



OPEN Machine learning regression algorithms to predict short-term efficacy after anti-VEGF treatment in diabetic macular edema based on real-world data

Ruijie Shi¹, Xiangjie Leng¹, Yanxia Wu^{1,2}, Shiyin Zhu¹, Xingcan Cai¹ & Xuejing Lu^{1,2,3,4}✉

The objective of this retrospective study was to predict short-term efficacy of anti-vascular endothelial growth factor (VEGF) treatment in diabetic macular edema (DME) using machine learning regression models. Real-world data from 279 DME patients who received anti-VEGF treatment at Ineye Hospital of Chengdu University of TCM between April 2017 and November 2022 were analyzed. Eight machine learning regression models were established to predict four clinical efficacy indicators. The accuracy of the models was evaluated using mean absolute error (*MAE*), mean square error (*MSE*) and coefficient of determination score (*R*²). Multilayer perceptron had the highest *R*² and lowest *MAE* among all models. Regression tree and lasso regression had similar *R*², with lasso having lower *MAE* and *MSE*. Ridge regression, linear regression, support vector machines and polynomial regression had lower *R*² and higher *MAE*. Support vector machine had the lowest *MSE*, while polynomial regression had the highest *MSE*. Stochastic gradient descent had the lowest *R*² and high *MAE* and *MSE*. The results indicate that machine learning regression algorithms are valuable and effective in predicting short-term efficacy in DME patients through anti-VEGF treatment, and the lasso regression is the most effective ML algorithm for developing predictive regression models.

Diabetes mellitus (DM) is a systemic disease associated with various health problems, which lower life quality and cause a higher mortality rate among patients¹. According to a global report on diabetes by World Health Organization (WHO), there were 451 million adults with DM worldwide in 2017 and this number was expected to increase to 693 million by 2045². Moreover, almost half of the population (49.7%) may live with undiagnosed DM. Diabetic retinopathy (DR) is a specific microvascular complication of DM, which has emerged as the primary cause of vision loss in the general population across various countries^{3–6}.

Diabetic macular edema (DME) is typified by retinal thickening, hard exudates, microaneurysms, and macular hemorrhage, can manifest at any stage of DR and often results in severe vision loss^{7–9}. In developing countries, the risk of vision loss is higher than developed country due to disparities in income levels and medical conditions¹⁰. The etiology and pathogenesis of DME are multifaceted. A critical factor in the pathogenesis of DR is the hyperglycemic state of DM, which triggers microangiopathy by upregulating various inflammatory and angiogenic mediators, particularly vascular endothelial growth factor (VEGF), in the vascular endothelium. These mediators induce pathological alterations in the blood-retinal barrier, resulting in fluid leakage into the extracellular space¹¹, leading to macular edema. Consequently, anti-VEGF therapy is a powerful and effective treatment method against DME¹².

Vitreous cavity injection of anti-VEGF drugs is believed as standard treatment options which can greatly save the patient's visual acuity (VA)^{13–16}. However, anti-VEGF therapy can cause adverse events such as cataract formation, elevated intraocular pressure, retinal artery occlusion, ocular hemorrhage, sterile intraocular inflammation,

¹Department of Ophthalmology, Eye College of Chengdu University of Traditional Chinese Medicine, Chengdu 610000, Sichuan, China. ²Ineye Hospital of Chengdu University of Traditional Chinese Medicine, Chengdu 610000, Sichuan, China. ³Key Laboratory of Sichuan Province Ophthalmopathy Prevention & Cure and Visual Function Protection with Traditional Chinese Medicine, Chengdu 610000, Sichuan, China. ⁴Retinal Image Technology and Chronic Vascular Disease Prevention & Control and Collaborative Innovation Center, Chengdu 610000, Sichuan, China. ✉email: luxuejing@cdutcm.edu.cn

infectious endophthalmitis^{17,18}. Meanwhile, anti-VEGF therapy is expensive and requires repeated treatments. Costs for treatment over a 5-year period were reported in the United States for diabetic retinopathy treated with anti-VEGF and estimated to be \$40,825, which may pose a huge financial burden on patients¹⁹. Meanwhile, not all patients respond equally to the standard treatment of intravitreal therapy²⁰. Therefore, predicting short-term efficacy after anti-VEGF treatment could have great clinical and economic significance.

Artificial intelligence (AI) is one of the most popular topics today and has been widely used in autonomous driving, face recognition, and intelligent robotics^{21,22}. Machine learning (ML) is a branch of AI which aims to learn patterns from data to improve performance in various tasks²³. Compared to traditional statistical methods, ML is more efficient, less susceptible to human factors, and can handle complex data. ML methods have been used extensively to solve various complex challenges in recent years in medical areas, such as prediction of disease efficacy, medical image processing, disease risk prediction, etc.^{24–26}. However, most of the above applications adopted classification algorithms, while the application of the regression algorithm is currently not developed well. Compared to the classification algorithm, the regression algorithm can handle continuous variables with higher precision and broader applicability, which confers it a strong potential and value for clinical development.

At present, the application of ML in DR, particularly in DME, is predominantly concentrated in the realm of diagnostics. The exploration of therapeutic effect predictions, which are based on regression algorithms, is still in the nascent stage. For DME, the implementation of an ML regression prediction model has the potential to furnish doctors and patients with more precise predictive data. This, in turn, can aid in clinical decision-making processes, mitigate clinical risks, and curtail superfluous treatment expenses. Consequently, the utilization of ML for prognostic predictions of DME has great clinical and economic value. In light of these considerations, the objective of this study is to establish ML regression models using real-world data, to predict the short-term efficacy of anti-VEGF treatment for DME.

Methods

All methods were carried out in accordance with all relevant guidelines and regulations, and the study was reviewed and approved by the Academic Committee and the Ethics Committee of Ineye Hospital of Chengdu University of Traditional Chinese Medicine (Ethics number: 2022yh-023). This study has obtained informed consent from all subjects or their legal guardian(s).

Source of data and participants

Patients diagnosed with DME in Ineye Hospital of Chengdu University of Traditional Chinese Medicine from April 2017 to November 2022 were included. After preparing valid data, ML was used to predict the short-term efficacy of DME patients after anti-VEGF treatment.

