



Deposited via The University of Sheffield.

White Rose Research Online URL for this paper:

<https://eprints.whiterose.ac.uk/id/eprint/245403/>

Version: Published Version

Article:

Tornaas, S., Wang, R., Emblem Moe, S.E. et al. (2026) Comparative analysis of human papilloma virus detection methods and prognostic significance of transcriptionally active viral load in oropharyngeal squamous cell carcinoma. *Oral Oncology*, 182. 108123. ISSN: 1368-8375

<https://doi.org/10.1016/j.oraloncology.2026.108123>

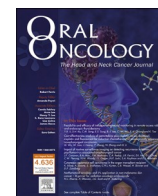
Reuse

This article is distributed under the terms of the Creative Commons Attribution (CC BY) licence. This licence allows you to distribute, remix, tweak, and build upon the work, even commercially, as long as you credit the authors for the original work. More information and the full terms of the licence here:

<https://creativecommons.org/licenses/>

Takedown

If you consider content in White Rose Research Online to be in breach of UK law, please notify us by emailing eprints@whiterose.ac.uk including the URL of the record and the reason for the withdrawal request.



Comparative analysis of human papilloma virus detection methods and prognostic significance of transcriptionally active viral load in oropharyngeal squamous cell carcinoma

Stian Tornaas^a, Ruoyu Wang^b, Svein Erik Emblem Moe^c, Siren Fromreide^a, Rammah Elnour^a, Åsa Karlsdottir^d, Marianne Brydøy^d, Hans Jørgen Aarstad^c, Olav Karsten Vintermyr^e, Nasir Rajpoot^b, Syed Ali Khurram^f, Daniela Elena Costea^{a,e,1,*}, Harsh Nitin Dongre^{a,1,*}

^a Center for Cancer Biomarkers CCBIO and Gade Laboratory for Pathology, Department of Clinical Medicine, University of Bergen, Norway

^b Tissue Image Analytics Centre, University of Warwick, UK

^c Department of Ear-Nose-and-Throat, Head and Neck Clinic, Haukeland University Hospital, Bergen, Norway

^d Department of Oncology, Haukeland University Hospital, Bergen, Norway

^e Department of Pathology, Haukeland University Hospital, Bergen, Norway

^f School of Clinical Dentistry, Faculty of Health, University of Sheffield, UK

ABSTRACT

Background: Human papilloma virus (HPV)-related head and neck squamous cell carcinoma (HNSCC) incidence is rising. Despite the prognostic and therapeutic implication of HPV status, standardised detection methods are lacking. This study compared multiple HPV detection methods and evaluated the prognostic value of viral transcriptional activity in archival HNSCC tissues.

Methods: HPV-DNA status and subtyping was determined by GP5+/GP6 + PCR. Multiplex immunohistochemistry (p16^{INK4}) and chromogenic in-situ hybridization (CISH for E6/E7 mRNA) was performed on tissue microarrays (TMAs) from a retrospective cohort of HNSCC (n = 207). Viral load was quantified using QuPath and an independent deep-learning HPV prediction algorithm was validated on H&E-stained tissue cores.

Results: The multiplex approach accurately detected active HPV in HNSCC primary tumours and metastases, including archival tissue > 20 years old. HPV-DNA was detected exclusively in oropharyngeal SCC and was absent in all other head and neck subsites. Receiver operating curve analysis confirmed all three tested methods match the diagnostic performance of the gold standard HPV-DNA PCR. Furthermore, viral load was identified as an independent predictor of overall survival, maintaining prognostic significance even after adjusting for age, stage and lifestyle factors.

Conclusions: This study underscores the importance of detecting transcriptionally active HPV to better clinical decision-making of HNSCC patients and supports broader adoption of CISH alongside emerging AI-based analysis in clinical practice.

Background

High-risk human papilloma viruses (HPV), particularly HPV-16 and HPV-18 types, are established drivers of the rising global incidence of head and neck squamous cell carcinoma (HNSCC) [1,2]. While HPV positivity, particularly in oropharyngeal squamous cell carcinoma (OPSCC), generally signals a favourable prognosis and suitability for treatment de-intensification [3], a subset of patients still experience poor outcomes [4]. Identifying these high-risk individuals currently requires advanced molecular techniques like digital droplet PCR (ddPCR); however, such methods remain cost-prohibitive and logistically challenging for routine practice [5].

Various methods can identify HPV-related HNSCC. While p16^{INK4} (p16) immunohistochemistry (IHC) is a common surrogate marker, its moderate specificity limits its accuracy in predicting OPSCC outcomes when used in isolation [6,7]. Consequently, p16 IHC alone is not recommended for selecting patients for treatment de-escalation [7]. Precise OPSCC diagnosis, particularly in p16+ cases, requires HPV-specific testing [8], and there is strong evidence that combining p16 IHC with HPV DNA PCR enhances specificity while maintaining sensitivity [9]. A meta-analysis showed that HPV DNA+/p16+ patients demonstrate significantly improved five-year overall-survival (OS) and disease-free-survival (DFS) compared to discordant (DNA-/p16+ or HPV DNA+/p16-) subgroups [10].

* Corresponding authors.

E-mail addresses: daniela.costea@uib.no (D.E. Costea), harsh.dongre@uib.no (H.N. Dongre).

¹ These authors contributed equally.

However, identifying HPV-infection by DNA analysis does not confirm an active oncogenic infection. Biologically significant HPV-related HNSCC is best defined by identifying cellular transformation driven by E6/E7 viral oncogenes [11,12]. Until recently, RT-qPCR was used for detection of E6/E7 mRNA, however, with the development of RNAscope technology, detection of E6/E7 mRNA by in situ hybridisation (ISH) has become more feasible for use even in clinical settings. ISH can also provide valuable information about the spatial localisation and abundance of viral infection.

We established a dual IHC (p16) and ISH (E6/E7 mRNA) protocol to quantify viral load in a well-characterised retrospective HNSCC cohort from a Norwegian cohort. We also aimed to determine if ISH-based viral load quantification improves risk stratification. Additionally, we validated a deep-learning HPV-prediction algorithm, originally developed for whole-slide H&E images, on tissue microarray cores [13]. Together, these approaches seek to refine HPV-related HNSCC classification and identify novel predictive tools for clinical outcome stratification.

Methods

Tumour samples

A retrospective cohort of 207 patients older than 18 years, diagnosed with HNSCC and treated with surgery and/or chemoradiotherapy at Haukeland University Hospital, Bergen, Norway between 2000-2016 was included in the study. Patients without informed consent, with neoadjuvant treatment, missing tissue blocks, and missing clinical information were excluded from the study. The clinical information (age, gender, smoking and alcohol use, localization, TNM stage (based on 7th Edition Union for International Cancer Control (UICC)/American Joint Committee on Cancer (AJCC) guidelines), co-morbidities, recurrence, last date of follow-up, survival) was obtained from patients' medical electronic journal. Formalin fixed paraffin embedded (FFPE) tissue blocks containing the tumour front with surrounding stroma and adjacent normal mucosa were selected for the study. Two cores of 1 mm² per patient were chosen for constructing tissue microarrays (TMAs). TMA sections of 3-4 µm thickness were cut using a microtome (HistoCore Biocut, Leica Biosystems, UK) and mounted on glass slides (Thermo Scientific Superforst Plus, USA). Further, the sections were allowed to adhere to the slides by incubating at 57°C for 2 hours and stored at 4°C until use. This study was approved by the Regional Ethical Committee for Western Norway (REK Vest 2010/481 and 680158) and followed REMARK criteria [14].

