

## **Symbolic regression analysis of interactions between first trimester maternal serum adipokines in pregnancies which develop pre-eclampsia**

Casper Wilstrup<sup>1,2</sup>, Paula L. Hedley<sup>1,3</sup>, Line Rode<sup>4</sup>, Sophie Placing<sup>1</sup>, Karen R. Wøjdemann<sup>5</sup>, Anne-Cathrine Shalmi<sup>6</sup>, Karin Sundberg<sup>7</sup>, Michael Christiansen<sup>1,8</sup>

<sup>1</sup> Department for Congenital Disorders, Statens Serum Institut, Copenhagen, 2300, Denmark

<sup>2</sup> Abzu, Copenhagen, 2200, Denmark

<sup>3</sup> Brazen Bio, Los Angeles, California, USA

<sup>4</sup> Department of Clinical Chemistry, Copenhagen University Hospital Rigshospitalet, Glostrup, 2600, Denmark

<sup>5</sup> Department of Gynecology and Obstetrics, Bornholm Hospital, Rønne, 3700 Denmark

<sup>6</sup> Department of Obstetrics and Gynecology, Nordsjællands Hospital, Hillerød, 3400, Denmark

<sup>7</sup> Center of Fetal Medicine, Department of Obstetrics, Copenhagen University Hospital Rigshospitalet, Copenhagen, 2100, Denmark

<sup>8</sup> Department of Biomedical Sciences, University of Copenhagen, Copenhagen, 2200, Denmark

**Words** (2,167 words); **Tables and figures** (5 figures, 2 tables)

**Running title:** Non-linear interactions in pre-eclampsia

**Corresponding author:**

Michael Christiansen\*, M.D., FRCPath

Chief physician, professor (adj)

Department for Congenital Disorders

Statens Serum Institut

5 Artillerivej DK2300S

Copenhagen, Denmark

Phone: +4532683657. Fax: +4532683878. E-mail: mic@ssi.dk

## Abstract

**Objectives:** Pre-eclampsia (PE) is an important cause of perinatal morbidity and mortality. Despite an elusive pathophysiology, PE has been associated with changes in maternal serum concentrations of adipokines in early pregnancy. We hypothesized, that symbolic regression might identify interactions between adipokines and PE, which may have eluded regression and Bayesian models.

**Methods:** In this nested case-control sub-study, of the Copenhagen First Trimester Screening Study, data regarding maternal weight and serum concentrations of PAPP-A, leptin (Lp), soluble leptin receptor (sLR), adiponectin, and resistin (Re) was available from 423 first trimester pregnancies (gestational week  $10^{+3}$ – $13^{+6}$ ), 126 of which developed PE. Symbolic regression with QLattice/Feyn 2.1 was used to identify models comprising two-interactions between up-to three markers.

**Results:** The optimal mathematical model exhibited a non-linear relation between Re and the combined effect of sLR and Lp. The model was dependent, in a Gaussian way, on the level of Re. The receiver operating characteristic (ROC) curve of the model viz. identification of PE cases in first trimester had an AUC of 0.81 and a modelled DR of 40 % for a FPR of 4 %. Symbolic regression outperformed logistic regression of the same parameters with a ROC with AUC = 0.77, and a DR of 7 % for a 3 % FPR.

**Conclusions:** Symbolic regression identified non-linear interactions between Lp, sLR and Re concentrations in first trimester pregnancy serum of pregnancies which later developed PE. Non-linear interactions suggest new pathophysiological pathways and may help in designing more efficient screening protocols for PE.

**Key words:** leptin; pre-eclampsia; pregnancy trimester, first; resistin; soluble leptin receptor; symbolic regression

## Introduction

Pre-eclampsia (PE) is a frequent (3-5 %) adverse outcome of pregnancy characterized by elevated blood pressure and proteinuria or other organ failure occurring from gestational week 20 [1]. The condition is an important contributor to maternal [2, 3] and perinatal — due to growth disturbance, placental abruption, and iatrogenic premature birth [4] — mortality and morbidity. The prevalence of PE is higher in low- and middle- income countries [5]. PE has been estimated to cause ~100K maternal deaths annually, through the development of symptomatic hypertension, eclampsia, and multiple organ involvement [6, 7]. It is of significance for outcome that high-risk patients are identified, carefully monitored, and offered preventive aspirin treatment, early [4, 8, 9]. PE is also associated with an increased occurrence of metabolic syndrome in both mother and foetus, emphasizing the significance of preventive strategies [1].

The aetiology of PE remains elusive, but several lines of evidence suggest that PE may be caused by defects in early trophoblast differentiation [10, 11]. This, in turn, may lead to impaired placentation [12], impaired trophoblast transformation of the uterine spiral arteries and placental hypoperfusion, which could result in endothelial dysfunction and elicit the maternal syndrome. Most of these processes can be monitored through biochemical markers found in maternal serum that, together with PE-associated demographic or clinical markers [13, 14], may be combined to create risk algorithms, and such algorithms have been constructed by FMF, ACOG and NICE and a number of other organisations [15], and clinically implemented with variable efficiency [16]. A more extensive use of electronic patient records, better clinical phenotyping, and use of machine learning (ML) may improve clinical usefulness of PE screening [17, 18].

A plethora of first trimester biochemical markers for the development of PE have been identified [19, 20], e.g. placenta growth factor [21], ADAM12 [22], PAPP-A [23], resistin (Re) [24], PP13 [25] and free leptin-index (leptin (Lp)/ soluble leptin receptor (sLR) [25, 26]. Mean arterial blood pressure is a frequently suggested marker [15] and Doppler ultrasonography of the uterine artery – in the form of a pulsatility index – has also been reported as a first trimester marker for the subsequent development of PE [27]. However, despite decades of research and an increasing number of first trimester PE markers, PE screening

performance is still poorer than first trimester Down syndrome screening [14, 16, 19]. The association between metabolic syndrome in the mother and PE makes it relevant to assess the role of biomarkers reflecting maternal metabolic status, e.g. maternal weight, leptin, adiponectin, resistin, PAPP-A, in first trimester screening for PE.