The inclusion criteria for the patients were: (1) clinical diagnosis of DME based on the Diabetic Retinopathy Preferred Practice Pattern 2019²⁷; (2) receipt of at least one anti-VEGF treatment; (3) age between 18 and 85 years; (4) follow-up period of no more than 3 months. The exclusion criteria were: (1) the presence of other eye disorders that may affect VA, such as glaucoma, age-related macular degeneration (AMD), retinal detachment, etc.; (2) lack of clinical data; (3) refractive interstitial opacification obscuring the macula; (4) undergoing intraocular surgery.

Data collection

We collected data including sex, age, type of anti-VEGF drugs, follow-up time, best corrected visual acuity (BCVA), intraocular pressure (IOP), central subfield thickness (CST), cube volume (CV), cube average thickness (CAT), macular thickness divided by ETDRS grid²⁸ (center, inner superior (IS), inner temporal (IT), inner inferior (II), inner nasal (IN), outer superior (OS), outer temporal (OT), outer inferior (OI), outer nasal (ON), and ganglion cell thickness (GC) analysis (average, minimum, superior, temporal superior, nasal superior, nasal inferior, inferior, temporal inferior). We included both the first and the second visits of BCVA, CST, CV, and CAT as the clinical predictors.

Data preprocessing

BCVA was converted from decimal notation to LogMAR notation for statistical analysis. Count finger, hand movement and light perception were recorded as 2.0, 2.3 and 2.6 in LogMAR notation respectively²⁹. No light perception should be recorded as infinity in LogMAR notation, for calculation, was recorded as 100 in LogMAR notation. 514 missing data were complementary by IBM SPSS statistics 27 using methods of regression to adjust residuals.

Machine learning models and training

We established eight ML regression models, including linear regression, polynomial regression, ridge regression, lasso regression, support vector machines (SVM), regression tree, multilayer perceptron (MLP) and stochastic gradient descent (SGD) regression. Using Python 3.9.0 and Scikit-learn 1.2.0 for training and testing. We used Matplotlib 3.5.2 to draw the figures. We split the data into 75% for training and 25% for testing.

Linear models are regression methods that assume a linear relationship between the features and the target. Ordinary Least Squares is a linear model that minimizes the sum of squared errors between the observed and predicted targets³⁰. Polynomial regression models y as an n th-degree polynomial in x . It can capture nonlinear relationships. It was influential in regression analysis history, with a focus on design and inference³¹. Ridge regression is a method of estimating the coefficients of multiple-regression models in scenarios where the independent variables are highly correlated³². It regularizes ill-posed problems and reduces multicollinearity in linear regression³³. In general, the method provides improved efficiency in parameter estimation problems in exchange

for a tolerable amount of bias³⁴. Lasso is a linear model with sparse coefficients. It reduces the number of features and can recover the exact non-zero coefficients under certain conditions³⁰. SVM is a young and practical branch of statistical learning theory. It transforms low-dimensional nonlinear functions into high-dimensional spaces via a smart nonlinear mapping, without requiring its explicit form^{35,36}. Classification and regression trees are nonparametric regression methods that recursively partition the feature space into rectangular areas³⁷, first proposed by Breiman et al.³⁸. They use binary recursive partitioning to split the data into smaller groups along each branch³⁷. MLP is a feedforward artificial neural network (ANN) with full connectivity and at least three layers: input, hidden and output. Each node except the input ones is a nonlinear neuron. MLP uses backpropagation for supervised learning^{39,40}. It differs from a linear perceptron by its multiple layers and non-linear activation. It can handle non-linearly separable data⁴¹. SGD is a simple yet very efficient approach to fit linear models. It is particularly useful when the number of samples (and the number of features) is very large. The method allows online/out-of-core learning^{30,42}.

Python 3.9.0 and Scikit-learn 1.2.0 were used for training and testing. Matplotlib 3.5.2 was used to draw the figures. The dataset was split into training dataset (75%) and testing dataset (25%).

Evaluating the performance of prediction models

Mean absolute error (*MAE*), mean square error (*MSE*), and coefficient of determination (*R*²) score were used as the evaluation metrics to assess the accuracy of the prediction models.

MAE is a risk metric that corresponds to the expected value of the absolute error loss or -norm loss *l*₁. On the other hand, *MSE* is a risk metric that corresponds to the expected value of the squared (quadratic) error or loss. *R*² value represents the proportion of variance (of *y*) that can be explained by the independent variables in the model, which indicates goodness of fit. A lower value of *MAE* and *MSE* indicates a higher accuracy of the prediction model. Conversely, a higher *R*² score, with the best possible score being 1.0, suggests a better fit of the prediction model. It's worth noting that the *R*² score can be negative if the model's performance is arbitrarily worse. Therefore, a score closer to 1 indicates a better fit of the prediction model^{30,42}.

For a sample of *n* observations *y* (*y*_{*i*}, *i* = 1, 2, . . . , *n*) and *n* corresponding model predictions \hat{y} , the *MAE*, *MSE* and *R*² score are

$$MAE = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i|$$

$$MSE = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2$$

$$R^2 = 1 - \frac{\sum_i (\hat{y}_i - y_i)^2}{\sum_i (\bar{y}_i - y_i)^2}$$

Weighted sum of *MSE*, *MAE* and *R*² score in BCVA, CAT, CST and CV for each model to tally the final score.

Data correlation

Regression coefficients (coef) are estimates of the unknown population parameters and describe the relationship between a predictor variable and the response. Which can be shown in four models (linear regression, ridge regression, lasso regression and SGD). The larger the absolute value of coef, the higher the correlation between the predictor variable and the response variable. A positive coef value indicates a positive correlation with the predicted outcome, while the opposite is a negative correlation.

Ethics approval

This study was reviewed and approved by the Academic Committee and the Ethics Committee of Ineye Hospital of Chengdu University of Traditional Chinese Medicine (Ethics number: 2022yh-023).

Results

Characteristics of dataset

279 eyes were included in our research, the overall characteristics of the dataset are shown in Table 1.

