OBSTETRICS

The role of cell-free DNA biomarkers and patient data in the early prediction of preeclampsia: an artificial intelligence model

Asma Khalil, MD; Giovanni Bellesia, PhD; Mary E. Norton, MD; Bo Jacobsson, MD; Sina Haeri, MD; Melissa Egbert, MS; Fergal D. Malone, MD; Ronald J. Wapner, MD; Ashley Roman, MD; Revital Faro, MD; Rajeevi Madankumar, MD; Noel Strong, MD; Robert M. Silver, MD; Nidhi Vohra, MD; Jon Hyett, MD; Cora MacPherson, PhD; Brittany Prigmore, MS; Ebad Ahmed, PhD; Zachary Demko, PhD; J. Bryce Ortiz, PhD; Vivienne Souter, MD; Pe'er Dar, MD

BACKGROUND: Accurate individualized assessment of preeclampsia risk enables the identification of patients most likely to benefit from initiation of low-dose aspirin at 12 to 16 weeks of gestation when there is evidence for its effectiveness, and enables the guidance of appropriate pregnancy care pathways and surveillance.

OBJECTIVE: The primary objective of this study was to evaluate the performance of artificial neural network models for the prediction of preterm preeclampsia (<37 weeks' gestation) using patient characteristics available at the first antenatal visit and data from prenatal cell-free DNA screening. Secondary outcomes were prediction of early-onset preeclampsia (<34 weeks' gestation) and term preeclampsia (≥37 weeks' gestation).

METHODS: This secondary analysis of a prospective, multicenter, observational prenatal cell-free DNA screening study (SMART) included singleton pregnancies with known pregnancy outcomes. Thirteen patient characteristics that are routinely collected at the first prenatal visit and 2 characteristics of cell-free DNA (total cell-free DNA and fetal fraction) were used to develop predictive models for early-onset (<34 weeks), preterm (<37 weeks), and term (≥37 weeks) preeclampsia. For the models, the "reference" classifier was a shallow logistic regression model. We also explored several feedforward (nonlinear) neural network architectures with ≥1 hidden layers, and compared their performance with the logistic regression model. We selected a simple neural network model built with 1 hidden layer and made up of 15 units.

RESULTS: Of the 17,520 participants included in the final analysis, 72 (0.4%) developed early-onset, 251 (1.4%) preterm, and 420 (2.4%) term preeclampsia. Median gestational age at cell-free DNA measurement was 12.6 weeks, and 2155 (12.3%) had their cell-free DNA measurement at ≥16 weeks' gestation. Preeclampsia was associated with higher total cell-free DNA (median, 362.3 vs 339.0 copies/mL cell-free DNA; $P<.001$) and lower fetal fraction (median, 7.5% vs 9.4%; $P<.001$). The expected, cross-validated area under the curve scores for early-onset, preterm, and term preeclampsia were 0.782, 0.801, and 0.712, respectively, for the logistic regression model, and 0.797, 0.800, and 0.713, respectively, for the neural network model. At a screen-positive rate of 15%, sensitivity for preterm preeclampsia was 58.4% (95% confidence interval, 0.569—0.599) for the logistic regression model and 59.3% (95% confidence interval, 0.578—0.608) for the neural network model. The contribution of both total cell-free DNA and fetal fraction to the prediction of term and preterm preeclampsia was negligible. For early-onset preeclampsia, removal of the total cell-free DNA and fetal fraction features from the neural network model was associated with a 6.9% decrease in sensitivity at a 15% screen-positive rate, from 54.9% (95% confidence interval, 52.9—56.9) to 48.0% (95% confidence interval, 45.0—51.0).

CONCLUSION: Routinely available patient characteristics and cell-free DNA markers can be used to predict preeclampsia with performance comparable to that of other patient characteristic models for the prediction of preterm preeclampsia. Logistic regression and neural network models showed similar performance.

Key words: cell-free DNA, early-onset preeclampsia, fetal fraction, linear neural network, noninvasive prenatal screening, nonlinear neural network, pregnancy, preterm preeclampsia, term preeclampsia

Introduction

Preeclampsia is a major contributing factor to maternal morbidity and mortality.[1] Low-dose aspirin started between 12 and 16 weeks of gestation has been shown to decrease the risk of early-onset preeclampsia in high-risk patients.[2] Accurate first-trimester assessment of preeclampsia risk enables identification of patients most likely to benefit from initiation of aspirin before 16 weeks' gestation when there is evidence for its effectiveness. Risk assessment is also important for guiding appropriate pregnancy care pathways and surveillance.

Current United States Preventive Services Task Force guidelines have limited the clinical ability to identify patients at risk for preeclampsia, and more accurate early pregnancy risk assessment has been identified as a priority.[3] Over 100 predictive models for preeclampsia have been reported using patient demographic and clinical characteristics, ultrasound measurements, and biomarkers, most of which have applied logistic regression (LR) or competing risk approaches to modeling.[4–7] However, there is a relative lack of published research on the use of neural networks (NNs) for early pregnancy prediction of preeclampsia[8–15] and few head-to-head comparisons of newer machine learning methods with more traditional approaches such as LR.[13,16,17]

Cite this article as: Khalil A, Bellesia G, Norton ME, et al. The role of cell-free DNA biomarkers and patient data in the early prediction of preeclampsia: an artificial intelligence model. Am J Obstet Gynecol 2024;231:554.e1-18.

0002-9378

https://doi.org/10.1016/j.ajog.2024.02.299

Click Video under article title in Contents at **ajog.org**

AJOG at a Glance

Why was this study conducted?

Preeclampsia risk assessment using patient characteristics alone has poor predictive performance, whereas models using biomarkers and ultrasound have superior performance but require additional resources.

Key findings

Artificial intelligence using a neural network model combining routinely collected clinical data and cell-free DNA markers from noninvasive prenatal screening had a 58% sensitivity for preterm preeclampsia at a 15% screen-positive rate. Predictive performance was similar for early-onset but lower for term preeclampsia. Cell-free DNA variables were associated with a substantial increase in sensitivity for *early-onset* preeclampsia compared with patient characteristics alone, but made a negligible contribution to prediction of preterm or term preeclampsia.

What does this add to what is known?

Routinely available patient characteristics and cell-free DNA markers can be used to predict preeclampsia without the added expense of obtaining biomarkers and ultrasound measurements.

Patient characteristics known to be associated with increased risk for preeclampsia[18] are routinely collected as part of prenatal care. In addition, first-trimester maternal and fetoplacental cell-free DNA (cfDNA) results, which may be associated with increased risk for development of preeclampsia,[19] are frequently available as part of noninvasive prenatal cfDNA screening, which is now the predominant screening test for fetal chromosomal abnormalities in the United States of America.[20]

The primary goal of this study was to develop and evaluate the performance of an artificial NN model for the prediction of preterm preeclampsia (<37 weeks' gestation) using patient characteristics available at the first antenatal visit and data from cfDNA screening. Secondary outcomes were prediction of early-onset preeclampsia (<34 weeks' gestation) and term preeclampsia (≥37 weeks' gestation).

Materials and Methods

This was a secondary analysis of the SNP-based Microdeletion and Aneuploidy RegisTry (SMART) study.[21] SMART was designed to evaluate the performance of single-nucleotide polymorphism (SNP)–based cfDNA screening in a general pregnant population undergoing screening as part of their clinical care. Twenty-one sites in 6 countries participated in the study between April 2015 and January 2019. The study was restricted to singleton pregnancies, ≥9 weeks' gestation, and with maternal age ≥18 years. Data on patient characteristics, cfDNA screening results, serum screening results, ultrasound, and pregnancy outcomes were collected by local research coordinators. Data were managed by the Data Coordinating Center at the Biostatistics Center at the George Washington University, Washington, DC. Ethics committee approval for SMART was obtained by each of the participating sites through their local institutional review boards.

Study population

A total of 20,887 women participated in the SMART study. For this analysis, pregnancies were excluded if there was a fetal chromosomal or major structural abnormality, termination of pregnancy, missing data on pregnancy outcome, participant withdrawal from the study (n=1604), or missing data on preeclampsia status or gestational age at birth (n=365) (Figure 1). Participants were also excluded if they had missing data for any of the variables required for model development (Figure 1). The final cohort used in model development comprised 17,520 participants: 671 with preeclampsia and 16,849 without preeclampsia (Figure 1). Of those with preeclampsia, 72 had early-onset preeclampsia, 251 had preterm preeclampsia, and 420 had term preeclampsia. Preeclampsia is a pregnancy-associated condition that consists of hypertension with or without evidence of organ dysfunction (eg, liver and kidney dysfunction), symptoms (headache and visual disturbance), and in severe cases, seizures ("eclampsia").[22] The presence of "preeclampsia" was based on a clinical diagnosis of preeclampsia documented in the participant's medical record.

Patient characteristics

Patient characteristics routinely collected at the time of enrollment and examined in the analysis included maternal age, body mass index (BMI), maternal height, maternal weight, country, self-reported race and ethnicity, parity, use of in vitro fertilization, chronic hypertension, prepregnancy diabetes, and cigarette smoking during pregnancy. Features obtained from SNP-based cfDNA screening comprised total cfDNA and fetal fraction (FF). Patient characteristics were compared between the groups of patients with and without preeclampsia. Comparison of continuous data was performed using the Wilcoxon rank sum test and comparison of independent categorical data using the chi-square test. *P* values were adjusted for multiple comparisons using the Holm method.