HPV DNA detection and subtyping

HPV DNA was detected using a protocol well established in our laboratory [15]. Briefly, three ten µm thick FFPE sections were collected in microcentrifuge tubes, deparaffinized using deparaffinization solution (Qiagen GmbH, Hilden, Germany), digested overnight in ATL buffer and Proteinase K (Qiagen) at 56°C. DNA was extracted using E.Z.N.A. tissue DNA Kit (Omega BioTek, Norcross, GA, USA) followed by DNA quantification using Qubit dsDNA BR assay kit (Thermo Fisher Scientific, Waltham, MA, USA) according to manufacturer's protocol. Extracted DNA was added to a PCR reaction mixture with standard Gp5+/Gp6+ primers, followed by DNA sequencing to determine HPV subtypes. Detailed protocol can be found in [supplementary materials](#).

p16 immunohistochemistry (IHC) and RNA in situ hybridization (ISH)

Four-micron TMA sections were evaluated for the expression of p16 by IHC. In brief, the sections were deparaffinized followed by antigen retrieval using Tris-EDTA pH 9 buffer (92°C for 20 min using a microwave). The expression of p16 in tissue was detected using anti-p16 antibody (clone BC42, Biocare, Concord, CA, USA). EnVision™ FLEX

HRP Magenta Substrate Chromogen (Agilent, Santa Clara CA, USA) was used for visualization.

Chromogenic RNA in situ hybridization (RNA CISH) for detection of high-risk HPV-7 mRNA (detecting HPV types 16, 18, 31, 33, 35, 52, and 58; E6/E7 mRNA; Cat. No. 312351, Advanced Cell Diagnostics (ACD), Hayward, CA, USA) was performed using the RNAscope® 2.5 HD Reagent Kit - Brown (Cat. No. 322310, ACD). The manufacturer's protocol was followed, including boiling of 4-µm TMA sections to unmask RNA targets followed by protease treatment. Samples were then incubated with HPV HR-7 probe for E6/E7 detection, along with positive (peptidylprolyl isomerase B - PPIB) and negative (a transcript from *Bacillus Subtilis*, DapB) controls. After six rounds of signal amplification, signals were visualized and digitally quantified as described previously [15].

Combined IHC and ISH and digital quantification

To simultaneously visualize HPV E6/E7 mRNA transcripts together with p16 protein expression, we employed the integrated co-detection protocol, combining RNAscope ISH with IHC according to the manufacturer's recommendations. Target retrieval was performed in a beaker on a hotplate using co-detection target retrieval solution (ACD) for 15 minutes at 98-102 °C. The slides were then incubated at 4°C overnight with p16 primary antibody (clone BC42) diluted 1:20 in Co-Detection Antibody Diluent (ACD). The next day, slides were washed in PBS-T and then fixed in 10 % neutral buffered formalin for 30 minutes at room temperature (RT). The slides were then treated with protease plus for 30 minutes at 40 °C and incubated with the HPV HR-7 probes for 2 hours at 40 °C, before signal amplification and visualization with DAB chromogen. After this, the co-detection blocker (ACD) was applied to the sections for 15 minutes at 40 °C, and then EnVision FLEX Mouse Linker (Agilent Dako) was applied for 15 minutes at RT, followed by EnVision FLEX HRP secondary antibody (Agilent Dako) for 20 minutes at RT. The antibody signal was developed using EnVision FLEX Magenta chromogen (Agilent Dako), and Haematoxylin was used as counterstaining.

Stained slides were digitally scanned using a Hamamatsu NanoZoomer-XR whole slide scanner (Hamamatsu Photonics) at x40 magnification and analysed using QuPath (version 0.3.2, Queen's University Belfast, UK) [16]. Tumour areas were manually annotated, and E6/E7 mRNA-positive cells were detected using colour deconvolution and threshold-based analysis. Viral load (VL) was quantified as the number of E6/E7 mRNA positive cells per µm² of tumour tissue using a QuPath groovy script (Supp material).

Deep learning algorithm for HPV prediction - external validation

The deep learning framework was based on the methodology established by Wang et al. [13], which employs a convolutional neural network to predict HPV status from routine H&E-stained histological images. The original model was developed and trained using two independent cohorts: the TCGA- HNSC cohort (431 diagnostic whole-slide images (WSIs) from 412 patients across 26 U.S. medical centres) and the Sheffield-OPSCC cohort (72 WSIs from 69 patients). In both training cohorts, HPV infection status was confirmed using combined p16 IHC and high-risk HPV DNA ISH as the reference standard. In the current study, we performed external validation of this trained model on 131 tissue microarray (TMA) cores from the Bergen HNSCC cohort, representing a different population (Norwegian vs. U.S./U.K.), tissue format (TMA cores vs. whole slides), and clinical setting. H&E-stained TMA sections (3-4 µm) were digitally scanned at 20× magnification using a Hamamatsu NanoZoomer scanner. Image preprocessing followed the original protocol: background removal using TIAToolbox [17] to exclude non- tissue areas, pen marks, and artifacts; extraction of 256 × 256 pixel image patches at 10× equivalent magnification from tissue regions; and application of Macenko stain normalization [18] to account for inter- slide and inter-scanner colour variability. The deep learning model architecture consists of a ResNet-based feature extractor followed

by an attention-based aggregation layer, trained using PyTorch 1.11.0 and Python 3.10 as described in Wang et al. [14]. The model outputs a probability score (0-1) for HPV-positive status, with a classification threshold of 0.5. For TMA cores, the entire core was processed as a single instance. Model performance was evaluated using area under the receiver operating characteristic curve (AUC), sensitivity, specificity, positive predictive value (PPV), and negative predictive value (NPV), with HPV DNA PCR as the reference standard. This external validation approach tests both the generalizability of the model to a new population and its robustness to different tissue sampling formats. The study followed TRIPOD AI guidelines for reporting clinical prediction models [19].

