tvst

Refractive Intervention

Predicting Intraocular Collamer Lens Vault in Myopic Patients With Shallow Anterior Chamber Using FedEYE Platform and ChatGPT

Zhanyu Su[1], Zhaoxiang Wang[2], Shanshan Wang[2], Meng Ni[2], Jinwei Hu[2], Ruiwen Wang[2], Jie Yang[2], Yunxia Bai[2], Yonghong Xu[2], Shuli Di[2], Yaqian Xu[2], Zheng Wang[1], Zhuoyue Chen[3], Julio Ortega-Usobiaga[4], Ming X. Wang[5], Xinlong Jiang[6], Yiqiang Chen[6], Weiwei Dai[1,7], Jiansu Chen[1], and Kangjun Li[1–3,8]

[1] Aier Academy of Ophthalmology, Central South University, Changsha, Hunan, China
[2] Xi'an Aier Eye Hospital, Northwest University, Xi'an, Shaanxi, China
[3] School of Medicine, Northwest University, Xi'an, Shaanxi, China
[4] Department of Cataract & Refractive Surgery, Clínica Baviera, Aier Eye Hospital Group, Bilbao, Spain
[5] Wang Vision Institute, Aier Eye Hospital Group, Nashville, TN, USA
[6] Institute of Computing Technology, Chinese Academy of Sciences, Beijing, China
[7] Institute of Digital Ophthalmology and Visual Science, Changsha Aier Eye Hospital, Changsha, Hunan, China
[8] Xi'an People's Hospital, Xi'an Fourth Hospital, Northwest University, Xi'an, Shaanxi, China

Correspondence: Kangjun Li, Xi'an People's Hospital, Xi'an Fourth Hospital, Northwest University, 6 Zhangba East Rd., Xi'an, Shaanxi 710000, China. e-mail: kj4907630@foxmail.com

Received: June 6, 2025
Accepted: September 11, 2025
Published: October 15, 2025

Keywords: myopia; ICL; shallow ACD; AI; ChatGPT

Citation: Su Z, Wang Z, Wang S, Ni M, Hu J, Wang R, Yang J, Bai Y, Xu Y, Di S, Xu Y, Wang Z, Chen Z, Ortega-Usobiaga J, Wang MX, Jiang X, Chen Y, Dai W, Chen J, Li K. Predicting intraocular collamer lens vault in myopic patients with shallow anterior chamber using FedEYE platform and ChatGPT. Transl Vis Sci Technol. 2025;14(10):20, https://doi.org/10.1167/tvst.14.10.20

Purpose: This study evaluated the effectiveness of the code-free platforms FedEYE and ChatGPT in predicting vault in patients with shallow anterior chamber depth (ACD) undergoing implantable collamer lens (ICL) surgery.

Methods: This retrospective study included 160 eyes with shallow ACD (<2.8 mm) and developed two code-free artificial intelligence (AI) models. An integrated neural network (INN) was developed using the FedEYE platform with ResNet and gradient-boosting trees, combined with optical coherence tomography (OCT) images and examination data, to determine how different data combinations achieve the best model performance and the correlation between parameters and vault. ChatGPT prompted with memory function and also used to validate its capability to predict vault based on these preoperative multimodal data.

Results: Based on the test set results, the INN model, using the combined OCT image and examination data, achieved favorable performance with accuracy of 0.875, sensitivity of 0.857, specificity of 0.889, weighted F1 score of 0.86, and area under the curve (AUC) of 0.873. Anterior chamber angle and spherical equivalent positively correlated with vault, and age and white-to-white diameter negatively correlated. Following memory prompting, ChatGPT improved itself and reached accuracy of 0.813, recall of 0.571, weighted F1 score of 0.73, and AUC of 0.865.

Conclusions: Code-free tools are particularly notable for their straightforward development and deployment. These two code-free AI models demonstrated encouraging performance in predicting vault in shallow ACD patients.

Translational Relevance: Using different AI models to predict vault after ICL implantation could be of clinical significance for improving surgical safety and efficacy in shallow ACD patients.

Introduction

The STAAR implantable collamer lens (ICL; STAAR Surgery, Lake Forest, CA) is a phakic posterior chamber intraocular lens used in refractive surgery. Since its approval by the U.S. Food and Drug Administration (FDA) in 2005, ICL implantation has proven to be a safe and effective method for correcting myopia and astigmatism.[1–3] ICL vault is a critical determinant of surgical safety; insufficient vault reduces the space between the ICL and the crystalline lens, potentially causing anterior subcapsular cataracts (ASCs).[4–7] Current guidelines recommend ICL implantation only in patients with an anterior chamber depth (ACD) > 2.8 mm and horizontal white-to-white (WTW) diameter > 10.5 mm, with the FDA advising against implantation in eyes with ACD < 3 mm.[8,9] These criteria are intended to ensure a postoperative vault within a safe range; however, they are primarily based on limited clinical observations rather than systematically validated data. With the widespread application of the STAAR ICL V4c model, with a central hole, the limitations of ICL surgery are constantly being explored and expanded in clinical practice.[10]

Due to anatomical differences, shallow ACD (<2.8 mm) is more frequently observed in East Asian populations.[11] Because ACD is positively correlated with ICL vault, patients with a shallow ACD are more likely to develop insufficient vault, which has been linked to an increased risk of anterior subcapsular opacification or even ASCs.[12,13] Nevertheless, our previous research demonstrated that, with appropriate lens size selection, implantation of the ICL V4c in shallow ACD patients remains both safe and effective, with 61.46% achieving vault within the normal range.[14] Therefore, a systematic model to predict postoperative vault based on preoperative anterior chamber parameters, rather than ACD alone, may provide more reliable guidance for patients, especially those with shallow ACD.

In recent years, the expansion of artificial intelligence (AI) technology has significantly enhanced its application across various medical domains.[15,16] Deep learning, a specialized branch of machine learning (ML), leverages multilayered neural networks and large datasets to excel in automatic learning and recognition. Convolutional neural networks, when trained on clinical data, are increasingly utilized for medical diagnostics and predictive tasks, including the diagnosis of keratoconus and selection of ICL models.[17,18] These methods provide automated, efficient, and precise predictions following rigorous training, surpassing the accuracy of traditional manual observations and formula-based calculations. Nevertheless, the scope of the current research in this area is mostly confined to models trained using solely images or examination data. Additionally, limitations in disciplinary backgrounds make it challenging for physicians and researchers to develop models, significantly restricting their widespread adoption and application.