Model performance on various clinical indicators

The results are shown in Table 2. For BCVA, the best model is regression tree, which has zero *MAE* and *MSE* and one *R*² for both sets. The worst model is SVM, which has the highest *MAE* and *MSE* for both sets. The other models have similar performance on the training set but vary on the testing set. For CST, the best model is lasso regression, which has the lowest *MAE* and *MSE* and the highest *R*² for both sets. The worst model is polynomial regression, which has very high *MAE* and *MSE* and negative *R*² for the testing set. The other models have similar performance on both sets. For CV, the best model is MLP, which has the lowest *MAE* and *MSE* and the highest *R*² for both sets. The worst model is polynomial regression, which has very high *MAE* and *MSE* and negative *R*² for the testing set. The other models have similar performance on both sets. For CAT, the best model is MLP, which has the lowest *MAE* and *MSE* and one of the highest *R*² for both sets. The worst model is polynomial

Characteristics	Value
Age, mean \pm SD, years	58.53 \pm 11.546
Sex, no. (%)	
Male	173 (62)
Female	106 (38)
Type of anti-VEGF drugs, no. (%)	
Ranibizumab	60 (22.5)
Conbercept	187 (67)
Aflibercept	32 (11.5)
Follow-up period, mean \pm SD, days	32.16 \pm 23.011
OCT, no. (%)	
CarlZeissMeditec	117(41.9)
SVisionImaging	162(58.1)
BCVA, mean \pm SD	2.55 \pm 13.196
CST, mean \pm SD, μ m	372.609 \pm 158.616
CV, mean \pm SD, mm ³	14.425 \pm 5.529
CAT, mean \pm SD, μ m	369.597 \pm 74.529
IOP, mean \pm SD, mm Hg	16.29 \pm 3.777
Macular thickness (ETDRS grid), mean \pm SD, μ m	
Center	386.92 \pm 152.966
IS	419.078 \pm 113.287
IT	418.482 \pm 127.267
II	409.239 \pm 108.215
IN	415.559 \pm 108.534
OS	380.325 \pm 100.977
OT	375.792 \pm 109.632
OI	353.743 \pm 90.237
ON	388.140 \pm 93.757
Ganglion cell layer thickness, mean \pm SD, μ m	
Average	80.615 \pm 26.761
Temporal superior	66.861 \pm 35.369
Superior	82.023 \pm 31.79
Nasal superior	86.159 \pm 37.422
Nasal inferior	83.378 \pm 37.38
Inferior	74.009 \pm 29.414
Temporal inferior	75.886 \pm 33.86
Minimum	42.639 \pm 36.302

Table 1. Baseline of characteristics.

regression, which has very high *MAE* and *MSE* and negative R^2 for the testing set. The other models have similar performance on both sets. Model performance on the test set is shown in Fig. 1.

Model performance

The visualization of prediction results is shown in Fig. 2. MLP has the highest R^2 (0.1626) and lowest *MAE* (0.0108). Regression tree and lasso regression have similar R^2 (0.1433 and 0.1428), with lasso having lower *MAE* (0.0109) and *MSE* (0.0044). Ridge regression, linear regression, SVM and polynomial regression have lower R^2 (0.1090, 0.0906, 0.0757 and 0.0228) and higher *MAE* (0.0116, 0.0118, 0.0122 and 0.0607). SVM has the lowest *MSE* (0.0043), while polynomial regression has the highest *MSE* (0.3930). SGD has the lowest R^2 (0.0226) and high *MAE* (0.0133) and *MSE* (0.0059).

Data correlation in prediction models

According to our results, the thickness of different areas of the macula and the type of anti-VEGF drug had a large impact on the predicted results. The specific results are as follows:

For BCVA prediction, the highest correlations for linear regression were CAT (0.765), Center (0.619), IS (−0.391) and OT (−0.386); for ridge regression were CAT (0.457), Center (0.286), IS (−0.243) and conbercept (0.232); for lasso regression were OT (−0.047), OS (−0.0420), sex (−0.009), and conbercept (0.005); for SGD were IOP (0.013), sex (−0.012), conbercept (0.011) and CV (0.009).

Clinical indicator	Model	Evaluation index					
		Training set			Testing set		
		MAE	MSE	R ²	MAE	MSE	R ²
BCVA	Linear regression	0.025	0.002	0.887	0.056	0.024	0.412
	Polynomial regression	0.000	0.000	1.000	0.055	0.033	0.182
	Ridge regression	0.022	0.002	0.874	0.048	0.019	0.529
	Lasso regression	0.015	0.004	0.768	0.027	0.014	0.664
	SVM	0.091	0.016	0.137	0.105	0.029	0.282
	Regression tree	0.000	0.000	1.000	0.003	0.000	1.000
	MLP	0.004	0.000	0.998	0.025	0.004	0.900
	SGD	0.029	0.017	0.086	0.051	0.038	0.055
CST	Linear regression	0.094	0.015	0.501	0.107	0.021	0.271
	Polynomial regression	0.000	0.000	1.000	0.554	1.184	− 40.008
	Ridge regression	0.094	0.015	0.500	0.106	0.021	0.273
	Lasso regression	0.098	0.016	0.458	0.103	0.020	0.309
	SVM	0.099	0.017	0.441	0.097	0.019	0.348
	Regression tree	0.000	0.000	1.000	0.145	0.043	− 0.478
	MLP	0.090	0.015	0.515	0.101	0.020	0.318
	SGD	0.122	0.024	0.198	0.117	0.025	0.118
CV	Linear regression	0.037	0.004	0.468	0.044	0.004	0.219
	Polynomial regression	0.000	0.000	1.000	0.144	0.109	− 22.028
	Ridge regression	0.037	0.004	0.462	0.041	0.003	0.312
	Lasso regression	0.035	0.005	0.373	0.036	0.003	0.428
	SVM	0.053	0.006	0.185	0.054	0.005	0.035
	Regression tree	0.000	0.000	1.000	0.043	0.008	− 0.678
	MLP	0.030	0.004	0.526	0.037	0.003	0.348
	SGD	0.036	0.006	0.191	0.045	0.004	0.095
CAT	Linear regression	0.042	0.004	0.660	0.059	0.009	0.289
	Polynomial regression	0.000	0.000	1.000	0.390	1.315	− 98.086
	Ridge regression	0.042	0.004	0.659	0.059	0.010	0.265
	Lasso regression	0.043	0.005	0.627	0.057	0.008	0.433
	SVM	0.053	0.006	0.558	0.064	0.008	0.433
	Regression tree	0.000	0.000	1.000	0.063	0.009	0.339
	MLP	0.032	0.003	0.786	0.058	0.009	0.326
	SGD	0.052	0.006	0.515	0.076	0.012	0.101

Table 2. Model performance on various clinical indicators.

