International Journal of
Molecular Sciences

MDPI

Article

Machine Learning and Blood-Targeted Proteomics Enable Early Prediction and Etiological Discrimination of Hypertensive Pregnancy Disorders

Natalia Starodubtseva [1,*], Alisa Tokareva [1], Alexey Kononikhin [1,2], Anna Bugrova [1,3], Maria Indeykina [2,3], Evgenii Kukaev [1,4,5], Alina Poluektova [1], Alexander Brzhozovskiy [1,2], Evgeny Nikolaev [2] and Gennady Sukhikh [1,6]

1 V.I. Kulakov National Medical Research Center for Obstetrics Gynecology and Perinatology, Ministry of Healthcare of Russian Federation, 117997 Moscow, Russia; a_tokareva@oparina4.ru (A.T.)
2 Project Center of Omics Technologies and Advanced Mass Spectrometry, 121205 Moscow, Russia
3 Emanuel Institute of Biochemical Physics, Russian Academy of Sciences, 119334 Moscow, Russia
4 V.L. Talrose Institute for Energy Problems of Chemical Physics, N.N. Semenov Federal Research Center for Chemical Physics, Russian Academy of Sciences, 119334 Moscow, Russia
5 Moscow Center for Advanced Studies, 123592 Moscow, Russia
6 Department of Obstetrics, Gynecology, Perinatology and Reproductology, Institute of Professional Education, Federal State Autonomous Educational Institution of Higher Education I.M. Sechenov First Moscow State Medical University of the Ministry of Health of the Russian Federation, 119991 Moscow, Russia
* Correspondence: n_starodubtseva@oparina4.ru

Abstract

Imperfect first-trimester screening for hypertensive disorders of pregnancy (HDP) means many high-risk women miss the window for preventive aspirin, and the biological heterogeneity of HDPs is overlooked. This study aimed to leverage first-trimester serum proteomics to create a more precise tool for predicting preeclampsia (PE) and differentiating it from other HDPs. A prospective nested case–control study (n = 172) was conducted using targeted liquid chromatography-multiple reaction monitoring-mass spectrometry (LC-MRM-MS) proteomic profiling of 115 proteins. Machine learning (ML) methods were used to develop classifiers from the proteomic data. The signature predictive of PE was characterized by dysregulation of the complement and coagulation cascades (*F10*, *C8A*, *C1QA*, *SERPING1*, *VTN*). The profile differentiating gestational hypertension (GAH) from chronic hypertension (CAH) was linked to lipid metabolism (*HRG*, *APOA4*, *APOC2*). An 18-protein support vector machine (SVM) model for predicting PE demonstrated exceptional performance, with 94% sensitivity and 100% specificity, significantly outperforming the standard Fetal Medicine Foundation (FMF) screening algorithm. Pathway analysis confirmed that PE is associated with early activation of innate immunity and coagulation pathways, while GAH is linked to a pregnancy-induced metabolic response. A targeted serum proteomic combined with ML approach represents a new perspective diagnostic tool with strong potential to personalize monitoring for women at the highest risk for specific hypertensive pregnancy complications.

Keywords: preeclampsia; gestational arterial hypertension; chronical arterial hypertension; proteome; serum; pregnancy; machine learning

Academic Editor: Camelia Quek

Received: 31 December 2025
Revised: 23 January 2026
Accepted: 28 January 2026
Published: 30 January 2026

Copyright: © 2026 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license.

1. Introduction

Hypertensive disorders of pregnancy (HDP)—including chronic hypertension (CAH), gestational hypertension (GAH), and preeclampsia (PE)—are leading causes of maternal and perinatal morbidity and mortality worldwide. PE, the most severe form, is characterized by systemic endothelial dysfunction and placental malperfusion, often leading to multi-organ complications [1,2]. While low-dose aspirin can reduce the incidence of preterm PE by over 60%, its efficacy hinges on early and accurate identification of high-risk women [3].

The first-trimester screening algorithm developed by the Fetal Medicine Foundation (FMF), which combines maternal factors, mean arterial pressure (MAP), uterine artery Doppler, and serum biomarkers (placental growth factor, PIGF and pregnancy-associated plasma protein-A, PAPP-A), demonstrates strong predictive performance in research settings. However, its sensitivity in real-world cohorts is more modest, ranging from 63% to 71% [4–6]. This performance gap, potentially attributable to operator-dependent variability in Doppler measurements [7], means that 25–40% of women who develop preterm PE are missed and do not receive aspirin prophylaxis.

The clinical challenge is compounded by the fact that HDPs are a spectrum of disorders with different underlying causes and risks [8–10]. While GAH increases the risk of severe maternal morbidity by 50%, CAH without PE presents a risk profile similar to normotensive pregnancies [11]. Furthermore, CAH is independently linked to placental complications [12]. The proven benefit of early blood pressure control, as demonstrated in the CHAP trial, highlights the critical need to identify and differentiate these conditions early [13]. Current screening falls short in this regard, as it is primarily designed to predict placental PE and cannot distinguish between the varied etiologies across the HDP spectrum.

High-throughput proteomics offers a promising path forward, with several studies demonstrating that first-trimester alterations in the maternal serum proteome can predict PE long before clinical onset [14–17]. However, most proteomic investigations to date have treated HDP as a unified group or compared PE only to normotensive controls, without differentiating between CAH and GAH [18,19].

To address this gap, this study aimed to derive and validate a first-trimester maternal serum proteomic signature capable of not only predicting HDP but also differentiating their underlying etiologies. By applying a rigorously validated targeted liquid chromatography-multiple reaction monitoring mass spectrometry (LC-MRM-MS) platform [20] together with machine learning, this approach successfully distinguished PE from GAH, pre-existing CAH, isolated intrauterine growth restriction (IUGR), and normotensive pregnancy.

2. Results

2.1. Analysis of First-Trimester Predictors and Delivery Outcomes

This prospective study included 172 women undergoing first-trimester prenatal screening at 11^{+2} to 14^{+2} weeks of gestation, comprising five groups: PE (n = 32), GAH (n = 22), CAH (n = 24), isolated IUGR (n = 11) without PE, and a control group (n = 83) (Figure 1). Within the PE group, most cases were late-onset (23/32, 72%) and moderate in severity (27/32, 84%); eleven women (34%) also had concomitant IUGR. Significant differences in maternal history, first-trimester biomarkers, and delivery outcomes were observed across these groups (Table S1).

https://doi.org/10.3390/ijms27031402

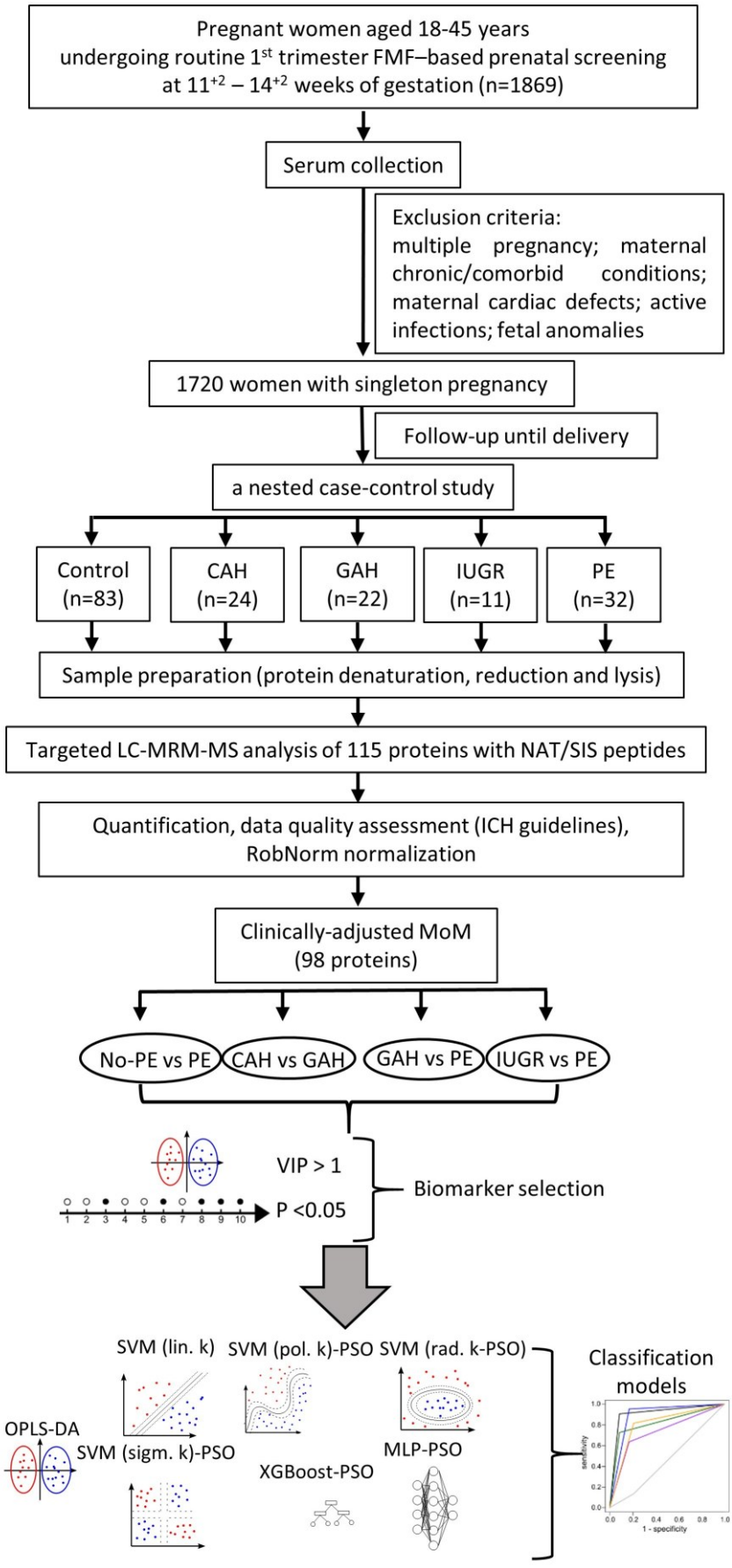


Figure 1. Overview of the experimental workflow from clinical sample collection to machine learning classifier development.

https://doi.org/10.3390/ijms27031402

Distinct maternal backgrounds were evident among the groups. A history of habitual miscarriage was significantly more common in the PE (38%, $p < 0.001$) and isolated IUGR (18%, $p = 0.048$) groups compared to the controls (0%). A prior history of PE was absent in both the control and GAH groups, but was present in the CAH (21%), isolated IUGR (18%), and PE (22%) groups ($p < 0.001$). Furthermore, a history of gestational diabetes mellitus (GDM) was significantly more frequent in women who developed PE (25%) compared to those without PE (1%) and to all other groups individually ($p < 0.001$). There were no differences between the groups in terms of parity ($p > 0.05$).

Women in the CAH and PE groups were significantly older than controls (mean 34.1 and 33.4 years vs. 30.3 years, $p = 0.006$ and $p = 0.002$, respectively). First-trimester body mass index (BMI) was also significantly higher in the CAH (median 24.9 kg/m^2, $p = 0.007$) and GAH (median 23.5 kg/m^2, $p = 0.003$) groups compared to controls (median 21.2 kg/m^2).

Signs of early placental dysfunction were prominent in the PE group. These women exhibited significantly lower first-trimester PAPP-A (2.07 vs. 3.09 mIU/mL, $p = 0.03$) and PlGF (15.7 vs. 26.6 pg/mL, $p < 0.001$), alongside a higher UtA-PI (2.00 vs. 1.56, $p < 0.001$) compared to controls. The CAH and GAH groups also showed significantly lower adjusted PlGF levels (0.63 and 0.58 multiples of the median (MoM), respectively) than controls (0.76 MoM, $p < 0.05$). Consequently, a high risk for PE from first-trimester screening was significantly more common in all hypertensive groups (CAH 29%, GAH 23%, PE 53%) than in controls (1%, $p < 0.001$).