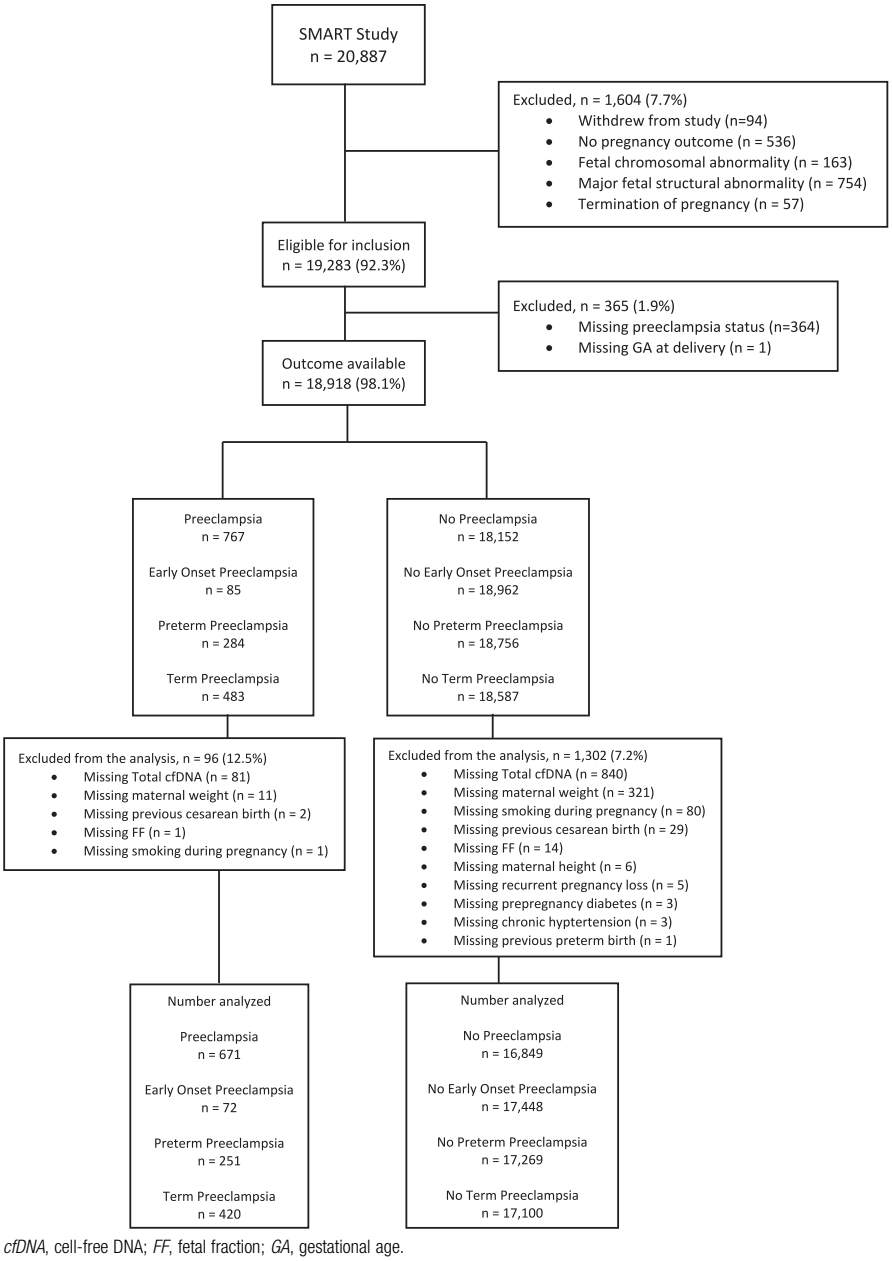
Outcome measures

The primary outcome was preterm preeclampsia defined as preeclampsia with birth <37 weeks' gestation. Secondary outcomes were early-onset preeclampsia (preeclampsia with birth <34 weeks' gestation) and term preeclampsia (preeclampsia with birth ≥37 weeks' gestation). Performance of the artificial intelligence models was also compared with that of LR models.

Features in the model

Variables in the predictive models included 13 patient characteristics available in the SMART study data and

FIGURE 1
Inclusion and exclusion flowchart of patients included in the analyses



cfDNA, cell-free DNA; *FF*, fetal fraction; *GA*, gestational age.

Khalil. Predictive model for preeclampsia using artificial neural networks. Am J Obstet Gynecol 2024.

identified as risk factors for developing preeclampsia based on the existing medical literature (Supplemental Table 1)[18] and 2 features obtained from SNP-based cfDNA screening (total cfDNA and FF). Five of the 15 features were encoded as continuous variables; the remaining 10 features were encoded as binary (Supplemental Table 1). FF was estimated through a probability model using SNPs that were deemed homozygous in the maternal alleles. The probability model used a maximum likelihood approach, based on the observed minor allele frequencies, to determine the FF value that maximizes the likelihood function.[23]

Including race and ethnicity in predictive modeling is controversial and has the potential to perpetuate racial inequities.[24,25] Race and ethnicity were therefore not included in model development. However, given the recent research supporting the importance of race in preeclampsia prediction,[14] a sensitivity analysis was performed to evaluate the effect of including race and ethnicity on the predictive accuracy of the model. Self-identified race and ethnicity information was added as a binary flag to the feature set in Supplemental Table 1. Four groups (Asian, Black, Latina, and White) were modeled separately using the NN model development and the cross-validation protocol described later in this Methods section. Each group was modeled one at a time, separately from other groups. To model a given ethnicity group, an additional binary feature was added to the model, with its value set to "1" for samples reported as being from the group in question and "0" otherwise. Area under the curve (AUC) was calculated for this augmented model and compared with the AUC for the standard model without race and ethnicity information.

Machine learning algorithms for classification

Separate networks were trained for the primary outcome (classifying preterm preeclampsia with respect to no preeclampsia) and the secondary outcomes (classifying early-onset preeclampsia with respect to no preeclampsia and term preeclampsia with respect to no preeclampsia). First, we considered an NN model equivalent to LR as a "reference classifier." The LR model is arguably the simplest NN architecture and directly connects the input features to a single sigmoid output neuron.

In addition, we explored a number of feedforward, dense, NN models incrementally adding $\geq$1 hidden layers of rectified linear units (ReLu)[26,27] to the LR model. The performance of the classifiers was evaluated on a number of metrics: AUC score, sensitivity, positive predictive value (PPV), and negative predictive value (NPV) at different values for the screen-positive rate metric.

All NN models were implemented and optimized in *keras.tensorflow*[28] using

TABLE 1
Baseline demographic and clinical characteristics according to presence of preeclampsia for participants with complete data (modeling cohort)

Characteristics	Preeclampsia	No preeclampsia	*P* value	Adjusted *P* value
n (%)	671 (3.8%)	16,849 (96.2%)		
Maternal age (y)			.022	.109
Median (25%ile, 75%ile)	34.3 (29, 37.9)	34.7 (30.7, 37.5)		
<20	14 (2.1%)	215 (1.3%)		
20–34	309 (46.1%)	7317 (43.4%)		
≥35	301 (44.9%)	8095 (48.0%)		
Body mass index (kg/m^2)			<.001	<.001
Median (25%ile, 75%ile)	28.7 (24.3, 33.5)	24.7 (22.2, 28.5)		
Underweight <18.5	8 (1.2%)	273 (1.6%)		
Normal 18.5–24.9	184 (27.4%)	8459 (50.2%)		
Overweight 25.0–29.9	197 (29.4%)	4720 (28.0%)		
Class I obesity 30.0–34.9	138 (20.6%)	1929 (11.4%)		
Class II obesity 35.0–39.9	67 (10%)	783 (4.6%)		
Class III obesity ≥40.0	69 (10.3%)	448 (2.7%)		
Height (cm)			<.001	<.001
Median (25%ile, 75%ile)	162.6 (157.5, 167.6)	165 (160, 170)		
Weight (kg)			<.001	<.001
Median (25%ile, 75%ile)	73.9 (63.7, 89.3)	66.7 (59.4, 76.2)		
Country			<.001	<.001
United States	502 (74.8%)	8620 (51.2%)		
Sweden	81 (12.1%)	3056 (18.1%)		
Ireland	70 (10.4%)	3875 (23%)		
England	8 (1.2%)	625 (3.7%)		
Australia	8 (1.2%)	422 (2.5%)		
Spain	2 (0.3%)	251 (1.5%)		
Race/ethnicity			<.001	<.001
Asian	51 (7.6%)	1441 (8.6%)		
Black	135 (20.1%)	1253 (7.4%)		
White	274 (40.8%)	10,939 (64.9%)		
Latina	187 (27.9%)	2665 (15.8%)		
Other	15 (2.2%)	376 (2.2%)		
Unknown/not reported	9 (1.3%)	175 (1%)		
Parity			<.001	<.001
Nulliparous	373 (55.6%)	7301 (43.3%)		
Multiparous	298 (44.4%)	9548 (56.7%)		
Previous preterm birth	80 (26.8%)	759 (7.9%)	<.001	<.001
Previous stillbirth	21 (7%)	269 (2.8%)	.023	.109
Previous cesarean delivery	134 (45%)	2849 (29.8%)	<.001	.002

Khalil. Predictive model for preeclampsia using artificial neural networks. Am J Obstet Gynecol 2024. *(continued)*

TABLE 1
Baseline demographic and clinical characteristics according to presence of preeclampsia for participants with complete data (modeling cohort) *(continued)*

Characteristics	Preeclampsia	No preeclampsia	*P* value	Adjusted *P* value
Recurrent pregnancy loss (≥3 before 20 wk)	20 (3%)	558 (3.3%)	.718	1
In vitro fertilization			.374	1
Yes	42 (6.3%)	908 (5.4%)		
Prepregnancy health				
Chronic hypertension	125 (18.6%)	604 (3.6%)	<.001	<.001
Prepregnancy diabetes	43 (6.4%)	155 (0.9%)	<.001	<.001
Cigarette smoking during pregnancy	13 (1.9%)	306 (1.8%)	.934	1
Fetal fraction			<.001	<.001
Median (25%ile, 75%ile)	7.5 (5.7, 10.6)	9.4 (7, 12.3)		
Fetal fraction percentile			<.001	<.001
Median (25%ile, 75%ile)	45.66 (22.90, 73.33)	56.62 (31.21, 78.26)		
Total cfDNA copies/mL			<.001	<.001
Median (25%ile, 75%ile)	362.3 (290.2, 463)	339 (275.8, 422.7)		
Fetal cfDNA copies/mL			<.001	<.001
Median (25%ile, 75%ile)	28.4 (20.7, 40.6)	31.4 (23.7, 42)		
Maternal cfDNA copies/mL			<.001	<.001
Median (25%ile, 75%ile)	333 (261.3, 431.3)	305.4 (244.6, 386.4)		
Serum screening			.004	.025
Yes	238 (35.5%)	5086 (30.2%)		
Gestational age at cfDNA screening (wk)			<.001	<.001
Median (25%ile, 75%ile)	12.7 (11.9, 14.6)	12.6 (11.4, 13.7)		

cfDNA, cell-free DNA.

