Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Risk factors for light perception or no light perception after vitrectomy for proliferative diabetic retinopathy: a retrospective cohort study

Xu Liang

hbliangxu1985@126.com

Tianjin Eye Hospital, Tianjin Key Lab of Ophthalmology and Visual Science, Tianjin Eye Institute

Jiaxing Wang

Emory Eye Center

Yupeng Li

Tianjin Medical University Chu Hsien-I Memorial Hospital & Institute of Endocrinology

Hui Zheng

Tianjin Eye Hospital, Tianjin Key Lab of Ophthalmology and Visual Science, Tianjin Eye Institute

Yajie Wang

Tianjin Eye Hospital, Tianjin Key Lab of Ophthalmology and Visual Science, Tianjin Eye Institute

Yue Zhang

Tianjin Eye Hospital, Tianjin Key Lab of Ophthalmology and Visual Science, Tianjin Eye Institute

Article

Keywords: proliferative diabetic retinopathy, vitrectomy, visual loss, prediction model, iris neovascularisation, diabetic nephropathy

Posted Date: November 18th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-11450786/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

Abstract

Background

This study aimed to identify preoperative and perioperative risk factors for light perception (LP) or no light perception (NLP) after vitrectomy for proliferative diabetic retinopathy (PDR) and, as an exploratory step, to develop and internally validate preliminary prediction models.

Methods

This study included 238 eyes of 238 patients undergoing vitrectomy for PDR; 14 candidate predictors were evaluated using Firth penalised logistic regression, least absolute shrinkage and selection operator (LASSO) regression, and random forest. Performance was assessed by the area under the receiver operating characteristic curve (AUC), calibration slope, Brier score and decision curve analysis, with bootstrap optimism correction and repeated five-fold cross-validation.

Results

Nine eyes (3.8%) reached the endpoint of LP or NLP. Preoperative iris neovascularisation (odds ratio 39.88, 95% confidence interval 6.29-405.45), higher glycated haemoglobin (odds ratio 1.92 per %, 1.05–4.35) and diabetic nephropathy (odds ratio 15.50, 1.21-2397.50) were independently associated with LP or NLP, whereas preoperative intraocular pressure and visual acuity were not. Optimism-corrected AUCs were 0.971, 0.975 and 0.988; with only nine events, these estimates are unstable and best regarded as unreliable upper bounds, and the Firth model was the best calibrated.

Conclusions

Iris neovascularisation, poor glycaemic control and diabetic nephropathy were the strongest predictors of LP or NLP. The best-calibrated Firth model, translated into a nomogram, is the most suitable candidate for risk stratification, pending external validation.

1. Background

Proliferative diabetic retinopathy (PDR) is the leading cause of blindness among working-age adults with diabetes [1, 2]. The major sight-threatening complications of PDR are vitreous haemorrhage, tractional retinal detachment (TRD) and fibrovascular proliferation. Vitrectomy is the main treatment for these complications and is widely performed in clinical practice [3, 4].

Although microincision vitrectomy has substantially improved anatomical outcomes, visual prognosis after pars plana vitrectomy (PPV) for PDR remains highly variable. A small but clinically important

proportion of eyes progress to light perception (LP) or no light perception (NLP), the worst visual outcome, which constitutes legal blindness in most countries [5, 6]. Identifying patients at high risk of LP or NLP before surgery or in the early postoperative period is therefore of considerable clinical value: it informs shared decision-making, refines patient counselling, guides the intensity of perioperative metabolic and anti-vascular endothelial growth factor (VEGF) management, and may help allocate follow-up resources.

Previous studies have linked poor visual outcomes after vitrectomy in patients with diabetes to neovascular glaucoma (NVG), preoperative visual acuity, TRD and systemic glycaemic control [7, 8]. However, most of these studies predate the microincision era and relied on univariable analyses rather than formal prediction models. In addition, potentially informative routine variables, including preoperative intraocular pressure (IOP), preoperative visual acuity and renal replacement therapy as a marker of end-stage systemic microangiopathy, have seldom been examined together within a unified modelling framework.

The aim of this study was therefore to identify risk factors for LP or NLP after PPV in patients with PDR using routinely collected preoperative and perioperative variables and, as an exploratory, hypothesis-generating step, to develop and internally validate preliminary prediction models, compare complementary modelling strategies, and translate the best-performing model into a nomogram as a candidate tool for future validation. The study is reported in accordance with the TRIPOD + AI statement, which covers prediction models developed using regression and machine-learning methods [9, 29]; the complete model-development and internal-validation pipeline, including predictor prespecification, penalisation, resampling and optimism correction, is described in Section 2.3. What this study adds is the joint quantification of ocular (NVI) and systemic (glycaemic control, diabetic nephropathy) markers for the most severe post-vitrectomy outcome, LP or NLP, in the microincision era, identifying the readily recognisable combination of NVI, poor glycaemic control and nephropathy as a profile that may warrant intensified postoperative surveillance.

2. Materials and methods

2.1. Study design and participants

This study reviewed the records of consecutive patients who underwent vitrectomy for PDR at Tianjin Eye Hospital between January 2018 and December 2020, with five years of postoperative follow-up. The prediction horizon was defined as the five years after vitrectomy. LP or NLP was classified as irreversible when best-corrected visual acuity of light perception or no light perception was documented at two consecutive examinations at least three months apart, after completion of all indicated treatment (including anti-VEGF injection, panretinal photocoagulation and management of neovascular glaucoma or other complications), with no recovery of visual acuity above light perception during that interval, and when the treating vitreoretinal surgeon judged that no further available intervention could reasonably be expected to restore form vision. Eyes whose visual acuity declined transiently to light perception or no

light perception in the early postoperative period but subsequently recovered above light perception were not counted as endpoint events. Follow-up for an eye ended at the visit at which irreversible LP or NLP was confirmed. The database was locked on 1 January 2026. The study followed the tenets of the Declaration of Helsinki and was approved by the institutional review board of Tianjin Eye Hospital, and all patients provided written informed consent.

Inclusion criteria were (i) age 18 years or older and (ii) vitrectomy for complications of PDR, including non-clearing vitreous haemorrhage persisting for more than 3 months, fibrovascular membranes or tractional retinal detachment. Exclusion criteria were (i) a history of glaucoma, (ii) ocular trauma, (iii) uveitis, (iv) intraocular injection other than preoperative anti-VEGF administered as part of the current surgical preparation, (v) other ocular diseases except mild refractive errors, and (vi) loss to follow-up before completion of the five-year visit.

Each patient underwent a complete preoperative ophthalmic examination. Best-corrected visual acuity (BCVA) was measured at a standardised distance (4 m) under standardised lighting, and Snellen fractions and low-vision grades were converted to logMAR values. Intraocular pressure was measured by applanation tonometry; slit-lamp biomicroscopy, binocular indirect ophthalmoscopy and B-scan ultrasonography were performed, and fundus fluorescein angiography was undertaken when clinically indicated. All PPVs were performed by one of two experienced surgeons (Dr Xu Liang or Dr Yue Zhang) using a 25-gauge three-port system. All eyes received an intravitreal anti-VEGF injection 5–7 days before vitrectomy as part of routine preoperative preparation. Endolaser panretinal photocoagulation was completed up to the ora serrata, and air, gas or silicone-oil tamponade was selected according to disease severity. HbA1c was reported in NGSP units (%); values could be converted to IFCC units (mmol/mol) using IFCC = 10.93 × NGSP (%) − 23.5 (for example, 9.9% ≈ 85 mmol/mol).