The burgeoning application of ChatGPT, powered by large language models (LLMs), is serving as a catalyst for innovations across various sectors. These models are trained on a comprehensive corpus that includes extensive medical literature, the latest advancements, and clinical case data.[19] Utilizing attention mechanism–based algorithms, LLMs can assimilate and process vast amounts of data to yield logically coherent outputs.[20] Primarily, these models strive to provide clear and structured responses. Nonetheless, the efficacy of LLMs in medical diagnostics and prognosis still warrants further empirical scrutiny. ChatGPT, one of the most popular LLMs, due to its newly developed image recognition and long-term memory capabilities, could potentially help researchers train more precise models tailored to individual needs.[21]

In this study, we adopted the code-free platform FedEYE to develop a multimodal integrated neural network (INN) model. We aimed to predict the ICL vault in patients with shallow ACD, aiming to identify those at risk of low vault by integrating preoperative optical coherence tomography (OCT) images with selected clinical examination data, and to examine the correlation between various risk factors and the postoperative vault. Additionally, we developed another model by prompting ChatGPT with memory functions and compared it with the model developed at the FedEYE platform.

Methods

Study Design

This retrospective study included 160 eyes with shallow ACD of myopic patients who underwent ICL V4c surgery at the refractive surgery department of Xi'an Aier Eye Hospital between October 2022 and June 2024. The study received approval from the Ethics Committee of Xi'an Aier Eye Hospital and adhered to the tenets of the Declaration of Helsinki. All patients underwent a comprehensive ophthalmic examination by a specialist physician before surgery. The inclusion criteria were (1) shallow ACD (<2.8 mm); (2) stable refractive status for 2 years before surgery; and (3) endothelial cell density > 2200 cells/mm^2. Exclusion criteria included a history of ocular disease or other related surgeries. All patients were closely followed up and examined postoperatively.

Data Preparation and Quality Control

All patients underwent OCT examinations before and after surgery (CIRRUS 5000; Carl Zeiss Meditec, Jena, Germany). Light intensity was measured using a T10A illuminance meter (Konica Minolta, Tokyo, Japan) to ensure that OCT measurements for all patients were conducted under consistent lighting conditions, thereby avoiding any bias in the ICL vault due to varying illumination.[22] The examination procedure followed previously published protocols.[14] Briefly, patients were seated, and the central pupil was used as the focal point for scanning. The right eye was scanned from the temporal to nasal side, and the left eye was scanned from the nasal to temporal side. Each scan was required to clearly delineate the anatomical boundaries of the anterior segment structures. For each patient, three consecutive images were obtained, and the clearest image was selected by a physician for storage and subsequent analysis. From preoperative OCT images, ACD, average nasal and temporal anterior chamber angle (ACA), WTW diameter, crystalline lens rise (CLR), and spherical equivalent (SE) were measured by two physicians who were blinded to the study design using ImageJ 1.53a (National Institutes of Health, Bethesda, MD). Postoperative OCT images were used to measure the ICL vault. Each parameter was measured independently by both physicians, and the mean of their measurements was used for subsequent analyses. Low ICL vault was defined as <250 µm, and all data were binarily labeled as either low ICL vault or normal ICL vault. Before inputting data into the AI algorithm, two clinicians measured and evaluated all case images and examination data to ensure consistency. Of the initial 183 eyes, 160 eyes were included in this study after applying strict exclusion criteria and quality control measures. All study eyes were randomly assigned to training, validation, and test sets in an 8:1:1 ratio. Independent validation and test sets were used to prevent poor generalization to unseen data.

Surgical Interventions

All procedures in this study were performed by the same experienced surgeon. Due to the unique physiological structure of patients with shallow ACD (<2.8 mm), we selected the 12.1-mm size ICL V4c for all patients. The ICL implantation and surgical procedures followed the same protocol as in our previous studies.[23,24] Briefly, following the induction of ciliary muscle paralysis and topical anesthesia, an approximately 2.8-mm clear corneal incision was made at the steepest meridian. Through this incision, the ICL was implanted into the posterior chamber, and the remaining ophthalmic viscosurgical device was completely washed out of the anterior chamber with a balanced salt solution. After surgery, a nonsteroidal anti-inflammatory drug (Pranopulin; Senju Pharmaceuticals, Osaka, Japan) and antibiotic medication (Cravit; Santen Pharmaceutical, Osaka, Japan) were administered topically four times daily for 2 weeks, and the dose was steadily reduced thereafter.

FedEYE Platform

As shown in the flowchart in Figure 1, in this study we developed an INN model on the FedEYE federated learning platform (https://fedeye.aierchina.com/). The platform allows physicians without coding experience to develop AI models using neural networks and ML algorithms. Within the platform, we trained multiple convolutional neural network models using the preoperative OCT images from patients included in the training set, including EfficientNet, SwinTransformer, SqueezeNet, MobileNet, ResNet, and GoogleNet (Supplementary Table S1). Ultimately, ResNet34 was selected for its superior performance in predicting outcomes from OCT images. The image-based predictions were used as input features, combined with clinical examination data including age, ACA, WTW diameter, CLR, ACD, and SE, and fed into a gradient-boosting tree model to generate the final predictions. The precision, recall, and accuracy of the model in predicting the postoperative vault in patients with shallow ACD following ICL surgery were evaluated using an independent test dataset that the INN model had not previously encountered.

Ablation Study

Ablation experiments assess the impact of individual or combined variables on the overall system. In this study, we performed ablation experiments on OCT images and six clinical examination data or risk factors by selectively omitting specific parameters. We then assessed the performance of the neural network using combinations of retained input parameters. This approach enabled us to document the maximum accuracy attainable on the FedEYE platform with and without certain features, thus establishing the significance of each parameter to the model and its impact on the postoperative vault. Accuracy heatmaps for various parameter combinations were generated using Prism 10 (GraphPad Software, Boston, MA).

