**Title Page**

1. Article Type: Original Paper

2. Title of paper:

Machine Learning Derived Prenatal Predictive Risk Model to Guide Intervention and Prevent the Progression of Gestational Diabetes Mellitus to Type 2 Diabetes

3. Authors’ names:

Mukkesh Kumar1,2,3, Li Ting Ang1,2, Cindy Ho1,2, Shu E Soh4, Kok Hian Tan5,6, Jerry Kok Yen Chan7,8,9, Keith M Godfrey10, Shiao-yng Chan1,7, Yap Seng Chong1,7, Johan G Eriksson1,7,11,12^, Mengling Feng3\*^, Neerja Karnani1,2,13\*^

\* Co-corresponding Authors: Mengling Feng, Neerja Karnani

^ Joint Senior Authors: Johan G Eriksson, Mengling Feng, Neerja Karnani

4. Institution(s) of origin:

1 Singapore Institute for Clinical Sciences, Agency for Science Technology and Research, Singapore, Republic of Singapore

2 Bioinformatics Institute, Agency for Science Technology and Research, Singapore, Republic of Singapore

3 Saw Swee Hock School of Public Health and Institute of Data Science, National University of Singapore and National University Health System, Singapore, Republic of Singapore

4 Department of Paediatrics, Yong Loo Lin School of Medicine, National University of Singapore, Singapore, Republic of Singapore

5 Division of Obstetrics and Gynecology, KK Women’s and Children’s Hospital, Republic of Singapore

6 Obstetrics and Gynecology Academic Clinical Programme, Duke-NUS Graduate Medical School, Singapore, Republic of Singapore

7 Department of Obstetrics and Gynaecology and Human Potential Translational Research Programme, Yong Loo Lin School of Medicine, National University of Singapore, Singapore, Republic of Singapore

8 Department of Reproductive Medicine, KK Women’s and Children’s Hospital, Republic of Singapore

9 Cancer and Stem Cell Biology, Duke-NUS Medical School, Singapore, Republic of Singapore

10 MRC Lifecourse Epidemiology Unit & NIHR Southampton Biomedical Research Centre, University of Southampton & University Hospital Southampton NHS Foundation Trust, UK

11 Department of General Practice and Primary Health Care, University of Helsinki, Finland

12 Folkhälsan Research Center, Helsinki, Finland

13 Department of Biochemistry, Yong Loo Lin School of Medicine, National University of Singapore, Singapore, Republic of Singapore

5. ORCID identifiers where possible:

Mukkesh Kumar: 0000-0001-6640-4284

Li Ting Ang: 0000-0003-2543-9055

Cindy Ho: 0000-0002-5748-8022

Shu E Soh: 0000-0002-0790-4200

Kok Hian Tan: 0000-0003-1945-0266

Jerry Kok Yen Chan: 0000-0002-0246-2159

Keith M Godfrey: 0000-0002-4643-0618

Shiao-yng Chan: 0000-0002-3530-3023

Yap Seng Chong: 0000-0002-7232-8473

Johan G Eriksson: 0000-0002-2516-2060

Mengling Feng: 0000-0002-5338-6248

Neerja Karnani: 0000-0002-5301-409X

6. Co-corresponding authors and email address:

Mengling Feng: ephfm@nus.edu.sg

Neerja Karnani: neerja\_karnani@sics.a-star.edu.sg

**Abstract**

***Background:***

The increasing prevalence of Gestational Diabetes Mellitus (GDM) is concerning as women with GDM are at high risk of later Type 2 Diabetes (T2D). The magnitude of this risk highlights the importance of intervening early to prevent progression of GDM to T2D. Rates of postpartum screening are sub-optimal, often as low as 13% in Asian countries. The lack of preventive care through structured postpartum screening in several healthcare systems and low public awareness are key barriers.

***Objective:***

In this study, we developed a machine learning model for early prediction of postpartum T2D following routine antenatal GDM screening. The early prediction of postpartum T2D during prenatal care would enable strategies to effectively implement diabetes prevention interventions. To our best knowledge, this is the first study applying machine learning for postpartum T2D risk assessment in antenatal populations of Asian origin.

