npj | digital medicine

Article

Published in partnership with Seoul National University Bundang Hospital

https://doi.org/10.1038/s41746-025-01582-6

Noninvasive early prediction of preeclampsia in pregnancy using retinal vascular features

Yuxuan Wu[1,8], Lixia Shen[2,8], Lanqin Zhao[1,8], Xiaohong Lin[2,8], Miaohong Xu[1], Zhenjun Tu[1], Yihong Huang[2], Lingyi Kong[2], Zhenzhe Lin[1], Duoru Lin[1], Lixue Liu[1], Xun Wang[1], Zizheng Cao[1], Xi Chen[1], Shengmei Zhou[1], Weiling Hu[1], Yunjian Huang[1], Shiyuan Chen[1], Meimei Dongye[1], Xulin Zhang[1], Dongni Wang[1], Danli Shi[3,4,5], Zilian Wang[2] ✉, Xiaohang Wu[1] ✉, Dongyu Wang[2] ✉ & Haotian Lin[1,6,7] ✉

Preeclampsia (PE), a severe hypertensive disorder during pregnancy, significantly contributes to maternal and neonatal mortality. Existing prediction biomarkers are often invasive and expensive, hindering their widespread application. This study introduces PROMPT (Preeclampsia Risk factor + Ophthalmic data + Mean arterial pressure Prediction Test), an AI-driven model leveraging retinal photography for PE prediction, registered at ChiCTR (ChiCTR2100049850) in August 2021. Analyzing 1812 pregnancies before 14 gestational weeks, we extracted retinal parameters using a deep learning system. The PROMPT achieved an AUC of 0.87 (0.83–0.90) for PE prediction and 0.91 (0.85–0.97) for preterm PE prediction using machine learning, significantly outperforming the baseline model ($p < 0.001$). It also improved detection of severe adverse pregnancy outcomes from 35% to 41%. Economically, PROMPT was estimated to avert 1809 PE cases and saved over $50 million per 100,000 screenings. These results position PROMPT as a non-invasive and cost-effective tool for prenatal care, especially valuable in low- and middle-income countries.

Preeclampsia (PE) is a life-threatening hypertensive disorder in pregnancy and a major cause of maternal and neonatal morbidity and mortality, which affects around 4.6% of pregnant women worldwide[1]. It results in an increased risk of adverse pregnancy outcomes (AO) including preterm birth (PTB), small for gestational age (SGA), and even stillbirth, ranking the second leading cause of maternal mortality in China[2]. Such outcomes highlight the need for accessible tools for identifying high-risk pregnancies before clinical presentation, especially in low-and-middle-income countries (LMICs)[3,4].

Assessment for pregnant women at high-risk for PE based on risk factors of maternal characteristics and obstetric history, such as advanced maternal age, chronic hypertension, renal disease, obesity, and PE history, has shown suboptimal performance and cannot quantify patient-specific risk[5–7]. Recent research has advanced the identification of novel biomarkers to predict PE in early pregnancy. The Fetal Medicine Foundation triple test, which combines maternal risk factors with biomarkers including mean arterial pressure (MAP), uterine artery pulsatility index (UtA-PI), placental growth factor (PlGF) and pregnancy-associated plasma protein-A (PAPP-A), has achieved moderate to high discrimination in different populations[7–11]. However, the need for a qualified sonographer to measure UtA-PI and additional blood sampling for serum markers limits their widespread application in clinical practice, especially in resource-limited areas. Analysis of circulating cell-free DNA (cfDNA) methylation and cell-free RNA also offers the possibility of early prediction of PE but is expensive and beyond the reach of clinical practice[12–14]. Exploring a

[1]State Key Laboratory of Ophthalmology, Zhongshan Ophthalmic Center, Sun Yat-sen University, Guangdong Provincial Key Laboratory of Ophthalmology and Vision Science, Guangdong Provincial Clinical Research Center for Ocular Diseases, Guangzhou, Guangdong, China. [2]The First Affiliated Hospital, Sun Yat-sen University, Guangzhou, Guangdong, China. [3]School of Optometry, The Hong Kong Polytechnic University, Kowloon, Hong Kong, China. [4]Research Centre for SHARP Vision (RCSV), The Hong Kong Polytechnic University, Kowloon, Hong Kong, China. [5]Centre for Eye and Vision Research (CEVR), 17W Hong Kong Science Park, Hong Kong, China. [6]Hainan Eye Hospital and Key Laboratory of Ophthalmology, Zhongshan Ophthalmic Center, Sun Yat-sen University, Haikou, Hainan, China. [7]Center for Precision Medicine and Department of Genetics and Biomedical Informatics, Zhongshan School of Medicine, Sun Yat-sen University, Guangzhou, Guangdong, China. [8]These authors contributed equally: Yuxuan Wu, Lixia Shen, Lanqin Zhao, Xiaohong Lin. ✉e-mail: wangzil@mail.sysu.edu.cn; wxhang@mail2.sysu.edu.cn; wangdy28@mail.sysu.edu.cn; linht5@mail.sysu.edu.cn

non-invasive, accessible test with comparable or superior predictive performance to current biomarkers is crucial for primary prenatal care in LMICs.

Mounting evidence has shown that the eye can serve as a 'window', i.e., the sentinel sensory organ, to observe microvascular alterations in systemic disorders[15–19]. Recent advances in artificial intelligence (AI) have further expanded the potential to identify systemic diseases through ophthalmic features, such as cardiovascular diseases, chronic kidney disease, hepatobiliary disease, Alzheimer's disease, etc[20–23]. In previous studies, retinal vessel spasms, tortuosity alterations, hemorrhage, and cotton-wool spots were detected in PE patients, and some serious cases might even suffer from retinal detachment and blindness[24–26]. Elevated blood pressure affects the retinal and choroidal circulation, and the mechanism is thought to be ischemia[24]. In obstetric clinical practice, the PE patients are referred to retinal examinations in case of ophthalmic complications. However, whether ocular manifestations occur before PE diagnosis has not been fully explored. We performed a prospective cohort to observe the early retinal features before PE occurrence and found that vascular changes were marked in early pregnancy before the onset of symptoms and clinical diagnosis. These observations enabled us to explore their potential to be novel and sensitive biomarkers for PE prediction.

In this study, we present a machine-learning (ML) based algorithm for PE prediction using retinal vascular features in a non-invasive and cost-effective approach. The model combines maternal risk factors, MAP, and one retinal photograph before 14 weeks of gestation, and aims to achieve better performance in PE prediction, AO detection, and economic evaluation than current practice (Fig. 1). Our study aims to deliver the novel prediction and screening biomarkers that are noninvasive, objective, low-cost and can be measured in primary care settings, thus suggesting new approaches to the early prediction and management of PE and related AO in under-resourced areas-lacking settings.