A formal sample size assessment was performed using the criteria of Riley et al. [10]. With nine events among 238 eyes (outcome proportion 3.8%), the events-per-variable ratio was 2.25 for the four-parameter Firth model, below the conventional threshold of 10, which would have required approximately 40 events (about 1,060 eyes at this outcome proportion) [11]. Under the criterion targeting a global shrinkage factor of 0.90, the maximum Cox-Snell R-squared attainable at this outcome proportion was 0.275; assuming an anticipated Cox-Snell R-squared of 0.15, a four-parameter model required at least 219 participants (238 available), whereas under a more conservative assumption of 0.10 the required sample size (340) exceeded the cohort size. The criterion for precise estimation of overall outcome risk was comfortably met (minimum n = 56). These considerations motivated penalised estimation, restriction of the final model to four predictors, bootstrap internal validation and framing of the models as hypothesis-generating. The effective complexity of the final Firth model was deliberately limited to four fitted parameters (2.25 events per parameter). As an additional assessment of model stability, the dispersion of out-of-fold performance across repeated cross-validation folds and the bootstrap percentile interval of the calibration slope were examined; as reported in Section 3.5, both were wide, and we present them transparently as a quantification of residual instability rather than as evidence of robustness.

2.2. Outcome and candidate predictors

Irreversible LP or NLP was determined according to the criteria defined in Section 2.1, and the endpoint date was the first examination at which these criteria were fulfilled. All eyes that did not reach the endpoint completed the full five years of follow-up, so outcome ascertainment was complete for the entire cohort and no eye was censored early; accordingly, the endpoint was analysed as a binary indicator of irreversible LP or NLP occurring at any time within the five-year horizon. The composite endpoint comprised eyes reaching light perception (LP) and eyes reaching no light perception (NLP), and the two components are reported separately in Table 1. Fourteen candidate predictors were selected a priori on the basis of clinical relevance and availability: age, sex, duration of diabetes, duration of hypertension, glycated haemoglobin (HbA1c), insulin therapy, diabetic nephropathy, maintenance haemodialysis, concomitant retinal macrovascular occlusion (RMO), preoperative iris neovascularisation (NVI), PDR severity (vitreous haemorrhage, fibrovascular proliferation or retinal detachment), intraocular tamponade agent (silicone oil, gas or air/water), preoperative BCVA (logMAR) and preoperative IOP (mmHg). RMO was defined as occlusion of a retinal macrovessel and/or the presence of retinal vascular white lines (white, sheathed, occluded retinal vessels), documented on preoperative fundus examination or intraoperatively [12]. There were no missing data for the outcome or any candidate predictor.

Table 1
Baseline characteristics of the study population according to postoperative visual outcome

Characteristic	Overall (n = 238)	LP or NLP (n = 9)	No LP or NLP (n = 229)	P value
Endpoint component, n	9	LP, 7; NLP, 2	0	
Age, years	57.0 (50.0–62.0)	60.0 (59.0–61.0)	57.0 (50.0–62.0)	0.090
Male sex, n (%)	103 (43.3)	5 (55.6)	98 (42.8)	0.506
Duration of diabetes, years	15.0 (10.0–20.0)	20.0 (16.9–23.0)	15.0 (10.0–20.0)	0.012
Duration of hypertension, years	1.0 (1.0–5.0)	10.0 (1.0–10.0)	1.0 (1.0–5.0)	0.026
HbA1c, %	7.4 ± 1.7	9.9 ± 1.5	7.3 ± 1.6	< 0.001
Insulin therapy, n (%)	157 (66.0)	3 (33.3)	154 (67.2)	0.066
Diabetic nephropathy, n (%)	95 (39.9)	9 (100.0)	86 (37.6)	< 0.001
Maintenance haemodialysis, n (%)	14 (5.9)	5 (55.6)	9 (3.9)	< 0.001
Retinal macrovascular occlusion, n (%)	166 (69.7)	6 (66.7)	160 (69.9)	1.000
Iris neovascularisation (NVI), n (%)	10 (4.2)	7 (77.8)	3 (1.3)	< 0.001
PDR severity, n (%)				< 0.001
Vitreous haemorrhage	157 (66.0)	0 (0.0)	157 (68.6)	
Fibrovascular proliferation	17 (7.1)	2 (22.2)	15 (6.6)	
Retinal detachment	64 (26.9)	7 (77.8)	57 (24.9)	
Tamponade agent, n (%)				0.048
Silicone oil	64 (26.9)	5 (55.6)	59 (25.8)	
Gas	35 (14.7)	2 (22.2)	33 (14.4)	
Air/water	139 (58.4)	2 (22.2)	137 (59.8)	
Preoperative BCVA, logMAR	2.0 (1.7-3.0)	2.0 (1.7-3.0)	2.0 (1.7-3.0)	0.477

Data are median (interquartile range), mean ± standard deviation, or n (%). LP or NLP = light perception or no light perception; HbA1c = glycated haemoglobin; PDR = proliferative diabetic retinopathy; BCVA = best-corrected visual acuity; IOP = intraocular pressure. The composite endpoint comprised eyes reaching LP and eyes reaching NLP; separate counts are shown above.

Characteristic	Overall (n = 238)	LP or NLP (n = 9)	No LP or NLP (n = 229)	P value
Preoperative IOP, mmHg	14.0 (12.0–16.0)	15.0 (13.0-16.7)	14.0 (12.0–16.0)	0.305

Data are median (interquartile range), mean ± standard deviation, or n (%). LP or NLP = light perception or no light perception; HbA1c = glycated haemoglobin; PDR = proliferative diabetic retinopathy; BCVA = best-corrected visual acuity; IOP = intraocular pressure. The composite endpoint comprised eyes reaching LP and eyes reaching NLP; separate counts are shown above.

2.3. Statistical analysis

2.3.1 Statistical analysis and Model development

Continuous variables were summarised as mean with standard deviation or median with interquartile range and compared between outcome groups using the t test or the Mann-Whitney U test, according to normality assessed by the Shapiro-Wilk test; categorical variables were compared using the chi-squared or Fisher's exact test. Univariable associations were quantified by Firth odds ratios (ORs) with 95% confidence intervals (CIs).

Three complementary modelling strategies were used. First, because only nine events occurred and diabetic nephropathy showed complete univariable separation (all eyes with LP or NLP had nephropathy), conventional maximum-likelihood logistic regression was not appropriate. Multivariable Firth penalised logistic regression [13, 14] was therefore fitted; this approach provides finite, bias-reduced estimates in the presence of separation and rare events. Predictors for the final Firth model were prespecified on the basis of prior clinical evidence and biological plausibility rather than data-driven selection: preoperative NVI, diabetic nephropathy and HbA1c have repeatedly been linked to severe visual loss or neovascular complications after vitrectomy for PDR [7, 8, 20–24, 26, 27], and preoperative IOP was retained a priori because higher preoperative IOP has consistently predicted post-vitrectomy neovascular glaucoma [23, 24]. No univariable P-value screening was used to select the variables in the final model; given the low events-per-variable ratio [11], the model was deliberately restricted to these four predictors. The contribution of each predictor was assessed with profile penalised likelihood tests and intervals. In addition, the LASSO and random forest analyses described below used all 14 candidate predictors without any univariable screening, and the concordance of variable rankings across these alternative strategies (Section 3.4) provided an informal check that the findings are not driven by any single selection strategy.