Notably, women with isolated IUGR and PE showed no significant differences in their medical history or most first-trimester screening parameters. The key exception was adjusted PlGF, which was significantly lower in the PE group (0.56 MoM) than in the isolated IUGR group (0.89 MoM, $p = 0.005$) (Table S2). In contrast, a direct comparison between the CAH and GAH groups revealed no statistically significant differences in their clinical first-trimester parameters (Table S3).

As expected, blood pressure was significantly elevated in all hypertensive disorders compared to controls, both in the first trimester and before delivery ($p < 0.05$). The PE and CAH groups presented the most severe hemodynamic profile, with the highest first-trimester MAP (91 and 93 mm Hg, respectively, $p < 0.001$) and the highest systolic/diastolic pressures before delivery (135/89 mm Hg and 123/80 mm Hg, $p < 0.001$ vs. controls). Moreover, an increase in blood creatinine, indicative of renal endothelial dysfunction, was found exclusively in the PE group (80.8 μM/L vs. 67.7 μM/L in controls, $p < 0.001$).

Evidence of placental insufficiency in PE was also reflected in umbilical cord Doppler parameters, showing an elevated pulsatility index (0.95 vs. 0.78 in controls, $p < 0.001$) and a reduced cerebro-placental ratio (1.5 vs. 1.84 in controls, $p < 0.001$) at delivery (Table S1). Consequently, birth weight and Apgar scores were significantly lower in the placenta-associated pathology groups (PE and isolated IUGR) compared to all others ($p < 0.001$).

Direct comparisons further clarified the distinctiveness of these conditions. When compared to GAH, PE was associated with worse placental function (lower first-trimester PlGF, higher cord artery PI, and lower cerebro-placental ratio at delivery, $p < 0.01$), significantly lower placental weight ($p < 0.001$), and poorer neonatal outcomes ($p < 0.01$) (Table S4). Placental dysfunction in the PE group was further evidenced at delivery by a significantly reduced placental weight (362 g vs. 455 g in non-PE groups, $p < 0.001$) and a high rate of preterm birth (44% vs. 1%, $p < 0.001$). Emergency cesarean section was also significantly more frequent in the PE group (59%) than in the non-PE groups combined (13%, $p < 0.001$) (Table S5). In contrast, the GAH and CAH groups had rates of term delivery comparable to controls.

Notably, only 47% of women in the PE group received prophylactic doses of acetylsalicylic acid, which may indicate that in more than half of these cases the risk of developing PE had been assessed as low. Low-molecular-weight heparins were prescribed

most frequently in the PE group compared with all other study groups (59%, $p < 0.001$). It should also be noted that at the time of sample collection, the use of antihypertensive medications was reported only in the chronic hypertension group (33%, $p < 0.001$).

The comparison of the PE group to all non-PE participants consolidated its identity as a severe placental disorder, linked to a high-risk maternal background, a severe course of pregnancy from the first trimester, and markedly worse perinatal outcomes (Table S5).

2.2. Targeted Proteome Profiling and Analytical Performance

Proteomic analysis quantified 115 proteins that met the pre-defined quality control criteria ($R^2 > 0.99$, precision and accuracy < 20% in over five of seven standard levels and in more than 66% of QC samples; Table S6). Twelve proteins were excluded from further analysis due to concentrations below the LLOQ or upper the HLOQ in over 50% of study samples. These were: Alpha-1-antitrypsin, Apolipoprotein A-I, Cholesteryl ester transfer protein, Clusterin, Complement C1q subcomponent subunit B, Fibrinogen alpha chain, Fibrinogen beta chain, Fibrinogen gamma chain, Kallistatin, Keratin type I cytoskeletal 9, Peroxiredoxin-2, and Serum paraoxonase/lactonase 3.

RobNorm normalization significantly reduced the median between-batch coefficient of variation from 6.6% (IQR: 5.0–9.6%) to 6.3% (IQR: 4.6–9.1%; $p < 0.001$, Mann–Whitney U test; Figure S1). An additional five proteins—Inter-alpha-trypsin inhibitor heavy chain H1, Inter-alpha-trypsin inhibitor heavy chain H2, L-selectin, Serum amyloid A-1 and A-2 proteins, and Zinc-alpha-2-glycoprotein—were excluded due to a high coefficient of variation (CV ≥ 20%).

Consequently, 98 proteins were included in the final analysis (Figure 1). The dynamic range spanned five orders of magnitude, from Protein deglycase DJ-1 (*PARK7*) at 2.5 (1.8; 3.6) nM to Serum Albumin (*ALB*) at 4.6×10^5 (4.2×10^5; 5.0×10^5) nM (Table S7, Figure S2). Following normalization and adjustment to MoM values, Carbon Anhydrase 1 (*CA1*) exhibited the lowest median level (0.67 MoM, IQR: 0.36–1.19), while Biotinidase (*BTD*) exhibited the highest (1.06 MoM, IQR: 0.94–1.15). Hemopexin (*HPX*) was the most stable protein with a CV of 9%, whereas Apolipoprotein(a) (*LPA*) showed the widest variation with a CV of 284% (Table S8, Figure S3).

Serum concentrations of PAPP-A measured by routine ELISA and LC-MRM-MS were strongly correlated (Pearson's R = 0.61, $p < 0.001$). This correlation remained virtually unchanged following MoM transformation (R = 0.60, $p < 0.001$) (Table 1).

Table 1. Pearson correlation analysis of first-trimester serum PAPP-A levels quantified by ELISA and LC-MRM-MS.

ELISA (FMF Screening)	LC-MRM-MS	R	*p*-Value
0.96 (0.65; 1.48) MoM	5.54 (3.80; 8.03) mM	0.61	$p < 0.001$
0.96 (0.65; 1.48) MoM	0.80 (0.47; 1.16) MoM	0.60	$p < 0.001$

2.3. Hypertension-Associated Serum Proteins

Proteomic profiling identified distinct protein signatures associated with the different clinical groups ($n = 172$). When comparing the isolated IUGR and PE groups, serum concentrations of Coagulation factor X (*F10*), Complement component C8 alpha chain (*C8A*), Complement C1q subcomponent subunit A (*C1QA*), Alpha-2-HS-glycoprotein (*AHSG*), Complement C1q subcomponent subunit C (*C1QC*), and Attractin (*ATRN*) were statistically significantly higher in the IUGR group, whereas Vitronectin (*VTN*), Ig gamma-1 chain C region (*IGHG1*), Ficolin-3 (*FCN3*), Transthyretin (*TTR*), and Plasma protease C1 inhibitor (*SERPING1*) were significantly higher in the PE group (Figure S4A, Table S9). An Orthogonal Projections to Latent Structures-Discriminant Analysis (OPLS-DA)

model demonstrated good discrimination power for separating the IUGR and PE groups with $R^2Y = 0.99$ and $Q^2Y = 0.46$ (Figure S4B, Table 2). The most important proteins for this separation, with a VIP >1, included Apolipoprotein(a) (LPA), F10, C8A, C1QA, AHSG, *SERPING1*, *C1QC*, *CA1*, *IGHG1*, Beta-Ala-His dipeptidase (*CNDP1*), *VTN*, *FCN3*, and *TTR* (Figure S4C). Consequently, *F10*, *C8A*, *C1QA*, *AHSG*, *C1QC*, *SERPING1*, *TTR*, *IGHG1*, *VTN*, and *FCN3* were selected as key discriminatory protein features. An OPLS-DA model based specifically on these potential markers demonstrated better predictive power ($Q^2Y = 0.67$) (Table 2).

Table 2. Quality of OPLS-DA discrimination models, based on all proteins and sets of proteins features.

	All Proteins			Proteins-Markers		
	R^2X	**R^2Y**	**Q^2Y**	**R^2X**	**R^2Y**	**Q^2Y**
IUGR vs. PE	0.46	0.99	0.46	0.57	0.76	0.67
CAH vs. GAH	0.19	0.89	0.14	1	0.64	0.55
GAH vs. PE	0.25	0.97	0.72	0.47	0.82	0.77
Non-PE vs. PE	0.14	0.89	0.79	0.62	0.84	0.8

In the comparison between GAH and CAH, concentrations of *HRG*, *APCS*, *APOA4*, and *SERPINA6* were statistically significantly higher, while *APOC2* and *A1BG* were statistically significantly lower in the GAH group (Figure S5A, Table S10). The OPLS-DA model based on the proteome showed high discrimination power ($R^2Y = 0.89$) (Figure S5B, Table 2). The most important proteins in this model according to VIP analysis are listed in Figure S5C. The proteins *HRG*, *APCS*, *APOA4*, and *SERPINA6*, along with *APOC2* and A1BG, were identified as discriminative for women with GAH versus CAH. An OPLS-DA model constructed using these features showed a substantial improvement in predictive power (Q^2Y from 0.14 to 0.55) (Table 2).

For the separation of PE from GAH, concentrations of *TTR*, *SERPING1*, *HPX*, *VTN*, *C1RL*, and *APOA2* were statistically significantly higher in the PE group, while concentrations of *F10*, *C8A*, *AHSG*, *C1QA*, *C1QC*, *VCAM1*, and *KNG1* were statistically significantly lower (Figure S6A, Table S11). The OPLS-DA model based on the proteome had high discriminative and predictive power ($R^2Y = 0.97$, $Q^2Y = 0.72$) (Figure S6B, Table 2). The most important serum proteins are shown in Figure S6C. The proteins *TTR*, *SERPING1*, *HPX*, *VTN*, *C1RL*, *F10*, *C8A*, *AHSG*, *C1QA*, *C1QC*, *VCAM1*, and *KNG1* were defined as the discriminative proteins. The OPLS model based on these features maintained high discriminative power ($R^2Y = 0.82$, $Q^2Y = 0.77$) (Table 2).

A consolidated analysis against non-PE groups defined a core first trimester serum protein signature for PE. Concentrations of *F10*, *C8A*, *AHSG*, *C1QA*, *C1QC*, *KNG1*, *FBLN1*, and *C1R* were statistically significantly decreased in PE, while concentrations of *SERPING1*, *IGHG1*, *TTR*, *AFM*, *VTN*, *APOA2*, *HPX*, *RBP4*, *APCS*, and *CNDP1* were statistically significantly increased (Figure 2A, Table S12). The OPLS-DA model based on the serum proteome showed high discriminative and predictive power ($R^2Y = 0.89$, $Q^2Y = 0.79$) (Figure 2B, Table 2). The most important proteins in the model according to VIP value are shown in Figure 2C. The final set of markers for the development of PE was defined as *F10*, *C8A*, *AHSG*, *C1QA*, *C1QC*, *FBLN1*, *KNG1*, *C1R*, *VTN*, *AFM*, *HPX*, *APOA2*, *RBP4*, *APCS*, *CNDP1*, *TTR*, *IGHG1*, and *SERPING1*. An OPLS-DA model based on these markers retained high discriminative and predictive power ($R^2Y = 0.84$, $Q^2Y = 0.80$) (Table 2).

