• Investigation •

Development and validation of a logistic regression model for predicting visual impairment in middle-aged and older adults with diabetes: results from the China Health and Retirement Longitudinal Study

Ri-Yan Zhang, Yu-Hao Liu, Duo-Duo Xie, Wei-Liang Fan, Zeng-Lin Bi, Qian Yao, Guan-Shun Yu, Zhong Lin, Jia Qu, Rong-Han Wu

National Clinical Research Center for Ocular Diseases, Eye Hospital, Wenzhou Medical University, Wenzhou 325027, Zhejiang Province, China

Co-first Authors: Ri-Yan Zhang and Yu-Hao Liu

Correspondence to: Jia Qu and Rong-Han Wu. National Clinical Research Center for Ocular Diseases, Eye Hospital, Wenzhou Medical University, Wenzhou 325027, Zhejiang Province, China. qujia@eye.ac.cn; wuronghan@wmu.edu.cn

Received: 2025-08-14 Accepted: 2025-12-18

Abstract

● **AIM:** To develop and validate a clinician-friendly logistic regression prediction model for self-reported visual impairment (VI) in middle-aged and older adults (≥45y) with diabetes.

● **METHODS:** Leveraging data from the China Health and Retirement Longitudinal Study (CHARLS), a model for VI among adults aged ≥45y with diabetes were developed. Feature selection involved LASSO regression and subsequent multivariable logistic regression. Eight machine learning algorithms were explored and compared for predictive performance. Logistic regression for its consistent performance, interpretability, and clinical usability was finally selected. A nomogram and interactive web-based tool were constructed to facilitate application.

● **RESULTS:** Totally 1918 participants (45.83% males) in CHARLS 2011 with aged ≥45y were analyzed in the training cohort and 1553 in CHARLS 2015 were in validation cohort. Among all participants in the training cohort, 39.6% reported VI. Seven variables were found to be independently associated with VI. The optimal model, logistic regression model, achieved area under the curve (AUC) of 0.702 and 0.706 in the training and validation cohorts, respectively. The model's potential for clinical application was supported by calibration and decision-curve analyses; the resulting nomogram and web calculator provided individualized risk prediction.

● **CONCLUSION:** We developed a clinically interpretable logistic regression model to predict the risk of VI in adults aged ≥45y with diabetes. The accompanying nomogram and web tool may assist with early identification and targeted vision care.

● **KEYWORDS:** visual impairment; diabetes; predictive model; elderly individuals

DOI:10.18240/ijo.2026.05.19

Citation: Zhang RY, Liu YH, Xie DD, Fan WL, Bi ZL, Yao Q, Yu GS, Lin Z, Qu J, Wu RH. Development and validation of a logistic regression model for predicting visual impairment in middle-aged and older adults with diabetes: results from the China Health and Retirement Longitudinal Study. *Int J Ophthalmol* 2026;19(5):974-986

INTRODUCTION

Diabetes is a chronic and increasingly prevalent disease that affects people around the world and poses significant challenges to healthcare systems[1-2]. Diabetes is associated with various microvascular complications that impact multiple systems, notably the cardiovascular, renal, neurological, and ocular systems[3-4]. Visual impairment (VI), which remains highly prevalent globally according to the Global Burden of Disease (GBD) 2021[5], is a clinically significant microvascular consequence involving various diabetes-related ocular problems, notably cataract, diabetic retinopathy, central retinal vein occlusion, and glaucoma[6-8]. Beyond reducing quality of life, diabetes-related VI can give rise to serious consequences such as increased likelihood of falls and cognitive decline, placing significant burdens on families and medical resources[9-11].

Previous studies have identified demographic, clinical, and behavioral factors that affect vision in populations with diabetes[12]. Advanced age, longer diabetes duration, and poor glycemic control have consistently been linked to retinal microvascular impairment and decreased vision[13]. The presence of complications such as hypertension and

dyslipidemia can aggravate eye damage by exacerbating both vascular and metabolic dysfunction[14-15]. Recent research also highlights the influence of psychological and functional factors, such as depression, sleep problems, and cognitive deficits, which increase the likelihood of vision loss in people with diabetes[16-18]. Despite these findings, most prior studies focus on individual variables, without integrating them into a unified predictive model.

In recent years, clinical risk prediction models have been increasingly applied across various medical specialties, including ophthalmology, to estimate disease risk based on personalized health data[19-20]. These models improve screening accuracy compared to traditional methods while reducing resource requirements and clinical workload[21]. For example, clinical prediction models for VI have been developed both for Chinese adults aged 45 and above and for patients with chronic kidney disease in the United States[22-23]. However, risk prediction tools for VI specifically targeting the population with diabetes are still lacking.

Therefore, we analyzed datasets from the China Health and Retirement Longitudinal Study (CHARLS, waves 1 and 3), a population-based national cohort, investigating risk factors linked to VI in individuals with diabetes aged 45y and above. Furthermore, we compared traditional logistic regression models with machine learning (ML) techniques to develop an accurate and interpretable prediction model. This tool can potentially promote early detection, timely intervention, and personalized strategies to prevent diabetes-related vision loss, thereby improving visual function and enhancing health-related quality of life.

PARTICIPANTS AND METHODS

Ethical Approval This study adhered to the ethical principles outlined in the Declaration of Helsinki and received approval from the Biomedical Ethics Committee of Peking University. Written informed consent was obtained from all participants, as approved by the university's Ethics Review Board (IRB00001052-11015).

Availability of Data and Materials The database was used from the China Health and Retirement Longitudinal Study. (http://charls.pku.edu.cn/).

Study Data and Population The analysis utilized CHARLS, a comprehensive national survey of adults aged 45y and above in China[24]. The baseline wave, conducted from June 2011 to March 2012, used a multi-level stratified, proportional cluster sampling design covering 28 provinces. Subsequent waves were administered at two to three-year intervals.

Two survey waves from CHARLS, conducted in 2011 and 2015, were used for this study, as these were the only CHARLS survey rounds that included measurements of blood-based biomarkers required for our analysis. The prediction model was developed based on a dataset from 2011, while the 2015 dataset served as a temporal validation. Both datasets cover detailed variables related to demographics, functional and health conditions, routine lifestyle patterns, and biological indicators, collected through structured questionnaires and physical examinations.

We enrolled adults ≥45y who satisfied established diagnostic criteria for diabetes. Participants with incomplete vision-related questionnaire data or insufficient information to confirm diabetic status were excluded from the analysis. After applying these criteria, we used 1918 eligible individuals from the 2011 CHARLS wave for model development, and 1513 participants were selected from the 2015 wave for temporal validation. To ensure the independence of the validation sample, we excluded those with duplicate IDs appearing in the 2011 dataset. Figure 1 outlines the participant selection procedure in detail.

Diagnosis of Diabetes In the CHARLS dataset, diabetes was identified through self-reports and laboratory measurements, without distinction between type 1 and type 2. Therefore, we applied the diagnostic guidelines established by the American Diabetes Association (ADA) in 2023[25]. Participants were considered with diabetes if they met one or more items from the list below: 1) fasting plasma glucose ≥126 mg/dL (7.0 mmol/L); 2) random plasma glucose ≥200 mg/dL (11.1 mmol/L); 3) glycated hemoglobin (HbA1c) ≥6.5%; 4) self-reported physician-diagnosed diabetes or hyperglycemia; 5) ongoing treatment of diabetes.

Definition of Visual Impairment The definition of VI was consistent with that used in previous CHARLS studies[26-27]. Visual function was evaluated through two self-reported items in the CHARLS questionnaire, assessing both distance and near vision. Self-reported vision status was assessed through the following items: 1) "How is your distance vision; For example, can you recognize someone across the street while wearing glasses"; 2) "How is your near vision; For instance, can you read a newspaper while wearing glasses". Responses were measured using a five-category scale, ranging from "excellent" to "poor", and those reporting "poor" on either item were classified as visually impaired. Participants who responded with the other four options to both questions were considered not visually impaired.

