A blood-based miRNA signature for early non-invasive diagnosis of preeclampsia

Siqi Bao1,2†, Tong Zhou1†, Congcong Yan2, Jiale Bao1, Fan Yang3, Shan Chao3, Meng Zhou2* and Zhangye Xu1*

Abstract

Background: Preeclampsia (PE) is a multisystemic maternal syndrome with substantial maternal and fetal morbidity and mortality. Currently, there is no clinically viable non-invasive biomarker assay for early detection, thus limiting the effective prevention and therapeutic strategies for PE.

Methods: We conducted a discovery–training–validation three-phase retrospective and prospective study with cross-platform and multicenter cohorts. The initial biomarkers were discovered and verified in tissue specimens by small RNA sequencing and qRT-PCR. A miRNA signature (miR2PE-score) was developed using Firth's bias-reduced logistic regression analysis and subsequently validated in two independent multinational retrospective cohorts and two prospective plasma cohorts.

Results: We initially identified five PE-associated differentially expressed miRNAs from miRNA sequencing data and subsequently validated two miRNAs (miR-196b-5p and miR-584-5p) as robust biomarkers by association analysis with clinical characteristics and qRT-PCR in tissue specimens in the discovery phase. Using Firth's bias-reduced logistic regression analysis, we developed the miR2PE-score for the early detection of PE. The miR2PE-score showed a high diagnostic performance with an area under the receiver operating characteristic curve (AUROC) of 0.920, 0.848, 0.864, and 0.812 in training, internal, and two external validation cross-platform and multicenter cohorts, respectively. Finally, we demonstrated the non-invasive diagnostic performance of the miR2PE-score in two prospective plasma cohorts with AUROC of 0.933 and 0.787. Furthermore, the miR2PE-score revealed superior performance in non-invasive diagnosis compared with previously published miRNA biomarkers.

Conclusions: We developed and validated a novel and robust blood-based miRNA signature, which may serve as a promising clinically applicable non-invasive tool for the early detection of PE.

Keywords: Biomarker, Early diagnosis, Liquid biopsy, MicroRNA, Non-invasive, Preeclampsia

†The authors wish it to be known that, in their opinion, Siqi Bao and Tong Zhou should be regarded as joint first authors.

*Correspondence: zhoumeng@wmu.edu.cn; xuzhangye@aliyun.com

1 Department of Gynecology and Obstetrics, The Second Affiliated Hospital and Yuying Children's Hospital of Wenzhou Medical University, Wenzhou 325027, People's Republic of China
2 School of Biomedical Engineering, Wenzhou Medical University, Wenzhou 325027, People's Republic of China

Full list of author information is available at the end of the article
Background

Preeclampsia (PE) is a multisystemic disease characterized by hypertension [1], endothelial cell dysfunction [2], poor extravillous cytotrophoblast invasion [3], and proteinuria and fetal growth retardation [4] after 20 weeks of gestation. This multisystem disorder affects approximately 5–8% of pregnancies [5, 6] and is one of the most severe pregnancy complications that pose a risk to both mother and baby. Each year, 70,000 women and 50,000 babies approximately die from this disorder worldwide [7].

Once PE occurs, the treatment available is limited to symptomatic treatment and the PE progression will not be effectively prevented, leading to severe maternal and fetal complications. Therefore, predicting PE in the early stage of disease, before symptoms present, is warranted to prevent or reduce the frequency and severity of PE. However, early diagnosis and effective treatment of PE remain a challenge as diagnostic criteria for PE are non-specific signs and symptoms, and its severity criteria are poorly correlated with adverse maternal and fetal outcomes. Over the past few years, there have been increasing efforts to improve the early diagnosis and treatment of PE. Angiogenesis-related factors, such as soluble fms-like tyrosine kinase-1 (sFlt-1), placental growth factor (PIGF), and soluble endoglin (sEng), could be used to predict PE [8–10]. However, diabetes mellitus, parity, twins, and advanced maternal age influenced the definition of normal parameters for these markers leading to higher variation in cutoff values and predictive efficacy across different ethnic and geographical populations [11–13]. However, despite these efforts, none of these biomarkers is adequate and satisfactory in managing PE. Therefore, there is an urgent need to identify clinically viable biomarkers and tools for the early diagnosis and personalized intervention of PE.

MiRNAs are single-stranded (~22 nucleotides), non-coding RNAs responsible for the mechanism of post-transcriptional gene expression regulation [14–17]. MiRNAs are characterized by a long half-life and high stability, and their stability is 10 times higher than that of mRNAs [18]. During pregnancy, placental trophoblast cells at the maternal–fetal interface produce a large number of miRNAs, and its expression level changes with the increase of gestational age and placental development, which highlights its importance in placental regulation. MiRNAs are highly stable in serum, plasma, and urine, making miRNA possible as non-invasive biomarkers for detecting and diagnosing PE [19]. Increasing evidence indicated the potential roles of miRNAs in the pathogenesis and treatment of PE [20, 21]. For example, a recent study has shown that miR-27a is overexpressed in PE placenta, and inhibition of miR-27a may be a probable treatment for PE [22]. Several miRNAs, such as Let-7a,
miR-125b, miR-133b, miR-384, miR-101, miR-206, and miR-203a-3p, have been recently identified as candidates that may be involved in the progression or suppression of PE [22–26]. However, the predictive value and clinical translation potential of miRNA for early diagnosis of PE remain uncertain and need further investigation.

