Prediction of gestational diabetes mellitus by multiple biomarkers at early gestation

Meng-Nan Yang[1,2†], Lin Zhang[3†], Wen-Juan Wang[1,4†], Rong Huang[2], Hua He[1], Tao Zheng[5], Guang-Hui Zhang[6], Fang Fang[1], Justin Cheng[2], Fei Li[1], Fengxiu Ouyang[1*], Jiong Li[1,7], Jun Zhang[1*] and Zhong-Cheng Luo[1,2*]

Abstract

Background It remains unclear which early gestational biomarkers can be used in predicting later development of gestational diabetes mellitus (GDM). We sought to identify the optimal combination of early gestational biomarkers in predicting GDM in machine learning (ML) models.

Methods This was a nested case-control study including 100 pairs of GDM and euglycemic (control) pregnancies in the Early Life Plan cohort in Shanghai, China. High sensitivity C reactive protein, sex hormone binding globulin, insulin-like growth factor I, IGF binding protein 2 (IGFBP-2), total and high molecular weight adiponectin and glycosylated fibronectin concentrations were measured in serum samples at 11–14 weeks of gestation. Routine first-trimester blood test biomarkers included fasting plasma glucose (FPG), serum lipids and thyroid hormones. Five ML models [stepwise logistic regression, least absolute shrinkage and selection operator (LASSO), random forest, support vector machine and k-nearest neighbor] were employed to predict GDM. The study subjects were randomly split into two sets for model development (training set, $n = 70$ GDM/control pairs) and validation (testing set: $n = 30$ GDM/control pairs). Model performance was evaluated by the area under the curve (AUC) in receiver operating characteristics.

Results FPG and IGFBP-2 were consistently selected as predictors of GDM in all ML models. The random forest model including FPG and IGFBP-2 performed the best (AUC 0.80, accuracy 0.72, sensitivity 0.87, specificity 0.57). Adding more predictors did not improve the discriminant power.

Conclusion The combination of FPG and IGFBP-2 at early gestation (11–14 weeks) could predict later development of GDM with moderate discriminant power. Further validation studies are warranted to assess the utility of this simple combination model in other independent cohorts.

†Meng-Nan Yang, Lin Zhang and Wen-Juan Wang contributed equally to this work.

*Correspondence:
Fengxiu Ouyang
ouyangfengxiu@xinhuamed.com.cn
Jun Zhang
junjimzhang@sjtu.edu.cn
Zhong-Cheng Luo
zc.luo@utoronnto.ca

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

Yang *et al. BMC Pregnancy and Childbirth*   (2024) 24:601

Page 2 of 9

Background

Gestational diabetes mellitus (GDM) is characterized by *de novo* glucose intolerance in the 2nd half of pregnancy affecting 7–15% of mothers worldwide [1, 2]. GDM increases the risks of pre-eclampsia, macrosomia and neonatal hypoglycemia [3, 4], and put both mothers and their offspring at increased risks of obesity, type 2 diabetes and cardiovascular diseases in later life [5–8]. Early interventions in pregnancy may reduce the risk of GDM and adverse pregnancy outcomes [9, 10]. However, GDM is routinely diagnosed between 24 and 28 gestational weeks by an oral glucose tolerance test (OGTT) [11, 12], missing the optimal window for interventions as fetal and placental development have already occurred [13]. Identifying subjects at early gestation who are destined to develop GDM may inform early intervention strategies in improving maternal and child health.

A number of early gestational biomarkers have been reported in predicting later development of GDM, mostly notably, fasting plasma glucose (FPG) [14, 15], sex hormone-binding globulin (SHBG) [16, 17], adiponectin [18–20], insulin-like growth factor I (IGF-I) and IGF binding protein 2 (IGFBP-2) [21, 22], high sensitivity C reactive protein (hs-CRP) [23, 24] and glycosylated fibronectin [25, 26]. However, no single biomarker has demonstrated sufficient discriminant power in clinical utility. It is unclear which combination is the optimal.

Most previous studies in predicting GDM employed conventional logistics regression models [27–29]. However, such models tend to ignore non-linear relationships and interactions between predictors, and the predictive performance may be suboptimal [30, 31]. Machine learning (ML) models may be helpful in alleviating these limitations [32]. In the present study, we sought to explore the optimal combination of early gestational biomarkers in predicting later development of GDM in ML models considering six main candidate biomarkers: FBG, SHBG, IGF-I, IGFBP-2, hs-CRP and glycosylated fibronectin.

Materials and methods

Study design and subjects

We conducted a nested matched (1:1) case-control study in the Early Life Plan (ELP) cohort. The ELP cohort was initiated in Xinhua Hospital - a university affiliated maternity and pediatric care hospital in Shanghai for studies on the determinants of pregnancy outcomes and child health. Pregnant women were recruited during routine first trimester antenatal care visits. Eligibility criteria were: (1) age 18+years; (2) residents in Shanghai; (3) planned to have prenatal care and delivery at Xinhua Hospital. Trained research staff collected data through face-to-face interviews and medical chart reviews using standardized structured questionnaires. The study was approved by the medical research ethical committee of Xinhua Hospital, School of Medicine, Shanghai Jiao Tong University (approval no. XHEC-C-2013-001). Written informed consent was obtained from all study participants. The data collection questionnaire is provided in the appendix - Early Life Plan Cohort Study Questionnaire.

All GDM and controls were from the ELP cohort in subjects recruited between 2016/6 and 2019/02. GDM was diagnosed according to the International Association of Diabetes and Pregnancy Study Groups (IADPSG) criteria [33]: if any one of the blood glucose values was at or above the following thresholds in the routine 75 g oral glucose tolerance test (OGTT) at 24–28 weeks of gestation: fasting 5.1 mmol/L, 1-hour 10.0 mmol/L and 2-hour 8.5 mmol/L. Controls were women with a euglycemic pregnancy.

The eligibility criteria for the present study were: (1) Han ethnicity (the majority ethnic group); (2) maternal age 20–45 years; (3) natural conception; (4) singleton pregnancy; (5) no severe pre-pregnancy illnesses (e.g., type 1 or type 2 diabetes) or life-threatening gestational complications (e.g., preeclampsia); (6) serum specimen at 11–14 weeks of gestation available for biomarker assays. All eligible GDM women with complete data and specimen collected by Feb 2019 (*n*=100) were included. Controls were randomly sampled among all eligible subjects recruited up to Feb 2019 in the ELP cohort, and matched to cases (1:1) by maternal age (within 1 year), pre-pregnancy BMI (within 1 kg/m^2) and gestational age (within 1 week) at blood sampling. Therefore, a total of 200 subjects (100 pairs of GDM/control) constituted the study sample. Figure S1 presents the flowchart in the selection of study subjects.