Candidate Predictors of Visual Impairment Informed by prior research and expert clinical judgment, this study incorporated a comprehensive set of potential predictors of VI[28-31]. 1) Sociodemographic variables: gender, age, family residence, marital status, education level, and public health insurance participation. 2) Lifestyle factors: average sleep duration, subjective sleep quality, smoking status, and alcohol consumption. 3) Health-related indicators: chronic pain,

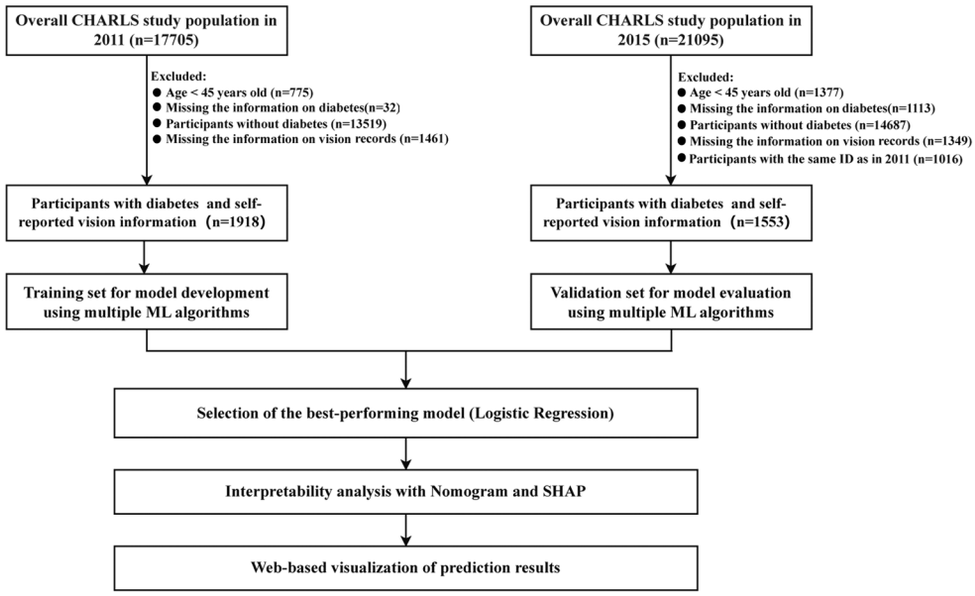

Figure 1 Flowchart of participant selection and study design CHARLS: China Health and Retirement Longitudinal Study; ML: Machine learning; SHAP: Shapley additive explanations.

hearing impairment, life satisfaction, self-rated expectation of future health, cognitive function score, activities of daily living (ADL), instrumental activities of daily living (IADL). 4) Medical comorbidities: self-reported diagnoses of hypertension, dyslipidemia, chronic lung disease, liver disease, heart disease, stroke, kidney disease, asthma, cancer, digestive disorders, memory-related conditions, arthritis, depression, and the total number of chronic diseases. 5) Objective clinical and laboratory measures: grip strength, waist circumference, body mass index (BMI), peak expiratory flow (PEF), white blood cell (WBC) and platelet (PLT) counts, fasting plasma glucose (GLU), HbA1c, total cholesterol (TC), triglycerides (TG), high-density lipoprotein cholesterol (HDL-C), low-density lipoprotein cholesterol (LDL-C), serum creatinine (SCr), uric acid (UA), blood urea nitrogen (BUN), C-reactive protein (CRP), and estimated glomerular filtration rate (eGFR).

Data Processing and Feature Selection Before model development, we conducted data preprocessing to ensure analytical quality. Variables with more than 20% missingness were excluded. For variables with partial missingness, imputation was performed utilizing a random forest-based approach implemented *via* the *mice* package in R. The 2011 CHARLS dataset served as the training cohort, and the 2015 wave was used for temporal validation. Participants with overlapping IDs between waves were excluded from the validation set to avoid data leakage.

To identify key predictors and minimize dimensionality, least absolute shrinkage and selection operator (LASSO) regression was conducted through the *glmnet* package in R. Ten-fold cross-validation was employed to identify the optimal regularization parameter, and λ_{1se} was for an optimal trade-off between model complexity and accuracy. Predictors retained by LASSO were subsequently entered into multivariable logistic regression analysis, and those with $P<0.05$ were regarded as statistically significant and included in the final model.

Model Development and Validation The CHARLS 2011 dataset was utilized as the training set for model development, while the 2015 wave served as a temporal validation cohort to assess model generalizability over time. An identical set of predictor variables was applied across both datasets to ensure consistency. Using the seven selected key predictors, we utilized eight ML algorithms to estimate the VI risk among individuals with diabetes. The algorithms included linear models [logistic regression (LR)], tree-based ensemble methods [gradient boosting machine (GBM), eXtreme gradient boosting (XGBoost), light gradient boosting machine (LightGBM), categorical boosting (CatBoost), and adaptive boosting (AdaBoost)], neural networks (NN), and support vector machines (SVM).

Model training and optimization were performed using grid search for hyperparameter tuning, in conjunction with 10-fold cross-validation to enhance generalizability and minimize overfitting. Model performance was assessed through multiple classification metrics, including accuracy, precision, and F1-score. To evaluate discriminatory ability, we assessed the discriminatory performance of each model by constructing receiver operating characteristic (ROC) curves. Predictive accuracy was quantified using the area under the curve (AUC). Discriminative performance was considered acceptable when the AUC reached or exceeded 0.70[32]. We further assessed model calibration by plotting predicted probabilities against actual outcomes. Calibration curves approaching the diagonal

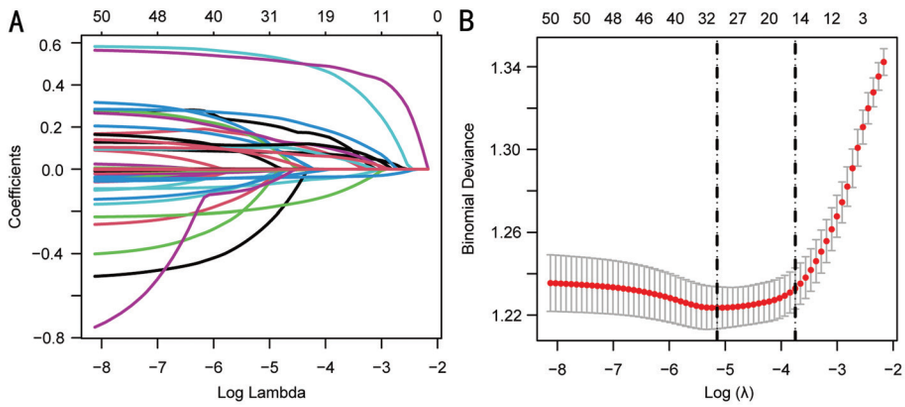

Figure 2 Predictor screening through the LASSO regression model A: According to the logarithmic (lambda) sequence, a coefficient profile was generated, and non-zero coefficients were produced by the optimal lambda; B: The optimal parameter (lambda) in the LASSO model was selected *via* 10-fold cross-validation using minimum criterion plus one standard error (right vertical line). LASSO: Least absolute shrinkage and selection operator.

reference line indicated better agreement. To evaluate practical applicability, we performed decision curve analysis (DCA) by quantifying the overall gain across varying probability thresholds.

Model Interpretation and Online Tool Deployment To improve the transparency and interpretability, we adopted the Shapley Additive exPlanations (SHAP) framework developed by Lundberg and Lee[33], which originates from cooperative game theory. This model-agnostic approach evaluates the individual contribution of each feature to the predicted outcome by quantifying its marginal effects through SHAP-based analysis. Features with higher SHAP values exert a more substantial positive impact on the predicted risk of VI, whereas lower values suggest a mitigating influence. Based on this method, we visualized and interpreted the contribution of the seven most influential predictors identified by the model.

In parallel, these seven independent risk factors were incorporated into a nomogram to represent individualized risk estimation visually. The nomogram facilitates intuitive, clinician-friendly interpretation of predicted probabilities based on patient characteristics.

