

Development, validation and clinical utility of a risk prediction model for adverse pregnancy outcomes in women with gestational diabetes

Cooray, Shamil D.; Boyle, Jacqueline A.; Soldatos, Georgia; Allotey, John; Wang, Holly; Fernandez-Felix, Borja M.; Zamora, Javier; Thangaratinam, Shakila; Teede, Helena J.

DOI:

[10.1016/j.eclim.2022.101637](https://doi.org/10.1016/j.eclim.2022.101637)

License:

Creative Commons: Attribution-NonCommercial-NoDerivs (CC BY-NC-ND)

Document Version

Publisher's PDF, also known as Version of record

Citation for published version (Harvard):

Cooray, SD, Boyle, JA, Soldatos, G, Allotey, J, Wang, H, Fernandez-Felix, BM, Zamora, J, Thangaratinam, S & Teede, HJ 2022, 'Development, validation and clinical utility of a risk prediction model for adverse pregnancy outcomes in women with gestational diabetes: The PeRsonal GDM model', *EClinicalMedicine*, vol. 52, 101637. <https://doi.org/10.1016/j.eclim.2022.101637>

[Link to publication on Research at Birmingham portal](#)

General rights

Unless a licence is specified above, all rights (including copyright and moral rights) in this document are retained by the authors and/or the copyright holders. The express permission of the copyright holder must be obtained for any use of this material other than for purposes permitted by law.

- Users may freely distribute the URL that is used to identify this publication.
- Users may download and/or print one copy of the publication from the University of Birmingham research portal for the purpose of private study or non-commercial research.
- User may use extracts from the document in line with the concept of 'fair dealing' under the Copyright, Designs and Patents Act 1988 (?)
- Users may not further distribute the material nor use it for the purposes of commercial gain.

Where a licence is displayed above, please note the terms and conditions of the licence govern your use of this document.

When citing, please reference the published version.

Take down policy

While the University of Birmingham exercises care and attention in making items available there are rare occasions when an item has been uploaded in error or has been deemed to be commercially or otherwise sensitive.

If you believe that this is the case for this document, please contact UBIRA@lists.bham.ac.uk providing details and we will remove access to the work immediately and investigate.

Development, validation and clinical utility of a risk prediction model for adverse pregnancy outcomes in women with gestational diabetes: The PeRsonal GDM model

Shamil D. Cooray,^{a,b} Jacqueline A. Boyle,^{a,c} Georgia Soldatos,^{a,b} John Allotey,^d Holly Wang,^b Borja M. Fernandez-Felix,^e Javier Zamora,^{d,e} Shakila Thangaratinam,^{e,f,1} and Helena J. Teede^{a,b,*1}

^aMonash Centre for Health Research and Implementation, School of Public Health and Preventive Medicine, Monash University, Clayton VIC 3168, Australia

^bDiabetes and Endocrinology Units, Monash Health, Clayton VIC 3168, Australia

^cMonash Women's Program, Monash Health, Clayton VIC 3168, Australia

^dWHO Collaborating Centre for Global Women's Health, Institute of Metabolism and Systems Research, University of Birmingham, Edgbaston, Birmingham B15 2TT, United Kingdom

^eCIBER Epidemiology and Public Health, 28029 Madrid, Spain

^fBirmingham Women's and Children's, NHS Foundation Trust, Birmingham, UK

Summary

Background The ability to calculate the absolute risk of adverse pregnancy outcomes for an individual woman with gestational diabetes mellitus (GDM) would allow preventative and therapeutic interventions to be delivered to women at high-risk, sparing women at low-risk from unnecessary care. We aimed to develop, validate and evaluate the clinical utility of a prediction model for adverse pregnancy outcomes in women with GDM.

Methods A prediction model development and validation study was conducted on data from an observational cohort. Participants included all women with GDM from three metropolitan tertiary teaching hospitals in Melbourne, Australia. The development cohort comprised those who delivered between 1 July 2017 to 30 June 2018 and the validation cohort those who delivered between 1 July 2018 to 31 December 2018. The main outcome was a composite of critically important maternal and perinatal complications (hypertensive disorders of pregnancy, large-for-gestational age neonate, neonatal hypoglycaemia requiring intravenous therapy, shoulder dystocia, perinatal death, neonatal bone fracture and nerve palsy). Model performance was measured in terms of discrimination and calibration and clinical utility evaluated using decision curve analysis.

Findings The final PeRsonal (Prediction for Risk Stratified care for women with GDM) model included body mass index, maternal age, fasting and 1-hour glucose values (75-g oral glucose tolerance test), gestational age at GDM diagnosis, Southern and Central Asian ethnicity, East Asian ethnicity, nulliparity, past delivery of a large-for-gestational age neonate, past pre-eclampsia, GWG until GDM diagnosis, and family history of diabetes. The composite adverse pregnancy outcome occurred in 27% (476/1747) of women in the development (1747 women) and in 26% (244/955) in the validation (955 women) cohorts. The model showed excellent calibration with slope of 0.99 (95% CI 0.75 to 1.23) and acceptable discrimination (c-statistic 0.68; 95% CI 0.64 to 0.72) when temporally validated. Decision curve analysis demonstrated that the model was useful across a range of predicted probability thresholds between 0.15 and 0.85 for adverse pregnancy outcomes compared to the alternatives of managing all women with GDM as if they will or will not have an adverse pregnancy outcome.

Interpretation The PeRsonal GDM model comprising of routinely available clinical data shows compelling performance, is transportable across time, and has clinical utility across a range of predicted probabilities. Further external validation of the model to a more disparate population is now needed to assess the generalisability to different centres, community based care and low resource settings, other healthcare systems and to different GDM diagnostic criteria.

Funding This work is supported by the Mothers and Gestational Diabetes in Australia 2 NHMRC funded project #1170847.

eClinicalMedicine

2022;52: 101637

Published online 5 September 2022

<https://doi.org/10.1016/j.eclinm.2022.101637>

eclinm.2022.101637

*Corresponding author at: Monash Centre for Health Research and Implementation, School of Public Health and Preventive Medicine, Monash University, Locked Bag 29 Clayton, VIC 3168, Australia.

E-mail address: Helena.Teede@monash.edu (H.J. Teede).

¹ Joint senior authors.

Copyright © 2022 The Author(s). Published by Elsevier Ltd. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>)

Keywords: Gestational diabetes mellitus (GDM); Prediction model; Risk-stratification; Prognosis; Pregnancy complications; Adverse pregnancy outcomes; Large-for-gestational age (LGA); Pre-eclampsia; Neonatal hypoglycaemia

Research in context

Evidence before this study

We previously published a systematic review of studies of prediction models for adverse pregnancy outcomes in women with gestational diabetes mellitus (GDM) to 16 August 2019. Five prediction modelling studies were identified, from which ten prediction models intended to predict adverse pregnancy outcomes related to GDM were developed. All models lacked external validation, performance was inadequately reported with no useful measures of calibration nor formal evaluation of clinical utility and methodologic limitations in statistical analysis limit generalisability.

Added value of this study

The PerSonal GDM (Prediction for Risk-Stratified care for women with GDM) model accurately predicts the individualised risk of adverse pregnancy outcome in pregnant women with GDM with excellent calibration and acceptable discrimination. The model was developed in an ethnically diverse population served by a universal and freely accessible health system and its performance maintained when evaluated in a more recently treated population. The model has clinical utility across a broad range of probability thresholds, to assist in shared decision-making for personalised and risk-stratified care.

