

Research Article

Development of a Pathomics-Based Model for the Prediction of Malignant Transformation in Oral Leukoplakia

Xinjia Cai^{a,b}, Long Li^c, Feiyan Yu^d, Rongrong Guo^d, Xuan Zhou^a, Fang Zhang^{d,*}, Heyu Zhang^{e,*}, Jianyun Zhang^{a,b,*}, Tiejun Li^{a,b,*}

^a Department of Oral Pathology, Peking University School and Hospital of Stomatology, National Center of Stomatology, National Clinical Research Center for Oral Diseases, National Engineering Research Center of Oral Biomaterials and Digital Medical Devices, Beijing, China; ^b Research Unit of Precision Pathologic Diagnosis in Tumors of the Oral and Maxillofacial Regions, Chinese Academy of Medical Sciences, Beijing, China; ^c Hunan Key Laboratory of Oral Health Research, Xiangya Stomatological Hospital, Central South University, Changsha, China; ^d Department of Oral Medicine, Shanxi Province Key Laboratory of Oral Diseases Prevention and New Materials, Shanxi Medical University School and Hospital of Stomatology, Taiyuan, China; ^e Central Laboratory, Peking University School and Hospital of Stomatology, Beijing, China

ARTICLE INFO

Article history:

Received 19 February 2023

Revised 14 April 2023

Accepted 2 May 2023

Keywords:

deep learning
histopathology
malignant transformation
oral leukoplakia
oral squamous cell carcinoma

ABSTRACT

Accurate prognostic stratification of oral leukoplakia (OLK) with risk of malignant transformation into oral squamous cell carcinoma is crucial. We developed an objective and powerful pathomics-based model for the prediction of malignant transformation in OLK using hematoxylin and eosin (H&E)-stained images. In total, 759 H&E-stained images from multicenter cohorts were included. A training set (n = 489), validation set (n = 196), and testing set (n = 74) were used for model development. Four deep learning methods were used to train and validate the model constructed using H&E-stained images. Pathomics features generated through deep learning combined with machine learning algorithms were used to develop a pathomics-based model. Immunohistochemical staining of Ki67, p53, and PD-L1 was used to interpret the black box of the model. Pathomics-based models predicted the malignant transformation of OLK (validation set area under curve [AUC], 0.899; testing set AUC, 0.813) and significantly identified high-risk and low-risk populations. The prediction performance of malignant transformation from dysplasia grading (validation set AUC, 0.743) was lower than that of the pathomics-based model. The expressions of Ki67, p53, and PD-L1 were correlated with various pathomics features. The pathomics-based model accurately predicted the malignant transformation of OLK and may be useful for the objective and rapid assessment of the prognosis of patients with OLK.

© 2023 United States & Canadian Academy of Pathology. Published by Elsevier Inc. All rights reserved.

Introduction

Oral squamous cell carcinoma (OSCC) is the most common oral and maxillofacial malignancy, accounting for more than 90% of

oral malignancies.¹ The incidence and mortality rates of oral cancer are increasing worldwide.² Although treatment strategies for OSCC are under development, the 5-year survival rate of advanced OSCC is only approximately 40%.³ Early diagnosis of OSCC may thus help contribute to improved patient survival.

Oral potentially malignant disorders are a group of precursor lesions of OSCC with significant malignant potential.⁴ Oral leukoplakia (OLK) is one of the most common oral potentially malignant disorders and has a malignant transformation rate of approximately 10%.^{5,6} Early identification of OLK in patients at a

* Corresponding authors.

E-mail addresses: litiejun22@vip.sina.com (T. Li); jianyunz0509@aliyun.com (J. Zhang); zhangheyu1983@sina.cn (H. Zhang); zhangfangdoctor@sxmu.edu.cn (F. Zhang)



high risk for malignant transformation may allow for early monitoring of OSCC in these patients. Currently, the clinical predictor of malignant transformation in OLK is the histopathologic grading of oral epithelial dysplasia (OED) using hematoxylin and eosin (H&E)-stained sections. Moderate-to-severe dysplasia is associated with an increased risk of development of OSCC.⁵⁻⁷ Although the World Health Organization has proposed 2 grading systems to improve the assessment of OED grading,⁸ histopathologic grading by pathologists still yields inconsistent results.^{9,10} Furthermore, the absence or lower grading of OED does not indicate an absence of malignant risk. Chaturvedi et al⁷ reported that 39.6% of oral cancers originated from OLK without dysplasia. Given the low agreement and specificity of OED grading, the identification of predictive biomarkers for malignant transformation of OLK is critical.

Deep learning- and artificial intelligence-based approaches are being widely developed for the prediction, differential diagnosis, and clinical management of cancers¹¹⁻¹⁷ Deep learning can acquire the pathomics-related features of biopsies from H&E-stained histopathology slides.¹⁸ The low agreement of grading by pathologists may be avoided using artificial intelligence-based models, and the development of big data models may improve the prediction performance.

