Development and multi-center validation of machine learning model for early detection of fungal keratitis

Zhenyu Wei,[a] Shigeng Wang,[b] Zhiqun Wang,[a] Yang Zhang,[a] Kexin Chen,[a] Lan Gong,[c] Guigang Li,[d] Qinxiang Zheng,[e] Qin Zhang,[f] Yan He,[g] Qi Zhang,[h] Di Chen,[i] Kai Cao,[a] Jinding Pang,[a] Zijun Zhang,[a] Leying Wang,[a] Zhonghong Ou,[b] and Qingfeng Liang[a,*]

[a]Beijing Institute of Ophthalmology, Beijing Tongren Eye Center, Beijing Tongren Hospital, Capital Medical University, Beijing Key Laboratory of Ophthalmology and Visual Sciences, Beijing, 100005, China
[b]Beijing University of Posts and Telecommunications, Beijing, 100876, China
[c]Department of Ophthalmology, Eye and ENT Hospital of Fudan University, Shanghai 200031, China
[d]Department of Ophthalmology, Tongji Hospital Affiliated to Tongji Medical College, Huazhong University of Science and Technology, Wuhan, 430030, China
[e]Eye Hospital, Wenzhou Medical College, Wenzhou, 325027, China
[f]Department of Ophthalmology, Key Laboratory of Vision Loss and Restoration, Ministry of Education, People's Hospital, Peking University, Beijing, 100044, China
[g]Department of Ophthalmology, The Second Xiangya Hospital, Central South University, Changsha, Hunan 410011, China
[h]Department of Ophthalmology, The First Affiliated Hospital of Chongqing Medical University, Chongqing 400016, China
[i]Department of Ophthalmology, Peking Union Medical College Hospital, Chinese Academy of Medical Sciences & Peking Union Medical College, Beijing, 100730, China

eBioMedicine
2023;88: 104438
Published Online xxx
https://doi.org/10.1016/j.ebiom.2023.104438

Summary

Background Fungal keratitis (FK) is a leading cause of corneal blindness in developing countries due to poor clinical recognition and laboratory identification. Here, we aimed to identify the distinct clinical signature of FK and develop a diagnostic model to differentiate FK from other types of infectious keratitis.

Methods We reviewed the electronic health records (EHRs) of all patients with suspected infectious keratitis in Beijing Tongren Hospital from January 2011 to December 2021. Twelve clinical signs of slit-lamp images were assessed by Lasso regression analysis and collinear variables were excluded. Three models based on binary logistic regression, random forest classification, and decision tree classification were trained for FK diagnosis and employed for internal validation. Independent external validation of the models was performed in a cohort of 420 patients from seven different ophthalmic centers to evaluate the accuracy, specificity, and sensitivity in real world.

Findings Three diagnostic models of FK based on binary logistic regression, random forest classification, and decision tree classification were established and internal validation were achieved with the mean AUC of 0.916, 0.920, and 0.859, respectively. The models were well-calibrated by external validation using a prospective cohort including 210 FK and 210 non-FK patients from seven eye centers across China. The diagnostic model with the binary logistic regression algorithm classified the external validation dataset with a sensitivity of 0.907 (0.774, 1.000), specificity 0.899 (0.750, 1.000), accuracy 0.905 (0.805, 1.000), and AUC 0.903 (0.808, 0.998).

Interpretation Our model enables rapid identification of FK, which will help ophthalmologists to establish a preliminary diagnosis and to improve the diagnostic accuracy in clinic.

Funding The Open Research Fund from the National Key Research and Development Program of China (2021YFC2301000) and the Open Research Fund from Beijing Advanced Innovation Center for Big Data-Based Precision Medicine, Beijing Tongren Hospital, Beihang University & Capital Medical University (BHTR-KFJJ-202001) supported this study.

Abbreviations: AK, Acanthamoeba keratitis; ANN, Artificial neural network; AUC, Area under curve; BCVA, Best-corrected visual acuity; BK, Bacterial keratitis; CI, Confidence interval; DL, Deep learning; FK, Fungal keratitis; IK, Infectious keratitis; IVCM, In vivo confocal microscopy; OR, Odds ratio; ROC, Receiver operating characteristic; VK, Viral keratitis

*Corresponding author.
E-mail addresses: liangqingfeng@ccmu.edu.cn, lqflucky@163.com (Q. Liang).

Keywords: Fungal keratitis; Diagnostic model; Slit-lamp microscopy; Machine learning

Research in context

Evidence before this study

The diagnosis of fungal keratitis (FK) was in a conflicting position between accuracy and timeliness. Although *in vivo* confocal microscopy (IVCM) and polymerase chain reaction could break this stalemate, equipment cost and operating threshold limited their application. Machine learning seem a "perfect answer", which had been successfully utilized for the diagnosis of eye diseases. Low cost and instantaneous feedback may increase patient compliance and alleviate the tremendous strain on healthcare resources. More than an assistive technology of decision support, this diagnostic modality that directly input medical images into the network for learning challenged the classical models of diagnosis. In the past, the determination of clinical signs depended on the physicians' knowledge and some clinical signs in infectious keratitis could overlap with other disorders. The correlation between clinical signs and disease could been furtherly validated. Machine learning could also utilize computerized algorithms to optimize the way of clinician diagnose.
The work of Standardization of Uveitis Nomenclature Working Group (2022) provided a reference. They collected the characteristics of cases with Fuchs uveitis syndrome and develop a database. And then using machine learning to screen signs and establish classification criteria. Based on this protocol, the FK diagnostic models could be established. A study (2021) using the largest international cohort of corneal specialists, identified that there had a gap of cognitive ability between experts practicing in India and other countries about FK, but not appears in bacterial keratitis. The result showed that FK may have some special signs, meanwhile, it could be validated by the laboratory examination.

Added value of this study

In this study, we had developed a machine learning model based on slit-lamp images of 1916 cases and manual labeling. This model had been validated in 7 sub-centers and detected diagnostic potency by corneal specialists. We had demonstrated that the model was reliable and solid in prediction of FK. Moreover, special features were extracted by machine learning model. The list of most valuable features identified will trigger future diagnostic modality.

Implications of all the available evidence

We hope that our results will help to inform future research in this area, including the development of guidelines for machine learning applications in clinical signs screening and FK diagnosis. By running the algorithm, confused signs might be more objective and could be distinguished in similar diseases.

Introduction

Fungal keratitis (FK) is a severe and sight-threatening corneal infection characterized by a high incidence, late diagnosis, and devastatingly poor prognosis. It is most prevalent in countries and regions with low socioeconomic status. In developing countries, approximately 30–40% of infectious keratitis (IK) is FK.[1–6] The incidence of FK also shows an increasing trend in developing countries due to improper care and cleaning of contact lenses. In the United States, 6–20% of corneal infections were caused by various fungal pathogens. A recent study revealed that the minimum annual incidence of FK was more than 1,051,787 cases globally.[7] The rate of perforation in FK is five to six times more likely than that in bacterial keratitis (BK).[8] About 8–11% of cases with FK must have their eyes removed, which represents an annual loss of a hundred thousand eyes.[9,10] Early diagnosis and prompt treatment are recognized as key strategies for improving prognosis of FK.[11]

Clinical symptoms and early signs of FK often overlap with other type of IK, so the diagnosis of FK is usually based on laboratory examinations including scraping and fungal culture. They are also regarded as the gold standards for FK diagnosis. However, difficult biopsy and low positive rate limited their clinical application.[12] *In vivo* confocal microscopy (IVCM) is a useful clinical tool for the diagnosis of FK with sensitivity and specificity of 88% and 91%. However, IVCM is not available in many underdeveloped regions where FK is common. Thus, clinical examination under slit-lamp microscopy is still the mainstay method for the early diagnosis of FK, which was verified by a recent multicenter study. The ophthalmologists among different areas (India vs. other countries) performed a significant variability of clinical diagnosis in FK under the same database (accuracy, 76% vs. 49%; AUC, 0.72 vs. 0.59; $P < 0.001$).[13] The ability of experience-based diagnosis suggested the presence of characteristic signs. But the previous method of determining clinical signs was subjective.