Blood samples

Maternal fasting blood samples were collected during the first prenatal care visits at 11–14 weeks of gestation. Trained technicians collected blood samples following standardized operating procedures. Fasting morning venous blood samples were collected in multiple tubes for serum (non-coagulant) and plasma (EDTA). The blood samples were centrifuged (Beckman Coulter Allegra X-15R, USA) at 4 ℃, 4000 r/min for 10 min. The separated serum and plasma samples were stored at −80 ℃ until assays.

Biochemical assays

In all biomarker assays, the lab technicians were blinded to the clinical status (GDM or not) of study subjects. Serum insulin-like growth factor (IGF)-I and insulin-like growth factor binding protein 2 (IGFBP-2) were measured by ELISA kits (Crystal Chem, Lllinois, USA) and the absorbance was determined using a microplate spectrophotometer (Beckman CX7, USA). Serum high sensitivity C reactive protein (hs-CRP) was measured by an ELISA kit (Anogen, Ontario, Canada). Serum HMW and total adiponectin were measured by an ELISA kit (ALPCO, Salem, NH, USA). Serum glycosylated fibronectin was measured by an ELISA kit (Anshlabs, Webster, USA). Serum sex hormone binding globulin (SHBG) was measured by an automated chemiluminescent assay (Kangrun Biotech, Guangzhou, China). The limits of detection were 2 ng/mL for IGF-I, 2 ng/mL for IGFBP-2, 0.03 mg/L for CRP, 0.034 ng/mL for HMW and total adiponectin, 2.85 ng/mL for glycosylated fibronectin, and 0.1 nmol/L for SHBG, respectively. The intra-assay and inter-assay coefficients of variation were in the ranges of 0.8-5.4% for IGF-I, 0.1-3.3% for IGFBP-2, 9.9-14.7% for hs-CRP, 0.5-4.5% for HMW and total adiponectin, 1.3-2.5% for glycosylated fibronectin, and <8% for SHBG, respectively.

Routine first-trimester prenatal blood test biomarkers included fasting plasma glucose (FPG), serum lipids [triglyceride, total cholesterols (TC)] and thyroid hormones [thyroid stimulating hormone (TSH); free thyroxine (FT4); free triiodothyronine (FT3)], and were measured in the clinical biochemistry laboratory of Xinhua hospital (a nationally accredited clinical biochemistry lab) following standardized operating protocols.

Selection of predictors

Potential predictors included the measured serum biomarkers (hs-CRP, IGFBP-2, IGF-1, SHBG, glycosylated fibronectin, total and HMW adiponectin) and available biomarkers from routine clinical tests (hemoglobin, FPG, TG, TC, TSH, FT4 and FT3) at 11–14 weeks of gestation, and maternal characteristics (parity, family history of diabetes or hypertension, systolic blood pressure (SBP), diastolic blood pressure (DBP)).

Four models were employed in the selection of predictors/features that distinguish GDM from euglycemic pregnancies, including stepwise logistic regression (LR), support vector machine-recursive feature elimination (SVM-RFE), the least absolute shrinkage and selection operator (LASSO) and random forest. The *MASS* package in R was used for stepwise LR. SVM-RFE is a ML technique that trains a subset of features from different categories to shrink the feature set and find the most predictive features [34]. The SVM-RFE was applied in the selection of predictive features via a 10-fold cross-validation, using the *e1071* package in R. LASSO is a regression method that performs both feature selection and regularization to enhance the prediction accuracy [35]. LASSO regression with a 10-fold cross-validation was performed to determine the optimal lambda using the *glmnet* package in R. The random forest method is a common method for ranking variable importance based on prediction performance [36]. For each trained tree, out-of-bag prediction performance was calculated before and after permutation of the values of each predictor variable, and the differences were then averaged over all trees. The *randomForest* package in R was used to implement the random forest algorithm. Ultimately, features that were selected by all algorithms were retained in the ML prediction models; other features that were selected by at least 2 algorithms were assessed on whether their additions could improve the predictions.

Prediction models

Conventional LR and five ML models (ML-LR, SVM, k-nearest neighbor (KNN), LASSO and random forest) were employed in predicting GDM. ML-LR iteratively identifies the strongest combination of variables with the greatest probability of detecting the observed outcomes. The SVM aims to create a decision boundary between two classes that enables the prediction from one or more feature vectors. KNN algorithm is a nonparametric approach for classification, and it integrates the information about the K neighbor points for the classification of subjects. The random forest model is a regression tree technique using bootstrapped samples and a random subset of features to achieve a high degree of predictive accuracy. In all ML models, we tuned the hyperparameters to select the optimal combination with the highest mean AUC using a 10-fold cross validation via GridSearchCV function in the "sklearn" package in Python: ML-LR (hyperparameters: penalty "L2", solver "liblinear", C "0.001"), KNN (hyperparameters: number of neighbors to use "14", weight function "uniform", algorithm used to predict the nearest neighbors "auto", leaf size "30"), SVM (kernel parameter "linear", C "1", gamma "scale"), and

random forest (hyperparameters: the number of trees in the forest "256", criterion to measure the quality of a split "gini", the maximum depth of the tree "3"), LASSO (hyperparameters: penalty"L1", solver "liblinear", C "0.5").

Of the 100 pairs of GDM/control subjects, 70 pairs (70%) were randomly selected as the training set, and the remaining 30 pairs (30%) were used as the testing set to evaluate the performance of the prediction models. To evaluate the diagnostic performance of each model, six metrics including area under the curve (AUC) in receiver-operating-characteristics (ROC), accuracy, precision, sensitivity, specificity and F1 score, were measured in the testing set. AUC (range 0 to 1) is a widely used index to describe a model's ability to predict outcomes. True positive (TP), true negative (TN), false positive (FP), and false negative (FN) numbers were used in calculating the accuracy, precision, sensitivity, specificity

and F1 score: accuracy = (TP+TN)/(TP+FP+TN+FN), precision=TP/(TP+FP), sensitivity=TP/(TP+FN), specificity=TN/(TN+FP), the F1 score is the harmonic mean of precision and sensitivity, F1=2TP/(2TP+FN+FP) (range 0–1).