Results

Participants and demographics

In total, 2114 pregnancies were enrolled initially, of which 302 were excluded in the analysis. The remaining 1812 participants consisted of 86 (4.7%) PE cases and 1726 (95.3%) healthy controls (HC) (Supplementary Fig. 1). The demographic, clinical, and retinal variables between the PE and HC groups were shown in Table 1. Compared with HC, the PE patients were more likely to be older (33.36 Vs. 32.22 years old), higher in body mass index (BMI) (23.77 Vs. 21.26 kg/m^2) and MAP (91.95 Vs. 83.88 mmHg) in the 11-13^{+6} gestational weeks. Additionally, PE patients showed increased rate of renal disease, chronic hypertension, and PE history. For retinal vascular features, the artery-to-vein ratio (AVR) and angle-based tortuosity (angtort) in the first trimester decreased significantly in the PE group. The retinal vessel skeleton density (VSD) and tortuosity density (tortD) increased in PE patients. A total of 393 (21.7%) pregnancies developed into AO, including 86 (21.9%) cases of PE, 206 (52.4%) cases of SGA, 146 (37.2%) cases of PTB. There were 98 (5.4%) cases of severe AO, including 29 (29.6%), 53 (54.1%) and 27 (27.6%) cases of preterm PE, severe SGA and early PTB in total.

Algorithm selection and variable selection

We used five machine learning algorithms to build PE prediction models based on clinical risk factors, MAP, and retinal parameters in 11-13^{+6} gestational weeks. The area under the receiver operating characteristic curve (AUC) of logistic regression (LR), random forest (RF), extreme gradient boosting (XGBoost), gradient boosting method (GBM), and supporting vector machine (SVM) models in the 10-fold cross validation (CV) were 0.863 ± 0.049, 0.863 ± 0.060, 0.858 ± 0.059, 0.860 ± 0.059, and 0.867 ± 0.063, respectively (Supplementary Table 1). Among all the models, LR had robust performance and highlighted a clear contribution of each variable, making it useful for the clinical implementation, thus being selected as the final model. Four models for PE prediction were constructed: (1)

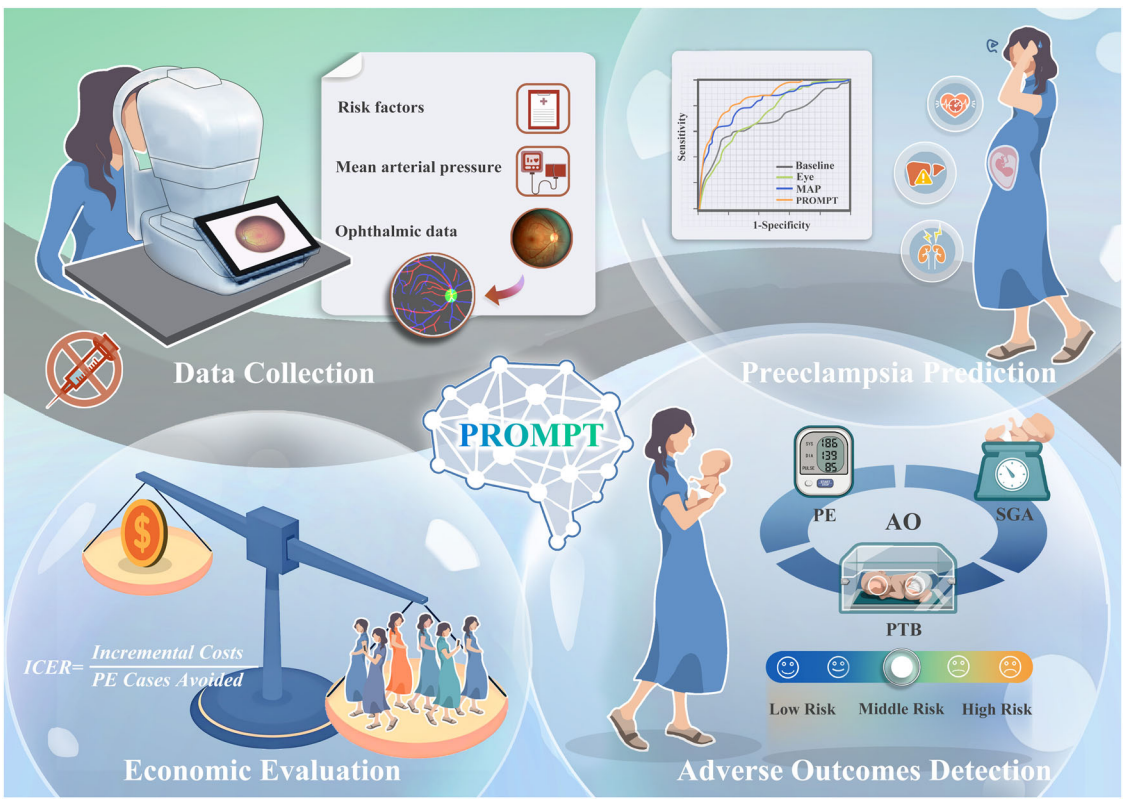

Fig. 1 | Overall workflow of the PROMPT model. The study includes data collection and types, model development and validation for PE prediction, detection of PE-related adverse pregnant outcomes, and cost-effectiveness analysis of the PROMPT model in population screening. MAP mean arterial pressure. PE preeclampsia, SGA small for gestational age, PTB preterm birth, AO adverse outcomes. ICER incremental cost effectiveness ratio.

Table 1 | Demographic information of participants

Variables	All (N = 1812)	HC (n = 1726)	PE (n = 86)
Age (years)	32.28 ± 4.19	32.22 ± 4.16	33.36 ± 4.28*
BMI (kg/m^2)	21.37 ± 3.14	21.26 ± 3.02	23.77 ± 4.29***
Overweight	289 (15.95)	246 (14.25)	43 (50.00)***
MAP (mmHg)	84.27 ± 6.21	83.88 ± 5.75	91.95 ± 9.37***
Renal disease	18 (0.99)	11 (0.64)	7 (8.14)***
PE history	7 (0.39)	4 (0.23)	3 (3.49)***
Chronic hypertension	16 (0.88)	10 (0.58)	6 (6.98)***
Autoimmune disease	49 (2.70)	46 (2.66)	3 (3.49)
Pregestational DM	18 (0.99)	15 (0.86)	3 (3.49)
Hypertension family history	362 (20.00)	133 (19.29)	29 (33.72)**
Conception by IVF	930 (51.32)	886 (51.33)	44 (51.16)
Nulliparous	1203 (66.39)	1146 (66.40)	57 (66.28)
AVR	0.876 ± 0.041	0.877 ± 0.041	0.853 ± 0.038***
VSD	0.263 ± 0.017	0.263 ± 0.017	0.270 ± 0.017***
TortD	0.497 ± 0.026	0.497 ± 0.026	0.504 ± 0.027*
Angtort	0.050 ± 0.017	0.050 ± 0.017	0.046 ± 0.016*
Angtort_std	0.231 ± 0.058	0.231 ± 0.057	0.237 ± 0.060

Mean ± SD or N (percentage). Independent t-tests or $\chi 2$ tests were performed between PE with HC. BMI was calculated as a person's weight (kg) divided by the square of the height (m). *$p < 0.05$, ***$p < 0.001$.