In addition, we developed an interactive web-based prediction tool using R, enabling clinicians and public health practitioners to easily estimate individual risk profiles in real-time screening or consultation scenarios.

Statistical Analysis All data analyses and model development were conducted in R (version 4.4.3). Continuous variables were described using medians with interquartile ranges (IQRs), and categorical ones using frequencies (*n*) and proportions (%). Group comparisons were analyzed using statistical methods suitable for the data type, such as the Mann-Whitney *U* test for continuous variables and Chi-square or Fisher's exact tests for categorical variables. A two-sided $P<0.05$ was considered statistically significant.

RESULTS

Baseline Characteristics We analyzed 1918 adults with diabetes in CHARLS 2011 (training; 39.6% with self-reported VI) and 1553 in CHARLS 2015 (validation; 36.3% with VI). Compared with non-VI participants, those with VI were older and less educated, reported worse self-rated health and more chronic pain, and showed more adverse psychological/functional profiles (depressive symptoms, ADL/IADL limitations, lower cognition and grip strength), sensory impairment (hearing), higher prevalence of several comorbidities (*e.g.*, arthritis, stroke, heart disease, asthma, digestive disorders), and less favorable laboratory profiles (lipids, Hb/CRP, renal indices, UA). Full estimates are provided in Table 1.

Feature Selection LASSO regression was implemented on the 2011 training dataset (n=1918) to identify the most relevant predictors of VI in individuals with diabetes (Figure 2). The variables selected by the LASSO included gender, age, education level, sleep quality, depression, arthritis, number of chronic conditions, ADL, IADL, life satisfaction, hearing impairment, chronic pain, self-expectations of health status, and cognition score.

These variables were incorporated into a multivariable logistic regression model, identifying seven significant factors independently associated with VI (Table 2). Female, older age, poor sleep quality, presence of hearing impairment, experience of chronic pain, lower self-expectations of health status, and lower cognition score emerged as final independent risk factors for VI among participants with diabetes (P<0.05). The selected predictors were used to develop multiple risk prediction models using various ML algorithms.

Model Development and Performance Eight ML models were developed using the training dataset to predict VI among adults aged ≥45y with diabetes, including LR, SVM, GBM, XGBoost, LightGBM, CatBoost, AdaBoost, and NN.

Table 1 Baseline characteristics of study participants in 2011 (*n*=1918)

Variable	Overall (*n*=1918)	Non-visual impairment (*n*=1158)	Visual impairment (*n*=760)	*P*
Gender				<0.001
Male	879 (45.83%)	592 (51.12%)	287 (37.76%)	
Female	1039 (54.17%)	566 (48.88%)	473 (62.24%)	
Age, y				<0.001
45-55	531 (27.69%)	370 (31.95%)	161 (21.18%)	
55-65	803 (41.87%)	466 (40.24%)	337 (44.34%)	
≥65	584 (30.45%)	322 (27.81%)	262 (34.47%)	
Marital				0.024
Married	1676 (87.38%)	1028 (88.77%)	648 (85.26%)	
Single	242 (12.62%)	130 (11.23%)	112 (14.74%)	
Family residence				0.218
Rural	1004 (52.35%)	593 (51.21%)	411 (54.08%)	
Urban	914 (47.65%)	565 (48.79%)	349 (45.92%)	
Education level				<0.001
Less than elementary school	885 (46.14%)	457 (39.46%)	428 (56.32%)	
Elementary school	400 (20.86%)	255 (22.02%)	145 (19.08%)	
Middle school	391 (20.39%)	275 (23.75%)	116 (15.26%)	
High school or above	242 (12.62%)	171 (14.77%)	71 (9.34%)	
Sleeping time				<0.001
≤6h	995 (51.88%)	556 (48.01%)	439 (57.76%)	
6–8h	774 (40.35%)	522 (45.08%)	252 (33.16%)	
>8h	149 (7.77%)	80 (6.91%)	69 (9.08%)	
Sleep quality				<0.001
Very good	886 (46.19%)	608 (52.50%)	278 (36.58%)	
Good	306 (15.95%)	181 (15.63%)	125 (16.45%)	
Fair	314 (16.37%)	177 (15.28%)	137 (18.03%)	
Poor	412 (21.48%)	192 (16.58%)	220 (28.95%)	
Smoking status				0.003
No	1197 (62.41%)	692 (59.76%)	505 (66.45%)	
Yes	721 (37.59%)	466 (40.24%)	255 (33.55%)	
Drinking status				0.011
No	1387 (72.31%)	813 (70.21%)	574 (75.53%)	
Yes	531 (27.69%)	345 (29.79%)	186 (24.47%)	
Depression				<0.001
No	405 (21.12%)	310 (26.77%)	95 (12.50%)	
Yes	1513 (78.88%)	848 (73.23%)	665 (87.50%)	
Hypertension				0.245
No	1071 (55.84%)	659 (56.91%)	412 (54.21%)	
Yes	847 (44.16%)	499 (43.09%)	348 (45.79%)	
Dyslipidaemia				0.339
No	1463 (76.28%)	892 (77.03%)	571 (75.13%)	
Yes	455 (23.72%)	266 (22.97%)	189 (24.87%)	
Cancer				0.285
No	1891 (98.59%)	1139 (98.36%)	752 (98.95%)	
Yes	27 (1.41%)	19 (1.64%)	8 (1.05%)	
Lung disease				0.109
No	1715 (89.42%)	1046 (90.33%)	669 (88.03%)	
Yes	203 (10.58%)	112 (9.67%)	91 (11.97%)	
Psych-problem				0.296
No	1894 (98.75%)	1146 (98.96%)	748 (98.42%)	
Yes	24 (1.25%)	12 (1.04%)	12 (1.58%)	
Asthma				<0.001
No	1820 (94.89%)	1115 (96.29%)	705 (92.76%)	
Yes	98 (5.11%)	43 (3.71%)	55 (7.24%)	

Table 1 Baseline characteristics of study participants in 2011 (*n*=1918) (continued)

Variable	Overall (*n*=1918)	Non-visual impairment (*n*=1158)	Visual impairment (*n*=760)	*P*
Liver disease				0.171
No	1823 (95.05%)	1107 (95.60%)	716 (94.21%)	
Yes	95 (4.95%)	51 (4.40%)	44 (5.79%)	
Heart disease				0.010
No	1520 (79.25%)	940 (81.17%)	580 (76.32%)	
Yes	398 (20.75%)	218 (18.83%)	180 (23.68%)	
Stroke				0.049
No	1818 (94.79%)	1107 (95.60%)	711 (93.55%)	
Yes	100 (5.21%)	51 (4.40%)	49 (6.45%)	
Kidney disease				0.126
No	1778 (92.70%)	1082 (93.44%)	696 (91.58%)	
Yes	140 (7.30%)	76 (6.56%)	64 (8.42%)	
Stomach or other digestive disease				<0.001
No	1519 (79.20%)	952 (82.21%)	567 (74.61%)	
Yes	399 (20.80%)	206 (17.79%)	193 (25.39%)	
Memory-related disease				0.118
No	1875 (97.76%)	1137 (98.19%)	738 (97.11%)	
Yes	43 (2.24%)	21 (1.81%)	22 (2.89%)	
Arthritis or rheumatism				<0.001
No	1240 (64.65%)	815 (70.38%)	425 (55.92%)	
Yes	678 (35.35%)	343 (29.62%)	335 (44.08%)	
Number of chronic conditions				<0.001
0	292 (15.22%)	216 (18.65%)	76 (10.00%)	
1	428 (22.31%)	277 (23.92%)	151 (19.87%)	
≥2	1198 (62.46%)	665 (57.43%)	533 (70.13%)	
Activities of daily living				<0.001
Non-disability	1508 (78.62%)	975 (84.20%)	533 (70.13%)	
Disability	410 (21.38%)	183 (15.80%)	227 (29.87%)	
Instrumental activities of daily living				<0.001
Non-disability	1429 (74.50%)	933 (80.57%)	496 (65.26%)	
Disability	489 (25.50%)	225 (19.43%)	264 (34.74%)	
Life satisfaction				<0.001
Not at all satisfied	48 (2.50%)	16 (1.38%)	32 (4.21%)	
Not very satisfied	237 (12.36%)	109 (9.41%)	128 (16.84%)	
Some what satisfied	1170 (61.00%)	721 (62.26%)	449 (59.08%)	
Very satisfied	415 (21.64%)	280 (24.18%)	135 (17.76%)	
Completely satisfied	48 (2.50%)	32 (2.76%)	16 (2.11%)	
Health insurance				0.062
No	138 (7.19%)	73 (6.30%)	65 (8.55%)	
Yes	1780 (92.81%)	1085 (93.70%)	695 (91.45%)	
Hearing impairment				<0.001
No	1619 (84.41%)	1038 (89.64%)	581 (76.45%)	
Yes	299 (15.59%)	120 (10.36%)	179 (23.55%)	
Pain				<0.001
No	1246 (64.96%)	857 (74.01%)	389 (51.18%)	
Yes	672 (35.04%)	301 (25.99%)	371 (48.82%)	
Self-expectations of health status				<0.001
Almost impossible	172 (8.97%)	72 (6.22%)	100 (13.16%)	
Not very likely	466 (24.30%)	231 (19.95%)	235 (30.92%)	
Maybe	670 (34.93%)	432 (37.31%)	238 (31.32%)	
Very likely	274 (14.29%)	186 (16.06%)	88 (11.58%)	
Almost certain	336 (17.52%)	237 (20.47%)	99 (13.03%)	
BMI category				0.800
Normal weight	69 (3.60%)	43 (3.71%)	26 (3.42%)	