All attempts to identify combinations of markers have used variations of linear, log-linear or Gaussian combinations of markers [15, 28]. Frequently based on assumptions about the distribution of markers as a function of gestational age and basic clinical information [14, 16, 19]. Machine learning (ML) models, with their promise of integrating multiple “omics” layers [29] and, potential to revolutionise individual patient treatment [30], have recently been introduced, albeit not yet in diagnostic use [31]. ML procedures are frequently of a “black-box” type, not revealing any underlying organization of parameters allowing for interpretability [32]. However, symbolic regression, exemplified by the QLattice method, works well with small datasets [33, 34], and returns a series of signal paths that are ordered by information content [35]. These signal paths reflect the interaction between the involved parameters [32], thus allowing for pathophysiologically relevant interpretation as well as construction of more efficient screening algorithms. This “white-box” characteristic makes symbolic regression an excellent exploratory research tool [18]. Thus, symbolic regression can be used to solve problems where the mathematical form of the data-generating process cannot be assumed a priori. This is in contrast with the typical regression problem where parameters are fitted to a fixed mathematical model, such as generalized linear models or logistic regression.

We hypothesized that a major reason for the inability to construct satisfactory risk algorithms is a lack of understanding of the complicated interactions between biochemical and clinical markers. Thus, the construction of statistical models based on generalized linear models may overlook the information content to be found in non-linear interactions between markers. In order to test this hypothesis we:

1. Examined whether previously reported metabolic maternal serum markers could be combined into a discriminatory mathematical function using symbolic regression.
2. Tested the screening performance of a symbolic regression based discriminatory function as compared to a log-linear discrimination.

## Materials and Methods

A nested case-control sub-study of the Copenhagen First Trimester Screening Study [36] was performed. Previously, several sub-studies of prenatal screening markers have been published [24-26, 37-40]. Out of 6,441 pregnant women for which serum samples were available; 160 verified PE pregnancies with proteinuria (2.5 %) were identified. Patients that developed PE were identified from the Danish National Patient Register [41, 42], and their files were retrieved and examined carefully for compliance with diagnostic criteria. Furthermore, to participate in the current study all patients had to have all the examined clinical parameters registered, and serum had to be available for the biochemical analysis. One hundred and twenty six PE pregnancies fulfilled these criteria [22]. Controls were selected at random but matched for maternal age, number of previous pregnancies and gestational age at biochemical examination. Gestational age was determined by crown-rump-length (CRL). Gestational age at birth was calculated using CRL and the date of birth [43]. All samples were collected in dry containers and kept at 4° C for a maximum of 48 hours until delivery to Statens Serum Institut. Samples were subsequently stored at – 20° C until analysis. Clinical and demographic information about patients and controls are given in Table 1. The biomarkers PAPP-A, resistin (Re), leptin (Lp), adiponectin and leptin soluble receptor (sLR) were quantitated in maternal serum as previously described [44-48].

As clinical information was collected in 1997 – 2001, we used the pre-2014 diagnostic criteria of PE of The International Society for the Study of Hypertension in Pregnancy (ISSHP) criteria to define PE [49]. Thus, the diagnosis of PE required hypertension, defined as either a systolic blood pressure > 140 mm Hg or a diastolic blood pressure > 90 mm Hg occurring in a previously normotensive woman after 20 weeks of gestation in combination with proteinuria (> 0.3 g per 24 hours or dipstick urine analysis >1+). In the current study we did not differentiate between mild and severe PE, early and late PE, or PE with and without IUGR.

### *Statistical analysis*

The analyses were performed using the QLatice/Feyn v 2.1 (Abzu, Copenhagen, Denmark) [35]. Briefly, the QLatice produces a collection of potential mathematical functions selected based on a categorical cross-

entropy loss function defined by the discrimination between controls and PE pregnancies. Symbolic regression was performed with the restriction that the maximum number of markers and functions was five. The markers included in the analysis were adiponectin, PAPP-A, Lp, Re, sLR and maternal weight. Logistic regression and receiver operating characteristic (ROC) curve creation are features of the QLattice/Feyn program and confidence intervals of modelled screening performance were established with the quantile bootstrap method [50], commonly known as reverse percentile bootstrapping.

### *Ethics*

Blood sampling was performed as part of the Copenhagen First Trimester Screening Study and approved by the Scientific Ethics Committee for Copenhagen and Frederiksberg Counties, (No.(KF) 01-288/97) and The Data Protection Agency [36].

## Results

The symbolic regression identified a model comprising the parameters Lp, sLR, and Re as the model giving the optimal discrimination between PE and controls, neither adiponectin, maternal weight nor PAPP-A improved the discriminatory ability. The model (Figure 1) comprises a Gaussian interaction between Re and Lp followed by an addition of sLR. The discriminatory function is given in Figure 1B. The Gaussian interaction between Re and Lp followed by an addition of sLR is reflected in the nomogram of Re and sLR versus the probability of PE (Figure 2A), where it is clear that increasing sLR decreases the risk of PE, but that the iso-risk level depends on the Re level (see below). The partial derivative risk curves, figures 2B – 2D, demonstrate the Gaussian relation between Re, Lp and risk for PE.

The symbolic regression model, with an AUC of 0.81, outperformed the equivalent logistic regression model (involving the same parameters) with an AUC of 0.77 (Figure 3). Using an estimated prevalence of 4 % for PE in the Danish population we modelled the expected performance of screening strategies for PE using the two models (Table 2) and found that the symbolic regression model offers a DR of 40 % for a FPR of 4 % as compared to a DR of 7 % for a FPR of 3 % for the logistic regression model. Indeed, the symbolic regression model out performs the logistic regression particularly well at low FPRs (Figure 3).

Maternal weight was – despite not being part of the final symbolic regression model – examined as a singular marker, and it was found to be associated (AUC = 0.62) in a Gaussian way with PE (Figure 4).

The three significant contributors to the symbolic regression model, sLR, Lp and Re were studied as singular markers (Figure 5A) and sLR was found to be exponentially related to PE with an AUC of the ROC of 0.76 (Figure 5B); Lp was log-linearly related to PE with an AUC of 0.71 (Figure 5B), whereas Re exhibited a Gaussian relation to PE with an AUC of 0.66 (Figure 5B). For comparison, the same ROC curves were depicted for the logistic regression model of each of the three markers. The AUCs for Lp (AUC = 0.71) and for sLR (AUC = 0.76) were identical to the values obtained in the SR model. This is not surprising, as the relation optimized through the symbolic regression procedure is either a linear or exponential function

(closely approximated by the logistic regression). This in contrast to Re, where the logistic regression results in an AUC of 0.56, as compared to the 0.66 in the SR model respectively (Figure 5). The symbolic regression model is thus superior in defining a discriminatory function for Re as compared to the logistic regression model, because it can identify the Gaussian interaction.



