RESEARCH

Open Access

Diabetes and prediabetes among women universally screened for gestational diabetes: a multi-ethnic, population-based, prospective study with eleven years follow-up

Christin W. Waage[1,2*], Anne Karen Jenum[1], Ibrahimu Mdala[1], Sindre Lee-Ødegård[3,4], Anja Maria Braend[1], Line Sletner[1], Jens Petter Berg[3,5] and Kåre I. Birkeland[3,4]

Abstract

Background Gestational diabetes (GDM) is a strong risk factor for later development of diabetes. However, data are scarce on the long-term risk for diabetes or prediabetes diagnosed by HbA1c, in non-selected, multi-ethnic populations universally screened for GDM using the WHO_{2013} criteria. We aimed to investigate the development of diabetes or prediabetes eleven years after the index pregnancy and identify risk factors in pregnancy or shortly after.

Methods A population-based cohort study of 360 women with complete eleven years follow-up data for diabetes (HbA1c ≥ 48 mmol/mol) or $prediabetes_{ADA}$ (HbA1c 39–47 mmol/mol). Women were enrolled in gestational week 15 and universally screened with an oral glucose tolerance test in week 28. We performed least absolute shrinkage and selection operator (LASSO) regression to identify predictors of future diabetes or $prediabetes_{ADA}$ and constructed a nomogram to predict individual risks.

Results Diabetes or $prediabetes_{ADA}$ combined, was found in 26.9%, and the prevalence was slightly higher in previous GDM compared with non-GDM women (35.6% versus 23.5%; $p = 0.019$). The relative risk (RR) for developing diabetes or $prediabetes_{ADA}$ was moderately elevated in GDM compared with non-GDM women (1.4 [1.0, 1.9], $p = 0.035$). Seven women (1.9%) had diabetes and all of these except for one, had previous GDM. Hence, the crude prevalence was 5.8% among GDM women vs. 0.4% among non-GDM women. The RR for developing diabetes was substantially higher in GDM vs. non-GDM women (14.8 [2.6, 277.1], $p = 0.012$). $Prediabetes_{ADA}$ was found in 25% and the RR for $prediabetes_{ADA}$ was not significantly increased for GDM compared to non-GDM women (1.3 [0.9, 1.8], $p = 0.143$). Among Europeans, 17.0% had diabetes or $prediabetes_{ADA}$, compared to 43.0% among South Asians ($p < 0.001$) and 34.4% among other ethnicities ($p = 0.002$). The most significant predictors identified from the LASSO were HbA1c measured in early pregnancy, ethnicity, and a family history of diabetes.

*Correspondence:
Christin W. Waage
c.w.waage@medisin.uio.no

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

Waage *et al. BMC Public Health*   (2025) 25:1264

Page 2 of 12

Conclusions The risk for developing diabetes was low, overall and among GDM women. Still GDM represented a strong risk for diabetes, but not for prediabetes$_{ADA}$. HbA1c early in pregnancy, non-European ethnicity, and a family history of diabetes were the strongest risk factors for developing diabetes or prediabetes$_{ADA}$.

Trial registration STORK G2 Women and Risk of Diabetes. NCT03870724 (ClinicalTrials.gov). February 27th, 2019.

Keywords Gestational diabetes, Type 2 diabetes, Ethnicity

Background

Gestational diabetes mellitus (GDM) and diabetes share genetic and environmental risk factors [1, 2]. GDM is considered as an early marker of disturbances in the glucose metabolism [3], and women diagnosed with GDM have increased risk of developing diabetes later in life [4]. However, the reported relative and absolute risks for developing diabetes after GDM vary widely, due to differences in diagnostic criteria and recommended screening procedures [5]. Populations examined also vary largely in ethnic and genetic background, exposure to unhealthy lifestyle and length of follow-up [6, 7]. Few studies have reported diabetes and prediabetes evaluated by HbA1c only, the currently most used diagnostic test.

The original criteria for screening and diagnosis of GDM primarily focused on identifying women at high risk for development of diabetes after pregnancy [8, 9]. However, the current criteria aim to prevent complications in mother and child in the actual pregnancy [10, 11]. The WHO$_{2013}$ diagnostic criteria for GDM have a lower threshold for fasting plasma glucose (FPG) compared to the previous WHO$_{1999}$ criteria. Hence, more women are diagnosed with GDM and are in the need for follow-up during and after pregnancy. Moreover, several stakeholders, including the American Diabetes Association (ADA), now recommend universal screening for GDM with an oral glucose tolerance test (OGTT) in gestational week 24–28 [12], in contrast to high-risk screening as previously used. The adverse effects of slightly elevated glucose in pregnancy for the mother and the newborn has been thoroughly described [13]. However, there is far less data on the long-term risk for diabetes or prediabetes in the mother after universal screening for GDM using the WHO$_{2013}$ criteria.

We have previously reported a high prevalence of GDM by the WHO$_{1999}$ criteria, and substantially higher by the WHO$_{2013}$ criteria, in the STORK Groruddalen (STORK G) cohort, set up to explore short and long-term ethnic differences in maternal health, primarily GDM, and effects of exposures on risk for type 2 diabetes, cardiovascular disease and other health issues [14]. In this population-based, multi-ethnic cohort of pregnant women, we aimed to investigate the development of HbA1c-diagnosed diabetes or prediabetes eleven years after the index pregnancy in which they underwent universal screening for GDM. We also aimed to identify risk factors that predicted later diabetes or prediabetes, as this information may be used to identify women who would benefit from targeted follow-up.

Methods

The primary outcome was diabetes (HbA1c ≥ 48 mmol/mol (≥ 6.5%)) or prediabetes$_{ADA}$ (HbA1c 39–47 mmol/mol (5.7–6.4%)) [15] eleven years after the index pregnancy, stratified by GDM status and ethnic group.

Baseline examinations 2008–2011

The population-based STORK G cohort was set up in 2008–2010 at three primary care Child Health Clinics in Groruddalen, Oslo, Norway, a district with a high proportion of ethnic minorities mostly from Asia and Africa [16]. To facilitate inclusion of ethnic minority women often excluded from research, information material and questionnaires were provided in eight languages: Arabic, English, Sorani, Somali, Tamile, Turkish, Urdu, and Vietnamese. Healthy pregnant women attending antenatal care were invited to participate, if they (1) were living in one of the districts, (2) planned to give birth at the two nearby study hospitals (Akershus University Hospital and Oslo University Hospital, Ullevål), (3) were in gestational week < 20, (4) were not suffering from diseases necessitating intensive hospital follow-up during pregnancy, (5) were not already included with a pregnancy lasting > 22 weeks, (6) could communicate in Norwegian or any of the other eight languages, and (7) were able to give informed consent [16]. A rich dataset including questionnaire data, physical measurements and fasting venous samples from gestational week 15 (visit 1), gestational week 28 (visit 2), and 14 weeks postpartum (visit 3) was collected by trained study personnel. Of the 823 healthy pregnant women included (74% of invited), 59% had ethnic minority background, and the sample was found to be representative of those invited [16, 17].