In medicine, significant progress of machine learning had been demonstrated in ophthalmology. As of this date, several models related to FK diagnosis were established previously,[14–19] which were all developed based on deep learning algorithm and used the images divided from the same dataset to validate. Small number

of studies provided the comparison of models and ophthalmologists. At the same time, existing researched hadn't validate independently by large prospective clinical cohort and lack of the furtherly interpretation of their input images and their results. In this study, we combined manual labeling and machine learning to quantitatively analyze the diagnostic potency of clinical signs and developed a diagnostic model of FK. The accuracy, sensitivity, and specificity of the model are validated by large prospective cohort and compared with previous models, which would help ophthalmologists to achieve an early diagnosis of FK when IVCM is not available.

Methods

Study design

Based on the inclusion criteria, 2890 patients with IK were recruited from Beijing Tongren Hospital, China, from January 2011 to December 2021. Following the exclusion criteria presented below, 877 patients were excluded because the slit-lamp images in diffuse illumination were missing. In addition to this, 517 patients were excluded because of poor quality of images (325 patients with out-of-focus images and 192 patients with images of incomplete lesion displaying). Finally, for the development dataset (n = 1047) and internal validation dataset (n = 449), patients were recruited. For the external validation cohort (n = 420, from Jan 2020 to Jul 2022), patients were recruited from seven eye centers of China. Patients diagnosed with FK, BK, or *Acanthamoeba* keratitis (AK) by clinical manifestations and laboratory tests (at least one positive of smears of corneal scraping or microbial cultures) were enrolled for the study. Enrolled patients with viral keratitis (VK) were diagnosed with their medical history, clinical manifestations, and effective outcomes of antiviral therapy.[20] Corneal scraping for microbiological tests was performed under slit-lamp microscopy by ophthalmologists, the samples were then placed on the glass slides for Gram and Giemsa staining. Blood agar medium, chocolate agar medium, potato dextrose medium, and Page's medium with *Escherichia coli* were used to culture the microorganism from cornea lesions, including bacteria, fungi, and *Acanthamoeba* individuals. Mass spectrometry and sequencing were used to identify some rare, isolated strains grown from the medium.

Exclusion criteria of the study included (1) patients with mixed ocular infections; (2) patients with a history of any ocular infection, ocular inflammation, ocular trauma, or eye surgery, which were irrelevant for the current keratitis; (3) patients without slit-lamp images in diffuse illumination or images with poor quality reviewed by two masked Ophthalmologists, including out-of-focus images and incomplete lesion displaying, etc.

Ethics

Ethical approval was obtained from the Medical Ethics Committee of Beijing Tongren Hospital (TRECKY2021-024) and all other participating hospitals, including Peking Union Medical College Hospital (No. zs-3516), the First Affiliated Hospital of Chongqing Medical University (No. 2014-76), the Second Xiangya Hospital of Central South University (No. LYF2021028), Eye Hospital of Wenzhou Medical College (No. 2020-191-K-174), and Tongji Hospital of Tongji Medical College of Huazhong University of Science and Technology (No. 2020–120), Peking University People's Hospital (No. 2017PHB167-01), Eye & ENT Hospital of Fudan University (No. 2021103). The study protocol fully adhered to the Declaration of Helsinki and the ARVO statement on human subjects. All subjects finished the written informed consent documents. If participants were under 18 years old, written informed consent should be got from their parents. The experimental design was performed following the flow chart (Fig. 1).

Data collection

A total of 1916 slit-lamp images (from 1916 eyes) were included in this study: 1229 images with FK and 687 images with non-FK (including 267 BK, 284 VK, and 136 AK). Among them, the clinical data of 713 FK images and 334 non-FK images (including 117 BK, 140 VK, and 77 AK) were collected to establish the development set and then produce a diagnostic model of FK. During the stage of model validation, 306 FK images and 143 non-FK images (diagnosed with 50 BK, 60 VK, and 33 AK) from the Beijing Tongren Hospital were applied to the internal validation dataset, and 420 images (including 210 FK, 100 BK, 84 VK, 26 AK) from seven different eye centers were included in the external validation dataset.

Basic information of patients, including duration from first symptom onset to attend our hospital, and risk factors of all enrolled patients were recorded. Ophthalmological evaluations were also collected during the patient's initial visit, including best-corrected visual acuity (BCVA), and clinical features with slit-lamp microscopy. Corneal photography was performed by three examiners (Z.W., Y.Z., K.C.) using Topcon SL-D7 slit-lamp microscopy (since Jan 2011), and Topcon SL-D701 (since Jan 2017), and Haag-Streit BX900 slit-lamp microscopy (since Oct 2020). A standard imaging protocol should be followed for each case, including ambient room light (90–180 lux), diffuse beam at maximal width (30 mm), and light intensity at a maximal level within patient tolerance. Photographs were taken at the 10 × magnification setting.

Image evaluation

Image evaluation was based on 12 common reported clinical signs of FK, including dense infiltrate,

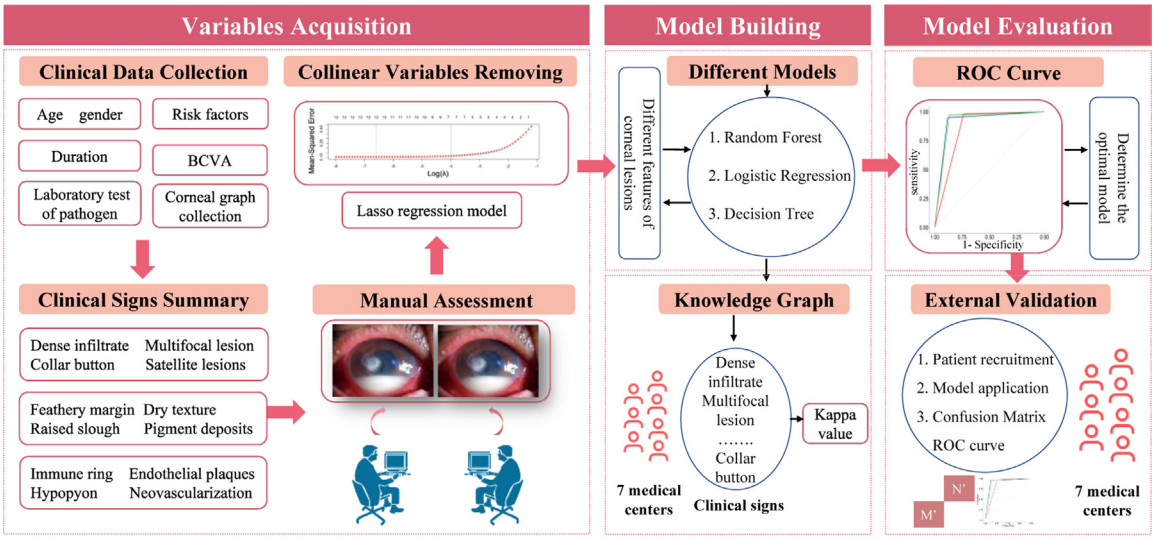

Fig. 1: Research design of diagnostic prediction model for fungal keratitis. The workflow consists of eight main steps (including clinical data collection, clinical signs summary, manual assessment, collinear variables removing, different model development, graph knowledge, internal validation and external validation). These steps could be summarized in three components (Variables acquisition, model building and model evaluation), which described in the flow diagram.