**P* values were adjusted for multiple comparisons using the Holm method.

Khalil. Predictive model for preeclampsia using artificial neural networks. Am J Obstet Gynecol 2024.

the *binary cross entropy* as the loss function.[27] The network's parameters were optimized via stochastic gradient descent, with a constant learning rate of 1e−2 and zero momentum.[27]

Training and internal validation

The training and validation process consisted of a series of steps that are listed and explained below.

Step 1: stratified split

The data set was randomly split into training and validation data sets of proportions 0.9 and 0.1, respectively. The random split was also stratified with regard to the outcome.

Step 2: dealing with an unbalanced data set

Of the study population, only 1.4% developed preterm preeclampsia. Hence, any machine learning model applied to such a data set would typically overpredict the majority class (patients without preterm preeclampsia) and be quite inefficient in predicting the "rare" events associated with the minority class (patients who developed preterm preeclampsia). To address this, oversampling was applied[27] to the positive samples in the training set, after the train and validation split, via uniform random sampling with replacement, yielding a training set with an equal number of positive and negative samples.

Step 3: neural network scaling and noise injection

A standard scaler was fitted against the oversampled training data set to appropriately scale the train and validation data sets. Gaussian noise with mean equal to zero and standard deviation equal to 1e−1 was added to the standardized training data set. The standard scaler and the Gaussian noise were applied only to the continuous input features. Noise injection was used as a form of both regularization[29] and data augmentation.[27]

Step 4: fit (training data set)

The fitting procedure was then performed on 100 epochs and a minibatch size of 32 using stochastic gradient descent with a constant learning rate of 1e−2.

Step 5: calculation of relevant metrics on the validation data set

Relevant performance metrics were calculated on the validation data set. Specifically, the predictive performance of the 2 classifiers was evaluated by calculating the AUC in addition to the sensitivity, PPV, and NPV for fixed screen-positive rates.

Internal validation

To minimize the bias intrinsic to a single instance of the random split in Step 1, Steps 1 through 4 were repeated 200 times. This Monte Carlo cross-validation[30] generated 200 estimations for each of the performance metrics of interest, corresponding to 200 different random training−validation splits. As a result, the collected metrics are de facto stochastic variables and can be analyzed as such. For each metric, the mean value and its 95% confidence interval (CI) were considered to evaluate the "expected behavior" of the predictive models.

Network regularization

In addition to the noise injection described above, regularization was also introduced in our models via the *kernel_regularizer*[28] using an *L2* penalty whose magnitude varied between 0.0 and 1e−2.

Hyperparameter optimization

To evaluate for overfitting, an incremental, although partial, grid-based hyperparameter search was performed, first considering and fixing the number of epochs and the batch size. Second, we explored the number and size of hidden layers and the following 3 regularization methods: L2 penalty magnitude in kernel regularization, percentage in layer dropout, and standard deviation in Gaussian noise injection in the training data set.[27]

TABLE 2
Diagnostic accuracy metrics for model performance

Early-onset preeclampsia	Logistic regression (LR)	Neural network (NN)
AUC score (95% CI)	0.782 (0.768−0.796)	0.797 (0.783−0.811)
Sensitivity (95% CI)		
10% screen-positive rate	0.475 (0.445−0.505)	0.486 (0.456−0.516)
15% screen-positive rate	0.545 (0.525−0.565)	0.549 (0.529−0.569)
20% screen-positive rate	0.621 (0.601−0.641)	0.651 (0.631−0.671)
Specificity (95% CI)		
10% screen-positive rate	0.902 (0.897−0.906)	0.902 (0.897−0.906)
15% screen-positive rate	0.852 (0.846−0.857)	0.852 (0.846−0.857)
20% screen-positive rate	0.802 (0.796−0.808)	0.802 (0.796−0.808)
PPV (95% CI)		
10% screen-positive rate	0.020 (0.014−0.028)	0.020 (0.014−0.028)
15% screen-positive rate	0.015 (0.011−0.020)	0.015 (0.011−0.020)
20% screen-positive rate	0.013 (0.009−0.017)	0.013 (0.010−0.018)
NPV (95% CI)		
10% screen-positive rate	0.998 (0.997−0.998)	0.998 (0.997−0.998)
15% screen-positive rate	0.998 (0.997−0.998)	0.998 (0.997−0.998)
20% screen-positive rate	0.998 (0.997−0.998)	0.998 (0.997−0.998)
Likelihood ratios (+, −)		
10% screen-positive rate	(4.85, 0.58)	(4.96, 0.57)
15% screen-positive rate	(3.68, 0.53)	(3.71, 0.53)
20% screen-positive rate	(3.14, 0.47)	(3.29, 0.44)
Preterm preeclampsia	**Logistic regression (LR)**	**Neural network (NN)**
AUC score (95% CI)	0.801 (0.795−0.807)	0.800 (0.794−0.806)
Sensitivity (95% CI)		
10% screen-positive rate	0.506 (0.491−0.521)	0.503 (0.488−0.518)
15% screen-positive rate	0.584 (0.569−0.599)	0.593 (0.578−0.608)
20% screen-positive rate	0.669 (0.654−0.684)	0.677 (0.662−0.692)
Specificity (95% CI)		
10% screen-positive rate	0.906 (0.901−0.910)	0.906 (0.901−0.910)
15% screen-positive rate	0.856 (0.851−0.862)	0.856 (0.851−0.862)
20% screen-positive rate	0.807 (0.801−0.813)	0.807 (0.801−0.813)
PPV (95% CI)		
10% screen-positive rate	0.071 (0.059−0.084)	0.071 (0.059−0.084)
15% screen-positive rate	0.054 (0.046−0.064)	0.055 (0.047−0.065)
20% screen-positive rate	0.047 (0.040−0.054)	0.047 (0.041−0.055)
NPV (95% CI)		
10% screen-positive rate	0.992 (0.991−0.994)	0.992 (0.991−0.994)
15% screen-positive rate	0.993 (0.992−0.994)	0.993 (0.992−0.995)
20% screen-positive rate	0.994 (0.993−0.995)	0.994 (0.993−0.996)

Khalil. Predictive model for preeclampsia using artificial neural networks. Am J Obstet Gynecol 2024. *(continued)*

TABLE 2
Diagnostic accuracy metrics for model performance *(continued)*

Preterm preeclampsia	Logistic regression (LR)	Neural network (NN)
Likelihood ratios (+, −)		
10% screen-positive rate	(5.38, 0.55)	(5.35, 0.55)
15% screen-positive rate	(4.06, 0.49)	(4.12, 0.48)
20% screen-positive rate	(3.47, 0.41)	(3.51, 0.40)
Term preeclampsia	Logistic regression (LR)	Neural network (NN)
AUC score (95% CI)	0.712 (0.707–0.717)	0.713 (0.708–0.718)
Sensitivity (95% CI)		
10% screen-positive rate	0.293 (0.283–0.303)	0.307 (0.297–0.317)
15% screen-positive rate	0.408 (0.398–0.418)	0.417 (0.407–0.427)
20% screen-positive rate	0.490 (0.480–0.500)	0.499 (0.489–0.509)
Specificity (95% CI)		
10% screen-positive rate	0.905 (0.900–0.909)	0.905 (0.901–0.909)
15% screen-positive rate	0.856 (0.851–0.862)	0.857 (0.851–0.862)
20% screen-positive rate	0.807 (0.801–0.813)	0.807 (0.801–0.813)
PPV (95% CI)		
10% screen-positive rate	0.070 (0.059–0.083)	0.074 (0.062–0.087)
15% screen-positive rate	0.065 (0.056–0.075)	0.067 (0.058–0.077)
20% screen-positive rate	0.059 (0.051–0.067)	0.060 (0.052–0.068)
NPV (95% CI)		
10% screen-positive rate	0.981 (0.979–0.983)	0.982 (0.979–0.984)
15% screen-positive rate	0.983 (0.981–0.985)	0.984 (0.981–0.986)
20% screen-positive rate	0.985 (0.983–0.987)	0.985 (0.983–0.987)
Likelihood ratios (+, −)		
10% screen-positive rate	(3.08, 0.78)	(3.23, 0.77)
15% screen-positive rate	(2.83, 0.69)	(2.92, 0.68)
20% screen-positive rate	(2.54, 0.63)	(2.59, 0.62)

Sensitivity, specificity, PPV, NPV, AUC with 95% CIs, and likelihood ratios for both the LR and NN models for predicting early-onset preeclampsia (<34 weeks), preterm preeclampsia (<37 weeks), and term preeclampsia (≥37 weeks' gestation).

AUC, area under the receiver operating curve; *CI*, confidence interval; *NPV*, negative predictive value; *PPV*, positive predictive value.

Khalil. Predictive model for preeclampsia using artificial neural networks. Am J Obstet Gynecol 2024.

Model performance

Model performance was assessed by the expected AUC score and the sensitivity for fixed screen-positive rates of 10%, 15%, and 20% for the primary and secondary outcomes.

Implementation in a real-world setting

Finally, examples were provided for how the NN model could be used to generate posttest probabilities for preeclampsia, which could then guide pregnancy care based on risk stratification.