HC healthy controls, *PE* preeclampsia, *BMI* body mass index. MAP mean arterial pressure. *DM* diabetes mellitus, *IVF* in vitro fertilization, *AVR* artery-to-vein ratio, *VSD* vessel skeleton density, *TortD* tortuosity density. *Angtort* angle-based tortuosity, *Std* standard deviation.

baseline model; (2) eye model; (3) MAP model; and (4) PROMPT model. The baseline model was built using risk factors from the clinical guidelines as reference to low-resource settings, including age over 35 years old, overweight, nulliparous, chronic hypertension, renal disease, autoimmune disease, pregestational diabetes mellitus, and PE history[27,28]. For the PROMPT model, we merged all the selected variables by least absolute shrinkage and selection operator (LASSO) logistic regression with minimum criteria (AUC maximum) to obtain a 9-feature set, namely, MAP, overweight, renal disease history, PE history, AVR, VSD, tortD, angtort, and the standard deviation (std) of angtort. The eye model was built with selected retinal parameters only (AVR, VSD, tortD, angtort, and std of angtort). The MAP model was built with selected maternal risk factors and MAP (overweight, renal disease history, PE history, and MAP).

The performance of PE prediction models

The LR model performances for the baseline, eye, MAP, and PROMPT models achieved AUC of 0.69 (0.63–0.76), 0.74 (0.69–0.79), 0.82 (0.78–0.87) and 0.87 (0.83–0.90), respectively (Fig. 2a). Decision curve analysis (DCA) showed that the PROMPT model performed the best across a wide range of risk thresholds (Fig. 2b). The PROMPT model for preterm PE and term PE prediction achieved the AUC of 0.91 (0.85–0.97) and 0.84 (0.80–0.89), respectively (Fig. 2c). In the subgroup analysis, the PROMPT model in parous and nulliparous groups achieved AUC of 0.90 (0.85–0.95) and 0.85 (0.80–0.90) (Fig. 2d). In the PROMPT model, the top five most important variables were MAP, overweight, tortD, VSD, and AVR (Fig. 2e). The lift plot demonstrated the discriminative performance of the PROMPT model (Fig. 2f). The nomogram was displayed as a visual and practical tool for bedside use to predict PE incidence (Supplementary Fig. 2). In this case, a patient was predicted to have PE based on more than half of contributions from retinal features, who eventually developed PE. The addition of retinal parameters to the PROMPT model yielded a significant improvement compared to the baseline model ($p < 0.001$) and MAP model ($p = 0.002$) using Delong test. Compared with the baseline and MAP model, the overall net reclassification index (NRI) for PROMPT model was 0.83 (0.63–1.04) and 0.60 (0.40–0.81), respectively. The NRI+ were 0.40 (0.20–0.60), 0.28 (0.08 to 0.48), and the NRI- was 0.43 (0.39–0.47), and 0.32 (0.28–0.37), suggesting that the inclusion of retinal parameters could improve the model performance both in case and non-case detection. The integrated discrimination improvement (IDI) indexes were 0.09 (0.05–0.14) and 0.03 (0.01–0.05), respectively (Table 2).

Detection rate (DR) of adverse pregnancy outcomes in screen-positive pregnancies

At false positive rate (FPR) of 5%, the DR for AO in baseline, MAP, and PROMPT models were 9.7%, 13.5%, and 16.8%, respectively (Table 3 and Supplementary Table 2). The DR rose to 16.5%, 22.4%, and 23.4% at FPR of 10%, and 23.9%, 28.2%, and 31.0% at FPR of 15%. For severe AO, the three models could achieve DR of 13.3%, 22.4%, and 27.6% at FPR of 5%, and 21.4%, 31.6%, and 35.7% at FPR of 10%. When it comes to FPR of 15%, the DR were 34.7%, 37.8%, and 40.8%, respectively, indicating a higher DR of AO and severe AO using the PROMPT model compared with the baseline and MAP models.

Economic evaluation of different PE prediction models

The total costs per person for no prediction, the baseline, MAP, and PROMPT prediction models were $2142.63, $2104.60, $2096.52, and $2091.29, respectively (Table 4). The PROMPT model was proved dominant to the other prediction methods, averting more PE cases while saving total societal costs. The PROMPT model could avoid 1809, 416, and 380 cases of PE per 100,000 pregnancy women screened in the frist trimester compared with no prediction, the baseline, and MAP prediction. An extensive sensitivity analysis was conducted to prove that the base-case results were robust to a broad range of parameter values. The one-way and probability sensitivity analysis showed the PROMPT model was cost-effective to other strategies, which were robust to randomly distributed parameters (Supplementary Fig. 3). The cost-effectiveness acceptability curve diagrams showed that PROMPT model was the best strategy, accounting for 96.3% of the simulations in total situations (Supplementary Fig. 4).

Discussion

This work provides evidence that the retinal vascular features are potential biomarkers for PE prediction and AO detection. These retinal biomarkers have several desirable properties. Compared with recent biomarkers reported, retinal photography is non-invasive, convenient, and accessible, especially in rural areas where quality resources are lacking. Besides, the PROMPT model is proved dominant in the cost-effective analysis, averting more PE cases while saving total societal costs, which presents an opportunity to promote equity of pregnancy health in LMICs from the societal perspective.

In this study, the metadata variables (MAP, overweight, renal disease history, and PE history), and the retinal parameters (AVR, tortD, etc.) included in the model were consistent with literature and clinical experience[6,27]. In the routine clinical practice, patients diagnosed with PE were referred to retinal examinations for complications like retinal hemorrhage, exudation, and detachment. However, the status quo focuses on ophthalmic manifestations after PE diagnosis and the quantification parameters are lacking. There was limited study on early ophthalmic characteristics prior to PE incidence and related pregnancy outcomes. Taking the advantage of deep learning algorithm, fully automate retinal microvasculature segmentation and quantification were achieved and we were able to utilize numeric retinal parameters for PE risk assessment[29]. Our study demonstrated that the retinal vascular alterations were found in the first trimester of pregnancy, much prior to the PE diagnosis, indicating the potential to be early and sensitive biomarkers. In the PROMPT model, five retinal vascular parameters (AVR, VSD, tortD, angtort, and angtort_std) were selected as variables by LASSO. AVR has been demonstrated to exhibit significant correlation with hypertension in prior studies[30–32]. VSD serves as