For CST prediction, the highest correlations for linear regression were Center (0.275), CV (−0.250), OI (−0.220) and IS (0.203); for ridge regression were conbercept (0.304), Center (0.281), ranibizumab (0.240) and CV (−0.235); for lasso regression were Center (0.189), CV (−0.188), ON (−0.096) and temporal superior (GC) (0.070); for SGD were conbercept (0.058), IS (0.056), Center (0.055) and IN (0.042).

For CAT prediction, the highest correlations for linear regression were OS (0.288), OT (0.213), temporal superior (GC) (0.165) and ON (0.158); for ridge regression were conbercept (0.260), OS (0.256), aflibercept (0.238) and ranibizumab (0.222); for lasso regression were OS (0.178), OT (0.127), temporal superior (GC) (0.072) and BCVA (0.006); for SGD were conbercept (0.083), OS (0.069), OI (0.070) and OT (0.067).

For CV prediction, the highest correlations for linear regression were CAT (0.295), IS (−0.287), Center (0.259) and conbercept (0.024); for ridge regression were conbercept (0.248), aflibercept (0.214), ranibizumab (0.207) and CAT (0.295); for lasso regression were CAT (0.082), temporal inferior (GC) (0.070), conbercept (0.032) and ON (−0.007); for SGD were conbercept (0.057), OI (0.039), OT (0.032) and OS (0.032). Complete results are shown in Supplementary 1.

Discussion

In this study, we developed eight ML regression models to predict the therapeutic effect of DME patients after anti-VEGF treatment. Our results demonstrate that ML regression algorithms have a high potential for efficacy prediction in diseases. We observed that the overall performance of the eight ML models was better on BCVA than on the other three clinical indicators. A possible reason for this is that BCVA measurements were performed independently by a standard procedure, while other clinical indicators were obtained by two OCT machines that varied in algorithm, clarity, and scan interval. Among all models, MLP and lasso regression models outperformed others. MLP is a type of ANN, which is essential for building nonlinear relationship models in

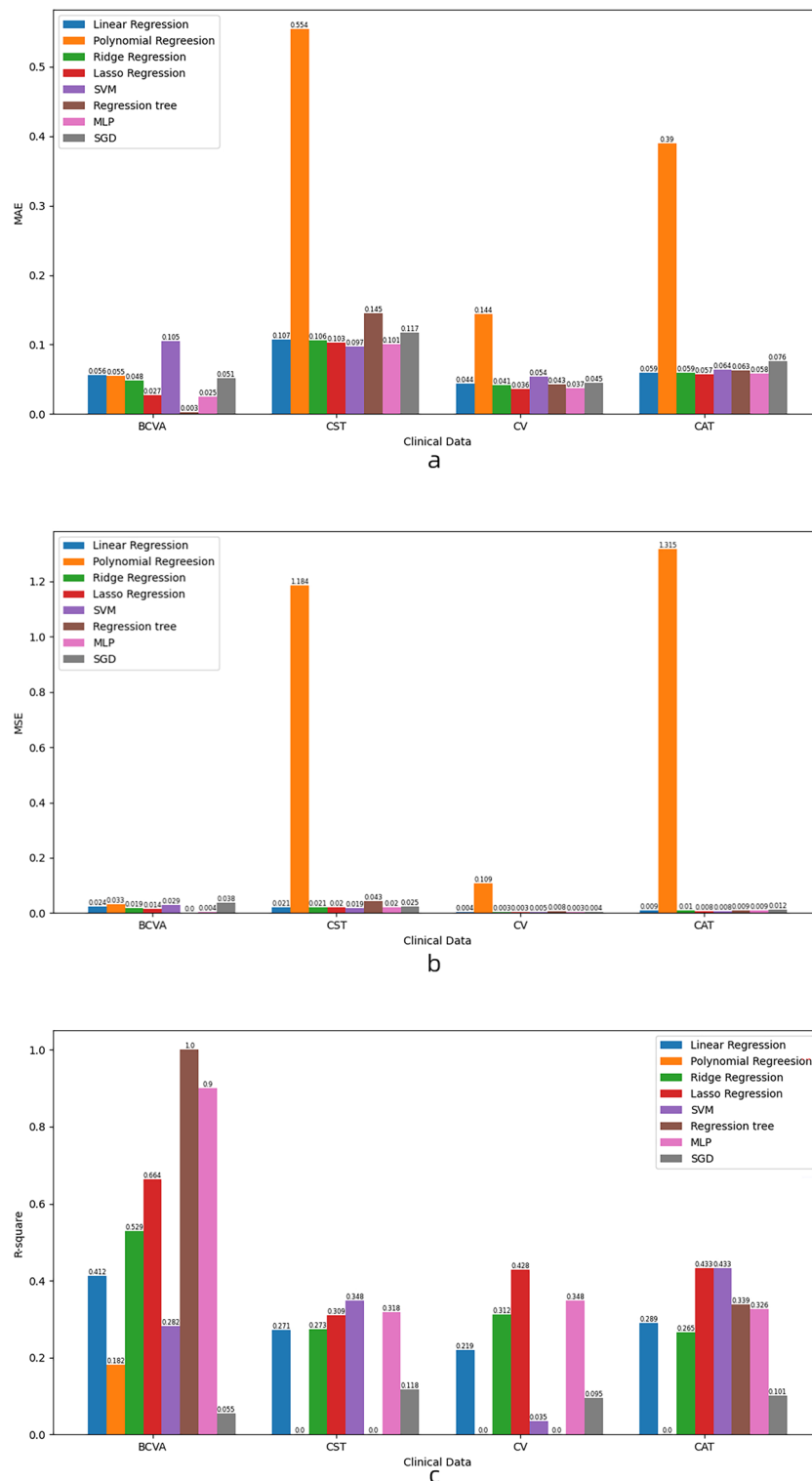


Figure 1. Model performance on the test set. (a) Performance of models with MAE as evaluation index. (b) Performance of models with MSE as evaluation index. (c) Performance of models with R^2 as evaluation index.

high-dimensional datasets⁴³. MLP adds one or more hidden layers on top of the single-layer neural network, which can be iterated. Therefore, it has a greater capacity to learn and generalize, to fit multiple classes of functions, and to predict nonlinear data. Lasso regression has been widely applied in general regression models to predict the risk of likely outcomes⁴⁴. Lasso regression performs both variable selection and regularization to improve the prediction accuracy and interpretability of the resulting model⁴⁵. Thus, lasso regression might be more appropriate for our regression task.