Statistical analysis

Mean±SD was presented for continuous variables, and frequency (%) was presented for categorical variables. Paired t-test and Chi-square test were used to examine the differences between GDM and control groups in continuous and categorical variables respectively. Generalized linear models were used to comparing circulating concentrations of biomarkers between GDM and control subjects adjusting for maternal and pregnancy characteristics. Logistic regression was used to explore the associations of GDM with biomarkers. A stepwise regression was used in the selection of the co-variables in final regression models (retaining co-variables with $P<0.2$ only).

ML prediction models were fitted using R packages (*MASS, e1071, glmnet, randomForest*) in R studio and Python (*Sklearn*). Missing data were rare in all biomarkers (0–4%), and were imputed/replaced by the respective mean values to maintain consistent sample sizes in all models. All data analyses were performed using R Studio (version 4.2) and Python (version 3.10.9). P value<0.05 (two-tailed) was considered statistically significant.

Results

Table 1 presents the characteristics of study subjects. There were no differences between GDM and control subjects in maternal age, primiparity, pre-pregnancy BMI, family history of diabetes or hypertension. FPG, SBP and DBP at 11–14 weeks of gestation were all higher in GDM vs. controls (all $p<0.05$), while there were no differences in serum levels of TSH, FT3, FT4, TG and TC. Comparing subjects in the training (70 pairs) vs. testing sets for ML models (Table S1), there were no differences in all maternal characteristics and biomarkers, except for that average FPG tended to be slightly higher in the training set (5.0±0.4 vs. 4.8±0.4 mmol/L, $P=0.027$).

Adjusting for family history of diabetes and gestational age at blood sampling (other covariates did not affect the comparisons), serum concentrations at early gestation (11–14 weeks) were lower in GDM vs. euglycemic pregnancies for IGFBP-2 (74.7±25.8 vs. 87.2±32 ng/ml, $P=0.005$), HMW adiponectin (2147.7±1434.1 vs. 2480.5±1309.3 ng/ml, $p=0.03$) and SHBG (157.0±42.6 vs. 168.5±40.3 nmol/L, $P=0.05$) (Table 2), but higher for hs-CRP (5.54±5.25 vs. 4.60±4.87 mg/L, $P=0.03$) and IGF-I/IGFBP-2 ratios (0.8±0.6 vs. 0.6±0.4, $p=0.03$). There were no differences in serum glycosylated fibronectin (381.1±211.5 vs. 375.7±193.8 µg/mL, $P=0.80$),

Table 1 Characteristics of study subjects in GDM and euglycemic (control) singleton pregnancies

	GDM (n = 100)	Control (n = 100)	P*
Age (years)	32.1±3.9	32.0±3.8	0.96
>35 y	27 (27.0)	25 (25.0)	0.75
Primiparity	54 (54.0)	53 (53.0)	0.89
Family history of diabetes	2 (2.0)	0 (0.0)	0.15
Family history of hypertension	37 (37.0)	33 (33.0)	0.55
Pre-pregnancy BMI (kg/m²)	23.5±3.1	23.5±3.1	1.00
Pre-pregnancy BMI category			0.68
BMI < 24	53 (53.0)	58 (58.0)	
24.0 =< BMI < 28	40 (40.0)	34 (34.0)	
BMI > 28	7 (7.0)	8 (8.0)	
GA at blood sampling	13.2±1.2	13.0±1.1	0.15
Routine blood tests 11–14 weeks			
FPG	5.1±0.4	4.8±0.3	**< 0.001**
SBP (mmHg)	121.5±12.5	115.6±12.0	**0.001**
DBP (mmHg)	73.3±8.5	70.6±8.2	**0.024**
Triglyceride (mmo/L)	1.6±0.6	1.5±0.6	0.29
TC (mmo/L)	4.9±0.8	4.7±0.8	0.08
hemoglobin (g/L)	123.4±8.8	123.3±7.9	0.95
TSH (uIU/ml)	1.4±1.0	1.2±0.8	0.21
FT4 (pmol/l)	15.2±2.3	15.2±2.1	0.92
FT3 (pmol/l)	4.7±0.6	4.6±0.6	0.11
75 g OGTT (24-28weeks)			
Plasma glucose (mmo/L)			
fasting	5.1±0.4	4.4±0.4	**< 0.001**
1 h	9.2±1.9	7.8±1.2	**< 0.001**
2 h	7.3±1.7	6.4±1.2	**< 0.001**

Data presented are Mean±SD or n (%)

GDM, gestational diabetes mellitus; GA, gestational age; BMI, body mass index; FPG, fasting plasma glucose; SPB, systolic blood pressure; DBP, diastolic blood pressure; TC, total cholesterol; TSH, thyroid stimulating hormone; FT4, free thyroxine; FT3, free triiodothyronine; OGTT, oral glucose tolerance test

*P-values in comparisons of the two groups in t-tests for continuous variables and Chi-square tests for categorical variables

P values in **bold**, $p<0.05$

Yang *et al. BMC Pregnancy and Childbirth*   (2024) 24:601

Page 5 of 9

Table 2 Serum biomarkers in the first trimester (11–14 weeks) in GDM and euglycemic (control) pregnancies

	GDM	Control	Crude	Adjusted
	(*n* = 100)	(*n* = 100)	*P*	*P*
Hs-CRP (mg/L)	5.54 ± 5.25	4.60 ± 4.87	0.19	**0.03**
GFN (µg/mL)	381.1 ± 211.5	375.7 ± 193.8	0.83	0.80
IGF-I (ng/ml)	52.2 ± 25.4	48.8 ± 19.8	0.32	0.48
IGFBP-2 (ng/ml)	74.7 ± 25.8	87.2 ± 32	**0.004**	**0.005**
IGF-I/IGFBP-2 ratio	0.8 ± 0.6	0.6 ± 0.4	**0.03**	**0.03**
SHBG (nmol/L)	157.0 ± 42.6	168.5 ± 40.3	0.07	**0.05**
Adiponectin				
HMW (ng/ml)	2147.7 ± 1434.1	2480.5 ± 1309.3	0.11	**0.03**
Total (ng/ml)	5489.8 ± 1989.4	6107.4 ± 2437.5	0.07	0.16

Data presented are mean±SD.