Second, least absolute shrinkage and selection operator (LASSO) logistic regression was fitted with the penalisation parameter selected by five-fold cross-validation repeated 10 times, optimising the AUC, after standardisation of continuous predictors and with class weights inversely proportional to class frequencies [15]. Third, a random forest classifier (500 trees, maximum depth 4, minimum of five observations per leaf, balanced class weights) was trained as a non-parametric benchmark [16]; variable

importance was summarised by the mean decrease in impurity. The LASSO and random forest models used all 14 candidate predictors.

2.3.2 Internal validation

A systematic search of public repositories (Kaggle, Figshare, Zenodo), publicly funded imaging datasets (EyePACS, MESSIDOR, IDRiD) and the Royal College of Ophthalmologists National Ophthalmology Database identified no publicly accessible dataset with individual-level data on vitrectomy for PDR together with the required predictors and an NLP outcome. External validation was therefore not feasible, and enhanced internal validation was performed instead.

Internal validation followed two complementary approaches. First, Harrell bootstrap optimism correction was applied with 200 bootstrap resamples [17]; in each resample the modelling pipeline was refitted, including re-selection of the LASSO penalisation parameter, and optimism-corrected performance was calculated as apparent performance minus mean optimism. Second, repeated five-fold cross-validation (10 repeats) generated out-of-fold predictions. Discrimination was quantified by the AUC (95% CIs from 2000 bootstrap resamples for apparent estimates; mean and standard deviation across repeats for cross-validation). Calibration was examined using quintile-based calibration plots, calibration curves derived from pooled bootstrap-model predictions on the original data, and the calibration slope, defined for each bootstrap resample as the slope from a logistic regression of the observed outcome on the logit of the resample-refitted model's predictions in the original dataset, summarised as the median and 2.5th-97.5th percentiles across the 200 resamples. The Brier score summarised overall performance, and clinical utility was assessed by decision curve analysis across threshold probabilities of 1–50% using bootstrap-corrected predictions [18]. A nomogram translating the Firth model into a point-based scoring system was constructed [19]. Analyses were performed in Python 3 (scikit-learn 1.9, statsmodels 0.14) and R 4.6.1 (https://www.r-project.org/). As an incremental-value analysis, a single-predictor Firth model including only NVI was fitted and compared with the four-predictor model in terms of optimism-corrected AUC and by a penalised likelihood-ratio test. Two-sided $P < 0.05$ was considered significant.

3. Results

3.1. Baseline characteristics

After exclusion of participants who failed to complete follow-up or did not satisfy the inclusion criteria, a total of 238 eyes of 238 patients (103 men, 43.3%) with a median age of 57.0 years (interquartile range 50.0–62.0) were analysed. Nine eyes (3.8%) reached the predefined endpoint of LP or NLP within the five-year horizon (7 eyes with LP and 2 eyes with NLP). Baseline characteristics by outcome are shown in Table 1. Eyes that developed LP or NLP had higher HbA1c (mean 9.9% vs 7.3%, $P < 0.001$), more frequently had diabetic nephropathy (100.0% vs 37.6%, $P < 0.001$), were more often receiving maintenance haemodialysis (55.6% vs 3.9%, $P < 0.001$), more often had preoperative NVI (77.8% vs 1.3%, $P < 0.001$), presented with more advanced PDR (retinal detachment in 77.8% vs 24.9%, $P < 0.001$), and had a longer duration of diabetes (median 20.0 vs 15.0 years, $P = 0.012$) and of hypertension (median

10.0 vs 1.0 years, P = 0.026). Silicone-oil tamponade was more frequent among eyes reaching LP or NLP at borderline significance (P = 0.048), whereas age, sex, preoperative BCVA, preoperative IOP, insulin therapy and RMO did not differ significantly between groups.

3.2. Univariable analysis

In univariable Firth logistic regression (Table 2), preoperative NVI was strongly associated with LP or NLP (OR 194.14, 95% CI 37.56-1432.91; P < 0.001), as were diabetic nephropathy (OR 31.52, 3.92-4082.19; P < 0.001) and maintenance haemodialysis (OR 28.37, 6.99-123.24; P < 0.001). More advanced PDR (OR 4.35 per grade, 1.93–14.74; P < 0.001) and higher HbA1c (OR 2.85 per %, 1.76–5.10; P < 0.001) were also associated with LP or NLP, together with longer duration of diabetes (OR 1.03 per year, 1.01–1.22; P = 0.002) and air or gas rather than silicone-oil tamponade (OR 0.45 per grade, 0.20–0.92; P = 0.028); insulin therapy showed an inverse association (OR 0.26, 0.06–0.95; P = 0.042). Age, duration of hypertension, preoperative IOP, preoperative BCVA, sex and RMO were not significantly associated (preoperative IOP, P = 0.414).

Table 2
Univariable and multivariable Firth penalised logistic regression for postoperative LP or NLP

Predictor	Univariable OR (95% CI)	P value	Multivariable OR (95% CI)	P value
Age, per year	1.05 (0.98–1.13)	0.142	-	-
Male sex	1.63 (0.45–6.21)	0.449	-	-
Duration of diabetes, per year	1.03 (1.01–1.22)	0.002	-	-
Duration of hypertension, per year	1.06 (0.99–1.12)	0.109	-	-
HbA1c, per %	2.85 (1.76–5.10)	< 0.001	1.92 (1.05–4.35)	0.034
Insulin therapy	0.26 (0.06–0.95)	0.042	-	-
Diabetic nephropathy	31.52 (3.92-4082.19)	< 0.001	15.50 (1.21-2397.50)	0.033
Maintenance haemodialysis	28.37 (6.99-123.24)	< 0.001	-	-
Retinal macrovascular occlusion (RMO)	0.80 (0.22–3.46)	0.751	-	-
Iris neovascularisation (NVI)	194.14 (37.56-1432.91)	< 0.001	39.88 (6.29-405.45)	< 0.001
PDR severity, per grade	4.35 (1.93–14.74)	< 0.001	-	-
Tamponade agent, per grade	0.45 (0.20–0.92)	0.028	-	-
Preoperative BCVA, per logMAR	0.81 (0.40–1.67)	0.553	-	-
Preoperative IOP, per mmHg	1.07 (0.90–1.24)	0.414	1.05 (0.71–1.45)	0.801

OR = odds ratio; CI = confidence interval; other abbreviations as in Table 1. P values and 95% CIs are from profile penalised likelihood. Tamponade grade: silicone oil < gas < air/water. PDR severity grade: vitreous haemorrhage < fibrovascular proliferation < retinal detachment.