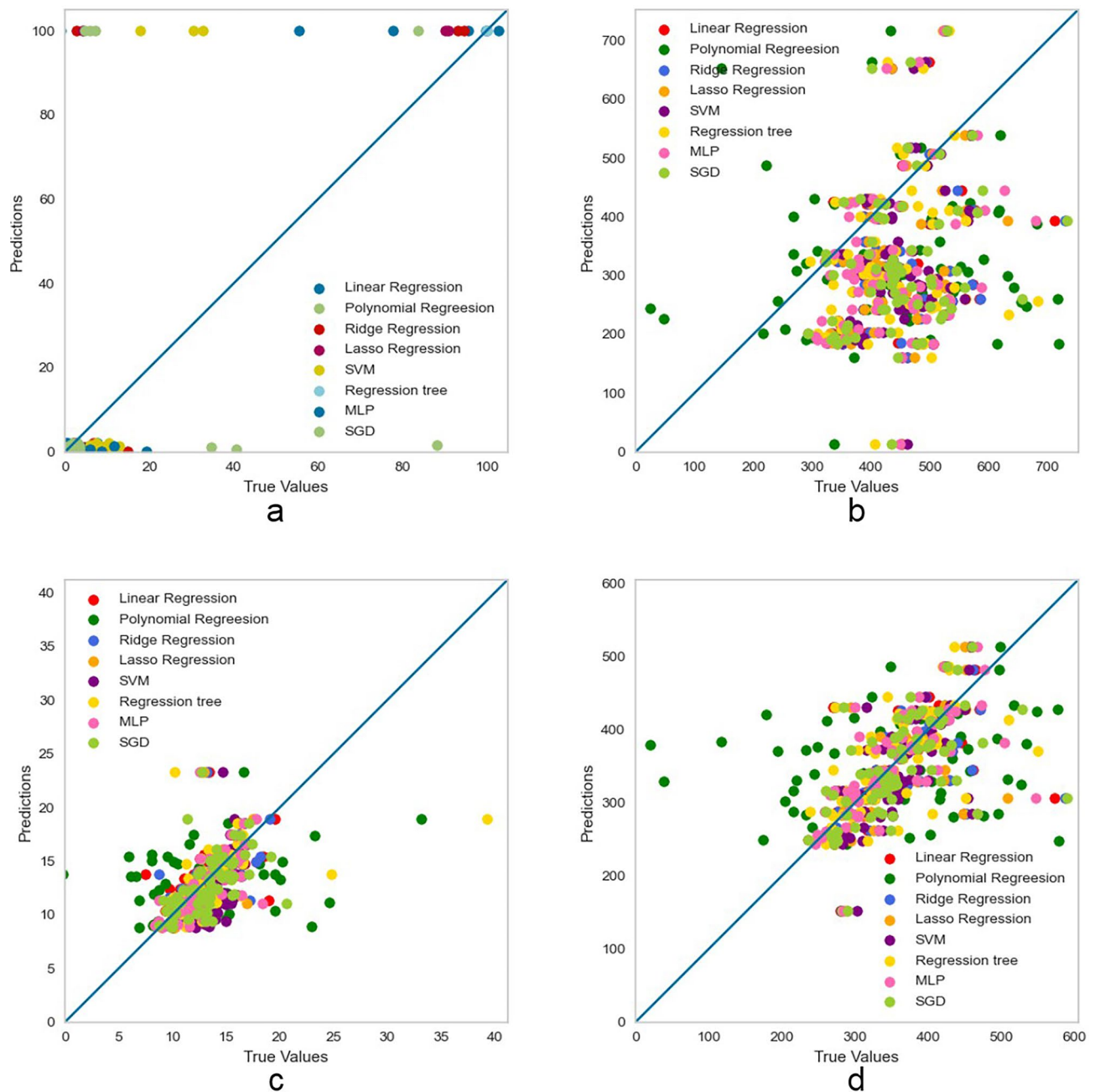


Figure 2. Visualization of prediction results. (a) Performance of each model when predicting BCVA, (b) performance of each model when predicting CST, (c) performance of each model when predicting CV, (d) performance of each model when predicting CAT.

Our results indicate that CAT and macular area thickness have a significant impact on BCVA prediction. DME can severely impair VA when retinal leakage accumulates in the OPL, which increases the overall macular thickness and may disturb the normal path of light from the inner retinal surface to the outer segments⁴⁶. Moreover, DME is also associated with vascular changes that have been linked to the degree of visual impairment^{47–50}. Meanwhile, CST prediction is more related to the overall macular thickness. In pathological conditions, altered organization and stability of junction proteins of RMG (e.g., ZO-1) can lead to cyst formation^{46,51}. RMG cell density is about five times higher in the fovea than in the periphery. Therefore, due to its thinness, the fovea is more susceptible to edema and has a greater influence on macular thickness⁵². The prediction of CAT and CV was more correlated with the thickness of the outer macular ring and ganglion cell layer because of their anatomical characteristics. The outer macula has a thick layer of ganglion cells while the central macular recess is the thinnest part of the entire macula⁴⁶. This may result in a higher likelihood of volume change in the outer macula affecting the thickness of different areas of the macula.

We also found that the type of anti-VEGF drugs was highly correlated with the predicted outcomes of BCVA, CST, CAT, and CV. Conbercept had a greater impact on these outcomes than ranibizumab and aflibercept.