In this study, we conducted a discovery–training–validation three-phase multicenter study with cross-platform retrospective and prospective cohorts to identify and validate a blood-based miRNA signature for early non-invasive diagnosis of PE.

Methods

Ethics statement
This study was approved by the human ethics committee at the Second Affiliated Hospital and Yuying Children’s Hospital of Wenzhou Medical University (approval number: LCKY2020-242), and informed consent was obtained from all patients and their families. And the study was performed in accordance with the regulations and guidelines established by this committee.

Study design and patient cohorts
This study was carried out in three phases: discovery, training, and validation. The details of the overall study are shown in Fig. 1. In the discovery phase, we conducted miRNA sequencing in a retrospective cohort (WMU cohort 1) for genome-wide screening of candidate miRNA biomarkers and then verified the expression levels of the candidate miRNA biomarkers in 20 PE patients and 20 healthy controls from WMU cohort 2 using qRT-PCR for identifying robust miRNA biomarkers. In the training phase, a miRNA signature for early diagnosis of PE (miR2PE-score) was developed in WMU cohort 1. In the validation phase, the performance of miR2PE-score was firstly evaluated in two independent multinational cohorts (Canada cohort and USA cohort) from the Gene Expression Omnibus (GEO) database (https://www.ncbi.nlm.nih.gov/geo).
Sample collection and preparation
We collected the samples at the Department of Obstetrics, the Second Affiliated Hospital and Yuying Children's Hospital of Wenzhou Medical University (Zhejiang, China) between June 2020 and December 2021. Blood samples were obtained from pre-cesarean women with overnight (≈10 h) fasting in 10-ml EDTA-coated Vacutainer tubes before any medical interventions. Five milliliters of blood for obtaining plasma was separated by centrifugation at 3000 rpm for 15 min at 20°C within 30 min after each collection, and then the supernatants were extracted to a new tube.

Leukocytes were purified from another 5 ml blood using red blood cell lysis solution (Solarbio, China) following the manufacturer's protocol. In brief, 5 ml blood was incubated with a threefold volume of red cell lysate on ice for 15 min and mixed evenly. The mixture was centrifuged at 450xg for 10 min at 4°C, the supernatant was discarded, and the leukocyte precipitate was obtained. The experiment described above was repeated three times. Finally, red cells were dissolved and leukocytes were washed with PBS.

Placental samples were rapidly processed as previously described [29]. In brief, the tissue was sectioned transversally (~5 cm) near the cord insertion site. Both the decidual layer along with the basal plate and the chorionic surface and membranes were removed by dissection. Each sample was rinsed thoroughly in cold physiological saline solution and frozen in liquid nitrogen.

All samples were stored at –80 °C freezer, where aliquots were stored until assay.

MiRNA discovery and quantification by next-generation sequencing (NGS)
A total amount of 1 μg total RNA per sample was used as input material for the small RNA library. Sequencing libraries were generated using TruSeq® miRNA Sample Prep Kit for Illumina®. Manufacturers’ recommendations and index codes were added to attribute sequences to each sample. The first-strand cDNA was synthesized using SuperScript II (ThermoFisher), and library quality was assessed on the Qubit system (Invitrogen). Raw data preparations were sequenced on a Nova6000 platform (Illumina). Fastp (v.0.20.1) was applied to remove low-quality reads and adapters of the raw sequencing reads. Using the miRdeep2 mapper.pl command, the filtered clean sequencing reads were mapped to human genome GRCh38 obtained from ENSEMBL (release 95). Then, the miRDeep2.pl command was executed to quantify known miRNAs, with miRNA reference files containing mature miRNA sequences and their corresponding precursor sequences downloaded from the miRBase database (https://www.mirbase.org/).

Quantitative reverse transcription polymerase chain reaction (qRT-PCR)
qRT-PCR analysis was used to verify miRNA expression. RNA was isolated from tissues and plasma using the miRcute miRNA Isolation Kit (Tian Biotech, China) and reverse transcribed using PrimeScript RT Master Mix for RT-PCR (Roche Diagnostics, USA) according to the manufacturer's protocol. For miRNA expression analysis, miRNA quantification was determined by using Bulge-Loop™ miRNA qRT-PCR Primer Sets (one RT primer and a pair of qRT-PCR primers for each set) specific for miR-196b-5p and miR-584b-5p, designed by RiboBio (RiboBio, China). All reagents for stem-loop RT mature miRNA assays were obtained from RiboBio (RiboBio, China). Real-time PCR was performed using SYBR Select Rremix Ex Taq II (TaKaRa Bio, Japan) on the Bio-Rad CFX96 Real-Time Detection System (Bio-Rad, USA). Data were normalized to levels of U6 (tissue) or cel-miR-39-3p (plasma) as appropriate, respectively. All analysis was carried out using the 2^−ΔΔCt method.