Implications of all the available evidence

The PerSonal GDM model can facilitate shared decision-making at the individual level and risk-stratified care at a health service level, ultimately, supporting more personalised care for women with GDM. To promote translation into clinical care, this model has been translated into an online clinical risk calculator allowing clinicians to calculate individualised risks of adverse pregnancy outcomes to facilitate shared decision-making on antenatal care and risk-stratified approaches to treatment.

GDM are diagnosed to have the condition using various arbitrary thresholds of glucose challenge tests, thereby dichotomising this continuous risk based on glucose values alone. Furthermore, lowering the thresholds for diagnoses with the newer diagnostic criteria has resulted in a significant increase in the proportion of women diagnosed with GDM, who are at varied risk of adverse pregnancy outcomes.

In addition to blood glucose levels, various factors have been associated with maternal and perinatal complications in women with GDM, such as maternal body mass index,^{3,4} ethnicity,⁵ and gestational weight gain (GWG).⁶ But current treatment strategies for planning obstetric management of GDM generally adopt a one-size-fits-all glucocentric approach, where women with GDM are generally treated as high-risk pregnancies with hospital-based care.^{7,8} This presents challenges given increased GDM prevalence and strain on health system resources,⁹ especially during and post COVID-19.¹⁰ It also retains a one-size fits all focus on all women with GDM with attendant individual healthcare and psychological burden¹¹ and economic costs.¹²

We need a robust risk-based approach to plan the management of women with GDM, enabling shared decision-making and more personalised care.¹³ The accurate identification of women with GDM at highest risk of adverse pregnancy outcomes would facilitate their targeted management with high intensity care, while those identified to be at low risk of complications can be managed within routine care pathways, or potentially in the community. Previous models to predict the risks are hampered by statistical methodological limitations which limit generalisability, such as inadequate power, dichotomisation of continuous predictor variables and predictor selection dependent on associations with the outcome in the development dataset.¹⁴ We aim to develop an individualised PerSonal (Prediction for Risk Stratified care for women with GDM) risk prediction model incorporating predictors for adverse pregnancy outcomes in women with GDM, and temporally validate its performance and determine its clinical utility.

Introduction

Gestational diabetes mellitus (GDM), defined as glucose intolerance diagnosed for the first time in pregnancy,¹ is on the increase worldwide with up to 15% of women diagnosed in some regions.² Currently, women with

Methods

We undertook a systematic review and critical appraisal of existing prediction models, finding that they are subject to a high risk of bias or are too limited with regard

to clinical applicability.¹⁴ Hence the decision was taken to develop a new model. We developed and validated the PeRsonal GDM model using a prospective study protocol,¹⁵ registered with the Australian New Zealand Clinical Trials Registry (ACTRN12620000915954).¹⁶ We reported our findings in line with the Transparent Reporting of a multivariable prediction model for Individual Prognosis or Diagnosis (TRIPOD) statement.¹⁷ This study was approved by the Human Research Ethics Committee of Monash Health (RES-19-0000713L). Consent was not required for the secondary use of non-identifiable data, as it was collected as part of routine clinical care for the primary purpose of quality improvement.

Model development and validation cohorts

We developed and validated the model using routinely collected prospective health data of pregnant women with GDM who gave birth in three metropolitan tertiary teaching hospitals serving South-East Melbourne, Australia between 1 July 2017 to 31 December 2018. The health network provides universal access to maternity care to over 9,000 ethnically diverse women annually. We linked obstetric and neonatal data to pathology data and additional clinical data from the parent health services. We developed the prediction model using pregnancies from the first 12 months, and used the last six months of data for temporal validation. Using the validation strategy the data was split into two parts: the first containing pregnant women treated earlier to develop the model and the second part containing more recently treated pregnant women to assess performance.¹⁸ Women with multiple pregnancy or repeat pregnancies within the study period were included. We included data from all women with GDM who were diagnosed using the International Association of Diabetes and Pregnancy Study Groups (IADPSG) criteria based on universal screening.¹⁹ Universal screening was performed at 24–28 weeks with a one-step 75g oral glucose tolerance test (OGTT).²⁰ For women with risk factors, screening was undertaken in early pregnancy using the OGTT. Treatment for GDM comprised of an initial 2-hour group education session with diabetes nurse educators and dietitians with lifestyle advice on diet, physical activity and weight management and on blood glucose monitoring. Women had regular antenatal appointments to see the diabetologists (every one to three weeks) and saw obstetricians at 28, 31, 34, 36, 38, 39 and 40 weeks' gestation. If on more than 20 units of insulin per day or clinically indicated, women had scans to assess fetal growth at 28, 32 and 34 weeks' gestation plus weekly biophysical profiles and umbilical artery doppler from 34 weeks. Insulin was commenced if the fasting glucose targets exceeded 5.5 mmol/L and 2-hour post-prandial value exceeded 7.0 mmol/L, despite dietary modification. Metformin was only used if there was evidence of significant insulin resistance or when

insulin use was not preferred due to concerns with potential psychological harm.

Candidate predictors

We selected the candidate predictors by systematically reviewing existing models for included predictors and other relevant prognostic studies.¹⁵ We evaluated the following predictors for inclusion in the model: maternal age, body mass index (BMI), parity, gestational age at GDM diagnosis, ethnicity, prior GDM, prior pre-eclampsia, prior large-for-gestational age (LGA) baby, prior shoulder dystocia, family history of diabetes, GWG, and OGTT glucose values (Supplementary Table S2). The predictors were assessed blinded to the outcome, due to the contemporaneous nature of this data in routine clinical practice. Self-reported ethnicity was classified into six categories according to the Australian Standard Classification of Cultural and Ethnic Groups.²¹

Outcome

The main outcome was a composite of adverse pregnancy outcomes that included: hypertensive disorders of pregnancy, birth of a large-for-gestational-age neonate, neonatal hypoglycaemia requiring intravenous treatment, shoulder dystocia, fetal death, neonatal death, bone fracture and nerve palsy. It was developed following extensive formative research (previously reported), to design a robust and clinically acceptable prediction model involving multidisciplinary engagement.¹⁵ This composite consisted of prioritised outcomes identified in a systematic review of existing models, the core outcome set for GDM treatment research and other relevant literature as previously described.¹⁵ The definitions of the components of the composite outcome are provided in Supplementary Table S1.

Sample size

The sample recruited in the development cohort was large enough to get an event per variable (EPV) ratio exceeding the common rule of 10, or even 20 EPV as suggested by other recommendations.²² We observed 476 composite events in the development cohort, and we evaluated 18 candidate predictors during development. For validating the model, we observed 244 events which again exceeds the recommendation of at least 100 events for validation cohorts.²³

Missing data

Missing data for ethnicity classification was recovered manually using available country of birth and language preference data where available. All remaining missing data was handled using multiple imputation by chained

equations (see Supplementary Appendix S1 for additional details).

Statistical analysis

Model development. We fit a multivariate logistic regression model to predict the adverse pregnancy composite outcome. The continuous predictors were kept as continuous. We used fractional polynomials to decide the functional form for continuous predictors. As several continuous variables were included in the model, we used the multivariable fractional polynomial algorithm combined with multiple imputation using the procedure described by Morris and colleagues.²⁴ We used a Least Absolute Shrinkage and Selection Operator (LASSO) method to build the predictive model.²⁵ For each of the 50 imputed datasets we repeated the LASSO procedure. We included as predictors in the final model those covariates that were selected in at least 90% of the LASSO models. Model coefficients were averaged across the 50 repetitions using Rubin's rule.²⁶ We used the 2.5 and 97.5 percentiles from 200 bootstrap samples as the limits of the 95% confidence intervals of model coefficients.