3.3. Multivariable Firth logistic regression

The final Firth model included preoperative NVI, diabetic nephropathy, HbA1c and preoperative IOP (Table 2). Preoperative NVI showed the strongest association (OR 39.88, 95% CI 6.29-405.45; P < 0.001), and diabetic nephropathy (OR 15.50, 1.21-2397.50; P = 0.033) and higher HbA1c (OR 1.92 per %, 1.05–4.35; P = 0.034) were also independently associated with increased risk; the wide confidence interval for nephropathy reflects the small number of events and its complete univariable separation. Preoperative IOP showed no independent association after adjustment (OR 1.05 per mmHg, 0.71–1.45; P = 0.801).

3.4. LASSO regression and random forest

LASSO regression retained 11 of the 14 candidate predictors, shrinking age, maintenance haemodialysis and tamponade agent to zero; the largest penalised coefficients were observed for NVI, diabetic nephropathy, male sex, HbA1c, insulin therapy (inverse) and PDR severity (Fig. 1A). Random forest variable importance ranked PDR severity highest (mean decrease in impurity 0.197), followed by NVI (0.178), HbA1c (0.175), diabetic nephropathy (0.146), age (0.058) and haemodialysis (0.053) (Fig. 1B), in broad agreement with the regression-based analyses.

The higher variable importance of PDR severity in the random forest largely reflects its strong correlation with NVI, which frequently coexists with advanced PDR stages.

3.5. Model performance and internal validation

Performance is summarised in Table 3 and Figs. 1–2. Apparent AUCs were 0.984 (95% CI 0.955-1.000) for the Firth model, 0.997 (0.990-1.000) for LASSO and 1.000 (1.000–1.000) for the random forest (Fig. 1C). Harrell bootstrap optimism was modest (0.013, 0.022 and 0.012, respectively), yielding optimism-corrected AUCs of 0.971, 0.975 and 0.988. Repeated five-fold cross-validation gave more conservative AUCs of 0.962, 0.898 and 0.969 (Fig. 1D), reflecting the instability of fold-based estimates with only nine events. The apparent Cox-Snell R-squared of the Firth model was 0.197, close to the maximum of 0.275 attainable at this outcome proportion, indicating that the model explains only a limited fraction of the outcome variation despite its high AUC. Calibration differed between models (Fig. 2A-2C): the bootstrap calibration slope was 0.95 (95% bootstrap interval 0.39–1.50) for the Firth model and 0.73 (0.27–2.06) for LASSO, both statistically compatible with the ideal value of 1, compared with 2.12 (1.14–3.32) for the random forest, indicating under-dispersed predictions that were not extreme enough: a calibration slope below 1 indicates overfitting with predictions that are too extreme, whereas a slope above 1 indicates under-dispersion or underfitting. Brier scores were 0.013, 0.022 and 0.021, respectively. In decision curve analysis using bootstrap-corrected predictions, the Firth model showed higher net benefit than the treat-all and treat-none strategies across threshold probabilities of approximately 2–50% in this cohort, whereas LASSO and random forest net benefit was less consistent at intermediate thresholds, reflecting their less stable calibration (Fig. 2D). Given the small number of events, this range is imprecisely estimated and should not be interpreted as a generalisable clinical decision interval. The NVI-only model achieved an optimism-corrected AUC of 0.871, compared with 0.971 for the four-predictor model (difference 0.10, 95% bootstrap interval 0.004–0.105; penalised likelihood-ratio test $\chi^2(3) = 16.17$, $P = 0.001$).

Table 3
Discrimination, calibration and overall performance of the three prediction models with internal validation

Model	Apparent AUC (95% CI)	Optimism	Corrected AUC	CV AUC, mean ± SD	Calibration slope (95% BI)	Brier
Firth logistic	0.984 (0.955-1.000)	0.013	0.971	0.962 ± 0.009	0.95 (0.39–1.50)	0.013
LASSO logistic	0.997 (0.990-1.000)	0.022	0.975	0.898 ± 0.029	0.73 (0.27–2.06)	0.022
Random forest	1.000 (1.000–1.000)	0.012	0.988	0.969 ± 0.008	2.12 (1.14–3.32)	0.021

AUC = area under the receiver operating characteristic curve; CI = confidence interval; CV = repeated five-fold cross-validation (10 repeats, out-of-fold predictions); BI = bootstrap percentile interval; SD = standard deviation. Corrected AUC = apparent AUC minus mean optimism across 200 Harrell bootstrap resamples, with the complete modelling pipeline refitted in each resample. Calibration slope is the median across 200 bootstrap resamples of the slope from a logistic regression of the observed outcome on the logit of each resample-refitted model's predictions in the original data; a slope below 1 indicates overfitting with predictions that are too extreme, whereas a slope above 1 indicates under-dispersion or underfitting.

3.6. Nomogram

The Firth model was translated into a nomogram (Fig. 2E) assigning points for NVI, diabetic nephropathy, HbA1c and preoperative IOP. For example, an eye with NVI (80 points), nephropathy (60 points), HbA1c of 10% (71 points) and preoperative IOP of 20 mmHg (12 points) scores 223 points, corresponding to a high model-estimated risk, whereas an eye without NVI or nephropathy, with HbA1c of 7% and IOP of 14 mmHg, scores about 35 points, corresponding to a low model-estimated risk. These examples are illustrative only: with nine events, individual point predictions carry substantial uncertainty and should not be quoted as precise probabilities. Accordingly, at this stage the nomogram must not be used to obtain precise individual risk estimates or to guide individual treatment decisions; it visualises a hypothesis-generating model built on only nine events.

4. Discussion

In this cohort of 238 eyes undergoing PPV for PDR, postoperative LP or NLP occurred in 3.8% of eyes, consistent with historical reports of severe visual loss after vitrectomy in patients with diabetes [5, 6]. Across three complementary modelling strategies, preoperative NVI showed the strongest ocular association with LP or NLP, and higher HbA1c and diabetic nephropathy were also associated after adjustment; maintenance haemodialysis, more advanced PDR stage and longer diabetes duration were associated in univariable analysis. The consistency of variable ranking across Firth regression, LASSO and random forest supports the robustness of these associations.

This rate, at the lower end of historical reports, may reflect modern surgical advances, including the widespread adoption of 25-gauge microincision vitrectomy, preoperative and intraoperative anti-VEGF therapy, and improved visualisation systems that allow more complete vitreous clearance and laser photocoagulation; together, these factors may have reduced the incidence of catastrophic visual outcomes compared with earlier surgical eras [5, 6].

The association of silicone-oil tamponade with LP or NLP in univariable analysis likely reflects confounding by indication: silicone oil is reserved for the most complex cases, with advanced PDR, extensive fibrovascular proliferation or combined tractional-rhegmatogenous retinal detachment. After adjustment for NVI, IOP and systemic factors, the tamponade association was no longer statistically significant, suggesting that silicone-oil tamponade is a marker of disease severity rather than an independent predictor.

Although maintenance haemodialysis showed a strong univariable association with LP or NLP, it was not retained in the final multivariable model because of its high collinearity with diabetic nephropathy (all patients receiving haemodialysis had nephropathy). The final model therefore retained diabetic nephropathy as the more parsimonious systemic marker, while acknowledging that dialysis dependence represents the most severe end of the renal disease spectrum.