The aim of this study was to develop a deep learning model to predict the malignant transformation of OLK. This study included a cohort of 759 cases of OLK with a median follow-up of 75 months. The pathomics-based model was developed with a large number of H&E-stained histopathology images and comprehensively evaluated using internal and external data sets.

Materials and Methods

Study Design and Data Collection

We developed a pathomics-based model to predict the risk of malignant transformation of OLK using H&E-stained images from multicenter cohorts. All OLK cases were diagnosed from biopsy and reviewed by 2 experienced head and neck pathologists following the criteria drafted by the World Health Organization.^{4,8} We retrospectively collected 699 cases of OLK from Peking University Hospital of Stomatology and randomly divided them into a training set cohort (n = 489) and a validation set cohort (n = 210) with a 7:3 ratio. We also collected 74 cases of OLK as a testing set cohort, which included 42 cases from Xiangya Stomatological Hospital of Central South University and 32 cases from Shanxi Medical University Hospital of Stomatology. After screening the H&E-stained sections of all samples and excluding sections with faded and unclear H&E stain, 759 H&E-stained images were included, with 489 H&E-stained images in the training set, 196 H&E-stained images in the validation set, and 74 H&E-stained images in the test set. The median follow-up time of the entire cohort was 75 months. By reviewing medical records, we obtained patient information such as age, sex, lesion site, smoking history, and alcohol consumption. The data of malignant transformation were recorded by performing follow-up. Malignant transformation was defined as the occurrence of OSCC at the same or adjacent site at least 1 month after the diagnosis of OLK. This study was approved by the Institutional Review Board of all involved hospitals.

Immunohistochemical Staining and Evaluation

We randomly selected 94 formalin-fixed paraffin-embedded OLK samples from the validation set cohort for immunohistochemical

(IHC) staining. The samples were sliced into sections (3 μm thick) and mounted on adhesive slides. IHC staining was conducted using a fully automated IHC instrument (BOND III; Leica Biosystems Melbourne Pty Ltd). The primary antibodies were as follows: mouse anti-Ki67 monoclonal antibody (1:200 dilution, MXB), mouse anti-p53 monoclonal antibody (1:200 dilution, MXB), and rabbit anti-PD-L1 monoclonal antibody (1:200 dilution, MXB). Ki67, p53, and PD-L1 expressions were assessed by scoring the percentage of positive cells as follows: score 0 (0%), 1 (1%–25%), 2 (26%–50%), 3 (51%–75%), and 4 (76%–100%). IHC evaluation was performed by 2 experienced pathologists.

Image Acquisition and Preprocessing

All H&E-stained slides were scanned as whole slide imaging (WSI) digital slides using a NanoZoomer and exported to JPG format by NDPView2 software. Because the digitalized images were large, the images were cut into small image tiles (512 \times 512 pixels) by a nonoverlapping sampling at 20 \times magnification (0.5 $\mu\text{m}/\text{pixel}$). We excluded images with a ratio of bright pixels larger than 0.8. The small tiles underwent color normalization with the Reinhard method and were further normalized using z scores on red, green and blue channels to obtain a standard normal distribution of image intensities as inputs. Data augmentations included random horizontal and vertical flipping.

Model Development and Pathomics Features

Our deep learning pipeline presented 2 predictions: patch-level and WSI-level. Because of the large image size and heterogeneity, the WSI was first divided into small patches and then the patch likelihoods were aggregated in an ensemble algorithm to obtain the WSI-level prediction. During the patch-level prediction, 4 SOTA convolutional neural networks (ResNet50, ResNet101, InceptionV3, and Densenet121) in ImageNet classification competition were trained to compute the patch likelihood in which the patches were assigned with the label of the WSI. Softmax cross-entropy loss was used to optimize the network using a mini-batch gradient descent method. The initial parameters of the model were transferred using the ImageNet data to generalize the model across cohorts with a high degree of heterogeneity. Transfer learning was conducted by reusing the model weights in the patch-level discriminators and then fine-tuning the weights using a small amount of labeled data. To enhance the generalization, we carefully set the learning rate ([Supplementary Methods](#)).

After training the deep learning model, all patches were predicted with a label and corresponding probability. The patch likelihoods were then aggregated in a classifier to obtain the WSI-level prediction. We developed 2 independent machine learning methods to aggregate the patch likelihoods, the Patch Likelihood Histogram (PALHI) pipeline and the Bag of Words (BoW) pipeline, which were inspired by the histogram-based method and the vocabulary-based method, respectively. In PALHI, a histogram of the occurrence of the patch likelihood was applied to represent the WSI. In BoW, each patch was mapped to a TF-IDF floating-point variable, and a TF-IDF feature vector was computed to represent the WSI. Traditional machine learning classifiers were then further trained using these feature vectors to predict the status of each WSI.