Role of the funding source

This study was a secondary analysis of data from the SMART study, which was funded by Natera, Inc., Austin, TX.

Results

Characteristics of the study population

Of 17,520 participants included in the final analysis, 671 (3.8%) developed preeclampsia: 72 (0.4%) early-onset preeclampsia, 251 (1.4%) preterm preeclampsia, and 420 (2.4%) term preeclampsia. Median gestational age at cfDNA measurement was 12.6 weeks, and 2155 (12.3%) had their cfDNA measurement at ≥16 weeks' gestation. The characteristics of the 17,520 participants in the study cohort (671 with preeclampsia and 16,849 without preeclampsia) are shown in Table 1. The racial and ethnic distribution of the study cohort, stratified by country, is shown in Supplemental Table 2. In addition, the preeclampsia rates (with 95% CIs around the point estimates) of the study cohort, stratified by country, are shown in Supplemental Table 3. The group that developed preeclampsia had a similar maternal age distribution to those who did not develop preeclampsia (34.3 vs 34.7 years; P=.109). The proportion of participants with a BMI >30 kg/m^2 was significantly higher in the preeclampsia group (40.8% vs 18.8%; P<.001). Participants in the United States and participants who were nulliparous, Black, or had a history of preterm birth, stillbirth, cesarean delivery, chronic hypertension, or prepregnancy diabetes were overrepresented in the group that developed preeclampsia. CfDNA parameters were also significantly different in the group that developed preeclampsia. Total cfDNA was significantly higher (median of 362.3 vs 339.0 DNA copies/mL; P<.001) and FF significantly lower (median of 7.5% vs 9.4%; P<.001) in the preeclampsia group (Table 1). We additionally calculated the Mahalanobis distance (MD), a measure of multivariate effect size comparing the standardized difference in cfDNA features between outcome groups. The MD comparing the preeclampsia and no-preeclampsia groups was 0.51, indicating considerable overlap between group distributions.

Model selection and performance

We analyzed and optimized a relatively large number of dense, feedforward NN models using the protocol described above. In more detail, we considered NN models with the number of hidden layers varying between 1 and 5 and the

number of neurons varying between 2 and 1024. On those NN models, we tested different regularization strategies such as L2 kernel regularization, dropout layers, and noise injection in the training set.[27] None of the NN models analyzed in the study outperformed the LR model.

The simplest NN model that performed on par with the LR model was an NN with 1 hidden layer composed of 15 ReLu units (resulting in a hidden layer with same dimension as the input) and with *L2* regularization implemented as Keras' *kernel_regularizer*.[28]

The expected AUC score, sensitivity, specificity, PPV, NPV, and likelihood ratios for fixed screen-positive rates of 10%, 15%, and 20% for each of the outcomes are shown in Table 2. For preterm preeclampsia at a set screen-positive rate of 15%, the NN model had a PPV of 0.055 and NPV of 0.993. As expected, the PPV was lower in early-onset preeclampsia, which had an even lower incidence (0.4% vs 1.4%). The numbers shown are mean values obtained from the optimization protocol described in the Methods, and represent the expected behavior of the model on unseen data.

The LR and NN models had similar predictive performance across each of the preeclampsia outcomes (early-onset, preterm, and term preeclampsia). The AUC scores for early-onset, preterm, and term preeclampsia were 0.78, 0.80, and 0.71, respectively, for the LR model, and 0.79, 0.80, and 0.71, respectively, for the NN model (Table 2). Sensitivity was highest for the primary outcome, preterm preeclampsia, with the LR model predicting 58% of positive cases and the NN model predicting 59% of positive cases at a screen-positive rate of 15% (Table 2). No differences were observed in predictive performance between the LR and NN models for the secondary outcomes (Table 2).

To assess the contribution of the cfDNA features to the performance of the models, we performed additional training–validation runs on a reduced version of the data set where the FF and total cfDNA features were omitted. In these NN models, the number of units in the hidden layer was rescaled to 13 to match the dimensionality of the input layer. The contribution of both total cfDNA and FF to the prediction of term and preterm preeclampsia was negligible (Table 3). For early-onset preeclampsia, removal of the total cfDNA and FF features from the model was associated with an approximately 7% decrease in sensitivity at a 15% screen-positive rate for both the LR and NN models (Table 3). Specifically, removal of the total cfDNA and FF features from the NN model decreased sensitivity for early-onset preeclampsia, at a 15% screen-positive rate, from 54.9% (95% CI, 52.9%–56.9%) to 48.0% (95% CI, 45.0%–51.0%) (Table 3).

The contribution of ethnicity-related information was evaluated via an additional set of models, estimating the AUC score for preterm preeclampsia and its 95% CI (Supplemental Table 4). Comparing these results with those listed in Table 2 (where race and ethnicity are not included in the models), race and ethnicity information did not contribute to the predictive performance of the models in our study.

Feature importance

The input features that were the major contributors to both the LR and NN models' predictions were maternal weight, chronic hypertension, prepregnancy diabetes, previous preterm birth, and previous stillbirth (Figure 2). Multiparity and increasing height showed a strong negative correlation with preterm preeclampsia in both models. When the cfDNA features were omitted from the model, the remaining 13 features

TABLE 3
Model performance with and without the cell-free DNA features

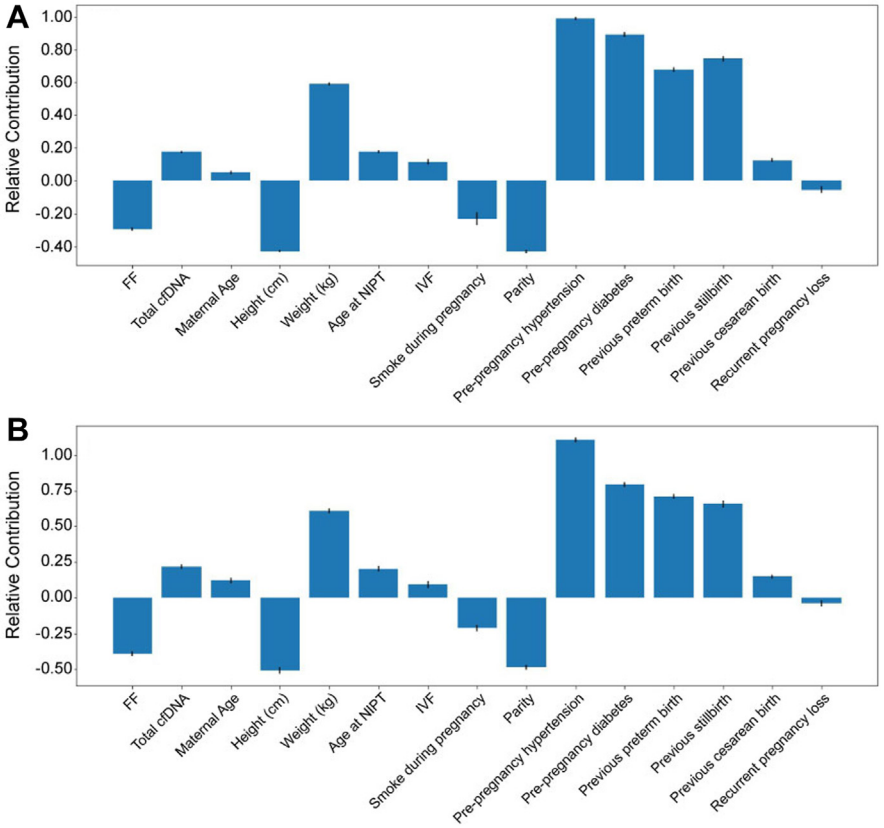
	AUC score	Sensitivity for 15% screen positive rate
Early-onset preeclampsia		
LR with MF+total cfDNA and FF	0.782 (0.768–0.796)	0.545 (0.525–0.565)
LR with MF only	0.768 (0.755–0.781)	0.474 (0.444–0.504)
Preterm preeclampsia		
LR with MF+total cfDNA and FF	0.801 (0.795–0.807)	0.584 (0.569–0.599)
LR with MF only	0.800 (0.794–0.806)	0.590 (0.575–0.605)
Term preeclampsia		
LR with MF+total cfDNA and FF	0.705 (0.702–0.708)	0.408 (0.398–0.418)
LR with MF only	0.704 (0.699–0.709)	0.398 (0.388–0.408)
Early-onset preeclampsia		
NN with MF+total cfDNA and FF	0.797 (0.783–0.811)	0.549 (0.529–0.569)
NN with MF only	0.768 (0.755–0.781)	0.480 (0.450–0.510)
Preterm preeclampsia		
NN with MF+total cfDNA and FF	0.800 (0.794–0.806)	0.593 (0.578–0.608)
NN with MF only	0.802 (0.798–0.808)	0.582 (0.567–0.597)
Term preeclampsia		
NN with MF+total cfDNA and FF	0.706 (0.703–0.709)	0.417 (0.407–0.427)
NN with MF only	0.707 (0.702–0.712)	0.413 (0.403–0.423)

Model performance for the LR and NN models with and without cfDNA features. The models where the cfDNA features (total cfDNA and FF) are omitted are termed "Maternal Factors (MF) only." Models "with MF+total cfDNA and FF" refer to the models trained and optimized on the full 15 features in Supplemental Table 1.

AUC, area under the receiver operating curve; *cfDNA*, cell-free DNA; *FF*, fetal fraction; *LR*, logistic regression; *MF*, maternal features; *NN*, nonlinear neural networks.

Khalil. Predictive model for preeclampsia using artificial neural networks. Am J Obstet Gynecol 2024.

FIGURE 2
Contribution of all features for the prediction of preterm preeclampsia



Significance and relative contribution of all 15 features to the final prediction of preterm preeclampsia for (**A**) the logistic regression model and (**B**) the nonlinear neural network model. The average contribution and 95% confidence intervals are shown. The magnitude and direction of the bars represent the contribution of each single feature to the definition of patients positive for preterm preeclampsia.

cfDNA, cell-free DNA; *FF*, fetal fraction; *IVF*, in vitro fertilization; *NIPT*, noninvasive prenatal testing.