GDM, gestational diabetes mellitus; Hs-CRP, High Sensitivity C Reactive Protein; GFN, glycosylated fibronectin; IGF-I, insulin-like growth factor-I; IGFBP-2, insulin-like growth factor binding protein 2; SHBG, sex hormone binding globulin; HMW, high molecular weight

Crude *p* values were from paired t-tests in the comparisons of log-transformed biomarker data. Adjusted *p* values were from generalized linear models adjusting for maternal family history of diabetes and gestational age at blood sampling; other co-variables were excluded since they were similar and did not affect the comparisons (all *p* > 0.2)

P values in **bold**, *p* < 0.05

IGF-I (52.2 ± 25.4 vs. 48.8 ± 19.8 ng/ml, *P* = 0.48) and total adiponectin (5489.8 ± 1989.4 vs. 6107.4 ± 2437.5 ng/ml, *P* = 0.16) concentrations in GDM and euglycemic pregnancies.

Higher serum IGFBP-2 and SHBG concentrations were associated with lower odds of GDM, while higher FPG, IGF-1/IGFBP-2 ratio, SBP and DBP were associated with higher odds of GDM (Table S2). Other observed biomarkers were not associated with GDM.

The four ML algorithms selected different unique sets of predictors. The LR identified 5 predictors (FPG, IGFBP-2, TSH, FT3, hemoglobin), the LASSO identified 2 predictors (FPG, IGFBP-2), SVM-RFE identified 5 predictors (FPG, IGFBP-2, TSH, Hs-CRP, hemoglobin), and random forest identified 5 predictors (FPG, IGFBP-2, Hs-CRP, IGF-I/IGFBP-2 ratio, HMW adiponectin) (Table S3). Two predictors (FPG, IGFBP-2) were selected in all the four ML methods. TSH, Hs-CRP and hemoglobin were each selected twice in the four ML models.

Prediction performance metrics were assessed in the training and testing sets in 4 different models: Model 1: FPG, IGFBP-2; Model 2: FPG, IGFBP-2, TSH; Model 3: FPG, IGFBP-2, hs-CRP; and Model 4: each ML algorithm selected set of predictors (Table 3). In the testing sets of all models, the combination of FPG and IGFBP-2 only (model 1) outperformed other combinations of features (Models 2, 3 and 4). The random forest model using FPG and IGFBP-2 was the optimal model with the highest AUC (0.8) among all the models with sensitivity = 87%. Other predictive performance metrics for these models are shown in Table 4. The addition of other predictors did not further improve the prediction performance.

The feature importance of variables for predicting gestational diabetes in random forest models is presented in Figure S2, and the SHAP (SHapley Additive exPlanations) summary plots illustrating the importance and impact of each feature in predicting gestational diabetes is presented in Figure S3. The FPG was the most important and impactful predictor, and IGFBP-2 ranked the second.

Discussion
Main findings
Our results suggest that at early gestation (11–14 weeks), a combination of FPG and IGFBP-2 could predict GDM with moderate discriminant power.

Most predictive biomarkers
ML model analyses showed that FPG and IGFBP-2 at early gestation were always selected in all ML models, and represented the minimal set of biomarkers in predicting GDM. FPG in the first-trimester is a known biomarker strongly predictive of GDM [14, 37, 38]. A prospective cohort study (*n*=450) reported an AUC of 0.61 for FPG [23], while in a large retrospective cohort (*n*=48,444), the AUC was 0.77 for first-trimester FPG in predicting GDM [15]. These consistent findings highlighted the importance of first trimester FPG in predicting GDM.

Circulating IGFBP-2 levels are lower in GDM vs. euglycemic pregnancies throughout the gestation [21, 39, 40]. Consistent with our results, low IGFBP-2 in early gestation was observed to be strongly predictive of later development of GDM in a previous study [21]. IGFBP-2 overexpression has been associated with reduced susceptibility to obesity and diabetes via inhibition of adipogenesis and stimulation of insulin sensitivity in mice [41, 42], suggesting a critical role of IGFBP-2 in metabolic homeostasis. There is a lack of post-prandial fluctuations in IGFBP-2 concentrations [43], rendering IGFBP-2 a promising early gestational biomarker of GDM.

Prediction models
In our study, the best prediction model was a combination of FPG and IGFBP-2, with moderate discriminant power (AUC=0.80) in identifying GDM. Some other observed biomarkers and features, although were associated with GDM, but their additions could not improve the predictions or discriminant power.

A number of studies have reported that biomarkers in early pregnancy may improve the accuracy in predicting GDM upon routine clinical risk factors. In a large multiethnic cohort, a model combining clinical risk factors (previous GDM, family history of diabetes, South/East Asian ethnicity, parity, maternal age, and BMI) and biomarkers (pregnancy-associated plasma protein A, uterine artery pulsatility index, MAP, and free β-human

Table 3 AUC and accuracy of conventional logistic regression (LR) and 5 machine learning (ML) models in predicting GDM by early gestational biomarkers

	Training set			Testing set		
	AUC		**Accuracy**	**AUC**		**Accuracy**
Model 1						
LR	0.78	0.72		0.73		0.67
ML-LR	0.76	0.66		0.77		0.75
KNN	0.77	0.70		0.77		0.68
SVM	0.76	0.68		0.77		0.73
RF	0.88	0.76		**0.80**		0.72
Lasso	0.76	0.67		0.77		0.73
Model 2						
LR	0.79	0.70		0.72		0.67
ML-LR	0.77	0.69		0.77		0.75
KNN	1.00	1.00		0.71		0.63
SVM	0.77	0.67		0.76		0.72
RF	1.00	0.99		0.71		0.65
Lasso	0.77	0.68		0.76		0.72
Model 3						
LR	0.77	0.70		0.72		0.65
ML-LR	0.76	0.66		0.75		0.70
KNN	0.76	0.71		0.74		0.65
SVM	0.77	0.69		0.76		0.68
RF	0.88	0.79		0.74		0.65
Lasso	0.76	0.66		0.75		0.68
Model 4						
LR	0.81	0.75		0.70		0.65
ML-LR	0.79	0.72		0.74		0.72
KNN	0.75	0.71		0.73		0.68
SVM	0.70	0.66		0.72		0.65
RF	0.91	0.81		0.77		0.73
Lasso	0.76	0.67		0.77		0.73

AUC, area under the curve; KNN, k-nearest neighbor; LR, logistic regression; SVM, support vector machine; RF, random forest; FPG, fasting plasma glucose; HMW, high molecular weight; IGFBP-2, insulin-like growth factor binding protein 2; hs-CRP, high sensitivity C reactive protein; thyroid stimulating hormone (TSH).