translational vision science & technology

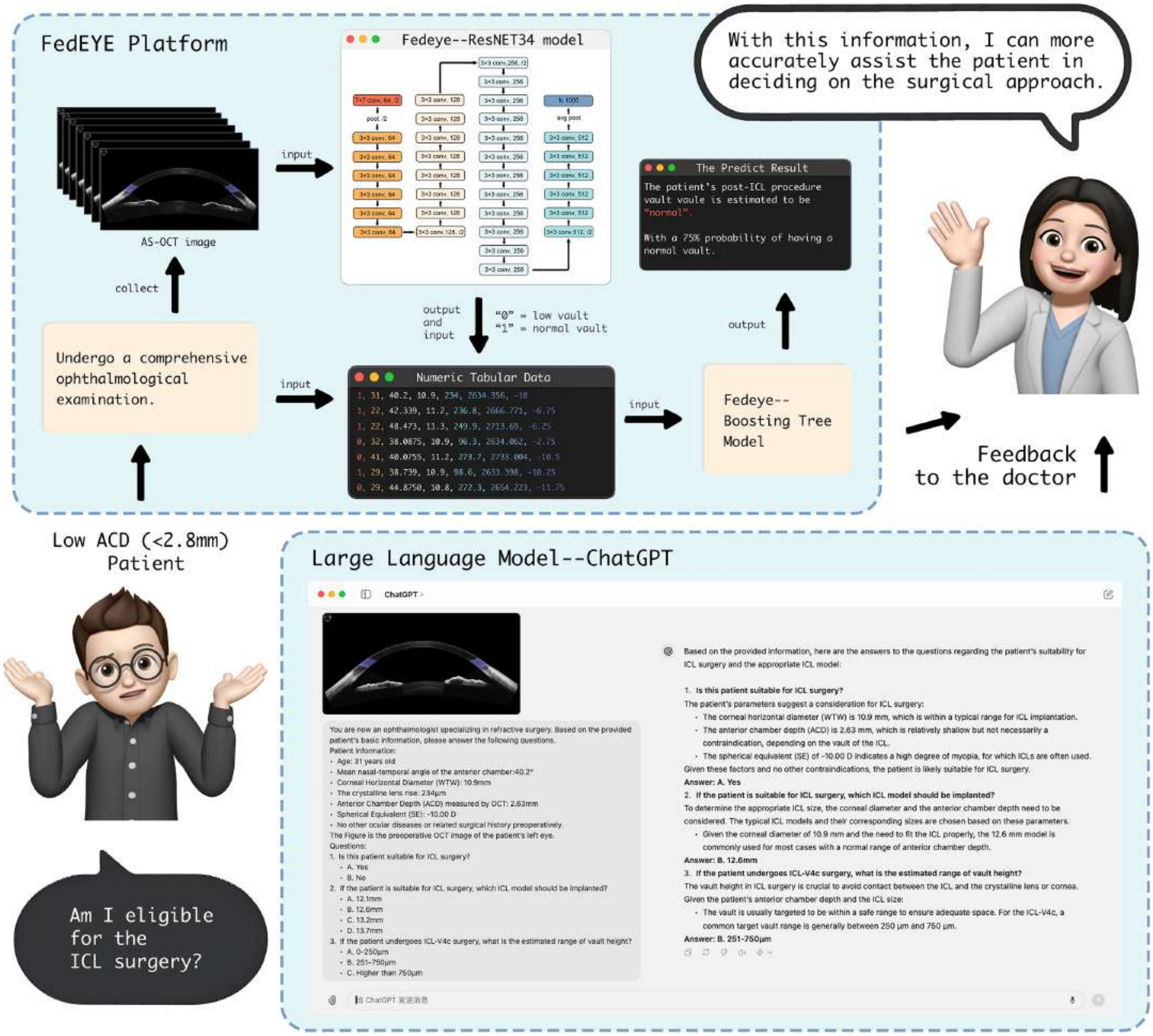

Figure 1. Patients with shallow ACD underwent a comprehensive ophthalmological examination to select and preprocess the training data, which were subsequently input into the FedEYE platform and ChatGPT. For FedEYE, OCT images were initially processed using ResNet34; the outputs were then merged with numerical tabular data and input into a gradient boosting decision tree for final result generation and performance evaluation. In the case of LLM, ChatGPT was employed to conduct case-based question-and-answer sessions using the same multimodal data, aimed at predicting postoperative vault and assessing the outcomes.

Principal Component Analysis

Principal component analysis (PCA) is a statistical method that utilizes orthogonal linear transformations to cluster data based on similarities within the dataset. This method efficiently summarizes the majority of the variance of the dataset into fewer dimensions, known as principal components (PCs), using transformed coordinates (scores).[25] In this research, PCA was applied to all OCT images along with associated clinical risk factors incorporated in the FedEYE platform and postoperative ICL vault within the training set. By analyzing the relationships and clustering trends among various components, we delineated the data structure, quantified the contribution of each parameter to the postoperative ICL outcomes, and confirmed the correlations between these parameters and the vault. The results were visually represented through biplot graphs generated in Prism 10.

ChatGPT

In this study, we utilized the official web versions of ChatGPT-4o and ChatGPT-4.5 provided by OpenAI (https://chatgpt.com), both of which support image input. All interactions were conducted using the default settings of the platform. To enhance the text generation capabilities of the LLMs for interpreting and responding to specific queries effectively, we refined the prompts used for ChatGPT. Relevant parameters were prepared based on the clinical examination data of the participants' affected eyes, and all information was structured in the form of case reports to ensure reproducibility under the same model version and prompt conditions (see example in Fig. 1, Supplementary Fig. S1). First, the training dataset previously used for the FedEYE model was incorporated into ChatGPT memory (Supplementary Fig. S2). The validation set was then used for prompt optimization and intermediate validation. Finally, an independent test set, neither stored in memory nor used for prompt optimization throughout the process, was employed to predict postoperative vault and associated probabilities (detailed workflow is shown in Supplementary Fig. S3). To standardize the influence of prediction probabilities on the final results, all predicted outcomes were normalized to derive a weighted predicted vault, which was subsequently used for analysis.

Statistical Analysis

In both models, the same training set was utilized for model training or prompting, followed by evaluation on the same validation set. Subsequently, an independent test set was used for further validation, data analysis, and result presentation. In this study, accuracy refers to the proportion of correctly predicted postoperative vault among all treated eyes. Recall, or sensitivity, is defined as the proportion of correctly identified cases of low postoperative vault out of all such cases. Similarly, specificity represents the proportion of correctly identified normal vault among all eyes with normal vaults. The weighted F1 score was used to evaluate the performance of the model by balancing precision and recall. Data processing and statistical analyses were performed using SPSS Statistics 27 (IBM, Chicago, IL) and Prism 10, and graphical illustrations were created with Prism 10 and Photoshop 2021 (Adobe, San Jose, CA). Receiver operating characteristic (ROC) curves were generated to evaluate the probabilities of postoperative vault predictions from various models using an independent test set. This methodology was applied to assess the performance of different models, including the one developed by the FedEYE platform and the other one by ChatGPT. $P < 0.05$ was considered statistically significant.