Table 1 Baseline characteristics of study participants in 2011 (*n*=1918) (continued)

Variable	Overall (*n*=1918)	Non-visual impairment (*n*=1158)	Visual impairment (*n*=760)	*P*
Obese	959 (50.00%)	588 (50.78%)	371 (48.82%)	
Overweight	718 (37.43%)	424 (36.61%)	294 (38.68%)	
Underweight	172 (8.97%)	103 (8.89%)	69 (9.08%)	
Grip, median (IQR)	30.90 (24.20-39.00)	32.00 (25.00-40.00)	29.00 (22.50-36.00)	<0.001
Cognition, median (IQR)	11.50 (8.50-13.50)	12.00 (9.50-14.00)	10.00 (7.50-13.00)	<0.001
Waist, median (IQR)	89.90 (82.10-97.00)	90.00 (82.93-97.00)	89.40 (81.10-97.00)	0.518
WBC, median (IQR)	6.10 (5.10-7.46)	6.04 (5.10-7.40)	6.10 (5.00-7.50)	0.720
PLT, median (IQR)	201.00 (159.00-251.00)	202.00 (158.25-251.00)	200.00 (160.00-251.00)	0.942
HbA1c, median (IQR)	5.70 (5.20-6.80)	5.70 (5.20-6.80)	5.80 (5.20-6.72)	0.567
HB, median (IQR)	14.40 (13.10-15.70)	14.45 (13.20-15.70)	14.30 (13.10-15.60)	0.020
GLU, median (IQR)	140.58 (127.08-177.07)	141.12 (127.44-177.26)	139.77 (126.00-176.67)	0.281
TC, median (IQR)	198.71 (172.04-226.16)	196.97 (171.75-226.16)	199.87 (172.04-225.97)	0.631
TG, median (IQR)	140.71 (95.58-225.68)	142.49 (98.24-234.52)	137.18 (89.39-212.62)	0.007
HDLC, median (IQR)	44.27 (35.57-53.74)	43.30 (35.18-53.35)	45.23 (36.63-54.51)	0.014
LDLC, median (IQR)	115.98 (89.30-141.40)	114.82 (88.24-139.47)	119.46 (90.08-144.20)	0.023
CRP, median (IQR)	1.35 (0.69-2.99)	1.43 (0.73-3.02)	1.28 (0.66-2.75)	0.023
UA, median (IQR)	4.45 (3.63-5.37)	4.53 (3.68-5.52)	4.35 (3.56-5.27)	0.002
BUN, median (IQR)	15.41 (12.94-18.51)	15.49 (12.94-18.51)	15.35 (12.79-18.50)	0.662
CR, median (IQR)	0.77 (0.67-0.90)	0.78 (0.67-0.92)	0.76 (0.66-0.89)	0.013
EGFR, median (IQR)	92.44 (79.56-100.94)	93.02 (79.89-101.69)	91.85 (79.05-99.68)	0.014
PEF, median (IQR)	75.25 (56.67-94.76)	78.00 (58.38-97.78)	71.29 (53.48-90.42)	<0.001

Median (interquartile range, IQR) was calculated for continuous variables, while frequencies and percentages were determined for categorical variables. The Wilcoxon rank-sum test was used to compare group differences for continuous variables, and Chi-squared tests were employed for categorical variables. BMI: Body mass index; WBC: White blood cell; PLT: Platelets; HbA1c: Glycated hemoglobin; HB: Glycated hemoglobin; GLU: Glucose; TC: Total cholesterol; TG: Triglycerides; HDLC: High density lipoprotein-cholesterol; LDLC: Low density lipoprotein cholesterol; CRP: C-reactive protein; UA: Uric acid; BUN: Blood urea nitrogen; CR: Creatinine; EGFR: Estimated glomerular filtration rate; PEF: Peak expiratory flow.

Regarding discrimination, GBM achieved the highest AUC in the training set (0.748; Figure 3A). However, in the validation set (Figure 3B), LR (AUC=0.706), XGBoost (AUC=0.705), and SVM (AUC=0.704) showed relatively more consistent performance across datasets. Calibration curves (Figure 3) indicated that LR, GBM, and LightGBM provided strong consistency between predicted outcomes and actual observations in training and validation cohorts.

Decision curve analysis (Figure 3) demonstrated that LR, CatBoost, and GBM offered favorable clinical utility over varying risk thresholds, particularly between 0.1 and 0.4. It suggests practical utility in risk-based decision-making.

Performance metrics summarized in Table 3 further highlight the capabilities and limitations inherent to each algorithm. LightGBM achieved the highest accuracy in the training set (0.703), while GBM showed higher specificity (0.737) and precision (0.611). In the validation set, LR delivered the most stable results, with overall accuracy (0.663), sensitivity (0.631), specificity (0.681), and F1-score (0.577), outperforming several complex models that exhibited signs of overfitting.

Taken together, LR demonstrated robust performance across all dimensions—discrimination, calibration, clinical utility, and generalizability—supporting its selection as the final model for deployment.

Model Explainability and Visualization SHAP was used to improve the interpretability of the final logistic regression model (Figure 4). The summary plot (Figure 4A) shows that pain, hearing impairment, and sleep quality were the most influential predictors, as indicated by their broad SHAP value distributions and substantial contributions to prediction.

The bar plot of mean SHAP values (Figure 4B) ranks the mean contribution of each variable to the prediction. Pain was the strongest contributor, followed by hearing impairment, sleep quality, and cognition score. Self-expectations of health status, gender, and age also had a moderate influence.

To illustrate feature-level contributions for individual predictions, SHAP force plots were used. Figure 4C presents a case with a high predicted risk, where poor sleep quality, low cognition score, and negative self-expectations of health status contributed substantially to the model output. In contrast, Figure 4D shows a case with low risk. Although the individual was older and reported pain, good cognitive function and positive self-expectations of health status helped mitigate the overall risk.