Development of the miRNA signature for early diagnosis of PE
Differential expression analysis of miRNAs was conducted between PE placental samples and normal placental samples using R package DESeq2 (V1.34.0) with the Wald test function. Those miRNAs with a fold change value greater than 2.0 or lower than 0.5 and a false discovery rate (FDR)-adjusted p-value lower than 0.05 were considered significantly differentially expressed miRNAs (DEmiRNAs). A Pearson correlation analysis was performed to filter DEmiRNAs related to clinical characteristics. To validate the sRNA-seq data, DEmiRNAs between the PE and normal groups were investigated by qRT-PCR in WMU cohort 2. Finally, the filtered DEmiRNAs were used to establish a miRNA signature (miR2PE-score) for PE diagnosis using Firth's bias-reduced logistic regression model. The miR2PE-score limits the risk score to 0–1, and the meaning of that is the perceived
probability of having PE according to our model. Using 0.5 as the threshold, samples with miR2PE-score > 0.5 are considered to have a high risk of PE, and vice versa.

**Functional enrichment analysis**

The target genes of miRNAs were extracted from miRTarBase (https://mirtarbase.cuhk.edu.cn) [30], with at least one experimental verification as the standard. Gene Ontology (GO) term and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway enrichment analysis was performed using Metascape (https://metascape.org/) [31]. All genes of the human genome were used as the enrichment background. Terms and pathways with p-value < 0.05, the number of genes enriched greater than 2, and enrichment factor > 1.5 were considered enriched terms and pathways. According to semantic similarity, similar enriched GO terms or KEGG pathways with kappa score > 0.3 were clustered to form a functional network. Each node in the functional network represents a GO term or KEGG pathway. The node size is proportional to the number of overlapping genes between the genes of interest and the genes belonging to the term or pathway, and the node color represents a cluster annotation. Similar term or pathway nodes are marked with the same color and entries of the representing nodes are used as cluster annotations.

**Statistical analysis**

All statistical analysis and generation of figures were performed with the statistical software R, version 4.0.5 (https://www.r-project.org), together with R packages “ggplot2,” “ggrepel,” “ggpubr,” “phetmap,” “corrplot,” “pROC,” “logistf,” “caret,” and “PRROC.” The Mann–Whitney U test was applied to compare outcomes between two independent groups. The area under the receiver operating characteristic curve (AUROC), precision–recall (PR) curve, sensitivity (TPR), specificity (TNR), positive predictive value (PPV), negative predictive value (NPV), and overall accuracy (ACC) were used to evaluate the diagnostic performance.

**Results**

**Characteristics of study populations**

The clinical and demographic data were obtained during routine visits and were recorded using standard datasheets. Gestational time was calculated with an algorithm based on a participant’s last normal menstrual period confirmed by early ultrasound. Patients with PE were diagnosed according to the technical bulletin of the American College of Obstetricians and Gynecologists [32]. All PE patients were nulliparous or multigravida (range 1–2). Subjects were excluded if they had abnormal 1-h glucose tolerance tests or pre-existing medical conditions such as hypertension, cardiovascular disease, diabetes mellitus, renal disease, or other chronic systemic diseases.

Maternal plasma samples and placental tissues were collected as part of four cohorts. The clinical characteristics of each cohort according to women with and without PE are presented in Table 1. Maternal age and BMI were statistically similar between healthy controls and PE patients (all p > 0.05). As expected, in all cohorts, pre-operative systolic blood pressure (PreSBP), preoperative diastolic blood pressure (PreDBP), and urine protein/urine CERA (mg/mmol) in patients with PE were significantly higher than those in healthy controls (all p < 0.05). There was an earlier median gestational age of delivery and lower mean birthweight in PE patients compared to healthy controls. Except for cohort 1, despite large variations among samples, albumin and blood urea nitrogen were statistically significant in the analysis with the other three cohorts included, respectively.

**Identification and verification of PE-associated miRNA biomarkers**

In the discovery stage, we first compared the expression profiles of 2888 miRNAs derived from miRNA sequencing between placental tissue samples of 5 PE patients and 5 healthy controls in WMU cohort 1 and identified five significantly differentially expressed miRNAs (DEmiRNAs) with the screening strategy of fold change > 2.0 or lower than 0.5 and FDR-adjusted p < 0.05 between the PE and the normal group (Fig. 2A). Unsupervised hierarchical clustering analysis showed that the expression patterns of five DEmiRNAs were able to discriminate PE samples from healthy controls (Fig. 2B). Among the five DEmiRNAs, two miRNAs (miR-19a-3p and miR-520f-5p) were found to be up-regulated and three miRNAs (miR-584-5p, miR-196b-5p, and miR-1299) were down-regulated in PE placental samples compared with normal placental samples (Fig. 2C). Functional enrichment analysis of GO and KEGG showed that the target genes of DEmiRNAs tended to be enriched in biological processes and pathways related to fetal growth and development and the pathogenesis of PE (Fig. 2D–G).