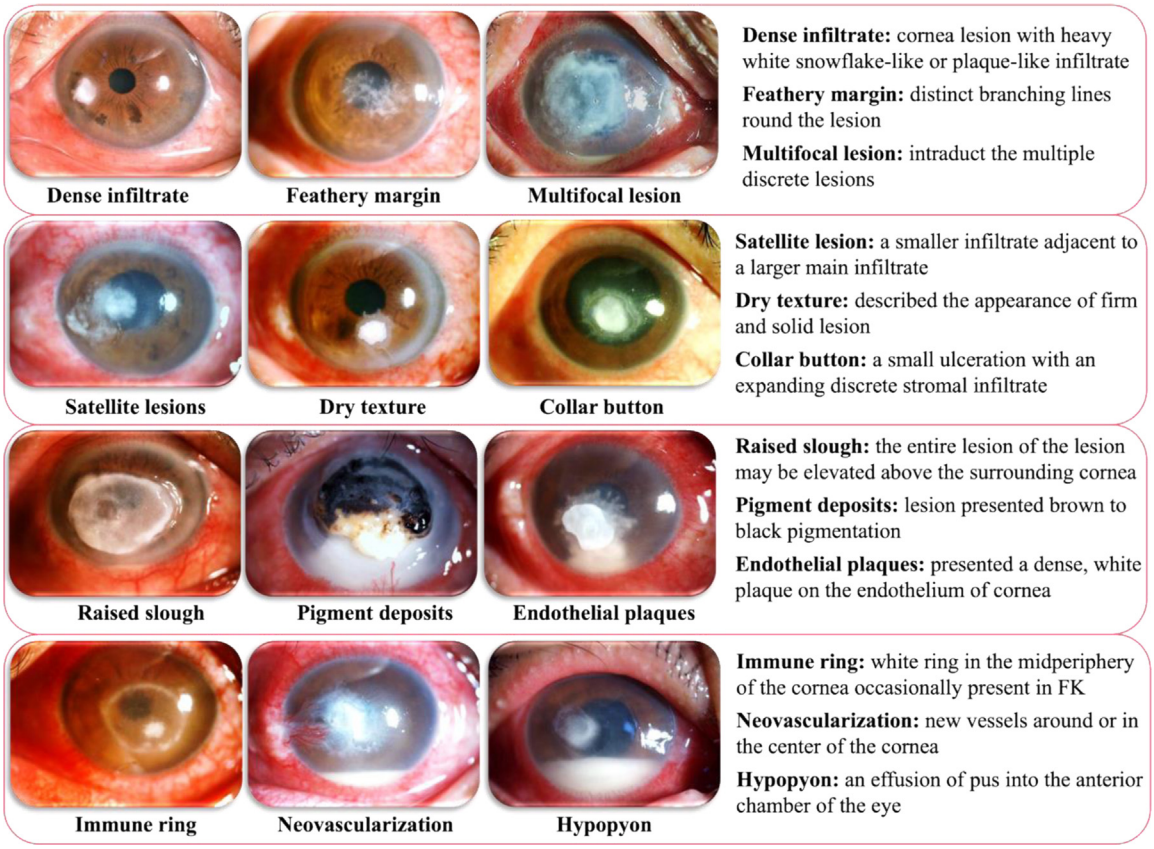

Fig. 2: All included clinical signs of fungal keratitis for manual annotation and machine learning model development in slit-lamp images. (magnification × 10). The text provides the word-to-picture interpretation at the right side of the figure.

multifocal lesion, satellite lesions, feathery margin, dry texture, "collar button" configuration, raised slough, pigment deposits, endothelial plaques, immune ring, neovascularization, and hypopyon.[21–31] (Fig. 2).

For each image, the presentation of 12 clinical signs was evaluated and recorded with "positive" or "negative" (positive was defined as 1, negative as 0) independently by two masked corneal specialty ophthalmologists. After that, the images were converted into numerical vectors and then inputted into machine learning models. When two ophthalmologists could not reach an agreement, the discrepancies were reviewed and decided by a senior corneal specialist. To evaluate the intra bias of each observer, 100 images (50 images from FK and another 50 images from non-FK) were randomly selected from the development dataset and pooled into a extra list. Intra- and inter-individual variations were calculated with analysis of variance.

Model development

The training dataset was established based on 1496 slit-lamp images from 1019 FK cases and 477 non-FK cases, which were randomly split into a development set (70%) and an internal test set (30%). Complete clinical information was collected and the characteristics of slit-lamp images from all cases were evaluated. The diagnostic model of FK was established with the development dataset. Before model building, the Lasso regression model was used to exclude collinear variables selection. Each slit-lamp image could provide the value of 12 clinical signs and the grouping situation (FK and non-FK). So, thirteen parameters were initially included in the Lasso regression model, which was built using the glmnet package in R software (https://www.r-project.org). After that, three different models (binary logistic regression, random forest classification, and decision tree classification) were established based on Python and Jupyter Notebook (version 1.0.0, https://jupyter.org). Specifically, binary logistic regression is a linear model for classification. In this model, the probabilities describing the possible outcomes of a single category are modeled using a logistic function (e.g., sigmoid function). As a non-parametric supervised learning method, decision tree model could predict the value of a target variable by learning decision rules inferred from the data features. Therefore, it has a strong interpretability. Random forest is an ensemble model based on randomized decision tree classifiers. The prediction of the ensemble is given as the averaged prediction of the individual classifiers. What's more, the random forest can evaluate the contribution of each feature to the overall model by averaging the contribution of each feature across each tree. Among them, binary logistic regression was performed in Stats models (version 0.12.0, https://www.statsmodels.org), and the other algorithms were achieved using packages Scikit-Learn (version 0.23.2, http://scikit-learn.org/). For binary logistic regression, the penalty term is L2 to regularize the weights. The solver is lbfgs to solve the optimization problem. The max_iter is 100 which controls the maximum number of iterations to solve. For decision tree, the criterion is entropy to measure the quality of a split. The max_depth is 4 to control tree depth and complexity. For random forest, the criterion is entropy, and the n_estimators is 1000 which determine the number of trees.

Internal validation

An internal evaluation dataset was established and applied to validate the diagnostic efficacy of three different models (binary logistic regression, random forest classification, and decision tree classification) for the diagnosis of FK. To evaluate the diagnostic potency of different models, the sensitivities, specificities, and area under the receiver operating characteristic (ROC) curves were computed by the internal test set with the R package "pROC". Based on sensitivity, specificity, and Delong's test of diagnostic indices provided in the ROC curve, the most valuable model was obtained.

External validation

The external validation dataset consisted of seven sub-datasets from seven different ophthalmic centers (30 FK cases and 30 non-FK cases in each center): **Sub-dataset 1**: Peking Union Medical College Hospital; **Sub-dataset 2**: The First Affiliated Hospital of Chongqing Medical University; **Sub-dataset 3**: The Second Xiangya Hospital of Central South University; **Sub-dataset 4**: Eye Hospital of Wenzhou Medical College; **Sub-dataset 5**: Tongji Hospital of Tongji Medical College of HUST; **Sub-dataset 6**: Peking University People's Hospital; **Sub-dataset 7**: Eye & ENT Hospital of Fudan University.

Before the establishment of the external validation dataset, training with 40-images of FK sign recognition was conducted in 7 sub-centers, and the extent of interobserver consistency was evaluated by Cohen's kappa coefficient (kappa) values. When its consistency was higher (Kappa value > 0.75), these cornea specialists of sub-centers were authorized to perform clinical validation. Otherwise, continued training until an agreement was obtained. After completing the training phase, 60 patients from each sub-center (30 FK cases and 30 non-FK cases) who met the inclusion criteria were recruited and an external validation dataset (420 cases totally) was established. The evaluation of clinical signs of all cases was conducted in each sub-center and the results from every sub-dataset were input to the diagnostic model of FK (determined by internal validation) individually. At the beginning, the performance of this model (e.g., sensitivity, specificity, accuracy, area under the ROC curve) was analyzed for data of each sub-centers. Furtherly, the performances were averaged, and obtained the model performance based on the

whole external dataset. At the same time, these ophthalmologists gave the disease prediction of each image. This part of the data would be used to compared with the selected model.