Results

We collected data for 160 eyes with shallow ACD, all of which underwent STAAR ICL V4c implantation. Among these, 65 eyes exhibited an ICL vault of <250 µm. Table 1 presents the distribution of these parameters across the training, validation, and test sets following random allocation. The average preoperative ACD for all patients was 2.61 ± 0.11 mm (range,

Table 1. Distribution of Parameter Characteristics

	Training Set ($n = 128$)		Validation Set ($n = 16$)		Test Set ($n = 16$)	
Parameter	Mean ± SD	Range	Mean ± SD	Range	Mean ± SD	Range
Age (y)	33.82 ± 5.93	20 to 48	32.69 ± 5.90	24 to 46	33.44 ± 6.47	24 to 42
ACA (°)	39.32 ± 4.50	29.18 to 52.69	38.28 ± 2.42	34.94 to 44.63	39.06 ± 4.73	31.23 to 49.40
WTW diameter (mm)	11.16 ± 0.34	10.4 to 12.2	11.09 ± 0.26	10.5 to 11.5	11.29 ± 0.47	10.4 to 12.2
CLR (µm)	269.36 ± 184.75	−171.6 to 683.1	240.55 ± 129.45	−11.5 to 426.9	289.48 ± 247.51	−203.72 to 643.74
ACD (mm)	2.60 ± 0.12	2.31 to 2.79	2.64 ± 0.81	2.46 to 2.76	2.63 ± 0.88	2.42 to 2.77
SE (D)	−8.14 ± 3.05	−15.00 to −2.75	−8.32 ± 3.31	−15.00 to −4.00	−8.05 ± 2.10	−11.50 to −3.75
Vault (µm)	290.93 ± 180.52	24.48 to 746.38	320.75 ± 181.51	59.47 to 672.48	295.31 ± 199.29	28.74 to 643.61

Table 2. Predictive Results of Models Using Solely Images, Data, and Combination of Both Images and Data on FedEYE Platform

Model	Accuracy	Sensitivity	Specificity	AUC	95% CI	*P*
OCT image	0.813	0.857	0.778	0.817	0.593–1.000	0.034*
Data	0.750	0.714	0.778	0.746	0.489–1.000	0.131
OCT image + data	0.875	0.857	0.889	0.873	0.676–1.000	0.013*

*$P < 0.05$, indicating a statistically significant difference from an AUC of 0.5 and robust diagnostic performance.

translational vision science & technology

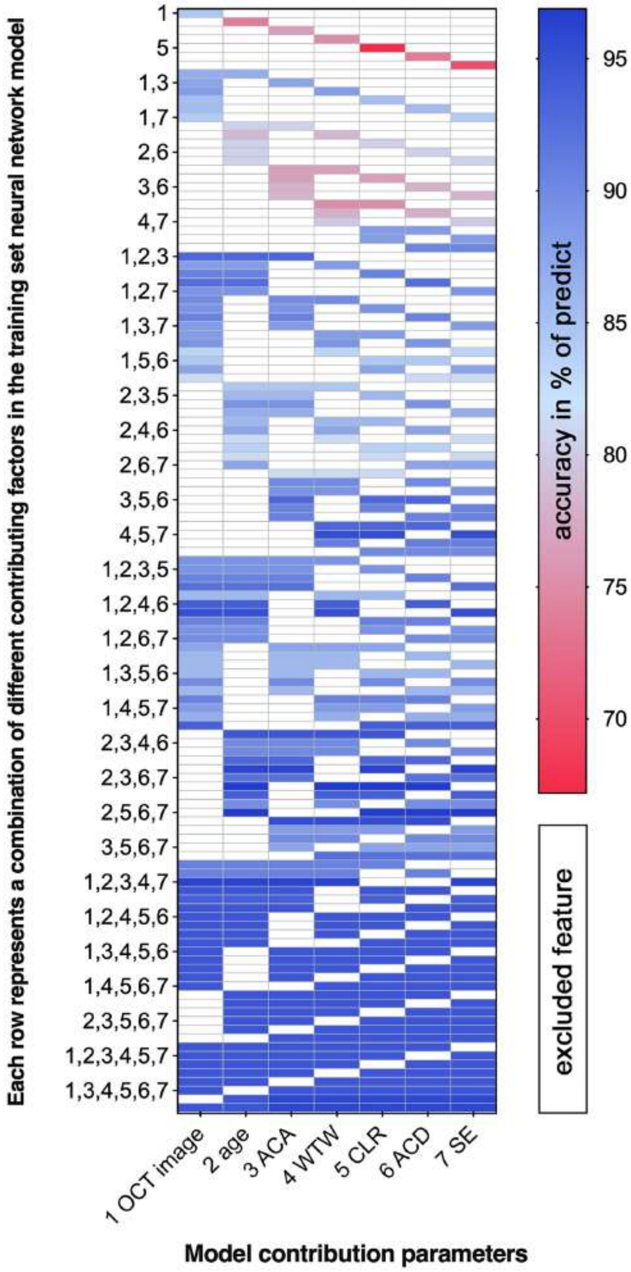

Figure 2. The training data incorporated into the FedEYE INN were visualized using a matrix diagram. This diagram displays 127 possible combinations derived from seven data types, with each row on the *y*-axis representing a distinct combination and the *x*-axis detailing the parameters included in the model. Excluded parameters are shown in *white*. The accuracy indicated in the diagram was automatically generated by the FedEYE platform, following the input and processing of each training set combination.

2.31–2.79). The mean postoperative vault measured 294.35 ± 181.56 µm (range, 24.48–746.38). There were no statistically significant differences in any parameters across the three datasets.

To accurately predict postoperative ICL vault in patients with shallow ACD, we developed an INN

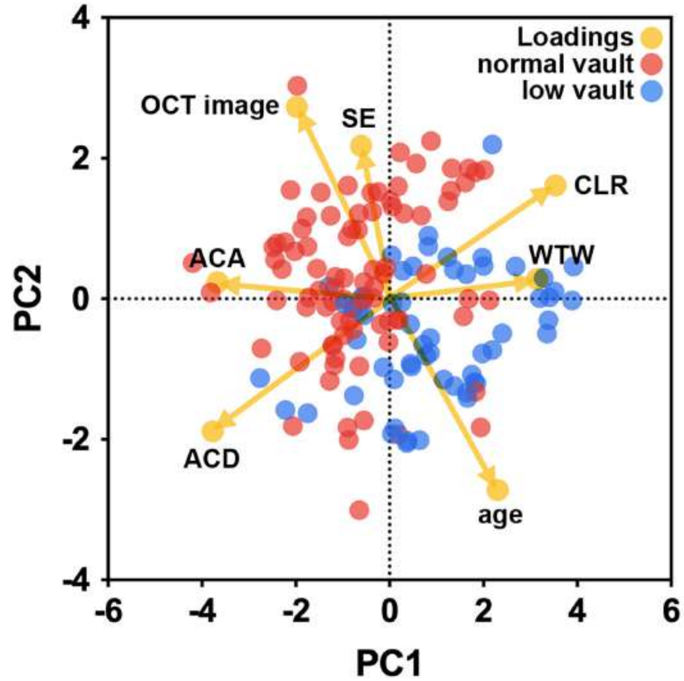

Figure 3. The relationship between the training data integrated into the model and the postoperative vault is visualized using a Biplot figure. The *yellow lines* represent the loadings of each parameter, and the scatter points denote the PC scores. *Red points* indicate normal vault, and *blue points* signify low vault.