Statistical analysis

Cohort description and correlations of clinicopathological parameters were performed using Descriptive Statistics and Cross tabs in Statistical Package for Social Sciences (SPSS), version 28 (IBM, NY, USA). Data are presented as mean ± standard error of mean. Survival curves were plotted using Kaplan-Meier (KM) method. Test of equality for survival distribution of the variables within the parameters was carried using Log Rank (Mantel-Cox). Risks of clinic-pathological parameters for OS (5 years) of HNSCC patients were further examined by univariate survival analysis using Cox’s proportional regression. Tests of independence of clinicopathological parameters with VL were assessed using Pearson’s chi-squared test. Tests of normality for the VL were checked by Kolmogorov-Smirnov and Shapiro-Wilk test, histogram and Q-Q plots. VL was dichotomized into higher and lower based on the mean expression value. Mann-Whitney U test was performed to find significant difference in VL in between high and low groups.

Results

Cohort description

A total of 207 cases of HNSCC cases (age range: 28-90; mean=61.7; median=62) were included in the final analysis comprising 153 males (73.9%) and 54 females (26.1%) (Table 1). Univariate Cox regression identified significantly lower 5-year OS for patients aged > 65 years, those with advanced tumour stage (stages III/IV), smokers and heavy alcohol consumers. Recurrence and HPV status (regardless of testing method i.e. p16 IHC, HPV mRNA ISH, HPV DNA PCR or HPV AI score) were also significantly associated with reduced survival (Fig. 1A and B). Multivariate Cox regression analysis of the significant univariate parameters revealed age >65, lymph node metastasis, recurrence and higher tumour stage as independent predictors of OS (p<0.05). When adjusted for age and tumour stage, only recurrence (p=0.003), and HPV status remained independent prognostic indicators, the latter across all modalities (DNA PCR- p=0.044, p16 IHC- p=0.007 and mRNA ISH- p=0.05) (Table 2).

Detection of HPV by various methods

HPV DNA was detected in 86 samples, all from oropharyngeal tumours, accounting for 41.5% of total HNSCC cases and 58.9% of total OPSCC cases (Fig. 2A). HPV DNA genotyping revealed presence of HPV 16/18 in 84 cases, and HPV 33 and HPV 58 in one case each. All oral cavity and tongue SCC lesions were HPV DNA-negative and hence were excluded from comparative HPV analysis. Considering only OPSCC, IHC for p16 was positive in 98/146 cases (67.1%) while E6/E7 mRNA ISH was positive in 76 cases (52.1%) (Fig. 2A). When p16 IHC and HPV mRNA detection were combined, 98.7% (75/76) of mRNA-positive cases were also p16-positive. A single discordant case (p16-negative/ mRNA-positive) was HPV DNA-positive, suggesting an alternative p16 inactivation mechanism. Deep learning analysis of 91 cases categorised 39 (42.9%) as HPV-positive and 52 (57.1%) as negative. This model,

Table 1
Cohort description. TNM: Tumour Node Metastasis staging from American Joint Committee on Cancer (AJCC 7th Edition); HPV: human papilloma virus; ISH: in-situ hybridization; IHC: Immunohistochemistry.

Parameters	Values (N = 207)
Age at diagnosis in years, median (range)	62 (28–90)
Gender Male	153 (73.9 %)
Female	54 (26.1 %)
TNM stage Stage I	23 (11.1 %)
Stage II	80 (38.5 %)
Stage III	36 (17.3 %)
Stage IV	68 (32.7 %)
Treatment Surgery alone	10 (4.8 %)
Surgery + radiotherapy	133 (63.9 %)
Surgery + radio-chemotherapy	64 (30.8 %)
Site Oropharynx	146 (70.2 %)
Oral cavity	45 (21.6 %)
Other	16 (7.7 %)
Lymph node metastasis No	75 (36.2 %)
Yes	132 (63.8 %)
Alcohol No	150 (72.1 %)
Yes	57 (27.4 %)
Smoking No	98 (47.1 %)
Yes	78 (37.5 %)
Yes (>20 cigarettes/day)	31 (14.9 %)
Recurrency No	143 (69.1 %)
Yes	64 (30.9 %)
HPV DNA No	121 (58.5 %)
HPV 16/18	84 (40.5 %)
HPV 33	01 (0.5 %)
HPV 58	01 (0.5 %)
High-risk HPV mRNA ISH No	131 (63.0 %)
Yes	76 (37.0 %)
p16 IHC Negative/Low	109 (52.4 %)
Positive	98 (47.1 %)

originally developed by Wang et al. [14] on whole-slide images, demonstrated robust generalisation to the Norway TMA cohort. KM survival analysis confirmed that HPV detection by all methods predicted superior 5-year OS (p<0.001 for DNA, p16 IHC, mRNA ISH, IHC+ISH combination, and AI detection) (Fig. 2B-F). Multivariate Cox regression demonstrated that the prognostic value of each test remained significant after adjusting for age, smoking, alcohol consumption, stage and recurrence (p<0.05 for all). ROC analysis showed high concordance across all detection methods (Fig. 2G). Using HPV DNA PCR as the reference standard, p16 IHC, mRNA ISH and the multiplex combination achieved areas under the curve (AUC) of 0.855, 0.938 and 0.929, respectively. The AI model yielded an AUC of 0.793 (95% CI: 0.698–0.888), with a sensitivity of 68.6% and specificity of 90%. Notably, mRNA ISH alone achieved the highest Youden index (Table S1). While the AI model’s AUC was lower than other detection methods, its ability to utilise routine H&E sections without consuming additional tissue, presents a distinct advantage for small biopsies or limited archival material.

Discordant cases between HPV detection methods

Across the various detection modalities, excellent concordance was observed in 87% (180/207) of cases; 74 were consistently HPV-positive and 106 consistently negative. Discordance was identified in 27 cases (13%), the majority of which (n=14; 6.7% of the total cohort) were p16-positive but mRNA- and DNA-negative, representing false-positive p16 immunostaining. Conversely, false-negative p16 staining was rare, occurring in only one case (0.5%) that remained mRNA- and DNA-positive. Two cases (1.0%) were negative for both p16 and mRNA but DNA-positive, potentially indicating latent HPV infection. A single case (0.5%) was positive for both mRNA and p16 but DNA-negative. Notably, nine cases (10.5% of DNA-positive samples) were p16-positive but mRNA-negative, categorising them as transcriptionally inactive HPV HNSCC (Fig. 3A-C). Survival analysis of these discordant groups

