Diabetologia (2016) 59:970–979
DOI 10.1007/s00125-016-3869-8

ARTICLE

First-trimester multimarker prediction of gestational diabetes mellitus using targeted mass spectrometry

Tina Ravnsborg[1,2] · Lise Lotte T. Andersen[3] · Natacha D. Trabjerg[4] · Lars M. Rasmussen[1,2,5] · Dorte M. Jensen[2,3,4,5] · Martin Overgaard[1,5]

Received: 1 July 2015 / Accepted: 22 December 2015 / Published online: 27 January 2016

Abstract
Aims/hypothesis Gestational diabetes mellitus (GDM) is associated with an increased risk of pre-eclampsia, macrosomia and the future development of type 2 diabetes mellitus in both mother and child. Although an early and accurate prediction of GDM is needed to allow intervention and improve perinatal outcome, no single protein biomarker has yet proven useful for this purpose. In the present study, we hypothesised that multimarker panels of serum proteins can improve first-trimester prediction of GDM among obese and non-obese women compared with single markers.
Methods A nested case–control study was performed on first-trimester serum samples from 199 GDM cases and 208 controls, each divided into an obese group (BMI ≥27 kg/m^2) and a non-obese group (BMI <27 kg/m^2). Based on their biological relevance to GDM or type 2 diabetes mellitus or on their previously reported potential as biomarkers for these diseases, a number of proteins were selected for targeted nano-flow liquid chromatography (LC) MS analysis. This resulted in the development and validation of a 25-plex multiple reaction monitoring (MRM) MS assay.
Results After false discovery rate correction, six proteins remained significantly different (p<0.05) between obese GDM patients (n=135) and BMI-matched controls (n=139). These included adiponectin, apolipoprotein M and apolipoprotein D. Multimarker models combining protein levels and clinical data were then constructed and evaluated by receiver operating characteristic (ROC) analysis. For the obese, non-obese and all GDM groups, these models achieved marginally higher AUCs compared with adiponectin alone.
Conclusions/interpretation Multimarker models combining protein markers and clinical data have the potential to predict women at a high risk of developing GDM.

Keywords Adipokine · Apolipoprotein · Early prediction · Gestational diabetes mellitus · Obesity · Targeted MS

Abbreviations

Apo	Apolipoprotein
CRP	C-reactive protein
GA	Gestational age
β-hCG	Beta chain human chorionic gonadotropin
GDM	Gestational diabetes mellitus
IADPSG	International Association of Diabetes and Pregnancy Study Groups
LC	Liquid chromatography
LGA	Large for gestational age
LLOQ	Lower limit of quantification
MS	Mass spectrometry
MRM	Multiple reaction monitoring
PAPP-A	Pregnancy-associated plasma protein A
PPV	Positive predictive value

Electronic supplementary material The online version of this article (doi:10.1007/s00125-016-3869-8) contains peer-reviewed but unedited supplementary material, which is available to authorised users.

✉ Martin Overgaard
Martin.Overgaard@rsyd.dk

[1] Department of Clinical Biochemistry and Pharmacology, Odense University Hospital, Sdr. Boulevard 29, 5000 Odense C, Denmark

[2] The Danish Diabetes Academy, Odense, Denmark

[3] Department of Obstetrics and Gynaecology, Odense University Hospital, Odense, Denmark

[4] Department of Endocrinology, Odense University Hospital, Odense, Denmark

[5] Institute of Clinical Research, University of Southern Denmark, Odense, Denmark

RBP-4 Retinol-binding protein 4
ROC Receiver operating characteristic
SHBG Sex hormone binding protein

Introduction

Gestational diabetes mellitus (GDM) is defined as glucose intolerance diagnosed during pregnancy; it is associated with an increased risk of obstetric and neonatal complications such as pre-eclampsia, caesarean section, macrosomia, shoulder dystocia, neonatal hypoglycaemia and the future development of type 2 diabetes mellitus in both mother and child [1–4]. GDM is a common complication of pregnancy, and its prevalence depends on the population and the diagnostic criteria applied. Obesity is a major risk factor for GDM, and obese pregnant women have an up to eight times higher risk of developing GDM compared with normal weight pregnant women [5].

In Denmark, pregnant women undergo a selective screening procedure for GDM based on risk factors; GDM diagnosis is based on a 75 g OGTT with a 2 h plasma glucose threshold of 9.0 mmol/l [6]. This threshold is in accordance with previous recommendations from the European Diabetic Pregnancy Study Group [7]. Recently, new diagnostic criteria for GDM were recommended by the International Association of Diabetes and Pregnancy Study Groups (IADPSG). These would lower the diagnostic 2 h threshold value of a 75 g OGTT to 8.5 mmol/l. Once implemented, this criterion is expected to greatly increase the prevalence of GDM not only in Denmark but also globally, thus challenging the health economy [8]. The IADPSG criteria are based on second trimester OGTT values and evidence of adverse perinatal outcomes. This recommendation suggests ruling out overt diabetes in early pregnancy by measuring fasting glucose (≥7.0 mmol/l), random plasma glucose (≥11.1 mmol/l on confirmation) or HbA_{1c} (≥6.5% [≥48 mmol/mol]) levels. It was suggested that fasting glucose in the range of 5.1–6.9 mmol/l (which exceeds the IADPSG GDM criteria but does not meet the criteria for overt diabetes) should be considered as diagnostic of GDM. A drop in fasting glucose occurs in early pregnancy because of natural physiological changes [9]; hence, the use of this GDM diagnostic threshold may not identify the same individuals in early and later pregnancy. Furthermore, fasting glucose tests and/or OGTTs are time-consuming and elevated fasting glucose in early pregnancy is not highly predictive of later GDM development [10]. For these reasons, a biomarker assay that enables the early and accurate prediction of GDM is warranted. This could be an important tool in future screening strategies to ensure that women at the highest risk undergo a diagnostic OGTT at gestational week 24–28.

At present, no serum biomarker with clinical utility has been identified and the potential utility in GDM screening has been investigated for only a limited number of proteins. Adiponectin is by far the best-studied protein biomarker for GDM but lacks the sensitivity and specificity required for clinical use [11–16]. Other proteins such as sex hormone binding protein (SHBG), retinol-binding protein 4 (RBP-4), resistin and C-reactive protein (CRP) have also been investigated as potential biomarkers of GDM, but conclusions have been variable [14, 17–23].