We reported the model performance in terms of discrimination and calibration. Discrimination refers to how well predictions separate between those participants who do and do not develop the composite adverse event. We reported the overall discriminatory ability of the developed model as the *c*-statistic (summarised using area under the receiver operating characteristic curve, range 0–1, values > 0.5 show discrimination) along with the 95% confidence interval (CI). Calibration examines the agreement between predicted and observed risks of the composite adverse event. Calibration was assessed graphically using a calibration plot. In order to avoid instability due to the number of cases we stratified the calibration plot for the validation cohort in quintiles. We also calculated calibration-in-the-large (predicted risks are under-estimated if > 0 or over-estimated if < 0) and the calibration slope (predicted risks that are too extreme if < 1 or not extreme enough if > 1). All measures were averaged over imputed samples using Rubin's rule.²⁶

Model validation. We temporally validated the model to assess transportability in the cohort who gave birth at a different time point and reported the model's predictive performance using the same measures of discrimination and calibration as described above. The development and validation data had similar eligibility criteria, predictors and outcome.

We conducted sensitivity analysis to evaluate the confounding effect of insulin treatment on predictor-

outcome associations. Full details of this method are reported in the Supplementary Appendix S1.

Decision curve analysis. To evaluate the clinical utility of the model we determined its net benefit. Net benefit is a measure which can be used to integrate the benefits and harms of using a model for clinical decision support. In this clinical context, the benefit of using the PerSonal GDM model is identifying women likely to experience an adverse pregnancy outcome and the avoidance of unnecessary diversion to a high-risk care setting leading to personal burden and decreased satisfaction with birth experience, cost and inconvenience and impact on health system resources. We determined the net benefit of the model over a range of probability thresholds,²⁷ instead of the alternatives approaches, either, managing all women with GDM as if they will have an adverse pregnancy outcome (high-risk care setting for all), or managing all women with GDM as if they will *not* have an adverse pregnancy outcome (routine antenatal care for all). Acknowledging that there isn't a single optimal probability threshold for risk-stratification for all settings, we represented the net benefit as a function of the probability threshold in a decision curve plot, and provided the results for multiple plausible risk thresholds.²⁸

The choice of probability threshold at which to stratify care is dependent on the local epidemiology, health service structure and resources and preferences of clinicians and patients in the local setting. A higher threshold would result in less women being stratified to a high-risk care setting which may be appropriate where there is a need to target interventions to those most likely to benefit due to resource limitations. This may be suitable in high prevalence or low resource settings. Conversely, at a lower threshold, more women would be referred to the high-risk care setting at the expense of more women receiving unnecessary care. We anticipate that an intermediate threshold may best address the clinical needs of many settings.

All statistical analyses were performed using Stata version 16.1 (College Station, TX: StataCorp LLC).

Role of the funding source

The funder had no role in the study design, data collection, data analysis, data interpretation, or writing of the report. All authors had full access to the data in the study and accept responsibility for the decision to submit for publication.

Results

Overall, we included 1747 pregnancies with GDM in the development cohort, and 955 pregnancies in the validation cohort for analysis (Supplementary Figure S1).

Characteristics of participants

Women in the development and validation cohorts had similar social and demographic factors, obstetric and family history and physical characteristics (Table 1). In both cohorts, the mean maternal age was 32 years. A similar proportion of women were of Caucasian origin (development, 19.6%; validation 20.7%); a greater proportion of women were classified as Southern and

Central Asian (46.2% vs 35.3%) and East Asian (21.6% vs 13.8%) in the validation than in the development cohort. There was a slightly higher proportion of women in the validation cohort with a history of prior GDM (25% vs 20%), and who required insulin therapy (40.2% vs 36.6%). The missing data per variable are provided in Table 1; there were no missing data for outcomes.

	Development cohort (n = 1747)		Validation cohort (n = 955)	
	Mean (SD) or n (%)	Number with missing data, n (%)	Mean (SD) or n (%)	Number with missing data, n (%)
Social / demographic factors				
Maternal age, years	32.2 (5.0)	0	32.2 (5.0)	0
Number of fetuses		0		0
Singleton	1683 (96.3)		924 (96.8)	
Twins	64 (3.7)		28 (2.9)	
Triplets	0 (0)		3 (0.3)	
Ethnicity		275 (15.7)		71 (7.4)
Caucasian	288 (19.6)		183 (20.7)	
Southern and Central Asian	520 (35.3)		408 (46.2)	
East Asian	203 (13.8)		191 (21.6)	
African	76 (5.2)		55 (6.2)	
Oceanian not Australian	48 (3.3)		25 (2.8)	
Other	337 (22.9)		22 (2.5)	
Obstetric and family history				
Nullipara	667 (38.2)	0	327 (34.2)	0
Prior GDM	330 (18.9)	0	232 (24.3)	0
Prior LGA birth	64 (3.7)	0	45 (4.7)	0
Prior pre-eclampsia	72 (4.1)	0	34 (3.6)	0
Prior shoulder dystocia	19 (1.1)	0	16 (1.7)	0
Family history of diabetes	886 (50.7)	0	506 (53.0)	0
Physical characteristics				
Pre-pregnancy body mass index, kg/m ²	28.4 (6.9)	2 (0.1)	28.8 (7.2)	3 (0.3)
Gestational weight gain to GDM diagnosis, kg	5.7 (4.9)	761 (43.6)	5.7 (4.7)	524 (54.9)
Disease characteristics				
Gestational age at GDM diagnosis, weeks	24.3 (5.7)	126 (7.2)	23.7 (6.6)	200 (20.9)
Fasting glucose from diagnostic OGTT, mmol/L	5.0 (0.7)	60 (3.4)	5.0 (0.7)	45 (4.7)
1-hour glucose from diagnostic OGTT, mmol/L	9.9 (1.8)	70 (4.0)	10.0 (1.8)	57 (6.0)
2-hour glucose from diagnostic OGTT, mmol/L	8.2 (1.8)	72 (4.1)	8.1 (1.8)	56 (5.9)
Treatment				
Insulin therapy	640 (36.6%)	0	384 (40.2%)	0
Composite adverse pregnancy outcome	476 (27)		244 (26)	
LGA > 90 th percentile	186 (10.6)	0	100 (10.5)	6 (0.6)
Hypertensive disorders of pregnancy	130 (7.4)	0	74 (7.7)	0
Neonatal hypoglycaemia requiring IV therapy	203 (11.6)	0	95 (9.9)	0
Shoulder dystocia	43 (3.2)	0	17 (2.4)	0
Neonatal fracture	3 (0.2)	0	0 (0.0)	0
Neonatal nerve palsy	4 (0.2)	0	2 (0.2)	0
Fetal or neonatal death	11 (0.6)	0	9 (0.9)	0

Table 1: Characteristics of the development and validation cohorts of the PerSonal GDM Pregnancy Model and the proportion with missing data.

Abbreviations: LGA, large-for-gestational age; IV, intravenous; GDM, gestational diabetes mellitus; BMI, body mass index; OGTT, oral glucose tolerance test.