## Discussion

The important finding in this study is that symbolic regression using the QLattice/Feyn procedure defines a non-linear relation between Re, Lp, sLR, and PE (Figure 1) that is superior to models based on logistic regression. The AUC increased from 0.77 using the logistic regression model to 0.81 using the symbolic regression model. More importantly, the symbolic regression model performed markedly better for low values of FDR when used as a screening algorithm (Figure 3 and Table 2). The reason for this difference is explained by the ability for the QLattice model to capture the Gaussian interaction between Re and Lp. Previous ML studies of biomarkers in PE have – in very small samples – obtained AUCs of 0.90 – 0.96, including many more markers [51-53]. In this study we have not, except for maternal weight, included maternal risk factors [15], so it must be expected that a clinical performance will be better. Furthermore, the large majority of our cases were non-severe pregnancies giving birth after gestational week 34 (Table 1), and these cases are more difficult to detect with screening markers [14], why one might speculate that the symbolic regression discrimination function would function much better in detecting early and severe pre-eclampsia. However, the real significance of the symbolic regression approach to defining screening algorithms will need to be assessed in prospective studies including other risk factors, i.e. clinical markers, differentiated socio-economical exposures, and health care policies [54].

Overfitting is a risk of all ML approaches, however the QLattice/Feyn2.1 has been shown empirically to exhibit an extent of overfitting comparable to that of linear regression [34]. Furthermore, the confidence intervals in Table 1 are of comparable size, reflecting that logistic regression and QLattice/feyn2.1 symbolic regression do indeed have similar degrees of overfitting. And a level of overfitting much better than other ML approaches [34].

The fact that Re – an inflammatory marker expressed in adipose tissue [24, 55] - is only a marker for PE for specific values of Lp and sLR – suggest that the Re inflammatory function is quenched in extremely low-weight or extremely high weight individuals. This finding may explain the U-shape of the relation between maternal weight and risk of PE [1]. Previous studies have differed in the Re levels measured in maternal

serum in PE pregnancies [24, 56, 57] , this may – in view of the non-linear interactions described above - possibly be explained by interference from other hormones. The “white-box” nature of symbolic regression may enable a much more detailed understanding of the complicated interplay between different endocrine axes, e.g. the GH – IGF axis [58], the IGF system [59], the placenta-brain-axis [60], in pregnancy. This knowledge may in turn aid in defining new therapeutic targets and screening algorithms.

The demonstration of the potential significance of including non-linear marker interactions should thus lead to a more widespread use of statistical models defined through the “white-box” characteristic of symbolic regression. With the availability of technically and financially sustainable methods to apply symbolic regression to identify non-linear interactions in physiological – and pathophysiological – systems we may well see a revolution in clinical areas where personalized medicine has been very difficult to introduce.

## Highlights

- Capturing non-linear relationships between markers can provide useful insight into physiology and pathophysiology.
- Symbolic regression can be used to produce superior models of the relation between adipokines and PE
- Symbolic regression may help produce more efficient screening algorithms for PE.

## Acknowledgments

We gratefully acknowledge the expert technical assistance of Pia Lind, and Pernilla Rasmussen. We also gratefully acknowledge the financial support of the Danish Medical Research Council, Copenhagen University, The John and Birthe Meyer Foundation, The Ivan Nielsen Foundation, The Else and Mogens Wedell-Wedellsborg Foundation, The Dagmar Marshall Foundation, The Egmont Foundation, The Fetal Medicine Foundation, The Augustinus Foundation, The Gangsted Foundation, The A.P. Møller Foundation, The Mads Clausens Foundation, The Copenhagen Hospital Corporation, SAFE Network of Excellence and Statens Serum Institut. This research has been conducted using the Danish national Biobank resource, supported by the Novo Nordisk Foundation.

## Tables

**Table 1.** Demographic and clinical and paraclinical parameters of PE cases and controls.

<b>Parameter</b>	<b>PE (n=126)</b>	<b>Controls (n=297)</b>	<b><i>p</i>-value</b>
<b>Maternal age (years), mean (<math>\pm</math>SD)</b>	30.6 ( $\pm$ 4.56)	30.3 ( $\pm$ 4.21)	0.631
<b>Maternal weight (kg), median (IQR)</b>	68.0 (62.0-75.0)	62.0 (57.0-70.0)*	<0.0002
<b>Maternal BMI (kg/m<sup>2</sup>), median (IQR)</b>	24.0 (21.7-26.4)	21.7 (20.1-24.0)*	<0.001
<b>Primigravida, n (%)</b>	74 (58.7)	165 (55.6)	0.75
<b>Primipara, n (%)</b>	101 (80.2)	242 (81.5)	0.79
<b>Smoking, n (%)</b>	16 (12.7)	41 (13.8)	0.72
<b>Spontaneous conception, n (%)</b>	120 (95.2)	278 (93.6)	0.29
<b>Spontaneous labor onset, n (%)</b>	32 (25.4)	241 (81.1)*	<0.001
<b>Crown rump length (mm), median (IQR)</b>	67.0 (61.0-72.0)	67.0 (61.0-74.0)	0.85
<b>Nuchal translucency (mm), median (IQR)</b>	1.6 (1.3-1.9)	1.5 (1.3-1.9)	0.51
<b>GA at birth (days), median (IQR)</b>	275.5 (264.0-283.0)	284.0 (278.0-290.0)*	4.49e-06
<b>GA at test (days), median (IQR)</b>	90.5 (87.0 – 93.0)	90.0 (88.0 – 94.0)	1
<b>GA below 34+0 weeks, n (%)</b>	10 (7.9)	0 (0.0)*	<0.001
<b>PAPP-A (mIU/L), median (IQR)</b>	3064 (1952 - 4576)	3874 (2605 - 5938)	0.028
<b>hCG (IU/L), median (IQR)</b>	42.90 (29.60 – 58.90)	41.75 (26.95 – 63.10)	0.975
<b>Leptin (ug/L), median (range)</b>	22.64 (5.00 – 70.14)	14.1 (2.0 – 54.2)*	1.47e-06
<b>Leptin sR (ug/L), median (range)</b>	27.8 (14.3 – 100.9)	36.9 (15.1 – 97.5)*	2.04e-13
<b>Adiponectin (mg/L), median (range)</b>	3.7 (1.6 – 10.2)	4.5 (1.4 – 17.3)*	7.58e-06
<b>Resistin (ug/L), median (range)</b>	21.89 (6.23 – 58.03)	24.38 (4.94 – 82.96)	0.177

**Table 2.** Comparison of discriminatory performance of the markers sLR, Re, and Lp using a symbolic regression versus a logistic regression. PPV was based on a 4 % prevalence of PE. Rates given in percentages.