Fasting blood samples were collected at all visits [1–3] after an overnight fast, analysed consecutively at Akershus University Hospital and the Hormone Laboratory, Oslo University Hospital, or were biobanked and stored at -80 °C. FPG values were measured from venous blood on gel tubes, allowed to clot for 30 min before cells were separated from serum, stored at +4 °C and daily shipped and handled at the Department of Multidisciplinary Laboratory Medicine and Medical Biochemistry,
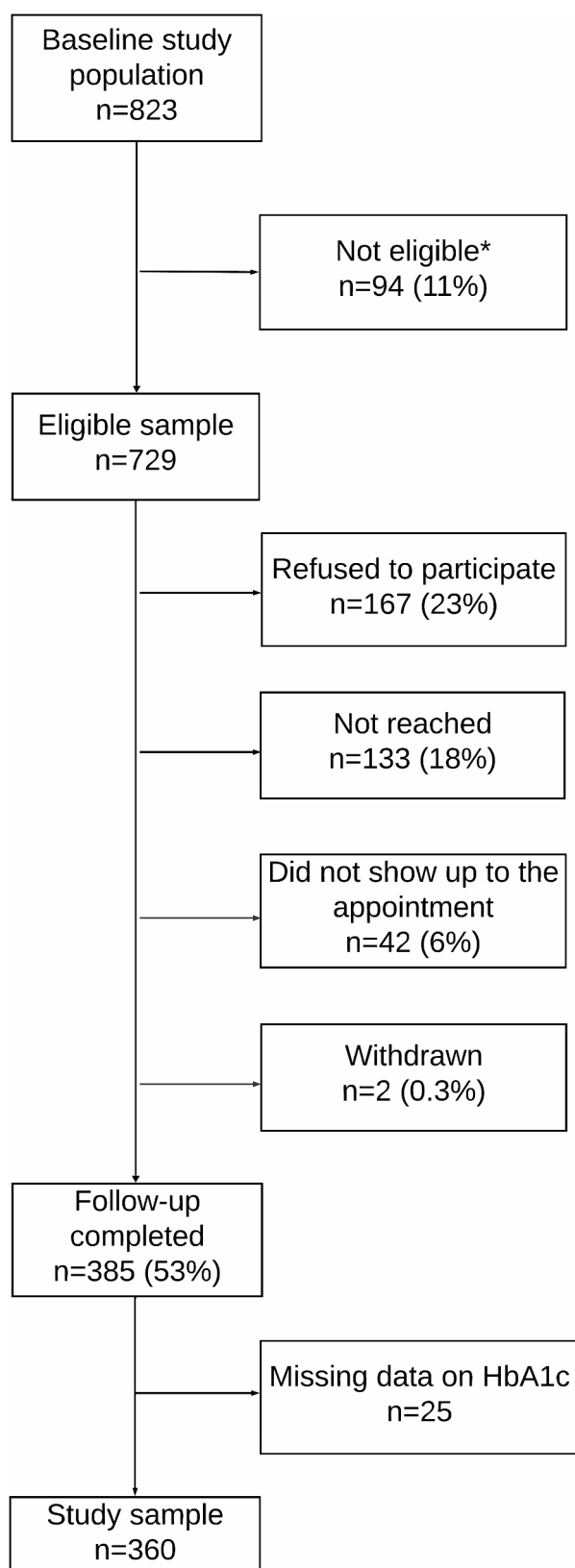
```
┌─────────────────┐
│ Baseline study  │
│   population    │
│     n=823       │
└─────────────────┘
         │            ┌──────────────────┐
         ├───────────▶│ Not eligible*    │
         │            │ n=94 (11%)       │
         │            └──────────────────┘
         ▼
┌─────────────────┐
│ Eligible sample │
│     n=729       │
└─────────────────┘
         │            ┌──────────────────────┐
         ├───────────▶│ Refused to participate│
         │            │ n=167 (23%)          │
         │            └──────────────────────┘
         │            ┌──────────────────┐
         ├───────────▶│ Not reached      │
         │            │ n=133 (18%)      │
         │            └──────────────────┘
         │            ┌──────────────────────┐
         ├───────────▶│ Did not show up to the│
         │            │ appointment          │
         │            │ n=42 (6%)            │
         │            └──────────────────────┘
         │            ┌──────────────────┐
         ├───────────▶│ Withdrawn        │
         │            │ n=2 (0.3%)       │
         │            └──────────────────┘
         ▼
┌─────────────────┐
│ Follow-up       │
│ completed       │
│ n=385 (53%)     │
└─────────────────┘
         │            ┌──────────────────────┐
         ├───────────▶│ Missing data on HbA1c│
         │            │ n=25                 │
         │            └──────────────────────┘
         ▼
┌─────────────────┐
│ Study sample    │
│    n=360        │
└─────────────────┘
```

Fig. 1 Flow chart. HbA1c based on capillary dried blood spot (*n* = 360, 94%)

Akershus University Hospital (Vitros 5.1 FS, Ortho Clinical Diagnostics, slide adapted colorimetric method). HbA1c was measured in venous EDTA blood with HPLC (Tosoh G8, Tosoh Corporation) [18]. Fasting total triglycerides were analysed in serum with a colorimetric method (Vitros 5.1 FS, Ortho clinical diagnostics). Leptin and C-peptide were measured at the Hormone Laboratory, Oslo University Hospital, leptin with radioimmunoassay from Millipore Corp (formerly from Linco Research Inc) and C-peptide by non-competitive immunofluorometric assays (DELFIA, PerkinElmer Life Sciences, Wallac Oy, Turku, Finland). Insulin resistance (HOMA-IR) and β-cell function (HOMA-β) were estimated by the Oxford University HOMA Calculator 2.2 from FPG and C-peptide concentrations [19].

In gestational week 28, all women were given a standard 75 g oral glucose tolerance test (OGTT) after an overnight fast [17]. Fasting at 2-h glucose was measured on site in venous EDTA blood, using a patient-near method (HemoCue 201+, Angelholm, Sweden). In total, 99 women were diagnosed with GDM by WHO_{1999} criteria of whom only five needed pharmacological glucose lowering treatment during pregnancy, and none had diabetes at the 14-week postpartum visit. In this study, however, we primarily use the new WHO_{2013} criteria to define GDM (FPG ≥ 5.1 mmol/l or 2-hour glucose ≥ 8.5 mmol/l), hereafter called $GDM_{WHO2013}$, but also report results for the WHO_{1999} criteria (FPG ≥ 7.0 and/or 2-hour glucose ≥ 7.8 mmol/l), and hereafter called $GDM_{WHO1999}$.

Maternal age at inclusion was calculated based on date of birth. Parity was categorised as nulliparous or multiparous (≥ 1), referring to status before the index pregnancy. Education was categorised as lower level (< 12 years) or higher level (≥ 12 years) [14]. Family history of diabetes was self-reported and categorized as yes or no. Height was measured to the nearest 0.1 cm using a fixed stadiometer at inclusion. Body weight was measured with a bioelectrical impedance analysis scale (Tanita-BC 418 MA, Tanita Corporation, Tokyo, Japan) [20], and body mass index (BMI) (kg/m^2) was calculated.

Ethnicity was defined based on self-reported country of origin of the woman or her mother, if the participant's mother was born outside Europe or North America [14]. We thereafter merged the largest groups and used the following categories: European, South Asian (primarily from Pakistan and Sri Lanka), and other ethnicities (mostly from Middle Eastern countries, Eastern Asia and Africa). Only 6.4% of non-Europeans (*n* = 23) were second-generation immigrants.

Follow-up examinations 2019–2022

Eleven years postpartum, 729 women (86% of the original cohort) still living in Norway and not excluded for reasons given in the flow chart (Fig. 1) were eligible for

a follow-up (visit 4), 385 (53%) met and 360 (49.4%) had a valid HbA1c [14] and were included in the present study. All eligible women living in the Oslo region were invited by telephone to a new visit at the same study sites, including an interviewer-administered digital questionnaire, clinical measurements, and non-fasting capillary dried blood spots were sampled and biobanked [14]. In addition, women were asked to attend another day for a fasting blood sample. Recruitment of study participants stopped temporarily during three periods during the COVID-19 pandemic. Eligible women who had moved outside the area ($n = 96$, 95 ethnic Norwegians), were invited to a less extensive data collection based on telephone interviews. These women were encouraged to take a dried blood spot themselves, supported by oral and written information, but not asked to visit for fasting samples due to long travel distances [14].