Predictors in the model	Coefficient	Bootstrap 95% CI	Odds ratio	95% CI
Maternal age	0.01	(0.00; 0.04)	1.01	(1.00; 1.04)
Pre-pregnancy BMI	0.04	(0.02; 0.06)	1.04	(1.02; 1.06)
Fasting glucose OGTT	0.32	(0.17; 0.50)	1.38	(1.19; 1.65)
1-hour glucose OGTT	0.06	(0.00; 0.13)	1.06	(1.00; 1.14)
Gestation at GDM diagnosis	−0.02	(−0.05; −0.00)	0.98	(0.95; 1.00)
Southern and Central Asian	−0.65	(−1.01; −0.39)	0.52	(0.36; 0.68)
East Asian	−0.14	(−0.64; 0.08)	0.87	(0.53; 1.08)
Nulliparity	0.17	(0.00; 0.47)	1.18	(1.00; 1.60)
Previous LGA baby	0.53	(0.00; 1.26)	1.70	(1.00; 3.53)
Previous pre-eclampsia	0.93	(0.41; 1.50)	2.53	(1.51; 4.48)
Gestational weight gain to GDM diagnosis per week	0.54	(0.02; 1.36)	1.71	(1.02; 3.90)
Family history of diabetes	−0.07	(−0.437; 0.00)	0.94	(0.65; 1.00)
Intercept	−4.11	(−5.53; −2.87)	0.02	(0.00; 0.06)

Table 2: Final model after LASSO selection with selected predictors, coefficients with bootstrap 95% confidence intervals and odds ratios.

Abbreviations: LASSO, Least Absolute Shrinkage and Selection Operator; CI, confidence interval; BMI, body mass index; OGTT, oral glucose tolerance test; LGA, large-for-gestational age; GDM, gestational diabetes mellitus.

Performance measures (95% CI)	Development cohort (n = 1747 women)	Validation cohort (n = 955 women)
C-statistic	0.68 (0.65; 0.71)	0.68 (0.64; 0.72)
Calibration slope	1.16 (0.96; 1.35)	0.99 (0.75; 1.23)
Calibration-in-the-large	0.01 (−0.10; 0.12)	−0.05 (−0.20; 0.11)
Expected: Observed ratio	1.00	1.03

Table 3: PerSonal GDM model performance.

Abbreviations: CI, confidence interval.

Model performance

In the development cohort, 27% (476/1747) of pregnant women with GDM had an adverse pregnancy outcome (Table 1). The most common outcome was neonatal hypoglycaemia requiring intravenous therapy (11.6%, 203/1747) followed by an LGA baby (10.6%, 186/1747) and hypertensive disorders of pregnancy (7.4%, 130/1747). The rate of shoulder dystocia was 3.2% (43/1747). Neonatal fracture (3/1747), nerve palsy (4/1747) and perinatal death were rare (11/1747).

The twelve predictors that were significantly associated with the composite adverse pregnancy outcome were included in the final model: maternal age, pre-pregnancy BMI, fasting and 1-hour OGTT values, gestation at GDM diagnosis, Southern and Central Asian ethnicity, East Asian ethnicity, nulliparity, previous LGA baby, previous pre-eclampsia, weight gain from booking until GDM diagnosis, and family history of diabetes (Table 2).

The final full prediction model is presented in Box 1. The final model c-statistic was 0.68 (95% CI 0.65 to 0.71) to discriminate between women with and without an adverse pregnancy outcome (Table 3). The calibration plot showed good agreement between predicted and observed risks overall, with a calibration slope of 1.16 (95% CI 0.96 to 1.35) and calibration-in-the-large of 0.01 (95% CI −0.10 to 0.12).

Box 1 Full prediction model to allow predictions for individuals.

The equation of the PerSonal GDM prediction model for risk of adverse pregnancy outcomes in women with GDM from a logistic regression model was as follows:

$$\text{percent risk} = \exp(Y) / (1 + \exp(Y)) * 100$$

where $Y = -4.11 + (0.04 * \text{pre-pregnancy body mass index in kg/m}^2) + (0.01 * \text{maternal age in years}) + (0.32 * \text{oral glucose tolerance test, fasting glucose in mmol/L}^\dagger) + (0.05 * \text{oral glucose tolerance test, 1-hour glucose mmol/L}^\dagger) - (0.02 * \text{Gestational age at GDM diagnosis in weeks completed}) - (0.65 * \text{South or Central Asian}) - (0.14 * \text{East Asian}) + (0.17 * \text{Nulliparous}) + (0.53 * \text{Past history of delivery of a large-for-gestational-age baby}) + (0.93 * \text{Past history of pre-eclampsia}) + (0.53 * \text{Gestational weight gain to GDM diagnosis per week in kilograms}) - (0.07 * \text{Family history of diabetes})$.

All variables are coded as binary (1 when present and 0 when absent) except for body mass index, maternal age, oral glucose tolerance test glucose levels and gestational age at GDM diagnosis which are continuous. † to convert glucose from conventional (mg/dL) to SI units (mmol/L), multiple by 0.06.

Temporal validation

In the validation cohort, 26% of pregnancies were complicated by an adverse pregnancy outcome (244/ 955), with the frequency of the component outcomes reported in [Table 1](#). The final model demonstrated acceptable discrimination when temporally validated, with a *c*-statistic in of 0.68 (95% CI 0.64, 0.72) ([Table 3](#)). The model also showed excellent agreement between predicted and observed risks overall, with a calibration slope of 0.99 (95% CI 0.75, 1.23) and calibration-in-the-large of -0.05 (95% CI -0.20, 0.11). Agreement was excellent over a predicted probability range of 0.17 to 0.55, which encompasses about 90% of the population ([Figure 1](#)). The calibration plot illustrates that the predicted probability was greater than the observed frequency at the higher end (an overestimate) of the range and less than the observed frequency at the lower end of the range (an underestimate) ([Figure 1](#)).

Clinical utility of model

Decision curve analysis demonstrated benefits of the PeRsonal GDM model in predicting women with GDM at risk of adverse pregnancy outcomes, compared to two alternative reference strategies: manage all as if they will or will not have an adverse pregnancy outcome ('treat all' or 'treat none'), for predicted probability thresholds between 0.15 and 0.85 ([Figure 2](#)). A threshold probability must be selected for each local setting that reflects the values of affected women and their clinicians and the resourcing and structure of the healthcare system. For probability thresholds < 0.15 and > 0.85 the model offers no net benefit compared to managing all women with GDM as if they will or will not have adverse pregnancy outcomes respectively ([Figure 2](#)).

In terms of the clinical utility, [Table 4](#) reports potential reduction in unnecessary intervention across probability thresholds. Applying the model can guide antenatal care with each threshold probability balancing two possibilities: treat as higher risk of adverse outcomes, yet ultimately these outcomes do not occur implying overtreatment (false positives), or treat as not at higher risk, ultimately with adverse outcomes presenting missed prevention opportunities (false negatives).

Sensitivity analyses

Sensitivity analyses using inverse probability treatment weighting demonstrates that the confounding effect of insulin treatment is minimal.

Discussion

The developed PeRsonal GDM model accurately predicts the risk of adverse pregnancy outcomes in women diagnosed with GDM. It includes twelve clinical predictors that are routinely available in clinical care: maternal

age, Southern and Central Asian ethnicity, East Asian ethnicity, pre-pregnancy or early pregnancy BMI, family history of diabetes, previous history of LGA baby, previous history of pre-eclampsia, gestational age at GDM diagnosis in current pregnancy, fasting and 1-hour glucose from the 75-g OGTT and GWG. The model shows excellent calibration and acceptable discrimination when temporally validated and identifies women with GDM at higher risk of adverse pregnancy outcomes over a broad range of clinically relevant predicted probability thresholds.