To further examine the association between expression levels of DEmiRNAs and PE-associated clinical characteristics (including age, BMI during pregnancy, PreSBP, PreDBP, and the content of platelet, ALB, GLOB, ALT, BUN, and BUN/CREA), we calculated Pearson correlation coefficients and found that four of the five DEmiRNAs were associated with at least one PE-associated clinical characteristics (Fig. 3A, B). Thus, we further verified the expression of these four DEmiRNAs in placental tissue samples of 20 PE patients and 20 healthy controls from WMU cohort 2 using qRT-PCR and validated miR-196b-5p and miR-584-5p as reliable and robust biomarkers (Fig. 3C).
| Characteristics | WMU cohort 1 | | WMU cohort 2 | | WMU cohort 3 | | WMU cohort 4 | |
|-----------------|-------------|------------------|------------------|------------------|------------------|------------------||
|                 | Healthy controls (n = 5) | PE patients (n = 5) | p-value | Healthy controls (n = 20) | PE patients (n = 20) | p-value | Healthy controls (n = 8) | PE patients (n = 10) | p-value | Healthy controls (n = 18) | PE patients (n = 19) | p-value |
| Mean maternal age (years) | 30.2 ± 5.7 | 31.0 ± 6.1 | 0.84 | 298 ± 5.1 | 304 ± 45 | 0.09 | 34.8 ± 4.2 | 31.1 ± 5.1 | 0.12 | 30.9 ± 4.8 | 31.3 ± 45 | 0.78 |
| Mean BMI (kg/m²) | 28.8 ± 2.9 | 30.6 ± 5.5 | 0.53 | 278 ± 1.9 | 280 ± 3.2 | 0.85 | 25.2 ± 2.5 | 276 ± 42 | 0.18 | 27.0 ± 2.7 | 278 ± 37 | 0.50 |
| PreSBP (mmHg) | 115.0 ± 70 | 164.6 ± 21.5 | < 0.01 | 117.6 ± 108 | 146.4 ± 17.2 | < 0.001 | 117.1 ± 10.1 | 152.9 ± 20.8 | < 0.001 | 117.4 ± 9.8 | 150.8 ± 17.7 | < 0.001 |
| PreDBP (mmHg) | 70.0 ± 63 | 107.2 ± 10.0 | < 0.001 | 72.0 ± 8.7 | 96.0 ± 13.1 | < 0.001 | 69.8 ± 5.3 | 100.6 ± 12.3 | < 0.001 | 722 ± 6.5 | 100.4 ± 11.0 | < 0.001 |
| Mean ALB (g/L) | 33.3 ± 3.2 | 29.8 ± 4.3 | 0.18 | 37.0 ± 3.4 | 30.4 ± 66 | < 0.001 | 35.6 ± 3.1 | 31.3 ± 49 | < 0.05 | 364 ± 2.8 | 325 ± 47 | < 0.001 |
| Mean ALT (IU/L) | 102.3 ± 39 | 20.6 ± 11.1 | 0.08 | 105.5 ± 5.0 | 145.73 | 0.05 | 9.6 ± 3.6 | 16.8 ± 12.7 | 0.14 | 10.2 ± 6.0 | 20.1 ± 11.6 | < 0.001 |
| Mean BUN (mmol/L) | 29.0 ± 5.6 | 44.4 ± 1.9 | 0.15 | 3.0 ± 0.8 | 4.1 ± 1.5 | < 0.01 | 3.1 ± 0.8 | 52 ± 29 | 0.07 | 3.0 ± 0.9 | 45 ± 2.2 | < 0.005 |
| BUN/CERA | 0.07 ± 0.02 | 0.08 ± 0.03 | 0.40 | 0.07 ± 0.02 | 0.08 ± 0.02 | 0.12 | 0.08 ± 0.08 | 0.10 ± 0.04 | 0.39 | 0.07 ± 0.03 | 0.09 ± 0.03 | 0.14 |
| Urine protein/urine CERA (mg/mmol) | 12.3 ± 3.5 | 517.9 ± 203.1 | < 0.001 | 119.4 ± 43 | 245.0 ± 307.6 | < 0.001 | 9.7 ± 3.9 | 508.4 ± 405.9 | < 0.001 | 11.2 ± 3.2 | 295.0 ± 327.0 | < 0.001 |
| Median GA delivery (week) | 38.2 ± 1.1 | 35.6 ± 2.9 | 0.10 | 38.4 ± 0.9 | 35.7 ± 26 | < 0.001 | 37.9 ± 0.6 | 32.9 ± 48 | < 0.05 | 38.5 ± 0.9 | 340 ± 41 | < 0.001 |
| No. preterm deliveries (%) | 0 (0.0%) | 3 (600%) | 0.17 | 0 (0.0%) | 12 (600%) | < 0.001 | 0 (0.0%) | 7 (700%) | < 0.01 | 0 (0.0%) | 12 (63.2%) | < 0.001 |
| Mean birthweight (g) | 3376.0 ± 405.3 | 22440.0 ± 968.8 | < 0.05 | 3412.0 ± 352.5 | 24610.0 ± 6898 | 0.001 | 3380.0 ± 343.3 | 17950.0 ± 8601 | < 0.001 | 3430.0 ± 432.8 | 21380.0 ± 8483 | < 0.001 |

Values are given as mean ± sd or n(%)

BMI: Body mass index during pregnancy, PreSBP: Preoperative systolic blood pressure, PreDBP: Preoperative diastolic blood pressure, ALB: Albumin, ALT: Alanine transaminase, BUN: Blood urea nitrogen, CREA: Creatinine, GA: Gestational age

p-values for comparisons between healthy controls with PE patients
Establishment and verification of a miRNA signature for PE diagnosis in two cross-platform retrospective cohorts