The current study took a targeted proteomics candidate biomarker approach and used multiple reaction monitoring (MRM) MS to assess the potential of a number of both previously suggested and novel serum biomarkers for GDM in obese and non-obese pregnant women. Specifically, we hypothesised that multimarker panels of serum proteins could improve first-trimester prediction of GDM in obese and non-obese women. We developed a 25-plex MRM-MS assay for simultaneously measuring serum proteins using a minimal amount of sample material and with low analytical imprecision. Subsequent multivariate modelling of the data improved the AUC values obtained by receiver operating characteristic (ROC) analysis.

Methods

Samples and clinical data All samples used in this nested case–control study were obtained from a biobank of ~20,000 first-trimester non-fasting serum samples (stored at −80°C) collected from the routine screening for Down syndrome at Odense University Hospital [24]. All samples used in this study were taken between gestational week 8 + 1 and 13 + 6, as corrected by crown-rump length using astraia version 1.23.4.1 (astraia software, Munich, Germany). The Danish selective screening strategy for GDM includes a set of risk factors: BMI of ≥27 kg/m^2, family history of diabetes and previous birth of a child with macrosomia. It is based on a large prospective study and has a sensitivity of 81% and a specificity of 65% [25]. Women with two or more of these risk factors or who have previously had GDM are offered an early diagnostic OGTT between gestational weeks 14 and 20 and again during week 27–30, whereas women with just one risk factor are only offered the late OGTT. In addition, glucosuria diagnosed at any gestational age (GA) also generates a referral for an OGTT. Altogether, 199 early pregnancy samples from women later diagnosed with GDM were identified in the biobank. Inclusion criteria were a singleton pregnancy and HbA_{1c} values of <6.5% (48 mmol/mol) at the time of GDM diagnosis. In addition, 208 random samples from women matched by BMI and year of sampling served as controls. Samples were divided into two groups in accordance with current Danish GDM screening guidelines: BMI of ≥27 kg/m^2 (obese group) and <27 kg/m^2 (non-obese group). For statistical power calculations, see electronic supplementary

material (ESM) Methods. Clinical data corresponding to all 407 samples were manually retrieved from patient medical records, entered into a database (SPSS version 21.0, IBM Danmark, Holte, Denmark) and analysed.

All aspects of the study were approved by the local ethics committee (S-20130092).

Sample preparation and MRM-MS analysis Samples were prepared for MRM analysis essentially as described by Overgaard et al (for a detailed description, see ESM Methods) [26]. Briefly, serum samples were diluted 1:20 in 50 mmol/l ammonium bicarbonate and 15 μl samples were denatured, reduced, alkylated and trypsinised. Individually adjusted amounts of heavy isotope labelled standard peptides were added to each sample to achieve a ratio 1:1 with the endogenous light peptides; all peptides were then purified using Oasis HLB 10 mg cartridges (Waters Danmark, Hedehusene, Denmark). Samples were dried and reconstituted in 0.1% formic acid. For each sample, 1 μg was run on an Easy-nLC II nano liquid chromatography (LC) system using a C18 trapping column (length 2 cm; internal diameter 100 μm) for desalting and a C18 analytical column (length 10 cm, internal diameter 75 μm) for peptide separation (Thermo Scientific Danmark, Hvidovre, Denmark). Peptides were eluted with a three-step 60 min gradient of 0.1% formic acid in acetonitrile at a flow rate of 300 nl/min. Peptides were ionised using a Nanospray Flex ion source (Thermo Scientific) and analysed on a TSQ Vantage triple quadruple mass spectrometer (Thermo Scientific) in selected reaction monitoring mode.

For intra- and interassay CV calculations, pooled serum quality control samples were prepared in triplicate and analysed in duplicate on the LC tandem MS system. Standard curves (ESM Fig. 1) were made in triplicate for each peptide by adding different concentrations of heavy isotope labelled standard peptide to the same pooled serum. The spiked samples were finally processed according to the protocol for clinical samples. The lower limit of quantification (LLOQ) for each peptide was derived from ESM Fig. 1 (shown in ESM Table 1).

Immunological assays Adiponectin and resistin levels were measured using the DRP300 and DRSN00 ELISA kits (R&D Systems, Abingdon, UK). Double testing for serum pregnancy-associated plasma protein A (PAPP-A) and free β chain human chorionic gonadotropin (β-hCG) was performed as part of the routine Down syndrome screening using an automated time-resolved fluoro-immunoassay (AutoDELFIA; PerkinElmer, Turku, Finland) on a 1235 AutoDELFIA analyser. Further details on immunological assays are provided in the ESM Methods.

Data analysis and statistics SRM raw files were processed using Pinpoint 1.3 (Thermo Scientific). Data for each peptide represents the peak area ratio between the endogenous light peptide and the heavy isotope labelled spiked peptide. All data (clinical, MRM and ELISA) were transferred to an SPSS 21.0 (IBM) database and statistical analyses was done using SPSS 21.0 and Excel 2003 (Microsoft Danmark, Hellerup, Denmark).

Prior to the analysis, the dataset for each group (i.e. obese and non-obese) was randomised into a training set (comprising 75% of the data) to be used for biomarker verification and model development, and a validation set (comprising the remaining 25% of the data) to be used for model validation. Within each of the four data subsets, there was no significant difference in BMI between GDM patients and controls.

A significance level of $p<0.05$ was applied to all statistical tests used in this study.

Results

An overview of the sample collection, method development and data analysis procedures is provided in Fig. 1. Altogether, the levels of 29 proteins in 407 first-trimester serum samples were measured and analysed along with the relevant clinical data to investigate their potential as biomarkers (either individually or combined in multimarker models) for the early prediction of GDM.

Baseline clinical data Clinical data for all patient groups are listed in Table 1. For both the obese and non-obese groups, women with GDM were significantly older than control participants. Women diagnosed with GDM also gave birth significantly earlier than those in the control group ($p<0.001$ for both groups). This difference is probably explained by the routine use of labour induction up to 2 weeks before term for women with GDM; in the obese group, this is reflected in a lower birthweight and a smaller birth length and abdominal circumference. After GA was taken into account, there was no significant difference in the number of large for GA (LGA; i.e. >90th percentile) children born to women with GDM and controls. In contrast, the frequency of pre-eclampsia and caesarean sections were significantly higher in the obese GDM group compared with obese controls ($p=0.032$ and $p=0.039$, respectively), whereas differences between the non-obese groups were not significant.