Model 1: FPG+IGFBP2; Model 2: FPG+IGFBP2+TSH; Model 3: FPG+IGFBP2+Hs-CRP;

Model 4: LR or ML-LR: FPG+IGFBP2+TSH+FT3+Hemoglobin, KNN: FPG+IGFBP2+hs-CRP+IGF-I/IGFBP-2 ratio+HMW adiponectin, SVM: FPG+IGFBP2+TSH+hs-CRP+Hemoglobin, Random forest: FPG+IGFBP2+hs-CRP+IGF-I/IGFBP-2 ratio+HMW adiponectin, Lasso: FPG+IGFBP2.

chorionic gonadotropin) achieved an AUC of 0.91 in predicting GDM [44]. In a large retrospective Chinese cohort, a 7-variables model (maternal age, family history of diabetes, multiple pregnancy, GDM history, FPG, HbA1c and triglyceride) achieved an AUC of 0.77 [45]. In a nested case-control study (GDM 80, control 300), the addition of SHBG and adiponectin might improve prediction beyond clinical risk factors (AUC: 0. 84 vs. 0.79) [46]. The AUC of a model including clinical factors (age, gestational age at sampling, BMI, ethnicity, family history of diabetes, and prior GDM) plus HDL cholesterol and tissue plasminogen activator was reported to be 0.86 in predicting GDM [47]. In a study of 1843 Belgian women, the AUC was 0.76 in a model including clinical factors, fasting plasma glucose, triglycerides and HbA1c [48]. A small study (*n*=20 GDM cases) reported that the combination of soluble CD163, insulin, tumor-necrosis factor alpha, placental protein 13, and pregnancy associated plasma protein A at the first trimester could predict DM with AUC=0.94 [49]. In a nested case-control study, the AUC was 0.70 in a model including chemerin, leptin, secreted frizzled-related protein 4 and adiponectin [50]. The AUC was 0.87 in predicting GDM in a model including total cholesterol, triglycerides, insulin, homeostatic model assessment, low density lipoprotein and tissue plasminogen activator in another small study (16 GDM cases) [51]. These variable findings may be partly attributable to the inclusion of different candidate biomarkers. In contrast, our model is much simpler, with 2-biomarkers only achieving an AUC of 0.80. We have considered the most promising candidate biomarkers according to up-to-date literature, and our study demonstrates the potential usefulness of a parsimonious model in predicting GDM using two biomarkers only (FPG and IGFBP-2).

Yang *et al. BMC Pregnancy and Childbirth*   (2024) 24:601

Page 7 of 9

Table 4 The performance of logistic regression and 5 machine learning models in predicting GDM by early gestational biomarkers

	TP	TN	FP	FN	Precision	Sensitivity	Specificity	F_1 Score
Model 1								
LR	23	17	13	7	0.64	0.76	0.57	0.70
ML-LR	25	20	10	5	0.71	0.83	0.67	0.77
KNN	25	16	14	5	0.64	0.83	0.53	0.72
SVM	25	19	11	5	0.69	0.83	0.63	0.76
RF	26	17	13	4	0.67	0.87	0.57	0.76
Lasso	24	20	10	6	0.71	0.80	0.67	0.75
Model 2								
LR	22	18	12	8	0.65	0.77	0.60	0.69
ML-LR	25	20	10	5	0.71	0.83	0.67	0.77
KNN	21	17	13	9	0.62	0.70	0.57	0.74
SVM	25	18	12	5	0.68	0.83	0.60	0.75
RF	22	17	13	8	0.63	0.73	0.73	0.68
Lasso	25	18	12	5	0.68	0.83	0.60	0.75
Model 3								
LR	24	15	15	6	0.62	0.80	0.50	0.70
ML-LR	23	19	11	7	0.68	0.77	0.63	0.72
KNN	22	18	12	8	0.65	0.73	0.60	0.69
SVM	23	18	12	7	0.66	0.77	0.60	0.71
RF	22	17	13	8	0.63	0.73	0.57	0.68
Lasso	23	18	12	7	0.66	0.77	0.60	0.71
Model 4								
LR	24	15	15	6	0.62	0.80	0.50	0.70
ML-LR	23	20	10	7	0.66	0.77	0.60	0.71
KNN	24	17	13	6	0.62	0.70	0.57	0.66
SVM	20	19	11	10	0.67	0.67	0.67	0.67
RF	25	19	11	5	0.61	0.83	0.47	0.70
Lasso	24	20	10	6	0.71	0.80	0.67	0.75

TP, true positive; TN, true negative; FN, false negative; FP, false positive; KNN, k-nearest neighbor; LR, logistic regression; MR-LR, machine learning logistic regression; SVM, support vector machine; RF, random forest; FPG, fasting plasma glucose; HMW, high molecular weight; IGFBP-2, insulin-like growth factor binding protein 2; hs-CRP, high sensitivity C reactive protein; thyroid stimulating hormone (TSH).

Model 1: FPG+IGFBP2; Model 2: FPG+IGFBP2+TSH; Model 3: FPG+IGFBP2+Hs-CRP;

Model 4: LR or ML-LR: FPG+IGFBP2+TSH+FT3+Hemoglobin, KNN: FPG+IGFBP2+hs-CRP+IGF-I/IGFBP-2 ratio+HMW adiponectin, SVM: FPG+IGFBP2+TSH+hs-CRP+Hemoglobin, Random forest: FPG+IGFBP2+hs-CRP+IGF-I/IGFBP-2 ratio+HMW adiponectin, Lasso: FPG+IGFBP2.

Conventional LR vs. ML models

A meta-analysis showed that LR was the most commonly used model in predicting GDM: 25 studies using conventional LR, vs. 5 studies using ML with an overall pooled AUC of 0.89 [29]. Our ML model requires the smallest number of predictors (only two). In a large prospective cohort, ML method (XGBoost) achieved a higher AUC than conventional LR in predicting GDM [52]. On the other hand, two other studies indicated that ML did not outperform conventional LR in predicting GDM [45, 53]. In our study, all ML models achieved higher AUC than conventional LR. ML models may perform differently in different datasets and algorithms [53–56]. We tested five common ML algorithms, and the random forest was identified as the best model with the highest AUC.