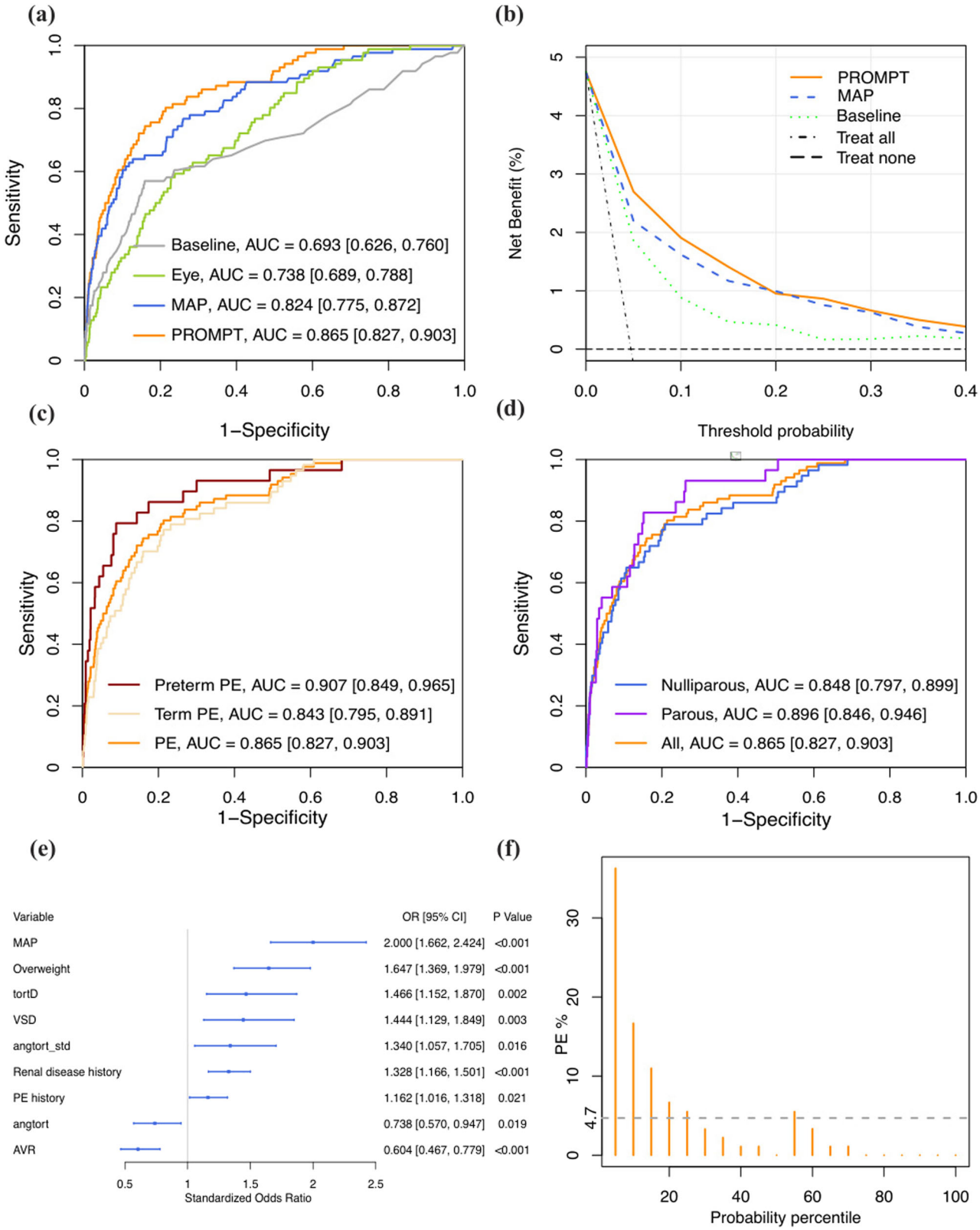

Fig. 2 | The PE prediction models, subgroup analysis, variable importance, and nomogram. a ROC curves for performance of PE prediction based on baseline, eye, MAP and PROMPT models. **b** The decision curve of the PROMPT model compared with the baseline model, MAP model, no treatment, and all treatment. **c** The prediction performance of preterm and term PE based on PROMPT model. **d** Subgroup analysis of PROMPT model in nulliparous and parous groups. **e** Variable importances in the PROMPT model. **f** The lift plot of the PROMPT model compared with MAP model. ROC curve Receiver-operating-characteristics curve, AUC area under the ROC, PE preeclampsia, MAP mean arterial pressure, TortD tortuosity density, VSD vessel skeleton density, Angtort angle-based tortuosity, AVR artery-to-vein ratio, Std standard deviation.

a sensitive biomarker for detecting capillary-level perfusion alterations[33]. For tortD and angtort, they collectively reflect pathological microvascular remodeling processes and/or ischemia-induced arterialization patterns in collateral micro vessels[30,34]. Angtort_std quantifies the temporal variability of angtort, representing dynamic vascular parameter fluctuations[34]. The clinical relevance of the retinal biomarkers included in the PROMPT model is consistent with the mechanisms underlying the pathogenesis of PE, which involve maternal generalized endothelial dysfunction and systemic vascular abnormalities[35]. These retinal biomarkers provide a non-invasive means of detecting early systemic vascular changes, potentially even before the clinical onset of PE.

The PROMPT model has comparable discrimination to the triple test for predicting PE and preterm PE. A large prospective cohort study in China reported an AUC of 0.86 and DR of 65.0% at 10% FPR for preterm PE using triple test. Recent ML-based study has achieved the DR of 52.9% at 10% FPR, with an AUC of 0.82 for PE prediction[36]. Our model could achieve AUC of 0.86 and 0.91 in PE and preterm prediction, and DR of 69.5% and 79.3% at 10% FPR for PE and preterm. Additionally, the integration of all biomarkers of the triple test in first-trimester screening still faces a series of clinical resources challenges. For instance, UtA-PI requires ultrasound examination performed by qualified sonographers, and PAPP-A and PlGF measurements involve costly blood sampling, which limits their widespread adoption in routine clinical practice, particularly in resource-constrained settings. A cross-sectional study included maternal features from 726 women and achieved a maximum accuracy of 0.79 using SVM[37]. Li et al. conducted a retrospective cohort using 39 parameters from electronic health records before 20 weeks of gestation, achieving an AUC of 0.955, with a shorter time window for intervention[38]. Another similar retrospective study reported an AUC of 0.79, albeit with a significant amount of missing data[39]. Liu et al. combined maternal characteristics, blood tests, and sonication parameters in a cohort of 1115 pregnant women, achieving a maximum AUC of 0.86 using RF methods[40]. These studies highlight the potential of ML in improving PE prediction, but also underscore the challenges associated with data quality, feature availability, and model generalizability.