MRM assay development We identified 54 candidate protein biomarkers from previous studies evaluating their potential as biomarkers for GDM and/or type 2 diabetes mellitus. Of these, 47 were selected for MRM-MS assay development because the characteristics of their trypsin-derived peptides (i.e. length, hydrophobicity, abundance and uniqueness) were theoretically compatible with the method. For a detailed description of the literature-based candidate search strategy and MRM assay

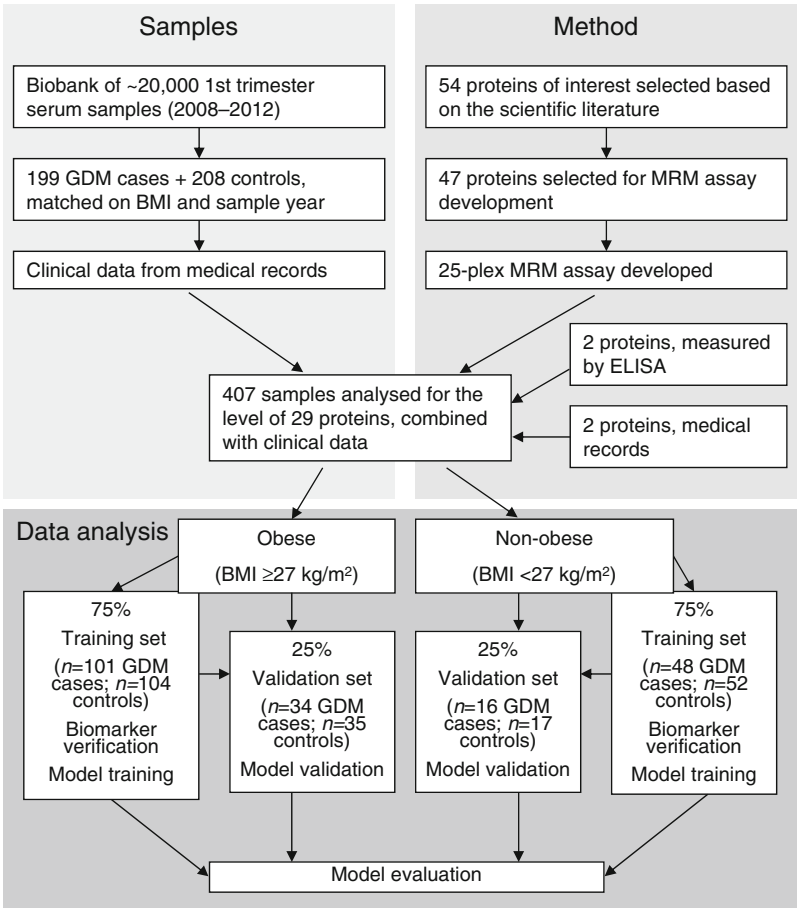

Fig. 1 Study workflow. Chart of the study workflow, illustrating its three stages: sample and clinical data collection; method development; and data analysis

peptide and transition selection, see ESM Fig. 2 and ESM Methods, respectively. The initial MRM-MS assay was tested on a pool of first-trimester serum and optimised using a spike of co-eluting heavy isotope labelled standard peptides to validate peptide identification and adjust for variations in sample preparation and LC-MS performance.

The final 25-plex MRM-MS assay measured the relative levels of 25 different proteins as represented by a total of 43 peptides (ESM Table 1). In this assay, 18 proteins were each represented by two peptides (for each protein, there was a high degree of correlation between peptides in the complete dataset; ESM Fig. 3) and seven proteins were each represented by a single peptide. Of the total of 25 proteins, 20 were represented by a minimum of five transitions. Standard curves verifying a linear correlation between amount of peptide and the intensity of the MS signal are shown in ESM Fig. 1. Peptide levels for all clinical samples were within the linear part of the standard curves and above the LLOQ shown in ESM Table 1. The interassay CVs for all peptides in the 25-plex MRM-MS assay were between 1.9% and 7.5% (ESM Table 1). Finally, the MRM assay was validated by a methods comparison analysis of three selected proteins for which either a clinical immunoassay or ELISA kit was available (ESM Fig. 4). The r values obtained for SHBG, apolipoprotein B (Apo B) and fibronectin were 0.97, 0.86 and 0.91, respectively.

Statistical analysis of the candidate biomarker panel In addition to the 25 proteins measured by MRM-MS, levels of two adipokines (adiponectin and resistin) were measured by ELISA and serum concentrations for PAPP-A and β-hCG along with maternal age were retrieved from patient medical records (ESM Table 2). The potential of each protein to differentiate between GDM cases and controls was evaluated by applying the Mann–Whitney U test separately to the obese and non-obese training sets. For the obese group, nine variables were significantly different between GDM cases and controls and seven remained so after false discovery rate correction by the Benjamini–Hochberg procedure (Table 2). Of these, five were measured in the 25-plex MRM assay. For the non-obese group, three variables were significantly different between GDM cases and controls; however, none of these differences remained significant after correction. For all variables that initially showed significant differences between cases and controls, ROC curves were constructed and AUC values evaluated (Table 2 and Fig. 2a, b). For the obese group, adiponectin was the best independent predictor of GDM (AUC 0.712 [95% CI 0.642, 0.783]) followed by