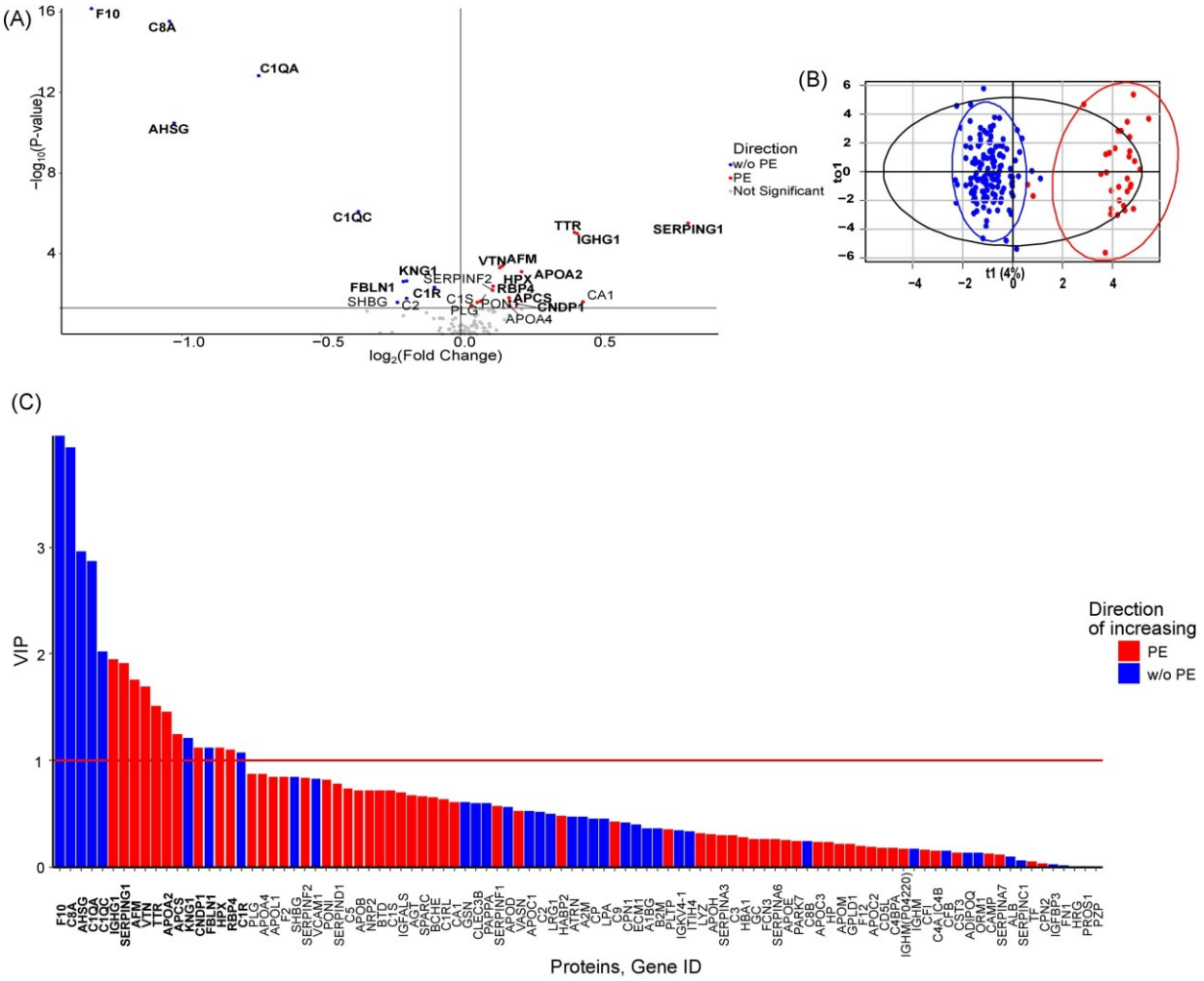


Figure 2. Discovery of first-trimester serum protein features predictive of PE. (**A**) Volcano plot illustrating differential protein abundance between the non-PE ($n = 142$) and PE ($n = 32$) groups. The fold change (PE/non-PE) is shown on the x-axis ($\log_2$ scale), and statistical significance ($-\log_{10}$ *p*-value, Mann–Whitney U test) is on the y-axis. Proteins identified as significant markers are labeled in bold. (**B**) OPLS-DA scores plot demonstrating the separation between the non-PE (blue) and PE (red) cohorts based on their proteomic profiles. (**C**) VIP scores for all proteins in the OPLS-DA model. Proteins with a VIP score > 1 (red line) significantly contribute to the separation and are highlighted in bold.

A final PLS-DA model for discriminating CAH ($n = 24$), GAH ($n = 24$), and PE ($n = 32$) had high descriptive power ($R^2Y = 0.73$) and medium predictive power ($Q^2Y = 0.38$). GAH and CAH samples demonstrated good separation from PE samples but not from each other (Figure 3A). The loadings revealed that proteins significantly decreased ($p < 0.05$) in PE (*F10*, *AHSG*, *C8A*, *C1QA*, *C1QC*, *VCAM1*) were aligned with the "Non-PE" direction, while proteins significantly increased ($p < 0.05$) in PE (*TTR*, *SERPING1*, *IGHG1*, *HPX*, *C1RL*) were aligned opposite (Figure 3B, Table 3). Furthermore, proteins differentiating GAH from CAH ($p < 0.05$), such as *A1BG* (increased in CAH) and *HRG*/*APCS* (increased in GAH), were associated with the vector separating these two conditions (Figure 3, Table 3). Finally, these PE protein features showed significant associations with clinical parameters, correlating with adverse neonatal outcomes on one hand and maternal morbidity on the other (Figure 4).

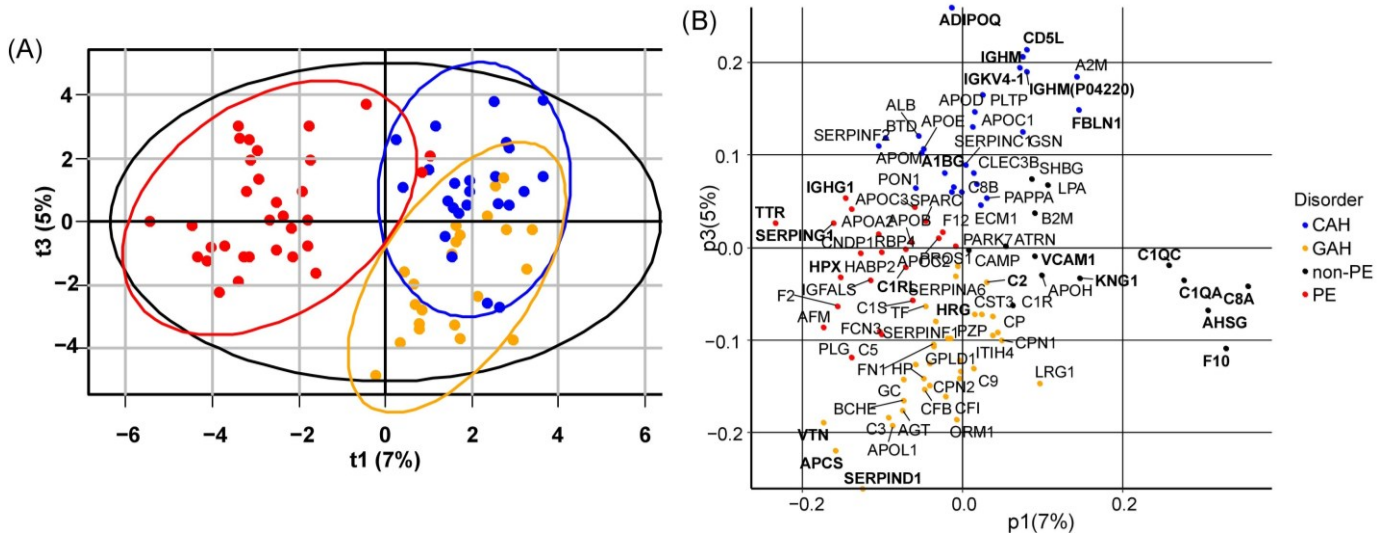


Figure 3. PLS-DA model distinguishing CAH (n = 24), GAH (n = 22), and PE (n = 32) samples. (**A**) Scores plot (components 1 vs. 3) showing sample distribution for CAH (blue), GAH (orange), and PE (red). (**B**) Corresponding loadings plot (components 1 vs. 3). Proteins with a VIP score >1 are displayed in bold. Black labels indicate proteins that specifically differentiate PE from non-PE conditions (CAH and GAH) but do not distinguish between CAH and GAH.

Table 3. The lists serum proteins that demonstrated significant alterations (VIP > 1 and $p < 0.05$) in first trimester identified by PLS-DA models. Arrows denote the direction of protein level's change in at least one comparative group: ↓ decrease; ↑ increase.

Protein Name	Gene Name	PE vs. IUGR	GAH vs. CAH	PE vs. GAH	PE vs. Non-PE
Afamin	*AFM*	-	-	-	↑
Alpha-1B-glycoprotein_VAR_018369	*A1BG*	-	↓	-	-
Alpha-2-HS-glycoprotein	*AHSG*	↓	-	↓	↓
Apolipoprotein A-II	*APOA2*	-	-	-	↑
Apolipoprotein A-IV	*APOA4*	-	↑	-	-
Apolipoprotein C-II	*APOC2*	-	↓	-	-
Beta-Ala-His dipeptidase	*CNDP1*	-	-	-	↑
Coagulation factor X	*F10*	↓	-	↓	↓
Complement C1q subcomponent subunit A	*C1QA*	↓	-	↓	↓
Complement C1q subcomponent subunit C	*C1QC*	↓	-	↓	↓
Complement C1r subcomponent	*C1R*	-	-	-	↓
Complement C1r subcomponent-like protein	*C1RL*	-	-	↑	-
Complement component C8 alpha chain	*C8A*	↓	-	↓	↓
Corticosteroid-binding globulin	*SERPINA6*	-	↑	-	-
Fibulin-1	*FBLN1*	-	-	-	↓
Ficolin-3	*FCN3*	↑	-	-	-
Hemopexin	*HPX*	-	-	↑	↑
Histidine-rich glycoprotein	*HRG*	-	↑	-	-
Ig gamma-1 chain C region	*IGHG1*	↑	-	-	↑
Kininogen-1	*KNG1*	-	-	↓	↓
Plasma protease C1 inhibitor	*SERPING1*	↑	-	↑	↑
Retinol-binding protein 4	*RBP4*	-	-	-	↑
Serum amyloid P-component	*APCS*	-	↑	-	↑
Transthyretin	*TTR*	↑	-	↑	↑
Vascular cell adhesion protein 1	*VCAM1*	-	-	↓	-
Vitronectin	*VTN*	↑	-	↑	↑

https://doi.org/10.3390/ijms27031402

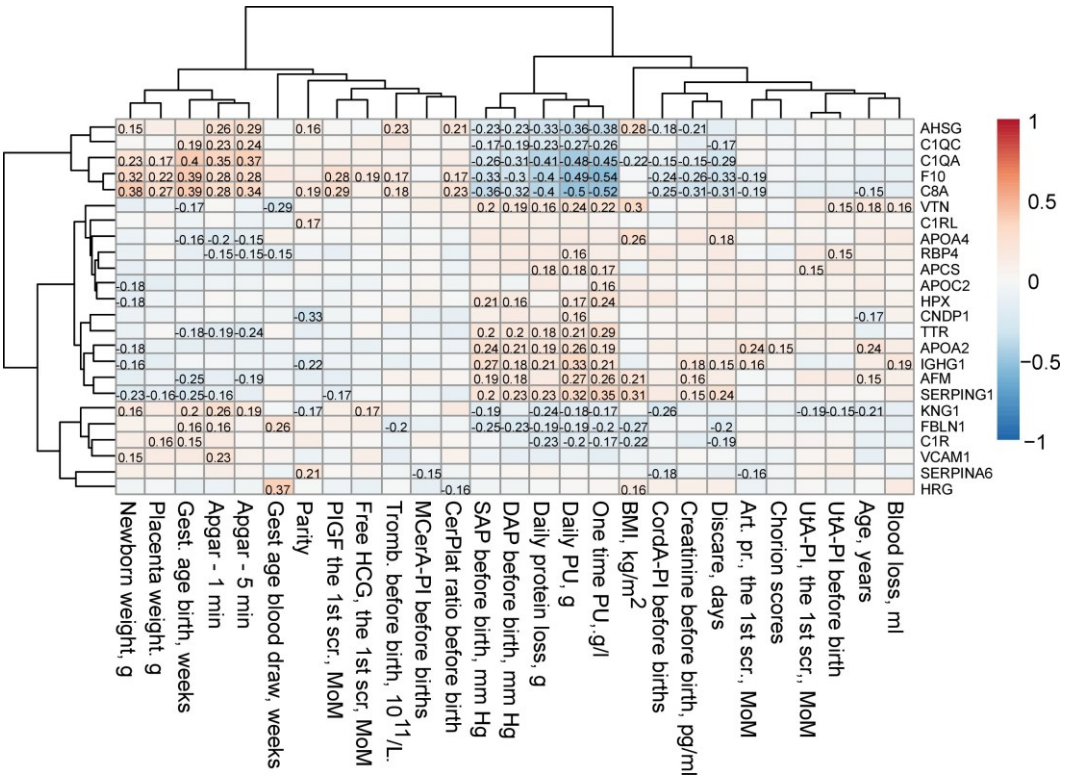


Figure 4. Association of protein markers with maternal, fetal, and neonatal clinical parameters. Cells display statistically significant correlation coefficients ($p < 0.05$).