To establish a miRNA signature for the detection of PE, Firth’s bias-reduced logistic regression method was used to construct the predictive model based on the expression levels of miR-196b-5p and miR-584-5p in WMU cohort 1. Finally, a miR2PE-score was calculated for each sample using the observed weights from the regression model and the expression value of miR-196b-5p and miR-584-5p with the equation:

\[
\text{miR2PE-score} = \frac{e^{\beta_0 + \beta_1 \log_{10}(\text{miR-196b-5p}) + \beta_2 \log_{10}(\text{miR-584-5p})}}{1 + e^{\beta_0 + \beta_1 \log_{10}(\text{miR-196b-5p}) + \beta_2 \log_{10}(\text{miR-584-5p})}}
\]

When miR2PE-score was applied to WMU cohort 1, the miR2PE-score and predicted labels of 5 PE placental samples and 5 normal placental samples were calculated and visualized in Fig. 4A. In WMU cohort 1, with the risk cutoff point of 0.5, five of five PE samples (100% sensitivity) and four of five normal samples (80.0% specificity) were correctly classified with an overall accuracy of 90.0% (nine of ten). The AUROC was 0.920 (95% confidence interval (CI) 0.736–1.000; Fig. 4A) and the area under the precision–recall curve (AUC-PR) was 0.919 (Additional file 1: Fig. S1A). To preliminarily verify cross-platform compatibility of miR2PE-score in PE diagnosis, the miR2PE-score was subsequently tested based on qRT-PCR data from WMU cohort 2. Figure 4B shows the expression levels of the two identified miRNAs, miR2PE-scores, and predicted labels and true labels of 20 PE and 20 control samples in WMU cohort 2. As expected, the miR2PE-score still works well in the qRT-PCR data and successfully identified 19 PE and 18 normal samples with an overall accuracy of 92.5%, an AUROC of 0.848 (95% CI 0.720–0.975; Fig. 4B), and PR-AUC of 0.815 (Additional file 1: Fig. S1B). These results initially confirmed that the miR2PE-score is reliable and compatible.

Independent validation of the miR2PE-score in external multicenter cohorts

To further evaluate the cross-platform compatibility and multinational robustness of the miR2PE-score, we tested the performance of the miR2PE-score in two external cohorts, including the Canada cohort with the miRNA sequencing platform and the USA cohort with the microarray platform. In the Canada cohort, with the same formula and the risk cutoff (0.5), 16 of 20 PE samples (80.0% sensitivity) and 15 of 21 normal samples (71.4% specificity) were correctly classified by the miR2PE-score with an overall accuracy of 75.6% (31 of 41; Fig. 5A). The two identified miRNA biomarkers, miR2PE-scores, and predicted labels and true labels of placental tissue samples in the Canada cohort are shown in Fig. 5B. As shown in Fig. 5C, the AUROC was 0.864 (95% CI 0.757–0.972; Fig. 5C) and the AUC-PR was 0.874 (Additional file 1: Fig. S2A). In the USA cohort, the miR2PE-score successfully identified 11 of 16 PE samples (68.8% sensitivity) and 12 of 16 normal samples (75.0% specificity) with an overall accuracy of 71.9% (23 of 32; Fig. 5D). The two identified miRNA biomarkers, miR2PE-scores, and predicted labels and true labels of placental tissue samples in the USA cohort are shown in Fig. 5E. Similarly, in the USA cohort, the miR2PE-score demonstrated an AUROC of 0.812 (95% CI 0.661–0.964; Fig. 5F) and an AUC-PR of 0.786 (Additional file 1: Fig. S2B). These results further confirmed that the predictive performance of miR2PE-score is reliable and universal across different technology platforms and countries.

Prospective validation of miR2PE-score as a non-invasive biomarker in whole blood samples from two independent cohorts