<b>Symbolic regression QLattice/Feyn v. 2.1.</b>			<b>Logistic regression</b>		
<b>DR (95% CI)</b>	<b>FPR (95% CI)</b>	<b>PPV %</b>	<b>DR (95% CI)</b>	<b>FPR (95% CI)</b>	<b>PPV %</b>
40 (31-48)	4 (1-6)	30	7 (2-11)	3 (1-5)	10
52 (43-60)	11 (7-15)	16	33 (25-41)	6 (4-9)	18
71 (63-79)	26 (21-31)	10	65 (57-74)	25 (20-29)	10

CI: confidence interval; DR: detection rate; FPR: false positive rate; PPV: positive predictive value

## Legends to figures

**Figure 1.** **A.** Model signal path in the symbolic regression model for PE. The numbers above the nodes are the Pearson correlation coefficient with the target after each mathematical operation. **B.** The mathematical equation describing the discriminatory function.

**Figure 2.** **A.** Nomogram depicting the risk of PE as a function of the combined Re and sLR level. **B.** Risk of PE as a function of Re level for different levels of Lp and sLR. **C.** Risk of PE as a function of sLR level for different levels of Re and Lp. **D.** Risk of PE as a function of Lp level for different levels of sLR and Re.

**Figure 3.** Modelled ROC curves of the discriminatory function using symbolic regression and an optimised discriminatory function defined through logistic regression.

**Figure 4.** **A.** Model signal path of maternal weight in the symbolic regression model for PE; **B.** The modelled ROC curve with the regressor function of maternal weight as marker for PE.

**Figure 5.** **A.** Model signal paths in the symbolic model for PE for each of the parameters Lp, sLR, and Re individually. **B.** Modelled ROC curves of the symbolic regression models for each of the parameters. **C.** Modelled ROC curves of the logistic regression models for each of the parameters.

## Reference List

1. Chappell LC, Cluver CA, Kingdom J, Tong S. Pre-eclampsia. *Lancet* 2021;398:341-354.
2. Say L, Chou D, Gemmill A, Tuncalp O, Moller AB, Daniels J, et al. Global causes of maternal death: a WHO systematic analysis. *Lancet Glob Health* 2014;2:e323-33.
3. Zhang J, Meikle S, Trumble A. Severe maternal morbidity associated with hypertensive disorders in pregnancy in the United States. *Hypertens Pregnancy* 2003;22:203-12.
4. von Dadelszen P, Payne B, Li J, Ansermino JM, Broughton Pipkin F, Cote AM, et al. Prediction of adverse maternal outcomes in pre-eclampsia: development and validation of the fullPIERS model. *Lancet* 2011;377:219-27.
5. Pregnancy WHOICSoHDo. Geographic variation in the incidence of hypertension in pregnancy. *American Journal of Obstetrics and Gynecology* 1988;158:80-83.
6. Mol BWJ, Roberts CT, Thangaratnam S, Magee LA, de Groot CJM, Hofmeyr GJ. Pre-eclampsia. *Lancet* 2016;387:999-1011.
7. Duley L. The global impact of pre-eclampsia and eclampsia. *Semin Perinatol* 2009;33:130-7.
8. Poon LC, Rolnik DL, Tan MY, Delgado JL, Tsokaki T, Akolekar R, et al. ASPRE trial: incidence of preterm pre-eclampsia in patients fulfilling ACOG and NICE criteria according to risk by FMF algorithm. *Ultrasound Obstet Gynecol* 2018;51:738-742.
9. Rolnik DL, Wright D, Poon LC, O'Gorman N, Syngelaki A, de Paco Matallana C, et al. Aspirin versus Placebo in Pregnancies at High Risk for Preterm Preeclampsia. *N Engl J Med* 2017;377:613-622.
10. Huppertz B. Placental origins of preeclampsia: challenging the current hypothesis. *Hypertension* 2008;51:970-5.
11. Phipps E, Prasanna D, Brima W, Jim B. Preeclampsia: Updates in Pathogenesis, Definitions, and Guidelines. *Clin J Am Soc Nephrol* 2016;11:1102-13.
12. Fisher SJ. Why is placentation abnormal in preeclampsia? *Am J Obstet Gynecol* 2015;213:S115-22.
13. Bartsch E, Medcalf KE, Park AL, Ray JG, High Risk of Pre-eclampsia Identification G. Clinical risk factors for pre-eclampsia determined in early pregnancy: systematic review and meta-analysis of large cohort studies. *BMJ* 2016;353:i1753.
14. Tan MY, Syngelaki A, Poon LC, Rolnik DL, O'Gorman N, Delgado JL, et al. Screening for pre-eclampsia by maternal factors and biomarkers at 11-13 weeks' gestation. *Ultrasound Obstet Gynecol* 2018;52:186-195.
15. Chaemsathong P, Sahota DS, Poon LC. First trimester preeclampsia screening and prediction. *Am J Obstet Gynecol* 2022;226:S1071-S1097 e2.
16. Tan MY, Wright D, Syngelaki A, Akolekar R, Cicero S, Janga D, et al. Comparison of diagnostic accuracy of early screening for pre-eclampsia by NICE guidelines and a method combining maternal factors and biomarkers: results of SPREE. *Ultrasound Obstet Gynecol* 2018;51:743-750.
17. Myatt L. The prediction of preeclampsia: the way forward. *Am J Obstet Gynecol* 2022;226:S1102-S1107 e8.
18. Christiansen M, Wilstrup, C., Hedley, P.,L. Explainable “White-box” machine learning is the way forward in pre-eclampsia screening. *American Journal of Obstetrics and Gynecology* 2022;(In press).
19. Grill S, Rusterholz C, Zanetti-Dallenbach R, Tercanli S, Holzgreve W, Hahn S, et al. Potential markers of preeclampsia--a review. *Reprod Biol Endocrinol* 2009;7:70.