2.4. Protein–Protein Interaction

The proteomic analysis of first-trimester maternal serum ($n = 172$) revealed significant insights into the molecular pathways associated with subsequent pregnancy complications. A predominant finding across most comparisons was the marked enrichment of pathways related to the complement system and coagulation cascades, underscoring their critical role in early disease pathophysiology. The discriminatory proteins for PE versus isolated IUGR, such as *VTN*, *FCN3*, *SERPING1*, *C8A*, *C1QC*, and *C1QA*, were predominantly localized to blood microparticles and the collagen-containing extracellular matrix (Figure S7A). Pathway analysis revealed a strong associated enrichment for the complement cascade, as detailed in Table S14. Furthermore, proteins including *SERPING1*, *C1QC*, and *C1QA* were linked to a pertussis pathway from the KEGG database, while a broader set comprising *VTN*, *TTR*, *FCN3*, *AHSG*, *SERPING1*, *C8A*, *C1QC*, and *C1QA* significantly enriched the Innate Immune System pathway from Reactome (Figure S7B, Table S13).

A distinct profile emerged for the discrimination between GAH and CAH. All six discriminatory proteins were extracellular, with five found in blood particles and the collagen-containing extracellular matrix (Figure S8A). The significantly enriched pathways were uniquely related to lipid metabolism, as proteins *APOA4* and *APOC2* were the sole drivers of pathways for familial hyperlipidemia and statin inhibition of cholesterol production in the Wikipathways database (Figure S8B, Table S14).

In the comparison of PE against GAH, the molecular signature again emphasized the involvement of blood microparticles and the extracellular matrix (Figure S9A). The pathway analysis confirmed a strong complement system activation, with six of the top ten enriched pathways being related to it (Table S15). This comparison was further distinguished by the significant enrichment of the kinin-kallikrein pathway by *SERPING1* and KNG1, and the intrinsic pathway of fibrin clot formation by *SERPING1*, *F10*, and *KNG1*,

pointing to a specific dysregulation of inflammatory bradykinin signaling and coagulation. Associations with pertussis (*SERPING1*, *C1QC*, *C1QA*) and allograft rejection (*C8A*, *C1QC*, *C1QA*) pathways were also noted (Figure S9B, Table S15).

Finally, the broadest comparison of PE versus non-PE groups reinforced these findings, showing the most robust activation of complement and coagulation pathways. Eleven of the eighteen protein markers were contained in blood microparticles and eight in the collagen-containing extracellular matrix (Figure S10A). Seven of the ten most enriched pathways were associated with the complement system (Table S16). The analysis again confirmed the significant involvement of the intrinsic pathway of fibrin clot formation (*SERPING1*, *F10*, *KNG1*) and pathways for pertussis (*SERPING1*, *C1QC*, *C1QA*, *C1R*) and allograft rejection (*C8A*, *C1QC*, *C1QA*) (Figure S10B, Table S16). The recurrent enrichment of these specific pathways across the PE-specific comparisons suggests that dysregulated immune tolerance, inflammation, and coagulation are central features of preeclampsia's early pathogenesis.

2.5. Classification Models

The list of serum proteins that demonstrated significant alterations (Table 3) in first trimester was used for building discrimination models based on OPLS-DA, Support Vector Machines (SVM, linear, polynomial, radial and sigmoid kernels), XGBoost and multilayer perceptron (MLP). SVM and OPLS-DA models demonstrated robust capability for discriminating between most clinical groups, with accuracies consistently above 90%. The exceptions were the comparisons between CAH and GAH, which were not discriminated as effectively (Table S17, Figure S11).

The sensitivity and specificity of the top-performing models for critical clinical distinctions are detailed in Table 4. Notably, SVM models employing polynomial and radial kernels perfectly distinguished IUGR from PE cases with 100% specificity and 94% sensitivity. The discrimination of PE from non-PE outcomes also achieved this high performance (100% specificity, 94% sensitivity) using similar SVM kernels. For distinguishing GAH from PE, an SVM model with a radial kernel showed balanced performance (95% sensitivity, 94% specificity). In contrast, the OPLS-DA model for differentiating CAH from GAH showed slightly lower, though still high, performance (91% sensitivity, 92% specificity).

Table 4. Diagnostic performance of optimal classification models for pregnancy disorders. Metrics include sensitivity, specificity, accuracy, positive predictive value (PPV), negative predictive value (NPV), and F1-score.

Task	Proteins	Model	Sens.	Spec.	Acc.	PPV	NPV	F_1-Score
IUGR vs. PE	F10, C8A, C1QA, AHSG, C1QC, SERPING1, TTR, IGHG1, VTN, FCN3	SVM, pol. kernel (degree = 3.7, $\gamma = 1.9 \times 10^{-7}$, $coef_0 = -310$) SVM, rad. kernel ($\gamma = 0.02$)	0.94	1	0.95	1	0.85	0.97
CAH vs. GAH	HRG, APCS, APOA4, SERPINA6, APOC2 A1BG	OPLS-DA	0.91	0.92	0.91	0.91	0.92	0.91
GAH vs. PE	TTR, SERPING1, HPX, VTN, C1RL, F10, C8A, AHSG, C1QA, C1QC, VCAM1, KNG1	SVM, rad. kernel ($\gamma = 0.01$)	0.95	0.94	0.94	0.97	0.91	0.96
PE vs. non-PE	F10, C8A, AHSG, C1QA, C1QC, FBLN1, KNG1, C1R, VTN, AFM, HPX, APOA2, RBP4, APCS, CNDP1, TTR, IGHG1, SERPING1	SVM, pol. kernel (degree = 3.6, $\gamma = 1.3 \times 10^{-6}$, $coef_0 = -70.4$) SVM, rad. kernel ($\gamma = 3.8 \times 10^{-3}$)	0.94	1	0.99	1	0.99	0.97

The 18-protein SVM model demonstrated superior diagnostic accuracy compared to the standard FMF first-trimester prenatal screening for PE. The model achieved a sensitivity of 94% and a specificity of 100%, substantially outperforming the standard screen's sensitivity of 53% and specificity of 91% (Figure S12, Table 5).

Notably, the standard screening yielded false positives in 1.2% of controls, 29.2% of chronic hypertension cases, and 22.7% of gestational hypertension cases, while failing to detect 47% of true PE cases. With an estimated PE prevalence of 4%, the FMF-based screening showed a positive predictive value (PPV) of 20% and a negative predictive value (NPV) of 98%. In contrast, the protein-based model achieved a PPV of 100% and an NPV of 99%, underscoring its enhanced reliability for both ruling in and ruling out the condition.

Table 5. Prediction of PE by the protein-based model versus the standard FMF first-trimester screening.

Clinical Group	No PE		PE	
	Proteome-Based Model	FMF	Proteome-Based Model	FMF
Control (*n* = 83)	83	82	0	1
CAH (*n* = 24)	24	17	0	7
GAH (*n* = 22)	22	17	0	5
IUGR (*n* = 11)	11	11	0	0
PE (*n* = 32)	2	15	30	17

The first-trimester serum proteomic profile, comprising 115 proteins (*n* = 172), was successfully validated against an independent pilot cohort profiled with the same LC-MRM-MS platform (BAK-125 kit, MRM Proteomics, Montreal, QC, Canada), which measured 125 proteins (*n* = 40) [21]. The two datasets showed strong analytical concordance, with the 55 overlapping proteins exhibiting closely matched MoM-adjusted concentrations (Figure S13) and intermingled distribution in principal component space (Figure S14A).

This validation confirmed *CNDP1* and *TTR* as common, reproducible protein biomarkers for PE across both cohorts (Figure S14B,C). Furthermore, the diagnostic model developed in the larger cohort refined the predictive performance, increasing sensitivity to 94% (from 87%) and specificity to 100% (from 95%) compared to the earlier model [21].

3. Discussion

This prospective nested case–control study establishes that the first-trimester maternal serum proteome, analyzed via a rigorously validated targeted LC-MRM-MS approach, holds significant potential for the early prediction and etiological differentiation of major obstetric syndromes. We identified and validated distinct protein signatures capable of not only predicting the subsequent development of PE but also effectively distinguishing it from other disorders, including GAH, CAH and isolated IUGR. A key methodological strength was the derivation of clinically adjusted MoM values from a well-characterized reference group of 83 healthy pregnancies, which minimized physiological confounders and enhanced the clinical translatability of the findings [20].

The diagnostic performance of the models developed in this study underscores their clinical potential. The SVM models, particularly those with polynomial and radial kernels, achieved exceptional performance in distinguishing IUGR from PE and, most critically, to identify future PE cases within a general obstetric population. The 18-protein model achieved a sensitivity of 94% and a specificity of 100%, meaning it correctly identified nearly all PE cases while producing no false positives among controls or women with other hypertensive or growth-related conditions. This high specificity is crucial, as it could

prevent unnecessary maternal anxiety and clinical interventions in low-risk women. The 94% sensitivity achieved here represents a substantial improvement over the current FMF screening algorithm, which failed to detect 47% of the PE cases in the study cohort.

This performance must be contextualized within the current landscape of PE prediction. A systematic review by Antwi et al. (2020) of 40 prediction models noted that most combine biomarkers with maternal characteristics, with PAPP-A and PlGF being the most common, and reported AUCs ranging from 0.70 to 0.98 [22]. The established FMF algorithm, which integrates maternal factors, mean arterial pressure, uterine artery Doppler, PAPP-A, and PlGF, achieves a detection rate of approximately 77% for preterm PE at a 10% false-positive rate in the ASPRE trial [23,24]. In contrast, the model developed here, with its superior sensitivity and specificity, outperforms this clinical benchmark.

Contemporary research also emphasizes the utility of accessible clinical data. For example, Chen et al. (2025) developed a pragmatic Hypertensive Disorders of Pregnancy (HDP) prediction model using first-trimester demographic and routine laboratory data (e.g., age, BMI, hemoglobin, lipids), achieving robust AUCs of 0.809 and 0.801 in modeling and validation cohorts, respectively [25]. Similarly, Zhou et al. (2024) developed a machine learning model using parameters like baseline BMI and blood pressure measured at 16–20 weeks, achieving a robust AUC of 0.85 [26]. While highly practical for resource-limited settings, Zhou et al.'s reliance on mid-pregnancy measurements shortens the window for early intervention. Although Chen et al.'s model uses first-trimester data, its foundation in standard clinical variables may not directly capture the underlying molecular pathophysiology of PE as effectively as a targeted proteomic approach [25].

This study's lies in its ability not only to identify high-risk individuals but also to pinpoint the underlying cause of their risk. While conventional models group patients into a broad "high-risk" category, proteomic analysis reveals distinct biological pathways at work within this group. Notably, PE was marked by dysregulation in complement and coagulation cascades, in contrast to GAH and CAH, which were differentiated by alterations in lipid metabolism. These findings suggest that proteomic stratification can classify high-risk women by underlying etiology, enabling more personalized monitoring and management strategies long before clinical symptoms emerge [27].