on the FedEYE platform. This network utilizes OCT images, clinical examination data, or a combination thereof. We modified the learning models and combinations for OCT images and examination data, ultimately employing a combination of the ResNet34 convolutional network and gradient-boosting trees (Supplementary Table S1). As shown in Table 2, according to the test set results, the multimodal INN model achieved a maximum accuracy of 0.875, sensitivity of 0.857, and specificity of 0.889. ROC curve results showed that approximately seven out of eight patients in our test data were accurately classified as correct postoperative vault (AUC = 0.873). This performance surpassed that of networks relying solely on OCT images (AUC = 0.817) and is markedly better than models using only clinical examination data (AUC = 0.746) (Table 2, Supplementary Fig. S4).

Predictive outcomes derived from multimodal data demonstrate enhanced accuracy over single data type models. Following this discovery, we further investigated the critical role of each parameter included in the FedEYE platform. As depicted in Figure 2, we employed ablation analysis to train a multilayer perceptron using 127 distinct parameter combinations on a consistent training set, documenting the accuracy of each combination as provided by the FedEYE

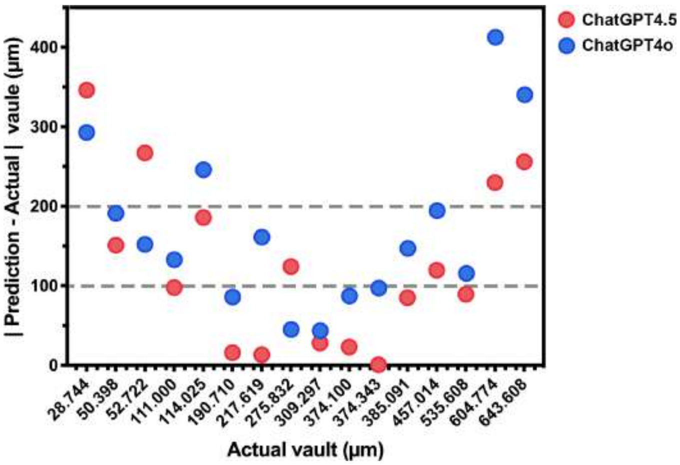

Figure 4. Scatterplots illustrate the ChatGPT predictions for vault. The vertical axis depicts the absolute differences between the weighted predictions and the actual values, whereas the horizontal axis indicates the actual vault from the test set. *Red* symbolizes ChatGPT-4.5, and *blue* signifies ChatGPT-4o.

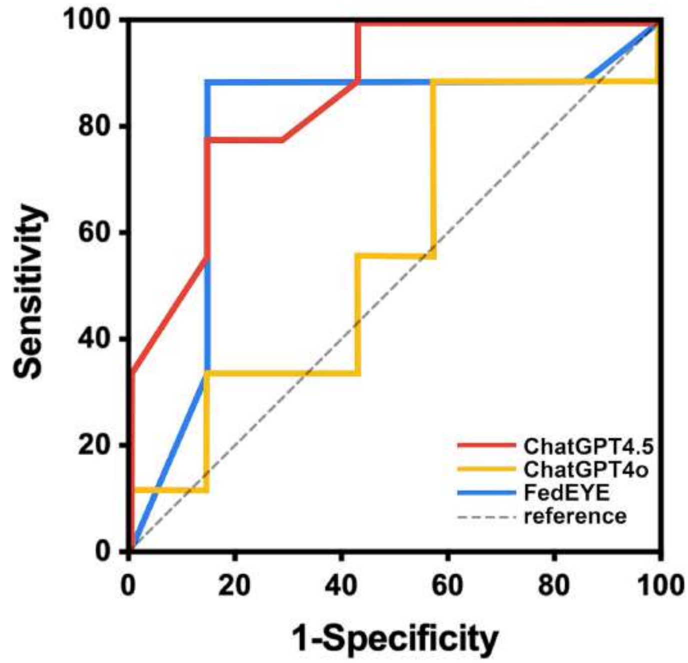

Figure 5. The ROC curves display the performance of three models: ChatGPT-4.5 (*red*), ChatGPT-4o (*yellow*), and FedEYE INN (*blue*).

platform. The findings revealed that integrating multiple parameters significantly improved accuracy; specifically, models incorporating OCT images and clinical examination data achieved the highest mean accuracy (0.96). Subsequently, as shown in Figure 3, we used PCA analysis to further explore the contribution of each parameter and observed the relationships between parameters and the ICL vault. OCT images and age were found to be particularly influential, whereas ACA and SE were crucial in predicting normal vault. Notably, older age and larger WTW diameter were associated with lower postoperative vault.

Given the constraints of limited sample sizes for the rapid development of more precise neural network models, this research leveraged a well-trained LLM for predicting postoperative vaults. It utilized the long-term memory features of ChatGPT to learn from training set samples and evaluated performance before and after prompting on the same test set. The prediction accuracy of ChatGPT-4.5 improved from 0.63 to 0.81, and its F1 weighted score improved from 0.25 to 0.73. Similar improvements were observed with ChatGPT-4o, as the accuracy improved from 0.38 to 0.56 and its F1 weighted score from 0.17 to 0.46. (Supplementary Table S2). As shown in Figure 4 and Supplementary Figure S5, after prompting, both GPT models achieved a 75% prediction accuracy for postoperative vault within ±200 μm of the actual values. In particular, ChatGPT-4.5 achieved 50% accuracy within ±100 μm, and ChatGPT-4o achieved 31.25% within this range. A comparative analysis between the trained GPT models and the INN model developed on the FedEYE platform showed that ChatGPT-4.5 achieved accuracy of 0.81, recall of 0.57, F1 weighted score of 0.73, and AUC of 0.865, significantly surpassing ChatGPT-4o, which scored accuracy of 0.56, recall of 0.43, F1 score of 0.46, and AUC of 0.57, and it slightly underperformed in comparison to the FedEYE model, which recorded an AUC of 0.873 (Table 3; Fig. 5).