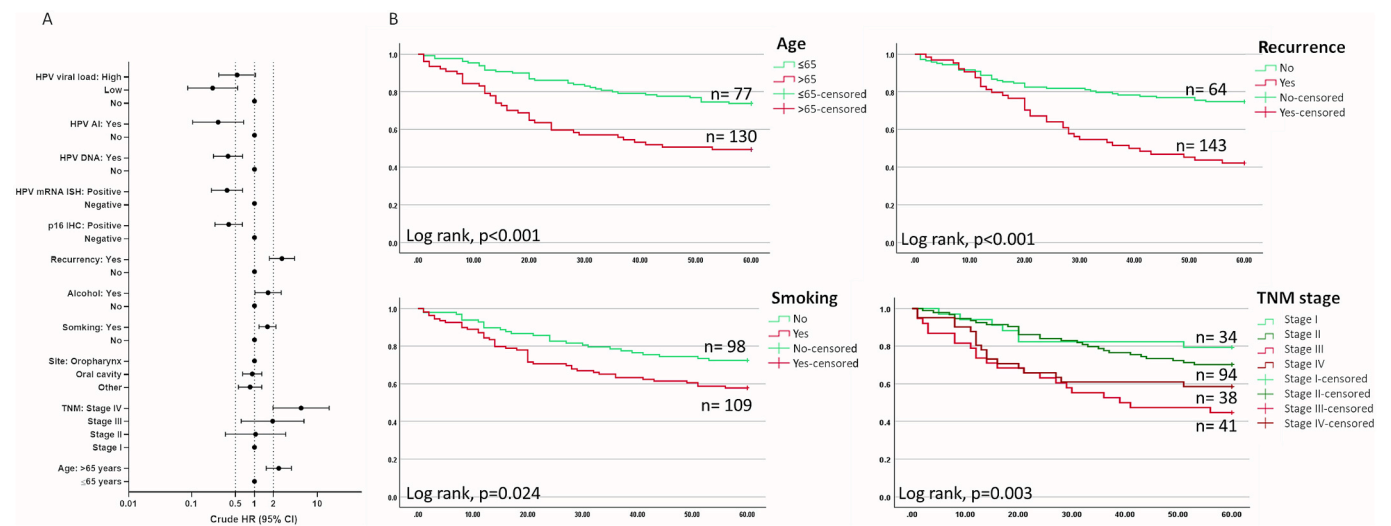


Fig. 1. Prognostic impact of clinicopathological and HPV-related parameters on overall survival (OS) in HNSCC. (A) Forest plot showing univariate Cox proportional hazards for 5-year OS in the HNSCC cohort. Hazard ratios (HRs) with 95 % confidence intervals (CIs) are shown for age, tumour stage, lymph node status, smoking, alcohol consumption, recurrence, and HPV status determined by different detection methods. (B) Kaplan-Meier survival curves illustrating 5-year OS stratified by selected prognostic factors: age at diagnosis (<65 vs ≥ 65 years); recurrence status; lymph node metastasis; and tumour stage. Numbers indicate the number of patients in each group. Survival differences were assessed using the log-rank (Mantel-Cox) test, with corresponding p-values displayed in each panel.

Table 2
Multivariate Cox regression analysis of significant parameters in univariate Cox regression for overall survival. Adjusted for tumour stage and age. Significant values are presented in bold.

Parameters	Hazard Ratio	95 % CI	p value
Smoking (yes)	1.37	0.846–2.219	0.201
Alcohol (yes)	1.602	0.99–2.591	0.055
Recurrence (yes)	2.045	1.276–3.280	0.003
HPV DNA PCR (yes)	0.566	0.326–0.984	0.044
p16 IHC (yes)	0.498	0.300–0.829	0.007
HPV mRNA ISH (yes)	0.565	0.314–1.01	0.05
Combined mRNA ISH + p16 IHC	0.578	0.322–1.04	0.06

revealed that p16-positive cases lacking HPV DNA or mRNA (p16+/HPV-) had significantly worse overall survival compared to truly HPV-positive (p16+/HPV+) cases (Log-rank, p = 0.038; Fig. 3D).

HPV viral load and patient survival

HPV viral load (VL) was quantified for all E6/E7 mRNA ISH-positive cases. The density of positive cells, normalised per μm^2 , ranged from 2.52 to 92.60 (mean: 30.43 ± 21.11 ; median: 28.96; Fig. 4A-B). Kaplan-Meier analysis demonstrated that a high VL (>30.43 cells/ μm^2) was significantly associated with increased 5-year mortality (p = 0.024; Fig. 4C). Furthermore, multivariate Cox regression identified high VL as an independent predictor of death in HPV-positive OPSCC after adjusting for age, smoking, alcohol consumption, stage, and recurrence (p = 0.03). These findings suggest that VL is a clinically relevant biomarker for risk stratification and may serve as a tool to guide treatment de-escalation in HPV-positive OPSCC.

Discussion

The increasing incidence of HPV-induced HNSCC necessitates robust and accurate detection methods for prognosis and treatment planning.

While various HPV detection techniques exist, including DNA-based PCR, p16 IHC and high-risk HPV E6/E7 mRNA ISH, the integration of AI-based prediction from routine H&E-stained sections represents a novel diagnostic frontier. This study compared these diverse modalities, evaluating the feasibility of multiplex IHC-ISH and AI-based analysis on archival OPSCC tissue, while exploring the prognostic significance of viral load quantified via ISH.

The development of a multiplex IHC-ISH approach was motivated by the clinical requirement for improved screening of transcriptionally active HPV-infection. Precise identification of active infection is critical, as it directly informs the selection of patients who may benefit from treatment de-escalation or more specific therapeutic interventions. Despite the generally superior prognosis of HPV-positive OPSCC, a significant subset of patients experiences unfavourable outcomes. Consequently, there is an urgent need for biomarkers capable of refined risk stratification. We propose that the presence of active viral transcription and the specific VL may provide the granular data necessary for informed, personalised treatment decisions that are currently lacking in standard OPSCC management. By integrating molecular, histopathological, and computational approaches, this study addresses both diagnostic accuracy and the biological relevance of HPV in clinical risk stratification.

Clinical implications of multiplex HPV detection

The multiplex p16 IHC-mRNA ISH assay represents a refined approach to identifying high-risk HPV in OPSCC. While p16 IHC is a highly sensitive and cost-effective surrogate for transcriptionally active high-risk HPV infection [20,21], overexpression is not pathognomonic for HPV-driven carcinogenesis [22]. We successfully optimised a multiplex protocol for the simultaneous detection of E6/E7 transcripts and p16 protein on a single slide. A key practical advantage of this method is its seamless integration into routine pathology workflows, utilising FFPE material and standard chromogenic readouts via bright-field microscopy. Our findings demonstrate that this approach is comparable to the HPV DNA “gold standard”. While a previous study involving 38 cases suggested that such multiplexing could enhance OPSCC stratification [23], our larger cohort indicates that the multiplex method does not significantly outperform individual DNA, p16 IHC, or mRNA ISH tests in predicting 5-year OS. Nevertheless, while the