Limitations

The study was based on a large pregnancy cohort. The proposed ML prediction model requires two predictors only, and that adding various clinical features or other biomarkers did *not* improve the predictions. All study subjects are Chinese. More studies in other ethnicity groups are required to understand the generalizability of the study findings.

In summary, the combination of FPG and IGFBP-2 at early gestation can predict later development of GDM with moderate discriminant power. Further validation studies are warranted to assess the utility of this simple combination model in other independent cohorts in the quest for a potentially useful clinical risk monitoring tool.

Abbreviations
GDM   Gestational diabetes mellitus
IGFBP-2   Insulin-like growth factor binding protein 2
OGTT   Oral glucose tolerance test

Supplementary Information

Yang *et al. BMC Pregnancy and Childbirth*   (2024) 24:601

Page 8 of 9

Supplementary Material 1

Data availability
Access to the de-identified participant research data must be approved by the research ethics board on a case-by-case basis, please contact the corresponding authors (zc.luo@utoronto.ca; ouyangfengxiu@xinhuamed.com.cn; junjimzhang@sjtu.edu.cn) for assistance in data access request.

Declarations

Ethics approval and consent to participate
The study was approved by the research ethics committee of Xinhua Hospital, School of Medicine, Shanghai Jiao Tong University (reference number XHEC-C-2013-001). Written informed consent was obtained from all study participants.

Consent for publication
Not applicable.

Competing interests
The authors declare no competing interests.

Author details
[1]Ministry of Education-Shanghai Key Laboratory of Children's Environmental Health, Department of Pediatrics, Xinhua Hospital, Early Life Health Institute, Shanghai Jiao-Tong University School of Medicine, Kong-Jiang Road, Shanghai 200092, China
[2]Prosserman Centre for Population Health Research, Department of Obstetrics and Gynecology, Mount Sinai Hospital, Faculty of Medicine, Lunenfeld-Tanenbaum Research Institute, University of Toronto, L5-240, Murray Street 60, Toronto, ON M5T 3H7, Canada
[3]Obstetrics and Gynecology, International Peace Maternity and Child Health Hospital, Shanghai Jiao-Tong University School of Medicine, Shanghai 200030, China
[4]Clinical Skills Center, School of Clinical Medicine, Shandong First Medical University, Shandong Academy of Medical Sciences, Jinan, China
[5]Obstetrics and Gynecology, Xinhua Hospital, Shanghai Jiao-Tong University School of Medicine, Shanghai 200092, China
[6]Department of Clinical Assay Laboratory, Xinhua Hospital, Shanghai Jiao-Tong University School of Medicine, Shanghai 200092, China
[7]State Key Laboratory of Reproductive Medicine, Department of Epidemiology, School of Public Health, Nanjing Medical University, Nanjing, China

References
1. Chiefari E, Arcidiacono B, Foti D, Brunetti A. Gestational diabetes mellitus: an updated overview. J Endocrinol Invest. 2017;40(9):899–909.
2. ACOG Practice Bulletin No. 190: gestational diabetes mellitus. Obstet Gynecol. 2018;131(2):e49–64.
3. Gabbe SG. Gestational diabetes mellitus. N Engl J Med. 1986;315(16):1025–6.
4. Yogev Y, Visser G. Obesity, gestational diabetes and pregnancy outcome. Semin Fetal Neonatal Med. 2009;14(2):77–84.
5. Patel S, Fraser A, Davey Smith G, Lindsay R, Sattar N, Nelson S, Lawlor D. Associations of gestational diabetes, existing diabetes, and glycosuria with offspring obesity and cardiometabolic outcomes. Diabetes Care. 2012;35(1):63–71.
6. Zhao P, Liu E, Qiao Y, Katzmarzyk P, Chaput J, Fogelholm M, Johnson W, Kuriyan R, Kurpad A, Lambert E, Maher C, Maia J, Matsudo V, Olds T, Onywera V, Sarmiento O, Standage M, Tremblay M, Tudor-Locke C, Hu G. Maternal gestational diabetes and childhood obesity at age 9–11: results of a multinational study. Diabetologia. 2016;59(11):2339–48.
7. Lowe W, Scholtens D, Lowe L, Kuang A, Nodzenski M, Talbot O, Catalano P, Linder B, Brickman W, Clayton P, Deerochanawong C, Hamilton J, Josefson J, Lashley M, Lawrence J, Lebenthal Y, Ma R, Maresh M, McCance D, Tam W, Sacks D, Dyer A, Metzger B. Association of gestational diabetes with maternal disorders of glucose metabolism and childhood adiposity. JAMA. 2018;320(10):1005–16.
8. Philipps L, Santhakumaran S, Gale C, Prior E, Logan K, Hyde M, Modi N. The diabetic pregnancy and offspring BMI in childhood: a systematic review and meta-analysis. Diabetologia. 2011;54(8):1957–66.
9. Song C, Li J, Leng J, Ma RC, Yang X. Lifestyle intervention can reduce the risk of gestational diabetes: a meta-analysis of randomized controlled trials. Obes Rev. 2016;17(10):960–9.
10. Tieu J, Shepherd E, Middleton P, Crowther CA. Dietary advice interventions in pregnancy for preventing gestational diabetes mellitus. Cochrane Database Syst Rev. 2017;1(1):Cd006674.
11. Metzger BE, Gabbe SG, Persson B, Buchanan TA, Catalano PA, Damm P, Dyer AR, Leiva A, Hod M, Kitzmiler JL, Lowe LP, McIntyre HD, Oats JJ, Omori Y, Schmidt MI. International association of diabetes and pregnancy study groups recommendations on the diagnosis and classification of hyperglycemia in pregnancy. Diabetes Care. 2010;33(3):676–82.
12. Kautzky-Willer A, Harreiter J, Bancher-Todesca D, Berger A, Repa A, Lechleitner M, Weitgasser R. Gestational diabetes mellitus. Wiener Klinische Wochenschrift. 2016;128(Suppl 2):S103–12.
13. Gude NM, Roberts CT, Kalionis B, King RG. Growth and function of the normal human placenta. Thromb Res. 2004;114(5–6):397–407.
14. Sesmilo G, Prats P, Garcia S, Rodríguez I, Rodríguez-Melcón A, Berges I, Serra B. First-trimester fasting glycemia as a predictor of gestational diabetes (GDM) and adverse pregnancy outcomes. Acta Diabetol. 2020;57(6):697–703.
15. Tong J, Wu L, Chen Y, Guan X, Tian F, Zhang H, Liu K, Yin A, Wu X, Prof J. Fasting plasma glucose in the first trimester is related to gestational diabetes mellitus and adverse pregnancy outcomes. Endocrine. 2022;75(1):70–81.
16. Thadhani R, Wolf M, Hsu-Blatman K, Sandler L, Nathan D, Ecker J. First-trimester sex hormone binding globulin and subsequent gestational diabetes mellitus. Am J Obstet Gynecol. 2003;189(1):171–6.
17. Smirnakis K, Plati A, Wolf M, Thadhani R, Ecker J. Predicting gestational diabetes: choosing the optimal early serum marker. Am J Obstet Gynecol. 2007;196(4):410.e1-6; discussion.e6-7.
18. Lacroix M, Battista M, Doyon M, Ménard J, Ardilouze J, Perron P, Hivert M. Lower adiponectin levels at first trimester of pregnancy are associated with increased insulin resistance and higher risk of developing gestational diabetes mellitus. Diabetes Care. 2013;36(6):1577–83.
19. Williams M, Qiu C, Muy-Rivera M, Vadachkoria S, Song T, Luthy D. Plasma adiponectin concentrations in early pregnancy and subsequent risk of gestational diabetes mellitus. J Clin Endocrinol Metab. 2004;89(5):2306–11.
20. Lain K, Daftary A, Ness R, Roberts J. First trimester adipocytokine concentrations and risk of developing gestational diabetes later in pregnancy. Clin Endocrinol (Oxf). 2008;69(3):407–11.
21. Zhu Y, Mendola P, Albert P, Bao W, Hinkle S, Tsai M, Zhang C. Insulin-like growth factor axis and gestational diabetes mellitus: a longitudinal study in a multiracial cohort. Diabetes. 2016;65(11):3495–504.
22. Wang X, Wang W, Yu X, Hua X, Ouyang F, Luo Z. Insulin-like growth factor axis biomarkers and gestational diabetes mellitus: a systematic review and meta-analysis. Front Endocrinol (Lausanne). 2019;10:444.
23. Ozgu-Erdinc A, Yilmaz S, Yeral M, Seckin K, Erkaya S, Danisman A. Prediction of gestational diabetes mellitus in the first trimester: comparison of C-reactive