The majority of Australian births (75%) occur in Australia's universal accessible health system.²⁹ The PeRsonal GDM model was developed using routinely collected data from the largest Australian health network servicing urban and regional, ethnically diverse and low SES populations. Predictors and composite outcome components were identified through systematic review and appraisal of existing models,¹⁴ and multi-disciplinary input from obstetricians, endocrinologists, biostatisticians and public health experts.¹⁵ Only predictors that are easily accessible in clinical practice were considered, including maternal characteristics, relevant family and past history, glycaemic parameters and GWG, to optimise feasibility and generalisability across settings. The composite adverse pregnancy outcome components were selected based on association with severe maternal and perinatal morbidity and mortality, and the general need for multi-disciplinary specialist antenatal care.

We used robust statistical methods to develop the model, including handling continuous predictors as such and avoiding dichotomisation, using multiple imputation to deal with missing data and considering non-linear predictor-outcome relationships using fractional polynomials. The LASSO method for predictor selection simultaneously penalises the model coefficients for over-optimism generating a model which is more likely to perform consistently in new populations. Finally, we reported the clinical utility of the model using decision curve analysis, informing healthcare professionals and health systems on management of GDM for various model generated risk probabilities.¹⁷

Addressing the treatment paradox can present a challenge in prediction modelling.³⁰ Here model performance may be affected by insulin therapy use in cases with the highest glucose levels, where clinicians subjectively perceive the highest risk of adverse outcomes. However, here the confounding effect of insulin treatment on predictor-outcome associations was explored in a sensitivity analysis and found to be limited. The population for model development included women from three hospitals in the health network, ranging from midwifery low-risk care to obstetric high-risk care, yet IADPSG GDM diagnostic criteria and GDM management were consistent. This model was developed and validated in the same setting, at different time points. The model's performance may vary in a different

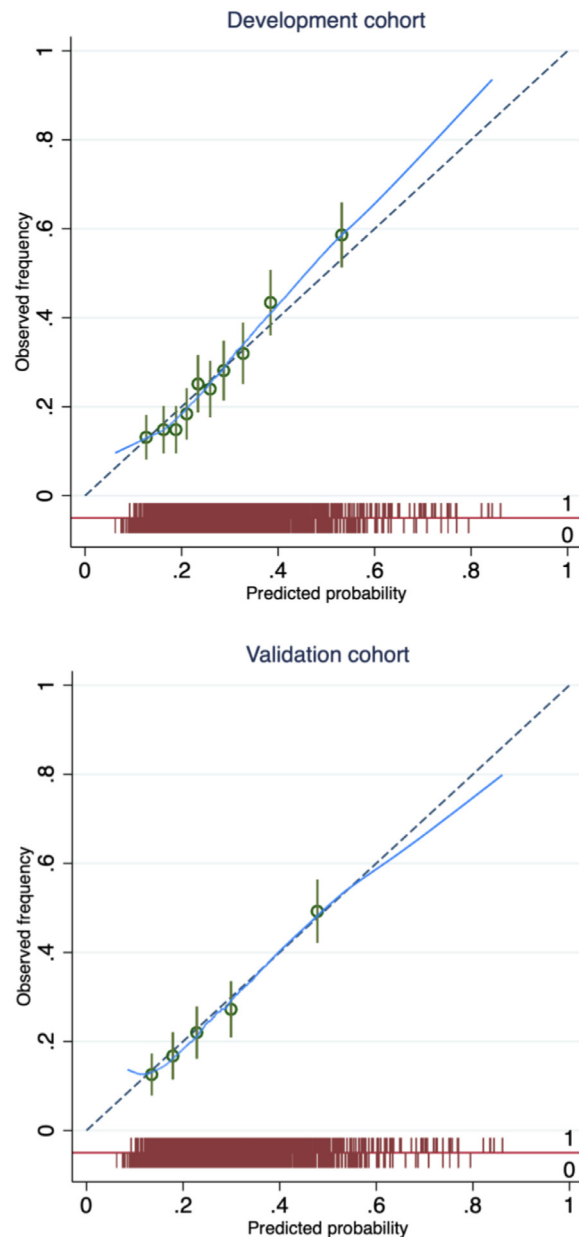


Figure 1. Calibration plot for the PerSonal GDM model. The predicted probability of adverse pregnancy outcome (x-axis) is compared to observed frequency (y-axis) in women with gestational diabetes in the development cohort (top panel) and the validation cohort (bottom panel). The plot is grouped by deciles and quintiles of the predicted risk (green circles) with 95% confidence intervals (green lines) in the development and validation cohort respectively, and supplemented by a smoothed (Lowess) line. A spike plot of the distribution of events (adverse pregnancy outcome) and non-events (red). Perfect predictions should lie on the 45 reference (dashed).

settings (e.g. community based care or low resource countries), by population characteristics, GDM diagnostic criteria and GDM management (e.g. metformin or insulin) without validation in similar populations.

We strongly support the principle of seeking to validate and, where possible, update existing models rather than developing new models de novo. Such an approach

focuses efforts on improving and ultimately implementing models into clinical practice. To date, five prediction modelling studies have developed ten models primarily intended to predict adverse pregnancy outcomes in women with GDM.¹⁴ Predictors in these models were most commonly fasting blood glucose levels, pre-pregnancy body mass index and maternal age. These models

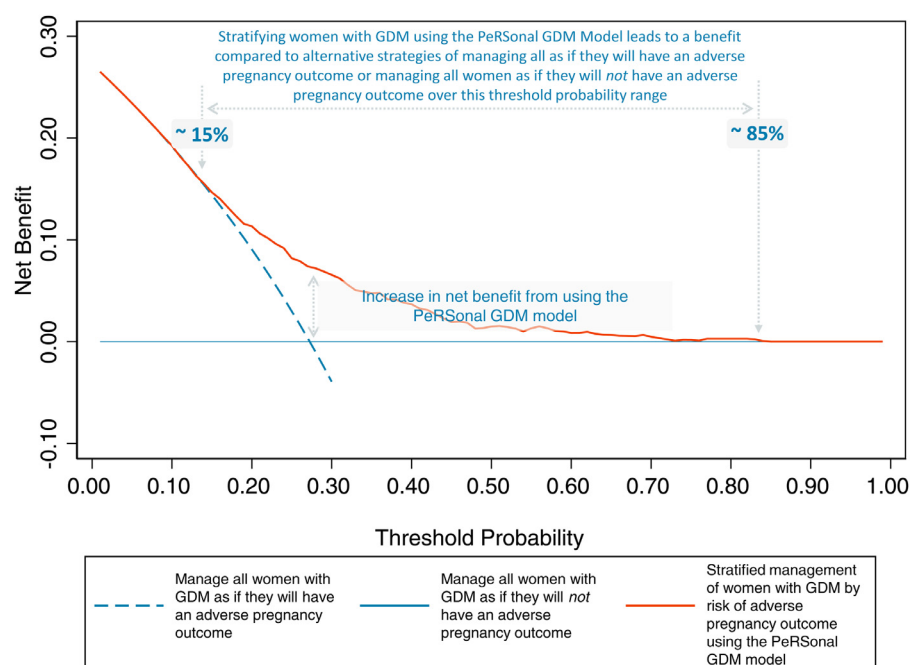


Figure 2. Decision curve analysis using the PerSonal GDM model to guide decisions at healthcare systems and individual shared decision making levels. As compared to a reference strategy where all women with GDM are managed as if they will have an adverse pregnancy outcome, the model offers no benefit at a threshold probability of < 0.15. Likewise, where all women with GDM are managed as if they will not have an adverse pregnancy outcome, the model offers no benefit at a threshold probability of > 0.85. Selection of a threshold probability between these thresholds, enables decisions on healthcare provision based on risk to be superior to the alternative reference strategies of managing all women as if they will or will not have an adverse pregnancy outcome.