Table 1 Maternal pregnancy and offspring characteristics

Characteristics	Obese			Non-obese		
	GDM cases ($n = 135$)	Controls ($n = 139$)	p value	GDM cases ($n = 64$)	Controls ($n = 69$)	p value
Maternal						
Age at delivery (years)	32.0 (29.0–36.0)	30.0 (27.0–33.0)	0.002[a]	31.0 (28.0–35.0)	29.0 (26.5–33.0)	0.024[a]
BMI at week 11–14 (kg/m^2)[b]	32.6 (30.5–36.0)	32.6 (30.2–35.9)	NS[a]	23.7 (21.8–25.5)	23.8 (21.8–25.4)	NS[a]
Smokers	25 (19)	17 (12)	NS[c]	11 (17)	8 (12)	NS[c]
Ethnicity						
White	113 (87)	129 (93)	NS[c]	50 (79)	63 (91)	NS[c]
Other	17 (13)	10 (7)		13 (21)	6 (9)	
GA at time of sample (days)	73 (68–79)	72 (68–79)	NS[a]	70 (67–76)	72 (66–78)	NS[a]
GA at time of GDM diagnosis[d]	201 (177–210)			209 (199–244)		
Nulliparous	51 (38)	56 (40)	NS[c]	26 (41)	40 (58)	0.046[c]
GDM treatment						
Diet	79 (59)	–		41 (64)	–	
Insulin	56 (41)	–		22 (34)	–	
Pregnancy outcome						
GA at delivery (days)	269 (265–278)	281 (272–288)	<0.0005[a]	269 (266–277)	280 (273–286)	<0.0005[a]
Induced delivery	70 (52)	70 (50)	NS[c]	47 (73)	19 (28)	<0.0005[c]
Mode of delivery						
Vaginal	81 (60)	101 (73)	0.039[c]	45 (70)	57 (83)	NS[c]
Caesarean section	52 (39)	38 (27)		19 (30)	12 (17)	
Pre-eclampsia[e]	17 (13)	7 (5)	0.032[f]	3 (5)	2 (3)	NS[f]
Sphincter injury	4 (3)	6 (4)	NS[f]	3 (5)	6 (9)	NS[f]
Shoulder dystocia	2 (1)	1 (1)	NS[f]	1 (2)	0 (0)	NS[f]
Abortion	1 (1)	0 (0)	NS[f]	0 (0)	0 (0)	NS[f]
Offspring						
Sex						
Male	71 (53)	66 (47)	NS[c]	33 (52)	33 (48)	NS[c]
Female	63 (47)	73 (53)		31 (48)	36 (52)	
Weight (g)	3,442 (3,084–3,788)	3,625 (3,335–3,969)	<0.0005[a]	3,407 (3,148–3,733)	3,460 (3,091–3,868)	NS[a]
Length (cm)	52 (50–53)	52 (51–54)	0.001[a]	51 (50–53)	52 (50–54)	NS[a]
Abd circumference (cm)	34 (31–35)	34 (33–35)	0.040[a]	34 (32–35)	33 (31–35)	NS[a]
Apgar						
7–10	132 (98)	135 (97)	NS[f]	63 (98)	69 (100)	NS[f]
1–6	1 (1)	4 (3)		1 (2)	0 (0)	
LGA	29 (21)	25 (18)	NS[f]	11 (17)	5 (7)	NS[f]
Birthweight SDS[g]	0.242 (−0.609 to 1.174)	0.114 (−0.539 to 0.814)	NS[a]	0.291 (−0.365 to 1.056)	−0.216 (−0.999 to 0.549)	0.011
Stillbirth	1 (1)	0 (0)	NS[e]	0 (0)	0 (0)	NS[e]
Perinatal death	0 (0)	0 (0)	NS[e]	0 (0)	0 (0)	NS[e]

Data are the median and interquartile range or n (%)

[a] Mann–Whitney U test

[b] Obese, BMI ≥27 kg/m^2 ; non-obese, <27 kg/m^2

[c] Pearson's χ^2 test

[d] GA at the time of GDM diagnosis had 10 and 6 missing values for the obese and non-obese groups, respectively

[e] Defined as a BP of >140/90 and proteinuria

[f] Fisher's exact test

[g] Birthweight adjusted for GA

Abd, abdominal; SDS, SD score

Apo M (AUC 0.633 [95% CI 0.557, 0.710]). For the non-obese group, adiponectin also showed the best predictive performance (AUC 0.671 [95% CI 0.563, 0.778]), followed by Apo D (AUC 0.659 [95% CI

Table 2 Variables displaying significant differences between GDM patients and controls, with the corresponding AUCs

Training set	GDM patients	Controls	*p* value		AUC (95% CI)[a]
			Mann–Whitney *U* test	Corrected[b]	
Obese (BMI ≥27 kg/m^2)					
n	101	104			
Adiponectin (μg/ml)	6.49 ± 4.34	8.80 ± 3.91	<0.0005	<0.015	0.712 (0.642, 0.783)
NGAL	1.04 ± 0.22	1.16 ± 0.27	0.002	0.03	0.627 (0.551, 0.704)
SHBG	0.53 ± 0.19	0.63 ± 0.23	0.002	0.03	0.626 (0.550, 0.702)
α2-macroglobulin	1.92 ± 0.39	2.12 ± 0.47	0.002	0.03	0.625 (0.549, 0.702)
Apo M	2.24 ± 0.45	2.42 ± 0.38	0.003	0.018	0.633 (0.557, 0.710)
Apo D	1.04 ± 0.25	1.12 ± 0.22	0.007	0.03	0.608 (0.531, 0.686)
Haptoglobin β chain	3.22 ± 1.12	2.90 ± 1.02	0.036	0.135	0.585 (0.507, 0.663)
Clusterin α chain	1.39 ± 1.44	1.35 ± 0.16	0.038	0.127	0.584 (0.505, 0.662)
Maternal age at delivery (years)	31.96 ± 5.19	30.14 ± 4.81	0.006	0.03	0.610 (0.533, 0.688)
BMI (kg/m^2)	33.49 ± 4.46	33.50 ± 4.61	NS	NS	NS
Non-obese (BMI <27 kg/m^2)					
n	48	52			
Adiponectin (μg/ml)	8.88 ± 4.47	11.13 ± 4.75	0.003	0.09	0.671 (0.563, 0.778)
Apo D	1.16 ± 0.23	1.29 ± 0.26	0.006	0.09	0.659 (0.551, 0.768)
Apo L1	1.42 ± 0.32	1.28 ± 0.33	0.021	0.21	0.634 (0.524, 0.744)
BMI (kg/m^2)	23.39 ± 2.35	23.46 ± 2.55	NS	NS	NS

MRM-MS data is presented as the ratio of endogenous light peptide to heavy isotope spiked peptide

Data are means ± SD

[a] Asymptotic significance in the ROC analysis

[b] Corrected for multiple hypothesis testing using the Benjamini–Hochberg method

0.551, 0.768]) and Apo L1 (AUC 0.634 [95% CI 0.524, 0.744]). These results indicate that although adiponectin is potentially a good universal predictor of GDM, apolipoproteins have distinct discriminative properties in subgroups of GDM patients.

Model development and validation To further investigate the potential use of the biomarker panel evaluated in this study, multivariate models for predicting GDM were constructed using different combinations of variables (ESM Tables 3–6). All models were developed using binominal logistic regression analysis of a training set for each group (obese or non-obese) and then tested using a smaller validation set to rule out any overfitting of the data. For the same reason, no more than five variables were permitted for each model. Despite different modelling strategies, a number of variables recurred in several of the final models: namely maternal age adiponectin, Apo M and Apo L1 for the obese group; and Apo D and Apo L for the non-obese group. All models, including a 4-plex core panel common to both study groups, were evaluated by ROC analysis and corresponding AUC values (Table 3, Fig. 2c–f and ESM Table 3). Of the models listed in Table 3, Model 2 and Model 7 had the best overall predictive performance and showed the least difference in AUC values when tested in the validation set of each group (obese or non-obese). Moreover, both performed better than any of the single variables alone, including adiponectin (Table 2).