To further examine whether the miR2PE-score can be detected in blood samples and used as a potential minimally invasive diagnostic biomarker, the expression levels of two identified miRNA biomarkers in peripheral blood leukocytes from WMU cohort 3 (including 10 PE and 8 normal samples) and WMU cohort 4 (including 19 PE and 18 normal samples) were analyzed with small RNA sequencing and qRT-PCR, respectively. As shown in Fig. 6A and B, two miRNA biomarkers (miR-196b-5p and miR-584-5p) revealed lower expression levels in peripheral blood leukocytes from PE patients than in healthy controls in both WMU cohort
Fig. 2 (See legend on previous page.)
3 and WMU cohort 4, which is consistent with observations from placental tissue-based validation. Using the miR2PE-score, 8 of 10 PE samples (80.0% sensitivity) and 5 of 8 normal samples (62.5% specificity) in WMU cohort 3 were correctly classified with an overall accuracy of 72.2% (13 of 18; Fig. 6C). The two identified miRNAs, miR2PE-scores, and predicted labels and true labels of placental tissue samples in WMU cohort 3 are shown in Fig. 6D. The AUROC was 0.787 (95% CI 0.565–1.000; Fig. 6D) and the AUC-PR was 0.823 (Additional file 1: Fig. S3A). In WMU cohort 4, 17 of 19 PE samples (89.5% sensitivity) and 15 of 18 normal samples (83.3% specificity) were correctly classified by the miR2PE-score with an overall accuracy of 86.5% (32 of 37; Fig. 6E). The two identified miRNAs, miR2PE-score, and predicted labels and true labels of placental tissue samples in WMU cohort 4 are shown in Fig. 6F. The AUROC was 0.933 (95% CI 0.844–1.000; Fig. 6F) and the AUC-PR was 0.953 (Additional file 1: Fig. S3B). Taken together, the above results suggest that miR2PE-score could serve as a minimally invasive diagnostic biomarker for the diagnosis of PE in the clinic.
The performance of the miR2PE-score was further compared with five recently proposed PE miRNA biomarkers (miR-210-3p, miR-19b-1-5p, miR-92a-1-5p, miR-486-5p, and miR-18a-5p) in different cohorts. As shown in Fig. 7A–C, the miR2PE-score and miR-210-3p performed very well in all three placental tissue-based cohorts compared with the other four previously published miRNA biomarkers. However, the miR2PE-score achieved the best predictive performance in blood-based WMU cohort 3 compared with the other five previously published miRNA biomarkers (Fig. 7D). These results demonstrated that the miR2PE-score holds great potential to become a more effective and robust non-invasive biomarker for the early diagnosis of PE.

**Discussion**

Preeclampsia is a pregnancy-specific disease with hypertension and proteinuria as the primary clinical manifestations and systemic arterial spasm as the main pathological features, affecting 2–4% of pregnancies [33, 34]; it poses a serious threat to the health of mothers and newborns. Prophylactic use of low-dose aspirin before the 16th gestational week appears to benefit pregnant women at high risk of preeclampsia from a perspective of preeclampsia prevention [35]. However, the clinical signs and symptoms of PE are non-specific and therefore pose a challenge to the early detection of women at the highest risk of developing PE. Although conventional diagnostic assessments for preeclampsia, such as blood pressure, proteinuria, lactate dehydrogenase, and other laboratory indicators, have been widely used, there is a lack of accurate testing criteria for preeclampsia. Their diagnostic performance may be affected by multiple clinical variables, such as race, body mass index, maternal comorbidities and/or obstetrical history, and unsatisfactory sensitivity and specificity [36, 37].

A growing body of evidence has showed that placental miRNAs play important functional roles in regulating placental pathophysiological processes and affecting adverse pregnancy outcomes [38–41]. The release of microRNAs from the placenta mainly occurs from the villous trophoblast cells, indicating that circulating microRNAs could serve as promising markers for monitoring trophoblast and placental function [42, 43]. Although abnormal expressions of circulating miRNAs have been found in pregnant women with preeclampsia, unfortunately, the clinical significance and predictive value of a few miRNAs have been systematically investigated and studied in PE diagnosis and outcomes [44–46]. Furthermore, a comparative study also indicated the unsatisfactory performance of potential miRNA markers previously reported in PE detection [47]. Therefore, the search for clinically viable non-invasive miRNA biomarkers for early detection of PE is still of utmost importance.
In the present study, we performed a high-throughput sequencing analysis for miRNAome and conducted a genome-wide screening for potential miRNA biomarkers. We identified five significantly differentially expressed miRNAs, and among which two miRNAs (miR-196b-5p and miR-584-5p) were finally selected by considering the association between the expression pattern and PE-associated clinical characteristics and qRT-PCR validation in tissues. From a functional viewpoint, the target genes of these two miRNA biomarkers have been shown to be involved in PE pathogenesis. Furthermore, a recent study by Mi et al. in PE models in vitro and in vivo revealed the pathological roles of the BHLHE40/miR-196a-5p/SNX16 axis in PE [48]. The dysregulated expression of miR-584 has also been observed in a previous placental microRNA expression study [49]. These existing pieces of knowledge further supported the critical roles of these two miRNA biomarkers identified in our study in PE pathogenesis, but their clinically applicable values in PE diagnosis remain questionable.