Methodologically, this study addresses key critiques from Antwi et al., who highlighted frequent suboptimal calibration and validation in existing models [22]. Similarly to the approach of Chen et al., rigorous internal validation was conducted [25]. The validity of this work is further strengthened by the application of a targeted, quantitative LC-MRM-MS platform and the establishment of MoM values, offering a level of analytical rigor and pathophysiological insight that purely clinical models cannot provide.

The most prominent molecular theme identified in the data was early dysregulation of the complement system and coagulation cascades, a feature consistently observed across all PE-focused comparisons. The significant enrichment of pathways like "Complement cascade," "Regulation of Complement cascade," and "Classical antibody-mediated complement activation" as early as the first trimester suggests that aberrant immune activation at the maternal-fetal interface is a pivotal, initiating event in PE pathogenesis. This is corroborated by the enrichment of the "Innate Immune System" and "Allograft rejection" pathways, painting a picture of maladaptive immune response resembling failed fetal tolerance. Concurrently, the involvement of the "Intrinsic Pathway of Fibrin Clot Formation" and the "Kinin-Kallikrein system" links this immune dysregulation to a procoagulant and pro-inflammatory state, potentially driven by proteins like *SERPING1*, *F10*, and *KNG1*. This early co-activation provides a plausible molecular substrate for the subsequent endothelial dysfunction that defines clinical PE [28].

The analysis also successfully delineates disorder-specific proteomic profiles. For instance, comparison between PE and isolated IUGR revealed markedly distinct signatures:

proteins like *F10*, *C8A*, *C1QA*, and *AHSG* were lower in PE, while *VTN*, *FCN3*, and *SERPING1* were elevated. This divergence points to separate pathogenic pathways from the first trimester, with PE exhibiting a more pronounced complement and inflammatory signature. Clinically, this translates to a more severe systemic response in PE, characterized by widespread endothelial activation and antiangiogenic imbalance that leads to maternal organ involvement, in contrast to the primarily fetoplacental restriction seen in isolated IUGR.

Similarly, the comparison between GAH and CAH revealed a profile centered on lipid metabolism, which is entirely distinct from these placental disorders. This confirms that despite their phenotypic similarity, GAH and CAH have different early-pregnancy etiologies. The GAH signature—marked by elevated *HRG*, *APCS*, *APOA4*, and *SERPINA6* and lower *APOC2* and *A1BG*—suggests a pregnancy-induced activation of metabolic and stress-response pathways, rather than a long-standing metabolic syndrome. This dynamic response is posited to occur in women with a latent vascular–metabolic vulnerability, who often present as young, primiparous, and without known risk factors [29]. The physiological stress of gestation unmasks this vulnerability, provoking the observed proteomic shifts [30]. In contrast, women with CAH exhibit a more stabilized profile, likely because pre-pregnancy treatment and ongoing medical surveillance have already attenuated part of their chronic metabolic burden.

The PLS-DA model visually confirmed this biological distinction, showing clear separation of PE from both GAH and CAH, while the latter two hypertensive conditions overlapped. However, the models developed here did not effectively discriminate between the CAH and GAH groups, presumably due to the shared clinical phenotype and the inherent diagnostic overlap between these conditions. The clinical distinction between chronic and gestational hypertension can be ambiguous, particularly in women with unrecognized preexisting hypertension, which may contribute to greater group homogeneity and reduced model separation.

The findings presented here consolidate and refine the existing proteomic landscape of PE, which has historically relied on semi-quantitative methods [14]. A substantial portion (40 out of 115) of the quantified proteins have been previously implicated as PE markers, as confirmed by recent meta-analysis [14], thus validating their involvement in the disease process. Notably, the direction of change for several key markers in this first-trimester cohort is aligns with studies from similar gestational windows. For instance, the observed decrease in AHSG and increase in AFM are consistent with findings from first- and second-trimester plasma studies [31,32]. Similarly, the downregulation of *FBLN1* [33], *SERPING1* [16], *KNG1* [17,33], alongside the upregulation of *RBP4* [33–35], *HPX* [36] and *TTR* [34], matches reports from third-trimester investigations of late-onset PE.

More importantly, critical temporal discrepancies were identified, revealing the dynamic nature of the maternal proteome. For several proteins, the alteration detected in the first trimester was the inverse of that reported at term. This is evident for *APCS* (elevated here versus decreased [33,37]), *F10* (decreased versus elevated [33]), *C1R* (decreased versus elevated [33]), and *APOA2* (decreased versus elevated [37]). These inversions highlight a fundamental principle: the first-trimester proteome likely reflects early, causative pathogenic processes, such as aberrant placentation, while the term proteome captures the downstream effects of systemic endothelial damage [28].

A detailed analysis suggests that inconsistencies in the literature often stem from differences in gestational age, PE subtype, and cohort size. For example, AFM was upregulated in this study and in a larger second-trimester cohort of mixed-onset PE [31], but downregulated in a smaller study of early-onset PE from the same research group [31]. Similar discrepancies exist for *FBLN1* and *VTN* [31–33]. These observations suggest that

PE subtypes may have distinct, time-sensitive proteomic fingerprints. The conflicting reports for markers like *RBP4*, *HPX* and *TTR* across different third-trimester late-onset PE studies [16,34–37] further illustrate the complex interplay of these factors. Despite these complexities, the concordance of core PE markers—including *AFM*, *AHSG*, *FBLN1*, *KNG1*, *SERPING1*, *RBP4*, *TTR*, and *HPX*—with at least one independent study [14] reinforces their validity and central role in PE pathophysiology.

This study has several limitations. First, while the nested case–control design is powerful for discovery, it requires validation in a large, prospective, unselected cohort to confirm generalizability. Second, although women with major comorbidities were excluded, the presence of other common conditions, such as obesity or gestational diabetes diagnosed later in pregnancy, was not systematically excluded or adjusted for, which may influence the findings. Third, the single-center design and specific population may limit the transferability of the proteomic signature and its performance thresholds to other populations with different ethnic backgrounds, healthcare protocols, or baseline risk profiles. Fourth, the sample size, though substantial for a detailed proteomic study, limited the power of some subgroup analyses, particularly for IUGR. Finally, the targeted nature of the proteomic panel, though a strength for validation, confined the analysis to the 115-plex assay, precluding the discovery of novel biomarkers outside this panel. Future studies with discovery-phase proteomics could uncover additional biomarkers.

The immediate priority following this study is the validation of this proteomic signature in a large, prospective, unselected cohort. This is essential to confirm generalizability and define clinical risk thresholds. Concurrently, efforts should focus on translating the 18-protein signature into a simplified, cost-effective assay suitable for high-throughput clinical screening, paving the way for routine implementation.

Looking ahead, the current findings open two transformative paths for research. First, the distinct etiological pathways identified here warrant investigation into personalized prevention strategies, where at-risk women receive interventions tailored to their specific proteomic profile (e.g., complement- vs. lipid-dominant). Second, integrating these proteomic markers with established clinical predictors could create a supremely accurate combined model, while functional studies on key proteins like *AFM*, *AHSG*, *SERPING1*, *F10*, *APOA2* and *APOA4* are needed to advance from biomarker discovery to a mechanistic understanding of disease pathogenesis and novel therapeutic targets.

4. Materials and Methods

4.1. Study Design

This prospective cohort study was conducted at the V. I. Kulakov National Medical Research Center for Obstetrics, Gynecology and Perinatology between January and December 2022. The primary objective was to investigate the association between first-trimester serum protein biomarkers and the subsequent development of hypertensive disorders of pregnancy. The study initially enrolled 1869 Caucasian pregnant women aged 18 to 45 years who attended their routine first-trimester prenatal screening performed according to the FMF protocol at 11^{+2} to 14^{+2} weeks of gestation. Participants were excluded based on the following maternal or fetal conditions: multiple pregnancy; maternal diabetes, cancer, autoimmune disease, decompensated kidney disease, or organ transplant history; maternal congenital heart defects or arrhythmia; acute or exacerbated chronic infections; or confirmed fetal chromosomal abnormalities and congenital malformations. From the initial screening population, a final cohort of 1720 (Figure 1) women with singleton pregnancies was established. From this cohort, a nested case–control study was performed for analytical precision. Case groups were defined by the development of specific obstetric complications: PE (n = 32), GAH (n = 22), CAH (n = 24), and isolated IUGR (n = 11) without PE.

Hypertensive disorders were classified according to the 2021 International Society for the Study of Hypertension in Pregnancy (ISSHP) criteria [38]. CAH was defined as systolic blood pressure ≥ 140 mmHg and/or diastolic blood pressure ≥ 90 mmHg present before pregnancy or first detected before 20 weeks' gestation. GAH was defined as new-onset systolic blood pressure ≥ 140 mmHg and/or diastolic blood pressure ≥ 90 mmHg arising after 20 weeks' gestation in the absence of proteinuria or other maternal organ dysfunction. PE was defined as hypertension after 20 weeks' gestation accompanied by one or more of the following: proteinuria; maternal organ dysfunction (renal, hepatic, hematologic, neurologic); or uteroplacental dysfunction, including fetal growth restriction.

A control group (n = 83) was selected from women who had uncomplicated singleton pregnancies, specifically excluding those who conceived via assisted reproductive technology or developed any hypertensive disorder, gestational diabetes (diagnosed by oral glucose tolerance testing), preterm delivery before 37 weeks, or IUGR (estimated fetal weight < 10th percentile).

Upon enrollment, written informed consent was obtained from all participants. A comprehensive first-trimester assessment was then conducted, which recorded standardized maternal weight, height, and blood pressure [39]. Ultrasonographic evaluation included transabdominal color Doppler measurement of the uterine artery pulsatility index (UtA-PI) [40]. Blood was drawn into Serum Z/9 tubes (Monovette, Sarstedt, Germany) and processed within 2 h of collection. After centrifugation during 20 min at room temperature with 300× *g*, the supernatant was aliquoted in vials. Serum was analyzed for PIGF and PAPP-A using the Delfia Xpress system (PerkinElmer, Shelton, CT, USA) as per the manufacturer's guidelines. Residual serum from these assays was aliquoted in cryo-tubes and stored at −80 °C to facilitate a subsequent targeted proteomic analysis of 115 first-trimester biomarkers [20]. Pregnancies were followed up until delivery. Key clinical characteristics during pregnancy, pregnancy outcomes, and placental morphological features were systematically recorded.

The study protocol was approved by the Institutional Review Board of the National Medical Research Center for Obstetrics, Gynecology and Perinatology (protocol No. 2, dated 9 March 2017). All procedures were performed in accordance with the ethical standards of the Helsinki Declaration and Good Clinical Practice guidelines.

4.2. Sample Preparation

This study utilized a targeted proteomic approach with 115 stable isotope-labeled (SIS) peptides as internal standards and their corresponding 115 unlabeled (NAT) analogues for calibration. All SIS and NAT peptides were synthesized and characterized in the Omics lab at Skoltech according to established protocols [41,42]. The assay was adapted from the commercially available BAK-270 kit (MRM Proteomics, Montreal, QC, Canada) [20,43].

Sample preparation followed standard procedures [20,41,44]. Briefly, 10 µL of serum was used for each sample. Proteins were denatured and reduced in a solution containing 9 M urea, 20 mM dithiothreitol, and 200 mM Tris × HCl (pH 8.0) for 30 min at 37 °C, followed by alkylation with 100 mM iodoacetamide for 30 min in the dark. For tryptic digestion, samples were diluted with 100 mM Tris × HCl (pH 8.0) to reduce the urea concentration below 1 M, incubated with TPCK-trypsin at a 20:1 protein-to-enzyme ratio, and digested for 18 h at 37 °C. The digestion was quenched by adding formic acid (FA) to a final concentration of 1.0% (pH ≤ 2), yielding a peptide concentration of approximately 1 mg/mL.