Both aflibercept and conbercept are recombinant decoy receptor groups of VEGF⁵³. They block free VEGF-mediated signalling through their cognate receptors, thus inhibiting the pro-inflammatory, hyperpermeable, and pro-angiogenic effects of VEGF in a similar way. However, aflibercept showed greater therapeutic effects than conbercept, including more VA improvement and anatomical recovery. It also had advantages over mAb drugs (e.g., bevacizumab and ranibizumab) in enhancing VA and reducing macular edema⁵⁴. Nevertheless, aflibercept is the most expensive among anti-VEGF drugs and may have low patient acceptability due to the high cost. Conbercept is a novel multi-targeted anti-VEGF drug that has demonstrated excellent efficacy in treating patients with DME and neovascular AMD. The induction of domain 4 distinguishes conbercept from aflibercept and enhances its VEGF binding capacity. Due to the structural change, conbercept may affect steadily in the vitreous humour⁵⁵. Conbercept requires fewer injections, has a lower risk of injection-related complications, and is more cost-effective than ranibizumab^{54,56}. Among the included patients, most of them chose conbercept. This may have been driven by financial considerations and better outcomes. However, this also introduces some degree of bias. The better efficacy for DME and the bias may have contributed to the higher predictive relevance of conbercept compared to other anti-VEGF drugs.

Several studies^{57–60} used classification algorithms with less than six models. We used regression algorithms innovatively and established up to eight models. However, the evaluation indicators of classification and regression algorithms are different, so we cannot directly compare their accuracy. We hope that future research will explore more on the regression algorithm. Moreover, most of these studies used public datasets as their data sources, while we used a self-built database based on real-world data. Therefore, our model may be more consistent with the current state of the population and have epidemiological significance.

Some limitations should be drawn out. First, we had limited access to patient information, such as blood glucose, blood pressure, insulin use, smoking history, family history, etc.⁶¹, that was not available in our outpatient system. This may have led to less input for model training and reduced model accuracy. We imputed missing data using regression methods, but this may differ from the actual data and affect the prediction results. Second, polynomial regression and regression trees showed overfitting due to a large number of parameters and complex structure. Both polynomial regression and regression trees are prone to overfitting because they both increase the complexity and degrees of freedom of the model which may learn too many features that do not reflect the general patterns of the data.

The application of artificial intelligence (AI) in the medical field is currently not mature enough. Machine learning (ML), as a subset of AI, is commonly used for classification tasks, while regression tasks, which are commonly used for quantitative analysis, are rarely addressed. However, ML regression algorithms have potential in clinical settings and the public health field, such as developing treatment plans and predicting post-operative complications^{62,63}. For example, for diabetic macular edema (DME) patients, clinicians can select the most suitable anti-VEGF drug by considering ML-predicted indicators and drug prices. The improved model can be better applied to medical institutions in underdeveloped areas and primary hospitals, and enhance the diagnostic efficiency and accuracy of clinicians, as well as help reduce the economic burden on patients. With the rapid development of computer science, deep learning has emerged as an extension of ML that is commonly used for prediction and classification tasks using images in the medical field. Combining ML regression algorithms with deep learning can greatly expand the type and amount of data available to make more accurate clinical predictions, this will help improve the medical level in underdeveloped areas and greatly reduce the pressure on clinicians. In general, AI will inevitably change the existing mode of clinical practice in the future. Although AI may not replace physicians, the “AI + physician” mode of diagnosis and treatment may not be far away.

Conclusion

ML regression algorithms are effective in predicting the short-term efficacy of anti-VEGF treatment in DME patients, which are valuable in clinical and public health settings. Our results show that BCVA has the best prediction result compared to CST, CAT and CV. Furthermore, our analysis suggests that the lasso regression algorithm is the most effective ML technique for developing predictive regression models.

Data availability

The raw data is in Supplementary 2, and other data are available from the corresponding author on reasonable request.

Received: 20 May 2023; Accepted: 26 October 2023

Published online: 31 October 2023

References

1. Baena-Díez, J. M. *et al.* Risk of cause-specific death in individuals with diabetes: A competing risks analysis. *Diabetes Care* **39**, 1987–1995 (2016).
2. Cho, N. H. *et al.* IDF Diabetes Atlas: Global estimates of diabetes prevalence for 2017 and projections for 2045. *Diabetes Res. Clin. Pract.* **138**, 271–281. <https://doi.org/10.1016/j.diabres.2018.02.023> (2018).
3. Wong, T. Y., Gemmy, C. C. M., Larsen, M., Sharma, S. & Rafael, S. Erratum: Diabetic Retinopathy. *Nat. Rev. Dis. Primers* **2**, 16012 (2016).
4. Cheung, N., Mitchell, P. & Wong, T. Y. Diabetic retinopathy. *Lancet* **376**, 124–136. [https://doi.org/10.1016/s0140-6736\(09\)62124-3](https://doi.org/10.1016/s0140-6736(09)62124-3) (2010).
5. Antonetti, D. A., Klein, R. & Gardner, T. W. Mechanisms of disease diabetic retinopathy. *N. Engl. J. Med.* **366**, 1227–1239 (2012).
6. Moss, S. E., Klein, R. & Klein, B. E. The 14-year incidence of visual loss in a diabetic population. *Ophthalmology* **105**, 998–1003 (1998).
7. Group E. T. D. R. S. R. Treatment techniques and clinical guidelines for photocoagulation of diabetic macular edema: Early Treatment Diabetic Retinopathy Study Report Number 2. *Ophthalmology* **94**, 761–774 (1987).