Khalil. Predictive model for preeclampsia using artificial neural networks. Am J Obstet Gynecol 2024.

adapted to compensate for their absence, as would be expected (Supplemental Figure).

Implementation in a real-world setting

The NN model could be implemented in a real-world setting by providing an early-pregnancy personalized risk prediction for preterm preeclampsia and using risk stratification to guide care in pregnancy. To demonstrate this, posttest probabilities were generated for 2 patients using a Bayesian approach and using the model inputs shown in the Boxes 1 and 2. The background incidence of preterm preeclampsia in the study cohort was 1.43%. Taking the output of the NN as the likelihood in a Bayesian model, the posttest probability of developing preterm preeclampsia was obtained as the posterior by applying a Bayesian update to the prior probability.

For the example provided in Box 1, the NN output was 0.92, which was then used to calculate the posttest probability based on the following 2 equations:

1. Posttest (posterior) probability of developing preeclampsia=k×likelihood of preterm preeclampsia×pretest probability of developing preterm preeclampsia
2. Posttest (posterior) probability of not developing preterm preeclampsia=k×likelihood of no preterm preeclampsia×pretest probability of not developing preterm preeclampsia

These resulted in the following:

1. Posttest (posterior) probability of developing preterm preeclampsia=$k\times0.92\times0.014$
2. Posttest (posterior) probability of not developing preterm preeclampsia=$k\times(1-0.92)\times(1-0.014)$

Solving these 2 equations yielded a posttest probability of 14.6% for developing preterm preeclampsia. Using the same approach, for the patient with the characteristics shown in Box 2, the posttest probability of developing preterm preeclampsia was only 0.32%.

These results have implications for clinical decision-making. In the case of the patient in Box 1, whose posttest probability was 14.6%, the clinician might recommend a higher dose of low-dose aspirin (150 instead of 75 mg per day), increased surveillance of blood pressure and fetal growth, and provide information for the patient about their increased risk for preterm preeclampsia and about symptoms that should prompt urgent medical evaluation. For the patient in Box 2, given the low posttest probability of preterm preeclampsia (0.32%), the clinician would likely recommend against low-dose aspirin.

Comment

Principal findings

Our study showed that an artificial NN model using patient characteristics routinely collected at the first antenatal visit and features available from SNP-based cfDNA screening (total cfDNA and FF) can predict preeclampsia, with performance comparable to that of previously reported models (expected AUC score, 0.800; 58% sensitivity at a 15% screen-positive rate and 50.3% at a 10% screen-positive rate).[9,14,31,32] Predictive performance was similar for early-onset preeclampsia (AUC,

BOX 1
Probability of preterm preeclampsia based on the neural network predictive model: high-risk case example

Posttest probability of preterm preeclampsia: 14.6%
HIGH RISK

Patient characteristics	Value
Total cfDNA	502.88
Fetal fraction percentile	7.41
Maternal age at delivery (y)	33.34
Height (in)	63
Weight (lb)	120
Gestational age at cfDNA screening	12 wk 2 d
In vitro fertilization	No
Smoking during pregnancy	No
Parity	Multiparous
History of hypertension	Yes
Prepregnancy diabetes	No
Previous preterm birth	Yes
Previous stillbirth	No
Previous cesarean delivery	Yes
Recurrent pregnancy loss	Yes

Probability of preterm preeclampsia based on the neural network predictive model. Patient characteristics and values for an individual with a high risk of preterm preeclampsia with a posttest probability from the neural network model of 14.6%.
cfDNA, cell-free DNA.
Khalil. Predictive model for preeclampsia using artificial neural networks. Am J Obstet Gynecol 2024.

0.797) but lower for term preeclampsia (AUC, 0.713). The contribution of cfDNA variables (total cfDNA and FF) to the prediction of *term* and *preterm* preeclampsia was negligible, but for *early-onset* preeclampsia (<34 weeks' gestation), the cfDNA variables were associated with a statistically significant 7% increase in sensitivity at a 15% screen-positive rate. PPVs for the model were low and NPVs were high, which is consistent with reporting for other preeclampsia prediction models.[5,32] PPV and NPV are dependent on the prevalence of the condition and will also vary according to thresholds for test positivity (typically set at 1 in 100 to 1 in 150 for preeclampsia prediction models).[5] PPV and NPV may be used in a clinical setting to determine the appropriate test positivity threshold for using a test as a rule-in or rule-out. For instance, in the case of preterm preeclampsia, given that the intervention is relatively unusual, clinicians may use the NPV and likelihood ratios as a "rule out" and minimize the chance of missing patients who are at risk of subsequently developing preeclampsia.

Notably, the NN models had very similar performance to the LR models. It is well known that LR, a type of machine learning algorithm in and of itself, might perform equally well as more complex models for variables with linear relationships. Therefore, our findings most likely reflect a lack of nonlinear relationships in the data studied and that the additional complexity provided by NNs was not needed to improve screening performance. This finding might not be surprising given that many studies have reported quite small or no performance advantage of NN over LR models in health care settings.[33] However, given that in other areas, NNs are adept at learning complex patterns and relationships in large data sets,[34] we might potentially observe better performance from NNs as data sets become larger, both in the number of variables and the number of data points. In addition, an NN may have better performance than LR models for preterm preeclampsia prediction as additional continuous variables are added to the model. Another potential benefit of using NN models is that NNs can allow us to use a "transfer learning"—based approach, whereby an NN trained on a different, large data set is applied to the current data set, resulting in a boost in performance, as opposed to an NN trained from scratch on the current data set.[35]

Results in the context of what is known

The value of early-pregnancy cfDNA for prediction of preeclampsia remains controversial.[36] In the current analysis, cfDNA features (total cfDNA and FF) were both associated with the risk for preeclampsia. However, cfDNA and FF contributed to model performance only for early-onset preeclampsia. This finding of better biomarker performance for earlier-onset disease may reflect a different pathogenic process underlying early- and later-onset preeclampsia, as has been previously postulated.[37]

There is a large body of research reporting clinical risk factors and predictive models for preeclampsia.[4,5,18] The Fetal Medicine Foundation (FMF) model, which is based on patient characteristics, mean arterial pressure, uterine artery Doppler, and maternal serum PlGF (placental growth factor) levels at 11 to 13 weeks of gestation, has been studied most extensively and has been reported to have sensitivities ranging from 75% to 90% at a 10% screen-positive rate.[38] Using patient characteristics without maternal serum markers or Doppler, the FMF model achieved a 48% sensitivity for preterm preeclampsia at an 11% false-positive rate.[39] Similarly, our LR and NN models, both based only on patient characteristics and medical history (not including total cfDNA and FF), achieved 51% and 52% sensitivity for preterm preeclampsia,

BOX 2
Probability of preterm preeclampsia based on the neural network predictive model: low-risk case example

Posttest probability of preterm preeclampsia: 0.32%
LOW RISK

Patient characteristics	Value
Total cfDNA	209.12
Fetal fraction percentile	97.28
Maternal age at delivery (y)	37.8
Height (in)	66
Weight (lb)	190
Gestational age at cfDNA screening	9 wk 6 d
In vitro fertilization	No
Smoking during pregnancy	No
Parity	Multiparous
History of hypertension	No
Prepregnancy diabetes	No
Previous preterm birth	No
Previous stillbirth	No
Previous cesarean delivery	No
Recurrent pregnancy loss	No

Probability of preterm preeclampsia based on the neural network predictive model. Patient characteristics and values for an individual with a low risk of preterm preeclampsia as determined by the neural network model with a posttest probability of 0.32%.
cfDNA, cell-free DNA.
Khalil. Predictive model for preeclampsia using artificial neural networks. Am J Obstet Gynecol 2024.

respectively, when the screen-positive rate was set at 11%.

The 2 most common computational approaches reported for model development are LR models or a competing-risks model based on the Bayes theorem.[16,40] Studies using artificial intelligence or machine learning approaches for the prediction of preterm preeclampsia are more limited,[8–15,41–43] and differ from our model (a subset of these articles are summarized in Table 4[8–15]). For example, Marić et al[9] used the elastic net algorithm that included mean arterial pressure, and routine laboratory results, as well as patient medications. This model resulted in an AUC of 0.89 for the prediction of preterm preeclampsia with a 72.3% sensitivity. In addition, Ansbacher-Feldman et al[14] used an artificial NN that included uterine artery pulsatility, as well as mean arterial blood pressure, PlGF, and pregnancy-associated plasma protein. When these biomarkers were included, the AUC for the model was 0.909, with a 75.3% sensitivity at a 10% screen-positive rate, for the prediction of preterm preeclampsia. The patient cohort used in the Marić et al[9] study was quite small (5245 patients), and the patient cohort from the Ansbacher-Feldman et al[14] study was recruited from only 2 centers in the United Kingdom. In contrast, our model, although limited to the variables available in the SMART data set, had the advantage of being trained in a large, diverse, and international data set. The study by Ansbacher-Feldman et al[14] reported similar performance for NN and LR models in their study cohort, consistent with the findings of our study. However, in contrast to our study, the authors reported that adding race and ethnicity made a significant contribution to their model performance. The reason why we observed no difference in model performance with the addition of race and ethnicity is uncertain but could reflect the diversity within the broad racial and ethnic groups among the different countries participating in the studies.

However, in both the United States and the United Kingdom, current national guidelines support a different methodological approach and recommend a rule-based algorithm to assess preeclampsia risk.[44–47] Applying these rules to a European pregnant population, O'Gorman et al[48] reported that the 2015 American College of Obstetricians and Gynecologists guidelines[44] identified 90% of cases of preterm preeclampsia at a 64.2% screen-positive rate and that the United Kingdom's 2010 National Institute for Health and Care Excellence guidelines[46] identified 39.0% of preterm preeclampsia at a 10.2% screen-positive rate. Our model's 50% sensitivity for preterm preeclampsia at a 10% screen-positive rate, although not as good as the performance reported for predictive models that include blood pressure and biomarkers, compares favorably with the current standard of care.