A uniform aliquot of the SIS peptide mixture was added to all samples—including the 172 clinical samples, calibration standards, and quality controls—at a final concentration of 100 fmol/µL in 10 µL. In contrast, the NAT peptides were added only to the seven-

point calibration curve (levels A–G) at concentrations spanning from the lower limit of quantification (LLOQ, level A) to the upper limit of quantification (HLOQ, level G). Following the addition of SIS peptides, all samples underwent solid-phase extraction (SPE) cleanup. The resulting peptides were reconstituted in 34 µL of 0.1% FA for subsequent LC-MRM-MS analysis.

Robust quality control was implemented using two distinct types of QC samples. The BSA-based QCs (QCA-C), prepared in a surrogate matrix at low, medium, and high concentrations, were used to monitor assay performance. A pooled serum QC (CLP) was included to track batch-to-batch reproducibility. Process blanks were analyzed to assess potential background contamination. All calibration standards and QC samples were processed identically to the clinical serum samples.

4.3. LC-MRM-MS Analysis

LC-MRM-MS analyses were performed in three randomized batches on an ExionLC UHPLC system (ThermoFisher Scientific, Waltham, MA, USA), interfaced with a SCIEX QTRAP 6500+ mass spectrometer (SCIEX, Toronto, ON, Canada). All 177 clinical samples were measured in randomized duplicate. The analytical sequence for each batch was designed to include a complete set of seven-point calibration standards, quality controls (BSA-based QCA-C and pooled serum CLP), and process blanks on every 96-well plate.

Peptide separation was carried out on a C18 column (2.1 × 150 mm, 1.7 µm) using a 53-min linear gradient of 2–45% acetonitrile in 0.1% FA at a constant flow rate of 0.4 mL/min [20,42,43]. The mass spectrometer was operated in positive electrospray ionization mode with a source voltage of 4000 V and temperature of 450 °C. Data acquisition utilized multiply reaction monitoring, with the specific precursor/product ion transitions provided in Table S18.

Robust quality assurance was embedded in the batch design. Each batch commenced with the calibration curve, while QC samples were interspersed at the beginning, middle, and end of the run to assess system stability and quantitative reproducibility throughout the entire analysis.

4.4. Data Preprocessing

Raw mass spectrometry data were processed and quantified using Skyline software (version 20.2.0.343) [45]. Quantification and data quality assessment adhered to the principles outlined in the ICH guidelines for Bioanalytical Method Validation [46]. Calibration curves were constructed from the seven-point standard series (levels A–G) using a $1/x^2$-weighted linear regression model to calculate peptide and corresponding protein concentrations in fmol/µL of plasma.

Assay performance was monitored using calibration standards and QC samples (QCA-C and CPL). The acceptance criteria mandated that the accuracy and precision of calculated concentrations be within ±20% of their theoretical values. A calibration curve was accepted if at least five of its seven points met these criteria. For an entire analytical batch to be considered valid, a minimum of 66% of all QC samples and no less than 90% of all peptide calibration curves were required to pass the acceptance criteria. Proteins with a calibration curve demonstrating an R^2 value greater than 0.99 were classified as "quantified" (Figure 1).

For study samples, measured concentrations below the lower limit of quantification (LLOQ, level A) were imputed with the LLOQ value, while those above the upper limit of quantification (HLOQ, level G) were assigned the HLOQ value. Furthermore, any protein with concentrations outside the quantifiable range (below LLOQ or above HLOQ) in more than 50% of the study samples was excluded from subsequent analysis.

4.5. Statistical Analysis

Data normalization was performed using the RobNorm method [47] to mitigate batch effects and analytical variance, consistent with a previous study of first-trimester serum proteomics in healthy pregnancy [20]. For each protein, the mean CV was calculated across standard and quality control samples between batches. Proteins exhibiting a CV greater than 50% were excluded from subsequent analysis. Furthermore, protein expression values were adjusted using MoM derived from the aforementioned reference study and subsequently adjusted for maternal age, body mass index, parity, fetal sex, the presence of uterine myomas, and gestational age at blood draw, based on models established in the same prior work [20].

The analysis focused on comparing specific pairs of pregnancy complications: PE versus isolated IUGR; GAH versus CAH; GAH versus PE; and PE versus all non-PE cases, which included CAH, GAH, the control group, and the IUGR group. Clinical parameters between each pair were compared using the Mann–Whitney U test for continuous variables and Pearson's chi-square test for categorical variables, with an alteration deemed statistically significant if the p-value was less than 0.05 and the test power exceeded 0.80. For an omnibus comparison across all five groups (Control, CAH, GAH, IUGR, PE), the Kruskal–Wallis test and Pearson's chi-square test were employed, followed by Dunn's test and pairwise chi-square tests with Bonferroni correction for post hoc analysis when the initial p-value was below 0.05.

For each comparative pair, discriminant analysis was conducted using OPLS-DA [48]. Models were considered to have good discrimination and predictive quality with $R^2Y > 0.5$ and $Q^2Y > 0.4$, respectively. Protein levels were also compared between groups using the Mann–Whitney U test. Potential protein markers for a complication were identified by combining a p-value < 0.05 from this test with a variable importance in projection (VIP) score greater than 1 from the OPLS-DA [49] (Figure 1). A new OPLS-DA model was then built using only these marker proteins. For each resulting marker set, protein–protein interaction networks and metabolic pathway enrichment analyses were performed using the STRING resource [50]. Pathways with a false discovery rate (FDR) below 0.01 were considered significantly enriched, and a final enrichment score was calculated as the weighted harmonic mean of the observed-to-expected protein ratio and the −log(FDR).

The proteomic associations between CAH, GAH, and PE were further explored using Partial Least Squares-Discriminant Analysis (PLS-DA) [51], with proteins classified as markers for each condition based on their loadings on latent components, prioritizing those with a VIP > 1. Associations between clinical parameters and the previously identified protein markers were assessed using Spearman's rank correlation with a significance threshold of $p = 0.05$.

Finally, the diagnostic potential of the identified protein markers was evaluated by constructing several classification models, including OPLS-DA, SVM with linear, polynomial, radial, and sigmoid kernels, XGBoost, and a MLP. Hyperparameters for the non-linear SVM kernels, XGBoost, and MLP were optimized using a particle swarm optimization method with 5-times repeated 5-fold cross-validation (Figure 1). Model performance was evaluated through 10-fold cross-validation to determine final sensitivity, specificity, and accuracy. The resulting classifier was then compared directly to the routine FMF-based test. For this comparison, PPV and NPV were calculated using the established global prevalence of PE (4%) [52].

All statistical analyses were performed using R version 4.3.2 [53] within RStudio 2023.09.1 [54]. The analysis leveraged numerous specialized packages including ropls 1.34.0 [55], effsize 0.8.1 [56], dunn.test 1.3.6 [57], jgsbook 1.0.7 [58], Desctools 0.99.60 [59], pwr 1.3-0 [60], xgboost 1.7.8.1 [61], e1071 1.7-16 [62], caret 7.0-1 [63], dplyr 1.1.4 [64] and keras 2.15.0 [65] for statistical modeling and machine learning, while ggplot2 3.5.2 [66],

reshape2 1.4.4 [67], forcats 1.0.0 [68], ggrepel 0.9.6 [69] and pROC 1.18.5 [70] were used for data visualization.

5. Conclusions

In conclusion, this study provides robust evidence that first-trimester maternal serum proteomics can serve as a powerful tool for the early prediction and discrimination between hypertensive pregnancy disorders. The identified protein signatures point to specific underlying biological pathways—primarily dysregulated complement and coagulation systems in PE and lipid metabolism in GAH—that are active long before clinical manifestation. The developed 18-protein model demonstrates a level of diagnostic accuracy that surpasses current screening methods, holding promise for a paradigm shift towards a more precise, etiology-based risk assessment in the first trimester. This would enable targeted monitoring and potentially early, preventative interventions for women at highest risk for preeclampsia, while reassuring those at low risk.

Supplementary Materials: The following supporting information can be downloaded at: https://www.mdpi.com/article/10.3390/ijms27031402/s1.

Author Contributions: Conceptualization, N.S., A.K., A.B. (Alexander Brzhozovskiy), E.N. and G.S.; data curation, N.S., A.P., A.T., E.K. and M.I.; formal analysis, A.B. (Alexander Brzhozovskiy), A.T. and E.K.; funding acquisition, E.N., A.B. (Alexander Brzhozovskiy), N.S. and G.S.; investigation, A.K., A.B. (Anna Bugrova), A.B. (Alexander Brzhozovskiy), M.I., E.K. and N.S.; methodology, A.K., A.B. (Anna Bugrova), A.T., E.K., A.P. and E.N.; project administration, E.N., N.S., A.K., A.B. (Alexander Brzhozovskiy) and G.S.; resources, N.S., G.S., E.N., A.B. (Alexander Brzhozovskiy) and A.K.; software, A.T., A.P., A.B. (Anna Bugrova) and M.I.; supervision, N.S., A.K., A.B. (Alexander Brzhozovskiy), E.N. and G.S.; validation, A.T., M.I., A.P. and E.K.; visualization, A.T., E.K., A.B. (Alexander Brzhozovskiy) and A.B. (Anna Bugrova); writing—original draft, N.S., A.T., M.I., A.B. (Anna Bugrova), A.P. and E.K.; writing—review and editing, A.K., A.B. (Alexander Brzhozovskiy), E.N. and G.S. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the Russian Science Foundation (No. 24-14-00140).

Institutional Review Board Statement: The study was approved by the Ethical Committee of the National Medical Research Center for Obstetrics, Gynecology and Perinatology named after Academician V.I. Kulakov (protocol No. 2, dated 9 March 2017).

Informed Consent Statement: Informed consent was obtained from all subjects involved in the study.

Data Availability Statement: The original contributions presented in this study are included in the article/Supplementary Materials. Further inquiries can be directed to the corresponding author.

Acknowledgments: The authors acknowledge the support of the Laboratory of Mass Spectrometry at Skolkovo Institute of Science and Technology for targeted proteomic analysis of serum blood samples of patients.

Conflicts of Interest: The authors declare no conflicts of interest.

Abbreviations

The following abbreviations are used in this manuscript:

BMI	Body mass index
CAH	Chronical arterial hypertension
CS	Cesarean section
CV	Coefficient of variation
DAP	Diastolic arterial pressure
GAH	Gestational arterial hypertension
HLOQ	The highest limit of quantification
IUGR	Intrauterine growth restriction
LC-MS	Liquid chromatography-mass spectrometry
MRM	Multiple reaction monitoring
LLOQ	The lowest limit of quantification
MAP	Mean arterial pressure
MoM	Multiply of medians
MS	Mass spectrometry
NAT	Natural synthetic proteotypic peptides
OPLS-DA	Orthogonal projections on latent structure discriminant analysis
PE	Preeclampsia
PSO	Particle swarm optimization
SIS	Stable isotope-labeled standards
SVM	Support vector machine
VIP	Variable importance projection
QC	Quality control