Body height (cm) was measured using a fixed stadiometer. Body weight (kg) and fat mass (kg) were measured with bioimpedance analysis using the Tanita-MC-780MA. Actual parity was categorised as primipara, para 2 or para 3+. Information on present educational status was categorised as primary school, high school or college/university. HbA1c was measured in dried blood spot samples by the Vitas laboratory (Vitas AS Oslo, Norway) [14]. One single dried blood spot punch for each woman was subjected to hemolysis for 30 min at room temperature using 600 µl of lysate buffer. Subsequently, an aliquot of the lysate was analyzed for HbA1c using ELISA. The absorbance was read at 660 nm using Spark 10 M microplate-reader from Tecan, Switzerland. The analytical coefficient of variance was < 3%. In addition, blood samples from women who agreed to meet fasting 1–7 days after the clinical examination ($n = 213$) were also analysed at a routine, high-volume, Clinical Chemistry Laboratory (Fürst Laboratory, Oslo), for FPG (ADVIA Glucose Hexokinase) and HbA1c (Tosoh G11 HPLC). The HbA1c measurements performed on filter-paper and by HPLC showed good agreement [14].

Power calculations

The original power calculations estimated that a sample size of 400 would provide a power of > 90% to detect significant differences for diabetes (2% vs. 8% in Europeans and South Asians, respectively) and prediabetes $_{WHO}$ (FPG 6.1–6.9 mmol/L), 6% and 20%, respectively). Due to unforeseen restrictions (the COVID-19 pandemic and lack of funding), the number included in the present study was somewhat smaller.

Statistical analyses

We summarized continuous data using the mean and standard deviation (SD) and compared mean differences between groups using the independent samples t-test. Categorical variables were described by frequencies and proportions and compared using the Chi-square test. We estimated relative risks (RR) using the generalized linear regression model with *binomial* family and a *log* link function.

We first performed univariate logistic regression analyses to identify potential predictors for our primary endpoint diabetes or prediabetes$_{ADA}$. Thereafter, we used the least absolute shrinkage and selector operator (LASSO) which selected a subset of the best predictors, by shrinking the regression coefficients of the least contributing independent variables to zero. This implies that all the independent variables that were set to zero were eliminated from the predictive model. To achieve this, a tuning parameter lambda, which determines the amount of shrinkage was selected using 10-fold cross validation.

Identified input variables were:

- from visit 1: age, ethnicity, parity, education level, body height (cm), body weight (kg), pre-pregnant BMI (kg/m²), BMI (kg/m²), total fat mass (kg), HbA1c (mmol/mol), FPG (mmol/L), HOMA-IR, HOMA-β, triglycerides (mmol/L), Leptin (ng/ml) and first-degree relatives with diabetes.
- from visit 2: body weight (kg), BMI (kg/m²), total fat mass (kg), Leptin (ng/ml), HbA1c (mmol/mol), FPG (mmol/L), HOMA-IR, HOMA-β, 2-hour plasma glucose (mmol/L) and GDM$_{WHO2013}$.
- from visit 3: body weight (kg), BMI (kg/m²), total fat mass (kg), triglycerides (mmol/L), Leptin (ng/ml), FPG (mmol/L), HOMA-IR and HOMA-β.

Due to some missing data particularly from the postpartum visit, we imputed missing data using the miceadds package in R and generated 10 different data sets, which we exported into Stata. We used predictive mean matching (pmm) to impute continuous variables, binary logistic regression (logreg) to impute binary variables and polytomous logistic regression (polyreg) to impute unordered factors. For each data set, a LASSO logistic regression model was fitted to the data. We compared the performances of the 10 fitted models and selected the model that included predictors that were also selected by the other nine LASSO models. More details about Model development and prediction of prediabetes or diabetes can be found in the Supplementary Box 1. However, since we did not have an external validation cohort to validate the prediction model, we considered a random sampling of the participants "with replacement" using bootstrapping. We used 500 bootstrap re-samplings.

The ability of the final predictive model in differentiating women with diabetes or prediabetes from those without was assessed using the area under the receiver operating characteristic curve (AUC). Threshold values

of AUC ranging from 50 to 69% indicates that the ability of the model to differentiate women with diabetes or prediabetes from those without is poor, an AUC of 70–80% represents an acceptable model, 80–90% an excellent model and above 90% a model with outstanding performance.

Finally, we developed a nomogram (Fig. 2, Supplementary Fig. 1a and 1b) based on the predictors selected by the LASSO logistic regression model to identify the risk for diabetes or prediabetes in individual women. The probability can be calculated as follows: scores for all predictors are obtained, summed to obtain a total score, which is then mapped on the probability graph.

All statistical analyses were performed in StataSE 18. A two-sided *P*-value < 0.05 was considered statistically significant (unless otherwise specified).

Results

Sample characteristics

Compared to women in the original cohort who did not meet at the eleven years follow-up, the 360 women included in the present study were somewhat older, more educated, and less likely to be from ethnic minorities (Supplementary Table 1). However, there were no differences in mean HbA1c or BMI between the two groups. The proportion of women fulfilling the GDM$_{WHO2013}$ criteria in the index pregnancy was 28.9% among those included and 32.8% among those who did not meet for follow-up (*p* = 0.254).

Overall, 206 (57.2%) of the participants were of European ethnic background, 86 (23.9%) of South Asian and 68 (18.9%) of other ethnic backgrounds. Of the 360 included at the 11 years follow up study, mean age was 41.9 (SD 4.8), mean BMI was 27.0 (5.0) kg/m^2 and 63.2% of the women had higher education (College/University). Hypertension (systolic blood pressure ≥ 140 mmHg or diastolic blood pressure ≥ 90 mmHg [21]) was found in 11.9% and morbid obesity (BMI ≥ 40 kg/m^2) was observed in 1.4% of participants. Further characteristics for participants with and without GDM$_{WHO2013}$ are shown in Table 1. At the baseline examination in gestational week 15, women with GDM were slightly older, more often had ancestral origin outside Europe, had higher parity and more often diabetes in their family. They had higher
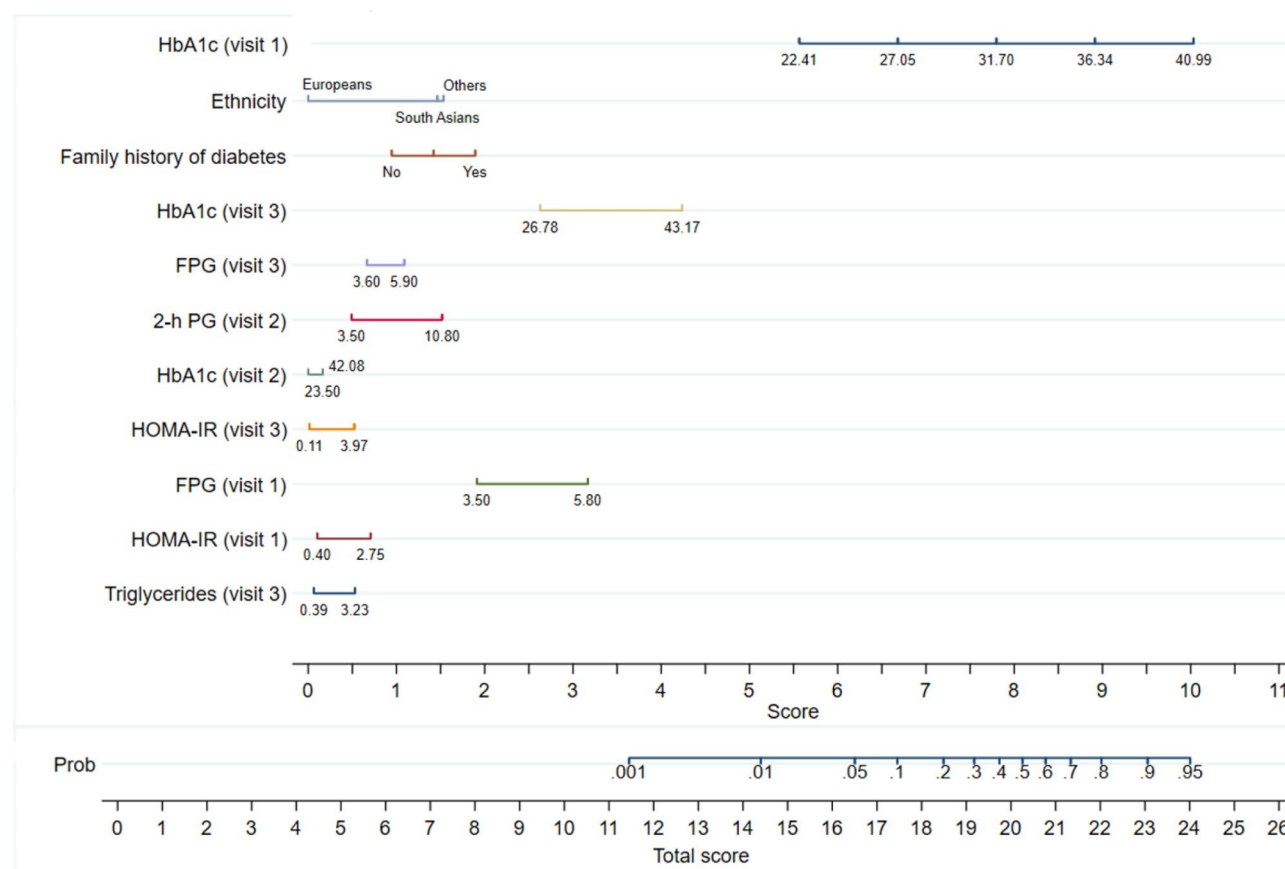