Table 3. Accuracy of Vault Predictions by ChatGPT-4.5, ChatGPT-4o, and FedEYE Model

AI Model	Accuracy (%)	Recall (%)	Weighted F1 Score	AUC	95% CI	*P*
ChatGPT-4.5+	81.3	57.1	0.73	0.865	0.681–1.000	0.015*
ChatGPT-4o+	56.3	42.9	0.46	0.571	0.274–0.869	0.634
FedEYE	87.5	85.7	0.86	0.873	0.676–1.000	0.013*

The use of memory is indicated by a plus sign (+).
*$P < 0.05$, indicating a statistically significant difference from an AUC of 0.5 and robust diagnostic performance.

Discussion

In this study, we utilized the FedEYE platform and ChatGPT to develop two code-free multimodal AI models using preoperative OCT images and a collection of examination data, including age, ACA, CLR, WTW diameter, ACD, and SE. An INN model developed using the FedEYE platform and ChatGPT accompanied by long-term memory was used to predict the postoperative vault in patients with shallow ACD. Through systematic validation of various network architectures on the FedEYE platform and an extensive evaluation involving 160 eyes, the FedEYE model achieved a favorable accuracy of 0.875 and AUC of 0.873. Simultaneously, ChatGPT, trained on the same dataset, achieved a very high degree of accuracy of 0.813 and AUC of 0.865.

In the assessment of ICL surgery, postoperative vault is a critical parameter for evaluating both surgical success and safety. Excessively high vault may damage endothelial cells and precipitate secondary glaucoma, whereas excessively low vault can lead to ASC.[26,27] In studies investigating ICL implantation in eyes with shallow ACD, Qian et al.[12] reported that 3.07% of patients developed anterior lens capsule opacification, and all of these patients had ICL vault < 250 µm. Our previous research also demonstrated that these patients are more prone to developing low vault. Thus, selecting an appropriate ICL size based on preoperative examinations and accurately predicting vault have received significant attention. With AI increasingly being incorporated into scientific research, challenges previously unresolved by conventional methods are now being overcome.[28] Chen et al.[29] utilized an optimized random forest (RF) regression model spanning sulcus to sulcus to predict postoperative vault, achieving an accuracy of 0.63 and AUC of 0.80. Similarly, Shen et al.,[30] using a random RF model to predict ICL vault and selected lens sizes, attained a maximum classification accuracy of 0.828 and an AUC of 0.765. These results are comparable to the results of our single-type data-driven model in the present study (Table 2), which recorded accuracy of 0.75 and AUC of 0.746. In addition to ML models, numerous studies have applied deep learning to automate recognition and prediction tasks using images.[17,31] As far as we know, however, there are still no predictive models currently existing for patients with shallow ACD, which limits the clinical evidence available for precise safety assessments by refractive surgeons in these shallow ACD patients.[32] Therefore, in this study, we retrospectively collected preoperative clinical data and the 1-month postoperative vault for this patient group and developed an AI model that is more suitable for these shallow ACD patients. We utilized the FedEYE platform, which supports the code-free development and deployment of various deep learning and ML models, such as SwinTransformer and ViT, simplifying research and enhancing collaboration among medical researchers.[33] Using OCT images and individual clinical examination data, we developed an INN model with ResNet34 and gradient boosting trees that achieved accuracy of 0.875 and AUC of 0.873. Compared to the models using only single-type data, their performance in terms of AUC is 0.817 and 0.746, respectively. This is consistent with numerous findings that multimodal data may yield more precise results compared to models using single data types.[34,35]

To evaluate the influence of various parameters on the model developed on the FedEYE platform and to elucidate the correlations between clinical examination data and the ICL vault, we performed an ablation analysis. This analysis assessed the accuracy of the model across 127 parameter combinations and revealed that incorporating imaging data markedly improved accuracy compared with the use of examination parameters alone. In terms of clinical risk factors, including age, ACA, WTW diameter, and SE, these all enhanced model accuracy, confirming reports in the existing literature on the relationship between these parameters and vault.[36,37] Additionally, we further analyzed these relationships using PCA analysis. Beyond OCT images contributing to increased predictive accuracy, patients with shallow ACD characterized by larger WTW diameter and smaller ACA were found to have lower postoperative vault, consistent with our previous findings. Age was also identified as a risk factor for lower vault, confirming prior research[26] and suggesting that older patients should be monitored more closely postoperatively.[38] Numerous studies have confirmed the correlation between ACD and CLR parameters with postoperative vault.[37,39] However, these parameters made relatively minor contributions compared to other factors in the model. We hypothesize that the fundamental characteristics of subjects included in different studies may be the primary reason for this. Eyes with shallow ACD can be characterized by lower corneal sagittal heights and higher CLR.[40] Most analyses of ICL surgery consider patients with ACD > 2.8 mm. In contrast, our study examined a more limited range of ACD (2.31–2.79 mm), leading to an insignificant correlation among ACD, CLR, and vault.

The scarcity of patients with shallow ACD poses challenges in acquiring large sample sizes quickly for prompting AI models. ChatGPT integrates extensive datasets to train on text and images, extracting crucial information and generating human-like text via natural language response algorithms.[19] These training sets are rich in medical records and imaging data, offering significant prospects for disease diagnosis and prediction.[41,42] Given their unique image recognition and analysis capabilities, we used ChatGPT-4.5 and ChatGPT-4o for result analysis in this study. ChatGPT results heavily depend on the questions posed at the end of case reports.[43,44] Consequently, we optimized the prompt to standardize case reports and questionnaires, enabling ChatGPT to provide accurate predictions and probability, thereby obtaining the weighted predictive vault for downstream analysis. Our findings suggest that ChatGPT-4.5 outperforms ChatGPT-4o, likely due to more refined algorithmic logic.[45] In this study, we used the "memory" feature of ChatGPT to enable sequential input of training sets into the memory of the model, thus enabling it to diagnose based on experience similar to a clinical physician. The ChatGPT-4.5 weighted F1 score improved from 0.25 to 0.73 after memory integration, and ChatGPT-4o showed similar improvement (from 0.17 to 0.46). This long-term memory function could enable more personalized prompting of LLM, enhancing their sensitivity and specificity. Currently, research on LLMs is limited to comparing different question-and-answer formats and model versions,[45–47] using it as a "trained" model alongside other deep learning models to provide more valuable insights for LLM applications. In this study, we input the same training set used for the FedEYE model into the ChatGPT "memory" and then applied identical validation and test sets for predicting postoperative vault. ChatGPT-4.5 with memory showed good predictive performance, achieving accuracy of 0.813 and AUC of 0.865, slightly lower than the FedEYE model (accuracy of 0.875, AUC of 0.873). In addition, OpenAI has released large-scale reasoning models for ChatGPT focused on complex tasks. We evaluated the performance of GPT-o1 in predicting vault in patients with shallow ACD using an independent test dataset that had not been incorporated into GPT memory (Supplementary Fig. S6). The weighted F1 score of its predictions was 0.28, surpassing those of no-memory-trained GPT 4.5 (0.25) and GPT-4o (0.17). These results suggest that combining memory and reasoning capabilities may enable more effective prompting of LLMs in the future.