Research and clinical implications

Simple models, such as ours, that use only routinely collected clinical data have practical and economic advantages for clinical implementation over more complex models (such as those that include ultrasound measurements), but at the expense of the accuracy of preeclampsia prediction. However, there is still potential for incremental improvement in performance by including additional routinely collected data. Mean arterial pressure has been shown to significantly increase predictive performance for preterm preeclampsia when added to a model based on patient characteristics alone.[38] Although blood pressure data were not collected as part of the SMART study, they are typically recorded as part of routine pregnancy care and could thus be added to the current model to potentially improve its performance. Future research could incorporate mean arterial pressure and potentially other routinely collected data, such as routine laboratory test results, into the model.

TABLE 4
Review of published artificial intelligence models that aim to predict preeclampsia early in pregnancy

Study (y)	Gestation at prediction (wk)	Outcome	Model	Features	AUC	Sensitivity	Screen-positive rate	External validation
Mello et al,[10] 2001	16	Pregnancy-induced hypertensive disorders	Artificial neural networks	Maternal age, BMI, urea, creatinine, uric acid, total proteins, hematocrit, iron, and ferritin	0.952	68.9%	14.85%	—
Jhee et al,[11] 2019	14–17	PE >34 wk	Stochastic gradient boosting (and additional models)	Maternal variables Routine laboratory results	0.924	77.1%	—	No
Sufriyana et al,[12] 2020		PE	Random forest (and additional models)	Maternal variables	0.76	—	—	Yes
Marić et al,[9] 2020	16	PE <34 wk Term	Elastic net algorithm (and additional models)	Maternal variables Routine laboratory results Medications	0.89 0.79	72.3% 45.2%	8.8% 8.1%	No
Li et al,[13] 2021	<20	PE ≥34 wk PE <37 wk Term	XGboost (and additional models)	Maternal variables Mean arterial pressure Routine laboratory results	0.93 0.98 0.92	— — —	— — —	No
Liu et al,[8] 2022	11–13+6	PE	Random forest (and additional models)	Maternal variables PAPP-A, B-HCG Uterine artery Doppler pulsatility index	0.86	42%	—	No
Lee et al,[15] 2022	<14	Pregnancy-associated hypertension	Graph-based semisupervised learning	Maternal variables, blood pressure	0.81 in test set	45.5% in test set	19.5%	No
Ansbacher-Feldman et al,[14] 2022	12–13	Preterm PE	Artificial neural network	Maternal variables Uterine artery pulsatility index Mean arterial blood pressure. PIGF, PAPP-A	0.816 0.909	53.3% maternal variables 75.3% with biomarkers	10%	Yes

AUC, area under the receiver operating curve; *B-HCG*, beta-human chorionic gonadotropin; *BMI*, body mass index; *PAPP-A*, pregnancy-associated plasma protein A; *PE*, preeclampsia; *PIGF*, placental growth factor.

Khalil. Predictive model for preeclampsia using artificial neural networks. Am J Obstet Gynecol 2024.

The population in the SMART study was international and diverse. However, predictive models do not always perform consistently across different populations. Future research would therefore also include evaluation of model performance in a different population (external validation). Even if predictive performance of a model is acceptable for clinical implementation, ongoing research is needed to determine if the information provided by the model produces improvements in clinical outcomes. Additional future research would include adding further features to the model, including blood pressure, routine laboratory test results, biomarkers, and longitudinal changes in cfDNA parameters over time.[9,17,37,38]

Strengths and limitations

The main strength of this study is the use of a large prospectively collected international data set that was used to train an NN model. The data included baseline patient characteristics and pregnancy outcomes. This is one of the largest studies to assess the contribution of total cfDNA and FF to the prediction of preeclampsia. However, although FF is frequently reported in prenatal cfDNA screening results, total cfDNA is not, which could impact implementation of the current model. In addition, FF estimation might not be consistent across laboratories.[49]

Although data on patient history and outcomes are robust, the lack of some clinical features, such as mean arterial blood pressure, as well as laboratory and ultrasonographic parameters, is a limitation. In addition, data were missing for some participants, resulting in exclusion of approximately 7.2% of cases; it is unknown if these data points were missing at random. Furthermore, we did not collect data on whether pregnant participants received aspirin during pregnancy. This limitation is common in screening studies on preeclampsia. In our study, this is in part due to the evidence from a randomized controlled trial on the benefits of low-dose aspirin for preeclampsia being published after the recruitment of participants.[50] Although some have argued for universal aspirin administration,[2,50,51] a predictive model for preeclampsia holds value given that some trials have found that universal aspirin administration was not associated with a reduction in the prevalence of preeclampsia in high-risk individuals.[52] Furthermore, others have reported that aspirin use is associated with increased rates of bleeding, particularly in the postpartum period.[53] An additional limitation is that the diagnosis of preeclampsia may not have been entirely consistent across clinical sites. However, the incidence of preeclampsia in this study is similar to that reported in the published literature using consensus guidelines. The optimal time to predict preeclampsia is before the end of the first trimester, when the greatest evidence exists for the effectiveness of low-dose aspirin. In this study, it was not possible to assess predictive accuracy according to the gestational age at cfDNA sampling because only 12.3% of participants had cfDNA evaluated at $\geq$16 weeks' gestation. Several limitations exist within our model, including the potential for overfitting in the NN due to the small data set and class imbalance (1%−1.5% positive preeclampsia cases). Further, at this time, the model has not undergone external validation, which is an important step to assess the model's transportability or ability to perform consistently in different health care settings. External validation is a future research goal. Last, NNs have the advantage of discerning nonlinear relationships that might exist among clinical and demographic data. However, they can be more difficult to interpret compared with a simple LR algorithm. Interpretability is especially important in a clinical setting because clinicians depend on the outcome of the model to make decisions regarding the health and treatment of patients. Thus, although NNs have exceptional computational advantages, their complexity hinders interpretability, which is a limitation of their utility in clinical settings.

Conclusions

Our study has developed a predictive model based only on easily acquired, routinely collected clinical data, all of which are part of routine clinical care early in pregnancy. Although the predictive ability of our model is not as good as that of models using other biomarkers and ultrasound, it has the advantage of maximizing routinely collected input features without adding the expense of biomarkers and ultrasound. It could therefore be incorporated into clinical care with minimal resource implications and inconvenience for patients. The predictive performance of the NN was very similar to that of the LR. A larger data set and more discriminatory input features are needed for future NN exploration. ■

References

1. Chinn JJ, Eisenberg E, Artis Dickerson SA, et al. Maternal mortality in the United States: research gaps, opportunities, and priorities. Am J Obstet Gynecol 2020;223:486–92.e6.
2. Roberge S, Bujold E, Nicolaides KH. Aspirin for the prevention of preterm and term preeclampsia: systematic review and metaanalysis. Am J Obstet Gynecol 2018;218:287–93.e1.
3. United States Preventive Services Task Force, Davidson KW, Barry MJ, et al. Aspirin use to prevent cardiovascular disease: US Preventive Services Task Force recommendation statement. JAMA 2022;327:1577–84.
4. Townsend R, Khalil A, Premakumar Y, et al. Prediction of pre-eclampsia: review of reviews. Ultrasound Obstet Gynecol 2019;54:16–27.
5. Chaemsaithong P, Sahota DS, Poon LC. First trimester preeclampsia screening and prediction. Am J Obstet Gynecol 2022;226:S1071–97.e2.
6. Chaiworapongsa T, Romero R, Gotsch F, et al. Preeclampsia at term can be classified into 2 clusters with different clinical characteristics and outcomes based on angiogenic biomarkers in maternal blood. Am J Obstet Gynecol 2023;228:569.e1–24.
7. Tarca AL, Taran A, Romero R, et al. Prediction of preeclampsia throughout gestation with maternal characteristics and biophysical and biochemical markers: a longitudinal study. Am J Obstet Gynecol 2022;226:126.e1–22.
8. Liu M, Yang X, Chen G, et al. Development of a prediction model on preeclampsia using machine learning-based method: a retrospective cohort study in China. Front Physiol 2022;13:896969.
9. Marić I, Tsur A, Aghaeepour N, et al. Early prediction of preeclampsia via machine learning. Am J Obstet Gynecol MFM 2020;2:100100.
10. Mello G, Parretti E, Ognibene A, et al. Prediction of the development of pregnancy-induced hypertensive disorders in high-risk pregnant women by artificial neural networks. Clin Chem Lab Med 2001;39:801–5.
11. Jhee JH, Lee S, Park Y, et al. Prediction model development of late-onset preeclampsia using machine learning-based methods. PLoS One 2019;14:e0221202.
12. Sufriyana H, Wu YW, Su ECY. Artificial intelligence-assisted prediction of preeclampsia: development and external validation of a nationwide health insurance dataset of the BPJS Kesehatan in Indonesia. EBioMedicine 2020;54:102710.
13. Li YX, Shen XP, Yang C, et al. Novel-electronic health records applied for prediction of pre-eclampsia: machine-learning algorithms. Pregnancy Hypertens 2021;26:102–9.
14. Ansbacher-Feldman Z, Syngelaki A, Meiri H, Cirkin R, Nicolaides KH, Louzoun Y. Machine-learning-based prediction of pre-eclampsia using first-trimester maternal characteristics and biomarkers. Ultrasound Obstet Gynecol 2022;60:739–45.
15. Lee SM, Nam Y, Choi ES, et al. Development of early prediction model for pregnancy-associated hypertension with graph-based semi-supervised learning. Sci Rep 2022;12:15793.
16. Sandström A, Snowden JM, Höijer J, Bottai M, Wikström AK. Clinical risk assessment in early pregnancy for preeclampsia in nulliparous women: a population based cohort study. PLoS One 2019;14:e0225716.
17. Schmidt LJ, Rieger O, Neznansky M, et al. A machine-learning–based algorithm improves prediction of preeclampsia-associated adverse outcomes. Am J Obstet Gynecol 2022;227:77.e1–30.
18. Bartsch E, Medcalf KE, Park AL, Ray JG; High Risk of Pre-eclampsia Identification Group. Clinical risk factors for pre-eclampsia determined in early pregnancy: systematic review and meta-analysis of large cohort studies. BMJ 2016;353:i1753.
19. Shree R, Kolarova TR, Mackinnon HJ, Lockwood CM, Chandrasekaran S. Association