Fig. 2 Nomogram estimating the probability of having prediabetes or diabetes as a function of 11 predictors. The predictors were selected by the Lasso logistic regression model. For a specific individual, each predictor value is given a score and a total score obtained by adding the individual scores. The total score is projected on total points axis to obtain the probability of prediabetes or diabetes

Waage *et al. BMC Public Health*   (2025) 25:1264

Page 6 of 12

Table 1 Characteristics of participants with and without gestational diabetes (GDM) (WHO$_{2013}$) in index pregnancy

		GDM n = 104 (28.9%)		No GDM n = 256 (71.1%)	*p* value
Gestational week 15	n		n		
Age at enrolment (years)	104	31.5 (4.8)	256	30.1 (4.5)	**0.007**
Parity (%)	104		256		**0.026**
Nulliparous	39	37.5	129	50.4	
Multiparous	65	62.5	127	49.6	
Education (%)	103		254		0.163
Primary school or less	12	11.6	21	8.3	
High school	43	41.8	87	34.2	
College/University	48	46.6	146	57.5	
Ethnicity (%)	104		256		**0.001**
Europe	45	43.3	161	62.9	
South Asia	37	35.6	49	19.1	
Other ethnicity	22	21.1	46	18.0	
Family history of type 2 diabetes (%)	101		253		**0.008**
Yes	30	29.7	43	17.0	
Pre-pregnant BMI (kg/m^2)	104	25.9 (5.2)	253	23.8 (3.9)	**< 0.001**
Body height (cm)	104	163.7 (7.2)	256	164.4 (6.8)	0.419
Body weight (kg)	104	71.6 (15.3)	256	65.8 (12.0)	**< 0.001**
BMI (kg/m^2)	104	26.7 (5.2)	256	24.3 (3.9)	**< 0.001**
HbA1c (mmol/mol)	102	33.9 (3.5)	250	32.7 (2.8)	**0.001**
HbA1c (%)	102	5.3 (0.3)	250	5.1 (0.3)	**0.001**
FPG (mmol/L)	103	4.7 (0.4)	253	4.3 (0.3)	**< 0.001**
11 years follow up					
Age at enrolment (years)	104	42.8 (4.9)	256	41.5 (4.7)	**0.022**
Time since index pregnancy (years)	102	10.8 (1.0)	254	10.9 (0.9)	0.104
Parity (%)	101		250		0.075
Primipara	8	7.9	27	10.8	
Para 2	43	42.6	132	52.8	
Para 3+	50	49.5	91	36.4	
Education (%)	102		254		0.149
Primary school	10	9.8	13	5.1	
High school	34	33.3	74	29.1	
College/University	58	56.9	167	65.8	
Body height (cm)	102	163.1 (7.4)	256	163.8 (6.9)	0.343
Body weight (kg)	104	75.2 (15.2)	254	70.9 (13.5)	**0.009**
BMI (kg/m^2)	104	28.2 (5.3)	254	26.4 (4.9)	**0.002**
HbA1c (mmol/mol)	104	38.6 (6.1)	256	36.8 (3.2)	**< 0.001**
HbA1c (%)	104	5.7 (0.6)	256	5.5 (0.3)	**< 0.001**
FPG (mmol/L)	68	5.2 (0.9)	133	4.7 (0.6)	**< 0.001**

BMI: body mass index. GDM$_{2013}$: FPG ≥ 5.1 or 2-hour PG ≥ 8.5 mmol/l. Other ethnicity: East Asia, Middle East, Africa. Participants were universally screened for GDM at mean gestational week 28. Values are presented as mean (sd) of frequencies (%). GDM versus no GDM were compared by two sample t-test or Pearson chi-squared test. P values < 0.05 are in bold
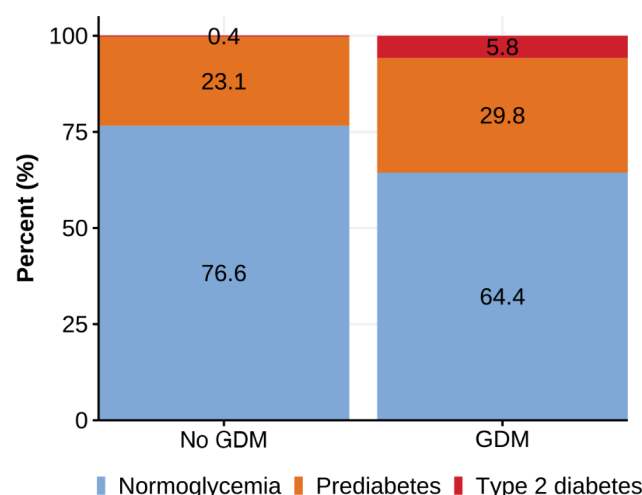
Fig. 3 Prediabetes and diabetes by GDM (WHO$_{2013}$) status

pre-pregnant and actual BMI, higher weight, HbA1c and FPG, compared with women without GDM. Also at the eleven years follow-up, actual BMI, HbA1c and FPG were higher in women with previous GDM than in those without (Table 1).

Diabetes and prediabetes at follow-up

Diabetes or prediabetes$_{ADA}$ combined (our primary outcome) was found in 26.9%, and the prevalence was slightly higher in women with GDM$_{WHO2013}$ compared with women without GDM in the index pregnancy (35.6% versus 23.5%; *p* = 0.019) (Fig. 3, Supplementary Table 3). The RR for developing diabetes or prediabetes$_{ADA}$ was moderately elevated in women with GDM in the index pregnancy compared to those without 1.4 (95% CI: 1.0, 1.9), *p* = 0.035.

Four women reported to have been diagnosed with diabetes prior to the study follow-up (two were treated with glucose lowering drugs). In addition, four women were detected with undiagnosed diabetes based on the HbA1c measured at the study visit. As one European woman with self-reported diabetes had a normal HbA1c (34 mmol/mol) without using glucose lowering drugs (since the postpartum visit she had 15 kg weight loss), a total of seven women (1.9%) had diabetes based on HbA1c, all with non-European ethnic origin (Supplementary Table 3). With a single exception, diabetes was only found in women with GDM$_{WHO2013}$ in the index pregnancy, representing a prevalence of 5.8% GDM women (Supplementary Table 3) compared with 0.4% in non-GDM women. The RR for developing diabetes was substantially higher in GDM$_{WHO2013}$ vs. non-GDM women 14.8 (95% CI: 2.6, 277.1), *p* = 0.012.