Statistical analyses

Statistical analysis was based on Python and Jupyter Notebook. Basic data processing was conducted with NumPy (version 1.19.2) and pandas (version 1.1.3, https://pandas.pydata.org). The normality of the data distribution was assessed using a Kolmogorov–Smirnov test. Normal variables were expressed as the mean and standard error of the mean (SEM). For non-normal data variables, the median and interquartile ranges were used. T-test and Chi-square tests were applied to normal and categorical data respectively. Based on Scipy (version 1.5.2, https://www.scipy.org) and SelectKBest algorithm (from Scikit-Learn), the demographics data and signs distribution of FK and non-FK groups were analyzed. For all results, a *P*-value less than 0.05 was considered significant. Multiple comparisons were controlled by Bonferroni correction.

Role of funding source

The funding did not participate in the study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Results

Medical data collection

The demographic data and clinical manifestations of these patients from the training dataset were shown in Table 1. The distribution of fungal species isolated from FK cases in the training dataset were provided in Table S1.

Clinical signs of all photographs in development set were recorded, and the distribution of clinical signs in FK and non-FK group were as follows: dense infiltrate (90.0% vs. 38.0%, $P < 0.001$), multifocal lesion (38.8% vs. 10.8%, $P < 0.001$), satellite lesions (12.6% vs. 0.3%, $P < 0.001$), feathery margin (76.4% vs. 10.2%, $P < 0.001$), dry texture (81.9% vs. 5.1%, $P < 0.001$), "collar button" configuration (7.3% vs. 1.8%, $P < 0.001$), raised slough (47.0% vs. 5.1%, $P < 0.001$), pigment deposits (4.5% vs. 2.7%, $P = 0.16$), endothelial plaques (2.8% vs. 0.3%, $P < 0.001$), immune ring (6.6% vs. 10.2%, $P = 0.04$), neovascularization (42.8% vs. 93.4%, $P < 0.001$), hypopyon (34.2% vs. 16.5%, $P < 0.001$) (all *P* value above used Chi-square tests). Significant differences were observed between FK and the control group (Fig. 3). The intra- and inter-variability of clinical signs between two observers were presented in Table S2.

Development of the diagnostic model

Before model development, the Lasso regression analysis was performed, and the deviance residuals were calculated to exclude collinear variables. Fig. S1 showed an increasing deviance as variables number dropped from 8 to 7 in this model. Therefore, eight candidate signs (dense infiltrate, multifocal lesion, feathery margin, dry texture, raised slough, endothelial plaques, neovascularization, hypopyon) were retained for further analysis. The variables coefficients of these variables in Lasso regression were shown in Table S3.

Three models were established based on binary logistic regression, random forest classification, and decision tree classification analysis. In the model of binary logistic regression, detailed information on 8 candidate variables, including coefficient, odds ratio (OR) value, and 95% confidence interval (CI) of OR value, were presented in Fig. 4A and B. In random forest model, the contributions of these variables were

Parameters	Development dataset			Internal validation dataset		
	FK group	Non-FK group	*P* value	FK group	Non-FK group	*P* value
Number	713	334		306	143	
Mean age (SD), years	54.1 (12.7)	43.9 (15.7)	<0.001*	52.4 (4.1)	43.1 (16.7)	<0.001*
Male, No. (%)	464 (65.1)	276 (82.6)	<0.001*	199 (65.0)	118 (82.5)	<0.001*
Risk factor						
Ocular injury (%)	346 (48.5)	54 (16.2)	<0.001*	148 (48.4)	23 (16.1)	<0.001*
Ocular surgery (%)	29 (4.0)	13 (3.9)	0.893	13 (4.2)	6 (4.2)	0.979
Tap water contact (%)	0 (0.0)	8 (2.4)	<0.001*	0 (0.0)	3 (2.1)	0.032*
Contact lens (%)	4 (0.6)	34 (10.2)	<0.001*	2 (0.7)	14 (9.8)	<0.001*
Unknown (%)	327 (45.9)	226 (67.7)	<0.001*	140 (45.8)	97 (67.8)	<0.001*
Mean duration (SD), days	22.7 (16.9)	20.8 (22.7)	0.523	22.9 (18.2)	20.4 (23.3)	0.622
Mean BCVA (SD)	1.18 (0.47)	0.70 (0.54)	<0.001*	1.23 (0.49)	0.76 (0.51)	0.012*
Scraping positive (negative)	536 (177)	127 (67)	0.007*	229 (77)	55 (28)	0.119
Culture positive (negative)	393 (320)	108 (86)	0.891	168 (138)	46 (37)	0.933

Note: *$P < 0.05$ was considered statistically significant; BCVA: Best corrected visual acuity.

Table 1: **Demographics and medical history of patients in the training dataset.**

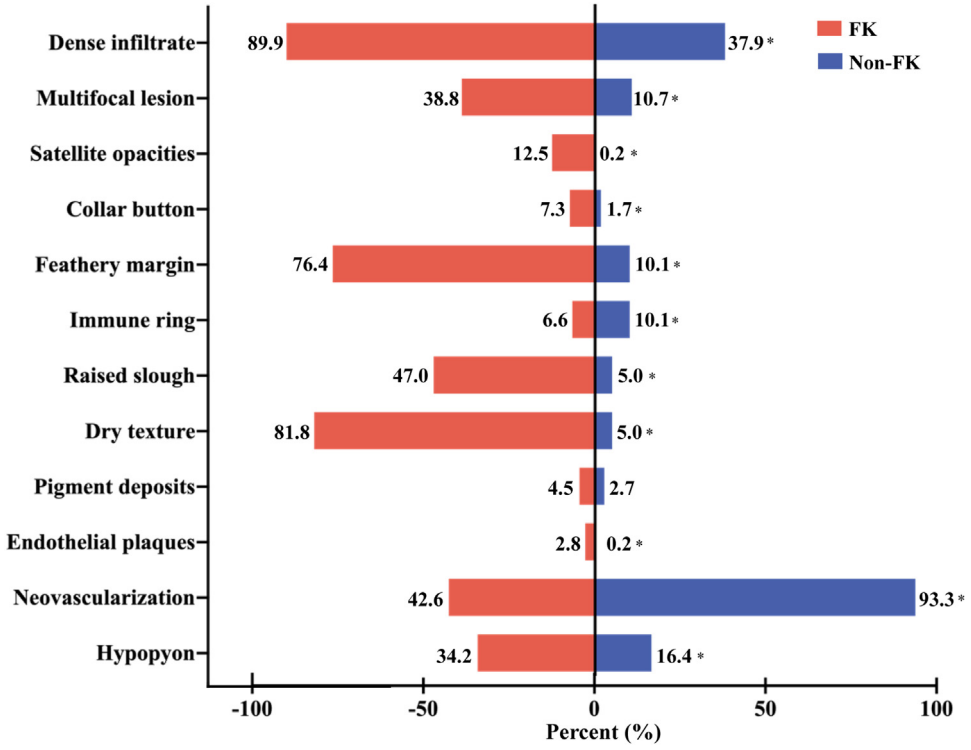

Fig. 3: The distribution of twelve clinical signs from FK group and non-FK group in development dataset. Note: * means $P < 0.05$, which was considered statistically significant.

as follows: dense infiltrate 0.288, dry texture 0.288, feathery margin 0.247, neovascularization 0.208, raised slough 0.071, multifocal lesion 0.038, hypopyon 0.025, endothelial plaques 0.015 (Fig. 4C). In the decision tree model, the contributions of these variables were as follows: dense infiltrate 0.561, dry texture 0.561, feathery margin 0.259, neovascularization 0.128, raised slough 0.012 (Fig. 4D).

Model performance and validation

The clinical manifestations of patients for the internal and external tests have been evaluated (Table S4). The confusion matrix and the ROC curves were calculated, and model validations were evaluated by internal dataset (Fig. 5). The sensitivity, specificity, accuracy, and the area under curve (AUC) of binary logistic regression, random forest classification, decision tree classification in internal validation were (0.948, 0.883, 0.928, 0.916), (0.971, 0.768, 0.939, 0.920), (0.980, 0.737, 0.905, 0.859), respectively.