References

1. Dimitriadis, E.; Rolnik, D.L.; Zhou, W.; Estrada-Gutierrez, G.; Koga, K.; Francisco, R.P.V.; Whitehead, C.; Hyett, J.; da Silva Costa, F.; Nicolaides, K.; et al. Pre-eclampsia. *Nat. Rev. Dis. Prim.* **2023**, *9*, 8. https://doi.org/10.1038/s41572-023-00417-6.
2. Roberts, J.M.; Hubel, C.A. The Two Stage Model of Preeclampsia: Variations on the Theme. *Placenta* **2009**, *30*, 32–37. https://doi.org/10.1016/j.placenta.2008.11.009.
3. Rolnik, D.L.; Wright, D.; Poon, L.C.; O'Gorman, N.; Syngelaki, A.; de Paco Matallana, C.; Akolekar, R.; Cicero, S.; Janga, D.; Singh, M.; et al. Aspirin versus Placebo in Pregnancies at High Risk for Preterm Preeclampsia. *N. Engl. J. Med.* **2017**, *377*, 613–622. https://doi.org/10.1056/nejmoa1704559.
4. Guerby, P.; Audibert, F.; Johnson, J.-A.; Okun, N.; Giguère, Y.; Forest, J.-C.; Chaillet, N.; Mâsse, B.; Wright, D.; Ghesquiere, L.; et al. Prospective Validation of First-Trimester Screening for Preterm Preeclampsia in Nulliparous Women (PREDICTION Study). *Hypertension* **2024**, *81*, 1574–1582. https://doi.org/10.1161/HYPERTENSIONAHA.123.22584.
5. Riishede, I.; Rode, L.; Sperling, L.; Overgaard, M.; Ravn, J.D.; Sandager, P.; Skov, H.; Wagner, S.R.; Nørgaard, P.; Clausen, T.D.; et al. Pre-eclampsia screening in Denmark (PRESIDE): National validation study. *Ultrasound Obstet. Gynecol.* **2023**, *61*, 682–690. https://doi.org/10.1002/uog.26183.
6. de Castro Rezende, K.B.; Bornia, R.G.; Rolnik, D.L.; Amim, J.; Ladeira, L.P.; Teixeira, V.M.G.; da Cunha, A.J.L.A. Performance of the first-trimester Fetal Medicine Foundation competing risks model for preeclampsia prediction: An external validation study in Brazil. *AJOG Glob. Rep.* **2024**, *4*, 100346. https://doi.org/10.1016/j.xagr.2024.100346.
7. Einig, S.; Monod, C.; Baumann, H.; Butenschön, A.; Engesser-Mussbah, J.; Reina, H.; Schoetzau, A.; Mosimann, B.; Manegold-Brauer, G. Impact of Sonographer Experience, Insonation Angle, and Bladder Filling on Uterine Artery Doppler Measurements in the First Trimester of Pregnancy. *J. Ultrasound Med.* **2024**, *43*, 2375–2383. https://doi.org/10.1002/jum.16572.
8. Yang, Y.; Xie, Y.; Li, M.; Mu, Y.; Chen, P.; Liu, Z.; Wang, Y.; Li, Q.; Li, X.; Dai, L.; et al. Characteristics and fetal outcomes of pregnant women with hypertensive disorders in China: A 9-year national hospital-based cohort study. *BMC Pregnancy Childbirth* **2022**, *22*, 924. https://doi.org/10.1186/s12884-022-05260-3.
9. Leonard, S.A.; Siadat, S.; Main, E.K.; Huybrechts, K.F.; El-Sayed, Y.Y.; Hlatky, M.A.; Atkinson, J.; Sujan, A.; Bateman, B.T. Chronic Hypertension during Pregnancy: Prevalence and Treatment in the United States, 2008–2021. *Hypertension* **2024**, *81*, 1716–1723. https://doi.org/10.1161/HYPERTENSIONAHA.124.22731.

10. Pembe, A.B.; Dwarkanath, P.; Kikula, A.; Raj, J.M.; Perumal, N.; Paulo, H.A.; Rajalakshmi, M.; Duggan, C.P.; Masanja, H.M.; Chopra, N.; et al. Hypertensive disorders of pregnancy and perinatal outcomes: Two prospective cohort studies of nulliparous women in India and Tanzania. *BMJ Glob. Health* **2025**, *10*, e016339. https://doi.org/10.1136/bmjgh-2024-016339.
11. Gunderson, E.P.; Greenberg, M.; Najem, M.; Sun, B.; Alexeeff, S.E.; Alexander, J.; Nguyen-Huynh, M.N.; Roberts, J.M. Severe Maternal Morbidity Associated With Chronic Hypertension, Preeclampsia, and Gestational Hypertension. *JAMA Netw. Open* **2025**, *8*, e2451406. https://doi.org/10.1001/jamanetworkopen.2024.51406.
12. Cohen, Y.; Gutvirtz, G.; Avnon, T.; Sheiner, E. Chronic Hypertension in Pregnancy and Placenta-Mediated Complications Regardless of Preeclampsia. *J. Clin. Med.* **2024**, *13*, 1111. https://doi.org/10.3390/jcm13041111.
13. Tita, A.T.; Szychowski, J.M.; Boggess, K.; Dugoff, L.; Sibai, B.; Lawrence, K.; Hughes, B.L.; Bell, J.; Aagaard, K.; Edwards, R.K.; et al. Treatment for Mild Chronic Hypertension during Pregnancy. *N. Engl. J. Med.* **2022**, *386*, 1781–1792. https://doi.org/10.1056/nejmoa2201295.
14. Starodubtseva, N.; Poluektova, A.; Tokareva, A.; Kukaev, E.; Avdeeva, A.; Rimskaya, E.; Khodzayeva, Z. Proteome-Based Maternal Plasma and Serum Biomarkers for Preeclampsia: A Systematic Review and Meta-Analysis. *Life* **2025**, *15*, 776. https://doi.org/10.3390/life15050776.
15. Than, N.G.; Posta, M.; Györffy, D.; Orosz, L.; Orosz, G.; Rossi, S.W.; Ambrus-Aikelin, G.; Szilágyi, A.; Nagy, S.; Hupuczi, P.; et al. Early pathways, biomarkers, and four distinct molecular subclasses of preeclampsia: The intersection of clinical, pathological, and high-dimensional biology studies. *Placenta* **2022**, *125*, 10–19. https://doi.org/10.1016/j.placenta.2022.03.009.
16. Odenkirk, M.T.; Stratton, K.G.; Gritsenko, M.A.; Bramer, L.M.; Webb-Robertson, B.J.M.; Bloodsworth, K.J.; Weitz, K.K.; Lipton, A.K.; Monroe, M.E.; Ash, J.R.; et al. Unveiling molecular signatures of preeclampsia and gestational diabetes mellitus with multi-omics and innovative cheminformatics visualization tools. *Mol. Omics* **2020**, *16*, 521–532. https://doi.org/10.1039/d0mo00074d.
17. Erez, O.; Romero, R.; Maymon, E.; Chaemsaithong, P.; Done, B.; Pacora, P.; Panaitescu, B.; Chaiworapongsa, T.; Hassan, S.S.; Tarca, A.L. The prediction of late-onset preeclampsia: Results from a longitudinal proteomics study. *PLoS ONE* **2017**, *12*, e0181468. https://doi.org/10.1371/journal.pone.0181468.
18. Zeng, S.; Han, M.; Jiang, M.; Liu, F.; Hu, Y.; Long, Y.; Zhu, C.; Zeng, F.; Gan, Q.; Ye, W.; et al. Serum complement proteomics reveal biomarkers for hypertension disorder of pregnancy and the potential role of Clusterin. *Reprod. Biol. Endocrinol.* **2021**, *19*, 56. https://doi.org/10.1186/s12958-021-00742-z.
19. Jiang, M.; Lash, G.E.; Zeng, S.; Liu, F.; Han, M.; Long, Y.; Cai, M.; Hou, H.; Ning, F.; Hu, Y.; et al. Differential expression of serum proteins before 20 weeks gestation in women with hypertensive disorders of pregnancy: A potential role for SH3BGRL3. *Placenta* **2021**, *104*, 20–30. https://doi.org/10.1016/j.placenta.2020.10.031.
20. Starodubtseva, N.; Tokareva, A.; Kononikhin, A.; Brzhozovskiy, A.; Bugrova, A.; Kukaev, E.; Poluektova, A.; Frankevich, V.; Nikolaev, E.; Sukhikh, G. Multiplexed Quantification of First-Trimester Serum Biomarkers in Healthy Pregnancy. *Int. J. Mol. Sci.* **2025**, *26*, 7970. https://doi.org/10.3390/ijms26167970.
21. Starodubtseva, N.; Tokareva, A.; Kononikhin, A.; Brzhozovskiy, A.; Bugrova, A.; Kukaev, E.; Muminova, K.; Nakhabina, A.; Frankevich, V.E.; Nikolaev, E.; et al. First-Trimester Preeclampsia-Induced Disturbance in Maternal Blood Serum Proteome: A Pilot Study. *Int. J. Mol. Sci.* **2024**, *25*, 10653. https://doi.org/10.3390/ijms251910653.
22. Antwi, E.; Amoakoh-Coleman, M.; Vieira, D.L.; Madhavaram, S.; Koram, K.A.; Grobbee, D.E.; Agyepong, I.A.; Klipstein-Grobusch, K. Systematic review of prediction models for gestational hypertension and preeclampsia. *PLoS ONE* **2020**, *15*, e0230955. https://doi.org/10.1371/journal.pone.0230955.
23. 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.* **2013**, *33*, 8–15. https://doi.org/10.1159/000341264.
24. Rolnik, D.L.; Wright, D.; Poon, L.C.Y.; Syngelaki, A.; O'Gorman, N.; de Paco Matallana, C.; Akolekar, R.; Cicero, S.; Janga, D.; Singh, M.; et al. ASPRE trial: Performance of screening for preterm pre-eclampsia. *Ultrasound Obstet. Gynecol. Off. J. Int. Soc. Ultrasound Obstet. Gynecol.* **2017**, *50*, 492–495. https://doi.org/10.1002/uog.18816. Erratum in *Ultrasound Obstet. Gynecol*. **2017**, *50*, 807. https://doi.org/10.1002/uog.18950.
25. Chen, Y.; Ma, J. The construction and validation of a prediction model of hypertensive disease in pregnancy. *Sci. Rep.* **2025**, *15*, 13406. https://doi.org/10.1038/s41598-025-98416-y.
26. Zhou, L.; Tian, Y.; Su, Z.; Sun, J.Y.; Sun, W. Risk factors and prediction model for new-onset hypertensive disorders of pregnancy: A retrospective cohort study. *Front. Cardiovasc. Med.* **2024**, *11*, 1272779. https://doi.org/10.3389/fcvm.2024.1272779.