of fetal fraction with hypertensive disorders of pregnancy incidence and disease severity. Am J Obstet Gynecol MFM 2022;4:100671.
20. American College of Obstetricians and Gynecologists' Committee on Practice Bulletins—Obstetrics; Committee on Genetics; Society for Maternal-Fetal Medicine. Screening for fetal chromosomal abnormalities: ACOG practice bulletin, number 226. Obstet Gynecol 2020;136:e48–69.
21. Dar P, Jacobsson B, MacPherson C, et al. Cell-free DNA screening for trisomies 21, 18, and 13 in pregnancies at low and high risk for aneuploidy with genetic confirmation. Am J Obstet Gynecol 2022;227:259.e1–14.
22. Sinkey RG, Battarbee AN, Bello NA, Ives CW, Oparil S, Tita ATN. Prevention, diagnosis, and management of hypertensive disorders of pregnancy: a comparison of international guidelines. Curr Hypertens Rep 2020;22:66.
23. Zimmermann B, Hill M, Gemelos G, et al. Noninvasive prenatal aneuploidy testing of chromosomes 13, 18, 21, X, and Y, using targeted sequencing of polymorphic loci. Prenat Diagn 2012;32:1233–41.
24. Buckley A, Sestito S, Ogundipe T, et al. Racial and ethnic disparities among women undergoing a trial of labor after Cesarean delivery: performance of the VBAC calculator with and without patients' race/ethnicity. Reprod Sci 2022;29:2030–8.
25. Huang J, Galal G, Etemadi M, Vaidyanathan M. Evaluation and mitigation of racial bias in clinical machine learning models: scoping review. JMIR Med Inform 2022;10:e36388.
26. Agarap AF. Deep learning using rectified linear units (relu). arXiv preprint arXiv:180308375; 2018. Available at: https://arxiv.org/abs/1803.08375. Accessed December 15, 2022.
27. Goodfellow I, Bengio Y, Courville A. Deep learning. Cambridge, MA: MIT press; 2016.
28. Abadi M, Agarwal A, Barham P, et al. Tensorflow: large-scale machine learning on heterogeneous distributed systems. arXiv preprint arXiv:160304467; 2016. Available at: https://arxiv.org/abs/1603.04467. Accessed December 15, 2022.
29. Camuto A, Willetts M, Simsekli U, Roberts SJ, Holmes CC. Explicit regularisation in gaussian noise injections. Adv Neural Inf Process Syst 2020;33:16603–14.
30. Xu QS, Liang YZ. Monte Carlo cross validation. Chemom Intell Lab Syst 2001;56:1–11.
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. Zwertbroek EF, Groen H, Fontanella F, Maggio L, Marchi L, Bilardo CM. Performance of the FMF first-trimester preeclampsia-screening algorithm in a high-risk population in the Netherlands. Fetal Diagn Ther 2021;48:103–11.
33. Issitt RW, Cortina-Borja M, Bryant W, Bowyer S, Taylor AM, Sebire N. Classification performance of neural networks versus logistic regression models: evidence from healthcare practice. Cureus 2022;14:e22443.
34. Poomka P, Kerdprasop N, Kerdprasop K. Machine learning versus deep learning performances on the sentiment analysis of product reviews. Int J Mach Learn Comput 2021;11:103–9.
35. He L, Li H, Wang J, et al. A multi-task, multi-stage deep transfer learning model for early prediction of neurodevelopment in very preterm infants. Sci Rep 2020;10:15072.
36. Rolnik DL, O'Gorman N, Fiolna M, van den Boom D, Nicolaides KH, Poon LC. Maternal plasma cell-free DNA in the prediction of preeclampsia. Ultrasound Obstet Gynecol 2015;45:106–11.
37. Myatt L. The prediction of preeclampsia: the way forward. Am J Obstet Gynecol 2022;226:S1102–7.e8.
38. Wright D, Tan MY, O'Gorman N, et al. Predictive performance of the competing risk model in screening for preeclampsia. Am J Obstet Gynecol 2019;220:199.e1–13.
39. Wright D, Syngelaki A, Akolekar R, Poon LC, Nicolaides KH. Competing risks model in screening for preeclampsia by maternal characteristics and medical history. Am J Obstet Gynecol 2015;213:62.e1–10.
40. Wright D, Wright A, Nicolaides KH. The competing risk approach for prediction of preeclampsia. Am J Obstet Gynecol 2020;223:12–23.e7.
41. Han Q, Zheng W, Guo XD, et al. A new predicting model of preeclampsia based on peripheral blood test value. Eur Rev Med Pharmacol Sci 2020;24:7222–9.
42. Huang KH, Chen FY, Liu ZZ, et al. Prediction of pre-eclampsia complicated by fetal growth restriction and its perinatal outcome based on an artificial neural network model. Front Physiol 2022;13:992040.
43. Bennett R, Mulla ZD, Parikh P, Hauspurg A, Razzaghi T. An imbalance-aware deep neural network for early prediction of preeclampsia. PLoS One 2022;17:e0266042.
44. Committee Opinion No. 638: first-trimester risk assessment for early-onset preeclampsia. Obstet Gynecol 2015;126:e25–7.
45. Practice advisory: low-dose aspirin use for the prevention of preeclampsia and related morbidity and mortality. The American College of Obstetricians and Gynecologists. 2021. Available at: https://www.acog.org/clinical/clinical-guidance/practice-advisory/articles/2021/12/low-dose-aspirin-use-for-the-prevention-of-preeclampsia-and-related-morbidity-and-mortality. Accessed February 1, 2023.
46. National Collaborating Centre for Women's and Children's Health (UK). Hypertension in Pregnancy: The Management of Hypertensive Disorders During Pregnancy. London: RCOG Press, 2010.
47. Hypertension in pregnancy: diagnosis and management. National Institute for Health Care Excellence. 2019. Available at: https://www.nice.org.uk/guidance/ng133/resources/hypertension-in-pregnancy-diagnosis-and-management-pdf-66141717671365. Accessed February 1, 2023.
48. O'Gorman N, Wright D, Poon LC, 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;49:756–60.
49. Becking EC, Linthorst J, Patton S, et al. Variability in fetal fraction estimation: comparing fetal fractions reported by noninvasive prenatal testing providers globally. Clin Chem 2023;69:160–7.
50. Rolnik DL, Wright D, Poon LCY, et al. ASPRE trial: performance of screening for preterm pre-eclampsia. Ultrasound Obstet Gynecol 2017;50:492–5.
51. Richards EMF, Giorgione V, Stevens O, Thilaganathan B. Low-dose aspirin for the prevention of superimposed preeclampsia in women with chronic hypertension: a systematic review and meta-analysis. Am J Obstet Gynecol 2023;228:395–408.
52. Shen L, Martinez-Portilla RJ, Rolnik DL, Poon LC. ASPRE trial: risk factors for development of preterm pre-eclampsia despite aspirin prophylaxis. Ultrasound Obstet Gynecol 2021;58:546–52.
53. Jiang Y, Chen Z, Chen Y, et al. Low-dose aspirin use during pregnancy may be a potential risk for postpartum hemorrhage and increased blood loss: a systematic review and meta-analysis. Am J Obstet Gynecol MFM 2023;5:100878.

Author and article information

From the Department of Obstetrics and Gynaecology, St. George's Hospital, St. George's University of London, London, United Kingdom (Dr Khalil); Natera Inc., Austin, TX (Drs Bellesia, Mses Egbert and Prigmore, and Drs Ahmed, Demko, Ortiz, and Souter); Department of Obstetrics, Gynecology & Reproductive Sciences, University of California San Francisco, San Francisco, CA (Dr Norton); Department of Obstetrics and Gynecology, Institute of Clinical Sciences, Sahlgrenska Academy, University of Gothenburg, Gothenburg, Sweden (Dr Jacobsson); Department of Obstetrics and Gynecology, Sahlgrenska University Hospital, Gothenburg, Sweden (Dr Jacobsson); Austin Maternal-Fetal Medicine, Austin, TX (Dr Haeri); Department of Obstetrics and Gynaecology, Rotunda Hospital, Royal College of Surgeons in Ireland, Dublin, Ireland (Dr Malone); Department of Obstetrics and Gynecology, Columbia University Irving Medical Center, New York, NY (Dr Wapner); Department of Obstetrics and Gynecology, New York University Grossman School of Medicine, New York, NY (Dr Roman); Department of Obstetrics and Gynecology, Saint Peter's University Hospital, New Brunswick, NJ (Dr Faro); Department of Obstetrics and Gynecology, Long Island Jewish Medical Center, Donald and Barbara Zucker School of Medicine at Hofstra/Northwell, New Hyde Park, NY (Dr Madankumar); Department of Obstetrics and Gynecology, Icahn School of Medicine at Mount Sinai, New York, NY (Dr Strong); Department of Obstetrics and Gynecology, University of Utah, Salt Lake City, UT (Dr Silver); Department of

Obstetrics and Gynecology, North Shore University Hospital, Donald and Barbara Zucker School of Medicine at Hofstra/Northwell, Manhasset, NY (Dr Vohra); Department of Obstetrics and Gynaecology, Royal Prince Alfred Hospital, Western Sydney University, Sydney, Australia (Dr Hyett); Biostatistics Center, George Washington University, Rockville, MD (Dr MacPherson); and Department of Obstetrics and Gynecology and Women's Health, Montefiore Medical Center, Albert Einstein College of Medicine, New York, NY (Dr Dar).