Delong's test showed that the random forest and logistic regression model had a better performance than the decision tree model ($Z = 4.120$, $P < 0.001$; $Z = 3.547$, $P < 0.001$), while the logistic regression model and random forest showed comparable performance ($Z = -0.324$, $P = 0.746$). To make a model simple and convenient for clinical application, the logistic regression model was determined as the diagnostic model for FK to continue the external validation. Based on the formula "$P = e^{logistic}/(1 + e^{logistic})$" from the logistic regression model, the formula of the diagnostic model of FK was generated which was "logistic = 0.482 + 1.035 × (dense infiltrate) + 1.015 × (multifocal lesion) + 2.526 × (feathery margin) + 1.294 × (raised slough) + 2.690 × (dry texture) - 4.130 × (neovascularization) + 0.332 × (hypopyon)". If a clinical sign was observed in the case, the variable was labelled "1", otherwise it was labelled "0". Cut-off value of the P-value was calculated as 0.772. When $P > 0.772$, cases with these specific signs should be suspected with FK. Based on the above decision result, this study applied this formula to external validation.

After the training of sign recognition of FK in the 7 sub-centers, the kappa values for interobserver consistency in the 7 sub-centers were 0.825, 0.800, 0.800, 0.775, 0.775, 1.000 and 0.900, respectively. Evaluation results of clinical signs in the external validation dataset were also provided in Supplementary Table S4. The performance of the diagnostic model of FK in the sub-centers was summarized in Table 2. It showed that the range of sensitivity, specificity, accuracy, and AUC were (0.793–1.000), (0.774–1.000), (0.817–0.933), (0.816–1.000). The mean and 95% CI of sensitivity and specificity were 0.907 (0.774, 1.000) and 0.899 (0.750, 1.000). The mean accuracy and AUC were 0.905 (0.805, 1.000) and 0.903 (0.808, 0.998).

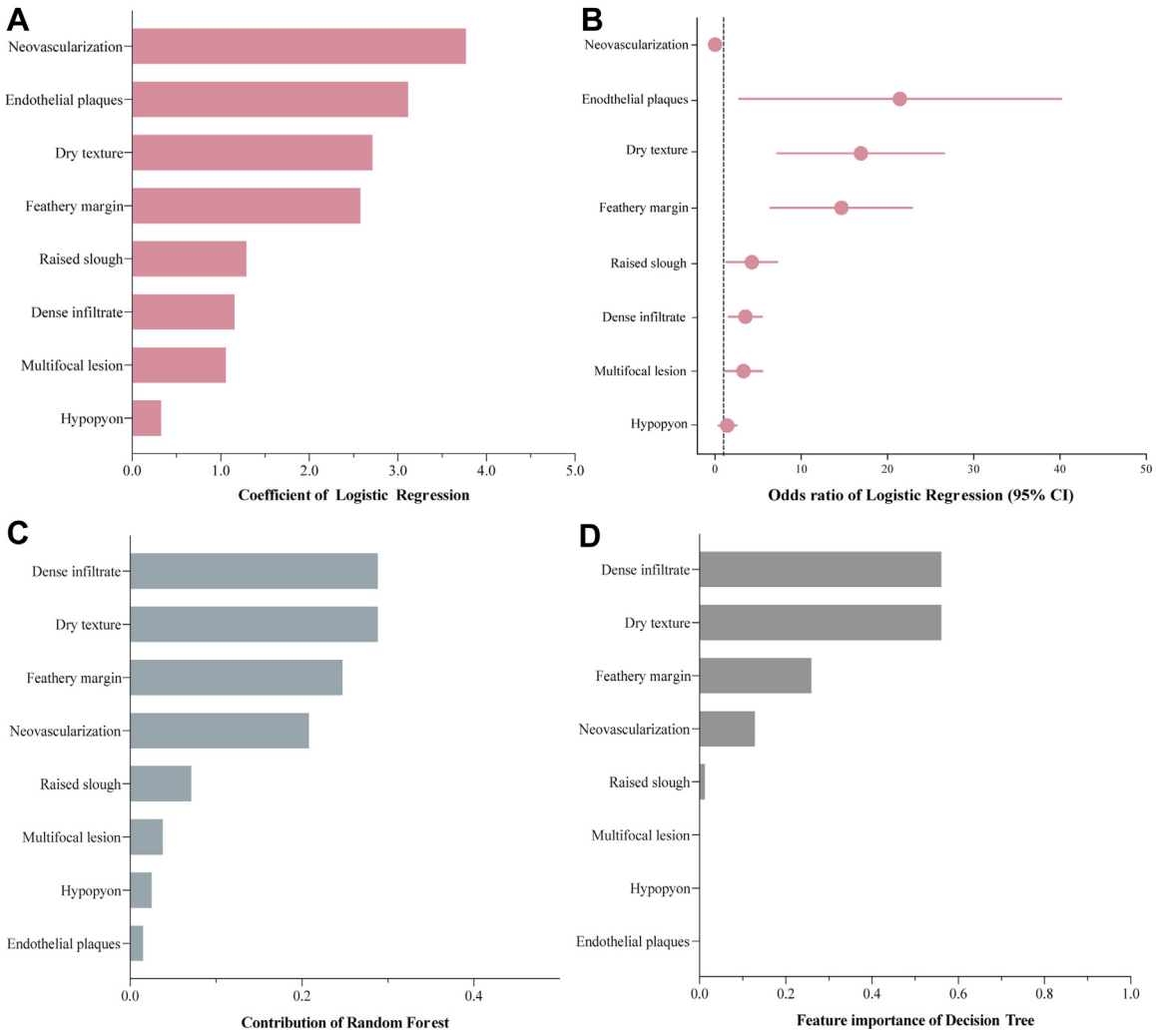

Fig. 4: The contribution of candidate variables in different classification models. (A) Regression coefficients for logistic regression classifiers. (B) Odds ratio for logistic regression classifiers. (C) Contribution degree of each indicator in the random forest. (D) Importance of features estimated by the decision tree.

Comparing prediction performance with corneal specialists and other models

The diagnostic sensitivity of corneal specialists clinically diagnosing FK ranged from 46.7% to 76.7%, and specificity was a bit higher (range from 52.0% to 83.3%). When comparing the performance with corneal specialists from the 7 sub-centers, the machine learning model showed the absolute advantage when classifying the same set of images (mean sensitivity: 90.7% vs. 69.1%; mean specificity: 89.9% vs. 71.7%). When comparing the performance of our model with previous models, the current model showed significantly higher sensitivity (90.7% vs. 65.8–77.0%) and accuracy (89.9% vs. 65.8–77.3%) (Table 3). In addition, most of the previous studies collected multiple images from one patient, and the sample size of their studies (63–580 cases) was smaller than that of Wang's (1923 cases) and our study (1916 cases). In this study, the prediction model has been established with slit-lamp images from four types of IK and the valuable signs for FK diagnosis has been screened, which would be useful for ophthalmologists to establish the preliminary diagnosis of FK during a hospital encounter.