***Methods:***

Prospective multi-ethnic data (Chinese, Malay and Indian ethnicities) from 561 pregnancies in Singapore’s most deeply phenotyped mother-offspring cohort study, Growing Up in Singapore Towards healthy Outcomes (GUSTO), was used for predictive modeling. The feature variables included were demographics, medical/obstetric history, physical measures, lifestyle information and GDM diagnosis. Shapley values were combined with CatBoost tree ensembles to perform feature selection. Our game theoretical approach for predictive analytics enables population subtyping and pattern discovery for data-driven precision care. The predictive models were trained using 4 machine learning algorithms: logistic regression, support vector machine, CatBoost gradient boosting and artificial neural network. We used 5-fold stratified cross validation to preserve the same proportion of T2D cases in each fold. Grid search pipelines were built to evaluate the best performing hyperparameters.

***Results:***

A high performance prediction model for postpartum T2D, comprising of 2 mid-gestation features (mid-pregnancy BMI after gestational weight gain; diagnosis of GDM) was developed [‘BMI\_GDM’ CatBoost model AUC: 0.86 (95% CI 0.72, 0.99)]. Pre-pregnancy BMI alone was inadequate in predicting postpartum T2D risk [‘ppBMI’ CatBoost model AUC: 0.62 (0.39, 0.86)]. 2-hour postprandial glucose [‘BMI\_2hour’ CatBoost model AUC: 0.86 (0.76, 0.96)] showed a stronger postpartum T2D risk prediction effect compared to fasting glucose [‘BMI\_Fasting’ CatBoost model AUC: 0.76 (0.61, 0.91)]. The ‘BMI\_GDM’ model was also robust when using a modified two-point IADPSG 2018 criteria for GDM diagnosis [‘BMI\_GDM2’ CatBoost model AUC: 0.84 (0.72, 0.97)]. Total gestational weight gain was inversely associated with postpartum T2D outcome, independent of pre-pregnancy BMI and diagnosis of GDM [*P*=.019, OR: 0.88 (0.79-0.98)].

***Conclusions:***

Mid-gestation weight gain effects combined with the metabolic derangements underlying GDM during pregnancy signal future T2D risk in Singaporean women. Further studies will be required to examine the influence of metabolic adaptations in pregnancy on postpartum maternal metabolic health outcomes. The state-of-the art machine learning model can be leveraged as a rapid risk stratification tool during prenatal care.

***Trial Registration:***

ClinicalTrials.gov NCT01174875

[449 words]

**Keywords**

Asian Populations; Diabetes Management; Digital Health; Gestational Diabetes Mellitus; Machine Learning; Prediction Models; Prenatal Care; Public Health; Risk Factors; Type 2 Diabetes

## Introduction

The prevalence of Gestational Diabetes Mellitus (GDM) is increasing globally, with 1 in 6 pregnancies being affected [1]. GDM has long-term implications as women with a history of GDM have a 10-fold higher risk of developing Type 2 Diabetes (T2D) compared to those with a normoglycemic pregnancy [2]. In the Growing Up in Singapore Towards healthy Outcomes (GUSTO) study, GDM women had a 12-fold higher risk of developing T2D 4-6 years post-delivery compared with non-GDM women [3]. From a public health perspective, early intervention in women with GDM could contribute to tackling the rising global health burden of T2D. The T2D epidemic is of particular concern in South-East Asia; 88 million adults are currently living with diabetes, but this is expected to increase to 153 million by 2045 [1]. 57% of the population with diabetes in South-East Asia are undiagnosed, increasing the risk of complications such as heart disease and stroke [1].

The American Diabetes Association (ADA) guidelines recommend that GDM women are tested 4-12 weeks postpartum using a 75g Oral Glucose Tolerance Test (OGTT) [4]. Further testing is recommended in those with normal postpartum OGTT every 1-3 years using fasting plasma glucose, HbA1c or a OGTT [4]. However, as GDM resolves post-pregnancy, postpartum surveillance of glycemia remains low across healthcare systems globally. The rate of postpartum diabetes screening can be as low as 13% in Asian countries [5]. Barriers to postpartum diabetes screening include lack of structured postpartum preventive care in healthcare systems, lack of patient awareness of future T2D risk and time restrictions due to maternal commitments [5, 6].

Machine learning models enable predictive population risk stratification. In a prospective metabolomics study by Allalou et al., 21 metabolites were identified at 6-9 weeks postpartum to predict the transition from GDM to T2D in women [7]. The metabolite model using decision trees performed well with an AUC of 0.77. In another GDM to T2D transition study by Joglekar et al., the inclusion of circulating microRNA (miR-369-3p) at 12 weeks postpartum enhanced the prediction of a clinical model (age, BMI, pregnancy fasting glucose, postpartum fasting glucose, cholesterol and triacylglycerols) from an AUC of 0.83 to AUC of 0.92 (logistic regression algorithm) [8]. In addition to low compliance of postpartum testing in GDM women, the other barriers for real world implementation of these two machine learning models includes cost and access to metabolomics assay & microRNA PCR during routine clinical visits.