The predominance of preoperative NVI accords with the wider literature [7, 20–25]. Severe visual decline to LP or NLP after vitrectomy for diabetic retinopathy has been characterised in dedicated series, in which anterior segment neovascularisation and markers of ocular ischaemia featured prominently among the risk factors [20, 21]. NVI is the precursor of neovascular glaucoma, which remains a frequent, sight-threatening sequela of vitrectomy for PDR, with reported incidences of 2% to more than 20% and established risk factors that include preoperative iris or angle neovascularisation, elevated preoperative IOP and intraocular tamponade [22–24]. Biologically, NVI marks an advanced stage of ocular ischaemia, in which iris neovascularisation coexists with profound retinal ischaemia and optic nerve compromise, and eyes that progress from NVI to NVG after vitrectomy frequently exhaust surgical and pharmacological options [7, 25].

The strong, graded association of HbA1c with LP or NLP, which persisted after adjustment, is biologically plausible: chronic hyperglycaemia drives systemic and retinal microangiopathy, and poorer glycaemic control has repeatedly been linked to worse visual outcomes after ocular surgery in patients with diabetes [8]. In both our previous study and a Japanese study, preoperative HbA1c levels were significantly higher in patients who developed postoperative NVG than in those who did not [26, 27]. HbA1c is also modifiable, making it an attractive target for perioperative optimisation.

The complete association of diabetic nephropathy with the LP or NLP endpoint and the strong univariable association of maintenance haemodialysis suggest that the systemic microvascular burden parallels the severity of ocular microvascular disease. Renal markers have been only inconsistently related to post-vitrectomy neovascular complications in previous studies [23, 24], suggesting that in this cohort renal disease was related to the LP or NLP endpoint itself rather than acting through NVG.

Nephropathy was retained in the final model as the broader renal marker, whereas dialysis, which is nested within advanced nephropathy, was attenuated after adjustment.

Retinal macrovascular occlusion was present in more than two-thirds of eyes (69.7%) in this cohort. This high prevalence is not unexpected: vascular white lines represent sheathed, occluded retinal vessels and are a characteristic end-stage finding of advanced PDR, reflecting cumulative retinal ischaemia and vascular shutdown rather than a discrete comorbid event [12]. The absence of an association between RMO and postoperative LP or NLP (univariable OR 0.80, P = 0.751) suggests that, once active neovascular drive and systemic microvascular burden are accounted for, this static anatomical marker of completed vascular closure adds little prognostic information, although extensive white-line formation may still serve as a useful visual cue of severe underlying ischaemia during preoperative assessment.

In contrast to two independent cohorts in which higher preoperative IOP predicted subsequent NVG after vitrectomy for PDR [23, 24], preoperative IOP showed no association with LP or NLP in this cohort, either in univariable analysis (OR 1.07 per mmHg, P = 0.414) or after adjustment (OR 1.05 per mmHg, P = 0.801). The preoperative IOP distribution was narrow and largely within normal limits, and any ischaemic signal carried by IOP may have been captured more directly by NVI itself; alternatively, IOP may mark earlier stages of the ischaemic-neovascular cascade that precede, rather than determine, the terminal LP or NLP endpoint.

No predictive value was observed for preoperative visual acuity: acuity was similar between outcome groups and was not retained by any modelling strategy. In PDR, BCVA is heavily influenced by media opacity (vitreous haemorrhage) rather than by underlying retinal function, and eyes with tractional retinal detachment may paradoxically present with better BCVA than those with vitreous haemorrhage alone, because the retina remains attached in some areas; measured acuity therefore reflects the degree of media opacity more than the severity of retinal ischaemia or structural damage. Once NVI and systemic microvascular status are accounted for, the preoperative functional visual level adds little information about the subsequent risk of LP or NLP, which appears to be driven more by ongoing ischaemic and neovascular activity than by the acuity measured before surgery.

Methodologically, this study illustrates the value of matching the estimator to the data. With nine events, classical logistic regression fails because of complete separation, as demonstrated by diabetic nephropathy; Firth penalisation yielded finite, bias-reduced, fully interpretable estimates suitable for nomogram construction [13, 14, 19]. Because no public dataset permits external validation of this endpoint, all models underwent enhanced internal validation. The Harrell bootstrap estimated modest optimism and corrected AUCs of 0.971–0.988, whereas repeated cross-validation gave lower, more variable estimates, a known behaviour of fold-based resampling when events are scarce [17, 28]. The models diverged on calibration, however: the Firth and LASSO models achieved calibration slopes compatible with unity, whereas the random forest, despite the highest corrected AUC, produced under-dispersed predictions (a calibration slope above 1, meaning that estimated risks were insufficiently spread between low and high), and decision curve analysis also favoured the Firth model within this

cohort. For rare, high-stakes outcomes, calibrated probabilities matter as much as rank discrimination; the Firth model is therefore the most appropriate candidate for preliminary clinical assessment and was translated into a nomogram. A formal comparison against an NVI-only model (Section 3.5) quantifies how much HbA1c, diabetic nephropathy and IOP add beyond the dominant NVI signal; given the wide uncertainty intervals, any incremental value should be interpreted cautiously and confirmed in larger cohorts. It should also be acknowledged that the three modelling strategies were not strictly comparable: the Firth model used four prespecified predictors, whereas LASSO and the random forest used all 14 candidates, so the superior calibration of the Firth model partly reflects its smaller predictor set rather than the estimation method alone. It should also be noted that the exact set of predictors retained by the LASSO was sensitive to the resampling procedure: with only nine events the cross-validated AUC profile was flat, and repeating the analysis with different resampling seeds retained between 6 and 11 of the 14 candidate predictors, although NVI, diabetic nephropathy, HbA1c, insulin therapy (inverse), male sex and PDR severity consistently showed the largest penalised coefficients; the precise retention count reported in Section 3.4 should therefore not be over-interpreted. The four-predictor model significantly outperformed the NVI-only model (penalised likelihood-ratio test $P = 0.001$; corrected AUC 0.971 vs 0.871), indicating that HbA1c, diabetic nephropathy and IOP add genuine predictive value beyond NVI alone, although the magnitude of this gain is imprecisely estimated with nine events.

In clinical terms, the decision curve analysis suggests that, within this cohort and across threshold probabilities of approximately 2–50%, using the Firth model to stratify patients by predicted risk yields greater net benefit than treating all or none as high risk; with only nine events this range is unstable and cohort-specific and should not be read as a universally applicable clinical benefit interval. The model can therefore help identify eyes that warrant priority for closer postoperative surveillance and more detailed counselling: for example, eyes combining NVI, poorly controlled diabetes (HbA1c > 9%) and diabetic nephropathy had among the highest model-estimated risks in this cohort, far above any threshold in this range. It must be emphasised, however, that this observational study does not demonstrate that intensified anti-VEGF therapy or tighter glycaemic control reduces the incidence of LP or NLP; the model supports risk stratification and surveillance planning rather than specific treatment recommendations, and any treatment modification guided by predicted risk requires prospective evaluation. Conversely, low-risk patients (predicted probability < 2%) might be managed with standard protocols, although this too requires prospective confirmation. These risk thresholds should be validated in prospective multicentre studies before routine clinical adoption.