Based on the ROC analysis of Model 2 and Model 7, predictive values were calculated for high, medium and low specificity using the cohort prevalence of GDM in obese (2%) and non-obese (0.5%) pregnant women (Table 4). Similar calculations were also performed to account for an estimated future prevalence of 10% (expected after implementation of the new diagnostic IADPSG criteria for GDM). As shown in Table 4, the obese and non-obese models have a similar performance but different optimal cut-off points. As expected, positive predictive values (PPVs) increase with an increased prevalence of GDM. This is also evident for the estimated prevalence of 10%, assuming the models perform equally well for the new diagnostic threshold.

Discussion

In the present nested case–control study, we used MRM-MS and ELISA to identify and validate two 5-plex multimarker models that discriminate between: (1) GDM in obese pregnant women and BMI-matched obese control participants; and (2)

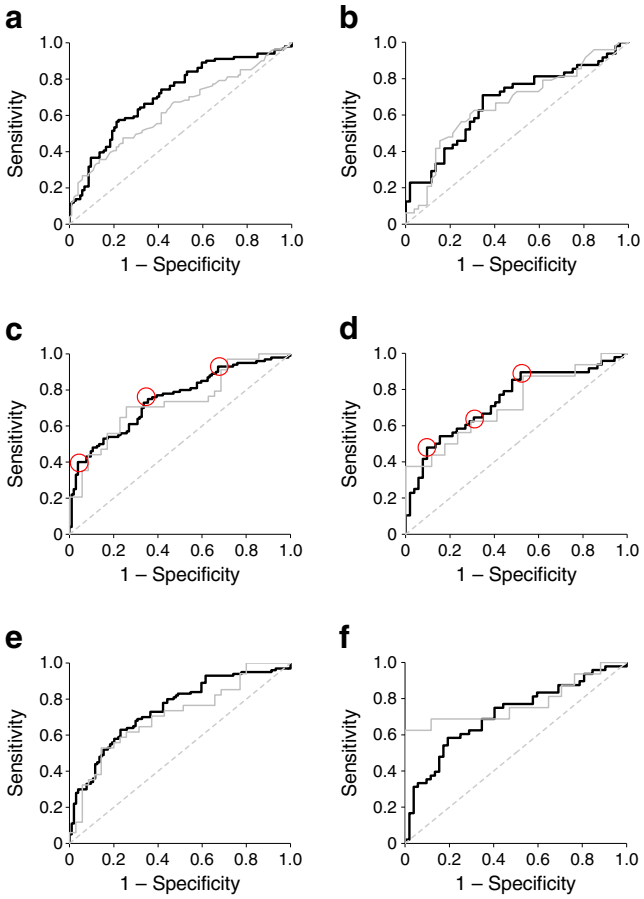

Fig. 2 ROC analysis of univariables and multivariate models. Univariate analysis of selected variables based on the training set of the (**a**) obese and (**b**) non-obese groups. Adiponectin (black line), Apo M (grey line in **a**) and Apo D (grey line in **b**). Multivariate analysis of (**c**) Model 2, obese group, (**d**) Model 7, non-obese group, (**e**) the core model obese group and the (**f**) core model non-obese group based on the training (black line) and validation (grey line) set. Red circles, the optimal cut-off points for high, medium and low specificity; dashed line; reference

GDM in non-obese pregnant women and BMI-matched non-obese control participants. Our data confirm that the adipocyte-derived hormone, adiponectin, is a specific independent predictor of GDM in both obese and non-obese women and directly compares its performance with those of a number of other previously suggested markers such as CRP, SHBG and the adipokines neutrophil gelatinase-associated lipocalin (NGAL), resistin and RBP-4 [14, 17–23, 27–32]. Of the novel candidate biomarkers included in our multiplex MRM-MS assay, Apo M, Apo D, Apo L1 and Apo AIV all appeared in the final multimarker models for predicting GDM.

In recent years, the MRM-MS method has gained immense popularity within the field of proteomics based on its application in biomarker development studies [33, 34]. Our study demonstrates how targeted MS applied to a classic nested case–control study, when adequately powered to detect subtle differences in serum protein levels, can support the development of multimarker models with better disease prediction. This approach also allowed us to directly compare a number of protein markers for which published results have been contradictory or inconclusive. A substantial advantage of MRM-MS assays is their compatibility with the often limited sample amount available in retrospective studies; our 25-plex MRM-MS assay uses less than 1 μl of crude serum. Although we used isotopically labelled peptide standards to correct for analytical and solid-phase extraction variances, the assay can potentially be affected by differences in trypsin digestion efficiencies. We chose a 21-h endpoint for the sample preparation protocol and validated the MRM assay by a methods comparison analysis of SHBG, Apo B and fibronectin; high r values were obtained for each (0.97, 0.86 and 0.91, respectively). As expected, the highest degree of correlation was obtained for the comparison between the MRM assay and an automated Siemens Immulite immunoassay. Apo B and fibronectin data were comparable between the MRM assay and manual ELISAs; for Apo B, the degree of correlation was similar to that previously obtained in a comparative study of LC-MS and nephelometry [35]. Taken together, we conclude that the MRM assay used in this study is fit for purpose but that further development of more focused marker panels with internal digestion controls would be valuable. A clear limitation of the method per se is its relatively low analytical sensitivity for serum or plasma (low μg/ml) when used without any prior reduction of sample complexity [26, 36]. Nevertheless, targeted MS offers an interesting analytical platform for future multimarker-based prenatal screening.

In this study, we divided the dataset into a training set for univariate and multivariate analyses and a separate validation set for testing the logistic prediction models and, thus, selection of those proteins with the best predictive performance. Of the 29 potential protein biomarkers, adiponectin is the best studied in relation to GDM [11–14, 16]. Here, adiponectin was the best performing independent predictor of GDM, regardless of BMI grouping. SHBG, RBP-4, CRP and resistin have also been described as potential biomarkers of GDM in numerous publications [14, 17–23, 28–31]. In this study, however, this potential was confirmed only for SHBG in obese women; in contrast, there was no significant difference in RBP-4, CRP and resistin levels between GDM patients and controls. This discrepancy may be attributable to sample collection at different GAs or statistical uncertainty due to a limited sample size.