We then fused the potential histogram and TF-IDF features. The final features were included in machine learning models such as logistic regression, support Vector machines (SVM), K-nearest

neighbor (KNN), random forest, Extra Trees, extreme gradient boosting (XGBoost), light gradient boosting machine (LightGBM), and multilayer perceptron (MLP) for risk model construction. The best model was selected based on the validation set. During the training of the WSI-level classifier, the hyperparameters were determined using grid search on the training set. Grad-CAM class localization maps were created by visualizing the gradients flowing into the final convolutional layer of the network, just before the fully connected layers.

Statistical Analysis

We used receiver operating characteristics (ROC) curve, sensitivity, specificity, positive predictive value, and negative predictive value to measure the model performance. ROC curves were generated by plotting the proportion of true positive cases (sensitivity) against the proportion of false-positive cases (1-specificity) on the basis of various predictive probability thresholds. A larger area under the ROC curve (AUC) indicated a better prediction performance. Experiments of model development were implemented with Python (version 3.7.12). The deep learning model used to extract pathology features was trained using PyTorch package (version 1.8.0). Some preprocessing, such as background removal and patch normalization, was conducted with OnekeyAI platform. All machine learning methods were performed using scikit-learn (version 1.0.2). The groups with high risk and low risk of malignant transformation predicted by the pathomics-based model were classified by the “survival” and “survminer” R packages. $P < .05$ indicated statistical significance. Figure visualizations were performed by GraphPad Prism (version 8.3.0), Python, and R software.

Results

Model Development and Evaluation

The 759 H&E-stained images from multiple academic medical centers were divided into the training set ($n = 489$ cases), validation set ($n = 196$ cases), and testing set ($n = 74$ cases). The baseline characteristics of the 3 data sets are presented in Table 1. A flowchart depicting the procedures for model development is shown in Figure 1.

We tried many different models, including resnet50, inception_v3, resnet101, and densenet121, and compared the algorithms using the same data sets. As shown in Supplementary Table S1, the resnet50 model obtained the best AUC (0.930 [95% CI, 0.929–0.931] in the training set and 0.656 [95% CI, 0.652–0.660] in the validation set). Overall, 106 pathomics features were generated by fusing the prediction and probably histogram (prob-0.5 to prob-1.0, pred-0, and pred-1) and TF-IDF (prob05 to prob10, pred0, and pred1) features together (Supplementary Table S2). Using the features, we then tested common machine learning algorithm models, such as logistic regression, SVM, KNN, random forest, Extra Trees, XGBoost, LightGBM, and MLP. The results are shown in Table 2 and Figure 2A, B. The LightGBM model showed the best AUC score in the max percentile with 0.899 (95% CI, 0.844–0.955) in the validation set and 0.813 (95% CI, 0.727–0.899) in the testing set. Thus, resnet50 combined with LightGBM was used to develop the pathomics-based model to predict the malignant transformation of OLK.

The pathomics-based model was then used to predict the risk of malignant transformation of OLK in the validation set cohort and testing set cohort. The pathomics-based model was used to

Table 1
Baseline characteristics

Data sets	Training set	Validation set	Testing set
No.	489	196	74
Age	55.09 ± 13.29	53.49 ± 12.06	50.74 ± 12.50
Sex			
Female	268 (54.81%)	97 (49.49%)	23 (31.08%)
Male	221 (45.19%)	99 (50.51%)	51 (68.92%)
Site			
Buccal	167 (34.15%)	71 (36.22%)	33 (44.59%)
Ventral of tongue	173 (35.38%)	52 (26.53%)	24 (32.43%)
Dorsum of tongue	59 (12.07%)	35 (17.86%)	6 (8.11%)
Gingiva	67 (13.70%)	31 (15.82%)	8 (10.81%)
Lip	7 (1.43%)	3 (1.53%)	2 (2.70%)
Palate	16 (3.27%)	4 (2.04%)	1 (1.35%)
Smoking history			
No	328 (67.08%)	119 (60.71%)	31 (41.89%)
Yes	161 (32.92%)	77 (39.29%)	43 (58.11%)
Alcohol drinking history			
No	348 (71.17%)	133 (67.86%)	45 (60.81%)
Yes	141 (28.83%)	63 (32.14%)	29 (39.19%)
Malignant transformation	72 (14.72%)	12 (6.12%)	6 (8.11%)

assess the probability of malignant transformation for individuals in the validation and test sets. By Kaplan-Meier curve analysis, the 2 sets were then each divided into 2 groups: high-risk population and low-risk population. As shown in Figure 3A, B, the risk of malignant transformation in the high-risk populations predicted by the pathomics-based model was significantly higher than that in low-risk populations both in the validation set (hazard ratio, 377.5; 95% CI, 19.0–7497.0) and testing set (hazard ratio, 9.2; 95% CI, 1.8–46.0). The malignant transformation rate of the low-risk population in the validation set was 4.81%, whereas the rate in the high-risk population was 33.33%. The malignant transformation rate of the low-risk population in the testing set was 0%, whereas the rate in the high-risk population was 18.18%.