Received May 12, 2023; revised Feb. 16, 2024; accepted Feb. 22, 2024.

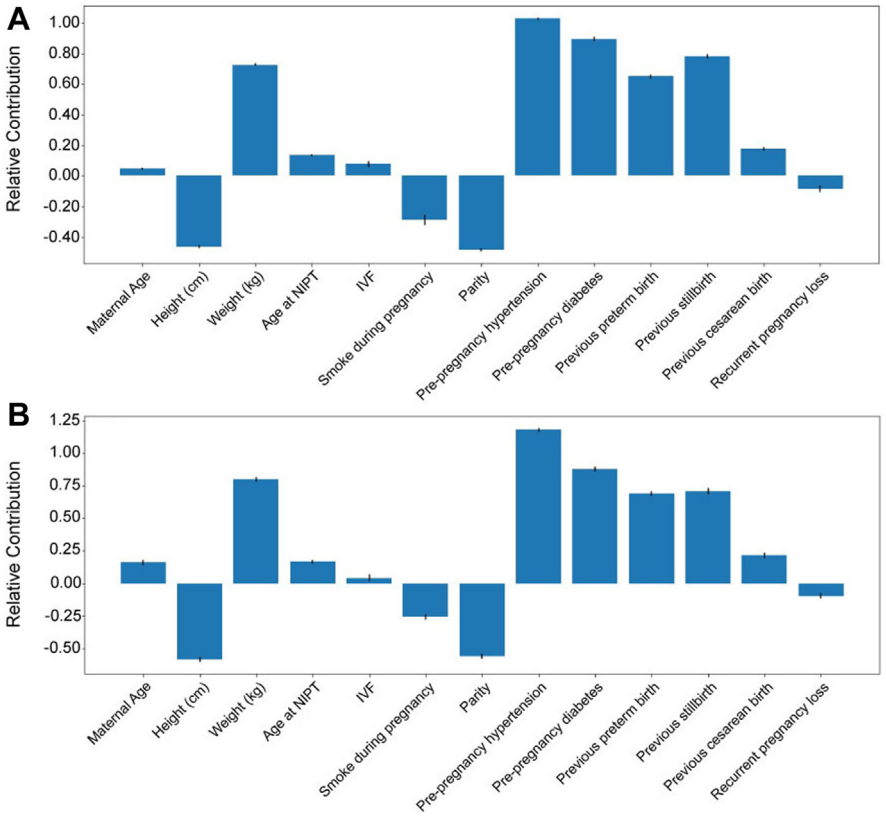
All site principal investigators (P.D., B.J., F.D.M., R.J.W., A.R., A.K., R.F., R.M., S.H., R.M.S., N.V., J.H., C.M., and M.E.N.) received institutional research support from the funding sponsor (Natera). G.B., M.E., B.P., E.A., Z.D., J.B.O., and V.S. are employed by the study's funding sponsor (Natera) and hold stock or options to hold stock. J.H. has an ongoing research collaboration that includes financial support for biochemical analytes from PerkinElmer, has earned honoraria and/or given talks that were compensated by Natera, Roche, and Canon, and has participated in Asian/Australasian expert consultancies for Natera and Roche. B.J. reports the following: research clinical diagnostic trials with Ariosa (completed), Vanadis (completed), Natera (ongoing), and Hologic (completed) with institutional expenditures reimbursed per patient and no personal reimbursements; clinical probiotic studies with product provided by FukoPharma (ongoing, no funding) and BioGaia (ongoing; also provided a research grant for the specific study); coordination of scientific conferences and meetings with commercial partners such as the European Spontaneous Preterm Birth Congress (ESPBC) 2016 and a Nordic educational meeting about noninvasive prenatal testing and preeclampsia screening. B.J. collaborated in the IMPACT study where Roche, PerkinElmer, and Thermo Fisher Scientific provided reagents for PlGF analyses. R.J.W. received research funding from the *Eunice Kennedy Shriver* National Institute of Child Health and Human Development and support from Illumina for research reagents. M.E.N. is a consultant to Invitae. The remaining authors report no conflict of interest.

This study was funded by Natera, Inc.

The findings from this study were presented as an oral poster presentation at the International Society of Ultrasound in Obstetrics and Gynecology 32nd Annual World Congress on Ultrasound in Obstetrics and Gynecology, London, United Kingdom, September 16—18, 2022.

Corresponding author: Asma Khalil, MD. asmakhalil79@googlemail.com

SUPPLEMENTAL FIGURE
Contribution of maternal factors for the prediction of preeclampsia



Significance and relative contribution of 13 features (maternal factors only, with total cell-free DNA and fetal fraction omitted) to the final prediction of preterm preeclampsia for (**A**) the logistic regression model and (**B**) the nonlinear neural network model. The average contribution and 95% confidence intervals are shown. The magnitude and direction of the bars represent the contribution of each single feature to the definition of patients positive for preterm preeclampsia.

IVF, in vitro fertilization; *NIPT*, noninvasive prenatal testing.

Khalil. Predictive model for preeclampsia using artificial neural networks. Am J Obstet Gynecol 2024.

SUPPLEMENTAL TABLE 1
Description and encoded type of input variables used in the predictive models

Description	Encoded type
Fetal fraction	Continuous
Total cfDNA	Continuous
Maternal age (y)	Continuous
Maternal height (cm)	Continuous
Maternal weight (kg)	Continuous
Gestational age at cfDNA screening	Continuous
Pregnancy conceived using in vitro fertilization	Binary [0,1]
Cigarette smoking during pregnancy	Binary [0,1]
Parity: nulliparous vs multiparous	Binary [0,1]
Prepregnancy hypertension	Binary [0,1]
Prepregnancy diabetes	Binary [0,1]
Previous preterm birth (<37 wk of gestation)	Binary [0,1]
Previous stillbirth	Binary [0,1]
Previous cesarean delivery	Binary [0,1]
Recurrent pregnancy loss (≥3 losses at <20 wk of gestation)	Binary [0,1]

cfDNA, cell-free DNA.
Khalil. Predictive model for preeclampsia using artificial neural networks. Am J Obstet Gynecol 2024.

SUPPLEMENTAL TABLE 2
Self-reported race and ethnicity of study participants stratified by country

Race and ethnicity N (%)	Total N=17,520	Australia N=430	England N=633	Ireland N=3945	Spain N=253	Sweden N=3137	United States N=9122
Asian	1492 (8.5%)	101 (23%)	64 (10%)	139 (3.5%)	0 (0%)	200 (6.4%)	988 (11%)
Black	1388 (7.9%)	1 (0.2%)	22 (3.5%)	7 (0.2%)	1 (0.4%)	21 (0.7%)	1336 (15%)
Latina	2852 (16%)	2 (0.5%)	1 (0.2%)	39 (1.0%)	8 (3.2%)	54 (1.7%)	2748 (30%)
White	11,213 (64%)	326 (76%)	510 (81%)	3744 (95%)	244 (96%)	2792 (89%)	3597 (39%)
Other	391 (2.2%)	0 (0%)	27 (4.3%)	15 (0.4%)	0 (0%)	69 (2.2%)	280 (3.1%)
Unknown/not reported	184 (1.1%)	0 (0%)	9 (1.4%)	1 (<0.1%)	0 (0%)	1 (<0.1%)	173 (1.9%)

Khalil. Predictive model for preeclampsia using artificial neural networks. Am J Obstet Gynecol 2024.

SUPPLEMENTAL TABLE 3
Preeclampsia rates stratified by country

Preeclampsia N (%) (95% CI)	Total	Australia	England	Ireland	Spain	Sweden	United States
Early-onset	72 (0.4%) (0.3–0.5)	0 (0%) (0–0.9)	0 (0%) (0–0.6)	7 (0.2%) (0.1–0.4)	0 (0%) (0–1.5)	4 (0.1%) (0–0.3)	61 (0.7%) (0.5%–0.9%)
Preterm	251 (1.4%) (1.3–1.6)	3 (0.7%) (0.2–2)	3 (0.5%) (0.2–1.4)	23 (0.6%) (0.4–0.9)	1 (0.4%) (0.1–2.2)	21 (0.7%) (0.4–1)	200 (2.2%) (1.9–2.5)
Term	420 (2.4%) (2.2–2.6)	5 (1.2%) (0.5–2.7)	5 (0.8%) (0.3–1.8)	47 (1.2%) (0.9–1.6)	1 (0.4%) (0.1–2.2)	60 (1.9%) (1.5–2.5)	302 (3.3%) (3–3.7)
Any preeclampsia	671 (3.8%) (3.6–4.1)	8 (1.9%) (0.9–3.6)	8 (1.3%) (0.6–2.5)	70 (1.8%) (1.4–2.2)	2 (0.8%) (0.2–2.8)	81 (2.6%) (2.1–3.2)	502 (5.5%) (5.1–6)

CI, confidence interval.

Khalil. Predictive model for preeclampsia using artificial neural networks. Am J Obstet Gynecol 2024.

SUPPLEMENTAL TABLE 4
Model performance for preterm preeclampsia prediction with the addition of race and ethnicity information

Race and ethnicity	LR AUC scores (95% CI)	NN AUC scores (95% CI)
Asian	0.796 (0.790–0.802)	0.797 (0.791–0.803)
Black	0.806 (0.799–0.813)	0.806 (0.800–0.812)
Latina	0.804 (0.798–0.810)	0.805 (0.799–0.811)
White	0.799 (0.792–0.806)	0.801 (0.794–0.809)

The AUC scores for the LR and NN models that did not include race and ethnicity were 0.801 and 0.800, respectively (Table 2).

AUC, area under the curve; *CI*, confidence interval; *LR*, logistic regression; *NN*, nonlinear neural networks.

Khalil. Predictive model for preeclampsia using artificial neural networks. Am J Obstet Gynecol 2024.