The many apolipoproteins included in the 25-plex MRM-MS assay were primarily chosen based on their previously reported potential as biomarkers of GDM and/or type 2 diabetes mellitus (ESM Fig. 2) [37–39]. The best independent predictive apolipoproteins were Apo M, Apo D and Apo L1 (Tables 2 and 3); the performance of Apo M and Apo L1 depended on BMI grouping. After logistic regression modelling, Apo CIII emerged as a significant contributory variable in two obese models (Table 3). In contrast, Apo M and Apo L1

Table 3 ROC analysis of multivariate models

Model	Variables included in the model	Training set		Validation set	
		AUC (95% CI)	*p* value[a]	AUC (95% CI)	*p* value[a]
Obese (BMI ≥27 kg/m^2)					
1	Adiponectin, Apo M, Mat age, Apo L1, SHBG	0.754 (0.687, 0.820)	<0.0005	0.694 (0.571, 0.817)	0.006
2	Adiponectin, Apo M, Mat age, Apo L1, Apo AIV	0.749 (0.682, 0.816)	<0.0005	0.731 (0.612, 0.850)	0.001
3	Adiponectin, Mat age, Apo L1, Apo CII, Apo CIII	0.749 (0.682, 0.815)	<0.0005	0.645 (0.514, 0.775)	0.039
4	Mat age, Apo L1, Apo CII, Apo CIII, transferrin	0.699 (0.628, 0.770)	<0.0005	0.668 (0.539, 0.797)	0.016
5	Adiponectin, Mat age	0.731 (0.660, 0.801)	<0.0005	0.560 (0.423, 0.697)	0.394
Core	Adiponectin, Apo M, Mat age, Apo L1	0.746 (0.679, 0.814)	<0.0005	0.707 (0.583, 0.830)	0.003
Non-obese (BMI <27 kg/m^2)					
5	Adiponectin, Mat age	0.694 (0.589, 0.799)	<0.001	0.732 (0.548, 0.916)	0.023
6	β-hCG, PAPP-A, Apo D, Apo L1	0.750 (0.653, 0.847)	<0.0005	0.717 (0.537, 0.897)	0.034
7	Adiponectin, Apo M, Mat age, Apo L1, Apo D	0.739 (0.640, 0.837)	<0.0005	0.721 (0.545, 0.896)	0.031
8	β-hCG, Apo D, Apo L1	0.729 (0.629, 0.829)	<0.0005	0.706 (0.526, 0.886)	0.044
Core	Adiponectin, Apo M, Mat age, Apo L1	0.712 (0.609, 0.814)	<0.0005	0.776 (0.603, 0.948)	0.007
All GDM					
2	Adiponectin, Apo M, Mat age, Apo L1, Apo AIV	0.724 (0.667, 0.781)	<0.0005	0.747 (0.651, 0.842)	<0.0005
5	Adiponectin, Mat age	0.711 (0.652, 0.770)	<0.0005	0.621 (0.512, 0.730)	0.036
7	Adiponectin, Apo M, Mat age, Apo L1, Apo D	0.731 (0.674, 0.787)	<0.0005	0.718 (0.617, 0.818)	<0.0005
Core	Adiponectin, Apo M, Mat age, Apo L1	0.723 (0.666, 0.780)	<0.0005	0.725 (0.626, 0.824)	<0.0005

Training set, obese: GDM patients (n = 101) + controls (n = 104); validation set, obese: GDM patients (n = 48) + controls (n = 52); training set, non-obese: GDM patients (n = 34) + controls (n = 35), validation set, non-obese: GDM patients (n = 16) + controls (n = 17); training set, all GDM: GDM patients (n = 135) + controls (n = 139), validation set, all GDM: GDM patients (n = 64) + controls (n = 69)

[a] Asymptotic significance

Mat, maternal

were included in both obese and non-obese models. This is therefore the first study to show that apolipoproteins have a predictive potential for GDM and a dependency on BMI.

Maternal age also proved to be significantly different between GDM cases and controls; thus, it was included in the final models (Table 3 and ESM Tables 3 and 4). No correlation was observed between maternal age and any of the protein markers included in the models. To evaluate the contribution of each marker to the AUC, we performed a 'leave one out' analysis (ESM Table 5). Apo M and adiponectin were the best single contributors in Model 2 (obese group), whereas Apo L1 made the greatest contribution to Model 7 (non-obese group). Based on these data, a 'core' panel for both groups was identified comprising adiponectin, Apo M, maternal age and Apo L1 (Table 3). The AUCs of the core panel were comparable with those of Model 2 and Model 7 when all data were analysed together regardless of BMI (all GDM group). In contrast, the core model performed marginally worse in the obese and non-obese groups. Taken together, the lack of enhanced discriminative power of each model in the obese and non-obese group compared with the all GDM group supports a universal rather than a BMI-stratified screening approach using different marker panels.

Although the predictive performance of both Model 2 (AUC 0.749) and Model 7 (AUC 0.739) was increased compared with the AUCs of single markers (Table 2), their performance in a low GDM prevalence population is only moderate. Moreover, although implementation of the IADPSG diagnostic criteria for GDM is expected to increase the prevalence to the vicinity of 10%, the associated test PPV of 19% for Model 2 at a medium cut-off point of 75% sensitivity and 64% specificity seems fair but not game changing. For a future screening regime, we envision the combined use of maternal history data, weight, age and biomarkers for risk calculation and, thus, stratification of the diagnostic OGTT. A similar multifactorial screening regime is widely used for the first-trimester detection of major aneuploidies including trisomy 21 [40], as well as for pre-eclampsia screening, which is currently being implemented at many prenatal screening clinics [41].