Therefore, we established a diagnostic signature (miR2PE-score) that comprised miR-196b-5p and miR-584-5p for the early detection of PE. To reduce the small sample bias of the maximum likelihood (ML) estimators of the logistic regression model, Firth's bias-reduced logistic regression was introduced to fit the miR2PE-score. Firth's bias-reduced logistic regression method is suitable for the problem of complete data separation caused by the extremely low number of cases. Furthermore, the miR2PE-score produced a scale of 0.0–1.0, which is beneficial to clinically risk probability assessment. Although the miR2PE-score was developed based on miRNA sequencing data, the performance of the miR2PE-score was also validated by qRT-PCR data,
highlighting the clinical cost-effectiveness and cross-platform compatibility. It is important to note that the miR2PE-score was developed and validated primarily in our retrospective Chinese cohorts. Therefore, the miR2PE-score was further tested in external multicenter cohorts comprised of non-Asian races and also showed high diagnostic performance similar to that of our retrospective Chinese cohorts. These independent cohorts from different centers were not integrated to form a cohort. Therefore, we did not do batch correction and normalization. Each cohort is used individually to validate the expression patterns of these two miRNA biomarkers and further evaluate the performance and robustness of the miR2PE-score.
Considering potential clinical application of the miR2PE-score in true clinical settings, we expanded the evaluation of the miR2PE-score as a non-invasive liquid biopsy assay to prospectively collected plasma cohorts. Although expression levels of two miRNAs in the miR2PE-score were measured using different technical methods (miRNA sequencing and qRT-PCR), the miR2PE-score still demonstrated effectiveness in distinguishing PE from healthy subjects. Blood collection was performed before the cesarean leading to potential gestational age-matched in control subjects and PE patients, which may introduce variation in data, and the scoring may not be accurate in the clinic when evaluating patients at different gestational ages. Therefore, we also examine the diagnostic performance of the miR2PE-score after adjusting for gestational ages through logistic regression, which demonstrated that the miR2PE-score still is a significant predictive factor even after adjusting for gestational ages (Additional file 1: Fig. S4). Furthermore, we showed that the miR2PE-score performed significantly better than any previously published miRNA biomarkers in blood-based cohorts and achieved the best and most stable predictive performance compared to other parallel methods across different cohorts (Additional file 1: Fig. S5). However, a potential limitation of our study is that although we tested the miR2PE-score in cross-platform
Conclusions
In conclusion, our study developed and validated a novel and robust blood-based miRNA signature for the early detection of PE in cross-platform retrospective and prospective multicenter cohorts. Investigations into non-invasive diagnostic performance in population-based prospective cohort studies warrant clinical transformation application in the future.

Abbreviations
AUCoR: Area under the receiver operating characteristic curve; FDR: False discovery rate; GO: Gene Ontology; KEGG: Kyoto Encyclopedia of Genes and Genomes; NGS: Next-generation sequencing; PE: Preeclampsia; PlGF: Placental growth factor; qRT-PCR: Quantitative reverse transcription polymerase chain reaction; sEng: Soluble endoglin; sFlt-1: Fms-like tyrosine kinase-1.

Supplementary Information
The online version contains supplementary material available at https://doi.org/10.1186/s12916-022-02495-x.

Received: 28 April 2022 Accepted: 25 July 2022
Published online: 13 September 2022

References
1. Kaaja R, Laivuori H, Kaasko M, Tikkanen M, Ylikorkala O. Evidence of a state of increased insulin resistance in preeclampsia. Metabolism. 1999;48(7):892–6.
2. Bramham K, Villa P, Joslin J, Laivuori H, Hämäläinen E, Kajantie E, et al. Predisposition to superimposed preeclampsia in women with chronic hypertension: endothelial, renal, cardiac, and placental factors in a prospective longitudinal cohort. Hypertens Pregnancy. 2020;39(3):326–35.
3. Leach R, Romero R, Kim Y, Chaiworapongsa T, Kilburn B, Das S, et al. Pre-eclampsia and expression of heparin-binding EGF-like growth factor. Lancet (London). 2002;360(9341):1215–9.
4. Ness R, Sibai B. Shared and disparate components of the pathophysiology of fetal growth restriction and preeclampsia. Am J Obstet Gynecol. 2006;195(1):40–9.
5. Mol BWJ, Roberts CT, Thangaratnam S, Magee LA, de Groot CJM, Hofmeyr GJ. Pre-eclampsia. Lancet. 2016;387(10022):999–1011.
6. Kolkova Z, Holubekova V, Grenard M, Nachajova M, Zubor P, Pribulova T, et al. Association of circulating miRNA expression with preeclampsia, its onset, and severity. Diagnostics (Basel). 2021;11(3):476.
7. Poon L, Shennan A, Hyett J, Kapur A, Hadar E, Divakar H, 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 Gynaecol Obstet. 2019;154(S1):1–33.
8. Karumanchi SA. Angiogenic factors in preeclampsia: from diagnosis to therapy. Hypertension. 2016;67(6):1072–9.
9. Dechend R, Luft FC. Angiogenesis factors and preeclampsia. Nat Med. 2008;14(11):1187–8.
10. Cin N, Kurdoğlu M, Ege S, Yoruk I, Yaman G, Yıldızhan R. An analysis on the roles of angiogenesis-related factors including serum vitamin D, soluble endoglin (sEng), soluble fms-like tyrosine kinase 1 (sFlt1), and vascular endothelial growth factor (VEGF) in the diagnosis and severity of late-onset preeclampsia. J Matern Fetal Neonatal Med. 2017;30(13):1602–7.
11. Sung KU, Roh JA, Eoh KJ, Kim EH. Maternal serum placental growth factor and pregnancy-associated plasma protein A measured in the first trimester as parameters of subsequent pre-eclampsia and small-for-gestational-age infants: a prospective observational study. Obstet Gynecol Sci. 2017;60(2):154–62.
12. Saleh L, Tahitu SIM, Danser AHJ, van den Meiracker AH, Visser W. The predictive value of the sFlt-1/PlGF ratio on short-term absence of preeclampsia and maternal and fetal or neonatal complications in twin pregnancies. Pregnancy Hypertens. 2018;14:222–7.
13. Lou WZ, Jiang F, Hu J, Chen XX, Song YN, Zhou XY, et al. Maternal serum angiogenic factor sFlt-1 to PI GF ratio in preeclampsia: a useful marker for differential diagnosis and prognosis evaluation in Chinese women. Dis Markers. 2019;2019:6270187.
14. Bartel D. MicroRNAs: target recognition and regulatory functions. Cell. 2009;136(2):215–33.
15. Wang X, Qian T, Bao S, Zhao H, Chen H, Xing Z, et al. Circulating exosomal miR-363-5p inhibits lymph node metastasis by downregulating PDGFB
and serves as a potential noninvasive biomarker for breast cancer. Mol Oncol. 2021;15(9):2466–79.