The random forest was included as a methodological benchmark; despite the highest corrected AUC it was poorly calibrated (calibration slope 2.12, indicating under-dispersion). Its results are therefore de-emphasised, and the well-calibrated Firth model is carried forward as the primary model.

The nomogram offers a transparent bedside approximation of the Firth model, and the machine-learning models could be integrated into clinical workflows after prospective, adequately powered external validation in an independent multicentre cohort. Given the very small number of events underlying the

model, the nomogram's points and predicted probabilities carry substantial uncertainty and should be regarded as illustrative rather than as precise patient-level probabilities.

5. Limitations

This study has several limitations. First, external validation was not performed because no accessible dataset contains individual-level data for this endpoint, and the models may perform less well in independent populations. Second, the number of events (nine LP or NLP eyes) was small, yielding wide confidence intervals. Formal sample size assessment indicated an events-per-variable ratio of 2.25, below the conventional threshold of 10 [11], and the cohort was adequate only under optimistic assumptions about model explanatory power: an anticipated Cox-Snell R-squared of 0.15 requires 219 participants, whereas a value of 0.10 would require 340 [10]. Consistent with this, apparent performance was high (Cox-Snell R-squared 0.197, close to the maximum of 0.275 attainable at this outcome proportion) and the bootstrap calibration slope, although close to unity, was imprecisely estimated (0.95, 95% interval 0.39–1.50), leaving residual overfitting possible; the models should therefore be regarded as hypothesis-generating pending external validation in an adequately powered cohort. Third, the retrospective design precludes causal inference and is subject to selection and information bias, including confounding by indication for silicone-oil tamponade. Fourth,Tianjin Eye Hospital is the largest ophthalmic referral centre in the region, which may limit generalisability.

Two further limitations warrant mention. The four predictors of the final Firth model were prespecified on prior clinical evidence without any univariable screening (Section 2.3.1); nevertheless, because they were chosen from among 14 candidates with knowledge of the cohort, some residual selection optimism cannot be excluded, although the concordant variable rankings of the unscreened LASSO and random forest analyses support the robustness of the selected set. In addition, the binary endpoint discards the timing of events within the five-year horizon; a time-to-event analysis was considered but is unreliable with nine events, although the complete five-year follow-up of all non-endpoint eyes means that censoring did not affect outcome ascertainment. Lastly, LP and NLP were combined into a single composite endpoint out of statistical necessity, because nine events do not permit separate analyses; although both grades constitute legal blindness, they differ subtly in residual visual function and psychological impact, and merging them slightly limits the clinical granularity of the endpoint. Furthermore, in this cohort of patients with advanced diabetes, some receiving maintenance haemodialysis, death during the five-year follow-up constitutes a competing risk. If patients who died before reaching the endpoint were analysed as non-events rather than censored, the incidence of LP or NLP would be systematically underestimated; a formal competing-risk analysis was not feasible with nine events and should be incorporated in future prospective studies. Moreover, the postoperative treatment pathway—including the development and management of neovascular glaucoma and any postoperative anti-VEGF therapy or additional panretinal photocoagulation—was not modelled; these factors are time-dependent mediators or treatments on the path from the baseline predictors to LP or NLP, and their omission limits mechanistic interpretation.

6. Conclusions

Preoperative iris neovascularisation, poor glycaemic control and diabetic nephropathy were the factors most strongly associated with LP or NLP after vitrectomy for PDR in this cohort, and haemodialysis, diabetes duration and disease severity were associated in univariable analysis; preoperative visual acuity and preoperative IOP added no predictive value after adjustment. Firth penalised logistic regression, LASSO regression and random forest all achieved high apparent discrimination under enhanced internal validation, although these estimates are unreliable with nine events, and the well-calibrated Firth model offers the most interpretable, nomogram-ready framework. These tools may support risk stratification and the identification of eyes warranting closer surveillance, but they do not by themselves justify changes in treatment, and they require prospective, adequately powered external validation before routine clinical adoption. Accordingly, these models should be regarded strictly as hypothesis-generating and are not yet mature enough for clinical application. Overall, this study is best regarded as a risk factor study with exploratory predictive modelling, providing candidate variables and hypotheses for future multicentre external validation.

Abbreviations

AUC
Area under the receiver operating characteristic curve
BCVA
Best-corrected visual acuity
CI
Confidence interval
CRediT
Contributor Roles Taxonomy
CV
Cross-validation
ETDRS
Early Treatment Diabetic Retinopathy Study
HbA1c
Glycated haemoglobin
IFCC
International Federation of Clinical Chemistry and Laboratory Medicine
IOP
Intraocular pressure
LASSO
Least absolute shrinkage and selection operator
LP
Light perception
NGSP

National Glycohemoglobin Standardization Program
NLP
No light perception
NVI
Iris neovascularisation
NVG
Neovascular glaucoma
OR
Odds ratio
PDR
Proliferative diabetic retinopathy
PPV
Pars plana vitrectomy
ROC
Receiver operating characteristic
RMO
Retinal macrovascular occlusion
SD
Standard deviation
TRD
Tractional retinal detachment
TRIPOD
Transparent Reporting of a multivariable prediction model for Individual Prognosis or Diagnosis
VEGF
Vascular endothelial growth factor

Declarations

Ethics approval and consent to participate

The study was approved by the institutional review board of Tianjin Eye Hospital and conducted in accordance with the Declaration of Helsinki; written informed consent was obtained from all participants.

Consent for publication

Not applicable.

Availability of data and materials

The de-identified dataset and analysis code are available from the corresponding author on reasonable request.

Competing interests

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Funding

This study and the article processing charges were funded by the Tianjin Key Medical Discipline (Specialty) Construction Project (No. TJYXZDXK-016A), Internal Funds of Tianjin Eye Hospital (YKQN2008), and Internal Funds of Tianjin Eye Hospital (YKZD1908).

Authors' contributions

Xu Liang: Conceptualization, Methodology, Investigation, Data curation, Formal analysis, Writing – original draft, Writing – review & editing, Supervision,Methodology,Project administration.

Jiaxing Wang: Formal analysis, Validation, Methodology; Software; Data curation

Yu-peng Li: Investigation, Resources, Assessment of Blood Glucose and Medical Conditions

Hui Zheng: Investigation, Writing

Yajie Wang: Investigation, Writing .

Yue Zhang: Conceptualization, Methodology, .

Acknowledgements

The authors thank the nursing and technical staff of the participating vitreoretinal units for their assistance with data collection and patient care throughout the study. We are grateful to [Name(s), Affiliation] for [methodological and statistical support with model development and validation]; to [Name(s)] for [critical revision of the manuscript / contribution of cases]; and to [English language editing service or colleague name] for assistance with English language editing. The content and writing of the manuscript are the responsibility of the authors alone.

