Ethics Committee. Seventy-nine of these patients participated in a translational side study and had received PIMO intravenously two hours before biopsy. The 58 patients in the present study were a subgroup of these 79 patients that had sufficient tissue available in the pathology archives of the RadboudUMC.

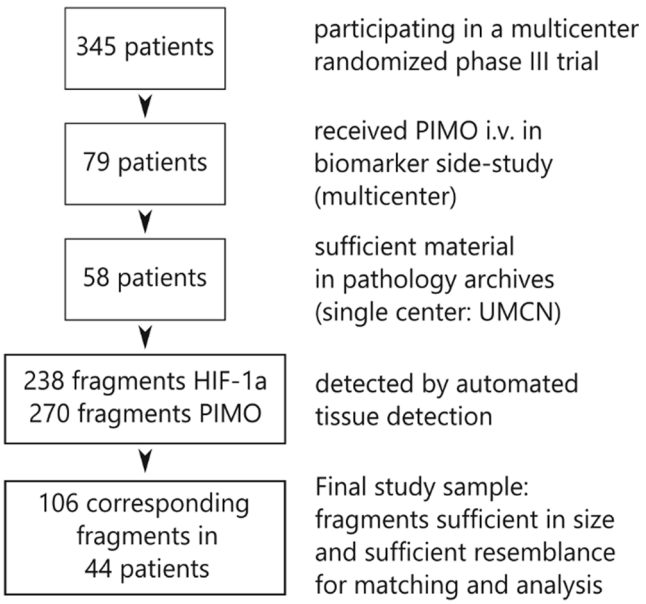

Fig. 1. Patient inclusion flowchart.

Immunohistochemical staining

Consecutive tissue sections were used for HIF-1α and PIMO-staining. For PIMO staining, sections were deparaffinized, followed by rehydration using Histosafe (Pathosafe, Selargius, Italy). Antigen retrieval was performed using a citrate buffer for 30 min. Endogenous peroxidase activity was blocked using a 3% H2O2 buffer in methanol. Sections were pre-incubated with Primary Antibody Diluent (PAD, Cat# BUF014, Bio-Rad, Veenendaal, the Netherlands) and blocked with normal donkey serum 5% in PAD. The primary antibody was a Mouse-anti-PIMO 1:50 in PAD overnight at four degrees Celsius (Lot# 9.7.11, HydroxyProbe, Massachusets USA), followed by a donkey-anti-mouse secondary antibody (Cat# 715–066-150, Jackson ImmunoResearchLaboratories Inc, Ely, UK). An ABC reagent was applied, followed by the DAB staining and counterstaining with hematoxylin.

Immunohistochemical staining for HIF-1α was performed as previously described [25]. In brief, sections were deparaffinized and rehydrated. Endogenous peroxidase activity was blocked using a 3% H2O2 solution in PBS. Antigen retrieval was performed by boiling the slides in a pH 9.0 EDTA buffer for 20 min. The Novolink kit (Leica Biosystems, Rijswijk, the Netherlands) was used according to the manufacturer's instructions for staining. Incubation with the primary antibody (Mouse-anti-HIF-1α, BD Biosciences, cat# 610959, lot 4 073 775, diluted 1:50 in PBS-BSA) was performed overnight at four degrees Celsius. For every staining batch, a renal cell carcinoma tissue section was used as a positive and negative control by incubation with the primary antibody or PBS-BSA to ensure similarity of the staining.

Digital image analysis

All sections were digitized using a Hamamatsu Nano Zoomer XD scanner at 40x magnification. Analyses were performed using QuPath

[24]. First, separate tissue fragments were automatically detected on each section and exported to separate image files (Figure. 2a, b). The corresponding HIF- and PIMO-stained tissue fragments were manually matched and automatically aligned in ImageJ using the TrakEM2 plugin (Figure 2c) [26]. When automatic alignment was not possible, alignment was performed manually in TrakEM2.

For each fragment, the tissue was automatically identified and annotated in QuPath. Areas of necrosis, healthy epithelium, scanning artifacts and areas of foldover were manually removed from the annotation. The overlapping area of the annotations in the corresponding HIF-1α and PIMO tissue fragments was used for analysis to ensure that the analyzed area was identical for both stains.

Cell detection was performed in QuPath using the settings in Supplementary info S1 (Figure 3). As the staining intensities of HIF-1α and PIMO were different, a different threshold was used for each stain. In most tumors HIF-1α staining was quite diffuse, although there were regions with more intense staining. Preliminary comparisons of HIF-1α and PIMO staining patterns suggested co-localization between PIMO staining and the intense regions of HIF-1α staining. To investigate this hypothesis only strong HIF-1α staining was taken into account in the following analyses. A cutoff for strong HIF-1α positivity was set by creating a composite (training) image from 17 different patients and visually identifying the best cutoff value. The researcher determining the threshold for HIF-1α positivity did not have knowledge of the PIMO expression of the corresponding areas (and vice versa) at that moment. Thus, thresholds were set in an unsupervised (blinded) manner. Cells with strong positivity for HIF-1α were defined as cells with an optical density (OD) of the DAB color > 0.65 in the nucleus. The staining intensity of PIMO was relatively weak and the cutoff for PIMO positivity was set at a DAB OD > 0.10 in the cytoplasm. Following positive cell detection, hypoxic regions were identified. This was done by automatically annotating regions with a high density of positive cells using a publicly available script [27]. The parameters used for hypoxic region detection were visually optimized. Identical settings were used for HIF-1α and PIMO and are provided in Supplementary info S1.

To convert the values from fragment to patient level, the total number of positive and negative cells for all tissue fragments within one patient were summed. On patient level, tumors with 2.6% or more PIMO positive cells were considered hypoxic, in concurrence with the original study [18]. For strong HIF-staining (with weak or moderate staining not taken into account) no staining cutoff is available in previous literature. Therefore, the median value of 6.2% strongly HIF-positive cells was used as a cutoff to classify a tumor as hypoxic.