Prediabetes$_{ADA}$ was found in 90 (25%) of the women (Supplementary Table 3) and the RR for developing prediabetes$_{ADA}$ for GDM$_{WHO2013}$ compared to non-GDM

women was not significantly increased 1.3 (95% CI: 0.9, 1.8), $p = 0.143$. Using the more conservative definition of prediabetes as recommended by the WHO, only 24 women (6.7%) had prediabetes$_{WHO}$, with a RR for of 3.2 (95% CI 1.6, 6.5), $p = 0.001$ in GDM vs. non-GDM women.

When investigating each ethnic group separately, 17.0% of women with European ethnicity had diabetes or prediabetes$_{ADA}$ at follow-up, compared to 43.0% of South Asians ($p < 0.001$) and 34.4% of women from other ethnicities ($p = 0.002$) (Supplementary Table 3). The proportions with diabetes or prediabetes$_{ADA}$ among GDM$_{WHO2013}$ and non-GDM women are shown in Fig. 4. We found no difference in the prevalence of diabetes or prediabetes$_{ADA}$ combined between the South Asian and the other ethnic minority groups, but European women had a substantially lower RR compared to the two ethnic minority groups combined (0.4 [0.3, 0.6], $p < 0.001$).

When applying the WHO$_{1999}$ diagnostic criteria for GDM in the index pregnancy, only 45 (12.5%) had GDM.

In these women, 12 (24.4%) had prediabetes$_{ADA}$ at follow-up, while 4 (8.9%) had diabetes, compared with 3 (0.95%) in non-GDM women (Supplementary Table 3).

Predictors for diabetes or prediabetes$_{ADA}$ combined

Supplementary Table 4 shows the relationship between potential predictors from visit 1 to 3 and the risk of developing diabetes or prediabetes $_{ADA}$ as observed. After analysing these factors with LASSO, the most significant predictors identified were HbA1c, ethnicity and a family history of diabetes, all recorded in early pregnancy (Table 2). Additionally, fasting FPG and HOMA-IR, also measured at 15 weeks of pregnancy, along with HbA1c, FPG, HOMA-IR, and triglycerides measured postpartum, were important predictors. While GDM$_{WHO2013}$ itself was not a direct predictor, HbA1c in early pregnancy and the 2-hour glucose levels from the OGTT at 28 weeks of pregnancy were (Table 2). The model's AUC was 0.80 (95% CI 0.78, 0.82), indicating a high accuracy in predicting the combined risk of our primary outcome
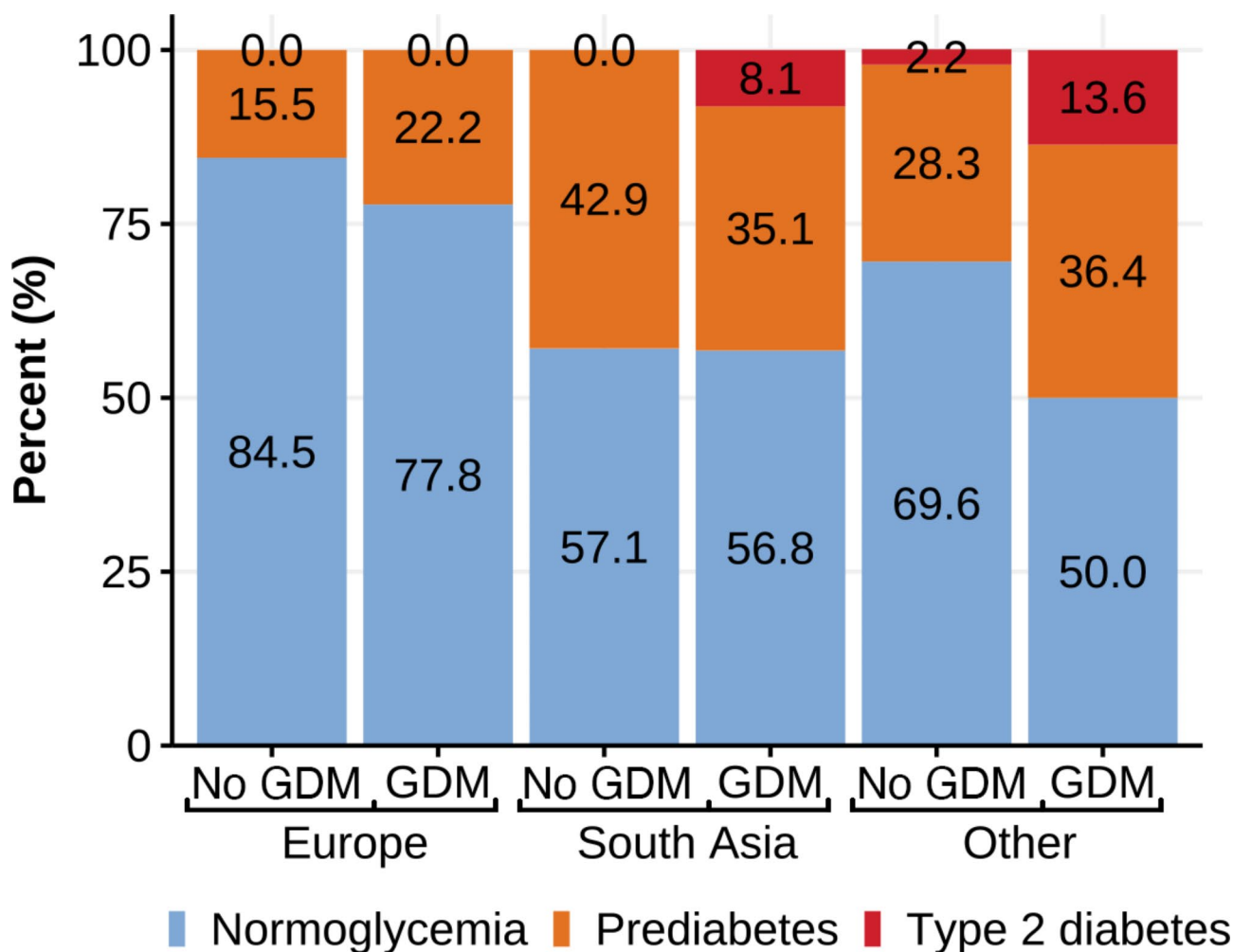
Fig. 4 Prediabetes and diabetes by GDM (WHO$_{2013}$) status by ethnic group

Table 2 LASSO[a] prediction model for prediabetes and diabetes after 11 years[b]

Predictors selected from visit 1–3	LASSO coefficient
Constant	-1.303
HbA1c (Visit 1)	0.424
Ethnicity (European)	-0.422
Family history of diabetes	0.152
HbA1c (Visit 3)	0.087
FPG (Visit 3)	0.062
2-hour PG (Visit 2)	0.043
HbA1c (Visit 2)	0.037
HOMA-IR (Visit 1)	0.037
FPG (Visit 1)	0.023
HOMA-IR (Visit 3)	0.019
Triglycerides (Visit 3)	0.015

[a]LASSO: least absolute shrinkage and selection operator

[b]Dysglycemia: prediabetes 39–47 mmol/mol and diabetes ≥ 48 mmol/mol

Visit 1: gestational week 15. Visit 2: gestational week 28. Visit 3: 14 weeks postpartum

Variables not selected by the model: education, age, parity, glucose (Visit 2), Pre-pregnant BMI, BMI (Visit 1–3), triglycerides (Visit 1 and 2), HOMA-IR (Visit 2), HOMA-β (Visit 1–3), Leptin (Visit 1–3), total fat mass (Visit 1–3), GDM$_{2013}$

(Fig. 5). Internal validation using bootstrapping confirmed the model's good predictive performance with an AUC of 0.80 (95% CI 0.79, 0.82) (Fig. 6). A nomogram for probability of developing diabetes or prediabetes$_{ADA}$ is presented in Fig. 2. The predictive performance of the nomogram gave an estimated C-index of ≥ 0.7. Examples of calculated risks for two participating women with different ethnicity are shown in Supplementary Fig. 1a and 1b.