Discussion

Delayed diagnosis remains the main reason for poor prognosis by deteriorating the lesions.[32] As such, improving clinicians' diagnostic ability and obtaining high-quality specimens as well as delicate microbiological examinations are fundamental for the early diagnosis of FK. To our knowledge, this is the first

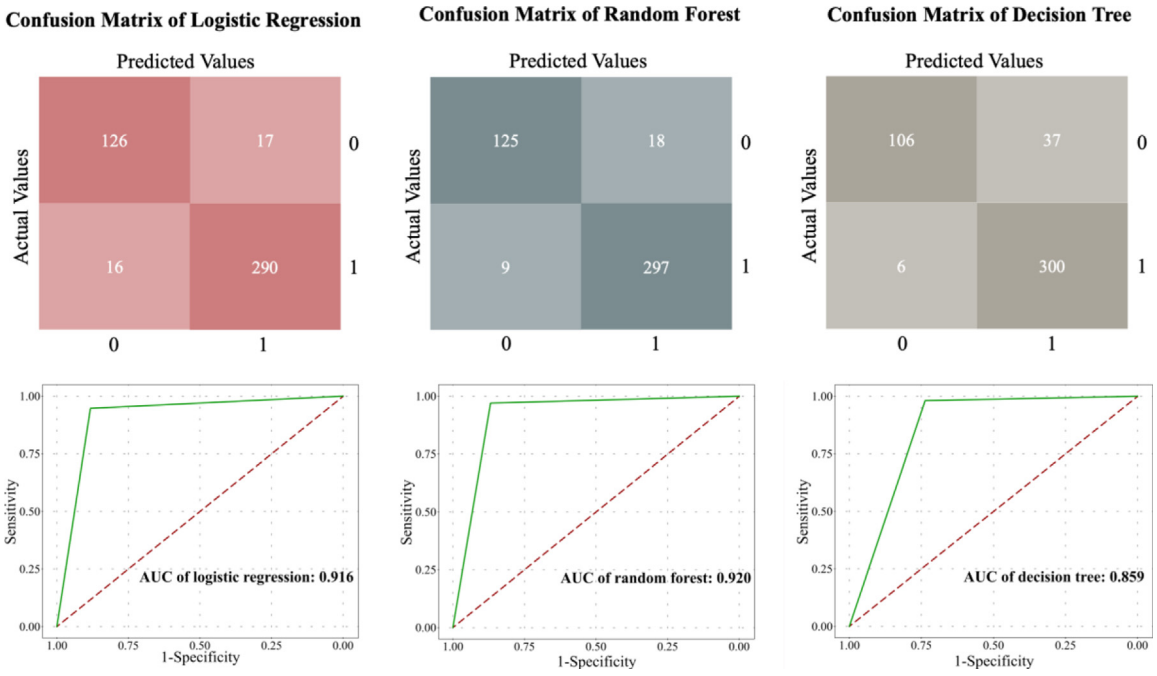

Fig. 5: Performance of the algorithms embodied by confusion matrix (upper) and ROC curve (lower) in different classification models. Left side of the picture: internal validation results of logistic regression model. Middle picture: internal validation results of random forest. Right side of the picture: internal validation results of decision tree.

diagnostic model that combines manual recognition of clinical signs and machine learning for FK diagnosis, which demonstrated a notably higher accuracy and sensitivity than any previously reported models. More importantly, the model was strictly and vigorously validated by prospective cohort of patients from multi-centers.

Previously, some retrospective studies had developed the diagnosis model. When Saini et al.[14] first combined the artificial neural network (ANN) and variables input in 2003, forty variables including predisposing factors and features of ulcers were inputted into the ANN. The lower sensitivity suggested that not all variables were specific for FK. Recently, Hung et al.[15] used convolutional neural networks (DenseNet161) to distinguish BK and FK. Although it showed high accuracy for BK identification, it had very low accuracy in identifying FK (26.3%–65.8%). The lack of observation of other types of IK (such as VK and AK) might also limit the application of clinical differential diagnosis. Kuo et al.[16] conducted a deep learning (DL) study of corneal photographs including all conditions of IK (141 BK, 174 FK, 21 VK, and 12 AK), and the average sensitivity of their model was higher than 70%, but the average specificity was lower than 70%. In their study, the accuracy of the DL group was higher than non-corneal specialty ophthalmologists, but lower than cornea specialists. The results demonstrated the unique strengths of cornea specialists. In our study, three cornea specialists were invited to participate in the sign-identification of FK for

Sub-centers	Sensitivity (95% CI)	Specificity (95% CI)	Accuracy (95% CI)	AUC (95% CI)
Center 1	0.917 (0.750, 1.000)	1.000 (1.000, 1.000)	0.967 (0.899, 1.000)	0.958 (0.877, 1.000)
Center 2	0.793 (0.655, 0.931)	0.839 (0.710, 0.968)	0.817 (0.717, 0.917)	0.816 (0.716, 0.916)
Center 3	0.931 (0.828, 1.000)	0.871 (0.742, 0.968)	0.900 (0.817, 0.967)	0.901 (0.825, 0.977)
Center 4	1.000 (1.000, 1.000)	0.774 (0.613, 0.903)	0.883 (0.800, 0.950)	0.887 (0.812, 0.962)
Center 5	0.827 (0.690, 0.966)	0.903 (0.774, 1.000)	0.867 (0.767, 0.950)	0.865 (0.778, 0.953)
Center 6	0.964 (0.893, 1.000)	0.906 (0.781, 0.893)	0.933 (0.867, 0.983)	0.935 (0.873, 0.997)
Center 7	0.917 (0.750, 1.000)	1.000 (1.000, 1.000)	0.967 (0.899, 1.000)	0.958 (0.877, 1.000)
Average	0.907 (0.774, 1.000)	0.899 (0.750, 1.000)	0.905 (0.805, 1.000)	0.903 (0.808, 0.998)

Table 2: External validation of diagnostic model for fungal keratitis in seven eye centers.

Author	Type of study	Sample size (case number)	Type of IK	Diagnostic model	Sensitivity (95% CI)	Specificity (95% CI)	Accuracy (95% CI)	AUC (95% CI)
Current study	Retrospective & prospective	1916 (1916)	FK, BK, VK, AK	Binary logistic regression	90.7 (77.4, 100.0)	89.9 (75.0, 100.0)	90.5 (80.5, 100.0)	90.3 (80.8, 99.8)
Saini et al.[14]	Retrospective	63 (63)	FK, BK	Artifical neural network	76.5	100.0	76.5	–
Hung et al.[15]	Retrospective	1330 (580)	FK, BK	DenseNet161	65.8 (41.5, 65.8)	87.3 (86.0, 95.3)	65.8	85.0
Kuo et al.[16]	Retrospective	288 (288)	FK, BK, VK, AK	DenseNet	71.1 (62.1, 78.6)	68.4 (61.1, 74.9)	69.4	65.0
Wang et al.[17]	Retrospective	1923 (1923)	FK, BK, VK	InceptionV3	–	–	77.3[a]	93.5
Ghosh et al.[18]	Retrospective	2167 (194)	FK, BK	DeepKeratitis	77.0 (81.0, 83.0)	–	–	90.4
Koyama et al.[19]	Retrospective	4306 (362)	FK, BK, VK, AK	ResNet50	–	–	83.0	85.6

[a]Means the data not obtained directly but could calculate based on the sufficient data provided in the paper; IK: infectious keratitis; AUC: the area under the receiver operating characteristic curve.

Table 3: **Prediction performance of different models for fungal keratitis.**

the model establishment. There is no denying that the identification process of signs was subjective. But the high score of intra- and inter-individual variations showed the stability of labeling process and the unity of the cognition of these signs. To further minimize the subjective bias and detect the implications of this variability on the capability of the model, another seven cornea specialists from sub-centers were invited to take a prospective external evaluation of the established model, which had not been available previously. The results revealed that the bias didn't affect model's capability, and the algorithm of machine learning could initially explore the values of these signs. At the same time, the best and most convenient diagnostic model was established with higher sensitivity (90.7%) and specificity (89.9%).