The Model Provided Better Performance than Oral Epithelial Dysplasia Grading

OED grading is currently the only clinical predictor of malignant transformation in OLK.⁹ We next examined OED grading results by pathologists and compared the results with the model. Two head and neck pathologists (10 and 20 years of experience, respectively) conducted OED grading of H&E-stained sections of 196 cases in the validation set, and a third head and neck pathologist (30 years of experience) consulted in cases of disagreement. We evaluated the performance of OED grading in predicting the malignant transformation of OLK, and the AUC was 0.743 (95% CI, 0.623–0.862). These results indicate that the pathomics-based model showed better performance in predicting malignant transformation compared with OED grading (Fig. 2A).

Pathomics Features Associated with Immunohistochemical Staining of Molecules

Deep learning and artificial intelligence algorithms have been described as a black box and lacking interpretability. To explore the interpretation of the black box of the pathomics-based model, we analyzed the relationship between pathomics features generated by the deep learning models and the expression of several

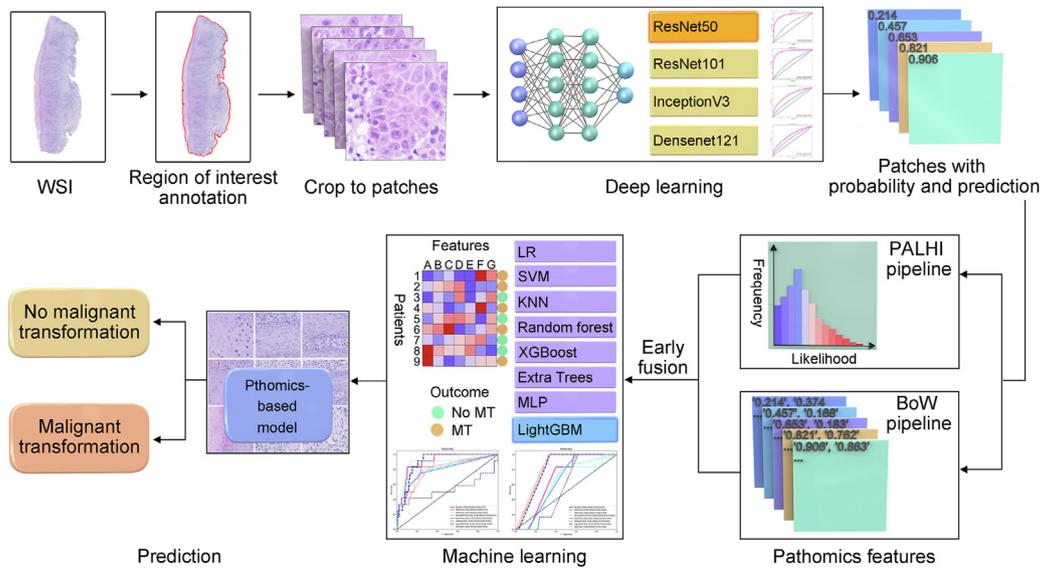


Figure 1.

Flowchart of the development of the pathomics-based model using hematoxylin and eosin staining images through deep learning and machine learning algorithms. BoW, Bag of Words; KNN, K-nearest neighbor; LightGBM, light gradient boosting machine; LR, logistic regression; MLP, multilayer perceptron; MT, malignant transformation; PALHI, Patch Likelihood Histogram; SVM, support vector machines; WSI, whole slide imaging; XGBoost, extreme gradient boosting.

prognostic-related molecules. Previous studies reported that the expressions of Ki67, p53, and PD-L1 were associated with the malignant transformation of OLK.¹⁹⁻²¹ We next performed immunohistochemical staining and evaluated the expressions of Ki67, p53, and PD-L1 in 94 OLK cases from the validation set. Spearman correlation analysis was used to explore pathology deep learning features related to the expression of these 3 molecules (Supplementary Table S3, Fig. 4). Among all deep learning features, 94 features were correlated with the expression of Ki67

($P < .05$). “prob-0.57,” “prob-0.63,” “prob052,” “prob057,” and “prob063” were the pathomics features most associated with expression of Ki67. Only 10 pathomics features were significantly associated with the expression of p53 ($P < .05$): “prob-0.86,” “prob-0.91,” “prob-1.0,” “prob091,” “prob095,” “prob10,” “pred-0,” “pred-1,” “pred0,” and “pred1.” In total, 71 features were related to the expression of PD-L1 ($P < .05$), and “prob-0.79,” “prob-0.8,” “prob058,” “prob079,” and “prob099” were the most associated pathomics features. These results indicate that the expression of