16. Bao S, Hu T, Liu J, Su J, Sun J, Ming Y, et al. Genomic instability-derived plasma extracellular vesicle-microRNA signature as a minimally invasive predictor of risk and unfavorable prognosis in breast cancer. J Nanobiotechnology. 2021;19(1):22.

17. Gebert LFR, MacRae I. Regulation of microRNA function in animals. Nat Rev Mol Cell Biol. 2019;20(1):21–37.

18. Gantier MP, McCoy CE, Rusinova I, Saulep D, Wang D, Xu D, et al. Analysis of microRNA turnover in mammalian cells following Dicer1 ablation. Nucleic Acids Res. 2011;39(13):5602–703.

19. Sayed AS, Xia K, Salma U, Yang T, Peng J. Diagnosis, prognosis and therapeutic role of circulating miRNAs in cardiovascular diseases. Heart Lung Circ. 2014;23(6):503–10.

20. Xu P, Zhao Y, Liu M, Wang Y, Wang H, Li Y, et al. Variations of microRNAs in human placenta and plasma from preeclamptic pregnancy. Hypertension (Dallas, Tex. 1979). 2014;63(6):1276–84.

21. Murphy M, Casselman R, Tayade C, Smith G. Differential expression of plasma micro RNA in preeclamptic patients at delivery and 1 year postpartum. Am J Obstet Gynecol. 2015;212(3):367.e1–9.

22. Zheng W, Chen A, Yang H, Hong L. MicroRNA-27a inhibits trophoblast cell migration and invasion by targeting SMAD2. Potential role in preeclampsia. Exp Ther Med. 2020;20(1):2626–29.

23. Cui J, Chen X, Li S, Li L, Fan J, Hou H, et al. MiR-101-containing extracellular vesicles bind to BRD4 and enhance proliferation and migration of trophoblasts in preeclampsia. Stem Cell Res Ther. 2020;11(1):231.

24. Licini C, Avellini C, Picciocchi E, Mensa E, Fantone S, Ramini D, et al. Pre-eclampsia predictive ability of maternal miR-125b: a clinical and experimental study. Tranls Res. 2021;228:13–27.

25. Wang D, Na Q, Song GY, Wang L. Human umbilical cord mesenchymal stem cell-derived exosome-mediated transfer of microRNA-133b boosts trophoblast cell proliferation, migration and invasion in preeclampsia by restricting SGK1. Cell Cycle. 2020;19(15):1869–83.

26. Zhou W, She G, Yang K, Zhang B, Liu J, Yu B. MiR-384 inhibits proliferation and migration of trophoblast cells via targeting PTBP3. Pregnancy Hypertens. 2020;21:132–48.

27. Awamleh Z, Gloor GB, Han VKM. Placental microRNAs in pregnancies with early onset intrauterine growth restriction and preeclampsia: potential impact on gene expression and pathophysiology. BMC Med Genomics. 2019;12(1):91 (https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE114349).

28. Brooks S, Martin E, Smeester L, Boggs K, Grace M, Fry R. GEO https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE84260. 2016.

29. Wyatt S, Kraus F, Roh C, Elchalal U, Nelson D, Sadovsky Y. The correlation between sampling site and gene expression in the term human placenta. Placenta. 2005;26(5):372–9.

30. Huang HY, Lin YC, Li J, Huang KY, Shrestha S, Hong HC, et al. miRTarBase 2020: updates to the experimentally validated microRNA-target interaction database. Nucleic Acids Res. 2020;48(D1):D148–54.

31. Zhou Y, Zhou B, Parche L, Chang M, Khodabakhshi AH, Tanaseichuk O, et al. Metascape provides a biologist-oriented resource for the analysis of systems-level datasets. Nat Commun. 2019;10(1):1523.

32. Gestational Hypertension and Preeclampsia: ACOG Practice Bulletin. Obstet Gynecol. 2020;135(6):e237–60.

33. Magee LA, Nicolaides KH, von Dadelszen P. Preeclampsia. N Engl J Med. 2019;380(21):2084–95.

34. Phipps EA, Thadhani R, Benzing T, Karumanchi SA. Pre-eclampsia: factors, diagnosis, management, and the cardiovascular impact on the offspring. J Clin Med. 2019;8(10):1625.

35. Rasmussen M, Reddy M, Nolan R, Camunas-Soler J, Khodursky A, Scheller NM, et al. RNA profiles reveal signatures of future health and disease in pregnancy. Nature. 2012;490(7419):422–7.

36. Rudov A, Balduni W, Caroli S, Perrone S, Buonocore G, Albertini MC. Involvement of miR-34a in placental alterations mediated by oxidative stress. Oxid Med Cell Longev. 2014;2014:103068.