Specific clinical signs of FK should be emphasized, especially those with higher coefficients (non-neovascularization, endothelial plaques, dry texture, and feathery margin). These signs may strongly correlate with the host immune response and the growth, metabolism, and activity of fungi. It was interesting that corneal neovascularization become a strongly negative factor in our diagnostic model. In the previous studies, neovascularization was presumed as an important sign of IK. But in this study, FK patients more commonly presented non-neovascularization. This result could be explained by Faraj's study which analyzed the etiology of corneal vascularization in 165 patients and found 41 patients (24.9%) caused by VK, 26 (15.8%) patients by BK, and 15 (9.1%) patients by AK. Only 1 (0.6%) case was diagnosed with FK.[33] Neovascularization also becomes a risk factor for VK or outcome after VK. Similar findings that neovascularization was a significant sign of the development stage and convalescence stage were observed in AK.[34] Gurung et al.[35] pointed out recurrent HSV stromal keratitis was associated with neovascularization and a well-defined factor (fibroblast growth factor-2) could drive and maintain progressive corneal neovascularization after HSV-1 infection. As another major portion of the non-FK group, BK always presented an acute onset pattern that VEGF-A would rapidly increase in the early stages of disease (Day 2) and promoted corneal angiogenesis.[36]

The endothelial plaque was the second associated variable.[37] As an identified factor of a diagnostic model to differentiate fungal or bacterial keratitis patients, the formation of endothelial plaque might be related to the vertical growth ability of fungal hyphae.[38] It could also be a marker of hyphae infiltrating Descemet's membrane.[39] Recently, an endothelial biopsy revealed a fungal element in the endothelial plaque.[40] So it may slightly differ in morphology compared with retro-corneal plaques of another microbial keratitis. In such cases, the retro-corneal plaques were related to the collections of inflammatory cells.[41] In our study, the large confidence interval of the parameter of endothelial plaque also suggested the difficulty in identifying endothelial plaque with corneal photography under slit-lamp microscope. Jin et al.[42] suggested that anterior segment optical coherence tomography and IVCM could be used for early corneal endothelial plaque detection. Different imaging investigations could compensate for the deficiency of slit lamps and furtherly improve the prediction accuracy.

Moisture is the essential condition for the growth of fungi. Decreased moisture can reduce the growth of fungi, then lead to a lower density of hypha.[43] Radiate pattern of fungal growth was related to the center area of lesion filling by the pathogen. When the balance between water volume and the fungal load was collapsed, the center of the lesion would present an extremely dry state, so dry texture becomes another important sign in FK. The same presentation also appeared in the solid–state medium, which was more realistic to investigate the growth characteristics of fungal species.[44] Feathery margins or irregular margins were considered strongly to be correlated with FK.[15,45] The unique morphology may be formed by the growth and move of the hypha. On the tip of the hypha, microtubules could facilitate the recycling of the local membrane.[46–48] Filamentation also

had a function of tissue penetration and escape from host immune cells.[49] Therefore, the ability to invade the host tissue by polarized growth of the hypha could determine the fungal pathogenicity.[46] A significant number of hyphae in the lesion periphery spreading and invading the cornea may be presented with the sign of feathery margin.

A recent study attempted to extract descriptors of IK from EHRs and validate the sensitivity of the natural language processing algorithm.[50] But that work was limited by the lack of standardization in recording IK examination elements. Standardizing these descriptors requires a large dataset and scientific method to screen and establish. Based on the current study, we hope to screen the valuable signs for future study and provide diagnostic proof for some clinicians before getting a laboratory test. In this study, we preferred to summarize the existing clinical signs for the diagnosis and developed a diagnostic model with manual labeling and machine learning technique, which would provide a basis for the establishment of diagnostic criteria in FK.

There were also several limitations in this study. Although this is a multicenter study, data were obtained only from Chinese patients and filamentous FK accounted for above 80%. This may induce a potential misclassification bias in the diagnosis process of a rare fungal pathogen keratitis. Future analyses should enlarge the dataset of rare pathogen and optimize performance. Additional prospective data from different ethnic populations are required to further validate our model before clinical application. Another bias was formed by the imbalance in age and gender between FK group and non-FK group. Though we focused on the sign extraction of slit-lamp images in this study, the bias cannot be ignored. It may be caused by the risk factor of different keratitis, as we mentioned in Table 1. Different living conditions, such as age, gender, geographic location, income and local culture, can affect the specific components and characteristics of IK. Enlarge dataset and balance the difference in age and gender may solve this problem. In the future study, we should continue to develop a new ensemble learning technique, which used different machine learning algorithms to subsequently combine into an ensemble model. This kind of model ensemble would have good discrimination and outperform the single models. At the same time, the integrated model could input more parameters, such as the age, gender, and risk factor mentioned above, and the model is closer to the real clinical situation and benefits the ophthalmologist's understanding, learning, and application.

In summary, this is currently the largest study that combines manual recognition and machine learning in the clinical diagnosis of FK. Based on eight reliable clinical signs of FK, a diagnostic model of FK was established using data from 713 FK patients and 334 non-FK patients and was validated with another set of patients, including 449 patients for internal validation and 420 patients for external validation. Higher sensitivity, specificity, and larger AUC of the diagnostic model was statistically better in predicting FK. This study could provide the basic information for ophthalmologists to formulate future diagnostic criteria for FK.

Contributors

Z.W., S.W., and Q.L. conceived the study and designed major experiments. Z.W., Y.H., G.L., Q.Z., Q.Z., Y.Z., K.C., Z.W., and Q.L collected the data and made the clinical diagnosis. S.W. performed machine learning. Z.W., K.C., and S.W. performed analysis and interpretation. The manuscript was written by Z.W. and Q.L.

Data sharing statement

Data are available from the corresponding author upon reasonable request.

Declaration of interests

The authors declare that they have no competing interests.

Acknowledgements

This study was supported by The Open Research Fund from the National Key Research and Development Program of China (2021YFC2301000); the Open Research Fund from Beijing Advanced Innovation Center for Big Data-Based Precision Medicine, Beijing Tongren Hospital, Beihang University & Capital Medical University (BHTR-KFJJ-202001).

Appendix A. Supplementary data

Supplementary data related to this article can be found at https://doi.org/10.1016/j.ebiom.2023.104438.

References

1 Jurkunas U, Behlau I, Colby K. Fungal keratitis: changing pathogens and risk factors. *Cornea*. 2009;28(6):638–643.
2 Bharathi MJ, Ramakrishnan R, Meenakshi R, et al. Analysis of the risk factors predisposing to fungal, bacterial & Acanthamoeba keratitis in south India. *Indian J Med Res*. 2009;130(6):749–757.
3 Pong JCF, Law R, Lai J. Risk factors for treatment of fungal keratitis. *Ophthalmology*. 2007;114(3):617.
4 Shi W, Wang T, Xie L, et al. Risk factors, clinical features, and outcomes of recurrent fungal keratitis after corneal transplantation. *Ophthalmology*. 2010;117(5):890–896.
5 Lalitha P, Prajna NV, Kabra A, et al. Risk factors for treatment outcome in fungal keratitis. *Ophthalmology*. 2006;113(4):526–530.
6 Alfonso EC, Miller D, Cantu-Dibildox J, et al. Fungal keratitis associated with non-therapeutic soft contact lenses. *Am J Ophthalmol*. 2006;142(1):154–155.
7 Brown L, Leck AK, Gichangi M, et al. The global incidence and diagnosis of fungal keratitis. *Lancet Infect Dis*. 2021;21(3):e49–e57.
8 Zaki SM, Denning DW. Serious fungal infections in Egypt. *Eur J Clin Microbiol Infect Dis*. 2017;36(6):971–974.
9 Burton MJ, Pithuwa J, Okello E, et al. Microbial keratitis in East Africa: why are the outcomes so poor? *Ophthalmic Epidemiol*. 2011;18(4):158–163.
10 Ali Shah SI, Shah SA, Rai P, et al. Visual outcome in patients of keratomycosis, at a tertiary care centre in Larkana, Pakistan. *JPMA*. 2017;67(7):1035–1038.
11 Niu L, Liu X, Ma Z, et al. Fungal keratitis: pathogenesis, diagnosis and prevention. *Microb Pathog*. 2020;138:103802.
12 McLeod SD, Kolahdouz-Isfahani A, Rostamian K, et al. The role of smears, cultures, and antibiotic sensitivity testing in the management of suspected infectious keratitis. *Ophthalmology*. 1996;103(1):23–28.
13 Redd TK, Prajna NV, Srinivasan M, et al. Expert performance in visual differentiation of bacterial and fungal keratitis. *Ophthalmology*. 2022;129(2):227–230.
14 Saini JS, Jain AK, Kumar S, et al. Neural network approach to classify infective keratitis. *Curr Eye Res*. 2003;27(2):111–116.