The early prediction of postpartum T2D during prenatal care would enable strategies to effectively implement diabetes prevention interventions. To date, there have been no studies applying machine learning for postpartum T2D risk assessment in antenatal populations of Asian origin. In this Singaporean study, we developed a machine learning model for early prediction of postpartum T2D during routine antenatal GDM screening. Our machine learning model was implemented using prospective GUSTO cohort study data (ClinicalTrials.gov NCT01174875).

## Methods

### Study Design

GUSTO is a prospective multi-ethnic mother-offspring cohort study (Chinese, Malay or Indian ethnicities). Mothers were recruited during early pregnancy from Singapore’s two major public maternity hospitals; National University Hospital (NUH) and KK Women’s and Children’s Hospital (KKH) between June 2009 to October 2010. The study have been reviewed by the National Healthcare Group (NHG) Domain Specific Review Board for ethics approval and SingHealth Centralized Institutional Review Board (CIRB/E/2019/2655).

Participants of mixed ethnicity or with self-reported T2D at recruitment were excluded from model training. 561 mothers had complete data on demographics, medical/obstetric history, physical measures, lifestyle information, antenatal OGTT and postpartum OGTT 4 to 8 years after delivery. World Health Organization (WHO) 1999 criteria were used to diagnose GDM [9] and WHO 2006 criteria to diagnose postpartum Impaired Glucose Tolerance (IGT), Impaired Fasting Glucose (IFG) and Type 2 Diabetes (T2D) [10]. Abnormal Glucose Metabolism (AGM) outcome comprises of IGT, IFG and T2D diagnoses.

### Feature Variables

Information on demographics (maternal age, maternal ethnicity) and medical/obstetric history (self-reported pre-pregnancy weight, family history of diabetes mellitus, family history of high blood pressure, family history of cardiovascular disease, previous history of gestational diabetes mellitus, previous history of gestational hypertension and parity) were derived from first trimester questionnaires. Systolic and diastolic blood pressure were recorded at mid-gestation (median = 26.7 weeks, IQR = 26.1-27.6 weeks) and obtained from hospital case notes. Mean arterial blood pressure was derived by doubling the diastolic blood pressure and adding to the systolic blood pressure, with the composite sum divided by 3. Maternal anthropometry was measured at mid-gestation (median = 26.9 weeks, IQR = 26.4-27.6 weeks). Maternal mid-upper arm circumference was measured to the nearest 0.1 cm, midway between acromion process and olecranon process (SECA 212). Maternal height was measured to the nearest 0.1cm (SECA 213). Maternal weight at mid-pregnancy was measured to the nearest 0.1kg (SECA 803), and BMI derived using weight divided by height squared (kg/m2). Total gestational weight gain was derived by subtracting first antenatal visit weight (median = 9.0 weeks, IQR = 7.3-11.0 weeks) from last antenatal visit weight (median = 38.1 weeks, IQR = 37.3-39.1 weeks). Lifestyle information on self-reported smoking, environmental tobacco smoke exposures and alcohol consumption were collected using questionnaires. GDM diagnosis was based on antenatal OGTT assessment (median = 26.9 weeks, IQR = 26.4-27.7 weeks).

### Machine Learning Methodology and Statistical Analyses

Our methodological novelty lies in combining coalitional game theory concepts with machine learning. SHapley Additive exPlanations (SHAP) framework was combined with CatBoost tree ensembles for feature selection and model explainability [11, 12]. SHapley Additive exPlanations (SHAP) framework connect optimal credit allocation with local explanations using the classic Shapley values from cooperative game theory. Lundberg and Lee have proposed SHAP as the only additive feature attribution method that satisfies two important properties of game theory - additivity (local accuracy) and monotonicity (consistency) [11]. In game theory, Shapley value is the average expected marginal contribution of one player across all possible permutation of players (average effects of team member composition and team size). Shapley value helps to determine a payoff for all the game players when each player might have contributed more or less than the others when working in coalition. In machine learning, game players are the features and collective payout is the model prediction. SHAP framework provides local explanations based on exact Shapley values to understand the global model structure. For each of all possible feature orderings, features are introduced one at a time into a conditional expectation function of the model’s output and changes in expectation are attributed to the introduced feature, averaged over all possible feature orderings in a fair manner. SHAP values represent a change in log odds ratio. Our game theoretical approach for predictive analytics enables population subtyping and pattern discovery for data-driven precision care.