Discussion

In this prospective, population-based cohort study of a multi-ethnic population of women universally screened for GDM$_{WHO2013}$, we found that one in three with GDM in the index pregnancy had the combined primary outcome diabetes or prediabetes$_{ADA}$ at the follow-up visit. We observed a high prevalence of prediabetes among ethnic minorities with women of South Asian background at highest risk. Important predictors for dysglycemia were HbA1c level from early pregnancy, ethnic minority background and family history of diabetes.
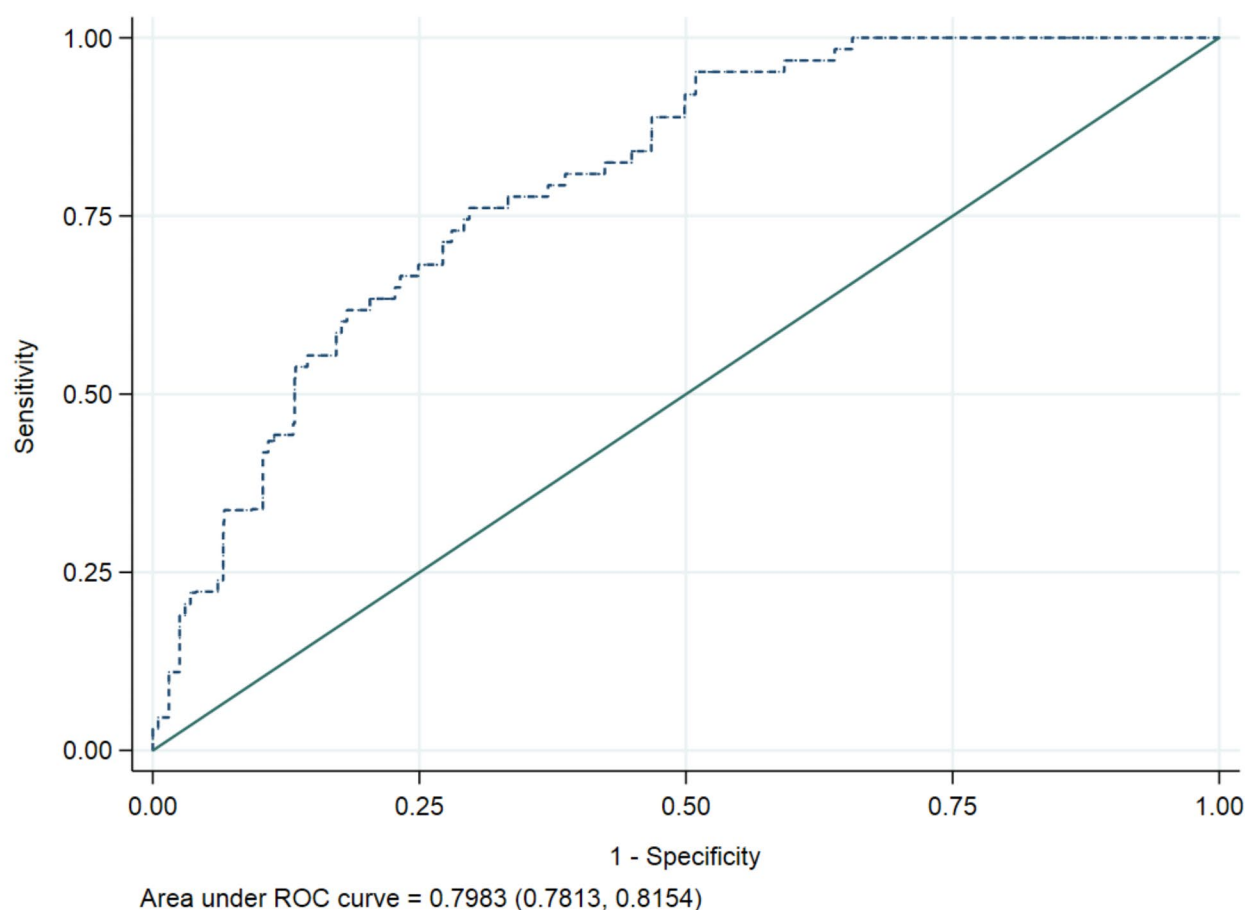

Fig. 5 ROC curve from LASSO for the performance of the least absolute shrinkage and selection operator risk score in identifying women with increased risk for prediabetes and diabetes. The AUC under the ROC curve is 0.7983 (95% CI 0.7813, 0.8154)
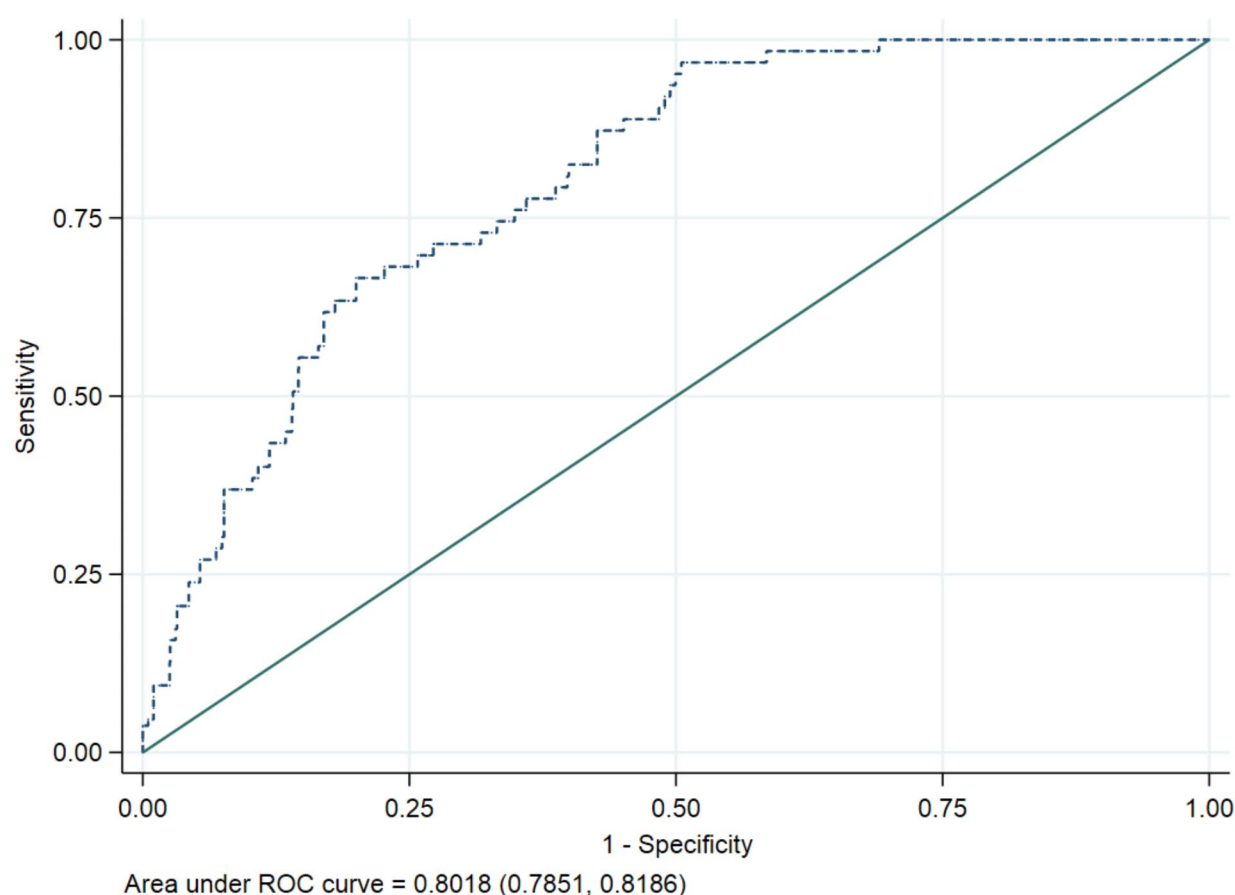
Fig. 6 ROC curve from Nomogram for the performance of the least absolute shrinkage and selection operator risk score in identifying women with increased risk for prediabetes and diabetes. The AUC under the ROC curve is 0.8018 (95% CI 0.7831, 0.8186)

Only very few of these women with a mean age of 42 years at follow-up had developed type 2 diabetes. We observed a slightly higher prevalence of dysglycemia in women with previous GDM. However, the RR for developing diabetes was substantially higher in previous GDM vs. non-GDM women.