8. Hwang, D.-K. *et al.* Smartphone-based diabetic macula edema screening with an offline artificial intelligence. *J. Chin. Med. Assoc.* **83**, 1102–1106 (2020).
9. Shahriari, M. H., Sabbaghi, H., Asadi, F., Hosseini, A. & Khorrami, Z. Artificial intelligence in screening, diagnosis, and classification of diabetic macular edema: A systematic review. *Surv. Ophthalmol.* (2022).
10. Teo, Z. L. *et al.* Global prevalence of diabetic retinopathy and projection of burden through 2045: Systematic review and meta-analysis. *Ophthalmology* **128**, 1580–1591. <https://doi.org/10.1016/j.ophtha.2021.04.027> (2021).
11. Miller, K. & Fortun, J. A. Diabetic macular edema: current understanding, pharmacologic treatment options, and developing therapies. *Asia-Pac. J. Ophthalmol.* **7**, 28–35 (2018).
12. Fogli, S. *et al.* Clinical pharmacology of intravitreal anti-VEGF drugs. *Eye (Lond)* **32**, 1010–1020. <https://doi.org/10.1038/s41433-018-0021-7> (2018).
13. Zhou, P., Zheng, S., Wang, E., Men, P. & Zhai, S. Conbercept for treatment of neovascular age-related macular degeneration and visual impairment due to diabetic macular edema or pathologic myopia choroidal neovascularization: A systematic review and meta-analysis. *Front. Pharmacol.* **12**, 696201. <https://doi.org/10.3389/fphar.2021.696201> (2021).
14. Chatziralli, I. & Loewenstein, A. Intravitreal anti-vascular endothelial growth factor agents for the treatment of diabetic retinopathy: A review of the literature. *Pharmaceutics*. <https://doi.org/10.3390/pharmaceutics13081137> (2021).
15. Uludag, G. *et al.* Efficacy and safety of intravitreal anti-VEGF therapy in diabetic retinopathy: What we have learned and what should we learn further?. *Expert Opin. Biol. Ther.* **22**, 1275–1291. <https://doi.org/10.1080/14712598.2022.2100694> (2022).
16. Bahrami, B., Hong, T., Gilles, M. C. & Chang, A. Anti-VEGF therapy for diabetic eye diseases. *Asia Pac. J. Ophthalmol. (Phila)* **6**, 535–545. <https://doi.org/10.22608/apo.2017350> (2017).
17. Cox, J. T., Elliott, D. & Sobrin, L. Inflammatory complications of intravitreal anti-VEGF injections. *J. Clin. Med.* <https://doi.org/10.3390/jcm10050981> (2021).
18. Ng, D. S. C., Ho, M., Iu, L. P. L. & Lai, T. Y. Y. Safety review of anti-VEGF therapy in patients with myopic choroidal neovascularization. *Expert Opin. Drug Saf.* **21**, 43–54. <https://doi.org/10.1080/14740338.2021.1952979> (2022).
19. Hutton, D. W. *et al.* Five-year cost-effectiveness of intravitreal ranibizumab therapy vs panretinal photocoagulation for treating proliferative diabetic retinopathy: A secondary analysis of a randomized clinical trial. *JAMA Ophthalmol.* **137**, 1424–1432. <https://doi.org/10.1001/jamaophthalmol.2019.4284> (2019).
20. Udaondo, P., Parravano, M., Vujosevic, S., Zur, D. & Chakravarthy, U. Update on current and future management for diabetic maculopathy. *Ophthalmol. Ther.* **11**, 489–502. <https://doi.org/10.1007/s40123-022-00460-8> (2022).
21. Pannu, A. Artificial intelligence and its application in different areas. *Artif. Intell.* **4**, 79–84 (2015).
22. Mintz, Y. & Brodie, R. Introduction to artificial intelligence in medicine. *Minim. Invasive Ther. Allied Technol.* **28**, 73–81. <https://doi.org/10.1080/13645706.2019.1575882> (2019).
23. Mitchell, T. M. & Mitchell, T. M. *Machine Learning* Vol. 1 (McGraw-Hill, 1997).
24. Shehab, M. *et al.* Machine learning in medical applications: A review of state-of-the-art methods. *Comput. Biol. Med.* **145**, 105458. <https://doi.org/10.1016/j.compbiomed.2022.105458> (2022).
25. Litjens, G. *et al.* A survey on deep learning in medical image analysis. *Med. Image Anal.* **42**, 60–88. <https://doi.org/10.1016/j.media.2017.07.005> (2017).
26. Fleuren, L. M. *et al.* Machine learning for the prediction of sepsis: A systematic review and meta-analysis of diagnostic test accuracy. *Intensive Care Med.* **46**, 383–400. <https://doi.org/10.1007/s00134-019-05872-y> (2020).
27. Flaxel, C. J. *et al.* Diabetic retinopathy preferred practice pattern. *Ophthalmology* **127**, P66–P145 (2020).
28. Group E. T. D. R. S. R. Grading diabetic retinopathy from stereoscopic color fundus photographs—An extension of the modified Airlie House classification: ETDRS report number 10. *Ophthalmology* **98**, 786–806 (1991).
29. Richman, J. *et al.* Importance of visual acuity and contrast sensitivity in patients with glaucoma. *Arch. Ophthalmol.* **128**, 1576–1582. <https://doi.org/10.1001/archophthalmol.2010.275> (2010).
30. Pedregosa, F. *et al.* Scikit-learn: Machine learning in Python. *J. Mach. Learn. Res.* **12**, 2825–2830 (2011).
31. Smith, K. On the standard deviations of adjusted and interpolated values of an observed polynomial function and its constants and the guidance they give towards a proper choice of the distribution of observations. *Biometrika* **12**, 1–85 (1918).
32. Hilt, D. E. & Seegrist, D. W. *Ridge, a Computer Program for Calculating Ridge Regression Estimates*. (Department of Agriculture, Forest Service, Northeastern Forest Experiment, 1977).
33. Kennedy, P. A *Guide to Econometrics* (Wiley, 2008).
34. Gruber, M. H. *Improving Efficiency by Shrinkage: The James-Stein and Ridge Regression Estimators* (Routledge, 2017).
35. Xian, G., Luo, X. & Xiao, Y. Statistics learning theory and support vector machine. *China Sci. Technol. Inf.* **12**, 178–181 (2005).
36. Huang, J.-C., Ko, K.-M., Shu, M.-H. & Hsu, B.-M. Application and comparison of several machine learning algorithms and their integration models in regression problems. *Neural Comput. Appl.* **32**, 5461–5469 (2020).
37. Strobl, C., Malley, J. & Tutz, G. An introduction to recursive partitioning: Rationale, application, and characteristics of classification and regression trees, bagging, and random forests. *Psychol. Methods* **14**, 323 (2009).
38. Breiman, L. *Classification and Regression Trees* (Routledge, 2017).
39. Rosenblatt, F. *Principles of Neurodynamics. Perceptrons and the Theory of Brain Mechanisms* (Cornell Aeronautical Lab Inc, 1961).
40. Rumelhart, D. E., Hinton, G. E. & Williams, R. J. *Learning Internal Representations by Error Propagation* (California University San Diego La Jolla Institute for Cognitive Science, 1985).
41. Cybenko, G. Approximation by superpositions of a sigmoidal function. *Math. Control Signals Syst.* **2**, 303–314 (1989).
42. Hodson, T. O. Root-mean-square error (RMSE) or mean absolute error (MAE): When to use them or not. *Geosci. Model Dev.* **15**, 5481–5487 (2022).
43. Renganathan, V. Overview of artificial neural network models in the biomedical domain. *Bratisl. Lek. Listy* **120**, 536–540. https://doi.org/10.4149/bll_2019_087 (2019).
44. Tibshirani, R. Regression shrinkage and selection via the lasso. *J. R. Stat. Soc. Ser. B (Methodological)* **58**, 267–288 (1996).
45. Ranstam, J. & Cook, J. A. LASSO regression. *Br. J. Surg.* **105**, 1348–1348. <https://doi.org/10.1002/bjs.10895> (2018).
46. Daruich, A. *et al.* Mechanisms of macular edema: Beyond the surface. *Prog. Retin. Eye Res.* **63**, 20–68. <https://doi.org/10.1016/j.preteyeres.2017.10.006> (2018).
47. Balaratnasingam, C. *et al.* Visual acuity is correlated with the area of the foveal avascular zone in diabetic retinopathy and retinal vein occlusion. *Ophthalmology* **123**, 2352–2367. <https://doi.org/10.1016/j.ophtha.2016.07.008> (2016).
48. Samara, W. A. *et al.* Quantification of diabetic macular ischemia using optical coherence tomography angiography and its relationship with visual acuity. *Ophthalmology* **124**, 235–244. <https://doi.org/10.1016/j.ophtha.2016.10.008> (2017).
49. Samara, W. A. *et al.* Quantitative optical coherence tomography angiography features and visual function in eyes with branch retinal vein occlusion. *Am. J. Ophthalmol.* **166**, 76–83. <https://doi.org/10.1016/j.ajo.2016.03.033> (2016).
50. Matet, A., Daruich, A., Dirani, A., Ambresin, A. & Behar-Cohen, F. Macular telangiectasia type I: Capillary density and microvascular abnormalities assessed by optical coherence tomography angiography. *Am. J. Ophthalmol.* **167**, 18–30. <https://doi.org/10.1016/j.ajo.2016.04.005> (2016).
51. Yanoff, M., Fine, B. S., Brucker, A. J. & Eagle, R. C. Pathology of human cystoid macular edema. *Surv. Ophthalmol.* **28**, 505–511. [https://doi.org/10.1016/0039-6257\(84\)90233-9](https://doi.org/10.1016/0039-6257(84)90233-9) (1984).
52. Distler, C. & Dreher, Z. Glia cells of the monkey retina—II. Müller cells. *Vis. Res.* **36**, 2381–2394. [https://doi.org/10.1016/0042-6989\(96\)00005-3](https://doi.org/10.1016/0042-6989(96)00005-3) (1996).