The supervised machine learning models were built using Anaconda’s distribution of Python v3.7.9 programming language in JupyterLab computational environment. The predictive models were trained using 4 machine learning algorithms to address algorithm bias; logistic regression (generalized linear model), support vector machine (linear support vector classification), CatBoost gradient boosting (tree-based) and artificial neural network (multilayer perceptron). We used 5-fold stratified cross validation to preserve the same proportion of AGM/T2D cases in each fold. Maximum absolute scaler was used as a preprocessor to scale each feature without destroying the sparsity. A grid search pipeline was built to evaluate the best performing hyperparameters for each machine learning model. Model performances were evaluated using the area under the receiver operating characteristic curve (AUC) with 95% confidence interval. Implementation details of machine learning algorithms are included in Appendix 1.

The feature selection model using clinical features at mid-gestation was trained on AGM outcome and top predictors with SHAP value magnitudes more than zero were included in AGM/T2D prediction models. Sensitivity analyses were performed to explore the prediction effects of diagnosing GDM using modified two-point IADPSG 2018 criteria [9] rather than WHO 1999 criteria (GUSTO study did not include a 1-hour glucose measurement), and of continuous fasting or 2-hour glucose measures and pre-pregnancy BMI. We also assessed the associations between total gestational weight gain and postpartum AGM/T2D outcomes. All association analyses were performed using Stata/MP 16.1 software (StataCorp LP, College Station, Texas, USA).

## Results

### The features significantly associated with T2D are aligned with the top features from SHAP feature selection model

The relationship between all feature variables and postpartum AGM/T2D outcomes are presented in Pearson correlation heatmap (Figures 1, 2). Diagnosis of GDM, mid-upper arm circumference and BMI are the best features for postpartum AGM/T2D machine learning model building.

Table 1 presents the univariate associations between mid-pregnancy features and postpartum AGM/T2D outcomes. Previous history of GDM, mean arterial blood pressure, mid-upper arm circumference, BMI and diagnosis of GDM were associated with later T2D. The top 4 features impacting the SHAP model outputs were mid-upper arm circumference, diagnosis of GDM, BMI and mean arterial blood pressure (Figure 3). The negative SHAP value for height implies that maternal height did not contribute to the prediction of AGM.

### Maternal adiposity during pregnancy and metabolic derangements underlying GDM signal future T2D risk

While showing the detailed training parameters and results for all machine learning models in Supplementary Tables 1a-1f (Appendix 2), we focus on describing the results of CatBoost machine learning models as this algorithm had the best overall performance. The results for each dataset of the 5-fold stratified cross validation and average of the cross validation are also provided in Supplementary Tables 1a-1f. Mid-upper arm circumference at mid-gestation [AUC: 0.78 (95% CI 0.71, 0.86)] and BMI at mid-gestation [AUC: 0.74 (0.53, 0.96)] had stronger predictive performances than GDM diagnosis [AUC: 0.73 (0.51, 0.95)] (Supplementary Table 1b). The addition of GDM diagnosis improved the performance of baseline models: ‘MUAC\_GDM’ model [AUC: 0.88 (0.79, 0.96)] and ‘BMI\_GDM’ model [AUC: 0.86 (0.72, 0.99)] (Supplementary Table 1d). Pre-pregnancy BMI alone was inadequate in predicting postpartum T2D risk [AUC: 0.62 (0.39, 0.86)] (Supplementary Table 1f).

Although there is a high correlation between mid-upper arm circumference and BMI (r = 0.91), BMI is more reliably and commonly assessed in clinical settings and hence a BMI-based pregnancy model is our proposed solution (Figure 4). Table 2 summarizes the detailed training parameters of logistic regression, support vector machine, artificial neural network & CatBoost gradient boosting algorithms and results of proposed postpartum T2D predictive model (comprising of mid-pregnancy BMI after gestational weight gain; diagnosis of GDM features). Total gestational weight gain was inversely associated with postpartum AGM and T2D outcomes, independent of pre-pregnancy BMI and diagnosis of GDM [*P*=.019, OR: 0.88 (0.79-0.98)] (Table 3).

Figures 5-7 present the validation curves obtained during the training of ‘BMI\_GDM’ CatBoost model. The hyperparameter candidates for CatBoost model were as follows.