References

1. Diabetic Retinopathy Vitrectomy Study Research Group. Early vitrectomy for severe vitreous hemorrhage in diabetic retinopathy. Two-year results of a randomized trial. Diabetic Retinopathy Vitrectomy Study report 2. *Arch. Ophthalmol.* **103**, 1644–1652 (1985).
2. Yau, J. W. Y. et al. Global prevalence and major risk factors of diabetic retinopathy. *Diabetes Care*. **35**, 556–564 (2012).
3. Cheung, N., Mitchell, P. & Wong, T. Y. Diabetic retinopathy. *Lancet* **376**, 124–136 (2010).
4. Wong, T. Y. et al. Guidelines on diabetic eye care: the International Council of Ophthalmology recommendations for screening, follow-up, referral, and treatment based on resource settings. *Ophthalmology* **125**, 1608–1622 (2018).
5. Diabetic Retinopathy Vitrectomy Study Research Group. Early vitrectomy for severe proliferative diabetic retinopathy in eyes with useful vision. Clinical application of results of a randomized trial - Diabetic Retinopathy Vitrectomy Study report 4. *Ophthalmology* **95**, 1321–1334 (1988).
6. Gupta, B. et al. Visual and anatomical outcomes following vitrectomy for complications of proliferative diabetic retinopathy: the DRIVE UK study. *Eye (Lond)*. **26**, 510–516 (2012).
7. Simo-Servat, O., Hernandez, C. & Simo, R. Diabetic retinopathy in the context of patients with diabetes. *Ophthalmic Res.* **62**, 211–217 (2019).
8. McCullough, P., Mohite, A., Virgili, G. & Lois, N. Outcomes and complications of pars plana vitrectomy for tractional retinal detachment in people with diabetes: a systematic review and meta-analysis. *JAMA Ophthalmol.* **141**, 186–195 (2023).
9. Collins, G. S., Reitsma, J. B., Altman, D. G. & Moons, K. G. M. Transparent Reporting of a multivariable prediction model for Individual Prognosis or Diagnosis (TRIPOD): the TRIPOD statement. *Ann. Intern. Med.* **162**, 55–63 (2015).
10. Riley, R. D. et al. Calculating the sample size required for developing a clinical prediction model. *BMJ* **368**, m441 (2020).
11. Peduzzi, P., Concato, J., Kemper, E., Holford, T. R. & Feinstein, A. R. A simulation study of the number of events per variable in logistic regression analysis. *J. Clin. Epidemiol.* **49**, 1373–1379 (1996).
12. Early Treatment Diabetic Retinopathy Study Research Group. Grading diabetic retinopathy from stereoscopic color fundus photographs—an extension of the modified Airlie House classification: ETDRS report number 10. *Ophthalmology* **98** (5 Suppl), 786–806 (1991).
13. Firth, D. Bias reduction of maximum likelihood estimates. *Biometrika* **80**, 27–38 (1993).
14. Heinze, G. & Schemper, M. A solution to the problem of separation in logistic regression. *Stat. Med.* **21**, 2409–2419 (2002).
15. Tibshirani, R. Regression shrinkage and selection via the lasso. *J. R Stat. Soc. Ser. B Stat. Methodol.* **58**, 267–288 (1996).
16. Breiman, L. Random forests. *Mach. Learn.* **45**, 5–32 (2001).
17. Harrell, F. E. Jr, Lee, K. L. & Mark, D. B. Multivariable prognostic models: issues in developing models, evaluating assumptions and adequacy, and measuring and reducing errors. *Stat. Med.* **15**, 361–387

(1996).
18. Vickers, A. J. & Elkin, E. B. Decision curve analysis: a novel method for evaluating prediction models. *Med. Decis. Mak.* **26**, 565–574 (2006).
19. Iasonos, A., Schrag, D., Raj, G. V. & Panageas, K. S. How to build and interpret a nomogram for cancer prognosis. *J. Clin. Oncol.* **26**, 1364–1370 (2008).
20. Mason, J. O. 3 et al. Visual outcome and risk factors for light perception and no light perception vision after vitrectomy for diabetic retinopathy. *Am. J. Ophthalmol.* **140**, 231–235 (2005).
21. Yorston, D. et al. Predictive clinical features and outcomes of vitrectomy for proliferative diabetic retinopathy. *Br. J. Ophthalmol.* **92**, 365–368 (2008).
22. Goto, A. et al. Frequency and risk factors for neovascular glaucoma after vitrectomy in eyes with proliferative diabetic retinopathy. *J. Glaucoma*. **22**, 572–576 (2013).
23. Sakamoto, M., Hashimoto, R., Yoshida, I., Ubuka, M. & Maeno, T. Risk factors for neovascular glaucoma after vitrectomy in eyes with proliferative diabetic retinopathy. *Clin. Ophthalmol.* **12**, 2323–2329 (2018).
24. Hwang, S. et al. Incidence and risk factor analysis of neovascular glaucoma following vitrectomy in patients with proliferative diabetic retinopathy. *Sci. Rep.* **14**, 22490 (2024).
25. Rodrigues, G. B. et al. Neovascular glaucoma: a review. *Int. J. Retina Vitreous*. **2**, 26 (2016).
26. Takayama, K. et al. Risk factors of neovascular glaucoma after 25-gauge vitrectomy for proliferative diabetic retinopathy with vitreous hemorrhage: a retrospective multicenter study. *Sci. Rep.* **9**, 14858 (2019).
27. Xu, L. et al. Frequency and risk factors for neovascular glaucoma after vitrectomy in eyes with diabetic retinopathy: an observational study. *Diabetes Ther.* **10**, 1801–1809 (2019).
28. Steyerberg, E. W. *Clinical Prediction Models: A Practical Approach to Development, Validation, and Updating* 2nd edn (Springer, 2019).
29. Collins, G. S. et al. TRIPOD + AI statement: updated guidance for reporting clinical prediction models that use regression or machine learning methods. *BMJ* **385**, e078378 (2024).