Clearly, the availability of complete maternal history data (i.e. previous GDM, previous birth of macrosomia child and diabetes in the family) for evaluating the selective screening regime both alone and in combination with serum markers would have strengthened this study. Furthermore, we used a nested case–control design, which can potentially lead to inflation of the AUC of the ROC compared with population-

Table 4 PPV and NPV calculations[a]

Optimal cut-off point	Obese (Model 2)			Non-obese (Model 7)		
	High	Medium	Low	High	Medium	Low
Sensitivity (%)	40.0	75.0	93.0	47.9	64.6	89.6
Specificity (%)	96.2	64.4	32.7	90.4	69.2	48.1
0.5% prevalence						
PPV (%)				2.4	1.0	0.9
NPV (%)				99.7	99.7	99.9
Accuracy (%)				90.2	69.2	48.3
2% prevalence						
PPV (%)	17.7	4.1	2.7			
NPV (%)	98.7	99.2	99.6			
Accuracy (%)	95.1	64.6	33.9			
10% prevalence						
PPV (%)	53.9	19.0	13.3	35.7	18.9	16.1
NPV (%)	93.5	95.9	97.7	94.0	94.6	97.7
Accuracy (%)	90.6	65.5	38.7	86.2	68.7	52.3

PPV, NPV and accuracy were calculated at the optimal cut-off points for high, medium and low specificity according to ROC analysis of Model 2 and Model 7 for the obese (BMI ≥27 kg/m^2) and non-obese (BMI < 27 kg/m^2) groups, respectively

[a] Fictive populations of 1,000 with different prevalences of GDM were used for the calculations

NPV, negative predictive value

based studies. The study design has nevertheless been shown to be an efficient alternative to cross-sectional studies for the exploration of potential predictors [42]. We acknowledge that the ROC evaluation of the prediction models in the non-obese group was limited by sample size, although it still reached statistical significance. Clearly, GDM screening performance using the suggested models should be validated in prospective studies and should preferably also include maternal history data.

Another limitation of the current study was that only obese women in the control group underwent a diagnostic OGTT, whereas lean control women were not tested. Thus, we have no proof of normoglycaemia in pregnancy for the latter group. However, all women had urine glucose measured at every visit to the outpatient clinic throughout pregnancy, and any level of glucosuria would have resulted in a subsequent OGTT. We therefore consider GDM to be unlikely in the lean control women. The use of HbA_{1c} measurement to rule out overt diabetes at the time of GDM diagnosis is another limitation to the study. HbA_{1c} measurement forms part of our routine clinical analysis to identify women at risk of diabetic complications such as retinopathy and nephropathy. HbA_{1c} levels decrease during pregnancy because of physiological changes [43]; thus, the cut-off level of 6.5% (48 mmol/mol) might be too high and so identify some women with more severe glucose intolerance.

In conclusion, this study has provided a comprehensive overview of the performance of protein biomarkers in early GDM prediction. Evidently, multivariate models using different combinations of predictors can improve the discriminative power over single markers.

Overall, the performance of the current models leaves room for additional metabolic and proteomic studies to identify novel early markers with even better discriminative power. Importantly, assays for novel candidate markers should detail molecular target specificity and be accessible to the research field. To this end, targeted MS offers an excellent platform for preclinical biomarker development to improve future prenatal care.

Acknowledgements We thank V. Jørgensen, Odense University Hospital, Denmark, for excellent technical assistance.

Funding This work was supported by a research grant from the Danish Diabetes Academy supported by the Novo Nordisk Foundation (to TR and DMJ) and by the Odense University Hospital Research Foundation (MO).

Duality of interest statement The authors declare that there is no duality of interest associated with this manuscript.

Contribution statement All authors made substantive contributions to the study conception and design, data acquisition or data analysis and interpretation. All authors contributed to drafting the manuscript and approved the final version of the article. In addition, TR, DMJ, LMR and MO conceived the study; TR and MO developed the 25-plex MRM assay; TR performed the sample handling and ELISA analyses; TR and MO performed the statistical analyses; DMJ and LLTA identified patients and resolved clinical issues; and TR and NDT extracted data from the medical records. MO and DMJ are the guarantors of this work.

References

1. Group HSCR, Metzger BE, Lowe LP et al (2008) Hyperglycemia and adverse pregnancy outcomes. N Engl J Med 358:1991–2002
2. Kim C, Newton KM, Knopp RH (2002) Gestational diabetes and the incidence of type 2 diabetes: a systematic review. Diabetes Care 25:1862–1868
3. Ovesen PG, Jensen DM, Damm P, Rasmussen S, Kesmodel US (2015) Maternal and neonatal outcomes in pregnancies complicated by gestational diabetes. A nation-wide study. J Matern Fetal Neonatal Med: 1–5
4. Coustan DR (2013) Gestational diabetes mellitus. Clin Chem 59: 1310–1321
5. Chu SY, Callaghan WM, Kim SY et al (2007) Maternal obesity and risk of gestational diabetes mellitus. Diabetes Care 30:2070–2076
6. Jensen DM, Damm P, Sorensen B et al (2003) Proposed diagnostic thresholds for gestational diabetes mellitus according to a 75-g oral glucose tolerance test. Maternal and perinatal outcomes in 3260 Danish women. Diabet Med 20:51–57
7. Lind T, Phillips PR (1991) Influence of pregnancy on the 75-g OGTT. A prospective multicenter study. The Diabetic Pregnancy Study Group of the European Association for the Study of Diabetes. Diabetes 40(Suppl 2):8–13
8. International Association of Diabetes and Pregnancy Study Groups Consensus Panel, Metzger BE, Gabbe SG et al (2010) International association of diabetes and pregnancy study groups