Learning rate: [‘0’ - 0.00001, **‘1’- 0.0001**, ‘2’ - 0.001, ‘3’ - 0.01, ‘4’ - 0.03, ‘5’ - 0.05, ‘6’ - 0.1, ‘7’ - 0.2, ‘8’ - 0.3]

L2 leaf regularization: [‘0’ - 1.0, ‘1’ - 2.0, ‘2’ - 3.0, ‘3’ - 4.0, ‘**4’ - 5.0**, ‘5’ - 6.0]

Random strength: [‘0’ - 1.0, ‘1’ - 2.0, ‘2’ - 3.0, ‘3’ - 4.0, ‘**4’ - 5.0**, ‘5’ - 6.0]

The CatBoost model was specified with 1000 iterations, maximum depth of 6 trees and symmetric tree growing policy. The hyperparameters tuned using grid search were learning rate of 0.0001, L2 leaf regularization of 5.0 and random strength of 5.0. The ‘BMI\_GDM’ CatBoost classifier is performing well under this optimal configuration.

### 2-hour postprandial glucose is a stronger predictor of postpartum T2D risk than fasting glucose

Modeling antenatal glucose measures as continuous features, 2-hour postprandial glucose [AUC: 0.86 (0.76, 0.96)] showed a stronger postpartum T2D risk prediction effect compared to fasting glucose [AUC: 0.76 (0.61, 0.91)] (Supplementary Table 1f). In the sensitivity analysis, predictive performance of ‘BMI\_GDM’ model was also robust when using the modified two-point IADPSG 2018 criteria [AUC: 0.84 (0.72, 0.97)] (Supplementary Table 1f).

## Discussion

### Principal Results

We have built an effective postpartum T2D predictive model by combining game theory-based feature selection with machine learning. SHAP values recovered predictive modeling features for optimal performance, aligning model interpretability with human intuition. Our ‘BMI\_GDM’ model achieved an excellent AUC of 0.86 with two mid-gestation features (BMI at mid-gestation; diagnosis of GDM by WHO 1999 criteria) for early prediction of postpartum T2D risk in a Singapore population. The model was also robust when using a modified two-point IADPSG 2018 criteria for GDM diagnosis (AUC: 0.84). The ‘BMI\_GDM’ machine learning model can be leveraged as a risk stratification tool during routine GDM screening to identify Asian women at high risk of developing T2D, enabling early intervention. The ‘BMI\_2hour’ model (AUC: 0.86) can be an alternative design during clinical implementation if GDM diagnosis feature is unavailable for patient. The trained classifier can be deployed using a web application, which can allow clinicians to identify women at T2D risk and develop a postpartum management plan.

The two-feature, mid-pregnancy BMI model (AUC: 0.86) performed better for postpartum T2D prediction than a preconception BMI model (AUC: 0.62), implying that mid-gestational weight gain effects combined with the metabolic derangements underlying GDM and fetoplacental unit signals future T2D risk. As pregnancy has a diabetogenic effect on metabolism [13], further studies will be required to examine the metabolic adaptations in pregnancy and postpartum maternal metabolic health outcomes.

In our ‘BMI\_GDM’ model sensitivity analysis, we observed that the 2-hour antenatal OGTT glucose peak was associated with a stronger prediction of postpartum T2D (AUC: 0.86) than fasting glucose (AUC: 0.76) in Singaporean women. Future studies with greater statistical power will be needed to confirm whether the postpartum T2D risk is heterogenous across the different thresholds of glucose tolerance for GDM diagnostic criteria.

### Limitations

This study has some limitations due to scarcity of longitudinal data. Postpartum OGTT at 4-12 weeks and further testing in those with normal postpartum OGTT every 1-3 years were not administered in GUSTO study, possibly underestimating the development of post-delivery dysglycaemia to a certain extent and inducing bias. However, the GUSTO mothers self-reported T2D status at two years after delivery and there were no self-reported T2D cases. Our prediction models were trained on a limited cohort of 561 pregnancies and require further validation using larger cohorts such as Electronic Health Record (EHR) databases. A sub-cohort analyses by individual ethnic groups can be trained with larger datasets.

### Comparison with Prior Work

Our early implementation of T2D risk prediction algorithm during prenatal care enables early engagement of patients and remote monitoring, compared to existing molecular biomarker-based T2D risk prediction algorithms [7, 8] developed for postpartum care. The two mid-gestation clinical features (mid-pregnancy BMI after gestational weight gain and diagnosis of GDM) discovered from our machine learning workflow is of low cost and easily accessible during routine antenatal GDM screening. The digital biomarkers identified from our work will guide antenatal research in preventing the progression of GDM to T2D.