Construction of a Risk Nomogram Given the stable performance of the traditional LR model in preliminary

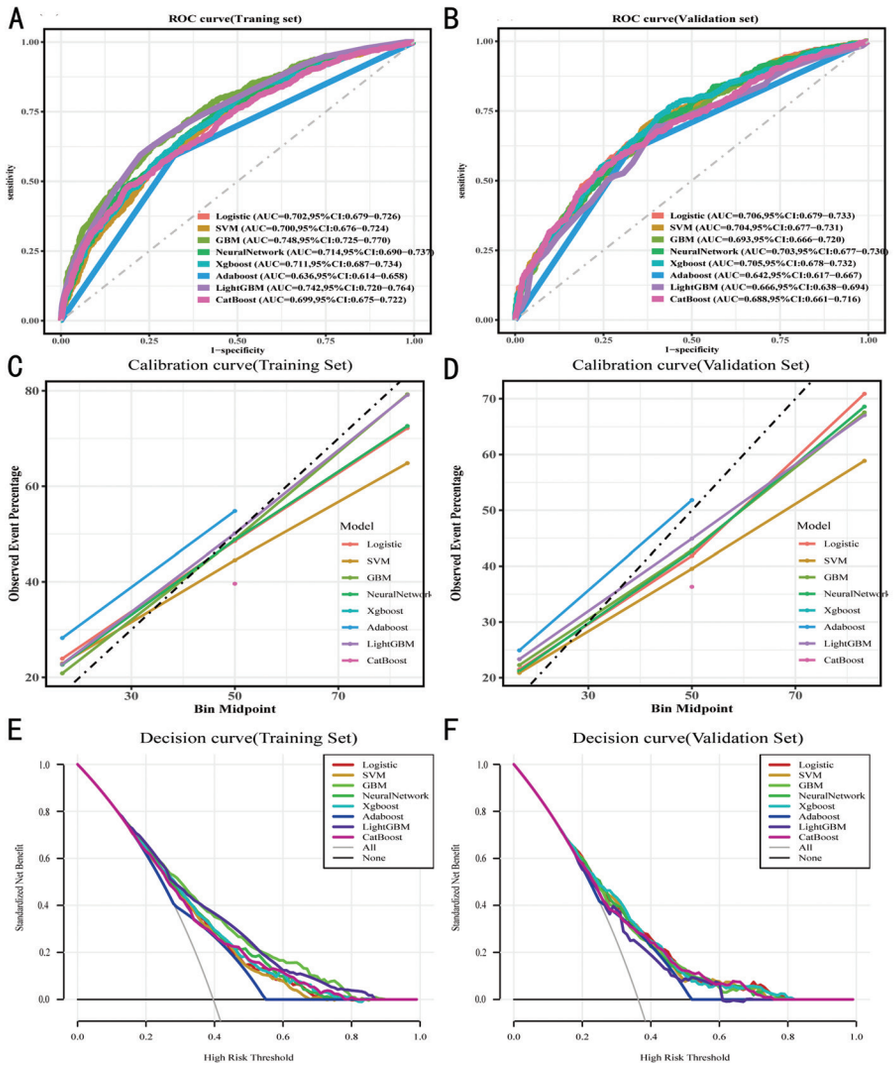

Figure 3 The performance of eight ML models for predicting visual impairment in diabetic individuals A: ROC curves for the training set; B: ROC curves for the validation set; C: Calibration curves for the training set; D: Calibration curves for the validation set; E: Decision curve analysis for the training set; F: Decision curve analysis for the validation set. ML: Machine learning; ROC: Receiver operating characteristic; AUC: Area under the curve; SVM: Support vector machines; GBM: Gradient boosting machine; XgBoost: eXtreme gradient boosting; AdaBoost: Adaptive boosting; LightGBM: Light gradient boosting machine; CatBoost: Categorical boosting.

analyses, we subsequently developed a nomogram based on the seven selected predictors (Figure 5). Each predictive factor contributes a specific point value, and the combined score reflects the estimated likelihood of VI, assisting in clinical decision-making and risk communication.

Deployment of an Online Prediction Tool To enhance the accessibility and practical utility of the final model, we deployed a web-accessible tool (https://liuyuhao123.shinyapps.io/make_web/), as illustrated in Figure 6. This tool integrates the seven independent risk factors—gender, age, sleep quality, reduced hearing ability, pain symptoms, anticipated health outlook, and cognition score—identified through model training. Clinicians can enter patient data to estimate the risk of VI quickly. This tool may assist in making timely decisions and encourage earlier referrals from diabetes care to eye specialists.

DISCUSSION

Diabetes is a major contributor to VI, primarily through its impact on the microvascular system[34]. Chronic hyperglycemia affects the microvasculature, including the retina, resulting in capillary leakage, blockage, reduced blood flow, and eventually causing retinal ischemia and hypoxia[35]. Such an environment of ischemia can lead to overproduction of proangiogenic mediators like vascular endothelial growth factor (VEGF), boosting abnormal neovascularization and fibrotic proliferation, thereby worsening the condition of diabetic retinopathy[36]. In addition, hyperglycemia may alter the metabolism of the lens, thus increasing the chances of developing cataracts[37]. Notably, many diabetes-related ocular complications, such as diabetic retinopathy and cataract, can be effectively managed in their early course to prevent further deterioration of visual function[38-39]. Therefore, it's imperative

Table 2 Multivariate logistic regression analysis of the risk factors of visual impairment in middle-aged and older adults with diabetes

Variables	Multivariate analysis OR (95%CI)	*P*
Gender		
Male	Reference	
Female	1.33 (1.08–1.64)	0.008
Age, y		
45-55	Reference	
55-65	1.47 (1.14–1.91)	0.003
≥65	1.44 (1.09–1.92)	0.011
Education level		
Less than elementary school	Reference	
Elementary school	0.84 (0.64–1.10)	0.213
Middle school	0.77 (0.56–1.04)	0.092
High school or above	0.95 (0.65–1.38)	0.79
Sleep quality		
Very good	Reference	
Good	1.22 (0.90–1.64)	0.193
Fair	1.06 (0.78–1.43)	0.726
Poor	1.37 (1.04–1.82)	0.027
Depression		
No	Reference	
Yes	1.34 (1.00–1.82)	0.054
Arthritis or rheumatism		
No	Reference	
Yes	1.16 (0.92–1.45)	0.211
Number of chronic conditions		
0	Reference	
1	1.17 (0.82–1.68)	0.383
≥2	1.33 (0.96–1.87)	0.092
Activities of daily living		
Non-disability	Reference	
Disability	1.15 (0.87–1.51)	0.329
Instrumental activities of daily living		
Non-disability	Reference	
Disability	1.09 (0.84–1.42)	0.503
Life satisfaction		
Not at all satisfied	Reference	
Not very satisfied	0.85 (0.42–1.67)	0.642
Somewhat satisfied	0.61 (0.31–1.15)	0.134
Very satisfied	0.51 (0.25–0.99)	0.051
Completely satisfied	0.48 (0.19–1.18)	0.115
Hearing impairment		
No	Reference	
Yes	1.81 (1.37–2.39)	<0.001
Pain		
No	Reference	
Yes	1.71 (1.37–2.14)	<0.001
Self-expectations of health status		
Almost impossible	Reference	
Not very likely	0.88 (0.60–1.28)	0.489
Maybe	0.64 (0.44–0.92)	0.017
Very likely	0.67 (0.44–1.04)	0.074
Almost certain	0.62 (0.40–0.94)	0.026
Cognition	0.96 (0.93–1.00)	0.036

OR: Odds ratio; CI: Confidence interval.