Concurrence in classification as normoxic versus hypoxic tumors between both hypoxia markers was investigated by Pearson correlation analysis and the McNemar test. Statistical significance was set at $p \leq$

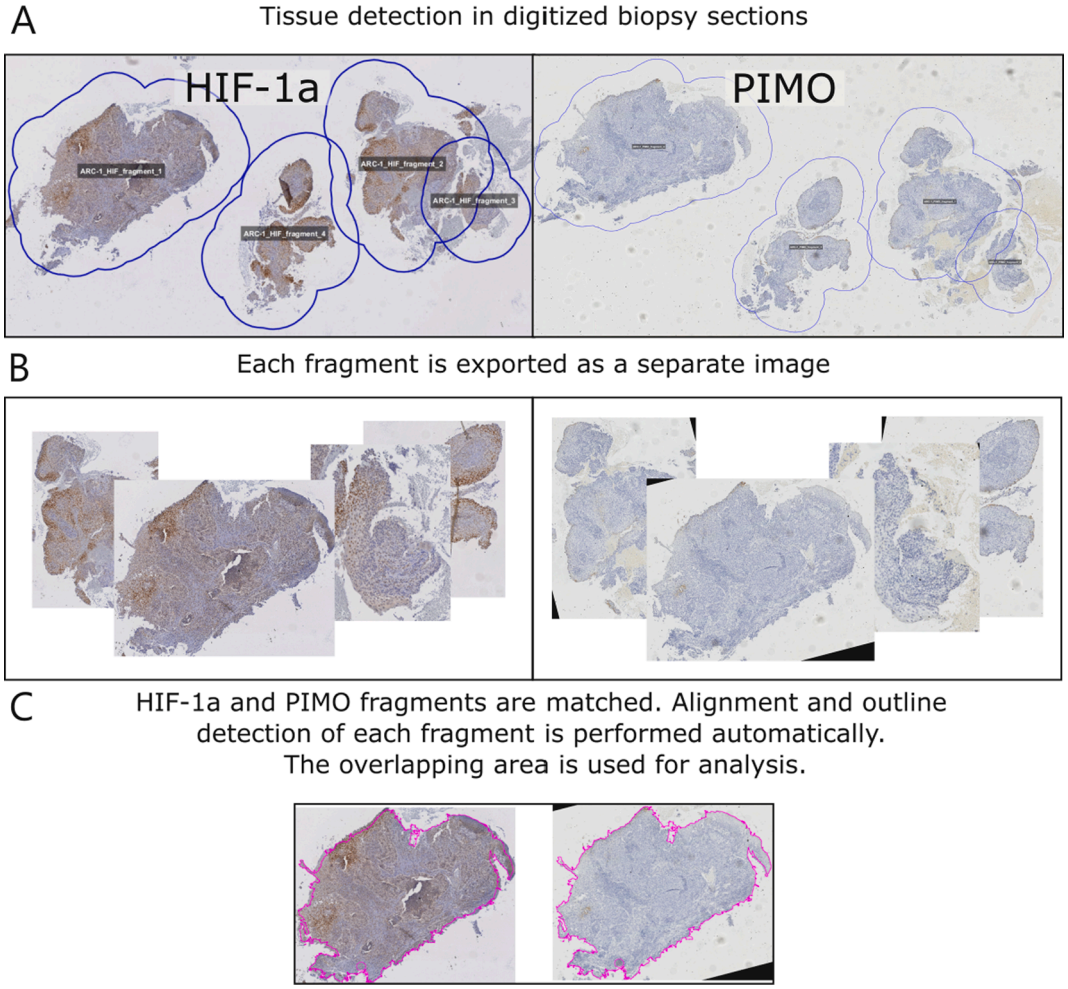

Fig. 2. Image preparation process. A shows a scanned tissue section containing four separate tissue fragments for HIF-1α and PIMO. Using the 'Thresholding' function in QuPath the tissue fragments were automatically detected. Each tissue fragment is exported to a separate file as shown in B. C shows matching fragments of HIF-1α (left) and PIMO staining (right). The tissue outline of each fragment was automatically detected and the overlapping area of both fragments (shown in pink) was used for the analyses. Analyses were performed on a per fragment level (i.e. for each corresponding fragment individually) and on a per patient level (i.e. by summing the data of the individual fragments belonging to each patient). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