Recent studies have also explored the use of cfDNA for PE prediction. A prospective cohort study evaluated the performance of artificial neural network models using maternal characteristics and cfDNA data in 17,520 cases, achieving an AUC of 0.713 and a DR of 59.3% at FPR of 15%[41]. Adil et al. investigated the utility of prenatal cfDNA screening using whole-genome sequencing data before 16 weeks of gestation in 1854 samples. The ML model achieved an AUC of 0.85 in the primary cohort and 0.84 in an external validation[42]. Although these results are encouraging, the high costs and limited accessibility of cfDNA testing currently limit its applicability in clinical practice. In contrast, the PROMPT model leverages retinal photography, a non-invasive, low-cost, and rapid examination that does not require experienced clinicians and laboratory test resources. Retinal imaging is increasingly accessible, particularly in rural areas due to the nationwide rollout of eye disease screening programs. This widespread availability positions retinal parameters as potential "point-of-care" biomarkers for systemic diseases screening (Supplementary Table 3).

In the cost-effectiveness analysis, the PROMPT model was proved dominant over no prediction, baseline, and MAP prediction. The superiority is mainly attributed to better performance of the PROMPT model, which avoids unnecessary monitoring costs and saves more PE-related medical expenses, thereby potentially compensating for the relatively low costs of retinal examinations. In the analysis, the PROMPT method emerges as the most cost-effective, holding the potential for widespread implementation in large population screening settings. It can help to avert a greater number of AO and result in significant savings in societal costs. Previous study showed that the ophthalmic artery doppler in the second or third trimester could improve PE prediction, providing initial evidence that eye can serve as a window to incident PE[43,44]. Compared with ophthalmic

Table 2 | Net reclassification index and integrated discrimination improvement results

	Number	Continuous NRI	IDI
PROMPT Vs. Baseline model			
PE	86	0.395 (0.205–0.600)	
HC	1726	0.430 (0.387–0.474)	
All	1812	0.825 (0.628–1.036)	0.093 (0.050-0.135)
PROMPT Vs. MAP model			
PE	86	0.279 (0.078–0.482)	
HC	1726	0.321 (0.279–0.367)	
All	1812	0.600 (0.402–0.812)	0.031 (0.007-0.055)

Table of NRI and IDI of PROMPT model Vs. baseline model and PROMPT model Vs. MAP model.
NRI net reclassification improvement, *IDI* integrated discrimination improvement.
PE preeclampsia, *HC* healthy controls, *MAP* mean arterial pressure.

Table 3 | Detection rate of adverse maternal and neonatal outcomes of the baseline, MAP and PROMPT model

Screening model/ FPR for PE	PE *n* = 86		PT PE *n* = 29	AO *n* = 393		Severe AO *n* = 98	
	DR%	PPV%	DR%	DR%	PPV%	DR%	PPV%
Baseline model							
5%	27.9 (19.5, 38.2)	21.2 (14.2, 29.7)	31.0 (17.3, 49.2)	9.7 (7.1, 13.0)	33.6 (25.6, 42.7)	13.3 (7.9, 21.4)	11.5 (6.8, 18.7)
10%	39.5 (29.9, 50.1)	16.6 (12.1, 22.3)	44.8 (28.4, 62.5)	16.5 (13.2, 20.5)	31.7 (25.7, 38.4)	21.4 (14.5, 30.5)	10.2 (6.8, 15.2)
15%	52.3 (41.9, 62.6)	15.0 (11.4, 19.4)	69.0 (50.8, 82.7)	23.9 (20.0, 28.4)	31.2 (26.3, 36.7)	34.7 (26.0, 44.5)	11.3 (8.2, 15.4)
MAP model							
5%	41.9 (32, 52.4)	29.5 (22.1, 38.1)	55.2 (37.5, 71.6)	13.5 (10.5, 17.2)	43.4 (35, 52.3)	22.4 (15.3, 31.7)	18.0 (12.2, 25.8)
10%	59.3 (48.7, 69.1)	22.9 (17.8, 28.8)	69.0 (50.8, 82.7)	22.4 (18.5, 26.8)	39.5 (33.3, 46)	31.6 (23.3, 41.4)	13.9 (10, 19.1)
15%	64 (53.4, 73.3)	17.7 (13.8, 22.3)	72.4 (54.3, 85.3)	28.2 (24, 32.9)	35.7 (30.6, 41.2)	37.8 (28.8, 47.6)	11.9 (8.8, 16)
PROMPT model							
5%	47.7 (37.4, 58.1)	32.3 (24.8, 40.8)	62.1 (44, 77.3)	16.8 (13.4, 20.8)	52 (43.3, 60.5)	27.6 (19.7, 37.1)	21.3 (15, 29.2)
10%	60.5 (49.9, 70.1)	23.1 (18.1, 29)	79.3 (61.6, 90.2)	23.4 (19.5, 27.8)	40.9 (34.7, 47.4)	35.7 (26.9, 45.6)	15.6 (11.4, 20.9)
15%	72.1 (61.8, 80.5)	19.3 (15.4, 24)	82.8 (65.5, 92.4)	31.0 (26.7, 35.8)	38.0 (32.9, 43.4)	40.8 (31.6, 50.7)	12.5 (9.3, 16.5)

Values in parentheses are 95% CI. AO was defined as any one or more of the following maternal and neonatal outcomes: PE, SGA (birth weight < 10th percentile), and PTB (< 37 weeks). Severe AO was defined as any one or more of the following maternal and neonatal outcomes: PT PE, severe SGA (birth weight < 3rd percentile), and early PTB (< 34 weeks).
FPR false positive rate, *PE* preeclampsia, *PT-PE* preterm preeclampsia, *AO* adverse pregnancy outcomes, *DR* detection rate, *PPV* positive predictive value, *MAP* Mean arterial pressure, *PROMPT* Preeclampsia Risk factors + Ophthalmic data + Mean arterial pressure Prediction Test.

Table 4 | Cost-effectiveness analysis of no prediction, baseline, MAP and PROMPT prediction

	Comparison for ICER calculation	Costs per person, $	Incremental costs per person screened, $	PE cases	PE cases avoided	ICERs
No prediction	\	2142.63	\	4700	\	\
Baseline prediction	\	2104.60	\	3307	\	\
MAP prediction	\	2096.52	\	3111	\	\
PROMPT prediction	No prediction	2091.29	−51.34	2891	1809	Dominant
	Baseline prediction	\	−13.31	\	416	Dominant
	MAP prediction	\	−5.23	\	380	Dominant

Costs are expressed in US dollars. Incremental costs, ICERs, and PE cases avoided are defined as values per 100,000 people. *ICER* incremental cost effectiveness ratio, *PE* preeclampsia, *MAP* mean arterial pressure.

artery doppler, retinal photography is cheaper, more time-efficient, and accessible in primary care settings, and the PROMPT method can be applied in the first trimester, hence creating longer time frame for PE intervention.