Table 2
The performance of the machine learning models

Models	Data sets	AUC (95% CI)	Sensitivity	Specificity	PPV	NPV
LR	Training set	0.918 (0.884–0.953)	0.542	0.978	0.813	0.925
	Validation set	0.906 (0.838–0.973)	0.833	0.897	0.345	0.988
	Testing set	0.685 (0.555–0.816)	0.833	0.612	0.156	0.976
SVM	Training set	0.951 (0.924–0.979)	0.556	0.990	0.909	0.928
	Validation set	0.533 (0.293–0.773)	0.333	0.984	0.500	0.957
	Testing set	0.558 (0.429–0.686)	1.000	0.388	0.125	1.000
KNN	Training set	0.955 (0.938–0.972)	0.556	0.976	0.800	0.927
	Validation set	0.819 (0.685–0.954)	0.750	0.884	0.273	0.982
	Testing set	0.626 (0.464–0.789)	0.833	0.563	0.135	0.973
RandomForest	Training set	1.000 (0.999–1.000)	1.000	0.998	0.986	1.000
	Validation set	0.788 (0.654–0.922)	0.833	0.683	0.145	0.984
	Testing set	0.839 (0.757–0.922)	1.000	0.900	0.207	1.000
ExtraTrees	Training set	1.000 (1.000–1.000)	1.000	1.000	1.000	1.000
	Validation set	0.787 (0.642–0.932)	0.750	0.761	0.170	0.979
	Testing set	0.651 (0.534–0.767)	1.000	0.403	0.128	1.000
XGBoost	Training set	0.994 (0.983–1.000)	0.931	1.000	1.000	0.988
	Validation set	0.819 (0.687–0.951)	0.833	0.761	0.185	0.986
	Testing set	0.826 (0.742–0.910)	1.000	0.687	0.214	1.000
LightGBM	Training set	0.971 (0.957–0.985)	0.514	0.995	0.949	0.922
	Validation set	0.899 (0.844–0.955)	1.000	0.739	0.200	1.000
	Testing set	0.813 (0.727–0.899)	1.000	0.672	0.207	1.000
MLP	Training set	0.922 (0.892–0.953)	0.514	0.986	0.860	0.922
	Validation set	0.899 (0.802–0.995)	0.917	0.870	0.314	0.994
	Testing set	0.594 (0.391–0.798)	0.833	0.537	0.135	0.973

AUC, area under the receiver operating characteristics curve; KNN, K-nearest neighbor; LightGBM, light gradient boosting machine; LR, logistic regression; MLP, multilayer perceptron; NPV, negative predictive value; PPV, positive predictive value; SVM, support vector machines; XGBoost, extreme gradient boosting.

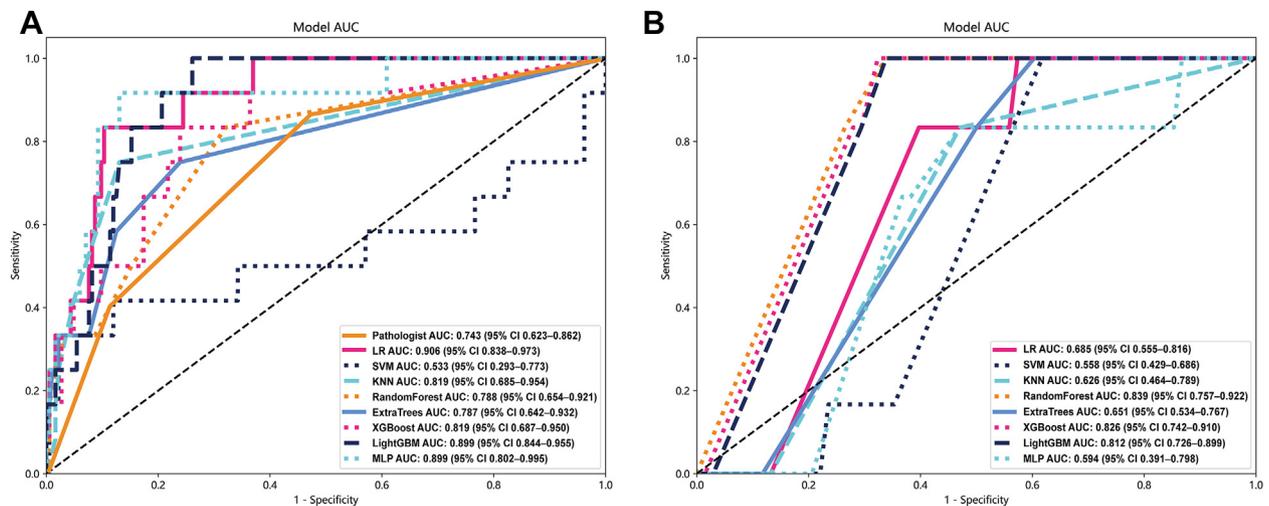


Figure 2.