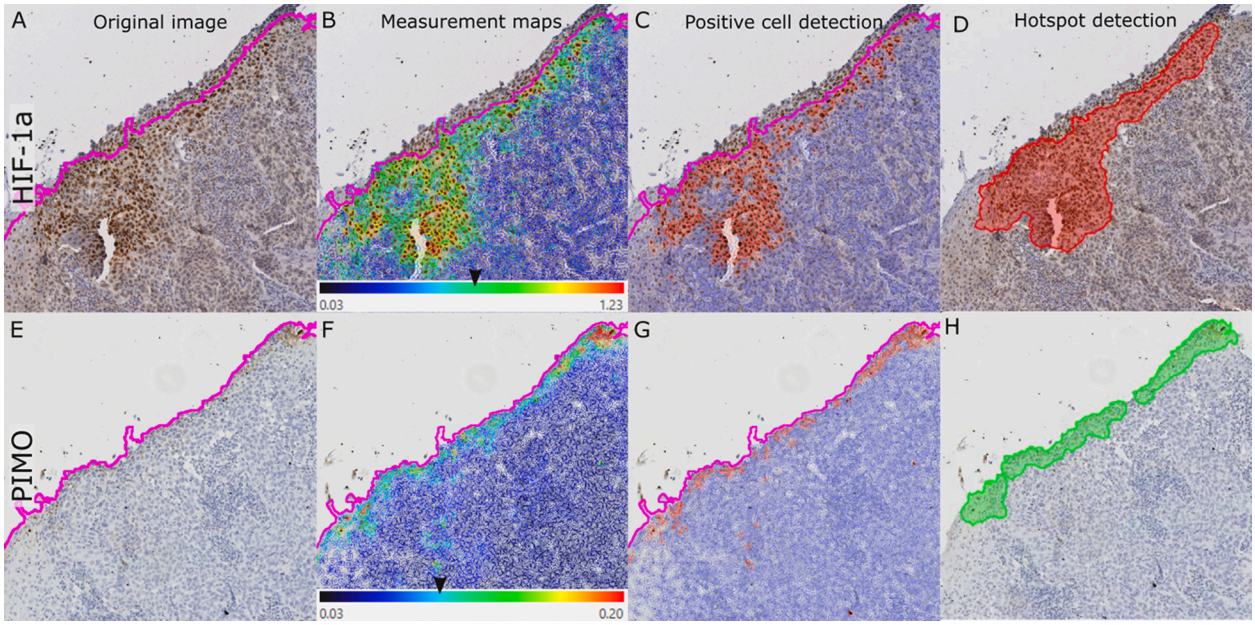

Fig. 3. Positive cell detection. A representative area of a biopsy is shown. Note that PIMO staining was less intense than HIF-1α staining. Each row depicts the original image with the tissue annotation (A and E), the measurement map with the corresponding legend (B and F), followed by the final positive cell detection with threshold 0.65 optical density for HIF-1α and threshold 0.10 for PIMO marked with an arrow (C and G). Here blue cells are negative and red cells are positive. The automatically annotated hotspots are shown in D and H. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

0.05. Overlap scores of the two stains were calculated using Dice-scores, positive predictive value (PPV) and conformity index (C-index). The Dice score describes the relationship of the overlapping area in relation to the total area [28]. It is calculated as Dice = 2 * overlapping area / (area of HIF-1α staining + area of PIMO staining). The C-index also describes the relationship of the overlapping area to the total area as does the Dice-score, albeit slightly different, and is often used to compare radiotherapy plans [29]. It is calculated as C-index = overlapping area / (area positive for HIF-1α + area positive for PIMO – overlapping area)). The PPV is a measure of how likely a positive score of a certain test relates to a positive score using the reference standard. In the present study the PPV was calculated with HIF-1α as the test and PIMO as the reference standard. It was calculated as PPV = true positive area / (true positive + false positive area).

Survival analyses

Of each patient, the date and type of local recurrence, regional recurrence or the occurrence of distant metastases as well as the date and reason of death and date of last follow-up visit were recorded. Follow-up was at least 5 years from the date of randomization for all surviving patients. For local control (LC) and regional control (RC), the incidence of such a recurrence was considered an event and patients were censored at the moment of the last follow-up or death. For disease-free survival (DFS), the date of any recurrence (local, regional or distant) or death were considered an event and patients were censored at the moment of last follow-up. The outcomes were compared using the log-rank test of a Kaplan-Meier survival analysis. All statistics were performed in SPSS Statistics version 26 (IBM).

Results

Patient characteristics

Biopsies of 58 patients were stained for HIF-1α and PIMO (Figure 1). After staining, digitization of the sections, and tissue fragment detection, 238 tissue fragments of HIF-1α and 270 tissue fragments of PIMO were identified, with an average of 4.25 fragments per patient. The final analysis was performed in the 106 adequately matched pairs of tissue fragments of 44 patients that were sufficient in size and contained sufficient amounts of tumor tissue (mean 2.33 fragments, range 1–9 fragments per patient). The baseline characteristics of these patients are shown in Table 1.

Table 1
Baseline patient characteristics.

	AR, n = 20	ARCON, n = 24
Mean age (SD)	61 (8.0)	64.6 (8.5)
Sex		
Male	14 (70)	18 (75)
Female	6 (30)	6 (25)
WHO performance status		
0	13 (65)	19 (79)
1	7 (35)	5 (21)
Site		
Glottic	5 (25)	6 (25)
Supraglottic	15 (75)	18 (75)
T-classification		
T2	6 (30)	3 (13)
T3	9 (45)	18 (75)
T4	5 (25)	3 (13)
N-classification		
N0	6 (30)	9 (38)
N1	7 (35)	3 (13)
N2a	1 (5)	1 (4)
N2b	1 (5)	2 (8)
N2c	5 (25)	9 (38)
N3	0 (0)	0 (0)
Outcomes		
Local recurrence	0 (0)	4 (17)
Regional recurrence	4 (20)	1 (4)
Distant metastasis	4 (20)	1 (4)
All-cause mortality	6 (30)	8 (33)
Events DFS	6 (30)	9 (38)

Characteristics shown as n (%), unless stated otherwise.
Abbreviations: AR = accelerated radiotherapy, ARCON = accelerated radiotherapy with carbogen breathing and nicotinamide, DFS = disease-free survival, WHO = World Health Organization.

Correlations between HIF-1α and pimonidazole positivity

On a per fragment level, the median percentage of positive cells was 4.1% (IQR 1.0 – 13.7) for strong HIF-1a staining and 2.6% (0.4 – 10.9) for PIMO. The correlation between these percentages was significant but weak ($r = 0.365$, $p < 0.001$, Figure 4). On a per patient level, the median percentage of positive cells was 6.2% (IQR: 2.4 – 12) for strong HIF-1α staining and 5.1% (IQR: 0.6 – 11) for PIMO staining. The correlation between these percentages was not statistically significant ($r = 0.176$, $p = 0.253$).

Of the 106 tissue fragments, there were 45 (42.5%) with 6.2% or more HIF-1α positive cells and 53 (50%) with 2.6% or more PIMO-positive cells (Table 2). The McNemar test shows that these positivity rates are comparable ($p = 0.253$). The concurrence rate (HIF+/PIMO + or HIF-/PIMO-) was 62.3%. On a per patient level, there were 22 patients (50%) with 6.2% or more HIF-1α positive cells and 29 (65.9%) patients with 2.6% or more PIMO positive cells. The McNemar test shows that these positivity rates are also comparable ($p = 0.167$), the concurrence rate was 56.8%.