20. Townsend R, Khalil A, Premakumar Y, Allotey J, Snell KIE, Chan C, et al. Prediction of pre-eclampsia: review of reviews. *Ultrasound Obstet Gynecol* 2019;54:16-27.
21. Akolekar R, Zaragoza E, Poon LC, Pepes S, Nicolaides KH. Maternal serum placental growth factor at 11 + 0 to 13 + 6 weeks of gestation in the prediction of pre-eclampsia. *Ultrasound Obstet Gynecol* 2008;32:732-9.
22. Laigaard J, Sorensen T, Placing S, Holck P, Frohlich C, Wojdemann KR, et al. Reduction of the disintegrin and metalloprotease ADAM12 in preeclampsia. *Obstet Gynecol* 2005;106:144-9.
23. Spencer K, Cowans NJ, Nicolaides KH. Low levels of maternal serum PAPP-A in the first trimester and the risk of pre-eclampsia. *Prenat Diagn* 2008;28:7-10.
24. Christiansen M, Hedley PL, Placing S, Wøjdemann KR, Carlsen AL, Jørgensen JM, et al. Maternal serum resistin is reduced in first trimester preeclampsia pregnancies and is a marker of clinical severity. *Hypertension in pregnancy* 2015;34:422-433.
25. De Villiers CP, Hedley PL, Placing S, Wøjdemann KR, Shalmi A-C, Carlsen AL, et al. Placental protein-13 (PP13) in combination with PAPP-A and free leptin index (fLI) in first trimester maternal serum screening for severe and early preeclampsia. *Clinical Chemistry and Laboratory Medicine (CCLM)* 2018;56:65-74.
26. Hedley PL, Placing S, Wojdemann K, Carlsen AL, Shalmi AC, Sundberg K, et al. Free leptin index and PAPP-A: a first trimester maternal serum screening test for pre-eclampsia. *Prenat Diagn* 2010;30:103-9.
27. Cnossen JS, Morris RK, ter Riet G, Mol BW, van der Post JA, Coomarasamy A, et al. Use of uterine artery Doppler ultrasonography to predict pre-eclampsia and intrauterine growth restriction: a systematic review and bivariable meta-analysis. *CMAJ* 2008;178:701-11.
28. O'Gorman N, Wright D, Poon LC, Rolnik DL, Syngelaki A, de Alvarado M, et al. Multicenter screening for pre-eclampsia by maternal factors and biomarkers at 11-13 weeks' gestation: comparison with NICE guidelines and ACOG recommendations. *Ultrasound Obstet Gynecol* 2017;49:756-760.
29. Perkel JM. Single-cell proteomics takes centre stage. *Nature* 2021;597:580-582.
30. Darcy AM, Louie AK, Roberts LW. Machine Learning and the Profession of Medicine. *JAMA* 2016;315:551-2.
31. Bertini A, Salas R, Chabert S, Sobrevia L, Pardo F. Using Machine Learning to Predict Complications in Pregnancy: A Systematic Review. *Front Bioeng Biotechnol* 2021;9:780389.
32. Christensen NJ, Demharter, S., Machado, M., Pedersen, L., Salvatore, M., Stentoft-Hansen, V., Iglesias, M., T. Identifying interactions in omics data for clinical biomarker discovery. *bioXRiv* 2022.
33. Wenninger S, Kaymakci, C., Wiethe, C. Explainable long-term building energy consumption prediction using QLattice. *Applied Energy* 2022;308.
34. Wilstrup C, Kasak, J. Symbolic regression outperforms other models for small data sets. *arXiv preprint 2104.15147* 2022.
35. Brøløs K, Machado, MV, Cave, C, Kasak, J, Stentoft-Hansen, V, Batanero, VG, Jelen, T, Wilstrup, C. An approach to symbolic regression using Feyn. *arXiv preprint 2104.05417* 2021.
36. Wojdemann KR, Shalmi AC, Christiansen M, Larsen SO, Sundberg K, Brocks V, et al. Improved first-trimester Down syndrome screening performance by lowering the false-positive rate: a prospective study of 9941 low-risk women. *Ultrasound Obstet Gynecol* 2005;25:227-33.