The performances of each machine learning model in the (A) validation set and (B) testing set. AUC, area under the receiver operating characteristics curve; KNN, K-nearest neighbor; LightGBM, light gradient boosting machine; LR, logistic regression; MLP, multilayer perceptron; SVM, support vector machines; XGBoost, extreme gradient boosting.

multiple cancer risk-associated molecules were significantly associated with deep learning model features, indicating the interpretability of the pathomics-based model.

Discussion

As described in the recent review and meta-analysis of OLK, there is currently no satisfactory single stratification method for prognosis prediction.^{22,23} OED grading is currently the only clinical predictor.²³ However, the low agreement and specificity of OED grading indicates its limited efficacy as a risk-predictive marker for invasive OSCC.^{10,24} Therefore, an efficient and repeatable method for malignant transformation prediction to stratify patients for treatment and management is required.

Deep learning based on H&E images can identify the pathomics features of patients, and artificial intelligence algorithms used in these specific features can distinguish benign and malignant lesions, identify molecular typing, and develop prognostic models.^{12,14-17,25,26} The implementation of artificial intelligence has provided the opportunity to improve precision medicine management and monitoring of OLK. In this study, we showed that a pathomics-based prediction model that was developed using H&E images achieved high predictive performance in a

training set, validation set, and test set derived from real-world data. The predictive performance was better than the previous prediction effect using the deep learning model of digital photos.²⁷ This study also used the pathomics features generated by 2 independent machine learning methods, the PALHI pipeline and the BoW pipeline, which greatly improved the prediction efficacy of the WSI-level of the deep learning model. Multiple deep learning and machine learning training allowed the selection and construction of the optimal pathomics model. Moreover, this pathomics-based model overcame the subjectivity of traditional OED grading and showed better performance compared with OED grading by pathologists. The model was developed with a large number of H&E images, and all patients were diagnosed by pathologic examination of H&E staining samples. The pathomics-based model may be helpful for the identification of malignant risk and monitoring of patients over the long-term; parameters can be adjusted on the basis of conditions. This model may be particularly advantageous in regions where pathologists are scarce. Additionally, the model may be useful to assist with the clinical management and monitoring of OLK cases.

The development and application of artificial intelligence algorithms is a focus of precision medicine. However, the black box issue of artificial intelligence should be considered. To explain the black box of this pathomics model, we performed

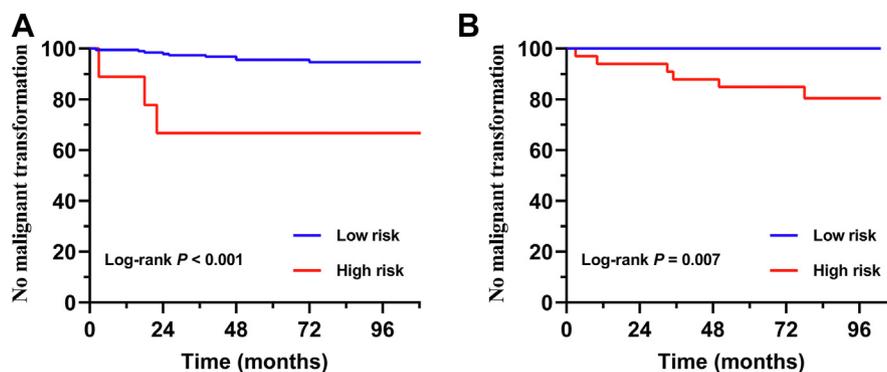


Figure 3.

The pathomics-based model was used to predict the risk of malignant transformation of oral leukoplakia in the (A) validation set cohort and (B) testing set cohort.

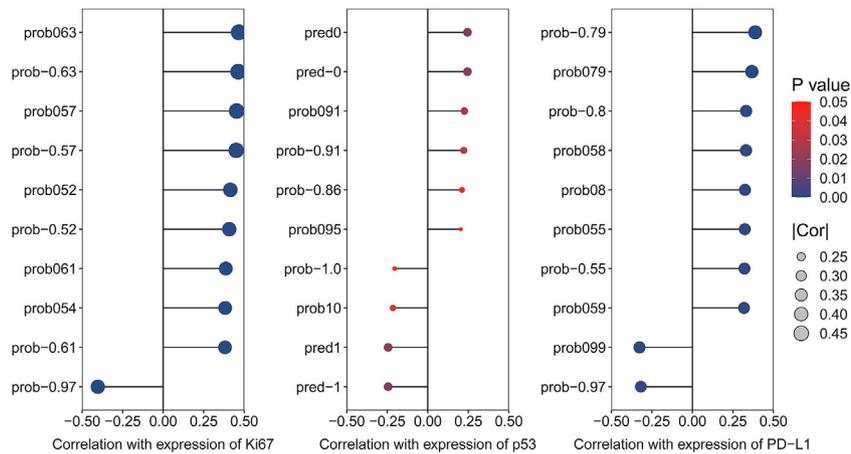


Figure 4.