Table 3 Comparison of the performance of predictive models for visual impairment in diabetes middle-aged and older adults with diabetes

Model	Accuracy	Sensitivity	Specificity	Precision	F1
Training set					
LR	0.650	0.659	0.644	0.549	0.599
SVM	0.640	0.704	0.598	0.534	0.608
GBM	0.695	0.632	0.737	0.611	0.621
NN	0.656	0.655	0.656	0.556	0.601
XGB	0.656	0.650	0.661	0.557	0.600
ADA	0.645	0.592	0.680	0.548	0.569
LGBM	0.703	0.595	0.775	0.634	0.614
CB	0.680	0.480	0.812	0.626	0.544
Validation set					
LR	0.663	0.631	0.681	0.531	0.577
SVM	0.655	0.691	0.634	0.519	0.593
GBM	0.679	0.534	0.762	0.562	0.547
NN	0.634	0.722	0.584	0.498	0.589
XGB	0.636	0.764	0.562	0.499	0.604
ADA	0.652	0.605	0.679	0.518	0.558
LGBM	0.634	0.695	0.599	0.497	0.579
CB	0.668	0.576	0.720	0.540	0.557

Accuracy denotes the proportion of correctly predicted samples among the total samples. Sensitivity reflects the model's ability to identify actual positive cases correctly. Specificity shows the model's ability to identify actual negative cases correctly. Precision denotes the fraction of correctly identified positive cases among all instances predicted to be positive. F1 Score represents the harmonic mean of precision and recall, indicating a balanced measure of a model's ability to identify both positive and negative cases accurately. LR: Logistic regression; SVM: Support vector machine; GBM: Gradient boosting machine; NN: Neural network; XGB: eXtreme gradient boosting; ADA: Adaptive boosting; LGBM: Light gradient boosting machine; CB: Categorical boosting.

to identify high-risk groups of diabetic VI as early as possible.

In this research, a predictive model for VI in adults with diabetes aged 45y and above was developed and validated using the CHARLS data. The prevalence of self-reported VI in this population with diabetes was approximately 36%–40%. Seven key predictors were identified through LASSO regression and multivariable logistic regression analysis. We compared the predictive performance of multiple ML algorithms. Although some ensemble models showed superior discrimination in the training set, the traditional LR model demonstrated more consistent predictive performance across training and validation cohorts. Accordingly, we developed a nomogram and web-based tool based on the LR model. These tools aim to facilitate early recognition of high-risk individuals with diabetes related VI and assist in timely clinical decisions, helping to prevent long-term ocular complications and reduce the broader healthcare and societal burden.

We identified several significant predictors of VI in individuals with diabetes, including pain, hearing impairment, poor sleep

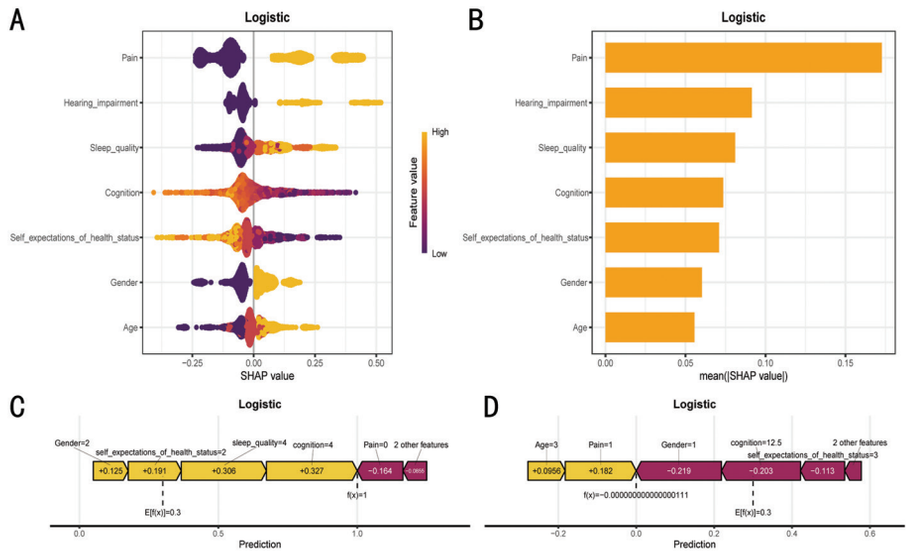

Figure 4 SHAP-based interpretation of the logistic regression model for VI prediction in individuals with diabetes A: Bar plot ranking the average SHAP values of features, indicating their relative importance; B: SHAP summary plot showing the impact and distribution of each feature on model output; C: Force plot of an example predicted as VI, showing how contributing variables increased the predicted probability; D: Force plot of an example predicted as non-VI, illustrating how each feature contributed to reducing the risk. SHAP: Shapley additive explanations; VI: Visual impairment.

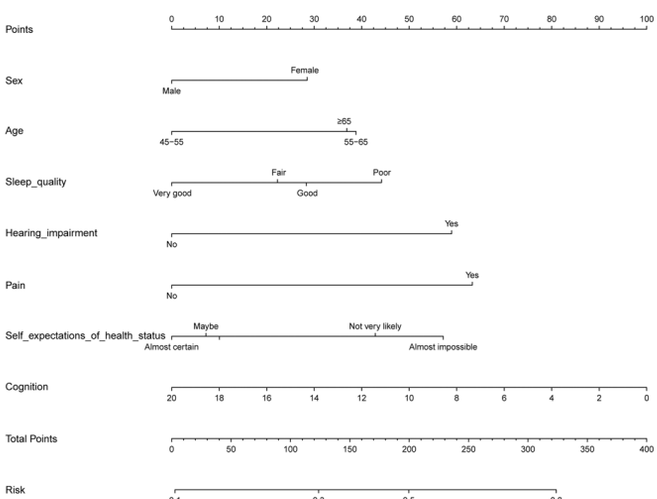

Figure 5 Nomogram for predicting VI risk in individuals with diabetes based on the logistic regression model VI: Visual impairment.

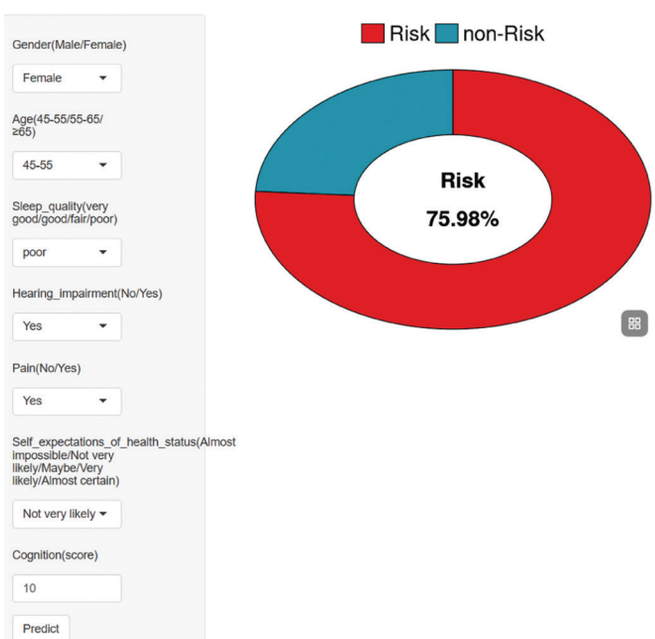

Figure 6 Web-based prediction tool for estimating the risk of VI in individuals with diabetes (PM-DVI) The interface allows users to input seven individual-level variables: gender, age, sleep quality, hearing impairment, pain, self-expectations of health status, and cognition score. Based on these inputs, the tool calculates the probability of VI and visualizes the result as a risk proportion. In this example, the estimated risk is 75.98%. VI: Visual impairment.

quality, lower cognition scores, negative self-expectations of health status, being female, and older age. Previous studies have extensively investigated these factors, revealing their stable contribution to vision decline. Studies have indicated a potential link between pain and VI[40]. Persistent chronic pain, as one of the symptoms of diabetes, often indicates underlying inflammation of the body and injury to the nerves. These changes could also facilitate the pathology of diabetic eye diseases[41-42]. Moreover, chronic pain can lead to sedentary lifestyles and decreased ability to care for oneself, which can delay eye care and routine examinations, thereby increasing the chances of loss of vision[43-44]. Additionally, hearing impairment in individuals with diabetes may contribute to visual decline due to overlapping microvascular damage and reduced sensory compensation. Chronic hyperglycemia can impair microcirculation in both the cochlea and retina[4]. When hearing is compromised, the visual system may bear a greater

perceptual load, accelerating functional deterioration[45]. Similarly, poor sleep quality is another significant factor linked to VI in individuals with diabetes. Sleep disturbances may accelerate systemic inflammation and metabolic dysregulation in the body, damaging the retina[46]. Emerging evidence suggests that poor sleep independently correlates with a higher risk of vision-threatening diabetic eye diseases[47-48].