Weight, false positive/ false negative	Threshold probability (%)	Net reduction in women unnecessarily managed as if they will experience an adverse pregnancy outcome (%)
1/5.6	15%	1.5
1/4	20%	9.1
1/3	25%	15.6
1/2.3	30%	24.5
1/1.8	35%	31.0
1/1.5	40%	37.4
1/1.2	45%	41.8
1/1	50%	47.0
1.2/1	55%	51.5
1.5/1	60%	55.2
1.8/1	65%	58.4
2.3/1	70%	61.3
3/1	75%	63.7
4/1	80%	66.0
5.6/1	85%	67.9

Table 4: Clinical utility of using the PerSonal GDM model compared to managing all women with gestational diabetes mellitus as if they will have an adverse pregnancy outcome, over the range of threshold probabilities. Intervention may include diversion away from routine antenatal care to a high-risk care setting with routine obstetric review and growth scans.

were not externally validated and model performance was inadequately reported with no useful measures of calibration nor formal evaluation of clinical utility limiting clinical application. Our initial plan was to externally validate and update these models for our population. Unfortunately, critical appraisal of these models identified gaps between the methods used and contemporary best practice.¹⁴ We therefore took the decision not to externally validate these existing models as model updating would require such significant structural changes that ultimately the benefits of using an existing model would be lost.

A model to predict the need for insulin therapy amongst women with GDM has been developed and externally validated³¹ and has been used to create a 'step down' approach to management in a metropolitan tertiary teaching hospital in Sydney, Australia.⁸ This approach focuses on organising care around the need for insulin therapy allowing targeting of diabetes clinician resources. But this approach negates issues related to the treatment paradox, and it may be less generalisable to settings where insulin therapy is not first-line or where glycaemic treatments targets differ. Moreover, a model designed specifically to predict adverse pregnancy outcomes, rather than glycaemic outcomes alone, may be more acceptable to affected women and clinicians. Acceptability by clinicians and women as well as shared decision making are areas of ongoing research.

Calibration, which quantifies the ability to accurately estimate the probability of a future event, is well suited to prognostic problems focused on the risk of a future event. Calibration must be considered alongside discrimination in prognostic clinical prediction model evaluation.³² Furthermore, in terms of clinical utility a prediction model which is well calibrated but has a lower discrimination than an alternative may be more clinically useful.³³ The PerSonal GDM model is well calibrated to this ethnically diverse cohort of IADPSG criteria diagnosed GDM i.e. observed risks are similar to predicted risks across various quintiles of risk thresholds and across time, over a broad predicted probability range (Figure 1). At the extremes, the model underestimates the risk of adverse pregnancy outcomes at the lower end of predicted probabilities (< 0.18), and overestimates the risk at the higher end (> 0.55).

Our model shows acceptable discrimination, accurately classifying women with and without adverse pregnancy outcomes. The negligible decrease in discrimination from the development to validation cohort suggests minimal overfitting and provides confidence in the overall robustness of the model. We note that interpretation of the *c*-statistic is challenging³⁴ and that the use of any threshold to determine a model's value is arbitrary. Given the excellent calibration characteristics of the PerSonal GDM model we argue that methodological limitations in the *c*-statistic are reflected here because of the distribution of risk in this

population, where we see little spread around the population average (27%) illustrated in the spike plot in Figure 1. This phenomenon has been empirically demonstrated and hence evaluating model performance on the *c*-statistic in isolation is ill advised.^{35,36} Here, the overall performance of this model based on discrimination and calibration is compelling and this is reflected in the decision curve analysis which demonstrates clinical utility across a range of predicted probabilities demonstrated (Figure 2).

Clinical acceptability of a prediction model is a prerequisite for translation into clinical care and the model must align with clinical understanding. The direction of association between predictors and the outcome was consistent with the literature, except for family history of diabetes. Glycaemic measures, BMI and maternal age are associated with adverse pregnancy outcomes³⁷⁻⁴¹ explained by the maternal hyperglycaemia-fetal hyperinsulinaemia hypothesis,⁴² elevations in maternal lipids⁴³⁻⁴⁵ and utero-placental dysfunction⁴⁶ respectively. Adverse pregnancy outcomes also vary by ethnicity⁴⁷ with immigrant ethnic Chinese and South Asian women at lower risk, compared to Australian-born Caucasian women⁴⁸ and white British women.⁴⁹ Earlier GDM diagnosis is associated with poor pregnancy outcomes^{50,51} potentially related to pathophysiological differences including higher insulin resistance related to adiposity.⁵² Unexpectedly, the model suggests that family history of diabetes is inversely correlated with adverse pregnancy outcome, potentially related to higher rates of early GDM screening, management and lower GWG. However the reasons for this association need further research.

Generally, health services manage all women with GDM as if they will have an adverse pregnancy outcome or adopt a 'step-up' model, where poor glycaemic control or the need for pharmacologic therapy is used as a surrogate for pregnancy risk. Our decision curve analysis demonstrates that in comparison to a reference 'treat all' strategy, the PerSonal GDM model offers a net benefit over a range of clinically relevant predicted probability thresholds.

To promote translation into clinical care, an electronic PerSonal GDM risk calculator has been developed allowing clinicians to calculate individualised risks of adverse pregnancy outcomes and to facilitate shared decision-making on antenatal care (available at <https://www.personalgdm.com/outcomes>). This also enables a risk-stratified approaches to treatment with those at highest risk recommended for more intensive monitoring and management and lower risk women offered less intensive models of care, with predefined escalation criteria where needed.

As resources and practice vary, we avoided recommending an arbitrary probability threshold and instead reported net-benefit estimates across the range of probability thresholds (Table 4), allowing nuanced local

management decisions. This risk-stratified approach with a threshold probability can be tailored to match women's and clinician's shared preference, health service structure, resources and capacity,²⁸ in consultation with service users and clinicians. Supplementary Box S1 presents an example of this approach, which can also be adapted to public health crises such as the COVID-19 pandemic. To reduce viral transmission and preserve limited resources, variation of thresholds can reduce referral to hospital based care.¹⁰

The clinical benefit of risk stratification based on the PeRsonal GDM model varies by risk threshold (Table 4) and impacts on personal burden, cost and convenience as well as health system resources. It also allows evolution away from a one-size-fits-all to a more personalised, risk-stratified approach to GDM care and can facilitate shared decision making. Further external validation of the PeRsonal GDM model to a more disparate population is now needed to assess the generalisability to different centres, community based care and low resource settings, other healthcare systems and to different GDM diagnostic criteria. It would also be beneficial if future external validation could be undertaken by independent investigators. To maximise usability and promote clinical application, an electronic risk calculator is needed along with an impact study to evaluate clinical, health service and health economic outcomes.