37. Pedersen N, Juul A, Christiansen M, Wøjdemann K, Tabor A. Maternal serum placental growth hormone, but not human placental lactogen or insulin growth factor  $\square$  1, is positively associated with fetal growth in the first half of pregnancy. *Ultrasound in Obstetrics and Gynecology* 2010;36:534-541.
38. Laigaard J, Sørensen T, Placing S, Holck P, Fröhlich C, Wøjdemann KR, et al. Reduction of the disintegrin and metalloprotease ADAM12 in preeclampsia. *Obstetrics & Gynecology* 2005;106:144-149.
39. Laigaard J, Gros Pedersen N, Olesen Larsen S, Hedley PL, Wøjdemann K, Gjerris AC, et al. ADAM12 in first trimester maternal serum from pregnancies conceived by assisted reproduction techniques (ART). *Prenatal Diagnosis: Published in Affiliation With the International Society for Prenatal Diagnosis* 2009;29:628-629.
40. Laigaard J, Larsen SO, Pedersen NG, Hedley PL, Gjerris AC. ADAM 12-S in first trimester: fetal gender, smoking and maternal age influence the maternal serum concentration. *Prenatal diagnosis* 2009;29:525-527.
41. Lyng E, Sandegaard JL, Rebolj M. The Danish National Patient Register. *Scand J Public Health* 2011;39:30-3.
42. Schmidt M, Schmidt SA, Sandegaard JL, Ehrenstein V, Pedersen L, Sorensen HT. The Danish National Patient Registry: a review of content, data quality, and research potential. *Clin Epidemiol* 2015;7:449-90.
43. Robinson HP, Fleming JE. A critical evaluation of sonar "crown-rump length" measurements. *Br J Obstet Gynaecol* 1975;82:702-10.
44. Frithioff  $\square$  Bøjsøe C, Lund MA, Lausten  $\square$  Thomsen U, Hedley PL, Pedersen O, Christiansen M, et al. Leptin, adiponectin, and their ratio as markers of insulin resistance and cardiometabolic risk in childhood obesity. *Pediatric diabetes* 2020;21:194-202.
45. Lausten-Thomsen U, Christiansen M, Fonvig CE, Trier C, Pedersen O, Hansen T, et al. Reference values for serum total adiponectin in healthy non-obese children and adolescents. *Clinica Chimica Acta* 2015;450:11-14.
46. Lausten-Thomsen U, Lund MAV, Frithioff-Bøjsøe C, Hedley PL, Pedersen O, Hansen T, et al. Reference values for leptin/adiponectin ratio in healthy children and adolescents. *Clinica Chimica Acta* 2019;493:123-128.
47. Thagaard IN, Hedley PL, Holm J-C, Lange T, Larsen T, Krebs L, et al. Leptin and Adiponectin as markers for preeclampsia in obese pregnant women, a cohort study. *Pregnancy hypertension* 2019;15:78-83.
48. Thagaard IN, Krebs L, Holm J-C, Lange T, Larsen T, Christiansen M. Adiponectin and leptin as first trimester markers for gestational diabetes mellitus: a cohort study. *Clinical Chemistry and Laboratory Medicine (CCLM)* 2017;55:1805-1812.
49. Brown MA, Lindheimer MD, de Swiet M, Van Assche A, Moutquin JM. The classification and diagnosis of the hypertensive disorders of pregnancy: statement from the International Society for the Study of Hypertension in Pregnancy (ISSHP). *Hypertens Pregnancy* 2001;20:IX-XIV.
50. Davison A, C., Hinkley, D., V. Bootstrap methods and their application (No.1): Cambridge University Press, 1997
51. Guo H, Wang J, Ao W, He Y. SGB-ELM: An Advanced Stochastic Gradient Boosting-Based Ensemble Scheme for Extreme Learning Machine. *Comput Intell Neurosci* 2018;2018:4058403.
52. Guo Z, Yang F, Zhang J, Zhang Z, Li K, Tian Q, et al. Whole-Genome Promoter Profiling of Plasma DNA Exhibits Diagnostic Value for Placenta-Origin Pregnancy Complications. *Adv Sci (Weinh)* 2020;7:1901819.

53. Nair TM. Statistical and artificial neural network-based analysis to understand complexity and heterogeneity in preeclampsia. *Comput Biol Chem* 2018;75:222-230.
54. Scott G, Gillon TE, Pels A, von Dadelszen P, Magee LA. Guidelines-similarities and dissimilarities: a systematic review of international clinical practice guidelines for pregnancy hypertension. *Am J Obstet Gynecol* 2022;226:S1222-S1236.
55. Lausten-Thomsen U, Christiansen M, Hedley PL, Nielsen TRH, Fonvig CE, Pedersen O, et al. Reference values for fasting serum resistin in healthy children and adolescents. *Clinica Chimica Acta* 2017;469:161-165.
56. Seol HJ, Oh MJ, Yeo MK, Kim A, Lee ES, Kim HJ. Comparison of serum levels and the placental expression of resistin between patients with preeclampsia and normal pregnant women. *Hypertens Pregnancy* 2010;29:310-7.
57. Banjac G, Ardalic D, Mihajlovic M, Antonic T, Cabunac P, Zeljkovic A, et al. The role of resistin in early preeclampsia prediction. *Scand J Clin Lab Invest* 2021;81:432-437.
58. Kaur H, Muhlhausler BS, Roberts CT, Gatford KL. The growth hormone-insulin like growth factor axis in pregnancy. *J Endocrinol* 2021.
59. Oxvig C. The role of PAPP-A in the IGF system: location, location, location. *J Cell Commun Signal* 2015;9:177-87.
60. Rosenfeld CS. The placenta-brain-axis. *J Neurosci Res* 2021;99:271-283.