Staining hotspot and overlap analysis

In the previous analysis, tumors or fragments were considered hypoxic when the percentage of positive cells was above a set threshold. In the following analysis, we defined tumors or fragments positive when hypoxic regions could be detected by QuPath. In the 106 tissue fragments, hypoxic regions of strong HIF-1α staining were detected in 84 fragments (77.8%) and hypoxic regions of PIMO staining were detected in 72 fragments (68%). On a per fragment level, the concurrence rate (HIF+/PIMO + or HIF-/PIMO-) was 77.3%. A median of 3 (IQR 1 – 8.5) hypoxic regions were detected for HIF-1α and a median of 2.5 (IQR: 0 – 7) hypoxic regions were detected for PIMO. When summed to a per patient level, a median of 11 (IQR: 3.3 – 22.5) and 7.5 (3.3 – 11) hypoxic regions were detected for HIF-1α and PIMO respectively. The number of detected hypoxic regions for both stains significantly correlated both on a per fragment level ($r = 0.511$, $p < 0.001$) and on a per patient level ($r = 0.705$, $p < 0.001$, Figure 4).

The hypoxic region detection analysis also meant that we were able to investigate overlap between hypoxic regions according to strong HIF-1α and PIMO-staining (Figure 5). In 16 fragments (15%) no hotspots of either HIF-1α or PIMO were detected. In 56 fragments (62%) of the remaining 90 fragments, the identified hotspots of HIF-1α and PIMO showed overlap. The median Dice score was 0.01 (IQR: 0 – 0.07), the median PPV was 0.12 (IQR: 0 – 0.35) and the median C-index was 0.07 (IQR: 0 – 0.19). Interestingly, there was a large number of outliers with a high PPV. This means that in these patients the hypoxic regions according to HIF-1α staining had a high likelihood to also be PIMO-positive.

Survival analyses

The patient outcomes are shown in Table 1. In the survival analyses, high HIF-1α expression, high PIMO binding or a combination of both were not significantly associated with a difference in LC, RC or DFS ($p > 0.05$). When patients were stratified for high or low PIMO binding, HIF expression, or positivity for one or both hypoxia markers, there was no survival difference between AR and ARCON treatment arms.

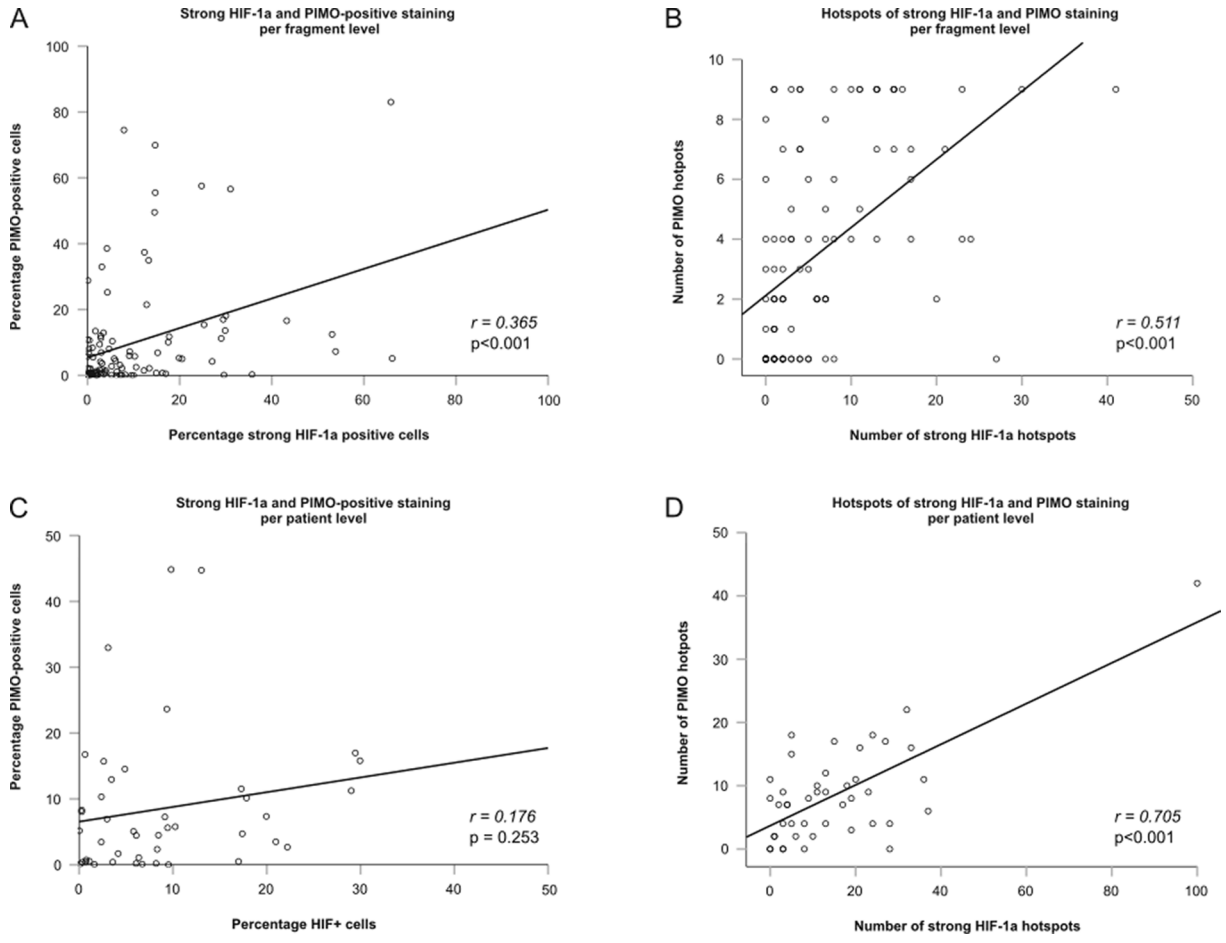

Fig. 4. Scatterplots of percentage positive cells (A, C) and number of detected staining hotspots (B, D) for strong HIF-1α and PIMO staining.

Table 2
Crosstabulation of strong HIF-positivity and PIMO-positivity.