protein, fasting plasma glucose, insulin and insulin sensitivity indices. J Matern Fetal Neonatal Med. 2015;28(16):1957–62.

24. Kansu-Celik H, Ozgu-Erdinc A, Kisa B, Findik R, Yilmaz C, Tasci Y. Prediction of gestational diabetes mellitus in the first trimester: comparison of maternal fetuin-A, N-terminal proatrial natriuretic peptide, high-sensitivity C-reactive protein, and fasting glucose levels. Arch Endocrinol Metab. 2019;63(2):121–7.

25. Huhn E, Fischer T, Göbl C, Todesco Bernasconi M, Kreft M, Kunze M, Schoetzau A, Dölzlmüller E, Eppel W, Husslein P, Ochsenbein-Koelble N, Zimmermann R, Bäz E, Prömpeler H, Bruder E, Hahn S, Hoesli I. Screening of gestational diabetes mellitus in early pregnancy by oral glucose tolerance test and glycosylated fibronectin: study protocol for an international, prospective, multicentre cohort trial. BMJ open. 2016;6(10):e012115.

26. Rasanen J, Snyder C, Rao P, Mihalache R, Heinonen S, Gravett M, Roberts C, Nagalla S. Glycosylated fibronectin as a first-trimester biomarker for prediction of gestational diabetes. Obstet Gynecol. 2013;122(3):586–94.

27. Benhalima K, Van Crombrugge P, Moyson C, Verhaeghe J, Vandeginste S, Verlaenen H, Vercammen C, Maes T, Dufraimont E, De Block C, Jacquemyn Y, Mekahli F, De Clippel K, Van Den Bruel A, Loccufier A, Laenen A, Minschart C, Devlieger R, Mathieu C. Estimating the risk of gestational diabetes mellitus based on the 2013 WHO criteria: a prediction model based on clinical and biochemical variables in early pregnancy. Acta Diabetol. 2020;57(6):661–71.

28. Falcone V, Kotzaeridi G, Breil MH, Rosicky I, Stopp T, Yerlikaya-Schatten G, Feichtinger M, Eppel W, Husslein P, Tura A, Göbl CS. Early assessment of the risk for gestational diabetes mellitus: can fasting parameters of glucose metabolism contribute to risk prediction? Diabetes Metab J. 2019;43(6):785–93.

29. Zhang Z, Yang L, Han W, Wu Y, Zhang L, Gao C, Jiang K, Liu Y, Wu H. Machine learning prediction models for gestational diabetes mellitus: meta-analysis. J Med Internet Res. 2022;24(3):e26634.

30. Lamain-de Ruiter M, Kwee A, Naaktgeboren CA, de Groot I, Evers IM, Groenendaal F, Hering YR, Huisjes AJ, Kirpestein C, Monincx WM, Siljee JE, Van 't Zelfde A, van Oirschot CM, Vankan-Buitelaar SA, Vonk MA, Wiegers TA, Zwart JJ, Franx A, Moons KG, Koster MP. External validation of prognostic models to predict risk of gestational diabetes mellitus in one Dutch cohort: prospective multicentre cohort study. BMJ. 2016;354:i4338.

31. Nombo AP, Mwanri AW, Brouwer-Brolsma EM, Ramaiya KL, Feskens EJM. Gestational diabetes mellitus risk score: a practical tool to predict gestational diabetes mellitus risk in Tanzania. Diabetes Res Clin Pract. 2018;145:130–7.

32. Obermeyer Z, Emanuel EJ. Predicting the future - big data, machine learning, and clinical medicine. N Engl J Med. 2016;375(13):1216–9.