Consequently, lower cognition scores are also related to an elevated risk of VI among individuals with diabetes[49]. Cognitive deterioration may impair one's capacity to recognize and cope with vision problems, leading to delays in seeking medical care and treatment[50]. Moreover, individuals with lower self-expectations of health tend to be pessimistic about the future, reducing their willingness to participate in preventive eye examinations or to respond to visual symptoms on time[51]. Notably, women have been found to exhibit increased susceptibility to VI[52]. Estimates from the GBD study indicate that women are more likely to be affected by blindness and visual problems than men, possibly related to multiple factors such as their longer life expectancy, differences in hormone levels, and lower participation in economic decision-making compared with men[53-54]. Finally, advancing age significantly increases VI risk[55]. As age grows, ocular degeneration such as lens opacification, macular degeneration, vitreous changes, and optic nerve degeneration frequently come to light[56-59].

Our investigation presents several strengths. To our understanding, it is the first to develop a nomogram and an interactive web-based tool based on the CHARLS dataset for estimating VI likelihood in populations with diabetes. Unlike prior models that depend mainly on imaging data[60], the model emphasizes simplicity and feasibility by incorporating seven readily available questionnaire-based predictors, enabling widespread application in primary care and public health contexts. This study has several limitations. First, VI was determined through self-reported data instead of an accurate clinical eye examination, which may lead to reporting bias. Second, we could not differentiate diabetes types due to the information limitations of the CHARLS questionnaire. However, the vast majority of Chinese adults with diabetes aged 45y and above are diagnosed with type 2, and thus our findings may apply primarily to this group. Third, because the outcome was based on self-reported visual impairment rather than clinical examination, the model cannot distinguish diabetes-related causes from other ocular conditions. As a result, causal inferences cannot be drawn. Fourth, because CHARLS represents a Chinese population, the model's performance in other regions or cultural settings remains uncertain. Validation in independent cohorts with different geographic or ethnic backgrounds will be an important next step. Lastly, some potentially important variables, including pre-existing ocular disease history, diabetes duration, genetic risk factors, and detailed ophthalmic findings, were not available in CHARLS and could not be included in the model.

In conclusion, this study applied multiple ML algorithms to develop and evaluate predictive models for VI among individuals aged ≥45y with diabetes. The results of multi-model comparison showed that LR performed best in prediction performance, stability, and clinical interpretability. Finally, seven key risk factors, including chronic pain, hearing impairment, poor sleep quality, lower cognitive performance, negative self-expectations of health status, being female, and advanced age, were identified. The constructed model can provide a simple and practical tool for clinicians, especially endocrinologists and primary diabetes management teams, to initially screen and evaluate the visual function risk of patients, and to refer high-risk individuals to ophthalmology in time for further evaluation.

ACKNOWLEDGEMENTS

Authors' Contributions: Zhang RY and Wu RH had full access to all of the data in the study and took responsibility for the integrity of the data and the accuracy of the data analysis. Zhang RY and Liu YH made contributed to concept and design; Xie DD, Fan WL, Bi ZL, Yao Q, and Yu GS contributed to acquisition, analysis, or interpretation of data; Lin Z, Qu J, and Yu GS contributed to project administration, supervision, and review. All authors read and approved the final manuscript.

Conflicts of Interest: Zhang RY, None; **Liu YH,** None; **Xie DD,** None; **Fan WL,** None; **Bi ZL,** None; **Yao Q,** None; **Yu GS,** None; **Lin Z,** None; **Qu J,** None; **Wu RH,** None.

REFERENCES

1 Lin X, Xu Y, Pan X, *et al*. Global, regional, and national burden and trend of diabetes in 195 countries and territories: an analysis from 1990 to 2025. *Sci Rep* 2020;10(1):14790.

2 Cole JB, Florez JC. Genetics of diabetes mellitus and diabetes complications. *Nat Rev Nephrol* 2020;16(7):377-390.

3 Yang Y, Zhao B, Wang Y, *et al*. Diabetic neuropathy: cutting-edge research and future directions. *Signal Transduct Target Ther* 2025;10(1):132.

4 Mauricio D, Gratacòs M, Franch-Nadal J. Diabetic microvascular disease in non-classical beds: the hidden impact beyond the retina, the kidney, and the peripheral nerves. *Cardiovasc Diabetol* 2023;22(1):314.

5 GBD 2019 Blindness and Vision Impairment Collaborators, Vision Loss Expert Group of the Global Burden of Disease Study. Causes of blindness and vision impairment in 2020 and trends over 30 years, and prevalence of avoidable blindness in relation to VISION 2020: the Right to Sight: an analysis for the Global Burden of Disease Study. *Lancet Glob Health* 2021;9(2):e144-e160.

6 Ramachandran N, Schmiedel O, Vaghefi E, *et al*. Evaluation of the prevalence of non-diabetic eye disease detected at first screen from

a single region diabetic retinopathy screening program: a cross-sectional cohort study in Auckland, New Zealand. *BMJ Open* 2021;11(12):e054225.

7 Zhang X, Sivaprasad S, Ting DSW. Editorial: Ocular complications associated with diabetes mellitus. *Front Endocrinol (Lausanne)* 2023;14:1193522.

8 Li Y, Mitchell W, Elze T, *et al*. Association between diabetes, diabetic retinopathy, and glaucoma. *Curr Diab Rep* 2021;21(10):38.

9 Thomas J, Almidani L, Ramulu P, *et al*. Falls and multiple falls among United States older adults with vision impairment. *Am J Ophthalmol* 2025;271:166-174.

10 Park YJ, Cho HJ, Kim K, *et al*. The association between stereoscopic vision and cognitive function on community-dwelling older adults: a cross-sectional study. *BMC Geriatr* 2025;25(1):359.

11 Seresirikachorn K, Somyanonthanakul R, Johnson M, *et al*. The impact of vision impairment on self-reported falls among older US adults: cross-sectional and longitudinal study. *JMIR Aging* 2025;8:e68771.

12 Zhang G, Chen W, Chen H, *et al*. Risk factors for diabetic retinopathy, diabetic macular edema, and sight-threatening diabetic retinopathy. *Asia Pac J Ophthalmol* (*Phila*) 2024;13(3):100067.

13 Sabanayagam C, Yip W, Ting DS, *et al*. Ten emerging trends in the epidemiology of diabetic retinopathy. *Ophthalmic Epidemiol* 2016;23(4):209-222.

14 Wei C, Huang Y, Xi P, *et al*. The non-linear association between remnant cholesterol/high-density lipoprotein cholesterol ratio and diabetic retinopathy: a cross-sectional study in type 2 diabetic patients. *Diabetol Metab Syndr* 2025;17(1):172.

15 Pieńczykowska K, Bryl A, Mrugacz M. Link between metabolic syndrome, inflammation, and eye diseases. *Int J Mol Sci* 2025;26(5):2174.

16 Rong H, Wang X, Lai X, *et al*. Associations between sleep duration and sensory impairments among older adults in China. *Front Aging Neurosci* 2022;14:910231.

17 Trott M, Driscoll R, Pardhan S. Associations between diabetic retinopathy, mortality, disease, and mental health: an umbrella review of observational meta-analyses. *BMC Endocr Disord* 2022;22(1):311.

18 Valluru G, Costa A, Klawe J, *et al*. Depression in individuals with diabetic retinopathy in the US national health and nutrition examination survey, 2005-2008. *Am J Ophthalmol* 2023;256:63-69.

19 Pang Y, Luo C, Zhang Q, *et al*. Multi-omics integration with machine learning identified early diabetic retinopathy, diabetic macula edema and anti-VEGF treatment response. *Transl Vis Sci Technol* 2024;13(12):23.

20 Ying B, Chandra RS, Wang J, *et al*. Machine learning models for predicting cycloplegic refractive error and myopia status based on non-cycloplegic data in Chinese students. *Transl Vis Sci Technol* 2024;13(8):16.