Pathomics features associated with the expression of Ki67, p53, and PD-L1.

immunohistochemical staining of 3 molecules that are involved in epithelial dysplasia and malignant transformation of OLK to assess the correlation between the expression of these molecules and pathomics features.¹⁹⁻²¹ The expression of Ki67, p53, and PD-L1 showed significant correlations with pathomics features. This suggested that our model does not occur in a vacuum. The weakly supervised deep learning methods based on H&E staining images might map the expression of a variety of molecules and microenvironment alterations involved in the malignant transformation of OLK.^{13,14,17,18,25} The increased expression of oncogenes and changes in the immune microenvironment are factors that influence the development of OSCC.²⁸⁻³⁰ Although the pathomics features of the model have not been fully analyzed, WSI scanning might help artificial intelligence to conduct systematic analysis and thus achieve good predictive performance.

This study has several limitations. First, this was a retrospective study, and the predictive performance of this pathomics-based model needs to be further investigated in prospective clinical trials. Second, the use of larger external cohorts might help the application of the model to patients in different regions.

In conclusion, the pathomics-based model developed in this study may represent an automatic prediction tool for malignant transformation of OLK, with an improved prediction efficiency and performance compared with conventional OED grading. The model will provide accessory opinions that may contribute to decision making in routine clinical management, especially in remote areas where medical resources are scarce. Moreover, the demonstrated high performance of this deep learning model in the internal validation set and external test set in this study warrants further studies in prospective clinical trials.

Author Contributions

L.T.J., Z.J.Y., Z.H.Y., and Z.F. performed conception and design, development of methodology, data acquisition, and revision of the manuscript. C.X.J., L.L., Y. F.Y., G.R.R., and Z.X. performed data acquisition, analysis, and interpretation and writing of the manuscript. All authors read and approved the final paper.

Data Availability

The data used and analyzed during the current study are available from the corresponding author upon reasonable request.

Funding

This work was supported by research grants from the National Nature Science Foundation of China (81671006, 81300894), CAMS Innovation Fund for Medical Sciences (2019-I2M-5-038), National Clinical Key Discipline Construction Project (PKUSSNKP-202102), and Innovation Fund for Outstanding Doctoral Candidates of Peking University Health Science Center (BMU2022BSS001).

Declaration of Competing Interest

None reported.

Ethics Approval and Consent to Participate

This study was approved by the institutional review board of all involved hospitals.

Supplementary Material

The online version contains supplementary material available at <https://doi.org/10.1016/j.labinv.2023.100173>

References

- Chi AC, Day TA, Neville BW. Oral cavity and oropharyngeal squamous cell carcinoma—an update. *CA Cancer J Clin.* 2015;65(5):401–421.
- Ren ZH, Hu CY, He HR, Li YJ, Lyu J. Global and regional burdens of oral cancer from 1990 to 2017: results from the global burden of disease study. *Cancer Commun (Lond).* 2020;40(2–3):81–92.
- Mody MD, Rocco JW, Yom SS, Haddad RI, Saba NF. Head and neck cancer. *Lancet.* 2021;398(10318):2289–2299.
- Warnakulasuriya S, Kujan O, Aguirre-Urizar JM, et al. Oral potentially malignant disorders: a consensus report from an international seminar on nomenclature and classification, convened by the WHO Collaborating Centre for Oral Cancer. *Oral Dis.* 2021;27(8):1862–1880.
- Cai X, Zhang J, Han Y, Tang Q, Zhang H, Li T. Development and validation of a nomogram prediction model for malignant transformation of oral potentially malignant disorders. *Oral Oncol.* 2021;123:105619.
- Aguirre-Urizar JM, Lafuente-Ibanez de Mendoza I, Warnakulasuriya S. Malignant transformation of oral leukoplakia: systematic review and meta-analysis of the last 5 years. *Oral Dis.* 2021;27(8):1881–1895.
- Chaturvedi AK, Udaltsova N, Engels EA, et al. Oral leukoplakia and risk of progression to oral cancer: a population-based cohort study. *J Natl Cancer Inst.* 2020;112(10):1047–1054.