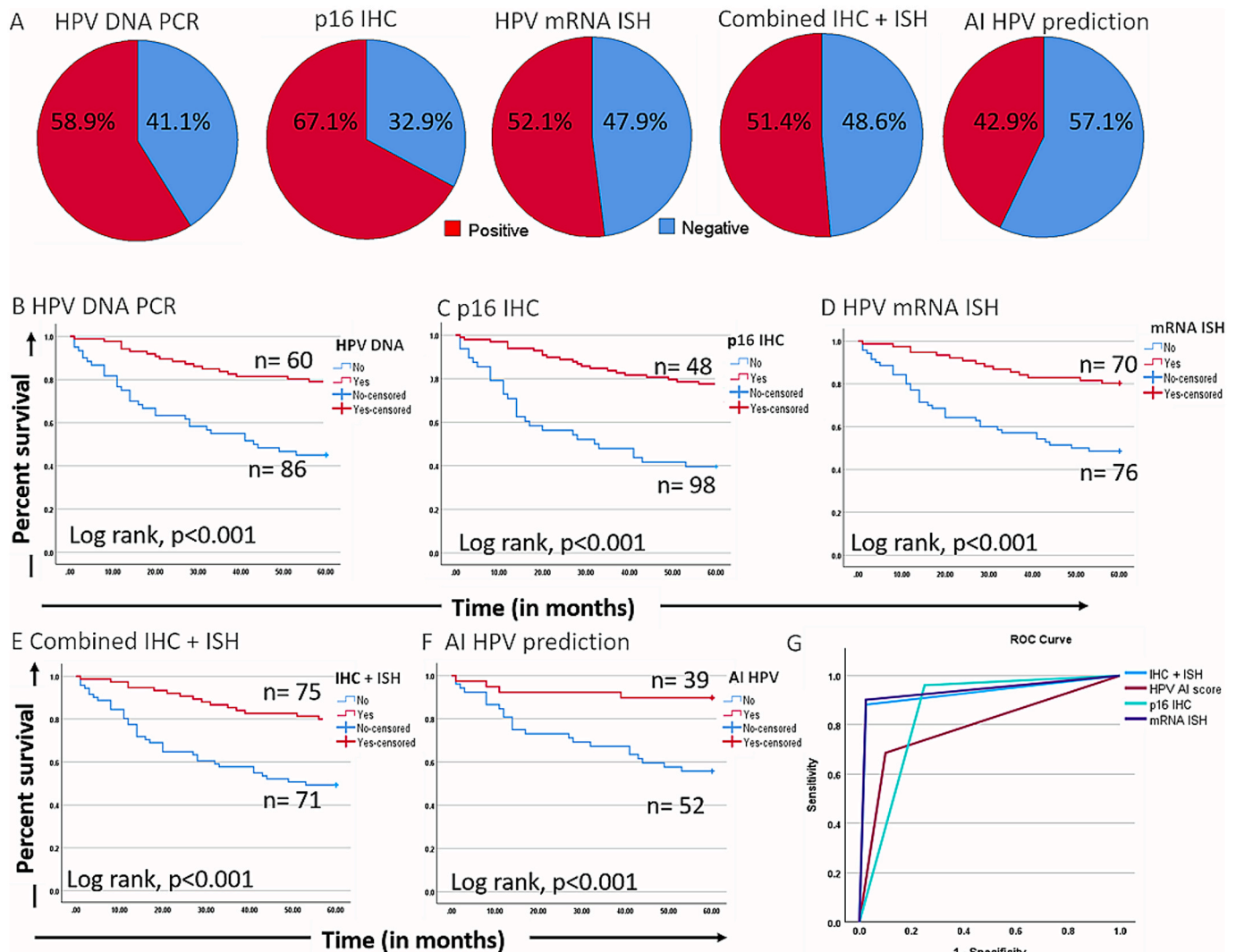


Fig. 2. Comparison of HPV detection methods and their prognostic performance in OPSCC. (A) Pie charts illustrating the proportion of HPV-positive (red) and HPV-negative (blue) cases in the OPSCC cohort as determined by different detection methods: HPV DNA PCR, p16 IHC, HPV E6/E7 mRNA CISH, combined p16 IHC and HPV E6/E7 mRNA ISH, and AI-based prediction. (B–F) Kaplan-Meier survival analyses showing 5-year overall survival (OS) stratified by HPV status according to each detection method: HPV DNA PCR, p16 IHC, HPV E6/E7 mRNA CISH, combined p16 IHC and HPV E6/E7 mRNA ISH, and AI-based HPV prediction. Numbers indicate the number of patients in each group. Survival differences were assessed using the log-rank (Mantel-Cox) test, with p-values shown in each panel. (G) Receiver operating characteristic (ROC) curve analysis comparing the performance of different HPV detection methods, using HPV DNA PCR as the reference standard. The area under the curve (AUC) illustrates the discriminatory ability of each method. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

predictive efficiency for overall survival may not be superior to individual tests, the multiplex assay remains clinically significant. It provides biologically definitive confirmation of active viral integration and transcription, potentially reducing the risk of misclassification in clinical decision-making, a critical factor as the field moves towards treatment de-escalation for HPV-positive patients.

Applying AI for clinical HPV stratification

The application of deep learning to predict HPV status from routine histology represents an emerging approach that complements molecular testing. In our external validation, the AI model achieved an AUC of 0.793 and demonstrated significant prognostic stratification ($p = 0.002$ for 5-year OS). While this performance is lower than that of more sophisticated molecular methods, several factors contextualize this finding.

Firstly, the model was originally optimized for whole-slide images but was applied here to 1 mm² TMA cores, which provide limited

architectural context and reduced sampling of tumour heterogeneity. Secondly, this study represents a rigorous external validation across a different population and tissue format, demonstrating model robustness despite a clear domain shift. Most significantly, the AI approach offers distinct practical advantages: (i) it requires only standard H&E sections without consuming additional tissue, which is critical for small biopsies; (ii) it provides rapid results without specialised molecular laboratory infrastructure; and (iii) it captures biologically relevant morphological features, as evidenced by its independent prognostic value.

We propose that AI-based HPV prediction serves not as a replacement for molecular confirmation, but rather as a first-pass screening tool or triage mechanism. In resource-limited settings or when tissue is scarce, the high sensitivity (93.8%) makes it suitable for ruling out HPV-negative cases, while positive predictions could subsequently be confirmed with molecular testing. The significant association with clinical outcome suggests that the model captures morphological surrogates of HPV-driven biology beyond the limits of conventional H&E histopathological assessment. Future work should validate the model on