Results in context – what's new

This is one of few long-term follow-up studies of a population-based, multi-ethnic cohort, universally screened for GDM, showing the long-term consequences for the mother of being diagnosed with GDM by the new WHO_{2013} criteria and the older WHO_{1999} criteria. In the total cohort, we have previously reported a GDM prevalence of 13.0% by the WHO_{1999} criteria [17], but most cases were mild, as only five were treated with glucose lowering drugs in pregnancy. The prevalence of GDM was, however, substantially higher (31.5%) by the WHO_{2013} criteria [17]. Those who met the WHO_{2013} criteria, but not diagnosed by the WHO_{1999} criteria and therefore not treated, had also mostly a mild GDM (a moderately elevated FPG and a 2-hour value < 7.8

mmol/l). Still, in a larger study where four Norwegian studies were pooled, this group had an increased risk of large for gestational age neonates, caesarean section and operative vaginal delivery compared to women without GDM [22].

Although the prevalence of prediabetes$_{ADA}$ was in line with other reports [23], the prevalence of diabetes was lower overall (1.9%) than estimated in our pre-specified power calculations, and in women with previous GDM (5.8%), substantially lower than reported in most previous studies (up to $> 50\%$) [24] and a recent meta-analysis [5]. Our results are, however, in line with a German study [25], which reported a diabetes prevalence of 6% after five years among women with previous GDM$_{WHO2013}$. We suggest that the major reason is that GDM detected by universal screening in a non-selected population identifies a group of women that differ from those included in many observational studies of GDM, that usually are based on referrals to secondary health care or on women that were subject to targeted screening in high-risk populations. Further, many studies have used OGTT-diagnosed diabetes as the outcome, while we used HbA1c

that is the recommended diagnostic test for diabetes today. The prevalence of diabetes diagnosed by HbA1c is generally lower than based on OGTT [26]. However, importantly, in the present study, one in three women with GDM in the index pregnancy identified with both the WHO_{1999} and the WHO_{2013} criteria had the primary outcome after universal screening, and GDM, particularly by the WHO_{2013} criteria represented a strong risk for diabetes, in line with other studies.

On the other hand, at follow-up, irrespective of the GDM definition used, $GDM_{WHO2013}$ constituted only a modest risk for the combined outcome (RR 1.4–3.6) and for prediabetes dependent on the definition (RR 1.3–3.2). Importantly, HbA1c at 15 weeks of pregnancy outperformed GDM diagnosed at gestational week 28 as predictor of the primary outcome in the LASSO regression model.

Few previous studies have such a rich dataset of potential predictors from pre-pregnancy and early pregnancy. The ethnic differences in diabetes and prediabetes$_{ADA}$ at the follow-up visit confirm previous findings. However, with information about family history of diabetes and ethnicity, HbA1c values from gestational week 15 can be used as a strong predictor for later diabetes or prediabetes$_{ADA}$. Other measures of glucose and HbA1c during later pregnancy and postpartum were also relevant and may reflect the effects of pregnancy more precisely than a diagnosis of GDM per se. Of note, no measures of baseline adiposity were included in the LASSO prediction model, probably because these factors are mediated by HbA1c and ethnicity already captured in the model.

Clinical relevance

As almost all pregnant women meet health care professional during pregnancy, they represent a unique potential for relevant health information and suggestions and support for prevention strategies. We find that simple clinical information like ethnic background and family history of diabetes combined with a single measurement of HbA1c early in pregnancy may provide the clinician a tool for relevant risk stratification regarding future diabetes. In this follow up study we chose to diagnose diabetes or prediabetes using only HbA1c and not an OGTT. This is in accordance with the recommendations in the Norwegian national guidelines for regular follow-up after GDM and also approved by the ADA [27].

Strengths and limitations

An important strength of the present study is the long-term follow-up of a cohort in an age-group highly relevant for preventive strategies. Further strengths include the population-based approach in a multi-ethnic population, also including "hard to reach" women (about 22% needed a translator at inclusion). We consider that a 53% (49% with HbA1c measured) attendance rate at the eleven years

follow-up visit despite the ongoing COVID-19 epidemic was decent, although obviously we planned for a higher participation rate. Universal screening with OGTT allowed for reliable estimates of GDM with different diagnostic criteria and the detailed data collected during and after pregnancy provided the opportunity for developing prediction models. To account for missing data, particularly at the postpartum visit, we performed multiple imputations to reduce the risk of biased estimates. The classification of diabetes or prediabetes as outcomes was based on HbA1c, which is the common measure used by clinicians today. However, HbA1c values in pregnancy must be interpreted with caution, as the levels are altered independent of glycaemia, due to changes in the red-cell turnover rate [28]. Further, ethnic background may influence HbA1c levels, as people of European decent have been reported to have lower HbA1c levels compared to other ethnic groups, e.g., African Americans and Asian Americans despite similar plasma glucose levels [29, 30].

Limitations include the fact that we were only able to recruit about half of the original cohort to this follow-up study, resulting in a relatively small sample size. Further, the number of individuals that developed diabetes was low, limiting the power for subgroup analyses. Further, as no OGTTs were performed at the eleven years visit, we cannot compare HbA1c-diagnosed and OGTT-diagnosed diabetes prevalence and how the prediction model would work for OGTT-diagnosed outcomes.

Conclusions

The prevalence of diabetes was low, both overall and among women with previous GDM. Still, GDM was a strong risk factor for diabetes, but not so for our primary combined endpoint of diabetes and prediabetes$_{ADA}$. Non-European ethnicity, a family history of diabetes and HbA1c early in pregnancy were the strongest risk factors for developing diabetes or prediabetes$_{ADA}$, as determined using an artificial intelligence algorithm.

Abbreviations

GDM	Gestational diabetes mellitus
OGTT	Oral glucose tolerance test
STORK G	Stork Groruddalen
HbA1c	Glycated haemoglobin
FPG	Fasting plasma glucose
HOMA-IR	Homeostatic model assessment for insulin resistance
HOMA-β	Homeostatic model assessment for beta-cell function
BMI	Body mass index
SD	Standard deviation
OR	Odds ratio
CI	Confidence interval
RR	Relative risk
LASSO	Least absolute shrinkage and selector operator
WHO_{2013}	World Health Organization criteria for diagnosing gestational diabetes, updated 2013
WHO_{1999}	World Health Organization criteria for diagnosing gestational diabetes from 1999

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s12889-025-22493-x.

Supplementary Material 1

Data availability
The editors can access data (in de-identified form) used in the manuscript, code book, and analytical code upon request. The project manager and the head of our department will contribute to the access being provided under appropriate conditions. However, research data for this publication include identifying health information subject to confidentiality. It is therefore not possible to share raw data.

Declarations

Ethics approval and consent to participate
Both the original Stork G study, and the follow-up study (Unique Protocol ID: 2018/2517) was conducted according to the Helsinki declaration and were approved by the Regional Committee for Medical and Health Research Ethics of South-Eastern Norway. Informed consent was obtained from all the participants and/or their legal guardians. All participants signed study-specific consent forms when included at baseline and at the eleven years follow-up.