recommendations on the diagnosis and classification of hyperglycemia in pregnancy. Diabetes Care 33:676–682
9. Mills JL, Jovanovic L, Knopp R et al (1998) Physiological reduction in fasting plasma glucose concentration in the first trimester of normal pregnancy: the diabetes in early pregnancy study. Metabolism 47:1140–1144
10. Corrado F, D'Anna R, Cannata ML, Interdonato ML, Pintaudi B, Di Benedetto A (2012) Correspondence between first-trimester fasting glycaemia, and oral glucose tolerance test in gestational diabetes diagnosis. Diabetes Metab 38:458–461
11. Ferreira AF, Rezende JC, Vaikousi E, Akolekar R, Nicolaides KH (2011) Maternal serum visfatin at 11–13 weeks of gestation in gestational diabetes mellitus. Clin Chem 57:609–613
12. Georgiou HM, Lappas M, Georgiou GM et al (2008) Screening for biomarkers predictive of gestational diabetes mellitus. Acta Diabetol 45:157–165
13. Lacroix M, Battista MC, Doyon M et al (2013) Lower adiponectin levels at first trimester of pregnancy are associated with increased insulin resistance and higher risk of developing gestational diabetes mellitus. Diabetes Care 36:1577–1583
14. Nanda S, Savvidou M, Syngelaki A, Akolekar R, Nicolaides KH (2011) Prediction of gestational diabetes mellitus by maternal factors and biomarkers at 11 to 13 weeks. Prenat Diagn 31:135–141
15. Ahnstrom J, Lindqvist PG, Walle U, Dahlback B (2010) Plasma levels of apolipoprotein M in normal and complicated pregnancy. Acta Obstet Gynecol Scand 89:1214–1217
16. Worda C, Leipold H, Gruber C, Kautzky-Willer A, Knofler M, Bancher-Todesca D (2004) Decreased plasma adiponectin concentrations in women with gestational diabetes mellitus. Am J Obstet Gynecol 191:2120–2124
17. Maged AM, Moety GA, Mostafa WA, Hamed DA (2014) Comparative study between different biomarkers for early prediction of gestational diabetes mellitus. J Matern Fetal Neonatal Med 27:1108–1112
18. Smirnakis KV, Plati A, Wolf M, Thadhani R, Ecker JL (2007) Predicting gestational diabetes: choosing the optimal early serum marker. Am J Obstet Gynecol 196:410, e1–6; discussion 410 e6–7
19. Caglar GS, Ozdemir ED, Cengiz SD, Demirtas S (2012) Sex-hormone-binding globulin early in pregnancy for the prediction of severe gestational diabetes mellitus and related complications. J Obstet Gynaecol Res 38:1286–1293
20. Su YX, Hong J, Yan Q et al (2010) Increased serum retinol-binding protein-4 levels in pregnant women with and without gestational diabetes mellitus. Diabetes Metab 36:470–475
21. Tepper BJ, Kim YK, Shete V, Shabrova E, Quadro L (2010) Serum retinol-binding protein 4 (RBP4) and retinol in a cohort of borderline obese women with and without gestational diabetes. Clin Biochem 43:320–323
22. Lewandowski KC, Stojanovic N, Bienkiewicz M et al (2008) Elevated concentrations of retinol-binding protein-4 (RBP-4) in gestational diabetes mellitus: negative correlation with soluble vascular cell adhesion molecule-1 (sVCAM-1). Gynecol Endocrinol 24:300–305
23. Qiu C, Sorensen TK, Luthy DA, Williams MA (2004) A prospective study of maternal serum C-reactive protein (CRP) concentrations and risk of gestational diabetes mellitus. Paediatr Perinat Epidemiol 18:377–384
24. Ekelund CK, Jorgensen FS, Petersen OB, Sundberg K, Tabor A, Danish Fetal Medicine Research Group (2008) Impact of a new national screening policy for Down's syndrome in Denmark: population based cohort study. BMJ 337:a2547
25. Jensen DM, Molsted-Pedersen L, Beck-Nielsen H, Westergaard JG, Ovesen P, Damm P (2003) Screening for gestational diabetes mellitus by a model based on risk indicators: a prospective study. Am J Obstet Gynecol 189:1383–1388
26. Overgaard M, Cangemi C, Jensen ML, Argraves WS, Rasmussen LM (2015) Total and isoform-specific quantitative assessment of circulating fibulin-1 using selected reaction monitoring MS and time-resolved immunofluorometry. Proteomics Clin Appl 9:767–775
27. Lou Y, Wu C, Wu M, Xie C, Ren L (2014) The changes of neutrophil gelatinase-associated lipocalin in plasma and its expression in adipose tissue in pregnant women with gestational diabetes. Diabetes Res Clin Pract 104:136–142
28. Chen D, Fang Q, Chai Y, Wang H, Huang H, Dong M (2007) Serum resistin in gestational diabetes mellitus and early postpartum. Clin Endocrinol 67:208–211
29. Kuzmicki M, Telejko B, Szamatowicz J et al (2009) High resistin and interleukin-6 levels are associated with gestational diabetes mellitus. Gynecol Endocrinol 25:258–263
30. Noureldeen AF, Qusti SY, Al-Seeni MN, Bagais MH (2014) Maternal leptin, adiponectin, resistin, visfatin and tumor necrosis factor-alpha in normal and gestational diabetes. Indian J Clin Biochem 29:462–470
31. Vitoratos N, Deliveliotou A, Dimitrakaki A et al (2011) Maternal serum resistin concentrations in gestational diabetes mellitus and normal pregnancies. J Obstet Gynaecol Res 37:112–118
32. D'Anna R, Baviera G, Corrado F, Giordano D, Recupero S, Di Benedetto A (2009) First trimester serum neutrophil gelatinase-associated lipocalin in gestational diabetes. Diabet Med 26: 1293–1295
33. Yassine H, Borges CR, Schaab MR et al (2013) Mass spectrometric immunoassay and MRM as targeted MS-based quantitative approaches in biomarker development: potential applications to cardiovascular disease and diabetes. Proteomics Clin Appl 7:528–540
34. Heegaard NH, Ostergaard O, Bahl JM et al (2015) Important options available–from start to finish–for translating proteomics results to clinical chemistry. Proteomics Clin Appl 9:235–252
35. Agger SA, Marney LC, Hoofnagle AN (2010) Simultaneous quantification of apolipoprotein A-I and apolipoprotein B by liquid-chromatography-multiple-reaction-monitoring mass spectrometry. Clin Chem 56:1804–1813
36. Kuzyk MA, Smith D, Yang J et al (2009) Multiple reaction monitoring-based, multiplexed, absolute quantitation of 45 proteins in human plasma. Mol Cell Proteomics 8:1860–1877
37. Kaur P, Rizk NM, Ibrahim S et al (2012) iTRAQ-based quantitative protein expression profiling and MRM verification of markers in type 2 diabetes. J Proteome Res 11:5527–5539
38. Li RX, Chen HB, Tu K et al (2008) Localized-statistical quantification of human serum proteome associated with type 2 diabetes. PLoS One 3, e3224
39. Kim SM, Park JS, Norwitz ER et al (2012) Identification of proteomic biomarkers in maternal plasma in the early second trimester that predict the subsequent development of gestational diabetes. Reprod Sci 19:202–209
40. Nicolaides KH (2011) Screening for fetal aneuploidies at 11 to 13 weeks. Prenat Diagn 31:7–15
41. Poon LC, Nicolaides KH (2014) Early prediction of preeclampsia. Obstet Gynecol Int 2014:297397
42. Biesheuvel CJ, Vergouwe Y, Oudega R, Hoes AW, Grobbee DE, Moons KG (2008) Advantages of the nested case–control design in diagnostic research. BMC Med Res Methodol 8:48
43. Nielsen LR, Ekbom P, Damm P et al (2004) HbA_{1c} levels are significantly lower in early and late pregnancy. Diabetes Care 27:1200–1201