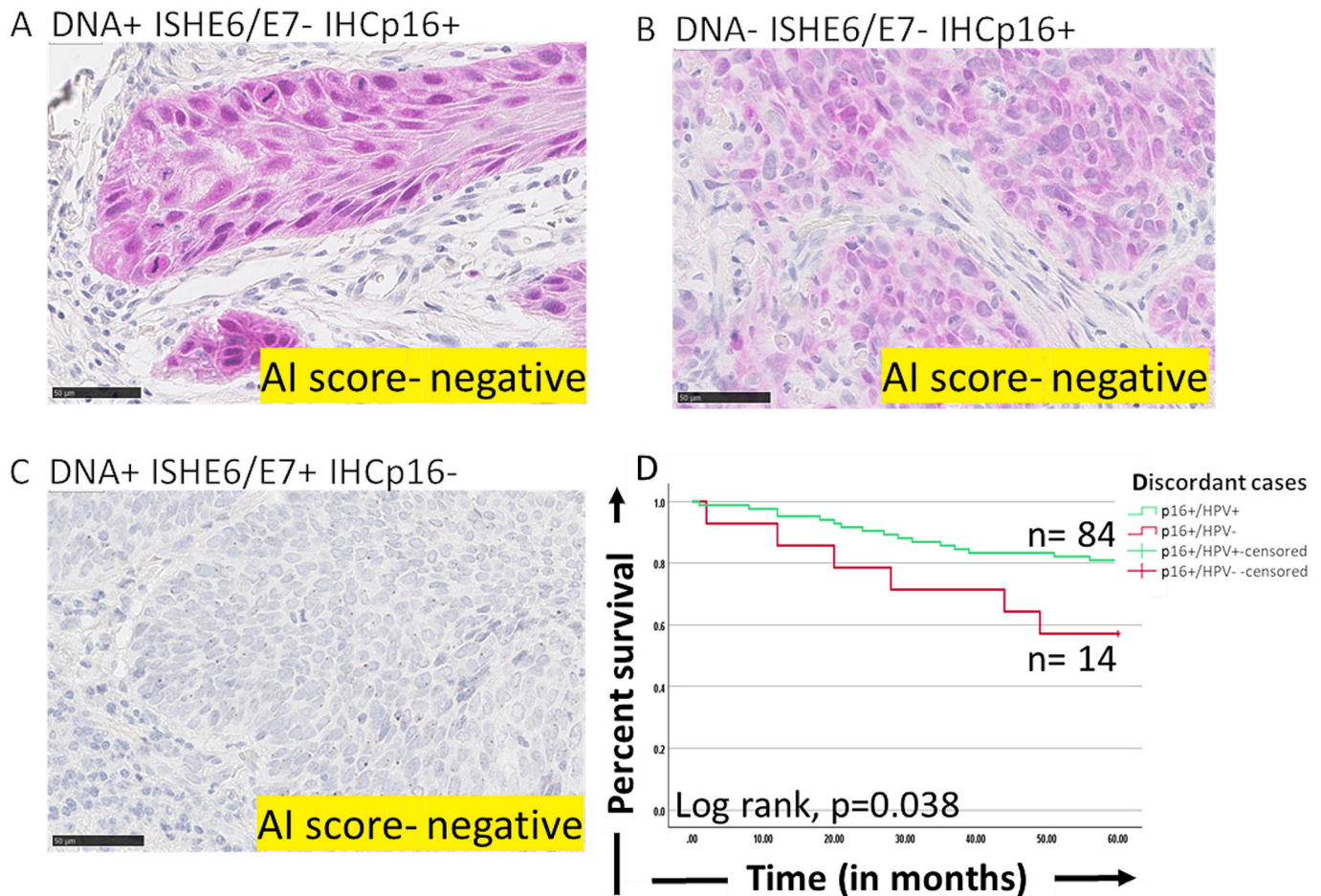


Fig. 3. Discordant HPV detection cases and associated survival outcome. (A-C) Representative haematoxylin and eosin (H&E)-stained images from discordant cases. Scale bars, 50 μ m. (D) Kaplan-Meier analysis of 5-year OS comparing discordant p16-positive/HPV-negative cases (p16+/HPV-; n = 14) with concordant HPV-positive cases (p16+/HPV+; n = 84). Survival differences were assessed using the log-rank (Mantel-Cox) test.

full tissue sections from this cohort to determine whether performance approximates molecular assays when complete architectural information is available. Additionally, integration of AI predictions with clinical parameters may yield multi-modal risk stratification tools that combine computational pathology with traditional biomarkers.

Clinical significance of HPV transcriptional activity

A recent study demonstrated that E6 mRNA expression is ubiquitous in HPV16 DNA-medium/high OPSCCs but absent in DNA-low cases [24]. This supports our observation of discordant cases and highlights the critical biological distinction between viral presence and transcriptional activity. Together, these findings underscore the importance of assessing not only the presence of HPV, but also the magnitude of viral oncogene expression within tumour cells.

ISH for E6/E7 enables refined stratification of HPV-positive OPSCC patients by allowing for direct assessment of transcriptional VL. While traditionally VL is measured via DNA-based techniques such as qPCR or digital PCR (where high VL often correlates with improved prognosis in cervical, anal, and oropharyngeal cancers (5, 20, 22-24), our results present a notable contradiction. In our cohort, a high transcriptional VL (>30 positive cells/ μ m²) was associated with significantly lower 5-year survival.

This discrepancy likely reflects the difference between genomic presence (DNA copy number) and active oncogenic signalling (mRNA levels). High transcriptional activity may indicate sustained oncogenic drive, increased tumour aggressiveness, or reduced dependence on host

regulatory mechanisms, thereby identifying a biologically distinct high-risk subgroup within the HPV-positive population. Although the use of ISH for VL quantification is relatively nascent [21], our findings suggest it may offer superior biological relevance compared to DNA-based metrics.

Limitations

This study has several limitations. First, as a single-centre retrospective analysis spanning nearly two decades, our findings may be subject to temporal and selection biases arising from evolving diagnostic practices and treatment protocols. Second, the RNA in situ hybridization assay targeted only selected high-risk HPV subtypes, potentially missing individually very uncommon subtypes [25]. Third, HPV DNA positivity was restricted to oropharyngeal tumours, limiting the generalisability of these findings to non-oropharyngeal HNSCC subsites. Fourth, the use of archival formalin-fixed paraffin-embedded (FFPE) tissue and Tissue Microarray (TMA) cores (while practically advantageous), may under sample tumour heterogeneity and introduce pre-analytical variability inherent to older material. Specifically, the external validation of the AI model on TMA cores rather than WSIs represents a significant domain shift; this reduced architectural context likely contributed to the model's lower AUC relative to molecular assays.

Furthermore, viral load was quantified using a threshold-based digital pipeline and dichotomised based on the cohort mean. Consequently, these results may be sensitive to specific parameter settings, cut-point selection, and observer-dependent annotations. Finally, the small