	Per tissue fragment			Per patient		
Method 1: regarded as positive when positive fraction is above cutoff[1]						
	PIMO-negative	PIMO-positive	Total	PIMO-negative	PIMO-positive	Total
Strong HIF-negative	37 (69.8)	24 (45.3)	61 (57.5)	9 (60)	13 (44.8)	22 (50)
Strong HIF positive	16 (30.2)	29 (54.7)	45 (42.5)	6 (40)	16 (55.2)	22 (50)
Total	53 (100)	53 (100)	106 (100)	15 (100)	29 (100)	44 (100)
Method 2: regarded as positive when hotspots of each staining are detected						
	PIMO-negative	PIMO-positive	Total	PIMO-negative	PIMO-positive	Total
Strong HIF-negative	16 (47.1)	6 (8.3)	22 (20.8)	2 (33.3)	2 (5.3)	4 (9.1)
Strong HIF positive	18 (52.9)	66 (91.7)	84 (79.2)	4 (66.7)	36 (94.7)	40 (90.9)
Total	34 (100)	72 (100)	106 (100)	6 (100)	38 (100)	44 (100)

Values are shown as n (%).
[1]Cutoff-value 6.2% positive cells for HIF (median value) and 2.6% for PIMO (based on previous literature).

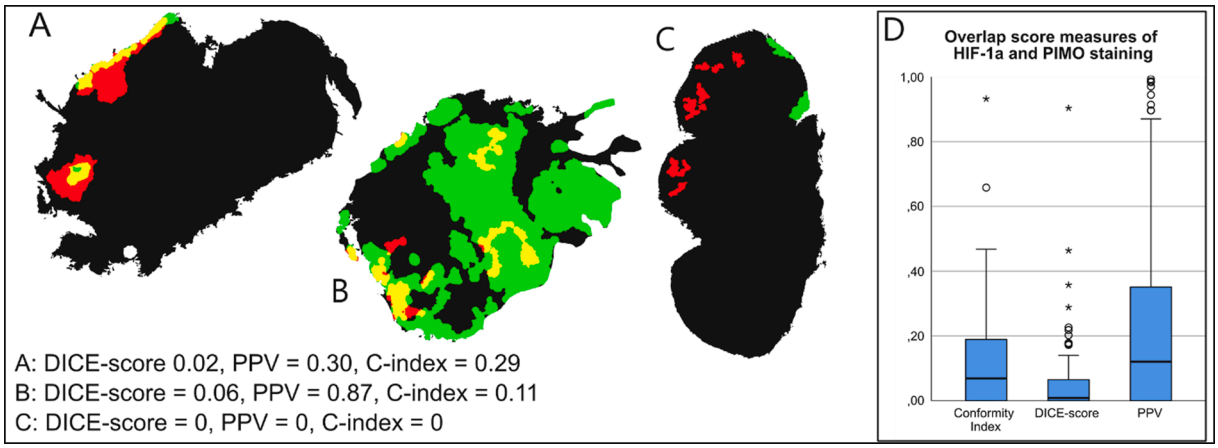

Fig. 5. Hotspot overlap analysis. Three example tissue fragments are shown with an average (A), low (B) and 0 Conformity-Index (C). Tissue outline is shown in black, HIF-1α hotspots are shown in red, PIMO hotspots are shown in green and overlapping areas are shown in yellow. These fragments are shown for illustration purposes and are not to scale. In D a boxplot is shown of the C-indices, PPV and DICE-scores of the 106 tissue fragments. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Discussion

In this digital, cell-based analysis of hypoxia in LSCC biopsies, we found a significant, but weak correlation between HIF-1α expression and PIMO binding in matching tissue fragments. This correlation was not significant when summed to a per patient level. In contrast there was a moderate correlation between the number of hypoxic regions on both a per fragment and per patient level. The spatial overlap between both stains was low. Furthermore, expression of HIF-1α, PIMO binding as well as positivity for both markers combined could not predict a benefit of ARCON over AR.

Comparison of HIF-1a and PIMO as hypoxia biomarkers

Hypoxia is an important issue in solid tumors because hypoxia increases aggressiveness, the potential for metastases, immune cell evasion and resistance to treatment [3]. While several treatment modifications are available for hypoxic tumors, the identification of hypoxic tumors for patient selection for such treatment is not a part of current clinical practice [16,30,31]. Should biomarkers be used to identify hypoxic tumors in a clinical setting, an endogenous marker such as HIF-1α would be preferable to an exogenous marker like PIMO as it would not require additional patient burden and costs.

There are distinct differences between forms of hypoxia and between hypoxia biomarkers such as HIF-1α and PIMO. For instance, PIMO is used as a hypoxia biomarker in many preclinical and animal studies and may be considered a specific marker of hypoxia, while HIF-1α may also be upregulated through other pathways than hypoxia [32,33]. In addition, HIF-1α upregulation may be detected after minutes of hypoxia and is quickly degraded after re-oxygenation [34]. Moreover, HIF-1α requires the cell's transcriptional activity to be functional, which may be inhibited during longer periods of deeper hypoxia. In contrast, PIMO binding requires a longer duration of hypoxia below 10 mmHg pO_2 [35]. And as PIMO-binding is irreversible it may be visualized both in situations where re-oxygenation has taken place and in cells that reach such critically low pO_2levels that they will soon undergo cell death or apoptosis [36]. Therefore, positivity for either hypoxia biomarkers may represent different forms of hypoxia in terms of pO_2-levels and duration.

Because of the need to administer PIMO to patients prior to biopsy, evidence on the effect of PIMO-binding on clinical patient outcome is relatively scarce. There is one study in HNSCC patients investigating the effect of PIMO binding on LRC and found a significant difference: a 2-year LRC 48% for high PIMO-binding and 87% for patients with low PIMO-binding tumors [17]. Because endogenous biomarkers can be studied more easily, and also in a retrospective manner, the body of evidence on the effect of HIF-1α on clinical outcome is larger [2].