In conclusion, the PeRsonal GDM model can accurately predict the absolute risk of adverse pregnancy outcomes in women with GDM. Temporal validation showed that the model is transportable across time. Decision curve analysis demonstrated that stratifying women with GDM using the PeRsonal GDM model offers clinical utility, compared with the current default strategy of managing all women with GDM as if they will have an adverse pregnancy outcome, over a broad range of predicted probabilities. The PeRsonal GDM model can therefore facilitate shared decision-making at the individual level and risk-stratified care at a health service level, ultimately, supporting more personalised care for women with GDM.

Contributors

Conceptualisation: SDC, GS, JAB, JZ, ST, HJT.

Data curation: SDC, BMFF, HW.

Formal analysis: SDC, BMFF, JZ.

Funding acquisition: SDC, HJT.

Investigation: SDC, JAB, GS, JZ, BMFF, JA, HW, ST, HJT.

Project administration: SDC, ST, HJT.

Resources: SDC, ST, HJT.

Software: BMFF.

Supervision: JZ, ST, HJT.

Validation: SDC, JZ, BMFF, JA, ST, HJT.

Visualisation: SDC, JZ, BMFF, JA, ST, HJT.

Writing – original draft: SDC, JZ, HJT.

Writing – review & editing: SDC, JAB, GS, JZ, BMFF, JA, HW, ST, HJT.

The underlying data has been verified by SDC, BMFF, JZ and HJT. All authors had full access to the data in the study and accept responsibility for the decision to submit for publication.

Data sharing statement

The data that support the findings of this study is based on routine clinical data. Data sharing is dependent upon permission from the health service and ethics approval.

Declaration of interests

SDC reports grants from the National Health and Medical Research Council (NHMRC), Diabetes Australia, the Australian Academy of Science and the Australian Government Department of Education and Training during the conduct of the study; JAB reports grants from the NHMRC during the conduct of the study; BMFF reports grants from CIBER (Biomedical Research Network in Epidemiology and Public Health, Madrid, Spain) during the conduct of the study and HJT reports grants from the NHMRC and the Medical Research Future Fund during the conduct of the study; no other relationships or activities that could appear to have influenced the submitted work. All the other authors report no conflict of interests.

Acknowledgments

This work is supported by the Mothers and Gestational Diabetes in Australia 2 NHMRC funded project #1170847. We thank Dr Alice Stewart for providing a neonatology perspective in the study steering committee. We thank Dr Jennifer Wong and A/ Prof Arul Earnest for their constructive feedback throughout this project. We thank Ms Amanda Kendell and Ms Michelle Knight from the Monash Women's Information Team for their assistance in extracting the perinatal data. Finally, we thank Ms Rebecca Muir, Mr Haozhe (Zach) Wang and Mr Saad Ahmad for their assistance in developing the prototype electronic risk calculator during the Monash Young Medtech Innovations Healthcare Innovation Summers School program.

Supplementary materials

Supplementary material associated with this article can be found in the online version at doi:[10.1016/j.eclinm.2022.101637](https://doi.org/10.1016/j.eclinm.2022.101637).

References

- 1 Classification and diagnosis of diabetes mellitus and other categories of glucose intolerance. National Diabetes Data Group. *Diabetes*. 1979;28(12):1039–1057.
- 2 McIntyre HD, Catalano P, Zhang C, Desoye G, Mathiesen ER, Damm P. Gestational diabetes mellitus. *Nat Rev Dis Primers*. 2019;5(1):47.