### Conclusions

The key strength of our study lies in applying machine learning-based predictive analytics during prenatal care for early prediction of postpartum T2D. This machine learning model can be leveraged as a risk stratification tool for preventive intervention.

**Tables**

**Table 1: Associations between mid-pregnancy features and postpartum AGM/T2D outcomes (4 to 8 years after delivery)**

|  |  |  |
| --- | --- | --- |
| **Feature** | **AGM (n = 139)** | **T2D (n = 32)** |
| **OR (95% CI)*****P*-value** | **OR (95% CI)*****P*-value** |
| Maternal age (years) | OR: 1.05 (1.01-1.09)*P*-value = .021\* | OR: 1.06 (0.99-1.14)*P*-value = .104 |
| Chinese vs Malay/Indian ethnicity | OR: 0.81 (0.55-1.19)*P*-value = .275 | OR: 0.71 (0.34-1.44)*P*-value = .339 |
| Malay vs Chinese/Indian ethnicity | OR: 1.20 (0.79-1.83)*P*-value = .400 | OR: 1.64 (0.78-3.43)*P*-value = .193 |
| Indian vs Chinese/Malay ethnicity | OR: 1.12 (0.68-1.84)*P*-value = .658 | OR: 0.87 (0.33-2.31)*P*-value = .777 |
| Family history of diabetes mellitus | OR: 1.72 (1.15-2.56)*P*-value = .008\* | OR: 1.55 (0.75-3.21)*P*-value = .239 |
| Family history of high blood pressure | OR: 0.88 (0.60-1.32)*P*-value = .545 | OR: 0.70 (0.33-1.51)*P*-value = .365 |
| Family history of cardiovascular disease | OR: 1.04 (0.57-1.90)*P*-value = .904 | OR: 0.51 (0.12-2.19)*P*-value = .367 |
| Previous history of gestational diabetes mellitus | OR: 5.96 (2.16-16.43)*P*-value = .001\* | OR: 7.98 (2.62-24.27)*P*-value < .001\* |
| Previous history of gestational hypertension | OR: 1.86 (0.66-5.21)*P*-value = .239 | OR: 2.45 (0.53-11.29)*P*-value = .249 |
| Parity | OR: 1.02 (0.69-1.50)*P*-value = .931 | OR: 1.38 (0.66-2.89)*P*-value = .387 |
| Mean arterial blood pressure (mm Hg) | OR: 1.05 (1.03-1.07)*P*-value < .001\* | OR: 1.07 (1.03-1.11)*P*-value < .001\* |
| Mid-upper arm circumference (cm) | OR: 1.18 (1.12-1.25)*P*-value < .001\* | OR: 1.23 (1.13-1.33)*P*-value < .001\* |
| Maternal height (cm) | OR: 0.96 (0.92-0.99)*P*-value = .013\* | OR: 0.96 (0.90-1.02)*P*-value = .196 |
| BMI (kg/m2) | OR: 1.14 (1.09-1.18)*P*-value < .001\* | OR: 1.16 (1.09-1.24)*P*-value < .001\* |
| Smoking during pregnancy | OR: 1.14 (0.30-4.36)*P*-value = .847 | -a |
| Environmental tobacco smoke exposure at home | OR: 1.07 (0.72-1.60)*P*-value = .729 | OR: 0.98 (0.46-2.08)*P*-value = .962 |
| Environmental tobacco smoke exposure at workplace | OR: 0.76 (0.38-1.51)*P*-value = .431 | OR: 1.37 (0.46-4.06)*P*-value = .572 |
| Alcohol consumption during pregnancy | OR: 1.14 (0.30-4.36)*P*-value = .847 | OR: 1.67 (0.21-13.50)*P*-value = .628 |
| Diagnosis of GDM (WHO 1999) | OR: 5.49 (3.51-8.58)*P*-value < .001\* | OR: 9.57 (4.45-20.55)*P*-value < .001\* |

\*Statistically significant features.

-a Fixed-effects regression estimates were not obtained as variable did not contribute to the likelihood estimation.