Consent for publication
Not applicable.

Competing interests
The authors declare no competing interests.

Author details
[1]Department of General Practice, Institute of Health and Society, General Practice Research Unit (AFE), University of Oslo, Post Box 1130 Blindern, Oslo N- 0318, Norway
[2]Department of Rehabilitation Science and Health Technology, Faculty of Health Sciences, Oslo Metropolitan University, Oslo, Norway
[3]Institute of Clinical Medicine, Faculty of Medicine, University of Oslo, Oslo, Norway
[4]Department of Endocrinology, Obesity and Preventive Medicine, Oslo University Hospital, Oslo, Norway
[5]Department of Medical Biochemistry, Oslo University Hospital, Oslo, Norway

References
1. Pervjakova N, Moen GH, Borges MC, Ferreira T, Cook JP, Allard C, et al. Multi-ancestry genome-wide association study of gestational diabetes mellitus highlights genetic links with type 2 diabetes. Hum Mol Genet. 2022;31(19):3377–91.
2. Yang J, Qian F, Chavarro JE, Ley SH, Tobias DK, Yeung E, et al. Modifiable risk factors and long term risk of type 2 diabetes among individuals with a history of gestational diabetes mellitus: prospective cohort study. BMJ. 2022;378:e070312.
3. Retnakaran R, Qi Y, Sermer M, Connelly PW, Hanley AJ, Zinman B. Glucose intolerance in pregnancy and future risk of pre-diabetes or diabetes. Diabetes Care. 2008;31(10):2026–31.
4. Bellamy L, Casas J-P, Hingorani AD, Williams D. Type 2 diabetes mellitus after gestational diabetes: a systematic review and meta-analysis. Lancet. 2009;373(9677):1773–9.
5. 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;369.
6. Zhu Y, Zhang C. Prevalence of gestational diabetes and risk of progression to type 2 diabetes: a global perspective. Curr Diab Rep. 2016;16(1):7.
7. McIntyre HD, Catalano P, Zhang C, Desoye G, Mathiesen ER, Damm P. Gestational diabetes mellitus. Nat Rev Dis Primers. 2019;5(1):47.
8. O'Sullivan JB, Mahan CM, CRITERIA FOR THE ORAL GLUCOSE. TOLERANCE TEST IN PREGNANCY. Diabetes. 1964;13:278–85.
9. Metzger BE, Buchanan TA, Coustan DR, de Leiva A, Dunger DB, Hadden DR et al. Summary and recommendations of the Fifth International Workshop-Conference on Gestational Diabetes Mellitus. Diabetes Care. 2007;30 Suppl 2:S251-60.
10. Metzger BE, Gabbe SG, Persson B, Buchanan TA, Catalano PA, Damm P, 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):676–82.
11. Diagnostic criteria and classification. Of hyperglycaemia first detected in pregnancy: a world health organization guideline. Diabetes Res Clin Pract. 2014;103(3):341–63.
12. 15. Management of diabetes in pregnancy: standards of care in diabetes—2024. Diabetes Care. 2024;47(Supplement1):S282–94.
13. Metzger BE, Lowe LP, Dyer AR, Trimble ER, Chaovarindr U, Coustan DR, et al. Hyperglycemia and adverse pregnancy outcomes. N Engl J Med. 2008;358(19):1991–2002.
14. Waage CW, Toftemo I, Brænd AM, Sletner L, Sommer C, Birkeland KI, et al. Cohort profile update: the Norwegian STORK Groruddalen (STORK G) pregnancy and birth cohort—the role of ethnicity and causal pathways for obesity, type 2 diabetes, cardiovascular disease and other health issues. BMJ Open. 2023;13(10):e076251.
15. ElSayed NA, Aleppo G, Aroda VR, Bannuru RR, Brown FM, Bruemmer D, et al. 2. Classification and diagnosis of diabetes: standards of care in diabetes—2023. Diabetes Care. 2023;46(Supplement1):S19–40.
16. Jenum AK, Sletner L, Voldner N, Vangen S, Mørkrid K, Andersen LF, et al. The STORK Groruddalen research programme: A population-based cohort study of gestational diabetes, physical activity, and obesity in pregnancy in a multiethnic population. Rationale, methods, study population, and participation rates. Scand J Public Health. 2010;38(5 suppl):60–70.
17. Jenum AK, Morkrid K, Sletner L, Vangen S, Torper JL, Nakstad B, et al. Impact of ethnicity on gestational diabetes identified with the WHO and the modified international association of diabetes and pregnancy study groups criteria: a population-based cohort study. Eur J Endocrinol. 2012;166(2):317–24.
18. Waage C, Jenum AK, Mdala I, Berg JP, Richardsen K, Birkeland K. Associations between gestational diabetes mellitus and elevated HbA1c early postpartum in a multi-ethnic population. Prim Care Diabetes. 2017;11(2):132–9.
19. Wallace TM, Levy JC, Matthews DR. Use and abuse of HOMA modeling. Diabetes Care. 2004;27(6):1487–95.
20. Ueda Y, Maruo M, Nakano H, Honda Y, Miyama T, Nishizawa M, et al. Estimation of body fat mass in pregnant women by a new method using bioelectrical impedance analysis with compensation for intrauterine component weight. Int J Body Compos Res. 2006;4(4):145.
21. Piepoli MF, Hoes AW, Agewall S, Albus C, Brotons C, Catapano AL, et al. 2016 European guidelines on cardiovascular disease prevention in clinical practice: the sixth joint task force of the European society of cardiology and other societies on cardiovascular disease prevention in clinical practice (constituted by representatives of 10 societies and by invited experts)Developed with the

special contribution of the European Association for cardiovascular prevention & Rehabilitation (EACPR). Eur Heart J. 2016;37(29):2315–81.

22. Rai AS, Sletner L, Jenum AK, Øverby NC, Stafne SN, Qvigstad E, et al. Adverse pregnancy outcomes among women in Norway with gestational diabetes using three diagnostic criteria. PLoS ONE. 2023;18(7):e0280750.

23. Mainous AG, Tanner RJ, Baker R, Zayas CE, Harle CA. Prevalence of prediabetes in England from 2003 to 2011: population-based, cross-sectional study. BMJ Open. 2014;4(6):e005002.

24. Kim C, Newton KM, Knopp RH. Gestational diabetes and the incidence of type 2 diabetes. Diabetes Care. 2002;25(10):1862–8.

25. Haschka SJ, Gar C, Sacco V, Banning F, Ferrari U, Freibothe I, et al. Pre-diabetes, diabetes and fluctuations of glucose tolerance after gestational diabetes mellitus: 5-year follow-up of a contemporary, prospective study in Germany. BMJ Open Diabetes Res Care. 2022;10(2):e002621.

26. Kaur G, Lakshmi P, Rastogi A, Bhansali A, Jain S, Teerawattananon Y, et al. Diagnostic accuracy of tests for type 2 diabetes and prediabetes: A systematic review and meta-analysis. PLoS ONE. 2020;15(11):e0242415.

27. 15. Management of diabetes in pregnancy: standards of care in diabetes—2025. Diabetes Care. 2025;48(Supplement1):S306–20.

28. Mosca A, Paleari R, Dalfrà MG, Di Cianni G, Cuccuru I, Pellegrini G, et al. Reference intervals for hemoglobin A1c in pregnant women: data from an Italian multicenter study. Clin Chem. 2006;52(6):1138–43.

29. Ziemer DC, Kolm P, Weintraub WS, Vaccarino V, Rhee MK, Twombly JG, et al. Glucose-independent, black–white differences in hemoglobin A1c levels: a cross-sectional analysis of 2 studies. Ann Intern Med. 2010;152(12):770–7.

30. Herman WH, Cohen RM. Racial and ethnic differences in the relationship between HbA1c and blood glucose: implications for the diagnosis of diabetes. J Clin Endocrinol Metab. 2012;97(4):1067–72.