Figures

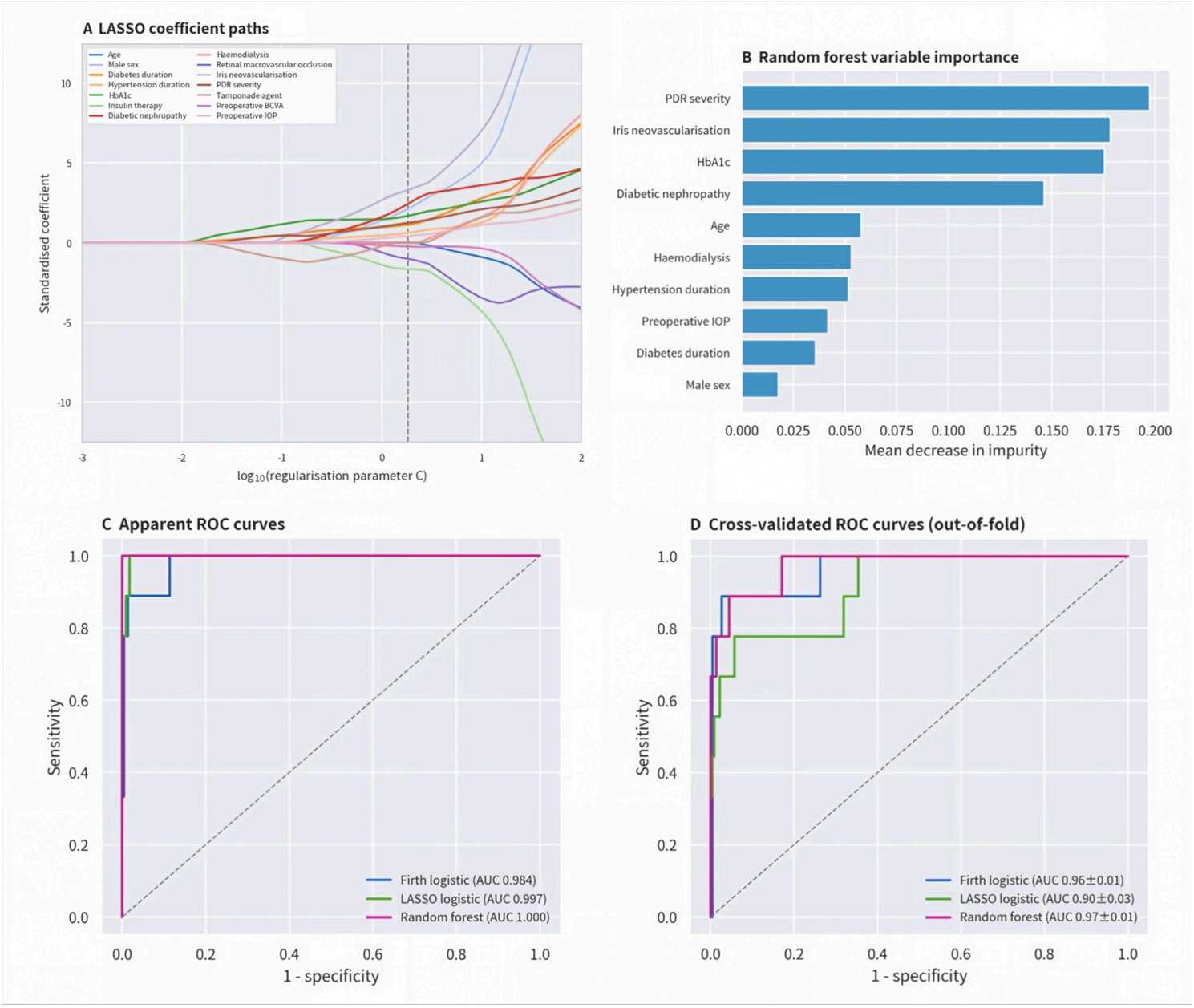


Figure 1

Variable selection, variable importance and discrimination of the three prediction models. (A) LASSO standardised coefficient paths across the penalisation sequence; the dashed line marks the cross-validated optimal penalty. (B) Random forest variable importance (mean decrease in impurity). (C) Apparent receiver operating characteristic (ROC) curves. (D) Cross-validated ROC curves based on pooled out-of-fold predictions (five-fold, 10 repeats; values are mean ± standard deviation across repeats). NVI = iris neovascularisation; DN = diabetic nephropathy; IOP = intraocular pressure; AUC = area under the ROC curve.

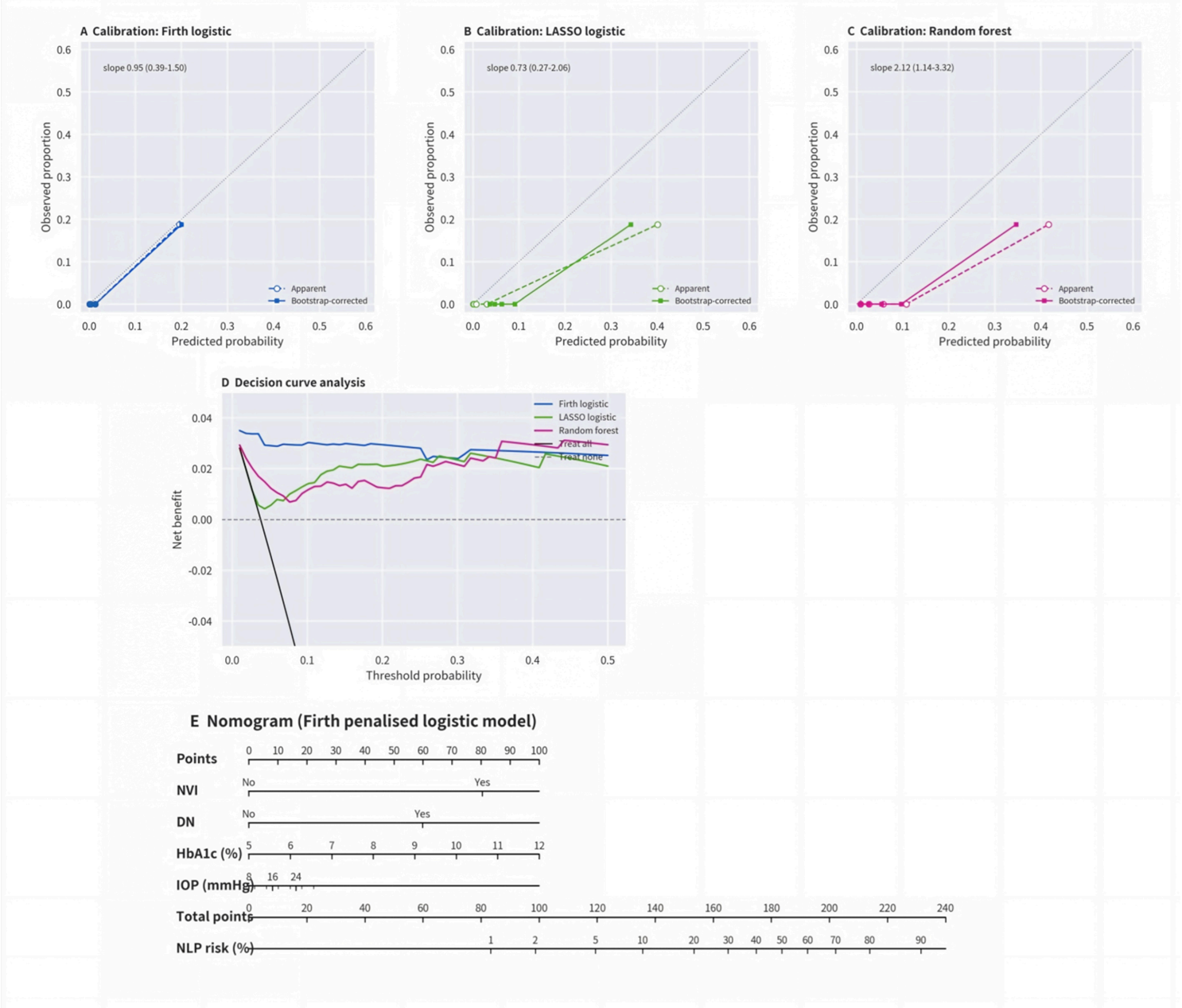


Figure 2

Calibration, clinical utility and nomogram. (A-C) Calibration of the Firth, LASSO and random forest models: apparent quintile-based calibration (circles, dashed) versus calibration derived from pooled bootstrap-model predictions on the original dataset (squares, solid); the dotted line denotes perfect calibration and the bootstrap calibration slope (95% interval) is shown. (D) Decision curve analysis using bootstrap-corrected predictions; net benefit is plotted against threshold probability. (E) Nomogram for predicting postoperative LP or NLP derived from the Firth penalised logistic model. NVI = iris neovascularisation; DN = diabetic nephropathy; IOP = intraocular pressure. The 2-50% threshold range in panel D is shown for completeness; with nine events it should not be interpreted as a generalisable clinical decision interval.