8. Muller S, Tilakaratne WM. Update from the 5th edition of the World Health Organization classification of head and neck tumors: tumours of the oral cavity and mobile tongue. *Head Neck Pathol.* 2022;16(1): 54–62.
9. Odell E, Kujan O, Warnakulasuriya S, Sloan P. Oral epithelial dysplasia: recognition, grading and clinical significance. *Oral Dis.* 2021;27(8): 1947–1976.
10. Peng JK, Dan HX, Xu H, Zeng X, Chen Q. [Agreement evaluation of the severity of oral epithelial dysplasia in oral leukoplakia]. *Zhonghua Kou Qiang Yi Xue Za Zhi.* 2022;57(9):921–926.
11. Yang SY, Li SH, Liu JL, et al. Histopathology-based diagnosis of oral squamous cell carcinoma using deep learning. *J Dent Res.* 2022;101(11):1321–1327.
12. Vanguri RS, Luo J, Aukerman AT, et al. Multimodal integration of radiology, pathology and genomics for prediction of response to PD-(L)1 blockade in patients with non-small cell lung cancer. *Nat Cancer.* 2022;3(10): 1151–1164.
13. Saldanha OL, Quirke P, West NP, et al. Swarm learning for decentralized artificial intelligence in cancer histopathology. *Nat Med.* 2022;28(6): 1232–1239.
14. Fujii S, Kotani D, Hattori M, et al. Rapid screening using pathomorphologic interpretation to detect BRAFV600E mutation and microsatellite instability in colorectal cancer. *Clin Cancer Res.* 2022;28(12):2623–2632.
15. Feng L, Liu Z, Li C, et al. Development and validation of a radiopathomics model to predict pathological complete response to neoadjuvant chemoradiotherapy in locally advanced rectal cancer: a multicentre observational study. *Lancet Digit Health.* 2022;4(1):e8–e17.
16. Boehm KM, Aherne EA, Ellenson L, et al. Multimodal data integration using machine learning improves risk stratification of high-grade serous ovarian cancer. *Nat Cancer.* 2022;3(6):723–733.
17. Klein S, Quaa A, Quantius J, et al. Deep learning predicts HPV association in oropharyngeal squamous cell carcinomas and identifies patients with a favorable prognosis using regular H&E stains. *Clin Cancer Res.* 2021;27(4): 1131–1138.
18. Lipkova J, Chen RJ, Chen B, et al. Artificial intelligence for multimodal data integration in oncology. *Cancer Cell.* 2022;40(10):1095–1110.
19. Xu SB, Wang MY, Shi XZ, et al. Influence of PD-1/PD-L1 on immune microenvironment in oral leukoplakia and oral squamous cell carcinoma. *Oral Dis.* Preprint. Posted online August 3, 2022. <https://doi.org/10.1111/odi.14332>
20. Ramos-García P, González-Moles MÁ, Warnakulasuriya S. Significance of p53 overexpression in the prediction of the malignant transformation risk of oral potentially malignant disorders: a systematic review and meta-analysis. *Oral Oncol.* 2022;126:105734.
21. Jayaraman S, Pazhani J, PriyaVeeraraghavan V, Raj AT, Somasundaram DB, Patil S. PCNA and Ki67: prognostic proliferation markers for oral cancer. *Oral Oncol.* 2022;130, 105943.
22. Monteiro L, Mello FW, Warnakulasuriya S. Tissue biomarkers for predicting the risk of oral cancer in patients diagnosed with oral leukoplakia: a systematic review. *Oral Dis.* 2021;27(8):1977–1992.
23. Celentano A, Glurich I, Borgnakke WS, Farah CS. World Workshop on Oral Medicine VII: prognostic biomarkers in oral leukoplakia and proliferative verrucous leukoplakia: a systematic review of retrospective studies. *Oral Dis.* 2021;27(4):848–880.
24. Pritzker KPH, Darling MR, Hwang JT, Mock D. Oral potentially malignant disorders (OPMD): what is the clinical utility of dysplasia grade? *Expert Rev Mol Diagn.* 2021;21(3):289–298.
25. Zheng X, Wang R, Zhang X, et al. A deep learning model and human-machine fusion for prediction of EBV-associated gastric cancer from histopathology. *Nat Commun.* 2022;13(1):2790.
26. Cheng NM, Yao JW, Cai JZ, et al. Deep learning for fully automated prediction of overall survival in patients with oropharyngeal cancer using FDG-PET imaging. *Clin Cancer Res.* 2021;27(14):3948–3959.
27. Ferrer-Sánchez A, Bagan J, Vila-Francés J, Magdalena-Benedito R, Bagan-Debon L. Prediction of the risk of cancer and the grade of dysplasia in leukoplakia lesions using deep learning. *Oral Oncol.* 2022;132, 105967.
28. Ruffin AT, Li H, Vujanovic L, Zandberg DP, Ferris RL, Bruno TC. Improving head and neck cancer therapies by immunomodulation of the tumour microenvironment. *Nat Rev Cancer.* 2023;23(3):173–188.
29. Sathasivam HP, Kist R, Sloan P, et al. Predicting the clinical outcome of oral potentially malignant disorders using transcriptomic-based molecular pathology. *Br J Cancer.* 2021;125(3):413–421.
30. Chen SH, Hsiao SY, Chang KY, Chang JY. New insights into oral squamous cell carcinoma: from clinical aspects to molecular tumorigenesis. *Int J Mol Sci.* 2021;22(5):2252.