**Table 2: Proposed postpartum T2D predictive model (mid-pregnancy BMI after gestational weight gain and diagnosis of GDM features)**

|  |  |  |
| --- | --- | --- |
| **Features: BMI at mid-gestation + Diagnosis of gestational diabetes mellitus (WHO 1999) (BMI\_GDM)** | **Hyperparameters tuned using grid search** | **Average AUC (95% CI)** |
| **Model Specifications**  |
| Logistic Regression (L2 regularization penalty, stochastic average gradient descent solver) | Inverse of regularization strength = 1.0 | 0.85 (0.72, 0.98) |
| Support Vector Machine (linear kernel, L2 regularization penalty) | L2 regularization penalty = 1.0Loss function = ‘squared hinge’ | 0.85 (0.72, 0.98) |
| Neural Network (3 hidden layers with 10 neurons each, ReLU activation function, Adam solver, 200 iterations)  | L2 regularization penalty = 0.01Initial learning rate = 0.1 | 0.85 (0.73, 0.97) |
| **CatBoost** (1000 iterations, maximum depth of 6 trees, symmetric tree growing policy) | L2 leaf regularization = 5.0Learning rate = 0.0001Random Strength = 5.0 | **0.86 (0.72, 0.99)** |

**Table 3: Association between total gestational weight gain and postpartum AGM/T2D outcomes (4 to 8 years after delivery)**

|  |  |  |
| --- | --- | --- |
| **Unadjusted analysis** | **AGM (n = 128)** | **T2D (n = 31)** |
| **OR (95% CI)*****P*-value** | **OR (95% CI)*****P*-value** |
| Total Gestational Weight Gain (kg) | OR: 0.87 (0.82-0.91)*P*-value < .001\* | OR: 0.79 (0.72-0.87)*P*-value < .001\* |

|  |  |  |
| --- | --- | --- |
| **Adjusted analysisb** | **AGM (n = 128)** | **T2D (n = 31)** |
| **OR (95% CI)*****P*-value** | **OR (95% CI)*****P*-value** |
| Total Gestational Weight Gain (kg) | OR: 0.93 (0.87-0.98)*P*-value = .014\* | OR: 0.88 (0.79-0.98)*P*-value = .019\* |

\*Statistically significant.

b Adjusted for maternal ethnicity, age, parity, family history of diabetes mellitus, pre-pregnancy BMI and diagnosis of GDM

**End Matter**

**Acknowledgements**

We thank the GUSTO study team for their help in acquiring the research data and their crucial work with the participants.

**Data Availability**

The data that support the findings of this research are available from the corresponding authors upon reasonable request.

The code generated to reproduce this research is available at GitHub page:

<https://github.com/mukkeshkumar/GUSTO_Type-2-Diabetes-Mellitus>

The postpartum T2D predictive model (CatBoost algorithm) have been deployed into a web application and can be accessed through the following URL:

<https://www.mornin-feng.com/all-projects-and-demos#gdm2>

**Funding**

The GUSTO birth cohort study is supported by the Translational Clinical Research (TCR) Flagship Program on Developmental Pathways to Metabolic Disease and Open Fund Large Collaborative Grant (OFLCG) Programmes, funded by the National Research Foundation (NRF) and administered by the National Medical Research Council (NMRC), Singapore (award numbers NMRC/TCR/004-NUS/2008, NMRC/TCR/012-NUHS/2014, OFLCG/MOH-000504). This research is supported by NMRC’s Open Fund - Large Collaborative Grant, titled ‘Metabolic Health in Asian Women and their Children’ (award number OFLCG19may-0033). KMG is supported by the UK Medical Research Council (MC\_UU\_12011/4), the National Institute for Health Research (NIHR Senior Investigator (NF-SI-0515-10042) and NIHR Southampton Biomedical Research Centre (IS-BRC-1215-20004)) and the British Heart Foundation (RG/15/17/3174). Additional funds for data analysis were supported by the Strategic Positioning Fund and IAFpp funds (H17/01/a0/005) available to NK through Agency for Science, Technology and Research (A\*STAR), Singapore (award number SPF 002/2013).

**Conflicts of Interest**

NK, KMG, SYC and YSC are part of an academic consortium that has received research funding from Abbott Nutrition, Nestec, BenevolentAI Bio Ltd. and Danone. MF was partially supported by the National Research Foundation Singapore under its AI Singapore Programme (Award Number: [AISG-GC-2019-001-2A]). Other authors declare no conflicts of interest.

**Contribution Statement**

MK contributed to research study design, data curation, machine learning modeling, statistical analyses, interpretation of results and writing of manuscript. LTA and CH contributed to clinical data curation. SES and SYC contributed to collection of phenotypic data in GUSTO cohort and critical reading of the manuscript. KHT, JKYC, KMG and YSC contributed to GUSTO cohort study design, data collection and critical reading of the manuscript. JGE contributed to interpretation of results, writing of manuscript and GUSTO cohort data collection. MF contributed to supervision of the study, interpretation of results and writing of manuscript. NK contributed to supervision of the study, interpretation of results, writing of manuscript and GUSTO cohort study data collection. MF and NK accepts full responsibility for the work, had access to the data, and controlled the decision to publish.