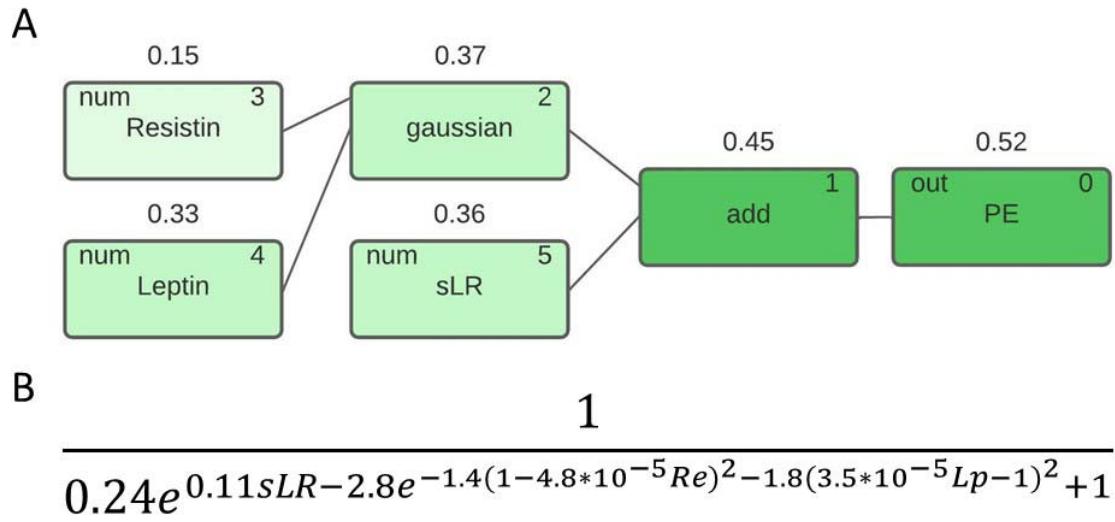


Figure 1

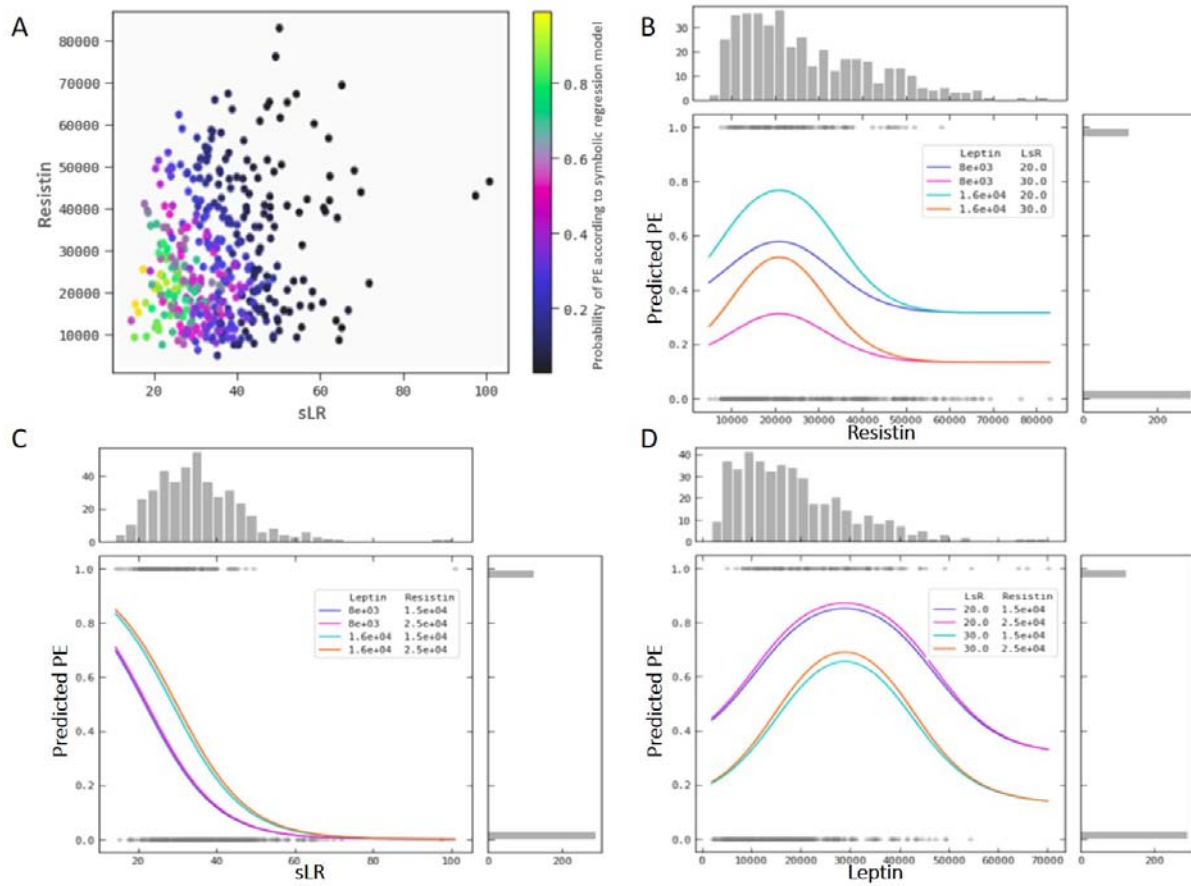


Figure 2

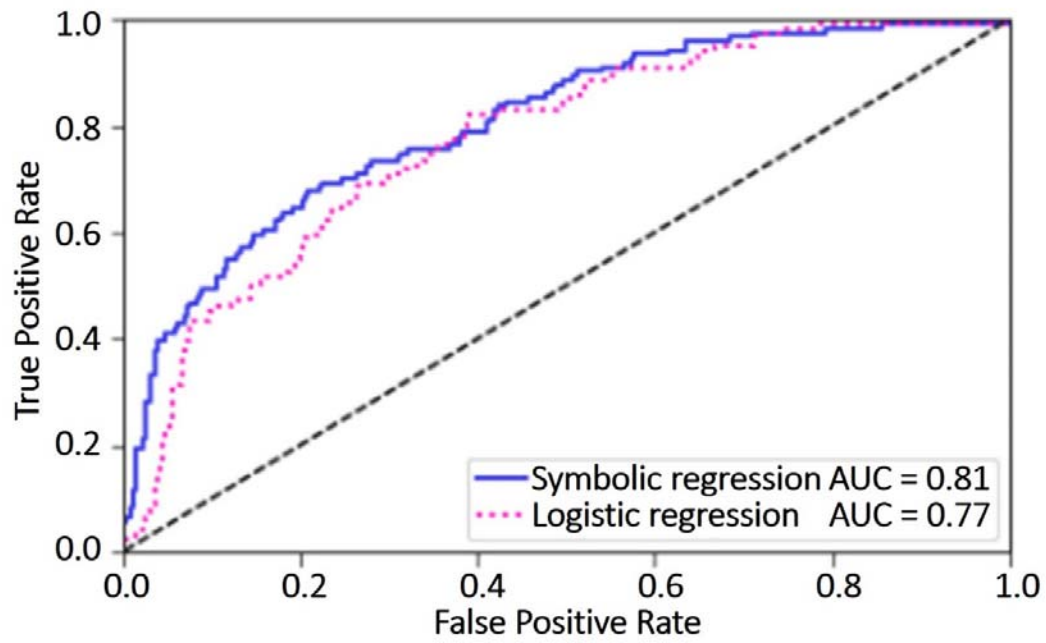