Our study has the following strengths. We prospectively collected ocular data in the first trimester of pregnancy, allowing us to observe the associations between retinal alterations before the onset of PE. In this study, we focused on parameters collected from 11-13^{+6} weeks of gestation, as this represents the earliest feasible timeframe to assess pregnant women's risk for PE. This time window aligns with the timing of most pregnancies' first prenatal visits and is consistent with both domestic guidelines and international screening strategies[27,45]. Early first-trimester screening also allows for timely prophylactic interventions, such as blood pressure monitoring and aspirin prescription. Then, we utilized DL-based system for retinal vessel segmentation and qualification, thus extracting the early and subtle vascular features before PE. Next, we employed ML algorithms to construct a model with good performance of PE and preterm PE prediction, in which only three types of data and nine variables were selected, making it easy to apply in clinical practice. Hence, it provides a non-invasive, accessible, and cost-effective way for prenatal care. Our study is subject to certain limitations, notably the need to increase both the sample size and data diversity. Due to low prevenance of PE, prospective cohort construction, and follow-up periods for pregnancy outcomes, the data acquisition faces a series of challenges. Considering the relatively small number of PE cases, we employed 10-fold CV to test the robustness of the model instead of a separate test set. Future study in other centers is needed to verify the robustness of the PROMPT model.

Finally, our work shows that advances in AI can provide new clinical insights by addressing important unsolved challenges and allowing for the development of new biomarkers. By quantifying the early retinal vascular characteristics using DL-based system, this study provides initial evidence of the retinal vascular biomarkers to improve PE prediction and AO detection. The PROMPT model, including four clinical variables and five ophthalmic parameters from one single retinal photograph, is a non-invasive, cost-effective, and convenient method for pregnancy health management in low-resource settings. The large-scale real-world evaluation of PROMPT model warrants further investigation before clinical implementation.

Methods

Study design and participants

This was a prospective cohort study conducted at the Department of Obstetrics in the First Affiliated Hospital (FAH), Sun Yat-sen University from August 2021 to November 2023. The inclusion criteria for the study were as follows: age ≥ 18 years old, singleton pregnancy and 11-13^{+6} weeks of gestation at the booking visit. The exclusion criteria were a history of ocular surgery history, cataract, severe opacity of refractive media, severe systemic illness, or significant mental illness or those for whom a major fetal abnormality was identified at the time of recruitment. All participants provided written informed consent. The study was conducted in accordance with the Declaration of Helsinki and was approved by the Institutional Review Board/Ethics Committee of Zhongshan Ophthalmic Center, Sun Yat-sen University (2021KYPJ098-2) and the Independent Ethics Committee for Clinical Research and Animal Trials of the First Affiliated Hospital of Sun Yat-sen University (2021-649). It was prospectively registered at the Chinese Clinical Trial Registry (ChiCTR2100049850) on Aug 10th, 2021.

Data acquisition and follow-up

Data of maternal characteristics and pregnancy outcomes were collected from the hospital maternity records. All participants underwent color fundus photography (CFP) at the same visit using Retinal Fundus Camera Reti-Cam3100 (SYSEYE, Chongqing, China). Only data from participants' oculus dexter (OD) were included. The data from Oculus Sinister (OS) was analyzed if the data quality of OD is poor. The detailed diagnosis criteria of other PE and other pregnancy outcomes were displayed in Supplementary Table 4.

Processing of retinal fundus parameters

We used a deep learning-based system (Retina-based Microvascular Health Assessment System, RMHAS) for segmentation and quantification of the retinal microvasculature[29]. Developed using a diverse dataset and a multi-branch U-Net architecture, the system could fully automatedly segment arteries, veins, and the optic disc and addresses the limitations of manual tools, providing a broad range of physical and geometric measurements. The image quality control module first assessed overall image quality, and the segmentation module generated artery, vein, and optic disc segmentation maps. Then the measurement module computed region-specific measurements including AVR, VSD, tortD, angtort, etc. The definitions and clinical significances of the extracted parameters were displayed in the Supplementary Table 5.

Algorithm selection

To forecast PE occurrence with clinical risk factors and retinal parameters, we tested five ML algorithms including LR, RF, XGBoost, SVM, and GBM. First, both clinical features and retinal fundus parameters were selected via LASSO with 10-fold CV. Next, grid search with automated 10-fold CV was used to tune the hyperparameters for each of the last four algorithms. Due to the low rate of PE, upsampling method was adopted to obtain balanced data. Then, a manual 10-fold CV was conducted to train and assess the five models. The entire dataset was randomly split into 10 subsets of nearly equal size, ensuring that the relatively even distribution of PE cases. Except for LR, the optimal hyperparameters and upsampling method were adopted for the other four algorithms. Specifically, at each iteration of 10, a training dataset including 9 subsets was upsampled by sampling randomly from positive cases with replacement, and then models were trained on the balanced training dataset with the optimal parameters, while the remaining subset was a validation dataset. The performance on the training dataset and validation dataset was calculated and recorded. Model performance was measured using AUC.

Model development and validation

Four models for PE prediction were constructed including the baseline model, eye model, MAP model, and PROMPT model. Except for the baseline model, LASSO with 10-fold CV was adopted to select features. Since non-significant features weren't helpful for LR model performance

and to ensure parsimony of the final models, both-direction stepwise was utilized to select features further. Then, the LR models with selected features were trained and assessed on the same 10-fold CV used in the algorithm selection. The final models with selected features were trained on the entire dataset for model interpretation and prediction on unseen datasets or future use. The PROMPT model was evaluated to predict preterm and term PE, and the subgroup analyses were performed in parous and nulliparous groups, respectively.

Detection of placenta-associated AO in screen-positive pregnancies

The DR for other placenta-associated AO were analyzed at fixed FPR of 5%, 10% and 15% for PE prediction. The AO included any one or more of: PE, SGA (< 10^{th} percentile), and preterm birth (< 37 weeks, PTB). The severe AO included any one or more of preterm PE (< 37 weeks), severe SGA (< 3rd percentile), and early PTB (< 34 weeks).

Cost-effectiveness analysis of different prediction models

Decision tree models were constructed for the economic evaluation of different prediction strategies, which were built on a simulated cohort of 100,000 pregnant subjects using no prediction, baseline model, MAP model, and PROMPT model, respectively (Supplementary Fig. 5). The data were collected from this study, real-world clinical routine, and a literature search of PE prevalence, PE prevalence reduction after using acetylsalicylic acid (ASA), side effect prevalence, etc[1,46]. The costs of screening, monitoring, medication, and birth were collected from the FAH (Supplementary Table 6). All costs were expressed in US dollars at the exchange rate as of 15 October 2024 (1 USD = 7.1 CNY). The primary outcome was incremental cost effectiveness ratio (ICER) calculating using the Eq. (1):

$$ICER = \frac{Incremental\ costs}{PE\ cases\ avoided} \quad (1)$$

We performed one-way and probabilistic sensitivity analyses to assess the robustness of the main outcomes. Fluctuation ranges of 10% (probability data including prevalence, sensitivity, specificity, etc.), 20% (costs of monitoring, medication and birth, etc.), and 50% (screening costs) were set for sensitivity analysis. Tornado diagrams showed the parameters that had the greatest influence on ICER, and the probabilistic sensitivity analysis evaluated the impact on the results by taking 10,000 random samples from the probability distribution of each parameter.