- 3 Scifres C, Feghali M, Althouse AD, Caritis S, Catov J. Adverse outcomes and potential targets for intervention in gestational diabetes and obesity. *Obstet Gynecol*. 2015;126(2):316–325.
- 4 Huet J, Beucher G, Rod A, Morello R, Dreyfus M. Joint impact of gestational diabetes and obesity on perinatal outcomes. *J Gynecol Obstet Hum Reprod*. 2018;47(9):469–476.
- 5 Yuen L, Wong VW, Simmonds D. Ethnic disparities in gestational diabetes. *Curr Diab Rep*. 2018;18(9):68.
- 6 Goldstein RF, Abell SK, Ranasingha S, et al. Association of gestational weight gain with maternal and infant outcomes: a systematic review and meta-analysis. *JAMA*. 2017;317(21):2207–2225.
- 7 National Institute for Health and Care Excellence. *Diabetes in Pregnancy: Management from Preconception to the Postnatal Period (NICE Guideline [NG3])*. 2015. <https://www.nice.org.uk/guidance/ng3/resources/diabetes-in-pregnancy-management-from-preconception-to-the-postnatal-period-51038446021>. Accessed 5 April 2021.
- 8 Sina M, Cade TJ, Flack J, et al. Antenatal models of care for women with gestational diabetes mellitus: vignettes from an international meeting. *Aust N Z J Obstet Gynaecol*. 2020.
- 9 Cade TJ, Polyakov A, Brennecke SP. Implications of the introduction of new criteria for the diagnosis of gestational diabetes: a health outcome and cost of care analysis. *BMJ Open*. 2019;9(1):e023293.
- 10 Thangaratnam S, Cooray SD, Sukumar N, et al. Endocrinology in the time of covid-19: diagnosis and management of gestational diabetes mellitus. *Eur J Endocrinol*. 2020;183(2):G49–G56.
- 11 Craig L, Sims R, Glasziou P, Thomas R. Women's experiences of a diagnosis of gestational diabetes mellitus: a systematic review. *BMC Pregnancy Childbirth*. 2020;20(1):76.
- 12 Farrar D, Simmonds M, Griffin S, et al. The identification and treatment of women with hyperglycaemia in pregnancy: an analysis of individual participant data, systematic reviews, meta-analyses and an economic evaluation. *Health Technol Assess*. 2016;20(86):1–348.
- 13 Cooray SD, Boyle JA, Soldatos G, Thangaratnam S, Teede HJ. The need for personalized risk-stratified approaches to treatment for gestational diabetes: a narrative review. *Semin Reprod Med*. 2020;38(06):384–388.
- 14 Cooray SD, Wijeyaratne LA, Soldatos G, Allotey J, Boyle JA, Teede HJ. The Unrealised Potential for Predicting Pregnancy Complications in Women with Gestational Diabetes: A Systematic Review and Critical Appraisal. *Int J Environ Res Public Health*. 2020;17(9):3048.
- 15 Cooray SD, Boyle JA, Soldatos G, et al. Protocol for development and validation of a clinical prediction model for adverse pregnancy outcomes in women with gestational diabetes. *BMJ Open*. 2020;10(11):e038845.
- 16 Australian and New Zealand Clinical Trials Registry [Internet]. *The Prediction Modelling for Risk-Stratified Care for Women with Gestational Diabetes (PeRSONal GDM) Study: Calculating the Individualised Risk of Adverse Outcomes for Women with Gestational Diabetes (ACTRN12620000915954)*. 2020. <https://www.anzctr.org.au/ACTRN12620000915954.aspx>. Accessed 25 September 2020.
- 17 Collins GS, Reitsma JB, Altman DG, Moons KG. Transparent Reporting of a multivariable prediction model for Individual Prognosis or Diagnosis (TRIPOD): the TRIPOD statement. *Ann Intern Med*. 2015;162(1):55–63.
- 18 Steyerberg EW. Clinical prediction models: a practical approach to development, validation, and updating. *Chapter 17, Validation of Prediction Models*. 2nd Ed. New York; London: Springer International Publishing; 2019:629–657.
- 19 International Association of Diabetes Pregnancy Study Groups Consensus Panel Metzger BE, Gabbe SG, et al. International association of diabetes and pregnancy study groups recommendations on the diagnosis and classification of hyperglycemia in pregnancy. *Diabetes Care*. 2010;33(3):676–682.
- 20 Nankervis A, McIntyre HD, Moses RG, et al. ADIPS consensus guidelines for the testing and diagnosis of hyperglycaemia in pregnancy in Australia and New Zealand. *Sydney: Australian Diabetes in Pregnancy Society*. 2014.
- 21 Australian Bureau of Statistics. 1249.0 - Australian Standard Classification of Cultural and Ethnic Groups (ASCEG). 2016. 18 July 2016; <https://www.abs.gov.au/ausstats/abs@.nsf/mf/1249.0>. Accessed 2 October 2019.
- 22 Wolff RF, Moons KGM, Riley RD, et al. PROBAST: a tool to assess the risk of bias and applicability of prediction model studies. *Ann Intern Med*. 2019;170(1):51–58.
- 23 Moons KG, Altman DG, Reitsma JB, et al. Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): explanation and elaboration. *Ann Intern Med*. 2015;162(1):W1–73.
- 24 Morris TP, White IR, Carpenter JR, Stanworth SJ, Royston P. Combining fractional polynomial model building with multiple imputation. *Stat Med*. 2015;34(25):3298–3317.
- 25 Tibshirani R. Regression Shrinkage and Selection via the Lasso. *J R Statistic Soc Ser B (Methodological)*. 1996;58(1):267–288.
- 26 Little RJA, Rubin DB. *Statistical Analysis with Missing Data*. 2nd Ed. Hoboken, NJ: Wiley; 2002.
- 27 Vickers AJ, Elkin EB. Decision curve analysis: a novel method for evaluating prediction models. *Med Decis Making*. 2006;26(6):565–574.
- 28 Wynants L, van Smeden M, McLernon DJ, et al. Three myths about risk thresholds for prediction models. *BMC Med*. 2019;17(1):192.
- 29 Safer Care Victoria. *Victorian Perinatal Services Performance Indicators 2015-2016*. Melbourne: Victorian Government; 2017.
- 30 Cheong-See F, Allotey J, Marlin N, et al. Prediction models in obstetrics: understanding the treatment paradox and potential solutions to the threat it poses. *BJOG*. 2016;123(7):1060–1064.
- 31 Barnes RA, Wong T, Ross GP, et al. A novel validated model for the prediction of insulin therapy initiation and adverse perinatal outcomes in women with gestational diabetes mellitus. *Diabetologia*. 2016;59(11):2331–2338.
- 32 Cook NR. Statistical evaluation of prognostic versus diagnostic models: beyond the ROC curve. *Clin Chem*. 2008;54(1):17–23.
- 33 Van Calster B, Vickers AJ. Calibration of Risk Prediction Models. *Med Decis Making*. 2014;35(2):162–169.
- 34 Royston P, Altman DG. Visualizing and assessing discrimination in the logistic regression model. *Stat Med*. 2010;29(24):2508–2520.
- 35 Cook NR. Use and misuse of the receiver operating characteristic curve in risk prediction. *Circulation*. 2007;115(7):928–935.
- 36 Janssens ACJW, Martens FK. Reflection on modern methods: revisiting the area under the ROC Curve. *Int J Epidemiol*. 2020;49(4):1397–1403.
- 37 Hapo Study Cooperative Research Group Metzger BE, Lowe LP, et al. Hyperglycemia and adverse pregnancy outcomes. *N Engl J Med*. 2008;358(19):1991–2002.
- 38 Black MH, Sacks DA, Xiang AH, Lawrence JM. The relative contribution of prepregnancy overweight and obesity, gestational weight gain, and IADPSG-defined gestational diabetes mellitus to fetal overgrowth. *Diabetes Care*. 2013;36(1):56–62.
- 39 Catalano PM, McIntyre HD, Cruickshank JK, et al. The hyperglycemia and adverse pregnancy outcome study: associations of GDM and obesity with pregnancy outcomes. *Diabetes Care*. 2012;35(4):780–786.
- 40 Hildén K, Hanson U, Persson M, Fadl H. Overweight and obesity: a remaining problem in women treated for severe gestational diabetes. *Diabet Med*. 2016;33(8):1045–1051.
- 41 Lean SC, Derricott H, Jones RL, Heazell AEP. Advanced maternal age and adverse pregnancy outcomes: a systematic review and meta-analysis. *PLoS One*. 2017;12(10):e0186287.
- 42 Pedersen J. *Diabetes and Pregnancy: Blood Sugar of Newborn Infants*, [doctoral thesis]. Copenhagen, Denmark: Danish Science Press; 1952.
- 43 Schaefer-Graf UM, Graf K, Kulbacka I, et al. Maternal lipids as strong determinants of fetal environment and growth in pregnancies with gestational diabetes mellitus. *Diabetes care*. 2008;31(9):1858–1863.
- 44 Catalano PM, Hauguel-De Mouzon S. Is it time to revisit the Pederesen hypothesis in the face of the obesity epidemic? *Am J Obstet Gynecol*. 2011;204(6):479–487.
- 45 Barbour LA, Hernandez TL. Maternal non-glycemic contributors to fetal growth in obesity and gestational diabetes: spotlight on lipids. *Curr Diab Rep*. 2018;18(6):37.
- 46 Lean SC, Heazell AEP, Dilworth MR, Mills TA, Jones RL. Placental dysfunction underlies increased risk of fetal growth restriction and stillbirth in advanced maternal age women. *Sci Rep*. 2017;7(1):9677.

- 47 Khalil A, Rezende J, Akolekar R, Syngelaki A, Nicolaides KH. Maternal racial origin and adverse pregnancy outcome: a cohort study. *Ultrasound Obstet Gynecol*. 2013;41(3):278–285.
- 48 Wan CS, Abell S, Aroni R, Nankervis A, Boyle J, Teede H. Ethnic differences in prevalence, risk factors, and perinatal outcomes of gestational diabetes mellitus: a comparison between immigrant ethnic Chinese women and Australian-born Caucasian women in Australia. *J Diabetes*. 2019;11(10):809–817.
- 49 Farrar D, Fairley L, Santorelli G, et al. Association between hyperglycaemia and adverse perinatal outcomes in south Asian and white British women: analysis of data from the Born in Bradford cohort. *Lancet Diabetes Endocrinol*. 2015;3(10):795–804.
- 50 Bartha JL, Martinez-Del-Fresno P, Comino-Delgado R. Gestational diabetes mellitus diagnosed during early pregnancy. *Am J Obstet Gynecol*. 2000;182(2):346–350.
- 51 Sweeting AN, Ross GP, Hyett J, et al. Gestational diabetes mellitus in early pregnancy: evidence for poor pregnancy outcomes despite treatment. *Diabetes Care*. 2016;39(1):75–81.
- 52 Bozkurt L, Göbl CS, Pfligl L, et al. Pathophysiological characteristics and effects of obesity in women with early and late manifestation of gestational diabetes diagnosed by the international association of diabetes and pregnancy study groups criteria. *J Clin Endocrinol Metab*. 2015;100(3):1113–1120.