When comparing HIF-1α expression and PIMO-binding in our patient cohort, we observed a significant but weak correlation between HIF-1α and PIMO in separate tissue fragments on consecutive slides. Summed to a per patient level, the correlation was not statistically significant. This may originate from a lack of statistical power, as we could analyze only 44 patients in contrast to 106 tissue fragments. This discrepancy may

also highlight tumor heterogeneity and an inherent difference between the two hypoxia markers; certain areas of a tumor may be more hypoxic than others, and the overall percentage or hypoxic cells may not be a good way to compare the two hypoxia markers.

Until now, manual scoring by a pathologist remains the reference standard for biomarker comparison, despite large inter-observer differences depending on the staining [20,21]. Comparison of HIF-1α and PIMO staining was previously done in 36 patients with cervical carcinoma where pathologists scored the presence of each biomarker using a visual, semi-quantitative system [19]. PIMO positivity was divided into 4 categories of percentage positive cells and HIF-1α positivity was divided into 6 categories that combined percentage positivity and staining intensity. The authors found a weak but significant correlation between these two measurements. As we performed a digital analysis, we were able to perform a more objective estimation of cell percentages and found a similar correlation when comparing corresponding tissue fragments, but not whole biopsies. Also, while we only considered strong HIF-1α positivity, the staining intensity by itself was not a part of our scoring system.

The correlation between the number of hypoxic regions on both stains was stronger than the percentage of positive cells for both stains. To our knowledge, comparison of hypoxia biomarkers in this manner has not been done previously. Therefore the (clinical) relevance of this finding is currently uncertain. We believe that this strong correlation illustrates that certain tumors have an architecture making them vulnerable to multiple forms of hypoxia, that may both be detected using HIF-1α and PIMO as hypoxia biomarkers. Possibly, the number of hypoxic regions is a better, or biologically more relevant, way than comparing positive cell fractions to compare these biomarkers. For instance, a tumor may contain small chronically deep hypoxic regions positive for PIMO surrounded by larger areas of 'mild' hypoxia positive for HIF-1α. In such cases there may be a discrepancy in positive cell percentages for both hypoxia biomarkers although two forms of hypoxia coexist.

While we identified a strong correlation between the number of hypoxic regions, the overlap between these areas was generally poor. Interestingly, we could identify subgroups with excellent and poor staining overlap. It may be possible that in patients with excellent overlap, the HIF-1α positivity was caused by hypoxia and that in the non-overlapping group HIF-1α positivity might have originated from transient hypoxia or oncogenic activation. We finally conclude that HIF-1α and PIMO may be considered complementary biomarkers for hypoxia.

HIF-1a and PIMO in relation to ARCON

In the original phase III trial of LSCC patients randomized between AR and ARCON, for patients with high PIMO positivity there was significant benefit of ARCON over AR in regional control and a trend toward better DFS with ARCON [18]. Because of the aforementioned reasons we hypothesized that tumors positive for both HIF-1α and PIMO represent tumors that are currently hypoxic but still vital. Such a subgroup could theoretically benefit most from the addition of hypoxia treatment modification or ARCON. We therefore investigated whether HIF-1α was better than or additive to PIMO in identifying patients that would benefit from the addition of carbogen breathing and nicotinamide to AR.

Unfortunately, we could not find such a role for PIMO, HIF-1α or combined positivity for these markers to predict a benefit of ARCON for LC, RC or DFS. In the original study of 79 patients, there was a significant benefit in RC in patients with high PIMO fractions [18]. In the present study, the low number of patients (44 of the 79 original patients) and events (deaths or recurrences) was a clear limitation to evaluate the predictive power for HIF-1α or a combination of HIF-1α and PIMO.

Our finding is in contrast to a previous, similar sized phase II study of patients with HNSCC treated with AR or ARCON where high PIMO fractions were associated with poor LRC and DFS in patients with HNSCC [17]. Moreover, high PIMO fractions predicted a stronger response to ARCON in this study. The number of events (recurrences and deaths) in this phase II trial was higher than in our study and the phase III trial. Therefore the required sample size to observe a statistically significant difference between two treatment groups was much lower in this study.

Of note, the use of tissue biomarkers in diagnostic biopsies for determining hypoxic status may be hampered by tumor heterogeneity. Especially in larger tumors, a single biopsy from the superficial part of the tumor may reflect the hypoxic status of the whole tumor less well. Hypoxia biomarker analysis in excised tumors or the use of specific hypoxia (PET-) imaging techniques may provide a better overview of hypoxia in a whole tumor [7,37]. Proof-of-principle has been demonstrated for [18] F-MISO-PET as a predictive biomarker for response to the hypoxic cytotoxin tirapazamine [38]. However, the optimal tissue- or imaging based predictive biomarker should be evaluated for each (hypoxia modifying) treatment specifically. In the case of tissue-based biomarkers, it remains to be further explored if a single biopsy will suffice or if multiple (deep) biopsies are required.

The use of a digital, cell-based analysis

The increase of digital pathology and free, open-source software for analysis is of great value to clinicians and researchers. Having an open-source platform is important for innovation and contribution to the software by users. Moreover, free software can also be used in resource-limited settings. QuPath offers a (positive) cell detection function, where cells are detected based on shape features for hematoxylin-DAB or immunofluorescence stained sections [24]. The DAB-staining intensity may then be determined separately in the nuclear and cytoplasmic cell compartments to classify a cell as positive or negative. This is a clear improvement over previous studies on hypoxia markers where the hypoxic fraction was determined by dividing the sum of positive pixels (positive tissue area) by the sum of total pixels (total tissue area) in immunofluorescence images [22,23]. This method does not take the cell size into account and is less suited to distinguish background staining compared to a cell-based approach. To our knowledge, the present study is the first study in LSCC or HNSCC that applies a cell-based, rather than pixel-based analysis to compare two hypoxia markers.

