PREDICTING HPV-ASSOCIATION USING REGULAR H&E SLIDES CAN IDENTIFY SUBGROUPS OF PATIENTS WITH FAVORABLE PROGNOSIS AT A HIGHLY DETAILED LEVEL

Jens P Klussmann1, Nora Wuerdemann2, Steffen Wagner2, Imke Demers1, Christopher Kopp1, Arthur Charpentier1, Shachi J Sharma1, Julie George1, Jochen Heiss1, Reinhard Bütter1, Ernst J Spee1, Christian Reinhardt2, Sebastian Klepp1,6,8
1 Department of Otolaryngology, Head and Neck Surgery, University Hospital Cologne; Cologne, Germany; 2 Department of Pathology, University Hospital Cologne; Cologne, Germany; 3 Department of Pathology, Maastricht University Medical Center; Maastricht, The Netherlands; 4 Department of Otolaryngology, Head and Neck Surgery, University Hospital Heidelberg; and German Cancer Research Center (DKFZ), Heidelberg, Germany; 5 Institute of Pathology, Medical Faculty, University Hospital Cologne, Cologne, Germany; 6 Department of Hematology and Stem Cell Transplantation, University St. Augustin-Krankenhaus, University Hospital Essen, Essen, Germany.

Introduction

Oropharyngeal squamous cell cancer (OPSCC) related to Human Papilloma Virus (HPV) is a subgroup of head and neck cancer that can be identified by p16 immunohistochemistry and HPV-DNA testing. Although the prognosis is generally favorable, there has been a lack of success in implementing therapy de-escalation, owing to the heterogeneity of the disease. This underscores the need for precise biomarkers to facilitate patient stratification. Recently, we developed a deep learning-based approach to predict the HPV association using scans of regular H&E stains. In this study, our objective is to develop a modular and computationally efficient algorithm designed to stratify OPSCC patients with higher accuracy compared to conventional HPV testing.

Results

Comparing the algorithm with HPV status, it showed good overall performance (AUROC=0.83; 95% CI=0.77-0.9, Fig 1D). In a subset of the validation cohort (n=339) the implementation of a fixed threshold for filtering resulted in increased AUROC to 0.88 (Fig 1E).

The algorithm was compared to the gold standard of HPV-testing in terms of its prognostic relevance and produced better results than the HPV test, indicated by its higher likelihood-ratio test value (LR, 49.23, p<0.001), higher concordance index (0.71), and higher 5-year overall survival rate (OS, 96%, 95% CI=90-100%, Fig 3A).

For the test cohort, filtered for primary tumors, there was a strong association between the combined score and overall survival by using a Cox proportional hazards model (HR=62.26, p < 0.001, employing a chi-square distribution, n=531; Fig 2A). Dividing the patients into three separate groups by multivariate analysis (high HR=0.17 [95% CI=0.10-0.29]; intermediate HR=0.49 [95% CI=0.37-0.66]; low: reference) with distinct five-year survival rates (high: 85% [95% CI=77–93%], intermediate: 56% [95% CI=50–63%], low: 34% [95% CI=25–45%]) demonstrated good discrimination (Fig 3A).

Particularly patients with early-stage disease (stage I/II, UICC 8th, n = 294) may qualify for potential treatment de-escalation strategies and in this subgroup the 5-year survival rates were 90% [95% CI = 82–99%] for patients classified as high, 73% [95% CI=65–82%] for patients classified as intermediate and 48% [95% CI = 35–65%] classified as low (Fig 3B).

Selecting only HPV-positive cases and using the same three-tier threshold patients could be stratified in a five-year overall survival rate of 92% [95% CI = 84–100%] for patients within the high group, 83% [95% CI=76–90%] for patients in the intermediate group and 63% [95% CI=47–84%] for the low group (Fig 3D).

Discussion

Our algorithm can identify patients with OPSCC who have a favorable prognosis using standard hematoxylin and eosin (H&E) histologic slides. In multiple scenarios, our stratification method performs better than the gold standard (p16/HPV-DNA) and could potentially be used to select patients for therapy de-escalation strategies.

Material & Methods

In this retrospective, multi-institutional study 906 patients were enrolled (Fig 1A). A deep learning algorithm was developed to analyze standard H&E stains for the calculation of a patient-level score associated with prognosis, comparing it to combined HPV-DNA and p16-status. A Feature Pyramid Network (FPN) with a ResNet-18 encoder was constructed to perform semantic segmentation of viable tumor regions. The extracted tumor tiles were analyzed on their HPV association using an additional ResNet-18 network. A training dataset of 267 patients from two centers and one database (Fig 1B) was used. The sensitivity of the model was maximized by applying a fixed threshold on the variance of the probability of the HPV-positive class. To resolve the prognostic value of predicting HPV association, we used a Cox proportional hazards models.

Fig. 1: A: Study population was derived from four centers and one database (n = 906); HPV association was defined as either dichotomous HPV-DNA and p16; B: Training cohort; C: Validation cohort; D: Area under the receiver operator curve (AUC) for six different patient populations; E: AUCOR for cases that were filtered by the threshold of variance.

Fig. 2: A: Hazard ratio plot of the predicted HPV association and overall survival. B: Histological image of three cases with corresponding classes of the combined score.