Statistical analysis

For sample size calculation, assuming the PE prevalence was 5% and nine variables would be selected to construct a logistic regression model, 1800 samples are needed for developing the model based on the rule of thumb "10 events per variable". Continuous features with normal distribution are presented as mean ± standard deviation and were compared with the *t*-test. Otherwise, they are presented as medians with interquartile ranges and compared with the Mann-Whitney *U* tests. Categorical variables are reported as frequency and percentages and compared with the chi-square test or Fisher's exact test. The discrimination of the models was assessed using the receiver operating characteristic (ROC) curves and AUC with 95% Delong confidence interval (CI). The paired DeLong test, continuous NRI and IDI were employed to evaluate the improvement in predicting PE by adding retinal vascular parameters. A brief explanation of the statistic metrics was displayed in Supplementary Table 7. Decision curve analysis was used to show the net benefit of using the models at different risk thresholds to assess the utility for decision making. For the PROMPT model, the lift plot was also utilized to demonstrate the discriminative performance, and forest plot and nomogram were adopted to improve the interpretation of the models. A two-tailed $p < 0.05$ was considered statistically significant. All models were developed and assessed using R version 4.1.1 (R Foundation for Statistical Computing, Vienna; Austria) and the cost-effectiveness analysis were conducted using Python version 3.10.14. The checklist of the Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) Statement was listed in Supplementary Table 8.

Data availability

Data can only be shared for noncommercial academic purposes and will require a formal data use agreement. Please email all requests for academic use of data to corresponding author at linht5@mail.sysu.edu.cn. For requests from verified academic researchers, access will be evaluated by the data access committee and be granted within 1 month. H.L. has directly accessed and verified the underlying data reported in the manuscript.

Code availability

The underlying code for this study is not publicly available but may be made available to qualified researchers on reasonable request from the corresponding author at linht5@mail.sysu.edu.cn and wxhang@mail2.sysu.edu.cn. For requests from verified academic researchers, access will be evaluated by the data access committee and be granted within one month.

Received: 7 December 2024; Accepted: 24 March 2025;
Published online: 05 April 2025

References

1. Dimitriadis, E. et al. Pre-eclampsia. *Nat. Rev. Dis. Prim.* **9**, 8 (2023).
2. Brown, M. A. et al. Hypertensive Disorders of Pregnancy: ISSHP Classification, Diagnosis, and Management Recommendations for International Practice. *Hypertens. Dallas Tex. 1979* **72**, 24–43 (2018).
3. Say, L. et al. Global causes of maternal death: a WHO systematic analysis. *Lancet Glob. Health* **2**, e323–e333 (2014).
4. Kassebaum, N. J. et al. Global, regional, and national levels and causes of maternal mortality during 1990–2013: a systematic analysis for the Global Burden of Disease Study 2013. *Lancet* **384**, 980–1004 (2014).
5. Tan, M. Y. et al. Comparison of diagnostic accuracy of early screening for pre-eclampsia by NICE guidelines and a method combining maternal factors and biomarkers: results of SPREE. *Ultrasound Obstet. Gynecol.* **51**, 743–750 (2018).
6. Yang, Y. et al. Preeclampsia Prevalence, Risk Factors, and Pregnancy Outcomes in Sweden and China. *JAMA Netw. Open* **4**, e218401 (2021).
7. Chaemsaithong, P. et al. Prospective evaluation of screening performance of first-trimester prediction models for preterm preeclampsia in an Asian population. *Am. J. Obstet. Gynecol.* **221**, 650.e1–650.e16 (2019).
8. Akolekar, R., Syngelaki, A., Poon, L., Wright, D. & Nicolaides, K. H. Competing risks model in early screening for preeclampsia by biophysical and biochemical markers. *Fetal Diagn. Ther.* **33**, 8–15 (2013).
9. Hu, J. et al. Prospective evaluation of first-trimester screening strategy for preterm pre-eclampsia and its clinical applicability in China. *Ultrasound Obstet. Gynecol.* **58**, 529–539 (2021).
10. Rolnik, D. L. et al. Aspirin versus Placebo in Pregnancies at High Risk for Preterm Preeclampsia. *N. Engl. J. Med.* **377**, 613–622 (2017).
11. Poon, L. C. et al. The International Federation of Gynecology and Obstetrics (FIGO) initiative on pre-eclampsia: A pragmatic guide for first-trimester screening and prevention. *Int. J. Gynecol. Obstet.* **145**, 1–33 (2019).
12. De Borre, M. et al. Cell-free DNA methylome analysis for early preeclampsia prediction. *Nat. Med.* **29**, 2206–2215 (2023).
13. Moufarrej, M. N. et al. Early prediction of preeclampsia in pregnancy with cell-free RNA. *Nature* **602**, 689–694 (2022).
14. Rasmussen, M. et al. RNA profiles reveal signatures of future health and disease in pregnancy. *Nature* **601**, 422–427 (2022).
15. Wang, S. B. et al. A spectrum of retinal vasculature measures and coronary artery disease. *Atherosclerosis* **268**, 215–224 (2018).