Digital analysis on a cellular level is relatively novel. While the reliability of scoring HIF-1α and PIMO expression has not been assessed in particular, other markers have already been validated. One study compared reproducibility for Ki-67 scoring (which is a nuclear DAB-staining like HIF-1α) and found a high inter-platform and inter-operator reproducibility for digital image analysis in general and in QuPath specifically [39]. Other studies have compared pathologist examination versus QuPath for CD8 + TILs and PD-L1 and found a high concordance between the two [40,41]. Still, automated scoring of HIF-1α and PIMO staining as well as the hotspot detection methods should be validated against scoring by an experienced pathologist.

A limitation of cell detection analysis in QuPath is that it can *detect* but not interpret staining. Indeed we identified some areas where the cell detection was not perfect. Because of this it was still necessary to perform manual corrections and to remove areas of necrosis, scanning artifacts and areas of foldover from the automatically established annotations. To address this issue, QuPath extensions are rapidly becoming available to improve cell detection for instance using deep-learning methods [42].

Conclusion

In summary, this is the first study to compare the immunohistochemical hypoxia markers HIF-1α and PIMO in LSCC using a digital, cell-based analysis. Our study shows that it is feasible to use open-source software to compute positive cell percentages and to investigate

overlap of the two biomarkers in digitized sections in LSCC. However, these digital methods should be further validated against the current reference standard of visual scoring by a pathologist. In this relatively small study, we were unable to identify HIF-1α, or a combination of HIF-1α and PIMO as a marker to predict response to the additional effect of ARCON over AR alone. To our knowledge, this is the first study that uses automated digital imaging technology to show spatial correlations of HIF-1α and PIMO staining in HNSCC. We found a weak correlation between positive cell fractions for the two biomarkers, and a moderate correlation between the number of hypoxic regions for each biomarker. The moderate strength of this correlation in combination with the poor overlap of HIF-1α and PIMO suggests distinct differences between these two hypoxia biomarkers and also highlights the co-existence of different forms of hypoxia in a single tumor.

Declaration of Competing Interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: The original Phase III trial was supported by the Dutch Cancer Society (KWF) Research Fund No. CKTO-2000-09 and KUN-2008-4088 and METOXIA (Metastatic Tumors Facilitated by Hypoxic Tumor Micro-Environments; European Community Grant No. FP7-HEALTH-2007-B222741. No additional funding was received for the present study.

Appendix A. Supplementary material

Supplementary data to this article can be found online at https://doi.org/10.1016/j.oraloncology.2022.105862.