53. Wang, Q. *et al.* Novel VEGF decoy receptor fusion protein conbercept targeting multiple VEGF isoforms provide remarkable anti-angiogenesis effect in vivo. *PLoS One* **8**, e70544. <https://doi.org/10.1371/journal.pone.0070544> (2013).
54. Cai, S., Yang, Q., Li, X. & Zhang, Y. The efficacy and safety of aflibercept and conbercept in diabetic macular edema. *Drug Des. Dev. Ther.* **12**, 3471–3483. <https://doi.org/10.2147/dddt.S177192> (2018).
55. Lu, X. & Sun, X. Profile of conbercept in the treatment of neovascular age-related macular degeneration. *Drug Des. Dev. Ther.* **9**, 2311–2320. <https://doi.org/10.2147/dddt.S67536> (2015).
56. Xu, Y., Rong, A., Xu, W., Niu, Y. & Wang, Z. Comparison of 12-month therapeutic effect of conbercept and ranibizumab for diabetic macular edema: A real-life clinical practice study. *BMC Ophthalmol.* **17**, 158. <https://doi.org/10.1186/s12886-017-0554-8> (2017).
57. Abraham, A. *et al.* Machine learning prediction of kidney stone composition using electronic health record-derived features. *J. Endourol.* **36**, 243–250. <https://doi.org/10.1089/end.2021.0211> (2022).
58. Shao, L. *et al.* Prediction of the fundus tessellation severity with machine learning methods. *Front. Med. (Lausanne)* **9**, 817114. <https://doi.org/10.3389/fmed.2022.817114> (2022).
59. Banna, H. U. *et al.* Evaluation of machine learning algorithms for trabeculectomy outcome prediction in patients with glaucoma. *Sci. Rep.* **12**, 2473. <https://doi.org/10.1038/s41598-022-06438-7> (2022).
60. Heo, J. *et al.* Prediction of intracranial aneurysm risk using machine learning. *Sci. Rep.* **10**, 6921. <https://doi.org/10.1038/s41598-020-63906-8> (2020).
61. Bandello, F. *et al.* Diabetic macular edema. *Dev. Ophthalmol.* **58**, 102–138. <https://doi.org/10.1159/000455277> (2017).
62. Anahtar, M. N., Yang, J. H. & Kanjilal, S. Applications of machine learning to the problem of antimicrobial resistance: An emerging model for translational research. *J. Clin. Microbiol.* **59**, e0126020. <https://doi.org/10.1128/jcm.01260-20> (2021).
63. Kherabi, Y., Messika, J. & Peiffer-Smadja, N. Machine learning, antimicrobial stewardship, and solid organ transplantation: Is this the future?. *Transpl. Infect. Dis.* **24**, e13957. <https://doi.org/10.1111/tid.13957> (2022).

Acknowledgements

The authors would like to thank Mr. Ruobing Liu from Delft University of Technology (TU Delft) for his help in language editing.

Author contributions

Conceptualization, S.R. and L.X.; methodology, S.R., L.X., W.Y., Z.S. and C.X.; software, L.X., S.R.; validation, W.Y., Z.S. and C.X.; formal analysis, S.R., L.X.; investigation, S.R., L.X., W.Y., Z.S. and C.X.; data curation, S.R., L.X., and W.Y.; writing-original draft preparation, S.R., L.X., and W.Y.; writing, review and editing, S.R., L.X., W.Y.; supervision, L.X.; funding acquisition, L.X. All authors have read and agreed to the published version of the manuscript.

Funding

This work was supported in part by the National Natural Science Foundation of China (NSFC) under Grant 82174444.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1038/s41598-023-46021-2>.

Correspondence and requests for materials should be addressed to X.L.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2023