21 Yuan M, Wang W, Kang S, *et al*. Peripapillary microvasculature predicts the incidence and development of diabetic retinopathy: an SS-OCTA study. *Am J Ophthalmol* 2022;243:19-27.

22 Mao L, Yu Z, Lin L, *et al*. Determinants of visual impairment among Chinese middle-aged and older adults: risk prediction model using machine learning algorithms. *JMIR Aging* 2024;7:e59810.

23 Tan YH, Li JQ, Sun XF. Development and validation of a visual impairment prediction nomogram in chronic kidney diseases: the National Health and Nutrition Examination Survey, 1999-2008. *Int J Ophthalmol* 2025;18(5):919-924.

24 Zhao Y, Hu Y, Smith JP, *et al*. Cohort profile: the China health and retirement longitudinal study (CHARLS). *Int J Epidemiol* 2014;43(1):61-68.

25 ElSayed NA, Aleppo G, Aroda VR, *et al*. 1. Improving care and promoting health in populations: *Standards of care in diabetes—2023*. *Diabetes Care* 2023;46(Supplement_1):S10-S18.

26 He J, Tang X, Jiang M, *et al*. Associations between vision and hearing impairment and cardiovascular diseases: a longitudinal cohort of middle-aged and older adults in China. *J Am Heart Assoc* 2024;13(21):e034851.

27 Yan JM, Zhang MZ, Yu HJ, *et al*. Residential greenness, air pollution and visual impairment: a prospective cohort study. *BMC Public Health* 2024;24(1):3332.

28 Sun M, Bo Q, Lu B, *et al*. The association of sleep duration with vision impairment in middle-aged and elderly adults: evidence from the China health and retirement longitudinal study. *Front Med* 2021;8:778117.

29 Gu Y, Cheng H, Liu X, *et al*. Prevalence of self-reported chronic conditions and poor health among older adults with and without vision impairment in China: a nationally representative cross-sectional survey. *BMJ Open Ophthalmol* 2023;8(1):e001211.

30 Shang X, Wu G, Wang W, *et al*. Associations of vision impairment and eye diseases with frailty in community-dwelling older adults: a nationwide longitudinal study in China. *Br J Ophthalmol* 2024;108(2):310-316.

31 Shang X, Zhu Z, Wang W, *et al*. Associations of vision impairment and eye diseases with memory decline over 4 years in China and the United States. *Am J Ophthalmol* 2021;228:16-26.

32 Mandrekar JN. Receiver operating characteristic curve in diagnostic test assessment. *J Thorac Oncol* 2010;5(9):1315-1316.

33 Lundberg SM, Lee SI. A unified approach to interpreting model predictions. In: *Proceedings of the 31st International Conference on Neural Information Processing Systems (NIPS'17).* Curran Associates Inc., Red Hook, NY, USA:4768-4777.

34 Cheung N, Mitchell P, Wong TY. Diabetic retinopathy. *Lancet* 2010;376(9735):124-136.

35 Aiello LP. Angiogenic pathways in diabetic retinopathy. *N Engl J Med* 2005;353(8):839-841.

36 Aiello LP, Avery RL, Arrigg PG, *et al*. Vascular endothelial growth factor in ocular fluid of patients with diabetic retinopathy and other retinal disorders. *N Engl J Med* 1994;331(22):1480-1487.

37 Pollreisz A, Schmidt-Erfurth U. Diabetic cataract-pathogenesis, epidemiology and treatment. *J Ophthalmol* 2010;2010:608751.

38 Borrelli E, Grosso D, Barresi C, *et al*. Long-term visual outcomes and morphologic biomarkers of vision loss in eyes with diabetic macular edema treated with anti-VEGF therapy. *Am J Ophthalmol* 2022;235:80-89.

39 Tang J, Li T, Wang X, *et al*. Evaluation of the effect of full-process blood glucose management on T2DM cataract patients based on visual electrophysiology and OCTA. *Sci Rep* 2024;14:27862.

40 Zeng J, Liao Z, Lin A, *et al.* Association between poor pain control and sensory impairment: a cross-sectional study. *J Pain Res* 2025;18: 1313-1323.

41 Ding M, Ding A, Zhu L, *et al*. The impact of diabetes on chronic pain in different body regions among adults aged 50 and older: a cross-sectional analysis. *Front Public Health* 2025;13:1520735.

42 van Hecke O, Austin SK, Khan RA, *et al*. Neuropathic pain in the general population: a systematic review of epidemiological studies. *Pain* 2014;155(4):654-662.

43 Krein SL, Heisler M, Piette JD, *et al*. The effect of chronic pain on diabetes patients' self-management. *Diabetes Care* 2005;28(1):65-70.

44 Hida M, Imai R, Nakamura M, *et al*. Investigation of factors influencing low physical activity levels in community-dwelling older adults with chronic pain: a cross-sectional study. *Sci Rep* 2023;13(1):14062.

45 Veenman E, Roelofs AAJ, Stolwijk ML, *et al*. Experiences of people with dual sensory loss in various areas of life: a qualitative study. *PLoS One* 2023;18(9):e0272890.

46 Sang D, Lin K, Yang Y, *et al*. Prolonged sleep deprivation induces a cytokine-storm-like syndrome in mammals. *Cell* 2023;186(25): 5500-5516.e21.

47 Mersha GA, Alemu DS, Ferede MG, *et al*. Association of poor quality of sleep with vision-threatening diabetic retinopathy: a matched case-control study. *Ophthalmol Ther* 2023;12(2):1141-1153.

48 Dutta S, Ghosh S, Ghosh S. Association of sleep disturbance with diabetic retinopathy. *Eur J Ophthalmol* 2022;32(1):468-474.

49 Shah HS, DeSalvo MN, Haidar A, *et al*. Characterization of cognitive decline in long-duration type 1 diabetes by cognitive, neuroimaging, and pathological examinations. *JCI Insight* 2025;10(5):e180226.

50 Yang J, Zhang Z, Zhang L, *et al*. Relationship between self-care behavior and cognitive function in hospitalized adult patients with type 2 diabetes: a cross-sectional study. *Diabetes Metab Syndr Obes* 2020;13:207-214.

51 Nam S, Chesla C, Stotts NA, *et al*. Barriers to diabetes management: patient and provider factors. *Diabetes Res Clin Pract* 2011;93(1):1-9.

52 Zetterberg M. Age-related eye disease and gender. *Maturitas* 2016;83:19-26.

53 Li M, Wang Y, Liu Z, *et al*. Females with type 2 diabetes mellitus are prone to diabetic retinopathy: a twelve-province cross-sectional study in China. *J Diabetes Res* 2020;2020:5814296.

54 Xu Y, Wang A, Lin X, *et al*. Global burden and gender disparity of vision loss associated with diabetes retinopathy. *Acta Ophthalmol* 2021;99(4):431-440.

55 Yonekawa Y, Varma R, Choudhury F, *et al*. Risk factors for four-year incident visual impairment and blindness: the Los Angeles Latino Eye Study. *Ophthalmology* 2011;118(9):1790-1797.

56 Gehrs KM, Anderson DH, Johnson LV, *et al*. Age-related macular degeneration—emerging pathogenetic and therapeutic concepts. *Ann Med* 2006;38(7):450-471.

57 Coleman-Belin J, Harris A, Chen B, *et al*. Aging effects on optic nerve neurodegeneration. *Int J Mol Sci* 2023;24(3):2573.

58 Chen SP, Woreta F, Chang DF. Cataracts: a review. *JAMA* 2025;333(23):2093-2103.

59 Sebag J. Vitreous and vision degrading Myodesopsia. *Prog Retin Eye Res* 2020;79:100847.

60 Tham YC, Anees A, Zhang L, *et al*. Referral for disease-related visual impairment using retinal photograph-based deep learning: a proof-of-concept, model development study. *Lancet Digit Health* 2021;3(1):e29-e40.