15 Hung N, Shih AK-Y, Lin C, et al. Using slit-lamp images for deep learning-based identification of bacterial and fungal keratitis: model development and validation with different convolutional neural networks. *Diagnostics*. 2021;11(7):1246.
16 Kuo M-T, Hsu BW-Y, Yin Y-K, et al. A deep learning approach in diagnosing fungal keratitis based on corneal photographs. *Sci Rep*. 2020;10(1):14424.
17 Wang L, Chen K, Wen H, et al. Feasibility assessment of infectious keratitis depicted on slit-lamp and smartphone photographs using deep learning. *Int J Med Inf*. 2021;155:104583.
18 Ghosh AK, Thammasudjarit R, Jongkhajornpong P, et al. Deep learning for discrimination between fungal keratitis and bacterial keratitis: DeepKeratitis. *Cornea*. 2022;41(5):616–622.
19 Koyama A, Miyazaki D, Nakagawa Y, et al. Determination of probability of causative pathogen in infectious keratitis using deep learning algorithm of slit-lamp images. *Sci Rep*. 2021;11(1):22642.
20 Darougar S, Wishart MS, Viswalingam ND. Epidemiological and clinical features of primary herpes simplex virus ocular infection. *Br J Ophthalmol*. 1985;69(1):2–6.
21 Dalmon C, Porco TC, Lietman TM, et al. The clinical differentiation of bacterial and fungal keratitis: a photographic survey. *Invest Ophthalmol Vis Sci*. 2012;53(4):1787–1791.
22 Ramappa M, Nagpal R, Sharma S, Chaurasia S. Successful medical management of presumptive pythium insidiosum keratitis. *Cornea*. 2017;36(4):511–514.
23 Kumar A, Khurana A, Sharma M, Chauhan L. Causative fungi and treatment outcome of dematiaceous fungal keratitis in North India. *Indian J Ophthalmol*. 2019;67(7):1048–1053.
24 Espósito E, Maccio JP, Monti R, et al. Alternaria keratitis and hypopyon after clear-cornea phacoemulsification. *J Cataract Refract Surg*. 2014;40(2):331–334.
25 Mascarenhas J, Lalitha P, Prajna NV, et al. Acanthamoeba, fungal, and bacterial keratitis: a comparison of risk factors and clinical features. *Am J Ophthalmol*. 2014;157(1):56–62.
26 Oldenburg CE, Prajna VN, Prajna L, et al. Clinical signs in dematiaceous and hyaline fungal keratitis. *Br J Ophthalmol*. 2011;95(5):750–751.
27 Kaufman HE, Wood RM. Mycotic keratitis. *Am J Ophthalmol*. 1965;59:993–1000.
28 Sharma N, Bagga B, Singal D, et al. Fungal keratitis: a review of clinical presentations, treatment strategies and outcomes. *Ocul Surf*. 2022;24:22–30.
29 Garg P, Vemuganti GK, Chatarjee S, et al. Pigmented plaque presentation of dematiaceous fungal keratitis: a clinicopathologic correlation. *Cornea*. 2004;23(6):571–576.
30 Zbiba W, Baba A, Bouayed E, et al. A 5-year retrospective review of fungal keratitis in the region of Cap Bon. *J Fr Ophtalmol*. 2016;39(10):843–848.
31 Farjo QA, Farjo RS, Farjo AA. Scytalidium keratitis: case report in a human eye. *Cornea*. 2006;25(10):1231–1233.
32 Hagan M, Wright E, Newman M, et al. Causes of suppurative keratitis in Ghana. *Br J Ophthalmol*. 1995;79(11):1024–1028.
33 Faraj LA, Said DG, Al-Aqaba M, et al. Clinical evaluation and characterisation of corneal vascularisation. *Br J Ophthalmol*. 2016;100(3):315–322.
34 Sun Y, Hong J, Zhang P, et al. Pathological characteristics of the different stages of Acanthamoeba keratitis. *Histopathology*. 2013;63(6):862–868.
35 Gurung HR, Carr MM, Bryant K, et al. Fibroblast growth factor-2 drives and maintains progressive corneal neovascularization following HSV-1 infection. *Mucosal Immunol*. 2018;11(1):172–185.
36 Narimatsu A, Hattori T, Koike N, et al. Corneal lymphangiogenesis ameliorates corneal inflammation and edema in late stage of bacterial keratitis. *Sci Rep*. 2019;9(1).
37 Jongkhajornpong P, Nimworaphan J, Lekhanont K, et al. Predicting factors and prediction model for discriminating between fungal infection and bacterial infection in severe microbial keratitis. *PLoS One*. 2019;14(3):e0214076.
38 Xie L, Zhai H, Shi W, et al. Hyphal growth patterns and recurrence of fungal keratitis after lamellar keratoplasty. *Ophthalmology*. 2008;115(6):983–987.
39 Qi X, Liu T, Du M, Gao H. Endothelial plaques as sign of hyphae infiltration of Descemet's membrane in fungal keratitis. *J Opthalmol*. 2020;2020:6083854.
40 Palioura S, Tsiampali C, Dubovy SR, Yoo SH. Endothelial biopsy for the diagnosis and management of culture-negative retrocorneal fungal keratitis with the assistance of optical coherence tomography imaging. *Cornea*. 2021;40(9):1193–1196.
41 Takezawa Y, Suzuki T, Shiraishi A. Observation of retrocorneal plaques in patients with infectious keratitis using anterior segment optical coherence tomography. *Cornea*. 2017;36(10):1237–1242.
42 Jin X, Jin H, Shi Y, Zhang N, Zhang H. Clinical observation of corneal endothelial plaques with fungal and bacterial keratitis by anterior segment optical coherence tomography and in vivo confocal microscopy. *Cornea*. 2022;41(11):1426–1432.
43 Liu R, Gong T, Zhang K, Lee C. Graphene oxide papers with high water adsorption capacity for air dehumidification. *Sci Rep*. 2017;7(1):9761.
44 Techaoei S, Jirayuthcharoenkul C, Jarmkom K, et al. Chemical evaluation and antibacterial activity of novel bioactive compounds from endophytic fungi in. *Saudi J Biol Sci*. 2020;27(11):2883–2889.
45 Thomas PA. Characteristic clinical features as an aid to the diagnosis of suppurative keratitis caused by filamentous fungi. *Br J Ophthalmol*. 2005;89(12):1554–1558.
46 Lenz JH, Schuchardt I, Straube A, Steinberg G. A dynein loading zone for retrograde endosome motility at microtubule plus-ends. *EMBO J*. 2006;25(11):2275–2286.
47 Wang A, Lane S, Tian Z, et al. Temporal and spatial control of HGC1 expression results in Hgc1 localization to the apical cells of hyphae in Candida albicans. *Eukaryot Cell*. 2007;6(2):253–261.
48 Wedlich-Söldner R, Bölker M, Kahmann R, Steinberg G. A putative endosomal t-SNARE links exo- and endocytosis in the phytopathogenic fungus Ustilago maydis. *EMBO J*. 2000;19(9):1974–1986.
49 Saville SP, Lazzell AL, Monteagudo C, Lopez-Ribot JL. Engineered control of cell morphology in vivo reveals distinct roles for yeast and filamentous forms of Candida albicans during infection. *Eukaryot Cell*. 2003;2(5):1053–1060.
50 Woodward MA, Maganti N, Niziol LM, et al. Development and validation of a natural language processing algorithm to extract descriptors of microbial keratitis from the electronic health record. *Cornea*. 2021;40(12):1548–1553.