Figure 3

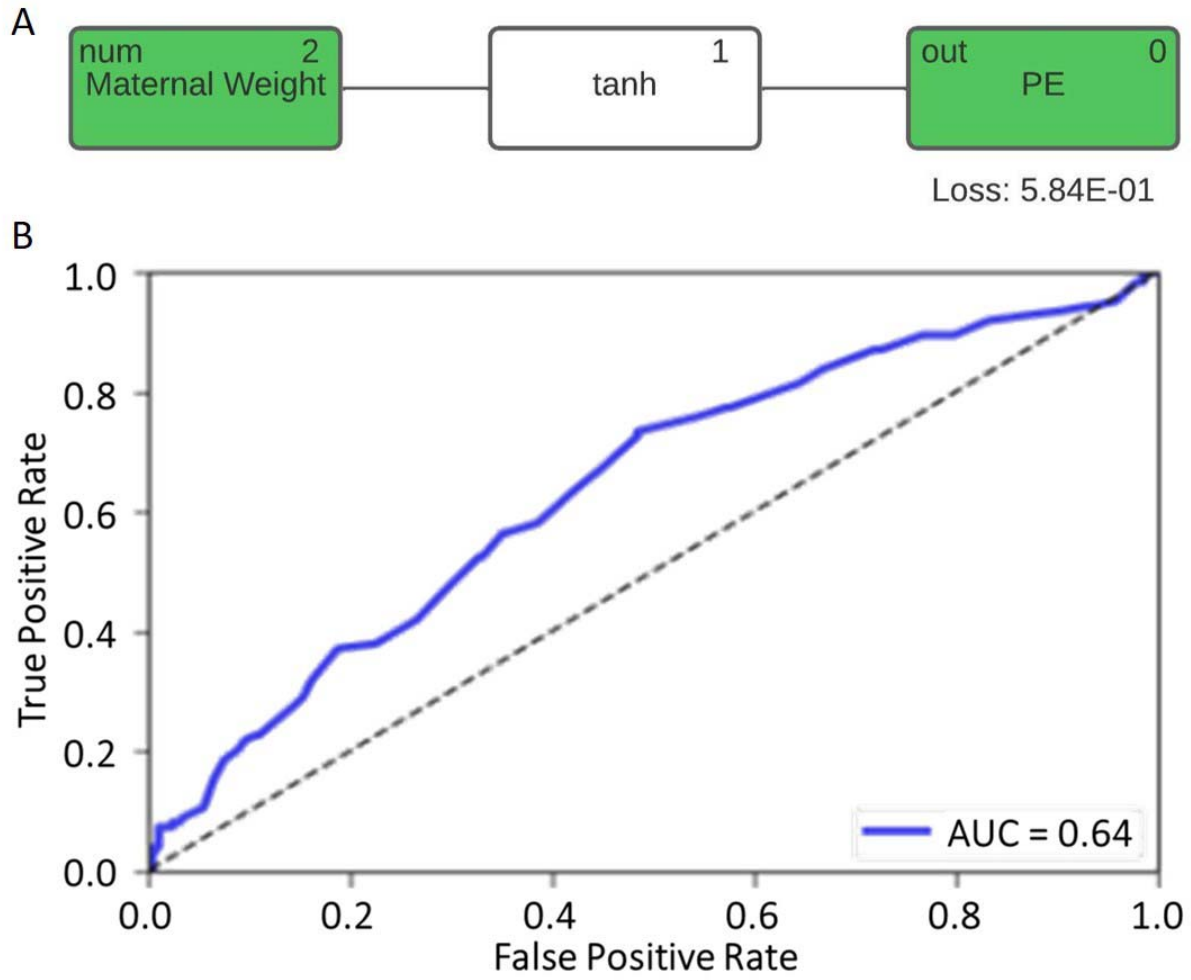


Figure 4

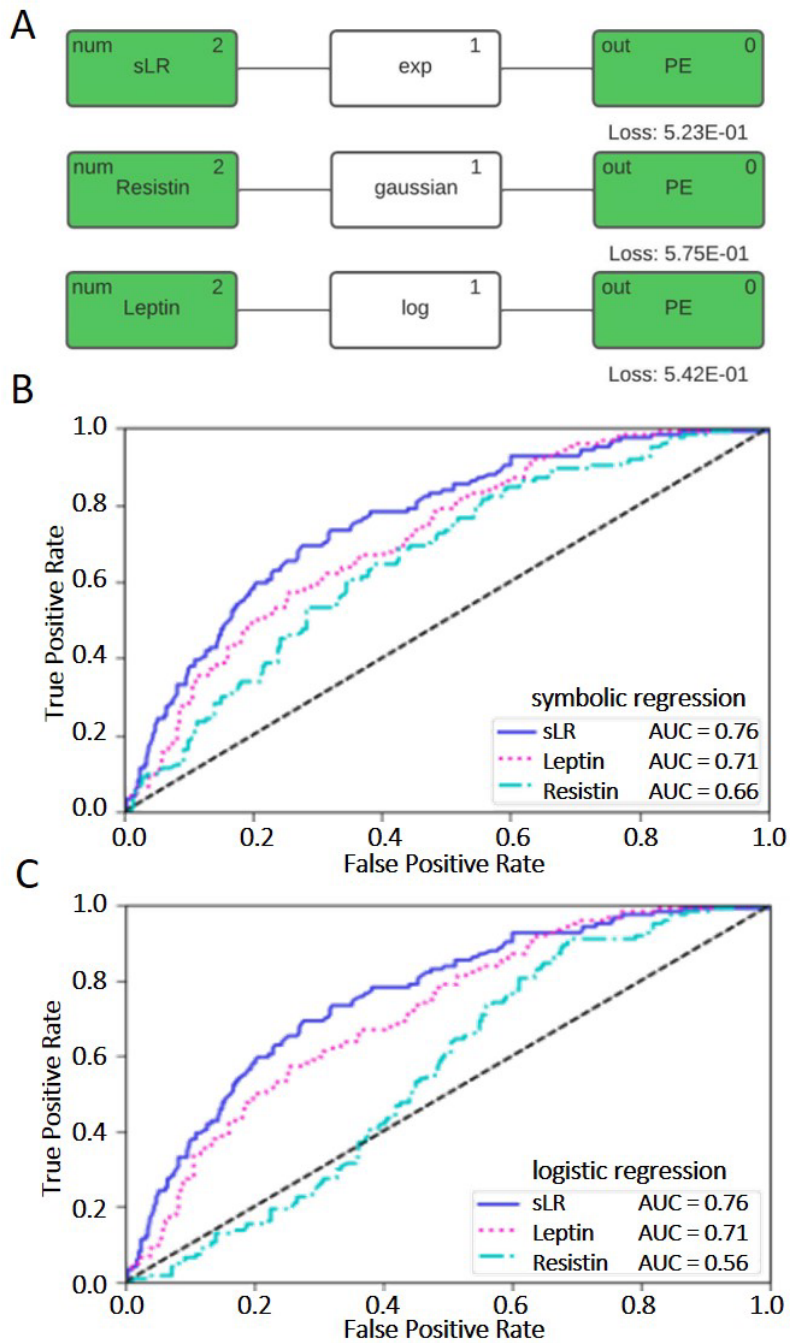


Figure 5