Overall, our proposed AI model significantly improved the predictability of postoperative vault in shallow ACD patients. Currently, we are conducting additional research to address several limitations of the present study. One limitation is the limited number of cases incorporated into our model, due to the relative rarity of patients with shallow ACD undergoing ICL surgery. In the future, we plan to conduct multicenter studies to establish a larger and more comprehensive database, while continuously assessing the performance of the model on external datasets and refining it through iterative optimization. It is noteworthy that we achieved excellent model performance with limited annotated data. In this study, we utilized the FedEYE platform, which allows clinical physicians to train models without coding. The federated nature of this platform ensures that various centers can easily be invited to join as federated parties to collaboratively train the federated model through convenient deployment methods. Through this, we also hope more centers performing ICL on patients with shallow ACD will join us in developing a larger sample size and therefore more accurate neural network. Furthermore, although ChatGPT benefits from extensive training data and sophisticated algorithms that mimic human thought processes, the reliability of its extensive data remains unverified. In this study, we have improved its accuracy with the "memory" function, and we will also conduct a more comprehensive evaluation of the predictions of LLM in the future.

In conclusion, through the FedEYE platform, we trained a multimodal model to predict the postoperative vault in patients with shallow ACD undergoing ICL surgery, using a code-free approach, and it exhibited excellent performance. Our analysis identified significant correlations among age, ACA, WTW diameter, and SE with vault. Additionally, we discovered that the "memory" function of ChatGPT-4.5 exhibits excellent performance in predicting postoperative outcomes, suggesting that developing specific LLMs holds substantial potential for clinical practice. Integrating multimodal data, with more convenient deployment methods and reduced reliance on annotated data, will be crucial for achieving superior model performance in the future.

Acknowledgments

Supported by grants from the Natural Science Foundation of Hunan Province (2021JJ30048), Science and Technology Innovation Program of Hunan Province (2024RC3249), and Aier Eye Group Research Project Fund (No. AR2209D3 AIG2303D04).

Disclosure: **Z. Su**, None; **Z. Wang**, None; **S. Wang**, None; **M. Ni**, None; **J. Hu**, None; **R. Wang**, None; **J. Yang**, None; **Y. Bai**, None; **Y. Xu**, None; **S. Di**, None; **Y. Xu**, None; **Z. Wang**, None; **Z. Chen**, None; **J. Ortega-Usobiaga**, None; **M.X. Wang**, None; **X. Jiang**, None; **Y. Chen**, None; **W. Dai**, None; **J. Chen**, None; **K. Li**, None

References

1. Sanders DR, Vukich JA, Doney K, Gaston M. U.S. Food and Drug Administration clinical trial of the Implantable Contact Lens for moderate to high myopia. *Ophthalmology.* 2003;110(2):255–266.
2. Packer M. Meta-analysis and review: effectiveness, safety, and central port design of the intraocular collamer lens. *Clin Ophthalmol.* 2016;10:1059–1077.
3. Choi JH, Lim DH, Nam SW, Yang CM, Chung ES, Chung TY. Ten-year clinical outcomes after implantation of a posterior chamber phakic intraocular lens for myopia. *J Cataract Refract Surg.* 2019;45(11):1555–1561.
4. Packer M. The Implantable Collamer Lens with a central port: review of the literature. *Clin Ophthalmol.* 2018;12:2427–2438.
5. Alfonso JF, Lisa C, Fernández-Vega L, Almanzar D, Pérez-Vives C, Montés-Micó R. Prevalence of cataract after collagen copolymer phakic intraocular lens implantation for myopia, hyperopia, and astigmatism. *J Cataract Refract Surg.* 2015;41(4):800–805.
6. Goukon H, Kamiya K, Shimizu K, Igarashi A. Comparison of corneal endothelial cell density and morphology after posterior chamber phakic intraocular lens implantation with and without a central hole. *Br J Ophthalmol.* 2017;101(11):1461–1465.
7. Guber I, Mouvet V, Bergin C, Perritaz S, Othenin-Girard P, Majo F. Clinical outcomes and cataract formation rates in eyes 10 years after posterior phakic lens implantation for myopia. *JAMA Ophthalmol.* 2016;134(5):487–494.
8. Hirabayashi M, Virdi G, Abramson AJ, Nasser T, Vanrachack A, Parkhurst G. Clinical outcomes of implantable collamer lenses for the treatment of myopia in eyes with anterior chamber depth (<3.0 mm) at a single center in the United States. *Clin Ophthalmol*. 2025;19:1985–1990.
9. Moshirfar M, Han KD, Jaafar MA, et al. Comparative evaluation of multiple nomograms for predicting postoperative vault after implantable collamer lens surgery. *J Cataract Refract Surg.* 2024;50(1):64–71.
10. Thompson V, Cummings AB, Wang X. Implantable collamer lens procedure planning: a review of global approaches. *Clin Ophthalmol.* 2024;18:1033–1043.
11. Yong KL, Gong T, Nongpiur ME, et al. Myopia in Asian subjects with primary angle closure: implications for glaucoma trends in East Asia. *Ophthalmology*. 2014;121(8):1566–1571.
12. Qian T, Du J, Ren R, et al. Vault-correlated efficacy and safety of Implantable Collamer Lens V4c implantation for myopia in patients with shallow anterior chamber depth. *Ophthalmic Res.* 2023;66(1):445–456.
13. Ouchi M. Vault of the phakic intraocular lens during vertical and horizontal fixation within patient comparison. *Sci Rep*. 2025;15(1):10002.
14. Su Z, Wang S, Li S, et al. Anterior segment characteristics changes after Implantable Collamer Lens V4c in low anterior chamber depth eyes: an optical coherence tomography study. *J Refract Surg.* 2025;41(4):e342–e355.
15. Hamet P, Tremblay J. Artificial intelligence in medicine. *Metabolism.* 2017;69S:S36–S40.
16. Wang H, Fu T, Du Y, et al. Scientific discovery in the age of artificial intelligence. *Nature.* 2023;620(7972):47–60.
17. Assaf JF, Reinstein DZ, Zakka C, et al. Deep learning-based estimation of implantable collamer lens vault using optical coherence tomography. *Am J Ophthalmol.* 2023;253:29–36.
18. Hartmann LM, Langhans DS, Eggarter V, et al. Keratoconus progression determined at the first visit: a deep learning approach with fusion of imaging and numerical clinical data. *Transl Vis Sci Technol.* 2024;13(5):7.
19. Thirunavukarasu AJ, Ting DSJ, Elangovan K, Gutierrez L, Tan TF, Ting DSW. Large language models in medicine. *Nat Med.* 2023;29(8):1930–1940.
20. Mahowald K, Ivanova AA, Blank IA, Kanwisher N, Tenenbaum JB, Fedorenko E. Dissociating language and thought in large language models. *Trends Cogn Sci.* 2024;28(6):517–540.
21. Liu J, Wang C, Liu S. Utility of ChatGPT in clinical practice. *J Med Internet Res.* 2023;25:e48568.
22. Gonzalez-Lopez F, Bouza-Miguens C, Tejerina V, et al. Dynamic assessment of variations in pupil diameter using swept-source anterior segment optical coherence tomography after phakic collamer lens implantation. *Eye Vis (Lond).* 2021;8(1):39.
23. Li K, Wang Z, Wang MX. Implantable collamer lens implantation (ICL) versus small incision lenticule extraction (SMILE) in low to moderate myopia: study protocol for a randomized, non-inferiority trial. *Trials.* 2022;23(1):910.
24. Li K, Wang Z, Zhang D, et al. Visual outcomes and corneal biomechanics after V4c implantable collamer lens implantation in subclinical keratoconus. *J Cataract Refract Surg.* 2020;46(10):1339–1345.