References

[1] Hoogsteen IJ, Marres HAM, Bussink J, van der Kogel AJ, Kaanders JHAM. Tumor microenvironment in head and neck squamous cell carcinomas: predictive value and clinical relevance of hypoxic markers A review. *Head Neck* 2007;29(6):591–604.
[2] 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.
[3] Vaupel P. Physiological Mechanisms of Treatment Resistance. In: Molls M, Vaupel P, C N, M A, eds. *The Impact of Tumor Biology on Cancer Treatment and Multidisciplinary Strategies. Medical Radiology*. Springer Berlin Heidelberg; 2009:273–290.
[4] Chapman JD. Measurement of tumor hypoxia by invasive and non-invasive procedures: a review of recent clinical studies. Radiother Oncol 1991;20:13–9.
[5] Gatenby RA, Coia LR, Richter MP, Katz H, Moldofsky PJ, Engstrom P, et al. Oxygen tension in human tumors: in vivo mapping using CT-guided probes. Radiology 1985;156(1):211–4.
[6] Rudat V, Vanselow B, Wollensack P, Bettscheider C, Osman-Ahmet S, Eble MJ, et al. Repeatability and prognostic impact of the pretreatment pO(2) histography in patients with advanced head and neck cancer. Radiother Oncol 2000;57(1):31–7.
[7] Marcus C, Subramaniam RM. Role of Non-FDG-PET/CT in Head and Neck Cancer. Semin Nucl Med 2021;51(1):68–78.
[8] Wiedenmann N, Grosu A-L, Büchert M, Rischke HC, Ruf J, Bielak L, et al. The utility of multiparametric MRI to characterize hypoxic tumor subvolumes in comparison to FMISO PET/CT. Consequences for diagnosis and chemoradiation treatment planning in head and neck cancer. Radiother Oncol 2020;150:128–35.
[9] Glover GH. Overview of functional magnetic resonance imaging. Neurosurg Clin N Am 2011;22(2):133–9.
[10] Semenza GL. Oxygen sensing, hypoxia-inducible factors, and disease pathophysiology. Annu Rev Pathol Mech Dis 2014;9(1):47–71.
[11] Semenza G. Hypoxia-Inducible Factors in Physiology and Medicine. Cell 2012;148(3):399–408.
[12] Mole DR, Blancher C, Copley RR, Pollard PJ, Gleadle JM, Ragoussis J, et al. Genome-wide association of hypoxia-inducible factor (HIF)-1α and HIF-2α DNA binding with expression profiling of hypoxia-inducible transcripts. J Biol Chem 2009;284(25):16767–75.
[13] Codony VL, Tavassoli M. Hypoxia-induced therapy resistance: Available hypoxia-targeting strategies and current advances in head and neck cancer. Transl Oncol 2021;14(3):101017. https://doi.org/10.1016/j.tranon.2021.101017.
[14] Wardman P. Chemical Radiosensitizers for Use in Radiotherapy. Clin Oncol 2007;19(6):397–417.
[15] Varia MA, Calkins-Adams DP, Rinker LH, Kennedy AS, Novotny DB, Fowler WC, et al. Pimonidazole: A novel hypoxia marker for complementary study of tumor hypoxia and cell proliferation in cervical carcinoma. Gynecol Oncol 1998;71(2):270–7.
[16] 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.
[17] Kaanders JHAM, Wijffels KIEM, Marres HAM, et al. Pimonidazole binding and tumor vascularity predict for treatment outcome in head and neck cancer. Cancer Res 2002;62(23):7066–74.
[18] Janssens GO, Rademakers SE, Terhaard CH, Doornaert PA, Bijl HP, van den Ende P, et al. Accelerated radiotherapy with carbogen and nicotinamide for laryngeal cancer: Results of a phase III randomized trial. J Clin Oncol 2012;30(15):1777–83.
[19] Hutchison GJ, Valentine HR, Loncaster JA, Davidson SE, Hunter RD, Roberts SA, et al. Hypoxia-inducible factor 1α expression as an intrinsic marker of hypoxia: Correlation with tumor oxygen, pimonidazole measurements, and outcome in locally advanced carcinoma of the cervix. Clin Cancer Res 2004;10(24):8405–12.
[20] Polley M-Y, Leung SCY, McShane LM, Gao D, Hugh JC, Mastropasqua MG, et al. An international ki67 reproducibility study. J Natl Cancer Inst 2013;105(24):1897–906.
[21] Sinclair W, Kobalka P, Ren R, Beshai B, Lott Limbach AA, Wei L, et al. Interobserver agreement in programmed cell death-ligand 1 immunohistochemistry scoring in nonsmall cell lung carcinoma cytologic specimens. Diagn Cytopathol 2021;49(2):219–25.
[22] Rademakers SE, Lok J, van der Kogel AJ, Bussink J, Kaanders JHAM. Metabolic markers in relation to hypoxia; staining patterns and colocalization of pimonidazole, HIF-1α, CAIX, LDH-5, GLUT-1, MCT1 and MCT4. BMC Cancer 2011;11(1):167.
[23] Brockton N, Dort J, Lau H, Hao D, Brar S, Klimowicz A, et al. High stromal carbonic anhydrase ix expression is associated with decreased survival in p16-negative head-and-neck tumors. Int J Radiat Oncol Biol Phys 2011;80(1):249–57.
[24] Bankhead P, Loughrey MB, Fernández JA, Dombrowski Y, McArt DG, Dunne PD, et al. QuPath: Open source software for digital pathology image analysis. Sci Rep 2017;7(1). https://doi.org/10.1038/s41598-017-17204-5.
[25] Swartz JE, Pothen AJ, van Kempen PMW, Stegeman I, Formsma FK, Cann EMV, et al. Poor prognosis in human papillomavirus–positive oropharyngeal squamous cell carcinomas that overexpress hypoxia inducible factor-1α. Head Neck 2016;38(9):1338–46.
[26] Cardona A, Saalfeld S, Schindelin J, et al. TrakEM2 software for neural circuit reconstruction. *PLoS One*. 2012;7(6).
[27] Nelson M. Hotspot detection. https://gist.github.com/Svidro/6171d6d24a85539d3af5d417bc928d50#file-hotspot-detection-0-2-0m8-groovy.
[28] Carass A, Roy S, Gherman A, Reinhold JC, Jesson A, Arbel T, et al. Evaluating White Matter Lesion Segmentations with Refined Sørensen-Dice Analysis. Sci Rep 2020;10(1). https://doi.org/10.1038/s41598-020-64803-w.
[29] Feuvret L, Noël G, Mazeron J-J, Bey P. Conformity index: A review. Int J Radiat Oncol Biol Phys 2006;64(2):333–42.
[30] Horsman MR, Mortensen LS, Petersen JB, Busk M, Overgaard J. Imaging hypoxia to improve radiotherapy outcome. Nat Rev Clin Oncol 2012;9(12):674–87.
[31] Bussink J, Kaanders JHAM, van der Kogel AJ. Tumor hypoxia at the micro-regional level: Clinical relevance and predictive value of exogenous and endogenous hypoxic cell markers. Radiother Oncol 2003;67(1):3–15.
[32] Palazon A, Goldrath A, Nizet V, Johnson R. HIF Transcription Factors, Inflammation, and Immunity. Immunity 2014;41(4):518–28.
[33] Scholz CC, Taylor CT. Targeting the HIF pathway in inflammation and immunity. Curr Opin Pharmacol 2013;13(4):646–53.
[34] Jewell UR, Kvietikova I, Scheid A, Bauer C, Wenger RH, Gassmann M. Induction of HIF–1α in response to hypoxia is instantaneous. FASEB J 2001;15(7):1312–4.
[35] Ljungkvist ASE, Bussink J, Kaanders JHAM, Rijken PFJW, Begg AC, Raleigh JA, et al. Hypoxic cell turnover in different solid tumor lines. Int J Radiat Oncol Biol Phys 2005;62(4):1157–68.
[36] Bennewith KL, Durand RE. Quantifying transient hypoxia in human tumor xenografts by flow cytometry. Cancer Res 2004;64(17):6183–9.
[37] Thorwarth D, Wack L-J, Mönnich D. Hypoxia PET imaging techniques: data acquisition and analysis. Clin Transl Imaging 2017;5(6):489–96.
[38] Rischin D, Hicks RJ, Fisher R, Binns D, Corry J, Porceddu S, et al. Prognostic significance of [18F]-misonidazole positron emission tomography-detected tumor hypoxia in patients with advanced head and neck cancer randomly assigned to chemoradiation with or without tirapazamine: A substudy of Trans-Tasman Radiation Oncology Group study 98.02. J Clin Oncol 2006;24(13):2098–104.
[39] Acs B, Pelekanou V, Bai Y, Martinez-Morilla S, Toki M, Leung SCY, et al. Ki67 reproducibility using digital image analysis: an inter-platform and inter-operator study. Lab Investig 2019;99(1):107–17.
[40] Humphries MP, Bingham V, Abdullahi Sidi F, Craig SG, McQuaid S, James J, et al. Improving the diagnostic accuracy of the pd-l1 test with image analysis and multiplex hybridization. Cancers (Basel) 2020;12(5):1114. https://doi.org/10.3390/cancers12051114.
[41] Jhun I, Shepherd D, Hung YP, Madrigal E, Le LP, Mino-Kenudson M. Digital image analysis for estimating stromal CD8+ tumor-infiltrating lymphocytes in lung adenocarcinoma. J Pathol Inform 2021;12(1):28. https://doi.org/10.4103/jpi.jpi_36_20.
[42] Schmidt U, Weigert M, Broaddus C, Myers G. Cell Detection with Star-Convex Polygons, Vol 11071. LNCS: Springer International Publishing; 2018.