16. Iao, W. C. et al. Deep Learning Algorithms for Screening and Diagnosis of Systemic Diseases Based on Ophthalmic Manifestations: A Systematic Review. *Diagn. Basel Switz.* **13**, 900 (2023).
17. Wu, Y. et al. Optical coherence tomography angiography for the characterisation of retinal microvasculature alterations in pregnant patients with anaemia: a nested case–control study. *Br. J. Ophthalmol.* **108**, 117–123 (2024).
18. Zhao, X. et al. Screening chronic kidney disease through deep learning utilizing ultra-wide-field fundus images. *Npj Digit. Med.* **7**, 275 (2024).
19. Chang, J. et al. Association of Cardiovascular Mortality and Deep Learning-Funduscopic Atherosclerosis Score derived from Retinal Fundus Images. *Am. J. Ophthalmol.* **217**, 121–130 (2020).
20. Cheung, C. Y. et al. A deep learning model for detection of Alzheimer's disease based on retinal photographs: a retrospective, multicentre case-control study. *Lancet Digit. Health* **4**, e806–e815 (2022).
21. Zhang, K. et al. Deep-learning models for the detection and incidence prediction of chronic kidney disease and type 2 diabetes from retinal fundus images. *Nat. Biomed. Eng.* **5**, 533–545 (2021).
22. Cheung, C. Y. et al. A deep-learning system for the assessment of cardiovascular disease risk via the measurement of retinal-vessel calibre. *Nat. Biomed. Eng.* **5**, 498–508 (2020).
23. Xiao, W. et al. Screening and identifying hepatobiliary diseases through deep learning using ocular images: a prospective, multicentre study. *Lancet Digit. Health* **3**, e88–e97 (2021).
24. Lee, C. S., Choi, E. Y., Lee, M., Kim, H. & Chung, H. Serous retinal detachment in preeclampsia and malignant hypertension. *Eye Lond. Engl.* **33**, 1707–1714 (2019).
25. Gilbert, A. L., Prasad, S. & Mallery, R. M. Neuro-Ophthalmic Disorders in Pregnancy. *Neurol. Clin.* **37**, 85–102 (2019).
26. Digre, K. B. Neuro-ophthalmology and pregnancy: what does a neuro-ophthalmologist need to know?. *J. North Am. Neuro-Ophthalmol. Soc.* **31**, 381–387 (2011).
27. Hypertensive Disorders in Pregnancy Subgroup, Chinese Society of Obstetrics and Gynecology, Chinese Medical Association. Diagnosis and treatment of hypertension and pre-eclampsia in pregnancy: a clinical practice guideline in China. *Zhonghua Fu Chan Ke Za Zhi* **55**, 227–238 (2020).
28. Magee, L. A. et al. The 2021 International Society for the Study of Hypertension in Pregnancy classification, diagnosis & management recommendations for international practice. *Pregnancy Hypertens.* **27**, 148–169 (2022).
29. Shi, D. et al. A Deep Learning System for Fully Automated Retinal Vessel Measurement in High Throughput Image Analysis. *Front. Cardiovasc. Med.* **9**, 823436 (2022).
30. Tapp, R. J. et al. Associations of Retinal Microvascular Diameters and Tortuosity With Blood Pressure and Arterial Stiffness. *Hypertens. Dallas Tex. 1979* **74**, 1383–1390 (2019).
31. Yan, S. et al. Deep Learning based Retinal Vessel Caliber Measurement and the Association with Hypertension. *Curr. Eye Res.* **49**, 639–649 (2024).
32. Ponto, K. A. et al. Retinal vessel metrics: normative data and their use in systemic hypertension: results from the Gutenberg Health Study. *J. Hypertens.* **35**, 1635–1645 (2017).
33. Chu, Z. et al. Quantitative assessment of the retinal microvasculature using optical coherence tomography angiography. *J. Biomed. Opt.* **21**, 66008 (2016).
34. Corliss, B. A., Mathews, C., Doty, R., Rohde, G. & Peirce, S. M. Methods to label, image, and analyze the complex structural architectures of microvascular networks. *Microcirc. N. Y. N. 1994* **26**, e12520 (2019).
35. Staff, A. C. The two-stage placental model of preeclampsia: An update. *J. Reprod. Immunol.* **134–135**, 1–10 (2019).
36. Ansbacher-Feldman, Z. et al. Machine-learning-based prediction of pre-eclampsia using first-trimester maternal characteristics and biomarkers. *Ultrasound Obstet. Gynecol.* **60**, 739–745 (2022).
37. Manoochehri, Z., Manoochehri, S., Soltani, F., Tapak, L. & Sadeghifar, M. Predicting preeclampsia and related risk factors using data mining approaches: A cross-sectional study. *Int. J. Reprod. Biomed.* **19**, 959–968 (2021).
38. Li, Y. et al. Novel electronic health records applied for prediction of pre-eclampsia: Machine-learning algorithms. *Pregnancy Hypertens.* **26**, 102–109 (2021).
39. Marić, I. et al. Early prediction of preeclampsia via machine learning. *Am. J. Obstet. Gynecol. MFM* **2**, (2020).
40. Melinte-Popescu, A. -S., Vasilache, I. -A., Socolov, D. & Melinte-Popescu, M. Predictive Performance of Machine Learning-Based Methods for the Prediction of Preeclampsia—A Prospective Study. *J. Clin. Med.* **12**, 418 (2023).
41. Khalil, A. et al. The role of cell-free DNA biomarkers and patient data in the early prediction of preeclampsia: an artificial intelligence model. *Am. J. Obstet. Gynecol.* **231**, 554.e1–554.e18 (2024).
42. Adil, M. et al. Preeclampsia risk prediction from prenatal cell-free DNA screening. *Nat. Med.* https://doi.org/10.1038/s41591-025-03509-w (2025).
43. Sarno, M. et al. Ophthalmic artery Doppler in prediction of pre-eclampsia at 35–37 weeks' gestation. *Ultrasound Obstet. Gynecol.* **56**, 717–724 (2020).
44. Sapantzoglou, I. et al. Ophthalmic artery Doppler in combination with other biomarkers in prediction of pre-eclampsia at 19–23 weeks' gestation. *Ultrasound Obstet. Gynecol.* **57**, 75–83 (2021).
45. American College of Obstetricians and & Gynecologists' Committee on Practice Bulletins Gestational Hypertension and Preeclampsia: ACOG Practice Bulletin. *Obstet. Gynecol.* **135**, e237–e260 (2020).
46. Rolnik, D. L., Nicolaides, K. H. & Poon, L. C. Prevention of preeclampsia with aspirin. *Am. J. Obstet. Gynecol.* **226**, S1108–S1119 (2022).

Acknowledgements

Funding was provided by the National Natural Science Foundation of China (92368205 and 82171035), Guangdong Provincial Natural Science Foundation for Progressive Young Scholars (2023A1515030170), Guangzhou Clinical Major Technology Project (2024P-ZD12), Sun Yat-Sen University Clinical Research 5010 Program (2022003), the National Key Research and Development Program of China (2021YFC2700700), and the Science and Technology Planning Projects of Guangdong Province (2018B010109008). The funders played no role in study design, data collection, analysis, and interpretation of data, or the writing of this manuscript.

Author contributions

Conception and design: Y.W., L.S., Z.W., X.W., D.W., and H.L.; Data collection: X.L., Y.H., L.K., L.L., Z.C., W.H., Y.H., S.C., M.D.Y., X.Z., and D.W.; Analysis and interpretation: L.Z., Z.T., Z.L., and D.S.; Figures L.Z., M.X., Z.T., and S.Z.; Manuscript draft and revisions: Y.W., L.S., D.L., X.W., X.C., XW, DW, and HL. Funding acquisition: Z.W., X.W., and H.L.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41746-025-01582-6.

Correspondence and requests for materials should be addressed to Zilian Wang, Xiaohang Wu, Dongyu Wang or Haotian Lin.