33. Metzger B, Gabbe S, Persson B, Buchanan T, Catalano P, Damm P, Dyer A, Leiva A, Hod M, Kitzmiler J, Lowe L, McIntyre H, Oats J, Omori Y, Schmidt M. International association of diabetes and pregnancy study groups recommendations on the diagnosis and classification of hyperglycemia in pregnancy. Diabetes Care. 2010;33(3):676–82.

34. Guyon I, Weston J, Barnhill S, Vapnik V. Gene selection for cancer classification using support vector machines. Mach Learn. 2002;46(1):389–422.

35. Tibshirani R. Regression shrinkage and selection via the lasso. J R Stat Soc Ser B Stat Methodol. 1996;58(1):267–88.

36. Rigatti SJ. Random forest. J Insur Med. 2017;47(1):31–9.

37. Zhu W, Yang H, Wei Y, Yan J, Wang Z, Li X, Wu H, Li N, Zhang M, Liu X, Zhang H, Wang Y, Niu J, Gan Y, Zhong L, Wang Y, Kapur A. Evaluation of the value of fasting plasma glucose in the first prenatal visit to diagnose gestational diabetes mellitus in China. Diabetes Care. 2013;36(3):586–90.

38. Riskin-Mashiah S, Younes G, Damti A, Auslender R. First-trimester fasting hyperglycemia and adverse pregnancy outcomes. Diabetes Care. 2009;32(9):1639–43.

39. Lappas M. Insulin-like growth factor-binding protein 1 and 7 concentrations are lower in obese pregnant women, women with gestational diabetes and their fetuses. J Perinatol. 2015;35(1):32–8.

40. Gęca T, Kwaśniewska A. The influence of gestational diabetes mellitus upon the selected parameters of the maternal and fetal system of insulin-like growth factors (IGF-1, IGF-2, IGFBP1-3)-A review and a clinical study. J Clin Med. 2020;9(10).

41. Hedbacker K, Birsoy K, Wysocki R, Asilmaz E, Ahima R, Farooqi I, Friedman J. Antidiabetic effects of IGFBP2, a leptin-regulated gene. Cell Metab. 2010;11(1):11–22.

42. Wheatcroft S, Kearney M, Shah A, Ezzat V, Miell J, Modo M, Williams S, Cawthorn W, Medina-Gomez G, Vidal-Puig A, Sethi J, Crossey P. IGF-binding protein-2 protects against the development of obesity and insulin resistance. Diabetes. 2007;56(2):285–94.

43. Clemmons D, Snyder D, Busby W. Variables controlling the secretion of insulin-like growth factor binding protein-2 in normal human subjects. J Clin Endocrinol Metab. 1991;73(4):727–33.

44. Sweeting A, Wong J, Appelblom H, Ross G, Kouru H, Williams P, Sairanen M, Hyett J. A novel early pregnancy risk prediction model for gestational diabetes mellitus. Fetal Diagn Ther. 2019;45(2):76–84.

45. Wu Y, Zhang C, Mol B, Kawai A, Li C, Chen L, Wang Y, Sheng J, Fan J, Shi Y, Huang H. Early prediction of gestational diabetes mellitus in the Chinese population via advanced machine learning. J Clin Endocrinol Metab. 2021;106(3):e1191–205.

46. Nanda S, Savvidou M, Syngelaki A, Akolekar R, Nicolaides K. Prediction of gestational diabetes mellitus by maternal factors and biomarkers at 11 to 13 weeks. Prenat Diagn. 2011;31(2):135–41.

47. Savvidou M, Nelson S, Makgoba M, Messow C, Sattar N, Nicolaides K. First-trimester prediction of gestational diabetes mellitus: examining the potential of combining maternal characteristics and laboratory measures. Diabetes. 2010;59(12):3017–22.

48. Benhalima K, Crombrugge PV, Moyson C, Verhaeghe J, Mathieu C. Estimating the risk of gestational diabetes mellitus based on the 2013 WHO criteria: a prediction model based on clinical and biochemical variables in early pregnancy. Acta Diabetol. 2020;57(Suppl 1).

49. Tenenbaum-Gavish K, Sharabi-Nov A, Binyamin D, Møller H, Danon D, Rothman L, Hadar E, Idelson A, Vogel I, Koren O, Nicolaides K, Gronbaek H, Meiri H. First trimester biomarkers for prediction of gestational diabetes mellitus. Placenta. 2020;101:80–9.

50. Schuitemaker J, Beernink R, Franx A, Cremers T, Koster M. First trimester secreted frizzled-related protein 4 and other adipokine serum concentrations in women developing gestational diabetes mellitus. PLoS ONE. 2020;15(11):e0242423.

51. Correa P, Venegas P, Palmeiro Y, Albers D, Rice G, Roa J, Cortez J, Monckeberg M, Schepeler M, Osorio E, Illanes S. First trimester prediction of gestational diabetes mellitus using plasma biomarkers: a case-control study. J Perinat Med. 2019;47(2):161–8.

52. Liu H, Li J, Leng J, Wang H, Liu J, Li W, Liu H, Wang S, Ma J, Chan J, Yu Z, Hu G, Li C, Yang X. Machine learning risk score for prediction of gestational diabetes in early pregnancy in Tianjin, China. Diabetes Metab Res Rev. 2021;37(5):e3397.

53. Ye Y, Xiong Y, Zhou Q, Wu J, Li X, Xiao X. Comparison of machine learning methods and conventional logistic regressions for predicting gestational diabetes using routine clinical data: a retrospective cohort study. Journal of diabetes research. 2020;2020:4168340.

54. Xiong Y, Lin L, Chen Y, Salerno S, Li Y, Zeng X, Li H. Prediction of gestational diabetes mellitus in the first 19 weeks of pregnancy using machine learning techniques. J Matern Fetal Neonatal Med. 2022;35(13):2457–63.

55. Zheng T, Ye W, Wang X, Li X, Zhang J, Little J, Zhou L, Zhang L. A simple model to predict risk of gestational diabetes mellitus from 8 to 20 weeks of gestation in Chinese women. BMC Pregnancy Childbirth. 2019;19(1):252.

56. Savona-Ventura C, Vassallo J, Marre M, Karamanos BG. A composite risk assessment model to screen for gestational diabetes mellitus among Mediterranean women. Int J Gynaecol Obstet. 2013;120(3):240–4.

Publisher's Note