translational vision science & technology

25. Jolliffe IT, Cadima J. Principal component analysis: a review and recent developments. *Philos Trans A Math Phys Eng Sci*. 2016;374(2065):20150202.
26. Manito SC, Trancón AS, Sierra OT, Baptista AM, Serra PM. Biometric and ICL-related risk factors associated to sub-optimal vaults in eyes implanted with implantable collamer lenses. *Eye Vis (Lond)*. 2021;8(1):26.
27. Moshirfar M, Webster CR, Ronquillo YC. Phakic intraocular lenses: an update and review for the treatment of myopia and myopic astigmatism in the United States. *Curr Opin Ophthalmol*. 2022;33(5):453–463.
28. Wang HC, Fu TF, Du YQ, et al. Scientific discovery in the age of artificial intelligence. *Nature.* 2023;620(7972):47–60.
29. Chen X, Shen Y, Jiang YJ, et al. Predicting vault and size of posterior chamber phakic intraocular lens using sulcus to sulcus-optimized artificial intelligence technology. *Am J Ophthalmol.* 2023;255:87–97.
30. Shen Y, Wang L, Jian WJ, et al. Big-data and artificial-intelligence-assisted vault prediction and EVO-ICL size selection for myopia correction. *Br J Ophthalmol.* 2023;107(2):201–206.
31. Nasser T, Hirabayashi M, Virdi G, Abramson A, Parkhurst G. VAULT: vault accuracy using deep learning technology: new image-based artificial intelligence model for predicting implantable collamer lens postoperative vault. *J Cataract Refract Surg.* 2024;50(5):448–452.
32. Lim DH, Lee MG, Chung ES, Chung TY. Clinical results of posterior chamber phakic intraocular lens implantation in eyes with low anterior chamber depth. *Am J Ophthalmol.* 2014;158(3):447–454.
33. Yan BJ, Cao DM, Jiang XL, et al. FedEYE: a scalable and flexible end-to-end federated learning platform for ophthalmology. *Patterns (N Y)*. 2024;5(2):100928.
34. Hartmann LM, Langhans DS, Eggarter V, et al. Keratoconus progression determined at the first visit: a deep learning approach with fusion of imaging and numerical clinical data. *Transl Vis Sci Technol.* 2024;13(5):7.
35. Chia MA, Antaki F, Zhou YK, Turner AW, Lee AY, Keane PA. Foundation models in ophthalmology. *Br J Ophthalmol.* 2024;108(10):1341–1348.
36. Alfonso JF, Fernández-Vega L, Lisa C, Fernandes P, Jorge J, Micó RM. Central vault after phakic intraocular lens implantation: correlation with anterior chamber depth, white-to-white distance, spherical equivalent, and patient age. *J Cataract Refract Surg.* 2012;38(1):46–53.
37. Trancón AS, Manito SC, Sierra OT, Baptista AM, Serra PM. Determining vault size in implantable collamer lenses: preoperative anatomy and lens parameters. *J Cataract Refract Surg.* 2020;46(5):728–736.
38. Ye YH, Zhao J, Niu LL, Shi WR, Wang XY, Zhou XT. Long-term evaluation of anterior lens density after implantable collamer lens V4c implantation in patients with myopia over 40 years old. *Br J Ophthalmol.* 2022;106(11):1508–1513.
39. Gonzalez-Lopez F, Bilbao-Calabuig R, Mompean B, Luezas J, Ortega-Usobiaga J, Druchkiv V. Determining the potential role of crystalline lens rise in vaulting in posterior chamber phakic collamer lens surgery for correction of myopia. *J Refract Surg*. 2019;35(3):177–183.
40. Sánchez-Trancón A, Manito SC, Sierra OT, Baptista AM, Serra PM. Influence of anterior chamber depth and vault on anterior chamber angle morphology after phakic posterior chamber intraocular lens implantation. *Int Ophthalmol*. 2024;44(1):15.
41. Delsoz M, Madadi Y, Munir WM, et al. Performance of ChatGPT in diagnosis of corneal eye diseases. *Cornea*. 2024;43(5):664–670.
42. Liu L, Qu SH, Zhao HY, et al. Global trends and hotspots of ChatGPT in medical research: a bibliometric and visualized study. *Front Med (Lausanne)*. 2024;11:1406842.
43. Huang X, Raja H, Madadi Y, et al. Predicting glaucoma before onset using a large language model chatbot. *Am J Ophthalmol.* 2024;266:289–299.
44. Huang X, Raja H, Madadi Y, et al. Reply to comment on: predicting glaucoma before onset using a large language model chatbot. *Am J Ophthalmol.* 2024;266:322–323.
45. Lim ZW, Pushpanathan K, Yew SME, et al. Benchmarking large language models' performances for myopia care: a comparative analysis of ChatGPT-3.5, ChatGPT-4.0, and Google Bard. *EBioMedicine.* 2023;95:104770.
46. Walker HL, Ghani S, Kuemmerli C, et al. Reliability of medical information provided by ChatGPT: assessment against clinical guidelines and patient information quality instrument. *J Med Internet Res.* 2023;25:e47479.
47. Pushpanathan K, Lim ZW, Er Yew SM, et al. Popular large language model chatbots' accuracy, comprehensiveness, and self-awareness in answering ocular symptom queries. *iScience.* 2023;26(11):108163.