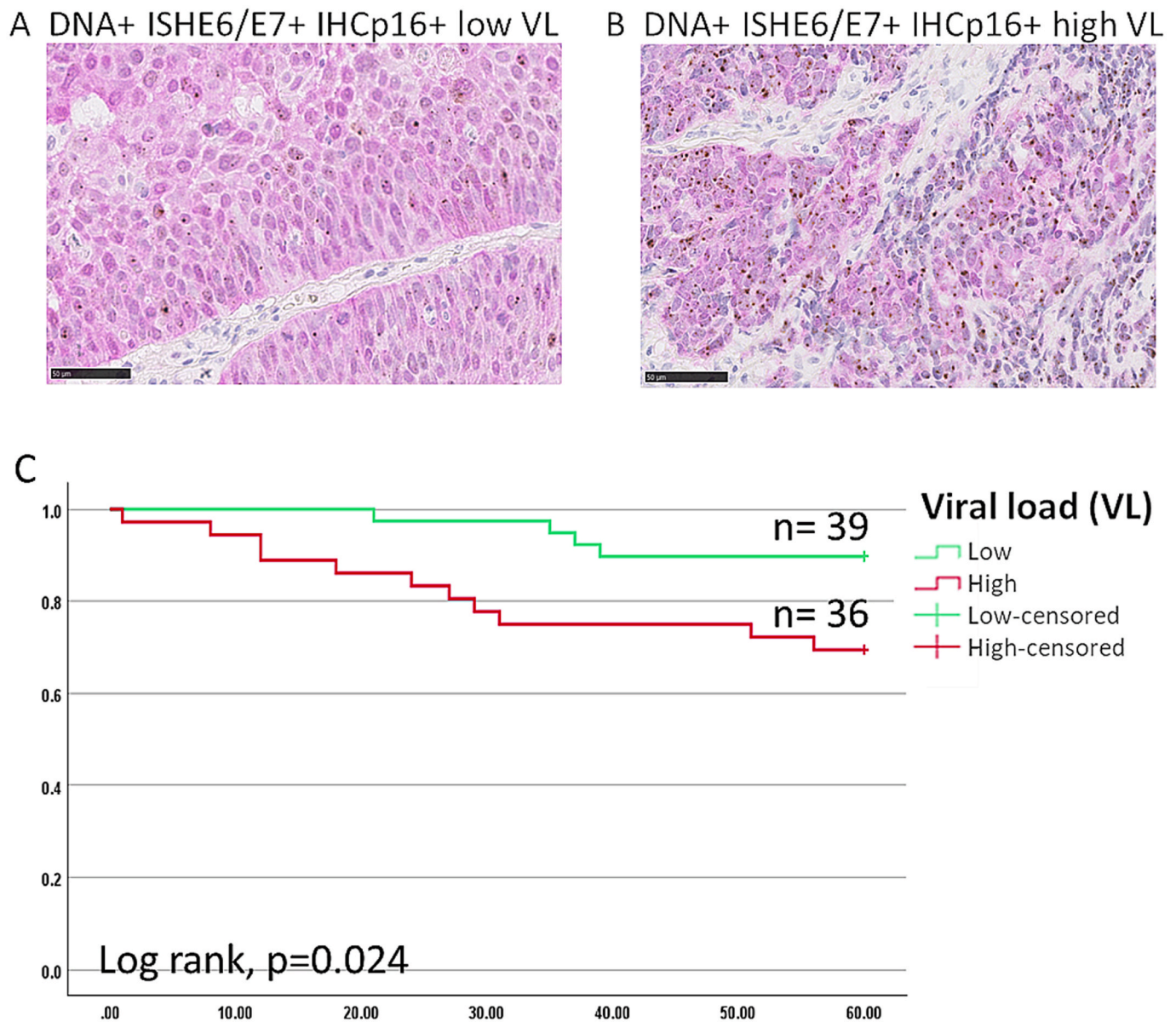


Fig. 4. HPV viral load assessed by E6/E7 mRNA ISH and association with overall survival (OS). (A–B) Representative chromogenic in situ hybridization (ISH) images showing HPV E6/E7 mRNA expression in OPSCC with low viral load (A) and high viral load (B), illustrating differences in the number of transcriptionally active HPV-positive tumour cells. Brown punctate signals indicate E6/E7 mRNA expression. Scale bars, 50 μm . (C) Kaplan-Meier analysis of 5-year OS in HPV-positive OPSCC stratified by viral load. Patients were categorized into low and high viral load groups based on the mean E6/E7 mRNA-positive cell density (30.43 cells/ μm^2). High viral load was associated with significantly reduced OS. Survival differences were assessed using the log-rank (Mantel-Cox) test.

sample sizes within certain discordant HPV-status subgroups limit statistical precision, necessitating cautious interpretation of these specific clinical outcomes.

Conclusions

In summary, this study demonstrates that multiple HPV detection modalities (including DNA PCR, p16 IHC, E6/E7 mRNA ISH, and AI-based prediction) provide valuable and largely concordant data for identifying HPV-driven OPSCC. While molecular assays remain the most accurate diagnostic approach, the multiplex IHC–ISH assay offers biologically definitive confirmation of transcriptionally active infection. Furthermore, AI-based analysis shows significant promise as a practical, tissue-sparing, low-cost and rapid screening tool in resource-limited settings. Crucially, viral load from mRNA ISH emerged as an independent prognostic marker, with high transcriptional activity identifying a

higher risk group. This underscores the clinical necessity of assessing the magnitude of viral load rather than mere viral presence. Together, these complementary methods enable more granular risk stratification in HNSCC and highlight the potential of integrating digital pathology with molecular diagnostics to inform future personalised treatment strategies.

Funding information

This work was supported by the Research Council of Norway through its Centres of Excellence funding scheme, (Grant No. 22325), The Western Norway Regional Health Authority (Helse Vest Grant No. 911902/2013, 912260/2019 and F-13105/2023), and The Norwegian Centre for International Cooperation in Education (project number CPEA-LT-2016/10106).

CRediT authorship contribution statement

Stian Tornaas: Writing – review & editing, Writing – original draft, Methodology, Investigation, Formal analysis, Data curation. **Ruoyu Wang:** Writing – review & editing, Validation, Software, Methodology, Formal analysis. **Svein Erik Emblem Moe:** Writing – review & editing, Project administration, Investigation, Formal analysis, Data curation. **Siren Fromreide:** Writing – review & editing, Methodology, Investigation, Formal analysis. **Rammah Elnour:** Writing – review & editing, Investigation, Formal analysis, Data curation. **Åsa Karlsdottir:** Writing – review & editing, Project administration, Formal analysis. **Marianne Brydøy:** Writing – review & editing, Project administration, Investigation. **Hans Jørgen Aarstad:** Writing – review & editing, Validation, Supervision, Project administration. **Olav Karsten Vintermyr:** Writing – review & editing, Supervision, Project administration, Investigation. **Nasir Rajpoot:** Writing – review & editing, Visualization, Supervision, Methodology, Formal analysis. **Syed Ali Khurram:** Writing – review & editing, Validation, Supervision, Resources, Project administration. **Daniela Elena Costea:** Writing – review & editing, Writing – original draft, Supervision, Project administration, Funding acquisition, Data curation, Conceptualization. **Harsh Nitin Dongre:** Writing – review & editing, Writing – original draft, Validation, Supervision, Project administration, Funding acquisition, Formal analysis, Data curation, Conceptualization.

Ethics approval and consent to participate

The study was performed in accordance with the Declaration of Helsinki. It was approved by the Regional Ethical Committee in West Norway (REK Vest 2010/481 and 680158). The AI-HPV model originally developed on WSIs had ethical approval from Health and Care Research Wales, UK (Sheffield REC reference-20/WS/0017). Patients were enrolled in the study only after they signed a written informed consent. Obtaining informed consent was waived from death patients.