**Multimedia Appendix**

Multimedia Appendix 1: Implementation details of machine learning algorithms

Multimedia Appendix 2: Supplementary Tables 1a-1f, detailed training parameters and results for all machine learning models

**References**

1. International Diabetes Federation. IDF Diabetes Atlas, 9th edn 2019.

2. Vounzoulaki E, Khunti K, Abner SC, Tan BK, Davies MJ, Gillies CL. Progression to type 2 diabetes in women with a known history of gestational diabetes: systematic review and meta-analysis. BMJ. 2020 May 13;369:m1361. PMID: 32404325. doi: 10.1136/bmj.m1361.

3. Chen L-W, Soh SE, Tint M-T, Loy SL, Yap F, Tan KH, et al. Combined analysis of gestational diabetes and maternal weight status from pre-pregnancy through post-delivery in future development of type 2 diabetes. Scientific Reports. 2021 2021/03/03;11(1):5021. doi: 10.1038/s41598-021-82789-x.

4. Classification and Diagnosis of Diabetes: Standards of Medical Care in Diabetes-2020. Diabetes Care. 2020;43(Supplement 1):S14-S31. doi: 10.2337/dc20-S002.

5. Nouhjah S, Shahbazian H, Amoori N, Jahanfar S, Shahbazian N, Jahanshahi A, et al. Postpartum screening practices, progression to abnormal glucose tolerance and its related risk factors in Asian women with a known history of gestational diabetes: A systematic review and meta-analysis. Diabetes Metab Syndr. 2017 Dec;11 Suppl 2:S703-s12. PMID: 28571777. doi: 10.1016/j.dsx.2017.05.002.

6. Sanderson H, Loveman E, Colquitt J, Royle P, Waugh N, Tan BK. Improving Uptake of Postnatal Checking of Blood Glucose in Women Who Had Gestational Diabetes Mellitus in Universal Healthcare Settings: A Systematic Review. J Clin Med. 2018 Dec 20;8(1). PMID: 30577481. doi: 10.3390/jcm8010004.

7. Allalou A, Nalla A, Prentice KJ, Liu Y, Zhang M, Dai FF, et al. A Predictive Metabolic Signature for the Transition From Gestational Diabetes Mellitus to Type 2 Diabetes. Diabetes. 2016;65(9):2529-39. doi: 10.2337/db15-1720.

8. Joglekar MV, Wong WKM, Ema FK, Georgiou HM, Shub A, Hardikar AA, et al. Postpartum circulating microRNA enhances prediction of future type 2 diabetes in women with previous gestational diabetes. Diabetologia. 2021 2021/07/01;64(7):1516-26. doi: 10.1007/s00125-021-05429-z.

9. World Health Organization. Diagnostic Criteria and Classification of Hyperglycaemia First Detected in Pregnancy. 2013.

10. World Health Organization; International Diabetes Federation (IDF). Definition and diagnosis of diabetes mellitus and intermediate hyperglycaemia. 2006.

11. Lundberg SM, Erion G, Chen H, DeGrave A, Prutkin JM, Nair B, et al. From local explanations to global understanding with explainable AI for trees. Nature Machine Intelligence. 2020 2020/01/01;2(1):56-67. doi: 10.1038/s42256-019-0138-9.

12. Prokhorenkova L, Gusev G, Vorobev A, Veronika Dorogush A, Gulin A. CatBoost: unbiased boosting with categorical features. arXiv e-prints [Internet]. 2017 June 01, 2017. Available from: <https://ui.adsabs.harvard.edu/abs/2017arXiv170609516P>.

13. Zeng Z, Liu F, Li S. Metabolic Adaptations in Pregnancy: A Review. Annals of Nutrition and Metabolism. 2017;70(1):59-65. doi: 10.1159/000459633.

**Abbreviations**

AGM: Abnormal Glucose Metabolism

AUC: Area under the Receiver Operating Characteristic Curve

BMI: Body Mass Index

GDM: Gestational Diabetes Mellitus

GUSTO: Growing Up in Singapore Towards healthy Outcomes

HbA1c: Hemoglobin A1c

IADPSG: International Association of Diabetes and Pregnancy Study Groups

OGTT: Oral Glucose Tolerance Test

SHAP: SHapley Additive exPlanations

T2D: Type 2 Diabetes

WHO: World Health Organization