27. Romero, R.; Erez, O.; Maymon, E.; Xu, Z.; Pacora, P.; Done, B.; Hassan, S.S.; Tarca, A.L.; Arbor, A.; Lansing, E. The Maternal Plasma Proteome Changes as a Function of Gestational Age in Normal Pregnancy: A Longitudinal Study. *Am. J. Obstet. Gynecol.* **2017**, *155*, 1683–1695. https://doi.org/10.1016/j.ajog.2017.02.037.
28. Jung, E.; Romero, R.; Yeo, L.; Gomez-Lopez, N.; Chaemsaithong, P.; Jaovisidha, A.; Gotsch, F.; Erez, O. The etiology of preeclampsia. *Am. J. Obstet. Gynecol.* **2022**, *226*, S844–S866. https://doi.org/10.1016/j.ajog.2021.11.1356.
29. Pettker, J.E.A.V.M. Clinical Management Guidelines for Obstetrician—Gynecologists. *Obstet. Gynecol.* **2020**, *133*, 168–186.
30. Sibai, B.M. Diagnosis and management of gestational hypertension and preeclampsia. *Obstet. Gynecol.* **2003**, *102*, 181–192. https://doi.org/10.1016/S0029-7844(03)00475-7.
31. Uchida, Y.; Higuchi, T.; Shirota, M.; Kagami, S.; Saigusa, D.; Koshiba, S.; Yasuda, J.; Tamiya, G.; Kuriyama, S.; Kinoshita, K.; et al. Identification and validation of combination plasma biomarker of afamin, fibronectin and sex hormone-binding globulin to predict pre-eclampsia. *Biol. Pharm. Bull.* **2021**, *44*, 804–815. https://doi.org/10.1248/bpb.b20-01043.
32. Kolialexi, A.; Tsangaris, G.T.; Sifakis, S.; Gourgiotis, D.; Katsafadou, A.; Lykoudi, A.; Marmarinos, A.; Mavreli, D.; Pergialiotis, V.; Fexi, D.; et al. Plasma biomarkers for the identification of women at risk for early-onset preeclampsia. *Expert Rev. Proteom.* **2017**, *14*, 269–276. https://doi.org/10.1080/14789450.2017.1291345.
33. Lim, J.H.; Lim, J.M.; Lee, H.M.; Lee, H.J.; Kwak, D.W.; Han, Y.J.; Kim, M.Y.; Jung, S.H.; Kim, Y.R.; Ryu, H.M.; et al. Systematic Proteome Profiling of Maternal Plasma for Development of Preeclampsia Biomarkers. *Mol. Cell Proteom.* **2024**, *23*, 100826. https://doi.org/10.1016/j.mcpro.2024.100826.
34. Atkinson, K.R.; Blumenstein, M.; Black, M.A.; Wu, S.H.; Kasabov, N.; Taylor, R.S.; Cooper, G.J.S.; North, R.A. An altered pattern of circulating apolipoprotein E3 isoforms is implicated in preeclampsia. *J. Lipid Res.* **2009**, *50*, 71–80. https://doi.org/10.1194/jlr.M800296-JLR200.
35. Chen, H.; Aneman, I.; Nikolic, V.; Karadzov Orlic, N.; Mikovic, Z.; Stefanovic, M.; Cakic, Z.; Jovanovic, H.; Town, S.E.L.; Padula, M.P.; et al. Maternal plasma proteome profiling of biomarkers and pathogenic mechanisms of early-onset and late-onset preeclampsia. *Sci. Rep.* **2022**, *12*, 19099. https://doi.org/10.1038/s41598-022-20658-x.
36. Liu, L.Y.; Yang, T.; Ji, J.; Wen, Q.; Morgan, A.A.; Jin, B.; Chen, G.; Lyell, D.J.; Stevenson, D.K.; Ling, X.B.; et al. Integrating multiple "omics" analyses identifies serological protein biomarkers for preeclampsia. *BMC Med.* **2013**, *11*, 236. https://doi.org/10.1186/1741-7015-11-236.
37. Lu, Q.; Liu, C.; Liu, Y.; Zhang, N.; Deng, H.; Zhang, Z. Serum markers of pre-eclampsia identified on proteomics. *J. Obstet. Gynaecol. Res.* **2016**, *42*, 1111–1118. https://doi.org/10.1111/jog.13037.
38. Magee, L.A.; Brown, M.A.; Hall, D.R.; Gupte, S.; Hennessy, A.; Karumanchi, S.A.; Kenny, L.C.; McCarthy, F.; Myers, J.; Poon, L.C.; et al. The 2021 International Society for the Study of Hypertension in Pregnancy classification, diagnosis & management recommendations for international practice. *Pregnancy Hypertens.* **2022**, *27*, 148–169. https://doi.org/10.1016/j.preghy.2021.09.008.
39. Poon, L.C.Y.; Zymeri, N.A.; Zamprakou, A.; Syngelaki, A.; Nicolaides, K.H. Protocol for measurement of mean arterial pressure at 11–13 weeks' gestation. *Fetal Diagn. Ther.* **2012**, *31*, 42–48. https://doi.org/10.1159/000335366.
40. Plasencia, W.; Maiz, N.; Bonino, S.; Kaihura, C.; Nicolaides, K.H. Uterine artery Doppler at 11 + 0 to 13 + 6 weeks in the prediction of preeclampsia. *Ultrasound Obstet. Gynecol.* **2007**, *30*, 742–749. https://doi.org/10.1002/uog.5157.
41. Kononikhin, A.S.; Starodubtseva, N.L.; Brzhozovskiy, A.G.; Tokareva, A.O.; Kashirina, D.N.; Zakharova, N.V.; Bugrova, A.E.; Indeykina, M.I.; Pastushkova, L.K.; Larina, I.M.; et al. Absolute Quantitative Targeted Monitoring of Potential Plasma Protein Biomarkers: A Pilot Study on Healthy Individuals. *Biomedicines* **2024**, *12*, 2403. https://doi.org/10.3390/biomedicines12102403.
42. Bugrova, A.E.; Strelnikova, P.A.; Kononikhin, A.S.; Zakharova, N.V.; Diyachkova, E.O.; Brzhozovskiy, A.G.; Indeykina, M.I.; Kurochkin, I.N.; Averyanov, A. V.; Nikolaev, E.N. Targeted MRM-analysis of plasma proteins in frozen whole blood samples from patients with COVID-19: A retrospective study. *Clin. Chem. Lab. Med.* **2024**, *63*, 448–457. https://doi.org/10.1515/cclm-2024-0800.
43. Bhardwaj, M.; Weigl, K.; Tikk, K.; Holland-Letz, T.; Schrotz-King, P.; Borchers, C.H.; Brenner, H. Multiplex quantitation of 270 plasma protein markers to identify a signature for early detection of colorectal cancer. *Eur. J. Cancer* **2020**, *127*, 30–40. https://doi.org/10.1016/j.ejca.2019.11.021.
44. Gaither, C.; Popp, R.; Mohammed, Y.; Borchers, C.H. Determination of the concentration range for 267 proteins from 21 lots of commercial human plasma using highly multiplexed multiple reaction monitoring mass spectrometry. *Analyst* **2020**, *145*, 3634–3644. https://doi.org/10.1039/c9an01893j.
45. MacLean, B.X.; Pratt, B.S.; Egertson, J.D.; MacCoss, M.J.; Smith, R.D.; Baker, E.S. Using Skyline to Analyze Data-Containing Liquid Chromatography, Ion Mobility Spectrometry, and Mass Spectrometry Dimensions. *J. Am. Soc. Mass Spectrom.* **2018**, *29*, 2182–2188. https://doi.org/10.1007/s13361-018-2028-5.

46. European Medicines Agency. *ICH Guideline M10 on Bioanalytical Method Validation and Study Sample Analysis*; European Medicines Agency: Amsterdam, The Netherlands, 2022; Volume 44.
47. Wang, M.; Jiang, L.; Jian, R.; Chan, J.Y.; Liu, Q.; Snyder, M.P.; Tang, H. RobNorm: Model-based robust normalization method for labeled quantitative mass spectrometry proteomics data. *Bioinformatics* **2021**, *37*, 815–821. https://doi.org/10.1093/bioinformatics/btaa904.
48. Trygg, J.; Wold, S. Orthogonal projections to latent structures (O-PLS). *J. Chemom.* **2002**, *16*, 119–128. https://doi.org/10.1002/cem.695.
49. Galindo-Prieto, B.; Eriksson, L.; Trygg, J. Variable influence on projection (VIP) for orthogonal projections to latent structures (OPLS). *J. Chemom.* **2014**, *28*, 623–632. https://doi.org/10.1002/cem.2627.
50. Szklarczyk, D.; Kirsch, R.; Koutrouli, M.; Nastou, K.; Mehryary, F.; Hachilif, R.; Gable, A.L.; Fang, T.; Doncheva, N.T.; Pyysalo, S.; et al. The STRING database in 2023: Protein-protein association networks and functional enrichment analyses for any sequenced genome of interest. *Nucleic Acids Res.* **2023**, *51*, D638–D646. https://doi.org/10.1093/nar/gkac1000.
51. Wold, S.; Sjöström, M.; Eriksson, L. PLS-regression: A basic tool of chemometrics. *Chemom. Intell. Lab. Syst.* **2001**, *58*, 109–130. https://doi.org/10.1016/S0169-7439(01)00155-1.
52. Vera-Ponce, V.J.; Loayza-Castro, J.A.; Ballena-Caicedo, J.; Valladolid-Sandoval, L.A.M.; Zuzunaga-Montoya, F.E.; Gutierrez De Carrillo, C.I. Global prevalence of preeclampsia, eclampsia, and HELLP syndrome: A systematic review and meta-analysis. *Front. Reprod. Health* **2025**, *7*, 1706009. https://doi.org/10.3389/frph.2025.1706009.
53. R Core Team. R: A Language and Environment for Statistical Computing. Available online: https://www.r-project.org (accessed on 27 January 2026).
54. R Team. *R Studio: Integrated Development for R*; RStudio, Inc.: Boston, MA, USA, 2016.
55. Thévenot, E.A.; Roux, A.; Xu, Y.; Ezan, E.; Junot, C. Analysis of the Human Adult Urinary Metabolome Variations with Age, Body Mass Index, and Gender by Implementing a Comprehensive Workflow for Univariate and OPLS Statistical Analyses. *J. Proteome Res.* **2015**, *14*, 3322–3335. https://doi.org/10.1021/acs.jproteome.5b00354.
56. Torchiano, M. effsize: Efficient Effect Size Computation. Available online: https://cran.r-project.org/package=effsize (accessed on 27 January 2026).
57. Dinno, A. dunn.test: Dunn's Test of Multiple Comparisons Using Rank Sums. Version 1.3.6. Avaliable online: https://CRAN.R-project.org/package=dunn.test (accessed on 27 January 2026).
58. Schlarmann, J. jgsbook: Package of the German Book "Statistik mit R und RStudio" by Joerg grosse Schlarmann. Version 1.0.3. Avaliable online: https://CRAN.R-project.org/package=jgsbook (accessed on 27 January 2026).
59. Signorell, A. DescTools: Tools for Descriptive Statistics. Version 0.99.60. Avaliable online: https://CRAN.R-project.org/package=DescTools (accessed on 27 January 2026).
60. Champely, S. pwr: Basic Functions for Power Analysis. Available online: https://github.com/heliosdrm/pwr (accessed on 27 January 2026).
61. Chen, T.; Guestrin, C. XGBoost: A scalable tree boosting system. In Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, San Francisco, CA, USA, 13–17 August 2016; Volume 13–17, pp. 785–794. https://doi.org/10.1145/2939672.2939785.
62. Meyer, D. Support Vector Machines. The Interface to libsvm in package e1071. *R News*, **2024**, *8*, e107.
63. Kuhn, M. Building predictive models in R using the caret package. *J. Stat. Softw.* **2008**, *28*, 1–26. https://doi.org/10.18637/jss.v028.i05.
64. Wickham, H.; François, R.; Henry, L.; Müller, K.; Vaughan, D. dplyr: A Grammar of Data Manipulation. Version 1.1.4. Avaliable online: https://CRAN.R-project.org/package=dplyr (accessed on 27 January 2026).
65. Kalinowski, T.; Falbe, D.; Allaire, J.; Chollet, F.; RStudio; Google; Tang, Y.; Van Der Bijl, W.; Studer, M.; Keydana, S.; et al. R Interface to "Keras". 2024. Available online: https://keras3.posit.co/index.html (accessed on 27 January 2026).
66. Wickham, H. *Elegant Graphics for Data Analysis: Ggplot2*; Springer: Berlin/Heidelberg, Germany, 2008.
67. Cai, C.; Zhang, Z.; Morales, M.; Wang, Y.; Khafipour, E.; Friel, J. Feeding practice influences gut microbiome composition in very low birth weight preterm infants and the association with oxidative stress: A prospective cohort study. *Free Radic. Biol. Med.* **2019**, *142*, 146–154. https://doi.org/10.1016/j.freeradbiomed.2019.02.032.
68. Wickham, H. forcats: Tools for Working with Categorical Variables (Factors). Version 1.0.1. Avaliable online: https://CRAN.R-project.org/package=forcats (accessed on 27 January 2026).
69. Slowikowski, K. ggrepel: Automatically Position Non-Overlapping Text Labels with "ggplot2". Available online: https://github.com/slowkow/ggrepel (accessed on 27 January 2026).

70. Turck, N.; Vutskits, L.; Sanchez-Pena, P.; Robin, X.; Hainard, A.; Gex-Fabry, M.; Fouda, C.; Bassem, H.; Mueller, M.; Lisacek, F.; et al. pROC: An open-source package for R and S+ to analyze and compare ROC curves. *BMC Bioinform.* **2011**, *8*, 12–77. https://doi.org/10.1007/s00134-009-1641-y.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.