Consent for publication

The manuscript does not contain any single individual person's data in any form.

Reporting guidelines

This is a tumour marker prognostic study and therefore adhered to the REMARK criteria. This study also included external validation of prediction model involving artificial intelligence and hence adhered to the TRIPOD+AI guidelines.

Generative AI and AI-assisted technologies were NOT used in the preparation of this work.

Declaration of competing interest

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

Acknowledgements

The authors would like to thank to study nurse Ellen Ørnen for the help in recruiting the patients and to pathology technician Bendik Nordanger for the help in preparing the TMA.

Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.oraloncology.2026.108123>.

References

- [1] Menezes FDS, Fernandes GA, Antunes JLF, Villa LL, Toporcov TN. Global incidence trends in head and neck cancer for HPV-related and -unrelated subsites: a systematic review of population-based studies. *Oral Oncol* 2021;115:105177.
- [2] McDermott JD, Bowles DW. Epidemiology of Head and Neck Squamous Cell Carcinomas: Impact on Staging and Prevention strategies. *Curr Treat Options Oncol* 2019;20(5):43.
- [3] You EL, Henry M, Zeitouni AG. Human papillomavirus-associated oropharyngeal cancer: review of current evidence and management. *Curr Oncol* 2019;26(2):119–23.
- [4] Ljokjel B, Haave H, Lybak S, Aarstad HH, Karlsdottir A, Vintermyr OK, et al. The impact of HPV infection, smoking history, age and operability of the patient on disease-specific survival in a geographically defined cohort of patients with oropharyngeal squamous cell carcinoma. *Acta Otolaryngol* 2014;134(9):964–73.
- [5] Villarme A, Ebran N, Pace-Loscos T, Schiappa R, Mignot A, Bozec A, et al. Prognostic impact of intra tumoral HPV-16 viral load in oropharyngeal squamous cell carcinomas. *Oral Oncol* 2024;159:107082.
- [6] Rietbergen MM, Brakenhoff RH, Bloemena E, Witte BI, Snijders PJ, Heideman DA, et al. Human papillomavirus detection and comorbidity: critical issues in selection of patients with oropharyngeal cancer for treatment De-escalation trials. *Ann Oncol* 2013;24(11):2740–5.
- [7] Wagner S, Prigge ES, Wuerdemann N, Reder H, Bushnak A, Sharma SJ, et al. Evaluation of p16(INK4a) expression as a single marker to select patients with HPV-driven oropharyngeal cancers for treatment de-escalation. *Br J Cancer* 2020;123(7):1114–22.
- [8] Robinson M, Schache A, Sloan P, Thavaraj S. HPV specific testing: a requirement for oropharyngeal squamous cell carcinoma patients. *Head Neck Pathol* 2012;6 Suppl 1(Suppl 1):S83–90.
- [9] Prigge ES, Arbyn M, von Knebel DM, Reuschenbach M. Diagnostic accuracy of p16 (INK4a) immunohistochemistry in oropharyngeal squamous cell carcinomas: a systematic review and meta-analysis. *Int J Cancer* 2017;140(5):1186–98.
- [10] Albers AE, Qian X, Kaufmann AM, Coords A. Meta analysis: HPV and p16 pattern determines survival in patients with HNSCC and identifies potential new biologic subtype. *Sci Rep* 2017;7(1):16715.
- [11] von Knebel DM. The causal role of human papillomavirus infections in non-anogenital cancers. It's time to ask for the functional evidence. *Int J Cancer* 2016;139(1):9–11.
- [12] Smeets SJ, Hesselink AT, Speel EJ, Haesevoets A, Snijders PJ, Pawlita M, et al. A novel algorithm for reliable detection of human papillomavirus in paraffin embedded head and neck cancer specimen. *Int J Cancer* 2007;121(11):2465–72.
- [13] Wang R, Khurram SA, Walsh H, Young LS, Rajpoot N. A Novel Deep Learning Algorithm for Human Papillomavirus Infection Prediction in Head and Neck Cancers using Routine Histology Images. *Mod Pathol* 2023;36(12):100320.
- [14] Altman DG, McShane LM, Sauerbrei W, Taube SE. Reporting Recommendations for Tumor Marker Prognostic Studies (REMARK): explanation and elaboration. *PLoS Med* 2012;9(5):e1001216.
- [15] Dongre HN, Elnour R, Tornaas S, Fromreide S, Thomsen LCV, Kolseth IBM, et al. TP53 mutation and human papilloma virus status as independent prognostic factors in a norwegian cohort of vulva squamous cell carcinoma. *Acta Obstet Gynecol Scand* 2024;103(1):165–75.
- [16] 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):16878.
- [17] Pocock J, Graham S, Vu QD, Jahanifar M, Deshpande S, Hadjigeorgiou G, et al. TIAToolbox as an end-to-end library for advanced tissue image analytics. *Communications Medicine* 2022;2(1):120.
- [18] Macekno M, Niethammer M, Marron JS, Borland D, Woosley JT, Xiaojun G, et al., editors. A method for normalizing histology slides for quantitative analysis. 2009 IEEE International Symposium on Biomedical Imaging: From Nano to Macro; 2009 28 June–1 July 2009.
- [19] Collins GS, Moons KGM, Dhiman P, Riley RD, Beam AL, Van Calster B, et al. TRIPOD+AI statement: updated guidance for reporting clinical prediction models that use regression or machine learning methods. *BMJ* 2024;385:e078378.
- [20] Schache AG, Liloglou T, Risk JM, Jones TM, Ma XJ, Wang H, et al. Validation of a novel diagnostic standard in HPV-positive oropharyngeal squamous cell carcinoma. *Br J Cancer* 2013;108(6):1332–9.
- [21] Jordan RC, Lingen MW, Perez-Ordóñez 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] Prigge ES, von Knebel DM, Reuschenbach M. Clinical relevance and implications of HPV-induced neoplasia in different anatomical locations. *Mutat Res Rev Mutat Res* 2017;772:51–66.
- [23] Zito Marino F, Ronchi A, Stilo M, Cozzolino I, La Mantia E, Colacurci N, et al. Multiplex HPV RNA in situ hybridization/p16 immunohistochemistry: a novel approach to detect papillomavirus in HPV-related cancers. a novel multiplex ISH/IHC assay to detect HPV. *Infect Agent Cancer* 2020;15:46.
- [24] Hashida Y, Higuchi T, Matsumoto S, Iguchi M, Murakami I, Hyodo M, et al. Prognostic significance of human papillomavirus 16 viral load level in patients with oropharyngeal cancer. *Cancer Sci* 2021;112(10):4404–17.
- [25] Araújo IS, Vasconcelos RAO, Gonçalves de Paiva JP, Lopes VSA, Santos-Silva AR, Lopes MA, et al. Global HPV genotype distribution and coinfections in oropharyngeal squamous cell carcinoma: systematic review and meta-analysis. *Oral Surg Oral Med Oral Pathol. Oral Radiol* 2026.